THREE EVOLUTIONS

Evolution of Universe, Evolution of Life, Evolution of Consciousness

Bernard Korzeniewski

KORONA, Kraków 1998 {Polish Edition}
Translated and Modified

Table of contents

INTRODUCTION.....3
Part I: PHYSICAL EVOLUTION.....7
I.1. Nature of physical evolution.....7
I.2. Evolution of the Universe.....33
I.2.1. Why do we believe the Universe to evolve?.....34
I.2.2. What does the evolution of the Universe consist in?.....36
I.2.3. What did in fact come into being at the beginning of the Universe?.....42
I.2.4. How did all this begin?.....52
I.2.5. Why does matter exist at all?.....59
I.3. Thermodynamic evolution.....69
I.3.1. Essence of thermodynamic evolution.....69
I.3.2. Thermodynamic evolution of the Universe.....87
I.3.3. Thermodynamic evolution of life.....102
Part II: EVOLUTION OF LIFE.....114
II.1. Classical formulation of biological evolution.....114
II.2. Cybernetic formulation of biological evolution.....150
II.3. The origin of life.....192
Part III: EVOLUTION OF CONCEPTUAL NETWORK.....245
III.1. What is conceptual network?.....248
III.2. Neural network as basis of conceptual network.....254
III.3. Evolution of conceptual network in individual development.....261
III.4. Evolution of conceptual network in biological evolution.....283
III.5. Culture and evolution of conceptual network.....297
III.6. Conceptual network and philosophy.....308

INTRODUCTION

The present book treats about the evolution of the Universe, the evolution of life and the evolution of self-consciousness. It attempts to answer the question concerning the essence of three levels of reality: the physical, the biological and the psychic one. It aims to investigate how a subject conscious of its own existence (i.e. simply self-consciousness) emerges from the purely biological functioning of human brain, how the purposefulness of structure and behaviour of living organisms emerges from the laws governing the behaviour of atoms, and finally, how the entire variety of objects that form the Universe emerges from nothingness.

In its significant part, the book is a popularising work that presents selected achievements of natural sciences, mainly physics and biology. Its ambitions, however, reach much further, in the direction of proposing totally new approach to some of the considered questions. In particular, I present my own views concerning the heart and the essence of life and consciousness. I try to reflect on all these problems in a wider, philosophical context, and to analyse the implications of the present state of knowledge of natural sciences for the general vision of reality in its diverse aspects. The philosophical approach starts to predominate in the third part of the book, focused on the evolution of the conceptual network, where I deal with the nature of human minds and the resulting implications concerning the status of our world vision.

I indicate evolution – physical, biological and psychical (conceptual) – as the main axis of this vision, integrating different levels of reality. In a sense, evolution is only a pretext for describing a uniform vision of the world as a whole. Yet the pretext is very handy and well in place, for the world, together with all the physical objects, living organisms and conscious subjects, not only is, but is becoming. Evolution is the other face of the existence of reality. Of the reality, that is "occurring" simultaneously at three basic levels, clearly different yet forming a unity, namely: the physical, the biological and the psychic one. Let it suffice as an explanation of the title of the book and an outline of the domains it tackles.

The first part of the book is an attempt to discuss selected aspects of physical evolution (its general attributes, the evolution of the Universe and the thermodynamic foundations of the evolution of life), especially these that bear significant philosophical implications and are important for the broad range of interdisciplinary knowledge. As the entire book, this part is addressed to readers who are not necessarily experts in physics or natural sciences in general. Therefore, certain simplifications, analogies and generalisations were unavoidable, although they might seem slightly far-fetched from the point of view of physicists "of pure breed". I do hope, however, that the resignation from the exact character of presentation is counterbalanced with its clarity and insight. It is important to keep an eye on the entire wood (the Universe as a multifarious entity) while entering between the trees (physical facts and theories). Moreover, each honest physicist will admit, that an attempt to answer certain questions – as "what does the existence of the Universe consist in?" "where does the arrow of time come from?" or "what was at the beginning?" – at the present stage of the development of knowledge verges inevitably on speculation. It does not mean that one should avoid posing such questions. Clear formulation of the scope of our knowledge and ignorance, as well as the incomplete character of the passing theories allow us to understand in a more complete manner the essence of the problem and enhances further attempts at finding solutions. In fact, what we still do not know is always what is most interesting.

The second part devoted to biological evolution focuses on the essence of the phenomenon of life, its evolution based on the mechanism of natural selection, as well as on the spontaneous emergence of life at the beginnings of the history of our planet. The thermodynamic paradigm in approaching the question of life (presented in part one) is followed in turn by the "classical" paradigm that approaches biological phenomena in the context of structure, function, inheritance of genetic information, reproduction and evolution. By realising the self-imposed duty of popularising the achievements of modern biology, I attempt again to situate the knowledge in a more general context and prepare in this way the ground for introducing my own conception concerning the very essence of the phenomenon of life. I formulate the conception within the framework of cybernetic paradigm that deals with broadly-understood regulative mechanisms. I look in particular for a minimal and necessary definition of life in a system of negative feed-backs subordinated to a governing positive feed-back. The presented criterion for distinguishing the living from the inanimate is then applied to a range of border phenomena, such as viruses, prions, cancerous formations and parasitic DNA, as well as to the process of spontaneous origination of life from inanimate matter.

Finally, the third part of the book concerns the evolution of the conceptual network – i.e. of what constitutes in my opinion the "substance" of the entire psychic domain and the related thought processes. This part presents mainly my own, original conception (discussed in detail in the book published in Polish and titled *Absolut - odniesienie urojone*). I assume that the nature of meanings in a conceptual network is a derivative of the general structure of the neural network that forms the ground for the former. In both networks, "meanings" of particular elements (neurones, concepts) mutually condition each other – some elements create an appropriate semantic context for others. Such structure of the conceptual network has enormous impact on the world view shaped within such network. The world view is to the same extent a derivative of the real world "in itself", as of the manner in which our brain integrates external signals. The conceptual network does not originate out of nothing. It grows and becomes more and more complex both during the development of an individual, and during the evolution of humankind. The factors conditioning the nature of the conceptual network – both those related to its biological origins, and those related to its neuronal functional foundation – bear great and to a great extent still undiscovered philosophical implications concerning the status and limitations of our cognition and our knowledge of the world.

The third part presents also a proposition concerning the essence of consciousness. In my opinion, consciousness emerges from the biological level (of brain functioning) – due to the fact, that the cognitive center in the brain, responsible for the creation of a picture of the surrounding world, becomes directed on itself (which means that apart a picture of the world, the center generates also a picture of itself) – just as the biological level emerges from the physical one when certain self-reproducing systems become self-oriented, i.e. aim at survival and expansion of their own identity. Thus, the "layering" of reality (and its evolution) into three principal levels is a derivative of one simple relation, namely the relation of being self-directed (self-applicable), which forms the basis of the liar's paradox, Russell's antinomy of classes or Gödel's proof.

Who is the book addressed to and who belongs to the potential group of its readers? It seems to me that the question is slightly misguided. It is primarily a book "from the author", that contains what I consider worth being presented in the subject in question, and it was written to all those interested in the mysteries of the world, those who happened to come into being for a short time in this world. Certainly, one cannot consider the book as pure popularisation, for I propose here a vision of human mind that is in many aspects a result of my own original work. The vision, however, is deeply rooted in the contemporary scientific knowledge, an my opinions converge in many places with those of other authors. The popularising (very important) aspect of this book is meant to make it available for non-specialists in the discussed domains. I would like to believe, however, that it will be of interest for biologists of different specialisations (e.g. neruophysiologists, ethologists, evolutionists), but also for a broad group of cognitively-minded philosophers and psychologists, interested in acquiring a broader and synthetic approach. At the same time, I tried to distinguish clearly the popularised scientific knowledge from my own views.

PHYSICAL EVOLUTION

NATURE OF PHYSICAL EVOLUTION

The term "evolution" is applied to a very broad spectrum of diverse phenomena, such as the evolution of the Universe, of the Earth crust, of living organisms, societies, of language(s) or somebody's opinions - to mention just a few. Attribution of the same label or linguistic name to such a broad range of frequently so different processes (concerning objects so diametrically opposed and based on so diverse mechanisms) rises the question of the essential meaning of the concept of evolution[1]. Undoubtedly, each of us has a more or less explicit intuitive understanding of this concept. Nevertheless, due to the specific character of particular cases, it is necessary each time to attune the concept of evolution before applying it to a particular process. The present book is an attempt at rendering - according my best understanding - the essence of the phenomenon of physical, biological and psychical evolution.

The present chapter will focus particularly on the question of what should be understood as the evolution of various physical systems. In order to determine the specific properties of the physical evolution, it would be convenient to find a reference point, that would allow one to characterise this process in a relatively transparent manner, by specifying its similarities and differences with respect to some other phenomena. Biological evolution can be regarded as a kind of "standard" for the concept of evolution in general, and for the physical evolution in particular. Biological evolution leads to the creation of something new and, as a rule, better and more complicated than its simpler and less perfect predecessors. It is of no consequence, whether such commonsensical opinions find complete corroboration in reality. There is no doubt that biological evolution, on one hand, exhibits certain similarities to physical evolution (at least this most trivial that both are changes occurring in time), while, on the other hand, it has a "surplus," lacking in various transformations taking place on the physical level, which brings one to face the question of what is evolution in the first place (if one assumes biological evolution to be the obligatory canon) and to what extent (if at all) the concept refers – in a heuristically fruitful manner – to purely physical systems.

[1] One could doubt if a concept of evolution understood so broadly has any non-trivial meaning at all, apart from being a synonym of the concept of "change in time".

The present chapter will attempt to analyse this problem by using the evolution of the living world as a relevant, intuitively comprehensible reference point. Such move will also be advantageous in the broader context of the entire book, as the evolution on the biological level (including the way this level emerges from the physical level) constitutes the subject matter of the second part of this book. Now, I will try to focus on some properties characteristic for biological evolution, attributable to particular kinds and cases of the **physical evolution** as well.

One frequently hears about the evolution of a certain physical (as well as chemical, geological *etc.*) system. The system can comprise the entire Universe, an object falling down in the gravitational field or gas in a box. All those "evolutions" have a certain common property (present in the case of the biological evolution), namely – the **change in time**. The Universe expands, galaxies and stars are getting formed in it and life develops on planets; a falling stone changes its position in space; gas in a box tends to the most probable state – homogeneous distribution of molecules. It is hardly necessary to provide more arguments proving that change in time constitutes an inseparable attribute of physical evolution.

One may ask, however, if the change in time itself suffices to formulate a univocal definition of physical evolution. In other words: is physical evolution simply a change in time (or, should each change in time on the physical level be acknowledged as evolution)?

Let us consider a system of two bodies, one of which revolves around the other in space under the influence of gravitational attraction, as in the case of the Earth revolving around the Sun. Let us assume, for the sake of simplicity, that the dimensions of both bodies approach the dimensions of a point, while the mass of the satellite is infinitely small (or, better, close to zero) – so that the center of gravity of the system coincides with the point representing the central body (with any required accuracy) – and that the gravitational interaction with other bodies is negligible (all these assumptions are necessary to avoid any disturbances that could influence the behaviour of our idealised system in any way). Such assumptions never reflect exactly the real situation, but they are frequently a sufficiently close representation of facts.

Let us see what happens in the case described above. The system will behave identically for an infinitely long time – the satellite will revolve around the central body with a certain constant period of revolution T. Thus, although there is a change in time here, the change is cyclic – after the period T the system returns to its starting point. Actually, it is not fully unequivocal in this case to state that there occurs a change in time either, and it would require the concept of "change" to be specified in a more detailed manner. Another example of such a system one finds in an (idealised) pendulum – a "device" moving periodically between two states of maximal deflection and zero velocity, between which it passes through the state of zero deflection and maximal velocity. The behaviour of the above systems is not only cyclic, but also reversible, i.e. stopping of the satellite or pendulum and setting them in motion in the opposite direction will change nothing in the dynamics of the system. The revolution of the satellite around the central body occurs in two dimensions, and therefore the reversal of the direction of its movement will transform the system in its mirror reflection (the satellite will revolve around the central body in the opposite direction). As the movement of the pendulum occurs principally in one dimension, setting the pendulum in motion in the reversed direction will not bring about even such a change in this system.

Do the systems described above evolve, as there occurs a change in time in them? Intuitively, one rejects such a diagnosis (let us remember that biological evolution constitutes for us the standard – the reference point). The satellite position obviously changes, if we consider a time interval comparable with T. If we choose the time scale larger by a couple orders of magnitude and average or "integrate" the spatial situation of the satellite in time, then it will turn out, that its situation is constant and nothing changes in the system. Thus, the answer to the question whether a given system changes in time (and therefore evolves, at least potentially) depends on the **time scale** in which the system is considered.

Let us take the example of electromagnetic radiation emitted by a laser. The laser ray is coherent, which means the electromagnetic waves concur in phase or that all waves in a given moment reach simultaneously the "peaks" or "valleys" or corresponding intermediate points of the sinusoid (electromagnetic waves are sequences of electric and magnetic waves inducing each other, and "hills" and "valleys" of the waves correspond to maxima and minima of intensity of respective fields). Our eyes detect electric and magnetic fields with periodically changing intensity. However, these changes are so quick that human eye perceives laser light as a constant (invariable in time) stream of light. The same reasoning applies to chaotic time courses like noise, for example. Here, if the parameters undergoing fluctuations are averaged in a sufficiently long time range, they reach a constant value. In these examples, we certainly do not deal with evolution.

By changing very slightly the initial conditions (velocities) and/or the edge conditions (influence of other bodies), it is possible to modify the above-described system (a satellite revolving around a central body) in such a way, that the distance between the satellite and the central body will decrease slightly with each subsequent revolution. In other words, the course of the satellite will no longer be a circle or an ellipse – the satellite will approach the central body along a very dense spiral, with particular loops located close to one another. The difference can be so small, that it becomes almost imperceptible. This system, however, will not keep returning cyclically to the starting point. It will be characterised by a continuous, **directional change** of some parameter(s), namely the distance between the satellite and the central body. This constitutes an unquestionable progress in the direction towards the accepted standard of biological evolution (which undoubtedly represents a directional change). Changes of a system tending in some particular direction not only correspond better to our intuitive imaginations of what evolution is, but also seem to represent a much more realistic situation. The above-presented example of an ideally cyclic process (the system satellite-central body) can exist only in the form of a mathematical model, as a completely isolated system. Real systems can only constitute its more or less imperfect approximation. Practically every process in the physical world – starting from the growth of crystals and ending with the transformations of the Universe – is directional, and therefore shows certain features of evolution. Does it not depend – one may wander – on the point of reference?

Let us return for a moment to the problem of time scale in physical evolution, and particularly in periodic processes. As it has been shown above, a given process may appear to be a periodically changing system or an invariable system if we consider, respectively, a time span comparable with the duration of oscillations or a much longer time span. If the difference amounts to at least one order of magnitude, only a small segment of a sinusoid describing the cyclic changes in the system will be found in the field of observation. The fragment, being approximately a straight segment, will not reveal the periodical nature of the process in question. In such a situation, an observer notices a continuous change of some parameter(s) of the system and concludes that a directional process occurs in this system. An organism born at dawn and dying one hour later would claim that the intensity of light increases constantly in its world, which suggests an evolution towards "lightness". This example is not so far removed from reality as one could suppose at first glance. Some cosmological models, assuming a positive curvature of space, allow for the possibility of a cyclic, oscillating Universe. In the first stage of a cycle, the Universe would expand very violently (the beginning of this expansion is the Big Bang) from the initial singularity (the state with zero dimensions and infinite density and temperature). The velocity of the expansion would slow down in time, like the speed of a stone thrown upwards, up to the moment when the gravitational forces stopped the expansion of the Universe. From this moment on, it would begin to collapse faster and faster, ending its life in the terminal singularity, which would be at the same time the initial singularity for the next cycle. Life would develop on many planets in each cycle, which would lead here and there to the origination of thinking creatures. Therefore, there would occur biological and cultural evolution. Both life and civilisations, however, would perish in the terminal singularity. It does not mean that subsequent universes would be exact replicas of the previous ones, but that a great similarity would occur between them, if the matter is viewed statistically (i.e. by taking into account billions of galaxies, forms of life and civilisations). Therefore, if the pulsating model of the Universe is true, then everything we call evolution in the exact sense (e.g. the biological evolution) would appear only a semblance of it, for it would be nothing, but a small fragment of a circular (periodical) process. Observed from outside in a sufficiently extensive time scale (but these are orders of magnitude escaping our comprehension), such an oscillating Universe would seem to be a system in which nothing changes, as in the case of the ray of laser, received by our eye as a constant stream of light. The slightly grotesque aspect of the described point of view results from our anthropocentrism – every day we deal with much smaller time scales, determined by our physical and biological properties (for example the length of our life) and we are primarily interested in them. Science has broadened this range – from subatomic phenomena to the time span of biological evolution and the duration of the Universe. Yet it is possible to go further. The above example illustrates again the fact that it depends on the applied time scale whether we recognise a given process as evolving or not. To avoid misunderstandings, for the present I take into account only two properties, two aspects of evolution: change in time and its directionality (it is by no means my aim to reduce the concept of evolution *ad absurdum*, but rather to inquire into its essence). As far as these properties are concerned, both biological and cultural evolutions – if only considered within sufficiently long time spans – may look for us as not evolving (of course, biological evolution is characterised by specific properties like natural selection, for instance, that allow to define this evolution in an unequivocal manner). The same applies also to what we call the evolution of the Universe. Nothing changes from the point of view of eternity.

The problem of time scale reappears in a juxtaposition of a sudden change or catastrophe with evolution as a continuous change. Let us look at a nuclear explosion – perhaps the most spectacular of all catastrophes known to us. Probably nothing seems more violent and discontinuous. However, if we "filmed" (metaphorically, since it is physically impossible) the interior of a lump of a radioactive isotope of uranium with the mass exceeding the critical level and replayed afterwards the film considerably slowed-down, we would see what follows: A spontaneously decaying nucleus of an unstable isotope of uranium produces three neutrons. By colliding with other nucleuses, the neutrons make them split and produce (in each case) three further neutrons. Because a neutron striking a nucleus results in the production of three neutrons, the number of these particles increases exponentially. We have therefore the growing sequence of 3, 9, 27, 71, 213, 639 and so on neutrons. This is the beginning of what we call chain reaction. It constitutes a typical example of positive feedback – the more neutrons have already been emitted, the more neutrons are being produced. The discussed process is therefore characterised by a continuous, directional change in time, which we have identified as one of the attributes of evolution.

Let us, in turn, look at the growth of a bacteria colony on an agar-agar culture medium. At the beginning, we place there a single bacterium. It feeds, grows and after some time divides into two bacteria. The descendent bacteria divide again after the same time into two bacteria each, which yields four cells, and afterwards eight cells, sixteen and so on. The amount of bacteria "evolves" therefore gradually in the direction of greater quantity. After a short time, their number already reaches millions and billions. Therefore, we see here exactly the same phenomenon we saw in the case of nuclear explosion –a positive feedback, manifesting itself in a proportionate correlation between the rate of the colony growth and its current size. If the film of the evolutionary growth of the bacteria colony is replayed a thousand or million times faster, one will watch what one witnessed in the case of the nuclear explosion – a sudden, violent catastrophe. Therefore, whether we classify a phenomenon as a discontinuous, jump-like change or as a continuous evolution depends also on the adopted time scale. The analogy between a nuclear explosion and the growth of a bacteria colony can be further extended. After all, neither the number of neutrons, nor the amount of bacteria grows infinitely, as the supply of uranium nucleuses gets exhausted, in the first case, and in the second case the medium is used up and the environment gets poisoned with products of metabolism. However, if not the limited amount of medium, the Earth would become in a short time a ball of bacteria expanding with the speed of light.

The above section presents a discussion of the model of an oscillating closed (i.e. with positive curvature) Universe (the next chapter considers different models of the Universe in more detail). If viewed in a sufficiently large time scale, such a Universe turns out not to evolve. And what about the infinite, open Universe? Like its closed counterpart, it begins with the Big Bang, the greatest catastrophe in the history of the Universe. However, if a sufficiently small time segment is chosen, would not the Universe take the aspect of a process? That is where, one faces a problem, for an infinite division of time into smaller and smaller segments is not valid within the framework of quantum mechanics – the indeterminacy principle imposes limits here. Moreover, there may be no sense in speaking about the existence of time at the beginning of the Universe, as Stephen Hawking suggests. In his conception, time emerges as one of physical parameters in some distance from the region of space-time which we call the beginning of time. Therefore, one cannot speak about time at the very beginning, like one cannot speak about the "beginning" of the terrestrial globe at the north pole. Space-time bends over its beginning, taking the shape of a cap without any borders or edges.

Let us therefore pass on to the other end of the scale. In the open model, any significant changes in the Universe (such as origination and evolution of galaxies, stars, planets and life) occur "only" during the first tens or hundreds of billions of years of its existence. Afterwards, stars burn out, the speed of the escape of galaxies gradually decreases, life dies out as thermodynamic gradients disappear, large part of matter becomes "imprisoned" in black holes. What follows is the "thermal death" of the Universe, in which nothing interesting happens any more (this conclusion meets however some interpretative problems, since the maximal possible entropy of the Universe grows faster than the actual entropy of the Universe). Therefore, if viewed within a sufficiently long time span, the history of the Universe appears to begin with the initial catastrophe – when all the existing resources of hydrogen and lighter elements "burn up" in nuclear synthesis, life develops and civilisations originate – followed by a period of stabilisation and immobility that lasts to infinity. Is there any space left for evolution at all?

Therefore, whether our Universe undergoes evolution or not depends, regardless of its geometry, on the order of magnitude of the time span taken into consideration. The properties of evolution imply not only a change in time and its directionality, but also a certain characteristic **time scale** in which the change occurs. This observation in its full sense applies to biological evolution as well. It does not occur in a period shorter than the time necessary for alteration of generations (needed for production of progeny), amounting in bacteria (in optimal conditions) to about twenty minutes, and in the case of man – to roughly twenty years. On the other hand, if this problem is viewed in the perspective of really astronomical time scales, the evolution of life will appear to be either a cyclic process (oscillating Universe) or an episode of the initial catastrophe (open universe, expanding to infinity).

In the face of all this, let us return again to intermediate time scales, that is the scales contained between the time of occurrence of subatomic phenomena and the period of the existence of the Universe. Here, the evolution of both particular physical systems (singled out better or worse) and the Universe as a whole can be clearly observed. Directionality was to be the determinant of evolution, differentiating this process from other changes in time. However, if it is not known in advance whether an observed phenomenon is a part of a circular course of event or not, there arise certain problems. For it is possible for one to err like the born-at-dawn ephemeral creature that observed the evolution of its world towards the "lightness" during the one hour of its life. However, the course of night-day changes – as a manifestation of the cyclic phenomenon of the rotation of the Earth around its axis – is itself a cyclic process as well. If the time of observation is much shorter than the duration of one cycle of a periodical phenomenon, the latter may be indistinguishable from a really directional process. Such a distinction seems to be important, for earlier in the text we excluded circular processes from the circle of evolutionary processes. Are we therefore doomed to eternal uncertainty as to the non-cyclic nature of some of the investigated processes[2]?

Fortunately, this is not the case. True directionality of a process is gauged by its **irreversibility**. One remembers the example of a satellite revolving around a central body. Nothing essential changed due to the change of direction. The satellite still revolved around the

[2] While the cyclic character of a process can be proven – it is enough to carry on an observation for the time equal to or greater than the time T, i.e. the duration of one period of a cyclic phenomenon – it is impossible to prove its non-cyclic character in an analogous way – the period T can be longer than any finite time of observation.

central body, but in the opposite direction. Exactly the same effect would be obtained by reversing the direction of the passage of time instead of the vector of velocity, which would be equivalent to making the system move in the direction of its own past. It means that the described system is symmetrical with respect to the axis of time, and the processes occurring in it are completely reversible. It applies also to the discussed above modification of the satellite-central body system, where a directional process occurred, i.e. the satellite was falling down on the central body along a tight spiral. The reversal of the velocity vectors (or of the arrow of time) would also yield a directional change, yet it would be occurring in the opposite direction and take the form of a slow movement of the satellite along a spiral course away from the center of the system. Therefore, the fact that a certain process is directional does not entail that it is irreversible as well.

The above-presented reversible dynamic system of two bodies mutually interacting through the gravitational force is described by physics by means of Newton's theory and its generalisation –Einstein's general relativity theory. It is interesting that the dynamic equations of these theories – as well as of another great theory: quantum mechanics – when applied to even more complicated systems, are symmetrical with respect to time, and therefore they are reversible in their essence. If these theories are correct and complete (i.e. able to describe exhaustively all phenomena in the Universe), they lead to the conclusion that all occurring processes are reversible. This contradicts openly our every-day experience.

Although one would not be particularly astonished to see the Moon revolving around the Earth in the opposite direction, one would find it glaringly unnatural for fragments of a broken cup to gather from the floor and form one object in order to land quietly on the table (as in a film replayed backwards), or for a well-mixed alcoholic drink to separate on one's very eyes into its ingredients. This is simply because most (if not all) processes in the Universe are **irreversible** and cannot proceed backwards.

Undoubtedly, irreversibility characterises biological evolution as well. It would be difficult to imagine man – a very recent product of evolution after all – commencing to move back to his former stages, climb back on trees in a few millions of years, and in a few hundred millions of years – transform back into a crossopterygian fish in the sea, to become a uni-cellular organism in a few billions of years and finally terminate his evolutionary march as a mixture of organic compounds in the primary ocean. Therefore, we have to agree that many phenomena in the surrounding world, both on the physical and biological level, are of irreversible character, and that this property constitutes an inseparable attribute of the evolution. The problem of the irreversibility of physical processes will be discussed in more detail in the chapter devoted to thermodynamic evolution.

While discussing physical evolution, either on the scale of the entire Universe, or in the perspective of the distances in space and time closer to our experience, or else, occurring in the microworld, at the level of atoms and elementary particles, it is impossible to avoid tackling the problem of **determinism**. Or, to be more precise, the problem of its absence, that is of indeterminism. One can talk about determinism when all future properties of a system can be predicted (at least in principle) completely and univocally on the basis of its past states and the knowledge of the laws of physics. On the other hand, indeterminism takes place when such an exact inference of future states from past states is not possible. If one looks at biological evolution, one must state – with a complete certainty – that its course is not determined. This is due to the accidental character of mutations, that is changes in the genetic information of living organisms, as well as to unpredictable fluctuations (or even catastrophic changes) of the properties of the environment inhabited by those organisms. If biological evolution on the planet Earth were set forth once again from the very beginning, then it would undoubtedly proceed along completely different route than did for the first time. Some general properties – for instance, the division of the living world into "plants" and "animals" – might be preserved, but species and greater systematic groups would certainly differ form their present counterparts. In particular, the probability of a repeated origination of the *Homo sapiens* species would tend practically to zero. The evolutionary history leading to the origination of man is due to so improbably huge number of accidental events, coincidences and strokes of luck, that the chance of its repetition for the second time should be completely excluded. The indeterminism of biological evolution seems therefore to be an unshakeable fact. But, what form does the problem of determinism take in physical evolution?

The presence or absence of determinism depends on the system taken into consideration. If one deals with a system composed of few (preferably two) astronomical bodies, then its behaviour approximates well the functioning of a deterministic system. In most realistic situations, however, it is an indeterministic (at least to some degree) behaviour that seems principally to take place. On the physical level, one can distinguish two completely independent kinds of indeterminism: **quantum indeterminism** and **chaotic (thermodynamic) indeterminism**. As it will appear later, the latter kind, due to the macroscopic behaviour of a system composed of a great amount of elements (atoms and molecules), is a close relative of the determinism of evolution on the biological level.

Let us regard the antagonism between the deterministic and indeterministic behaviour in our Universe in its real form. The classical Newton's theory of gravitation as well as Einstein's relativity theory are deterministic theories. Within their framework, movements of any physical objects, stars, stones and atoms are ruled by strict mathematical laws, determining the changes in the location of those objects in space with a potentially infinite accuracy. The similar state of affairs is presented by the classical electrostatic theories or Maxwell's wave theory of light. If one, like the famous demon of Laplace, gained an accurate knowledge about the situation and velocity of all particles in the Universe at the present moment, than he would automatically possess detailed knowledge about the entire future (and the past as well) of our Universe. In other words, it is enough in deterministic theories to know the initial conditions (location and velocity vectors of particles) and physical laws to have before one's eyes, at least in principle, the entire Universe from its beginning to the end, the Universe whose elements advance along determined trajectories, like trains along rails. No wander therefore, that in the XVIII and XIX centuries – when Newton's theory achieved its greatest triumphs – the Cosmos was compared to a regularly functioning clockwork mechanism.

The situation takes a different aspect in quantum mechanics. Here, we cannot know at **the same time** the position and momentum (a physical parameter being the result of velocity and mass) of a given elementary particle, for example an electron. If one has accurate data concerning the position of a particle, then the data concerning its velocity are so undetermined that it can in a moment find itself very far from the place in which it is presently situated, and inversely, the exact determination of the velocity of a particle makes it impossible to locate it exactly in space. The famous Heisenberg's indeterminacy (uncertainty) principle describes this correlation. What quantum mechanics describes with the so-called wave function is only a **probability** of finding an elementary particle, e.g. an electron, in different places of space at a given time moment. A free electron disperses in space in the form of a wave of matter, analogous to some extent to the wave formed on water by a stone striking its surface. The evolution of the spatial distribution of the wave function in time proceeds in a strictly determined manner. However, we cannot determine accurately (till the moment of the detecting the electron in the act of measurement) where such an electron is located. We only know that there exists a greater chance of finding the electron (if we detect it) at the peak of the wave, which corresponds to the greatest probability, than in the valley of the wave. Therefore, it is spatial distribution of the probability of finding it in just this and not any other place in space that "evolves" in the case of an electron moving in space[3]. One can conclude that it is not the electron itself that moves before being detected by a measuring device, but the chance of its detection in different points of space (where the probability of finding an electron was small before a moment, it can be very great now, and vice versa).

A similar situation takes place in the case of an electron bound within an atom. Here, its wave function has the form of an orbital. Orbitals (their shapes, sizes and spatial distribution) come into being due to the fact that the peaks and valleys of the wave related to an electron do not extinguish each other only for some strictly determined energies (and therefore wavelengths of the waves of matter, inversely proportional to energy). Instead, they intensify each other, leading to the formation of stable "structures". Speaking metaphorically, an electron "revolves" around an atomic nucleus in such a way that the "peaks" of the wave of matter related to this electron converge in subsequent revolutions. If a "peak" met a "valley", one would cancel the another out. The "circumference" of orbitals must correspond, in a sense, to the wavelength[4] of the wave of matter of a given electron (the distance between its subsequent "peaks"), due to which spatial structures in the form of orbitals are formed. These structures delimit the space volume with the greatest probability of finding an electron. Most frequently orbitals take the form of a sphere with the atomic nucleus in its center, or the form of two, four or six "tears" arranged in pairs opposite one another and connected by their sharp ends in the atomic nucleus. The area delimited by the surface of orbitals (presented graphically as such a surface) encloses the fragment of space where the chance of finding a given electron adopts a certain determined value, for example 50 % or 90 %; the greater the assigned probability, the larger the orbital to be drawn. Why do not we deal, however, with the boundaries of an orbital corresponding to the probability of 100 %, where the electron related to this orbital is located for certain? Because the area described by these boundaries covers not less, but the entire Universe. In other words, an orbital has no boundaries (they are used only for the purpose of graphical presentation). It can be imagined in the form of a small cloud, with highest density in the center (at least in the case of a spherical orbital), rarefying

[3] This state of affairs results, for example, in the fact that it is possible for an electron to pass **simultaneously** through two slits in a diaphragm on its way.

gradually as one moves away from the center, but nowhere reaching the level of zero density. Therefore, while it is very probable that an electron is located in the vicinity of the atomic nucleus, there exists a certain – although astronomically small – chance that it is anywhere in the Universe at a given moment.

Thus, both a free electron "moving" in space in the form of a wave of matter, and an electron bound within an atom and occupying a particular orbital (also described with a wave function), do not have a strictly determined position in space (until detected), while the probability of finding them in a given volume of space is determined by a local value of the wave function corresponding to each electron. A fundamental question arises here: is this delocalisation of an electron – i.e. its blurred presence in a certain volume of space (opposed to a strict location at a certain point of space) – a matter of our ignorance (and therefore an epistemological problem), resulting from our inability to determine the accurate location of an electron that does posses such a position, or is the delocalisation a fundamental (ontological) fact reflecting a "real" blurring of an electron, and not a derivative of our ignorance concerning some lower level of matter in the reductionistic sequence? Mathematical formalism of quantum mechanics is by itself unable to give an unequivocal answer to this question. This is the reason for the emergence of different physical interpretations of the existing mathematical structure.

To the present day, the most commonly accepted interpretation among physicists is the Copenhagian interpretation, elaborated by Niels Bohr, among others. It says that the Heisenberg's uncertainty principle is an ontological fact, and therefore the evolution of physical systems is fundamentally indeterministic. For it is not possible, if the position and velocity of each elementary particle is not determined accurately, to predict precisely their further movement in space, as well as their interactions with other particles (equally undetermined), and therefore to predict the future of the entire system. The claim that 90 % of electrons would behave in a given situation in a particular way says nothing about the behaviour of a particular electron.

An electron, however, may be precisely located at last by an act of measurement. At the moment an electron is detected, there occurs the so-called reduction of the wave function. Until then, the electron was blurred and the distribution of probability of finding it at different points in space was described by the wave function itself. At the moment an electron is detected by a measuring device (e.g. as a shining point on the scintillating screen), its position becomes, of

[4] Or to its multiplication.

course, strictly determined. It means that the probability of its occurrence at this point adopts the value 1, while the probability of finding it somewhere else is reduced to zero. The act of an accurate localisation of an electron is strictly related to the act of an experimental measurement. Until the measurement, an electron is in a state of superposition of different possible states (determined, for instance, by the position, momentum and the electron spin direction), while at the moment of measurement, the electron "decides to chose" one of these states. Therefore, while it was possible before the electron detection to say that there was slightly more of it at one place, and significantly less at another place, or that it had the velocity X in 20 % and the velocity Y in 5 %; after the detection one can attribute to it a univocally determined position and velocity. Pure chance decides which of the superposing states will be chosen. The only factor favouring some choices and limiting the probability of others consists in the probability of states determined by the wave function.

The dependence of the wave function reduction – that is of the choice of one of the superposing states – on the act of measurement arises serious philosophical reservations. It may not be hard to believe that the fact of superposition of different states (e.g. different positions in space) before the act of detection refers to elementary particles, including electrons (physics has taught us to adopt more unparalleled and intuitively absurd truths), for the question concerns the microworld, alien to our everyday experience. However, there do exist possibilities – at least theoretically – of building a bridge between the micro- and macroworld, which extends the application of quantum paradoxes onto the latter. The Schrödinger's cat is probably the best known example of the kind of thought experiment in question (Schrödinger was the "inventor" of the wave function). Let us imagine that we close in an ideally isolated box a lump of a radioactive substance, the atoms (their nuclei) of which can undergo a spontaneous decay and emit determined elementary particles at the same time. Such an emission is a quantum phenomenon and therefore an unpredictable process – it is a matter of a chance if and when a particle will be sent out. Let us put in the box a detector of these particles, connected to a mechanism that releases a poisoning gas from a vial at the moment of detecting such a particle. Finally, let us imagine a cat in this container. According to the Copenhagian interpretation, until the moment of opening the box (and performing of a measurement in this way), one would deal with two superposing states: a particle has been emitted and a particle has not been emitted. What is worse, there would also come into being a superposition of the following macro-states: poison in a vial – poison released, as well as:

alive cat – dead cat. Therefore, if somebody accepts the Copenhagian interpretation of quantum mechanics, he has to assume that the cat can be at the same time both dead and alive, till the moment of observation – the act that would force the cat to "decide" in an unequivocal manner in favour of one of those states.

As it accepts the dependence of a physical phenomenon – i.e. the wave function reduction due to an electron detection – on the act of observation, the Copenhagian interpretation of quantum mechanics seems to be highly unsatisfactory from the philosophical point of view, since a physical state constitutes here a derivative of a certain **psychological** process, for such is undoubtedly the act of observation. It means relegating to subjectivist positions the best theory describing the microworld we know of, which is equal to throwing away several-hundred-years' long tradition of successes in natural sciences, aiming at working out an objective description of the world. We had already had time to learn – at the cost of tens and hundreds of years of revolutions in our views of the world – that matter is independent of spirit, and sciences to the present day are proud of their methodology, independent of the prejudices and turbulences of our psyche. Now, the Copenhagian interpretation proposes not less, but resignation from this heritage obtained by hard toil.

Let us place a man instead of a cat in the container isolated from the rest of the world, and replace the poison, for both humanitarian and technical reasons (a dead observer is not much of an observer), with a cheering gas. The quantum phenomena, unpredictable in principle, will again decide whether the gas is released or not. Let us assume that there took place a radioactive decay during one minute, or it did not. Thus, the man in a box will **either** fall into a joyful euphoria, **or** he will **not**, but he will be aware of his state anyway (and therefore able to fulfil the criterion of a valid observer). He will therefore come into possession of the univocal information concerning the actual course of quantum phenomena, and therefore the wave function reduction will occur in relation to this man.

The situation will take a completely different aspect for another man, who is **outside** the container. If we assume (this is after all the condition of our thought experiment) that no signal escapes from the box, then the external observer will not gain information whether the gas has been released or not (and therefore, whether the quantum radioactive decay has taken place). Thus, the reduction of the wavelength will not occur in relation to this observer. For him, the physical state inside the box will correspond to a superposition of the following states: the gas is in the vial – the gas is released, and: the inhabitant of the box is serious – the inhabitant of the box is joyful. From the point of view of the man outside the box, the man in its interior will be, to some extent, both extremely joyful and mortally solemn **at the same time**. For each of those observers, completely different things will occur inside the box, and the "actual" physical state will appear to differ totally. It will be possible to determine this state only **in relation** to a particular observer, but it will become impossible, or even senseless, to speak about any state inside the box universal for the entire Universe and the observers inhabiting it.

This **relativity principle** included in the Copenhagian interpretation of quantum mechanics seems even much more radical than the "original" relativity principle in Einstein's relativity theory. In the latter, the relativity principle states that, for example, event A may occur before event B for one observer, while event B may just as well precede in time event A for another observer. In principle, however, one of the observers recording both events could, at least potentially, calculate – on basis of the known physical laws – how the events would present themselves for another observer, located at such and such position and moving with a determined velocity. The observer outside the box, however, will not be able to determine in any way the state of his colleague in the box, until he opens the box (and *vice versa*, if the complete isolation of the box means an impossibility of transfer of any signals in both directions). This seems to contradict the most elementary laws of logic, for a man cannot be **simultaneously** unequivocally joyful and, **at the same time**, find himself in the state constituting a superposition of joy and gravity. The above-presented picture demolishes completely our conviction concerning objective character of the real world, and therefore many physicists (and philosophers) regard the Copenhagian interpretation with a great scepticism, and treat quantum mechanics as a transitional theory, that will be replaced in the future by a fully deterministic and objectivistic theory[5].

[5] Some experiments confirmed recently the idea of **decoherence** proposed by Zurek. According to this idea, the quantum behaviour (superposition of states and so on) passes gradually into the classical behaviour, when we pass from systems involving small amounts of energy to greater systems, involving more energy. If we deal with a single, isolated electron, it preserves its quantum character, because it does not interact with its environment. On the other hand, electrons and atoms in more complex systems interact incessantly with other electrons (atoms), due to which entire system looses its quantum properties and adopts classical properties (constitutes one state and not a superposition of different states). The interactions with the surroundings are equivalent here to the act of measurement. In this case, however, the interaction or "measurement’ is a fully objective phenomenon and therefore it does not fall under the Schrödinger’s cat paradox.

The second – probably even more fundamental – reason why some scientists do not accept the Copenhagian interpretation as the ultimate expression of the laws ruling the microworld is simply the above-discussed immanent indeterminism resulting from this interpretation. The fact that, on its deepest level, the world is subject to laws that are statistical in their essence (determining only probabilities of certain states) can seem to be very difficult to accept for intellectual, aesthetic and philosophical reasons. Einstein – one of the most determined and certainly the best known antagonist of the Copenhagian interpretation – expressed this view in the famous statement that God does not play dice with the world. Many others followed him, sharing the opinion that the indeterminism of quantum mechanics is due to our lack of knowledge concerning some hidden parameters, describing in a deterministic way the behaviour of elementary particles on some deeper, hitherto unknown level of reality. Therefore, scientists have been attempting to formulate alternative and competitive interpretations of quantum mechanics.

The paradox that the behaviour of physical objects (elementary particles, in particular) depends on a their detection by an observer (an therefore, in fact, on a psychological act) does not enter into the interpretation of the mathematical formalism of quantum mechanics proposed by David Bohm. The most important property of this interpretation consists in its determinism. An elementary particle, for instance an electron, is here "carried" by the front of the wave of matter related to this particle. The propagation of this wave in space is viewed in the same way as in the Copenhagian interpretation. Here however, the location and velocity of a particle are accurately defined, and its behaviour – strictly determined. Thus, the condition of determinism and objectivism is fulfilled. Unfortunately, this is not the whole story. Although it manages to avoid the probabilistic approach as well as the above-mentioned paradoxes (Schrödinger's cat), Bohm's interpretation is not so simple and less elegant then the Copenhagian interpretation. It does not fulfil the Occam's razor principle, for it assumes the existence of (interpretative) entities that fall beyond necessity. These entities do not possess any rational justification, apart from being helpful in removing the paradoxes from quantum mechanics. For this reason, it is not generally accepted by physicists. It may be that the rejection is motivated by an additional factor, consisting in an unwillingness to return to the 19th century view of the Universe as a Clock.

However, we already know beyond any doubt that such a return is not possible and this knowledge does not depend on the assumed interpretation of quantum mechanics. The problem of quantum (in)determinism may be solved at last in the framework of a new theory, that will take over the place of quantum mechanics in the future, by reaching a deeper, ultimate layer of physical reality. However, even if it turned out that particular elementary particles behave in a strictly determined manner, sufficiently numerous **sets** of those particles – or atoms built of them – will be doomed to **practical** indeterminism. It means that such sets can behave in a completely unpredictable way (from our point of view), unless we are able to determine the position and velocity of each particle with **infinite** accuracy, which is practically impossible, for such knowledge requires infinite amount of information. Complex dynamic systems tend in certain conditions to behave chaotically, at least in longer periods of time. This second, macroscopic indeterminism (as opposed to a conceivable microscopic indeterminism of quantum mechanics that would concern single particles) can be named "**chaotic indeterminism"**.

The chaotic indeterminism comes into view when one passes from single elementary particles, atoms and chemical molecules to their sets containing billions of billions of billions… etc. of particles or atoms. As it is obviously impossible to follow the movements of each particle (atom), a statistical approach is applied to systems composed of many particles. This approach replaces the parameters describing behaviour of single particles (such as position, velocity or kinetic energy) with parameters characterising the state of the entire system, e.g. temperature or entropy (degree of disorder). Observation of such systems in evolution brings one to the conclusion that they behave chaotically. It is interesting, that the arrow of time – the unidirectional flow of time from the past to the future – appears only in macroscopic systems composed of many elements.

The arrow of time (i.e. the above-mentioned irreversibility of many processes) is not accounted for by the greatest two theories of modern physics – quantum mechanics and relativity theory. They describe the behaviour of single objects (elementary particles, heavenly bodies in the gravitational field) or of systems composed of a very small number of such objects at best (relativity theory – like its predecessor, i.e. Newtonian dynamics – faces enormous problems when trying to predict the evolution in time of a system composed of only **three** gravitating bodies, as its behaviour cannot be calculated in a strict, analytical way). The dynamic equations describing the movement of objects (translocation of bodies attributed with mass, expansion of a wave of matter) within the framework of the relativity theory and quantum mechanics are symmetrical with respect to time[6]. It means that if all particles in the Universe were stopped at some moment and set in motion with the same speed, yet in the opposite direction, then the Universe would begin to tend towards its past. It might not be surprising, on the one hand, to see the Earth revolving around the Sun in the opposite direction. On the other hand, however, such phenomena as rains falling upward or light rays emitted by our eyes, reflected by different objects to land finally in the Sun, certainly do not belong to the world we live in. In the light of the quoted great physical theories, these events are equally probable as the phenomena actually occurring in our world. Relativity theory and quantum mechanics are certainly incomplete attempts of a unified, homogeneous and universal description of reality on the physical level, for they do not account for what belongs to every day experience – the passage of time (speaking strictly: the irreversibility and unidirectionality of the flow of time).

There exists a branch of physics that deals with systems composed of a great number of atoms or particles and takes into account the unidirectionality of time passage. This is thermodynamics. It says that the entropy of an isolated system – its degree of disorder, or simply **chaos** – increases **with the passage of time**. The universal increase of disorder can, however, lead to formation of local organised structures, such as convective currents in a liquid in a heated-up vessel, or chemical reactions where colour changes cyclically, or finally the phenomenon of life. Chaos (more strictly: the tendency to maximise it) can therefore generate order.

A system behaving in a strictly determined manner is the quintessence of order. That used to be a generally accepted opinion until the discovery of fractals and "strange" (chaotic) atractors. The former are objects that look identical, regardless of the scale in which they are viewed (watching a fractal through a microscope one sees the same picture as with the naked eye). We are accustomed to see a line as uni-dimensional, a sheet of paper has two dimensions, while the sculpture of Venus of Milo is a three-dimensional object. Although it is difficult to accept it, the number of dimensions in fractals is not an integer. This is so, because the edge of a fractal drown on a plane tends towards infinitely great complication (twisting), when the observation tends to regard infinitely small dimensions. The edge of a fractal fills up a fragment of a plane (a two-dimentional object) in an infinitely dense manner. Hence, such fractals are objects with the number of dimensions between 1 and 2 (there also exist fractals with a greater number of

[6] In my opinion, if the quantum decoherence mentioned in the previous footnote was asymmetrical in relation to time, this could form a bridge between dynamics and thermodynamics.

dimensions, although this number is never an integer). Fractals might have remained just a brilliant toy in the hands of mathematicians (they are frequently presented as splendidly colourful objects), but for their kinship with "strange" attractors, that can – as it seems – describe many physical systems. An attractor is the "goal" towards which a given system tends (e.g. the hole in a sink is the attractor for molecules of water in it). A "strange" attractor (it shares similar mathematical formalism with fractals) is characterised by the fact that an infinitely small external disturbance can set the evolution of a system on a completely unpredictable course, although the mathematical description of this course is deterministic. This applies to a great extent to meteorological phenomena. The so-called **butterfly effect** is a good example of the principle, which allows a movement of the wings of a butterfly in Japan to cause a hurricane in Florida. One can conclude therefore, that order can generate chaos.

This is the so-called deterministic chaos. It means that the behaviour of chaotic systems (atmospheric phenomena, turbulent whirls in liquids, etc.) is ruled by laws that are **in principle** deterministic. **In practice** however, the non-linear character of the dynamic equations describing such systems results in their extreme sensitivity to even the smallest differences in the initial state of the system (this property is responsible for "turbulent" behaviour of a system, due to an extremely quick growth and expansion – generated by way of auto-stimulation – of even the most microscopic disturbances and fluctuations, such as whirls in liquids)[7]. As it is obviously impossible to gather data on any system and perform calculations on their basis with infinite accuracy, one is doomed to face practically complete unpredictability of the further evolution of such a system, if sufficiently long periods of time are considered.

This type of unpredictability is perfectly applicable to biological evolution, where a minute accidental disturbance (mutation in the genetic material) can be enormously "enhanced" and result in commencing the line of a new systematic (taxonomic) group, for example. Each pair of organisms living presently on the Earth has got a common ancestor (this is equally true with respect to a lizard and a snake, as well as to a man and a cauliflower). Further evolution of the progeny of such a common ancestor was directed by nothing more but single mutations, that resulted in putting some descendent individuals on the course towards plants, for instance, while

[7] In linear systems, a small disturbance of the initial conditions causes a small change in further evolution of the system. On the other hand, even the smallest disturbance in non-linear systems leads to a completely disproportionate effect, which puts the evolution of the system to a totally different course.

others evolved towards animals. A mutation in a single gene could cause unicellular organisms to refrain from separating completely after a division, which could open the way for the origination of multicellular organisms. The emergence of great systematic (taxonomic) units in the process of biological evolution could be therefore initiated by microscopic disturbances in the system (mutations). The entire structure of each living organism is – to a large extent – a record of billions of such single "decisions", taken in the accidental way in the past.

The present form of particular groups of plants and animals is – to a huge extent – the work of chance. It is possible to imagine billions functional schemes of a living organism structure, equally good or even more efficient than those represented among organisms on our planet (e.g. insects, chordates, germ plants). Only few of them, however, have become "the elect" – principally by way of accident and natural selection – during the natural history of life on the Earth. Biological evolution is therefore in its nature a chaotic process, such as meteorological phenomena. Its course cannot be predicted for longer time spans, and the smallest disturbance can set it on a completely different course.

It was possible for the arrow of time (irreversibility), chaos (unpredictability) as well as – on the other pole – spatially and temporally organised structures (to be discussed further on) to gain presence in our conceptual apparatus thanks to a passage from dynamics (mechanics) to thermodynamics, that is from single objects to systems composed of many objects. This forced the deterministic approach to be replaced with an essentially statistical theory – thermodynamics, which treats a macroscopic state as the result of averaging microscopic states of a huge amount of particles. In spite of the presence of the arrow of time within its framework, thermodynamics does not, in principle, contradict dynamics, that is e.g. quantum mechanics or relativity theory. The latter two theories are simply **incomplete** in their descriptions of reality on the so-called physical level[8]. **Something entirely new** emerges when in the passage from the behaviour of single objects to their sets. The "something" simply escapes the mathematical structure and conceptual apparatus of dynamic theories (again: quantum mechanics and relativity theory). Similarly – to be shown in the second part of this book – the essence of the biological level escapes completely the system of concepts offered by physics. The above discussion, however, shows clearly that physics faces huge problems even on its own ground, when it attempts at a unified description of reality.

[8]In my opinion, physics is inherently unable to deal with **complexity** which constitutes as objective aspect of the reality as space, time, matter or casuality.

In fact, the so-called physical level (by "level" I understand a certain "layer of reality" that can be described in the framework of a unified system of concepts) appears to be an illusion, existing only due to historical reasons. The framework of physics comprises at least two levels: the dynamic one and the thermodynamic one. They are not mutually contradictory – the thermodynamic level constitutes a kind of an "superstructure", built over the dynamic level. The crucial fact is that it is **not possible to describe –** within the framework of mathematical formalism and concepts taken only from the dynamic level – the manner in which the former emerges from the latter. To the present day, this remains a mystery to a great degree. The emergence (along whichever route) of the thermodynamic level from the dynamic one constitutes a very important step on the way towards the biological level and the evolution of the living world. The insights in the irreversibility and unpredictability of biological evolution that are offered by thermodynamics have already been discussed. Now, it is time for a short remark concerning spontaneous formation of macroscopic structures, organised spatially and temporally (this subject matter will be treated in more detail in the chapter devoted to thermodynamic evolution).

Living organisms are undoubtedly such dynamic structures, highly organised in space and time. There is no need to convince anybody about their complexity. A man, for example, is composed of an enormous amount of molecules of various organic compounds with complicated structure and specific functions, such as proteins, fats, sugars and nucleic acids. They are building blocks of higher-order structures, for example ribosomes and chromosomes, and further – cellular organelles (nucleus, mitochondria, endoplasmatic reticulum), that add up to the entity called a cell. Various types of cells compose tissues, organs and systems of organs, characterised by various structures appropriate for the performed functions, to form finally the entire living organism. One of such organs, the brain, probably still remains the most complex structure in the Universe known to us. Many inanimate systems are structured as well; for example crystals, convective currents in liquids (e.g. in water boiling in a vessel), our planet and galaxies[9].

Formation of organised structures of living organisms (let us limit our considerations to this class for a moment) – either during biological evolution, or in the course of ontogenetic (individual) development – seems to remain in glaring contradiction with the universal tendency toward chaos, indicated by thermodynamics. The second law of thermodynamics states that

[9]I will show in the chapter devoted to the thermodynamic evolution about the origins of the structure of astronomical objects, existing thanks to the force of gravity.

entropy (the measure of disorder or chaos) must **grow**, or, having reached the maximal value, remains unchanged in each isolated system. The inevitable increase of entropy is precisely the principle responsible for determining the arrow of time and requires the majority of processes occurring in the Universe to be irreversible. The most surprising and interesting aspect of the principle consists in the fact, that this global tendency towards disorder can, however, generate order locally, which appears not only to avoid contradicting the principles of thermodynamics, but also to be an inevitable conclusion to be drawn after a deeper analysis of its rules.

The second law of thermodynamics refers only to isolated systems. Living organisms, however, are open systems, exchanging matter, energy and information with the environment. Only when viewed together with their surroundings, they constitute an isolated system. The law requiring entropy to increase in time applies in an obligatory manner only to such entity. As the development of the structure of an organism (e.g. during ontogenesis) is accompanied by an increase of the degree of organisation, and therefore with a decrease of entropy, the inevitable conclusion can be drawn that the entropy of the surroundings has to increase even more, so that the total change in the entropy of the system consisting of an organism + surroundings should also assume a positive value. And this is so, in fact. For, in order to grow and remain alive, a man, for instance, must oxidise different nutritive substances (e.g. glucose) with participation of oxygen, which yields carbon dioxide and water as products of the process. The increase of entropy during the process of oxidation of glucose is **greater**[10] than its respective fall due to the development of the body structure, accumulation of memory (information) in the brain and so on. Thus, the second law of thermodynamics remains valid, for local formation of complex structures appears not to contradict this principle.

The mere non-contradiction does not explain, however, the very origins of organised dynamic structures, such as convective currents, whirls in liquids or gases (e.g. cyclones), or living organisms. Why should production of entropy – that is dissipation of energy in the form of

---

[10]The arrangement of e.g. atoms in a six-carbon molecule of glucose is highly ordered and contains much information. If we know the location of one of those atoms, then we possess the information that in the nearest neighbourhood there are the remaining five atoms, arranged additionally in a given concrete configuration. A set of glucose molecules has therefore a low entropy. On the other hand, one-carbon molecules of carbon dioxide, formed in the result of combustion of glucose, are scattered randomly in space, and no correlation takes place between their spatial locations. In the face of such a chaotic distribution of molecules of carbon dioxide,

heat – lead to the formation, or even require the existence, of forms organised macroscopically in space and time?

Theoretical mathematical analysis of the laws of thermodynamics has led to the conclusion that if a system is near the state of thermodynamic equilibrium (its entropy is close to the maximal entropy), then the speed of the production of entropy (dissipation of energy) reaches the lowest possible value and no structures should be expected to appear in this situation. This, however, does not apply to living organisms, that "feed on" the order of their environment – far from the state of thermodynamic equilibrium and therefore characterised by entropy significantly lower than the maximal.

At greater thermodynamic gradients (e.g. differences in temperature), **dissipation** of energy proceeds in a turbulent manner, and not quietly and gradually, as it is the case in the states close to equilibrium. If the bottom of a vessel filled with water is warmed up slightly, the heat will be conduced from the bottom to the top layer of water by transfer of kinetic energy (velocity of movement) from the molecules located lower to the molecules located higher up. However, when the difference of temperature between the bottom and the top layer of water exceeds a certain critical value (different for different liquids and dependent on their viscosity), this mechanism of dissipation of energy becomes insufficiently effective. The simple conduction of heat is replaced by upward and downward convective currents. The upward currents are composed of molecules with high kinetic energy, gained by heating up the bottom of the vessel. These molecules themselves move upwards, instead of transferring their energy to the molecules located higher up. When on top, they give away their energy to the air (in the process of evaporation), and thus undergo cooling down and return in downward currents to the bottom of the vessel, where the entire cycle closes. This process of energy dissipation (called **convection**) through huge amounts of molecules organised into microscopic dynamic structures – convective currents – is much faster than conduction. An analogous situation is found in the case of water flowing through a pipe. At a slow speed, the flow of water has a unvarying character. When a certain limit velocity is exceeded, it becomes turbulent and there appear "structures" in the form of vortices. A convective current or vortex is initiated by a microscopic fluctuation – a disturbance that expands very quickly to "enormous", i.e. macroscopic dimensions. Once such a fluctuation appears, it

---

the knowledge of the situation of one atom of carbon does not give us any information as to the situations of the others. The entropy of the system adopts therefore a high value.

'organises' its surroundings in a proper way (by forming vortex or convective current oriented in a determined way – turned left or right, moving upwards or downwards). The system retains in this way the memory about its history, i.e. about the fact that there has appeared such and such fluctuation in this and not any other place. These facts can bring the association with the process of biological evolution, characterised by the dependence of large-scale histories on microscopic and unpredictable disturbances, as well as the ability to retain the memory of this history.

The fact that the above-described **dissipative structures** will form when the thermodynamic gradient (amount of energy that "requires to be dissipated") exceeds a certain limit value was predicted theoretically by Ilia Progogine on the basis of the derived by him properties of non-equilibrium thermodynamics. It can be concluded that, although thermodynamics claims that each isolated system tends to the state of the greatest entropy (the smallest degree of order), it also predicts that there can form highly organised structures on the way taken by a system to achieving this state. Surprisingly enough, the function of the structures consists in speeding up the growth of the general disorder. A bridge has thus been built between physics (especially thermodynamics) and biology, for life – from the thermodynamic point of view – is nothing else but one more process on the surface of the Earth participating in the dissipation of the energy of the visible electromagnetic radiation emitted by the Sun (this energy is finally dissipated in the form of heat radiation sent into the cosmic void). As the Earth dissipates sufficiently great amounts of energy in a unit of time (produces sufficient amount of entropy), various dissipative structures form on its surface, such as living organisms and atmospheric phenomena (e.g. cyclones). Whatever else life is, it is **also** a dissipative structure or, if somebody prefers, a complex of such structures, that can exist only in systems intensively dissipating energy (producing entropy), such as the surface of the Earth. However, its being a dissipative structure does not characterise **completely** the phenomenon of life (neither convective currents nor turbulences are alive, after all), yet it touches the heart of its essence. Therefore, while thermodynamics quotes a necessary condition (one out of many) for the existence of living organisms as dynamic structures organised in space and time, it does not formulate criteria that would be sufficient to define biological systems univocally. Such criteria must belong to the laws ruling a level of reality that is 'higher' (superior) with respect to the dynamic or even thermodynamic one, namely the biological level.

This will finish a brief review of various properties of physical evolution that are found at the sources of analogous properties of biological evolution. The major properties belonging here are: change in time, directionality of this change, specific time scale in which a given type of evolution occurs, irreversibility, unpredictability (indeterminism) as well as formation of dynamic structures, ordered in space and time. The above convergences have an important role, as they show that many inseparable properties of biological evolution find their roots in purely physical evolution, which proves the phenomenon of life to be deeply settled in the physical level of reality. On the other hand, this does not exclude the existence of the core of biological evolution – determining the essence of transformations in the living world – which **does not find an equivalent** on the physical level and cannot be reduced to it. The second part of this book will be devoted to its presentation. Selected aspects of thermodynamic evolution and their implications for biological evolution will be discussed in a more detailed manner in the last chapter of this part of the book devoted to physical evolution. Now, I will focus for some more time on the most spectacular, yet still purely physical evolution of everything that exists (in the physical sense), namely the entire Universe.

EVOLUTION OF THE UNIVERSE

For centuries, people were convinced about stability of the Cosmos. Such an opinion – probably equally obvious as the medieval belief in the revolution of the Sun around the Earth – derived its strength mainly from the psychological need for constancy in the world, and also from the observed invariability of the "sphere of fixed stars". This conviction was so strong that the "discoveries" of biological evolution proceeding by way of natural selection (Darwin) or of the evolution of the Earth's crust (e.g. the drift of continents – Wagener's theory) preceded in time the recognition of changeability of the Universe as a whole. The last of the discoveries took place in the 20-ies and 30-ies of the 20th century, due to the discovery by Hubble of the so-called escape of galaxies. A marvellous opportunity to predict theoretically the phenomenon of the evolution of the Universe was let slip by Albert Einstein, probably the greatest genius in the history of science. His general theory of relativity (1915) implied that our Universe is unstable – it should either expand or contract. In order to save the Universe from evolution, Einstein introduced to his equations the so-called cosmological constant, corresponding to a certain universal repelling force (negative gravitation). When Alexander Friedman demonstrated that even a Cosmos "equipped" with the cosmological constant is unstable and proposed a few basic models of the evolution of the Universe, Einstein initially did not believe those results, suggesting an error in calculations. Therefore, even his genius was not sufficient to overcome the orthodox patterns of thought and to accept the conclusions resulting from his own theory. It was necessary for Hubble to discover that galaxies (strictly speaking – their clusters) move away from each other the faster, the more distant they are (1929), in order to convince the majority of scientists that the Universe actually undergoes expansion. Ultimately, Einstein acknowledged the introduction of the cosmological constant as the greatest mistake of his life.

Today, the fact that the Universe evolves is no longer questioned by the majority of scientists. Cosmology (the science about the Universe) has become a vigorous, strongly developed branch of science, finding confirmation not only in theoretical models derived from the general relativity theory and quantum mechanics, but also supported by a broad range of observations. The next section will present briefly the most important data corroborating the fact the Cosmos is changeable in time.

Why do we believe the Universe to evolve?

The first and, as yet, the most direct evidence of the evolution of the Universe was Edwin Hubble's discovery of the **escape of galaxies**, that is the fact that galaxies (clusters of galaxies) move away one from another, and therefore all galaxies move away from all other galaxies. Moreover, the greater the distance between two galaxies, the faster they move away from each other (the proportionality between the velocity of escape and the distance is determined by the so-called Hubble's constant). It means that the entire space of the Universe expands, but the matter contained in it does not (if matter – and thus, the observer as well – were undergoing together with space the process of "swelling", then it would not be possible to detect this fact in any way, for both the expanding object and the "ruler" used to measure the expansion would change simultaneously!). Therefore, the dimensions of the entire Universe grow – if it is finite, or else, the dimensions of each finite part of it expand – if it is infinite. However, the above statement applies only to sufficiently large areas – the gravitational force overcomes the expansion of space in smaller scales of magnitude, and therefore such objects as planets, stars, galaxies and their clusters retain an approximately constant size. On the other hand, the distances between clusters of galaxies increase. The escape of galaxies is indicated by the reddening of their spectrum, that is a shift of the absorption lines of particular elements (e.g. hydrogen) towards greater wavelengths of electromagnetic radiation, resulting from the so-called Doppler effect.

The second important evidence for the evolution of the Universe is the so-called **background radiation**. Its discovery arose many controversies concerning the principles of awarding the Nobel Prize.

Therefore, I will venture a short digression at this point.

If you turn on a radio receiver and hear noises and sizzles, inform about this the Nobel Prize Committee of the Swedish Academy of Science as soon as possible. No doubt, this prize will be awarded to you. Additionally, you must only learn from some expert why your radio receiver does sizzle. You do not believe it? In 1965, while Penzias and Wilson were testing an antenna recording radio waves, that had served earlier for communication with satellites, they discovered a noise coming from all directions, which they could not eliminate in any way. Initially, they suspected that pigeons, leaving their excrement in the apparatus, were responsible. However, removing them (i.e. both excrement and pigeons) helped little. The noise *remained. Therefore, they stated that there existed the noise, interpreted this noise as a fact and explained that the noise is inexplicable. Which was the basis for awarding them the Nobel Prize several years later.*

What they interpreted as noise was just the famous background radiation, constituting the crucial argument supporting the Big Bang hypothesis and a marvellous confirmation of the models describing early stages of the evolution of the Universe. The models were constructed mainly by intellectual effort of Gamow, Dicke and Peebles. The latter two learned about the Penzias and Wilson's troubles with the device from an accidental telephone call and immediately provided a correct interpreted of the observations. Who actually "discovered" the background radiation? Does the essence of a scientific discovery consist in chance, or rather in gaining new knowledge on the basis of the knowledge already possessed? Noises in a radio receiver can be discovered by anyone, but only few are able to explain why the radio receives the noise.

I will finish the digression on this.

Background radiation is a microwave radiation that reaches the observer from all directions in a homogeneous form. It is the residue of an early phase of the history of the Universe, when it became transparent and the radiation was released due to a recombination of protons with electrons into hydrogen atoms. As mentioned above, its presence is predicted by theoretical models based on the physics of elementary particles, that assume the Cosmos to have been small, dense and hot at the beginning. Until now, no other reasonable explanation for background radiation has been found.

The fact that the Universe once possessed different properties than at present is also testified by the **appearance of its distant regions**. As light needs time to reach us from very distant objects, one sees them in earlier stages than the objects situated in our nearest neighbourhood; the observer watches them as they were billions years ago. Such distant objects may already have ceased to exist a long time ago, while the electromagnetic radiation reaching our eyes is only their residue. The more distant are the objects aimed at with telescopes, the deeper in the past of the Universe the observer reaches. If it was invariable in time, distant objects should not differ from those near-by. However, this is not the case. All quasars discovered until now are very distant objects, that existed at least a couple of billions of years ago. Moreover, the shapes of distant galaxies – observed through visible and radio waves – differ more from the shapes of the galaxies surrounding us, if the distance increases. One can conclude that matter in the Cosmos adopted different forms in the past than at present. This is equivalent to the evolution of the Universe.

The last piece of evidence, testifying to the evolution of our Universe is the **abundance of helium** in the Cosmos. All chemical elements heavier than hydrogen and helium are produced from lighter elements in nucleosynthesis taking place in star interiors. They can be thrown out into space afterwards in explosions of Supernovae. This mechanism, however, does not lead to releasing significant amounts of helium, and cannot explain the abundance of this element in the cosmic space (it constitutes about 23 % of all elements, while the rest is almost entirely taken by hydrogen). On the other hand, cosmological models describing the early stages of the evolution of the Universe, when enormous pressures and temperatures prevailed, account for the production of amounts of helium that are very similar to those actually observed. Therefore, it constitutes a further confirmation of the hypothesis that the Universe passed in its youth through a hot stage.

What does the evolution of the Universe consist in?

Generally speaking, two aspects can be distinguished in the evolution of the Universe: the geometrical aspect and the material-energetic aspect. The first one is related to the changes in the size of the entire Universe as well as the kind and degree of the curvature of the space filling it (more precisely: space together with matter and energy contained in it simply **is** the Universe). The second aspect refers to transformations of matter and energy, that adopt different forms, entering into ever new configurations.

Let us begin with a discussion of the geometrical aspect. As mentioned before, the Universe expands in the sense that, with the passage of time, three-dimensional space expands (like the surface of a balloon due to blowing it up), which is why particular clusters of galaxies move away from each other (like small spots painted on the surface of a balloon being blown up). On the other hand, galaxy clusters themselves, as well as galaxies, stars and planets, do not get enlarged. If the Universe expands, it means that once it was smaller. Moreover, we have reasons to think that gravitation restrains the escape of galaxies in time, which allows one to believe that the pace of the growth of the dimensions of the Universe used to be even larger. By going back in time, one comes at the end to a moment – about 15 billion years ago – when the dimensions of the Universe were very small (or zero), while the density of matter, pressure and temperature were very high (or infinite). This moment can be regarded as the beginning of the Universe. One of the subsequent subchapters will restate the problem what in fact ought to be understood by such a beginning.

Since the moment when each finite area of our Universe was of extremely small size (if the Universe is infinite – which we do not know – there is no sense to speak about the dimensions of it as a whole), the Universe expands, and the rate of this expansion diminishes due to the restraining influence of gravitational force, operating between conglomerations of matter that are moving away from each other (i.e. clusters of galaxies). Initially, the rate of the Universe expansion was huge, which is why the moment of its origination is called the Big Bang. The future of the Universe as well as its general geometry depend on the density of matter contained in it. If this density exceeds the **critical value**, i.e. the minimal density needed for bringing the expansion to a halt, the space of the Universe possesses a positive curvature and finite dimensions, and the Universe itself will find its end in the far future. The forces of gravitation will finally halt its expansion and, after some time, it will start to collapse faster and faster, to finish its existence in a terminal stage, very similar to the initial stage, as its dimensions will be very small, while density and temperature – very large. It means destruction of all formations, such as galaxies, stars, planets and living creatures. This scenario is represented by the bottom curve in Fig. 1.1. An analogy can be found in a stone thrown up from the surface of the Earth. If we give it an insufficient speed, it will cease to raise, stop its movement and begin to fall down, back onto the Earth.

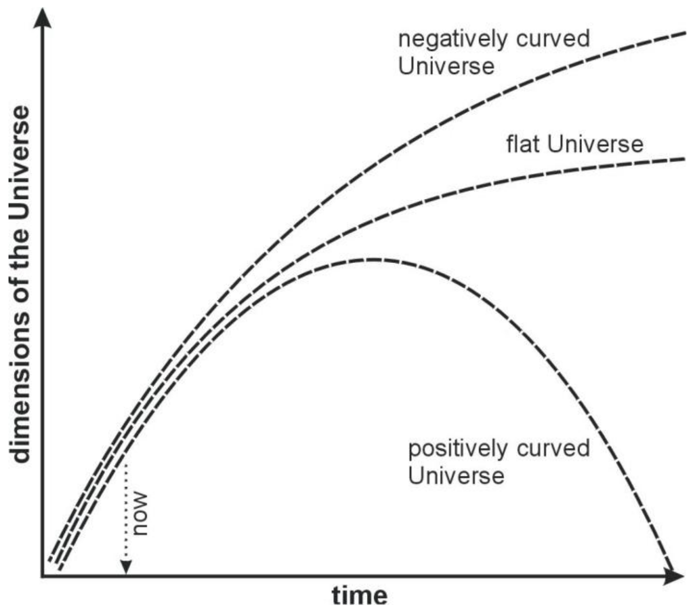

Fig. 1.1. Three models of the evolution of the Universe for three different average densities of matter: greater than the critical density (space curved positively), equal to the critical density (flat space) and smaller than the critical density (space curved negatively).

If the density of matter in the Universe is lower than the critical value, the Universe contains an insufficient amount of matter to stop the expansion (the force of gravitation is proportional to mass) and it will expand eternally. This is illustrated by the upper curve in Fig. 1.1. Similarly, if a stone is thrown up with a sufficiently great speed, it will never fall down onto the surface of the Earth, and continue infinitely moving away from it. The rate of escape will systematically decrease, tending asymptotically to a value greater than zero. The Universe will always expand with a finite velocity. In this model of the Universe, an infinite space is curved negatively and it ends (after hundreds billions years) in a thermal death due to burning out of stars and to the increase of disorder (chaos), measured with entropy (life would not be possible in such a universe either). To be sure, this conclusion meets some interpretative difficulties, since the maximal possible entropy of the Universe grows quicker than the actual entropy, but a broader elaboration of this subject matter goes beyond the frame of the present chapter.

If the density of matter is exactly equal to the critical density, the space of the Universe is infinite and flat. In this case, the Universe will also expand eternally, and eventually reach the "thermodynamic death", although both processes will be slower than in the case of the space with a negative curvature (the middle curve in Fig. 1.1). The rate of escape of galaxies will tend asymptotically to zero, and therefore it will be infinitely small in the infinitely distant future. Due to these similarities (infinite dimensions and eternal expansion), the models of the Universe with a negative or zero curvature (flat universe) are called "open", while the models with a positive curvature referred to by the name "closed" (the finite space "closes" here as the surface of a sphere).

It remains to be explained what is the meaning of the qualifications: positively curved, flat and negatively curved. The space of the Universe constitutes a three-dimensional formation, the curvature of which is difficult to imagine (although it can be strictly defined mathematically), which is the reason why two-dimensional analogies are frequently used to understand the problem. A flat surface is an ordinary, "flat" plane. It possesses infinite dimensions and two straight lines intersecting at some angle at one point on the plane will continue moving away at the same angle at any distance from the intersection. The same applies to flat three-dimensional space.

An example of a two-dimensional surface curved positively one finds in the surface of a sphere. It has finite dimensions (surface area) – a finite amount of paint suffices to cover it, although it lacks any edges or borders. Straight lines derived from a given point at some angle will display – in longer distances – a tendency to diverge at smaller and smaller angles, until they bend towards each other and finally meet like meridians on a globe (a "straight line" in a curved space is defined in a slightly different manner than a "normal" straight line known in common beliefs). If one moves "straight ahead" on the surface of a sphere, for example along the equator, one finally comes to the starting point. By analogy, a positively-curved three-dimensional space has a finite volume (a finite amount of paint suffices to **filling** it **up**), and at the same time, it does not possess any edges (limits) (such a space is finite, but unlimited). Straight lines originating at one point display a tendency to converge. A movement "straight ahead" from place finally leads to the starting point. This should be interpreted literally: in the Universe with a positive curvature of space-time, a space-ship setting off from the Earth and moving all the time away from it in the same direction, would come back to our planet after a sufficiently long time, but from the opposite direction!

A two-dimensional surface with a negative curvature finds its analogue in an idealised horse saddle (or a mountain pass) with its edges infinitely extended. Such a surface is of course infinite and two lines derived from one point at some angle display a tendency to diverge (bend outward) faster (at a greater angle), the further they move away from this point. Therefore, if it is difficult to get lost while travelling in the positively-curved space, because all directions tend to converge, then, even a small deviation from a route in the space with a negative curvature can lead us quickly very far from the intended destination. Of course, the effects of an eventual curving of space manifest themselves in much greater scales (the order of billions of light years) than those known from our every-day experience. Therefore, it would be very difficult to prove the existence of this curvature by direct observation, just as we do not perceive the curving of the terrestrial globe while standing on its surface.

It remains to be seen which of the above models reflects best the present state of knowledge[11]. All the three types of models of the Universe look very similarly in the present stage of evolution and in the scales of distance accessible for observation (all three curves in Fig. 1.1 run very close to each other at the point in time indicated as "now"). It is therefore not possible to determine the curvature of space by a simple measurement. However, it is possible to try to estimate the density of matter in the Universe and to deduce its geometry and future from the gathered data. Direct measurements of the mass contained in the Universe in all known forms of matter yield a density over ten times smaller than the critical density. This favours the model of an infinite Universe with a negative curvature of space, extending eternally as an already dead corpse, in which all processes have ceased. Such a future is not very comforting from the human point of view, which is why many scientists exhibit a strong preference in favour of the model of a closed (finite) Universe, positively curved, having its end in the Big Crunch – the counterpart of the Big Bang – a kind of *finis mundi*, much more picturesque than a cold death of the Universe. Though, it seems to be rather incomprehensible, that so many physicists enjoy this kind of philosophical preferences, for if they are right, life will sooner cease to be possible in a closed Universe than in its open alternative. Probably, they simply prefer – for purely aesthetic reasons – a crematory-like funeral of the Universe to the type resembling a "slow decomposition of a dead body".

On the other hand, the so-called conception of **inflation** – assuming a huge increase of the dimensions of the Universe during the first fractions of seconds due to a repelling (!) force of gravity in the so-called false vacuum – the only "component" of the Universe at that moment – implies that the present density of matter is very near the critical density, and therefore the space

[11]There exists a fourth possibility: some recent observations suggest that the rate of the expansion of the Universe even accelerates in time.

is (almost) flat. The latter two models face the problem of "lack of mass" that could increase the density of matter to the level assuring that the critical density is achieved or exceeded. Great hopes are deposited in neutrinos – weakly affecting elementary particles, traditionally regarded as devoid of mass – that unceasingly penetrate the space of the Universe. However, if it were shown that neutrinos possess a small mass, its sum total could suffice to "close" the curvature of space and restrain the expansion of the Universe[12].

It is time now to discuss shortly the material-energetic aspect of the evolution of the Universe, i.e. the changes in the form adopted by matter and energy that occurred along with the spatial expansion and were, in a sense, imposed by the expansion of our Universe, due to which there was more and more space for matter (and radiation) to fill in, which resulted in a decrease in density, pressure and temperature. As the behaviour of elementary particles and characteristics of their interactions depend principally on temperature (i.e. kinetic energy possessed by particles), the expansion and cooling off of the Universe caused mutual transformations of different forms adopted by matter and energy. Therefore, the growing dimensions of the Universe were – to some extent – the moving force behind these changes. What follows is a brief enumeratation of the most important stages corresponding to important transformations of predominant forms of matter and energy, accompanied by an estimated time from the Big Bang when a given transformation or state came into being.

a. The beginning of the Universe. The laws of the present physics do not reach the "very beginning". Therefore, we do not know what forms did matter and energy have at that moment, if the terms had any sense at all. In accordance with the so-called standard cosmological model, the Universe started from an initial singularity with zero dimensions and infinite density and temperature. It is not possible to say anything reasonable about matter and energy in such conditions. However, physicists believe that a future theory of quantum gravity will remove the unpleasant infinities, that undoubtedly constitute a disruption of the aesthetic elegance of the standard model. Next chapters will say more about possible shape of the very beginning of the Universe.

b. The first fractions of a second. There exists a "false" vacuum, which differs from the "normal" vacuum by a higher energetic state (other singularities of quantum vacuum appearing in quantum mechanics – apart from different energetic states – will be discussed further on). The

[12] Some recent observations indeed suggest that a neutrino has a non-zero mass.

"traditional" matter is lacking. The "false" vacuum exerts a negative (repelling) force of gravitation, which leads in a short time to an enormous increase of the dimensions of the Universe. This is the phenomenon of the so-called **inflation**. At a certain moment, there occurs something like a "phase transition" (analogous to freezing of water, i.e. to the transition from the liquid phase to the solid phase), due to which the "false" vacuum is transformed into the "normal" vacuum, while the released energetic surplus undergoes transformation into matter and radiation. From that moment on, the Universe expands at a much slower pace than in the phase of inflation.

c. First seconds. There exists a homogeneous "soup" of radiation and elementary particles mixed with their antiparticles (matter and antimatter). The following particles exist already: electron and proton (and their antiparticles: positron and antiproton) as well as neutron. Speaking more strictly, in such a high temperature, quarks, that combine in triplets to compose protons and neutrons, "swim" freely in the soup of particles. Radiation and matter are in thermal equilibrium, which means that pairs particle-antiparticle are created at the same rate as they are annihilated and transformed into energy (radiation).

d. First tens of seconds. All (almost) free neutrons join with a part of protons (that outnumber neutrons several times) to form nuclei of helium. Radiation is too "rarefied" (has too little energy) to maintain the creation of the particle-antiparticle pairs, due to which the reverse transformation predominates and almost entire matter undergoes annihilation by meeting with antimatter, which produces huge amounts of radiation. It remains in thermodynamic equilibrium with matter. As the density of matter is still sufficiently great, there occurs unceasing absorption and emission of quanta of electromagnetic radiation (photons).

e. 300 000 (three hundred thousands) years. The few protons and electrons that have survived annihilation recombine (join) to form atoms of hydrogen, while electrons and helium nuclei form atoms of helium. The Universe becomes transparent for radiation (the above-mentioned background radiation is released).

f. 1 000 000 000 (one billion) years. As a result of gravitational attraction, there form conglomerations of matter in different scales of magnitude – origins of stars, galaxies and clusters of galaxies. Planets built of elements heavier than hydrogen and helium are lacking, because these elements simply do not exist yet.

g. 5 000 000 000 (five billions) years. This is the time of evolution of the first generations of stars. Heavier elements are produced in their interiors in the process of nucleosynthesis and are thrown out into space during Supernovae explosions. Neutron stars (pulsars) and black holes come into being.

h. 10 000 000 000 (ten billions) years. This is the time of evolution of next generations of stars and of the formation of planets, containing large amounts of heavier elements. This is also the time of origination of life.

i. 15 000 000 000 (fifteen billions) years. The Universe achieves its present state.

What did in fact come into being at the beginning of the Universe?

The existence of the Universe (the existence of anything in general) can be opposed to nonentity, whatever the last term could denote. In other words, the origination of the Universe (that is the transition from its non-existence to existence) would consist in its emergence from nonentity, illustrated in the following scheme:

nonentity → Universe

To be sure, we have no certainty even as to the question whether the concept of nonentity has any well-defined, real sense, or if it is an empty, or even an intrinsically contradictory concept. There are good reasons to suspect that this concept is a consequence of the structure of the system of concepts in our brains (based on opposing contraries), rather than a reflection of any real aspect of the world. Leaving this problem aside, we will treat nonentity as a certain intuitively comprehensible philosophical term that will serve as a reference point for determining what the Universe is, as the opposite to its existence.

Common-sensical understanding expresses nonentity as the fact that "there is nothing". With reference to the Cosmos, this does not mean only that nothing fills an empty space, and therefore that an absolute vacuum devoid of matter prevails. Quantum mechanics teaches us that vacuum (the so-called quantum vacuum) is not so empty at all. It is characterised by its energetic states, and every moment witnesses spontaneous creation of virtual elementary particles in particle-antiparticle pairs, that vanish again instantly as a result of annihilation. Thus, the very concept of a total vacuum, a "pure" space containing completely nothing, is a product of our idealisation. It does not exist in nature and, moreover, nothing sensible corresponds to the concept. Although it can sound paradoxical, (quantum) vacuum possesses its structure and properties. It constitutes a dynamic formation, able to change in time, shrink, expand, undergo fluctuations *etc*.

Some scientists go so far as to claim that the entire present physics is in fact a science about vacuum, and that such objects as elementary particles or waves are simply a form of existence of quantum vacuum that grants masses to the particles created in it. Moreover, it may be vacuum that determines the fundamental constants of nature, such as the gravitational constant.

However, we have to be more radical in the quest of the true nonentity – not only for the above-mentioned reasons – and go further than the removal of matter and energy from space, leaving only empty vacuum, for it is also necessary to annihilate space itself (including vacuum, of course) as well as time. Only then will it be possible to say in a totally consequent manner that "there is nothing" (there appears, however, a subsequent question: can there exist – in the absence of anything real – at least the laws of nature alone, as a kind of platonic ideas; personally, I believe that the laws of nature are an inseparable part of the real Universe).

Speaking about any properties of thus understood nonentity verges on the absurd. For sake of argumentation, let us assume however (which seems to be quite a natural assumption) that nonentity has zero dimensions (as it does not allow the existence of space and time) as well as zero energy and mass (for, probably, obvious reasons).

On the other hand, the Universe consists of time and space, coupled into space-time, filled by a huge variety of formations: galaxies, stars, planets, moons, living organisms – generally speaking, matter in different forms. It is built of elementary particles, such as electrons, protons and neutrons. The latter two combine into atomic nuclei of particular elements, while electrons form their electron shells, composed of different types of orbitals that condition the possibility of joining atoms into various configurations, and therefore determine all properties of chemical compounds, minerals, living organisms *etc.* The picture is completed by different forces (gravitational, electromagnetic, strong nuclear, weak nuclear), and various forms of energy, e.g. kinetic energy, potential energy and so on. Therefore, when speaking about the existence of the Universe, one speaks about the existence of space together with different objects. It is objects – the building blocks of the Universe – that had to come into being, in one form or other, at the moment of the emergence of the Universe from nonentity. It turns out, however, that all this apparent variety can be ordered and unified in a simple way, i.e. all objects located in space (space-time) can be reduced to a consistent form, which will allow to determine in a more univocal manner what came into being at the beginning of the Universe and what its existence consists in.

First of all, various forms of energy are mutually equivalent in the sense that some energy forms can be transformed into others, while the entire amount of energy must remain constant (the law of conservation of energy). Therefore, different sorts of energy can be regarded as different manifestations of the same quality, while the system of categories in which we see the world – the system is a derivative of our brain mechanisms of sensory data integration, rather than of this world "in itself" – seems to be responsible for their apparent dissimilarity. Matter is a particular form of energy, perceived by us in a distinguished manner. We have good reasons to suspect, however, that matter and various forms of energy (for example energy of radiation) are simply forms of quantum vacuum organised in different manner or, if one prefers, different forms adopted by a certain universal physical field. The famous Einstein's equation defines the equivalence of ("traditional") energy and mass (which is an inseparable attribute of matter):

$$\mathbf{E = mc^2}$$

where **E** denotes energy, **m** – mass, and **c** – speed of light. Matter is therefore nothing more than a certain form of energy. Moreover, it is a **positive** (greater than zero) energy. It is attributed with the sign "+" within the framework of the accepted convention. The energy equivalent to matter (mass) contributes fundamentally to the overall energy of the Universe.

The energy resulting from interactions is another very important kind of energy in the scale of the Cosmos. Gravitational attraction is the only interaction important in large scales. To be sure, other forms like, for example, the force of electrostatic interaction between objects charged differently (attraction) or identically (repulsion) greatly exceed the attraction due to gravitational force. On the macroscopic level (in the scale of celestial bodies), however, all objects in the Universe possess an exactly equal number of positive and negative charges, so that the bodies are electrically neutral and their overall charge is equal to zero. Therefore, there is no electrostatic interaction between them. If, for instance, there existed at least a small 'surplus' of one charge in relation to the opposite charge on the Earth and the Moon, the electrostatic interaction between the bodies would exceed the force of gravity many billions times, while with the equal number of charges of both signs, the force of attraction between charges with opposite signs is ideally counterbalanced by the force of repulsion between charges with the same sign. As both the entire Universe and particular astronomical objects are perfectly neutral electrostatically, the electrostatic interactions are "turned off" (attractions cancels repulsion) in large scale distance, and the much weaker gravitational force gains prominence. It is this force that constitutes practically the only interaction contributing significantly to large-scale properties of the Universe.

The force of gravitation is correlated with **negative** potential energy. It can be remembered, that energy is an ability to perform work. If one lifts a stone over the surface of the Earth and lets it fall, the stone will be able to perform some work, for example to break a sheet of ice. The higher the stone is lifted, the greater work is it able to perform, as its negative potential energy is greater (more negative). Of course, even a stone lying on the surface of the Earth possesses an ability to perform work. If a vertical shaft in our planet were bored, the stone would fall down to the center of it. Generally speaking, two masses attracting each other would have a zero gravitational energy only when they were situated at the same point. To move them away from each other against the gravitational force, it is necessary to use some energy, which, as the distance increases, is transformed into the potential energy. An appropriate example can be found in the potential energy of a stone hung above the surface of the Earth. Therefore, the negative energy of gravitation would come into being as a result of pushing bodies attributed with a mass apart **in space**.

It turns out that the entire varied "furniture" of the Universe can be reduced to two basic categories of "entities": matter with positive energy and space linked with the negative energy of gravitation. Now, we come to the decisive step in our reasoning, for it also turns out – at least due to some estimations and I decidedly sympathise with them – that, probably, **the positive energy equivalent to mass is exactly equal to the negative potential energy of gravitation**. Therefore, these energies counterpoise each another and **the total energy of the Universe is zero!** This rather well established (as far as I know) possibility applies both to the entire closed (finite) Universe, and to each sufficiently large region of the open (infinite) Universe. Earlier in the text, it was intuitively accepted as obvious that nonentity entails also lack of any energy. Therefore, if everything that exists (matter and even space, in a sense) constitutes a certain form of energy, and the total energy is equal to zero, then the following question arises: what did in fact originate at the moment when the Universe emerged from nonentity? Does it make any sense at all to state if the Universe exists?

It is possible to try to give different, more or less speculative answers to this paradoxical question. I will propose below one of them, accompanied by a warning that it has a purely speculative character, since physics at the present stage of development cannot say anything reasonable about the problem under consideration. If nonentity is assumed to be a point of zero dimensions (lack of space), zero mass (lack of matter) and zero total energy, then, the origination of the Universe would be equivalent to splitting of nonentity into matter (mass) with positive energy and space as a "carrier" of the negative energy of gravitation. One can find an analogy in the formation of pairs of complementary colours (such as red-green, yellow-violet or orange-blue), due to "splitting" of whiteness – complete "absence of colours" (white colour is not treated in such a formulation as a colour). Generally, white light contains potentially all the colours of a rainbow, and they can be extracted, for instance, by means of a prism, splitting white light into the entire spectrum of colours – from red to blue-violet. For such an operation to become possible, however, there must exist a certain deeper layer of reality than that governed by the laws of our vision. In this case, it is possible to split the nonentity in the domain of "colours" due to the existence of a determined level, lower than the (phenomenological) level of colours, namely the physical level. In physical description, what we subjectively perceive as colours, corresponds to electromagnetic waves with a particular length. Is it possible to compare the origination of the Universe to the stratification of some white "colour nonentity" by means of a certain "prism" into particular colours, corresponding to space, matter, energy, interactions, elementary particles, radiation, stars and galaxies? If it is, then, is a certain deeper, yet unknown level of reality responsible for this?

Immediately, there arises a question, what is it – apart from the sign – that differentiates positive energy from negative energy, that is mass from space. It is known, that the differences in the perception of particular colours result from different stimulation of visual cells in our eyes by electromagnetic waves with a different wavelength. In the case of the Universe, however, it is not possible to say – without any knowledge of the postulated deeper layer of reality – why nonentity stratified into such and not other beings. It may be a metaphysical problem, and as such, it is nonsensical in the light of any scientific methodology. There may actually exist its solution – in the sphere of "things in themselves" – but we will never reach it, because of the limitations of our cognitive apparatus formed during the process of biological evolution. In this case, the "matter-space" differentiation would come as much from the real world as from the manner of processing signals from the world by our brain (problems of the type are discussed in a more detailed manner in my book *Absolut – odniesienie urojone* [Absolute – imaginary reference], Oficyna Literacka, Krakow 1994). Perhaps, that deeper level is represented by a relatively simple arithmetic formula, that "generates" the entire Universe, as in the case of another simple mathematical procedure generating the immensely complicated Mandelbrot's set – close relative of fractals.

We do not know. In my opinion, however, space cannot exist without matter (and *vice versa*), nor can plus without minus or red colour without green colour. These phenomena are, in a sense, opposites. They originated via stratification of nonentity and may annihilate into nonentity after some time (if the Universe is finite in space and time). The force of gravity seems to be the factor that separates and joins them at the same time (gravitation is identical with the curvature of space, according to the general relativity theory). The existence of mass – constituting both the source (subject) of this force and the object of its action – is inseparable from the existence of space – not only the stage for gravitational interactions, but also their "carrier" – that is the condition for the concept of gravitation to make any sense. The existence of space as the "carrier of gravitation" is inseparably correlated with matter, while the existence of matter as a set of masses attracting each other requires the presence of space. Therefore, both space and matter represent two aspects of the same thing, and it is logically impossible to separate them. This situation resembles two hands drawing one another in the well-known Escher's drawing, or two men, each existing only in the dream of the other. It is obvious that either they must exist together, or nonentity prevails.

Up to now, the conclusions from the above considerations have been rather negative, i.e. they indicated that the Universe lacks any principal properties differentiating it from nonentity. At least, we are unaware of such properties. However, in my opinion, there exists a certain feature of the Universe that certainly cannot be attributed to our philosophical concept of nonentity, and constitutes a sign-post marking the direction that leads to understanding of the essence of the Universe. The crucial feature is **information**. Information about the distribution of positive and negative energy, that is matter and space. This is the point where mechanics (dynamics) that ignores information could be unified[13] in the future with thermodynamics that includes information as its important component.

A plain sheet of paper does not contain any information at all – its information is none. However, if some formulas or symbols are written on the sheet, there will appear a certain amount

[13] Now, dynamics and thermodynamics cannot be co-ordinated, they are contradictory with each other (since, for example, thermodynamics takes into account the arrow of time, lacking in mechanics). I will say more about the differences between dynamics and thermodynamics in the next chapter, devoted to the thermodynamic evolution.

of information. The more information will be there, the less chaotic and more ordered will be the contents of the sheet. If a thousand points is scattered at random over the paper, the operation will possibly result in a system with the smallest amount of information for this number of points. We can increase the amount of information in a simple way, for instance, by grouping all the points on one half of the sheet. On the other hand, if the points formed a drawing of a human face (lines and spots also constitute sets of points), then the informational content of the sheet would reach a very high value. The system "a-sheet-with-one-thousand-points" can therefore contain different amounts of information, ranging from the minimal value of information (chaotic, random distribution of points) to its maximal value (points distributed in an extremely ordered manner).

On the other hand, if a single point is taken instead of a sheet of paper, then, there exists only one manner for the point to be "arranged" – one state which it can adopt. The minimal amount of information for a single point is at the same time the maximal value, and both values equal zero, if this case allows to define reasonably an amount of information at all. So, if the sheet with one thousand points is compared to the Universe, then, nonentity corresponds to a single point (yet it should be stressed that – due to the existence of gravitation – the smallest amount of information in the Universe is possessed not by a set o masses randomly scattered in space, but by masses concentrated in one place). The Universe differs from nonentity, like the sheet with a thousand points from a single point, by the amount of information – both the present and the maximal value (the greatest it can potentially assume). It is so, because there exist two kinds of "beings" on the sheet (in the Universe), namely, the white background and black points (space and matter), which allows them to be combined in different ways. In the case of a single point (nonentity), one can speak either about only one "being", or about no "beings" at all, depending on the accepted convention. The existence of information in the Universe seems therefore to be inseparably correlated with its differentiation into various types of objects. However, it still remains a mystery where these differences come from and what their essence consists in.

Concluding this part of argumentation, it can be claimed that for the Universe to come into being was to some extent identical with the origination of information concerning the distribution of positive and negative energy, that is matter and space. The Verb did not become the Body. If the Universe exists primarily as information and not energy (including matter), until today, it remains first of all the Verb. The course of reasoning followed in the present subchapter is briefly recapitulated in Table 1.

Tab. 1. Summary of the considerations concerning the essence of the existence of the Universe.

What has in fact come into being together with the origination of the Universe in the Big Bang?

nonentity →	**Universe**
there is *nothing*	**there exist: *space*, *matter*, *energy*, *interactions***
	equivalence of matter and energy **$E = mc^2$** ***matter = positive energy***
	+
	the most important force in the scale of the Universe: *gravitation* ***gravitation = negative energy***
	withdrowal from each other of masses in *space* ↓ **potential energy of gravitation**
	sum: **positive energy of matter = negative energy of gravitation**
ENERGY = 0	**ENERGY = 0** **The total energy of the Universe is zero**

origination of the Universe:

♥ **stratification of 'nonentity' on positive energy (matter) and negative energy (space = distance between bodies in the gravitational field).**

♠ **appearence of INFORMATION related to the separation of positive and negative energy**

At the end, I would like once more to stress that, at least in my opinion, the existence of the Universe consists in a "stratification" of nonentity into two opposite beings, signs or "senses"; since space acquires some "sense " only when related to matter, and vice versa. The sense or essence of matter and space co-define each other to some extent on the way of connotation. This is of great significance, because – as will be seen in the second and third part of this book – on the biological and psychological levels, "senses" arise in a similar manner. In living organisms – that, functionally speaking, are constituted by certain systems of regulatory mechanisms, ensuring survival and production of offspring – one regulatory mechanism is invested with a "sense", a "purpose" or a "meaning" only in relation to the remaining mechanisms (mainly negative feedbacks). The meanings of concepts within the framework of a conceptual network – in my opinion, constituting the "substance" of psyche – also are realised by way of connotation, i.e. in the process of mutual definition of some concepts by others, or, in an equivalent terminology, by a stratification of opposed meanings along semantic axes. This amazing parallelism of the essence of the physical, biological and psychological level – although it still remains to a great extent an enigma – seems to indicate some very important and fundamental property of the real world.

The property of a certain system consisting in being self-directed on itself is the second (postulated by me) among the elementary properties applicable to the emergence of the biological level from the physical level, as well as of the psychological level from the biological one (this will be discussed in detail further on). If this analogy were transferred onto the emergence of the physical level from nonentity, this property of being self-directed on itself could be translated into a recursion (self-reference) of a certain mathematical function (recurrence formula) that "generates" the Universe (compare the procedure generating the Mandelbrot's set). Although this supposition should be considered as a pure speculation at the present stage of knowledge, it may be taken as an indication as to where to look for the true "theory of everything", capable of explaining (preferably: generating) the fundamental properties of the Universe on the physical level (e.g. the masses of elementary particles).

How did all this begin?

The proceeding part discussed the problem **what** in fact originated together with the appearance of the Universe. Now I will try to analyse the question **how** our Universe happened to have a beginning and whether we can attribute any well defined sense to the concept of the beginning. Scheme 1 in Fig. 1.2 presents symbolically the question of the beginning of the Universe. The passage of time (t) is represented here by the horizontal axis, while one of the three spatial dimensions (x) is represented by the vertical axis. We can see that the size of the Universe along dimension x (and of course along the remaining two spatial dimensions y and z) increases in time, while the speed of the expansion of the Universe, very large at the beginning, slowly diminishes afterwards. On the other hand, we do not know what in fact initiated this expansion, which is indicated symbolically by a question mark in the place which we would be inclined to define as the beginning.

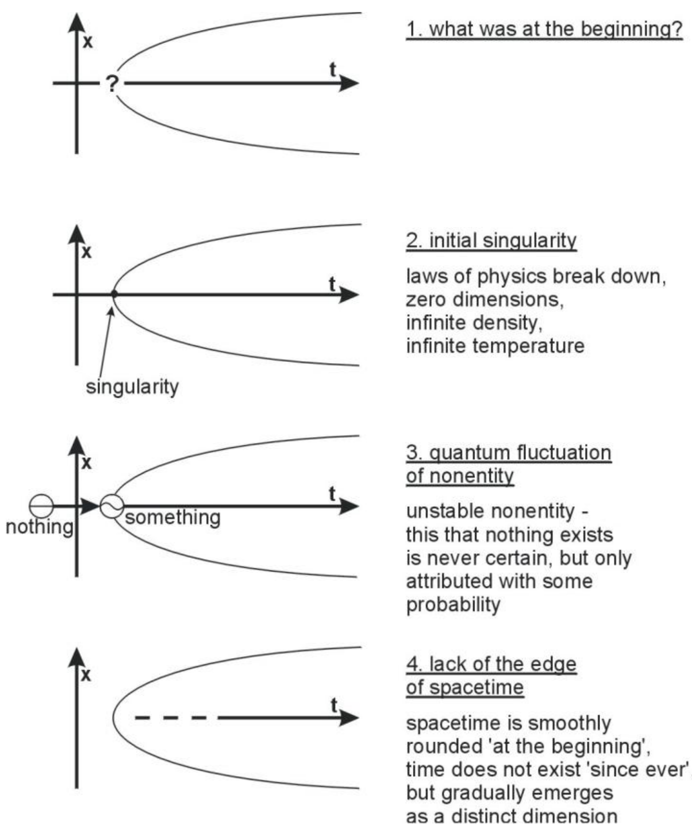

Fig. 1.2. Graphical presentation of the question: "how did the Universe originate" and of three proposed answers to this question.

It can be remembered, that the fact of the expansion of the Universe has led to the conclusion that the Universe was very small about 15 billions years ago. If Einstein's general relativity theory, describing the force of gravitation in terms of the curvature of space – the best physical theory describing phenomena on the macroscopic level – is used to describe the overall geometry of the Universe (what is more macroscopic than the entire Universe, after all!?), it appears that by moving back in time one comes at the end to the so-called initial singularity, that is to the state of zero dimensions and infinite density and temperature. Similar singularities exist now at the center of black holes, and the Universe – if it is curved positively – will finish its existence, after the phase of contraction, also in the stage of singularity, the so-called terminal singularity. All the known to us laws of physics cease to apply in such singularities. The evolution of the Universe from the stage of the initial singularity is described by the so-called **standard cosmological model**, based on the general relativity theory. This was presented in scheme 2, Fig. 1.2, where the Universe starts its existence from a single point of zero dimensions.

Due to the infinities incorporated in them, the singularities are very undesirable elements of any cosmological scenario, which is why scientists tend to eliminate them. There keep appearing still new propositions and interpretations following this tendency. The already mentioned concept of **inflation** is one of those that follow closely the guidelines formulated by the general relativity theory. It is proved that the initial singularity must appear at the beginning of the Universe, if the force of gravity was always a positive force of attraction. On the other hand, the inflationary scenario postulates, that "false" vacuum exerted a negative gravitational pressure (i.e. it acted repulsively) in the first fractions of the first second of the existence of the Universe. Before the stage of inflation, the Universe could extend very slowly; moving back in time we would state that its dimensions tended asymptotically to zero, yet they never reached the value.

Conceptions tending in the direction of the so-called quantum gravity theory seem to be more radical. What is interesting, this theory does not exist yet, but its construction in the future is widely anticipated as very probable. Such theory would be formulated as a unification of general relativity and quantum mechanics. These two greatest theories of the modern physics – the first of which describes large-scale properties of the Universe, while the second deals with the laws ruling the microscopic world on the atomic and subatomic level – share one important disadvantage: they cannot be both correct at the same time. They simply contradict each another. Each one describes perfectly a certain aspect of reality, but the language of one theory cannot be translated into the language of the other. One of them (and most probably both) will have to be modified, if they are to be unified, i.e. turned into one coherent theory. Such a uniform and unified theory was given in advance the name of quantum gravity.

It is easy to understand the need of such a theory for the description of the earliest stages of the evolution of the Universe, when one realises that the Universe was so small and dense at the beginning that its description as a whole (the domain of the general relativity theory) was at the same time a description of the subatomic level (the applicability area of quantum mechanics). At some moment, the entire observable Universe (the area from which light has had enough time to reach us since the Big Bang) possessed dimensions smaller than those of an atom! Although the theory of quantum gravitation has not been created yet, scientists believe that they can predict some properties that should be possessed by this theory, which is the basis to attempt answering questions concerning the beginning of the Universe. In particular, they extend the principle of quantum uncertainty onto the entire Universe in its early stages.

According to quantum mechanics, it is not possible to determine **simultaneously** the values of certain physical parameters concerning elementary particles in some very small scales (characterised by the so-called Planck constant), for example the position of particles in space and their momentum (velocity multiplied by mass). This property is described by Heisenberg's indeterminacy principle, which allows quantum mechanics to talk only about **probabilities** of certain states, for example, the probability that a particle is found at a given point of space, or that it possesses a particular energy. Even vacuum (called here quantum vacuum) behaves in a very strange way within the range of uncertainty (i.e. in the scales studied by quantum mechanics). The vacuum witnesses continuous origination of particle-antiparticle pairs (so-called virtual particles), created spontaneously and vanishing immediately afterwards in the process of annihilation. The energy and degree of spatial curvature are not determined with an infinite accuracy in quantum vacuum either – their values undergo unceasing fluctuations. Therefore, quantum vacuum looks, in a sense, as the surface of boiling water.

The quantum indeterminacy illustrates a very important point, by providing an example at the level of elementary particles. It is not possible to say with complete certainty that there is nothing in a given place of space. Due to unpredictable fluctuations of quantum vacuum, such a statement can be attributed with a certain, sometimes very high probability, yet it is never exactly equal to unity. Quantum mechanics is not able, in principle, to describe the world in a deterministic way. The so-called wavefunction can only determine the probability that a state S_0 at a moment t_0 will be followed by a state S_1 at a next moment t_1. It states, for example, that if the chance of a finding a given electron at a given point amounts 51 % at the present moment and under such and such conditions, then it will be equal to 73 % in one billionth of second.

So, the present attempts at the theory of quantum gravity included endeavours to construct the so-called wavefunction of the Universe, determining the probability of the Universe being in state S_1 at a given moment, if it was in state S_0 a while before. The properties of such a provisional wavefunction of the Universe turned out to be very interesting. It shows that if one substitutes "zero" or "nonentity" for S_0, there exists a **non-zero** probability that there will emerge "something" from this nonentity at the next moment of time, namely space, matter and energy! The scenario of this kind can be called a **quantum fluctuation of nonentity**. "Nonentity" as a physical state is therefore unstable and it becomes finally inevitable for the Universe to originate from it. In this sense, nonentity appears to be a physically forbidden state, the existence of which is not possible, and even logically inadmissible as intrinsically contradictory. Therefore, there must exist "something", and this something turns out to be the Universe. Why just the Universe known to us, possessing strictly determined properties and not a completely different object? We do not know. Maybe the response to this question lies again in a deeper level of reality, yet unknown. Scheme 3 in Fig. 1.2 presents the concept of the emergence of the Universe as a quantum fluctuation of nonentity.

If the explanation of the problem of the beginning of the Universe is concerned, the radical Hawking and Hartle's proposition, known as the concept of lack of an edge (border) of space-time (related to quantum gravity theory discussed above) goes probably further than most. In its essence, this conception denies that the concept of "beginning" in its normal sense is sensible at all. The postulate of absence of any edge of space-time means that four-dimensional space-time bends smoothly in the region of its beginning, like the surface of the Earth in the neighbourhood of the north pole. Therefore, there are no favoured points there, such as the initial singularity in the standard cosmological model. Due to the continuous bending of space-time, time as such does not exist in the region of the "beginning", where it possesses the character of an additional space dimension, indistinguishable from the other three. Time "emerges" (the dimension corresponding to time acquires its specific character) only at a certain distance from the rounded "tip" of space-time, and this transformation proceeds in a continuous manner (scheme 4 in Fig. 1.2). Let us take a closer look at the concept of lack of the edge of space-time (I will present and interpret this concept according to my understanding of it).

Traditionally, time is regarded as a dimension fundamentally different from spatial dimensions. Its passage occurs only in one direction (the so-called arrow of time), while space possesses no distinguished ("privileged") directions. Another difference – more mathematical in character and therefore further removed form intuitive comprehension – consists in the fact that, during calculation of distances in space-time, the distances in space are taken into account differently than distances in time. Time constitutes, to some extent, a parameter that serves for describing changes occurring in space.

If we refer the traditional concept of time to the evolution of the Universe, it turns out that the increase of the dimensions of the Universe (of the space filling it) occurs in time, as it is shown in Fig. 1.3. A closed Universe in subsequent moments of time is represented here by one-dimensional circles. Their sequence creates the surface of a figure called paraboloid, which corresponds to curved space-time. As the space expands, the circumference (and diameter) of those circles increases. Moreover, nothing prevents the movement back in time, to the very beginning of the Universe, when its dimensions were zero, and therefore the corresponding circle was tightened to a point. In principle, it is possible to go even further and follow the passage of time (whatever this could mean) **before** the beginning of the Universe, until minus infinity. This is due to the fact that time is treated here as something in a sense **external** in relation to our paraboloid – it always runs (according to the convention used in Fig. 1.3) vertically, from top to bottom, in opposition to the horizontally positioned "circles of space". The above outline leads to various logical paradoxes, and among them to the paradoxes reflected upon by Kant (this philosopher maintained that time, and the Universe, could neither have a beginning, nor last eternally).

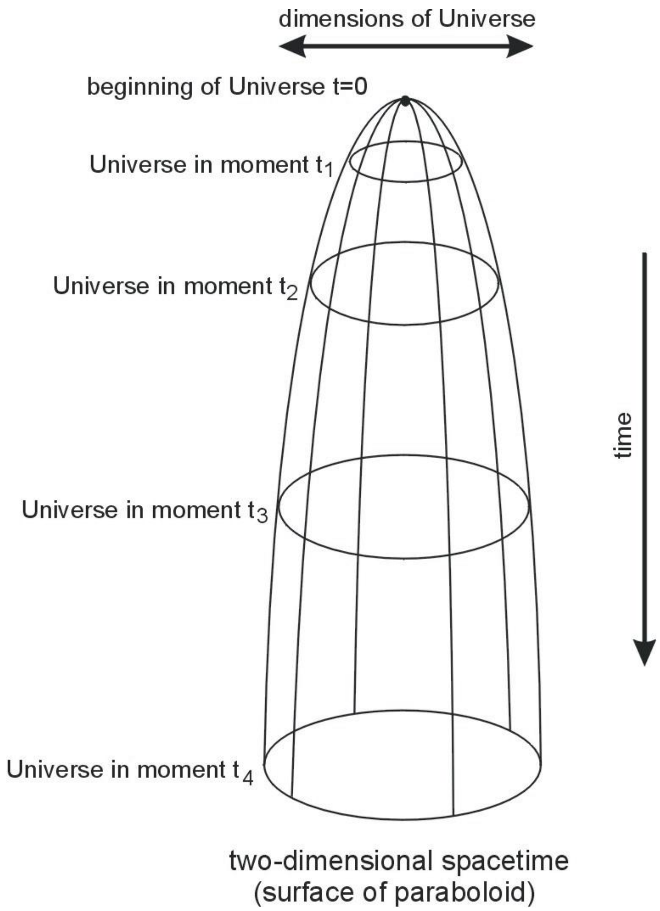

Fig. 1.3. Evolution of the Universe in time where time is independent of the Universe and flows "eternally". The Universe originates in a point at a certain moment of time t=0 and continues its evolution as an expanding one-dimensional, positively curved object – a circle. The two remaining space dimensions have been omitted.

What will happen, however, if time is considered to be an immanent part of the paraboloid? Let us assume that the surface of this paraboloid (two-dimensional by definition) "contains" two perpendicular dimensions: x and t, as it is illustrated in Fig. 1.4. We do not indicate in advance which one is spatial and which is temporal. Instead, we define the horizontal dimension as a dimension of spatial character, while the vertical dimension is assumed to represent time direction. The slanting dimension will possess properties that combine characteristics of both the spatial and temporal dimension. Assuming this new point of view, while moving down from the tip of the paraboloid, we will observe not the flow of time, but something, that can be called – more generally – the evolution of the Universe.

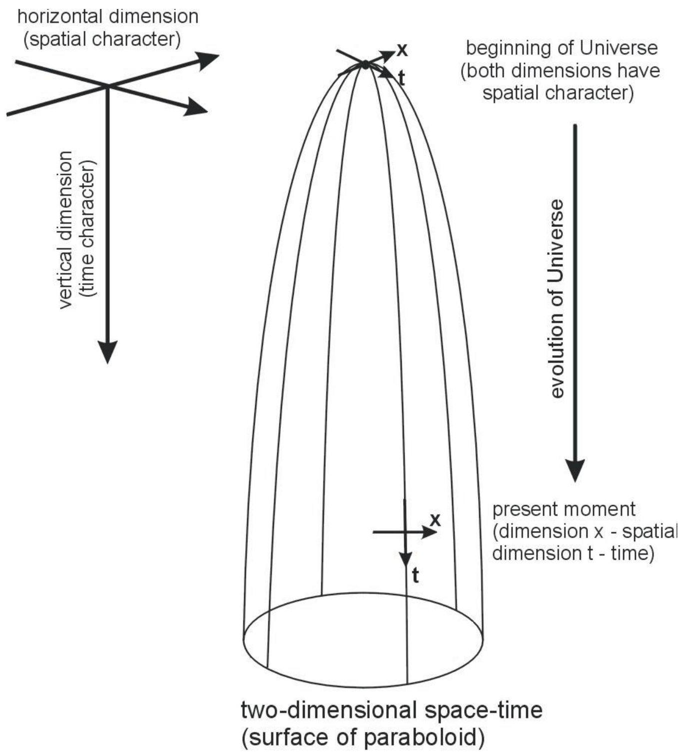

Fig. 1.4. Evolution of the Universe where time constitutes an integral part of the Universe and "begins" together with its existence. The space-time does not possess any "edge" in the form of the initial singularity, and the two presented dimensions (t and x) are of spatial character at the (smoothly curved) "beginning", since both are "horizontal". The further from the beginning, the more clearly one of the dimensions, t, becomes "vertical" and therefore acquires temporal character, while the dimension x (and the omitted dimensions y and z) remains spatial ("horizontal").

At the very tip, which can be identified with the "beginning" of the Universe, the surface of the paraboloid is exactly horizontal. Therefore, in accordance with the given definition, both dimensions x and t assume there spatial character. At the very "beginning", therefore, there is no passage of time! If one moves down the paraboloid along one of its "meridians", then dimension t becomes gradually vertical, acquiring temporal character, while dimension x still remains a spatial dimension, retaining its "horizontality". At a distance sufficiently far from the tip, t is decidedly a temporal dimension, different from the spatial dimension x (and its "siblings" – dimensions y and z). This is the situation we deal with presently.

The conception of lack of the edge of space-time does not allow to move back in time beyond the "beginning", corresponding to the tip of the paraboloid in the above analogy. Moreover, the nearer this "beginning", the more clearly time ceases to be time, and becomes space instead (its additional dimension). Simplifying the matter, one can say that the passage of time is slower closer to the "beginning", until it stops completely at the top! Because there was no time at all at the "beginning", the concept of a beginning loses completely its traditional meaning, and starts to refer only to a certain point of a smoothly curved space-time that does not stand out among other points. The situation is analogous to the case of the north pole on the terrestrial globe. This place differs by nothing special from other places on the surface of our planet, and the particular rank attributed to it comes rather from our way of describing reality, than from the reality itself.

The space-time devoid of an edge, where time – initially a spatial dimension – emerged only at some distance from the stipulated beginning, can be imagined as a certain general object, which simply exists and is everything that exists. Such a Universe is not hung in space, since it contains entire space. It does not exist in time either, since time is also its part. Its "evolution" would be an illusion due to the movement down from the tip of the paraboloid. This evolution would be equivalent to changes of space (and also of matter, energy and so on) in time, except for the neighbourhood of the very beginning, where time is not sufficiently well defined.

Problems with intuitive understanding of these "exotic" properties of the Universe – so alien to our common sense – come from the biological genesis of the human brain. The logic of our mind, the most fundamental categories of the human cognitive apparatus originated in the process of biological evolution, formed in the world where a univocally determined flow of time can be observed. The further we move away from our common experience – either in the direction of the subatomic world of quantum mechanics, or in that of large-scale curvatures of space described by general relativity – the less the structure of the "conceptual network" in our brain adheres to the structure of the world. There is nothing strange in this. Evolution formed our brain for effective functioning in the world surrounding us, and not for entering deeply into the mysteries of nature. Therefore the concept of time becoming space seems to us to be so odd and difficult to understand intuitively.

Why does matter exist at all?

Each elementary particle, except photon[14], possesses its antiparticle of the same mass, yet some of its properties are opposite, e.g. charge. The antiparticle of a negatively charged electron is a positron, attributed with an electrical charge of the same magnitude, but opposite (positive) sign. Similarly, to a positively charged proton, there corresponds a negatively charged antiproton. Generally speaking, particles are the building blocks of matter, while antiparticles form antimatter.

The properties of antimatter are, in principle, identical with the properties of matter, if one disregards the above-mentioned reversal of signs of electrical charges in antiparticles. If, for example, an atom of hydrogen – composed of a proton and an electron revolving around it – is replaced with an "anti-atom" by substituting an antiproton for a proton and a positron for an electron, then, the properties of this newly formed system should be indistinguishable from the initial system, for the elementary particle with a smaller mass and certain electrical charge will keep revolving around a particle of a greater mass and an opposite charge. The same should be expected in the case of the entire Solar System, together with the Earth and people living on it. The replacement of matter with antimatter should remain imperceptible.

This far-reaching symmetry between the properties of matter and antimatter stands in a flagrant contradiction with the fact that the entire observable Universe is, in principle, built of matter and not of antimatter (the latter can be obtained in small quantities in terrestrial laboratories). It seems natural to expect that there should exist equal amounts of matter and antimatter or that entire matter should annihilate after a contact with antimatter, which would fill the empty Universe with the resultant radiation.

To understand this apparent paradox, one must move back again to the earliest stages of the evolution of the Universe. According to the above explications, it was very small, dense and hot during first seconds of its existence. A highly condensed mixture of matter (and antimatter) and radiation filled the space, while remaining in thermal equilibrium, which means that there proceeded a very quick, bi-directional transformation of radiation into matter and antimatter, and vice versa. The rate of the creation (synthesis) of particles and antiparticles from radiation was exactly equal to the rate of the annihilation of particles and antiparticles that resulted in emission of radiation, so that the amount of particles, antiparticles and radiation remained unaltered in time. However, after some tens of seconds, due to the expansion of the Universe, the density and

[14] The antiparticle for photon in simply photon

temperature of radiation, and therefore its energy, became too low to allow effective production of particles and antiparticles. The reaction of annihilation started to predominate over the creation and, after a short time, the entire antimatter disappeared, combined with matter, and the energy related to their mass was released in the form of radiation.

What happened with matter? In the process of creation from radiation, there are produced exactly equal amounts of matter and antimatter, and therefore both should undergo complete annihilation, due to which an empty Universe should be filled today only with radiation. What was therefore the cause that triggered the existence of galaxies, stars, planets and ourselves? Where did the necessary matter come from?

It seems that the perfect symmetry between matter and antimatter accepted above is only an approximation, although an unexpectedly good one. The problem is that this approximation is not ideal. The properties of matter and antimatter differ very insignificantly; so insignificantly, that it is very difficult to notice these differences. Yet, they were decisive for the process of origination of the Universe in the form known to us.

Matter and antimatter are formed (in equal amounts) not only in the process of the creation from radiation. They can also transform one into another, where the so-called particles X mediate the transformation. The essence of this problem lies in the fact that matter forms from antimatter **slightly faster**, than antimatter from matter. This difference is so small that the advantage of matter over antimatter amounted only one part per ten billions (10^{10}) at the moment when the rate of annihilation commenced to predominate over the rate of creation. Therefore, only an extremely microscopic "surplus" of matter survived the process of annihilation with antimatter. The ratio corresponds more or less to that of one millimetre to ten thousand kilometres (one fourth of the circumference of the equator of the Earth). However, everything that exists – namely galaxies, stars and planets – has formed from this unimaginably small surplus.

We would be willing to expect that if the properties of matter and antimatter are almost identical, then, why in fact should they not be exactly the same? However, our intuition, our need for simplicity and aesthetics completely fails at this point. Fortunately. For we owe the existence of all sorts of matter in the Universe just to this small, unaccountable "defect" or imperfection in the ideal symmetry between matter and antimatter. It does not mean that further development of science will not demonstrate this apparent blemish in the ideal symmetry in the laws of the Universe to constitute a manifestation of some deeper logic, in its essence more consequent than the logic of man, conditioned by the structure of junctions between neural cells in his brain. This new logic will probably mean a general unification theory – a turning-point similar to the breakthrough of quantum mechanics and relativity theory that took place at the beginning of our century.

Apart from the case of matter-antimatter, probably the best known problem connected with the lack of symmetry in the Universe is the direction of time flow. Each of us experiences the passage of time in everyday experience. People are born, grow up, age and die, rains fall downwards, heat flows from places with higher temperature to places with lower temperature, the Universe undergoes expansion. In each of cases, the sequence of events is determined and cannot be reversed. We would instantly recognise a film replayed backwards. A view of a broken cup, forming spontaneously from pieces into one object and whole and flying up from the floor onto the table, would astonish with its unnatural character. Anyway, our psychological introspection – in which the future always follows the present, preceded in turn by the past – convinces us of unidirectional flow of time, that is of the so-called arrow of time.

It may seem strange in this context that the two greatest physical theories – relativity theory and quantum mechanics, being an unquestionable pride of the modern science – simply ignore the fact of the existence of the arrow of time! Mathematical equations appearing in them are symmetrical with respect to time, unable to the distinguish between the future and the past (the reduction of the wavefunction in quantum mechanics constitutes here an exception; this is, in my opinion, a conceptual monstrosity, as well as a flagrant evidence of our ignorance, rather than reflection of the actual state of things). It means that if all particles in the Universe were stopped and set going in the opposite direction with the same velocity, then the course of all events would become ideally reversed. We would observe apples "falling upwards" from the earth onto branches, hear thunders resulting in lightnings, while light rays would run out from our eyes and, reflected from different objects, land finally on the Sun.

The above contradiction – appearing to be an obvious absurd – originates from the fact that both relativity theory (together with its predecessor – Newtonian dynamics) and quantum mechanics were created for the purpose of describing single objects (such as celestial bodies or elementary particles), or at best, simple systems, composed of only few elementary objects. And actually, if we consider the Earth revolving around the Sun, or a system composed of a proton and an electron (atom of hydrogen), then the reversal of the direction of time passage would change nothing. The Earth revolving the Sun in the opposite direction would not be something in any way unnatural.

On the other hand, the arrow of time appears in thermodynamics, a physical theory describing mutual transformations of different forms of energy (e.g. mechanical, electrical, heat energy). The famous second law of thermodynamics states that entropy (the degree of disorder) grows with time passage. Thermodynamics differs from relativity theory and quantum mechanics first of all by the fact that it deals not with single objects, but with complex systems containing billions – I am saying this as a figure of speech – in reality much greater numbers are concerned – elements (e.g. atoms, chemical molecules, elementary particles). For obvious reasons, thermodynamics cannot analyse the fate of each of such elements. Instead, it represents a statistical approach, describing the average behaviour of elements in a system. It can be calculated, for example, that the mean velocity of molecules in a gas amounts x at a given temperature, while the per cent of molecules with a certain determined velocity is equal to y. Thermodynamics states also that in a system of two chambers filled with gas and connected by a pipe, the gas will flow from the chamber with greater pressure to the chamber with lower pressure, **regardless of the behaviour of particular molecules**. It means that it is possible to imagine **theoretically** such pattern of velocities and situations of molecules, that would result in the opposite behaviour, but the probability of such a pattern to occur is attributed with so extremely small number, that it is not worth taking it into consideration at all[15]. Let us notice that the flow of gas in time proceeding "against" the gradient of pressure is exactly identical with the "normal" flow of gas (proceeding from the place of greater pressure to the place of lower pressure) backwards in time, according to the normal direction of passage. Thermodynamics states therefore not less than that the occurrence of certain phenomena (balancing of the pressure of gas in the chambers connected by a pipe) is much more probable in the direction of the time passage than in the opposite direction. Thus, there appears a well defined arrow of time. The effects of this type are possible thanks to statistical approach, adequate in the systems composed of many elements. The emergence of the arrow of time in the transition from simplicity to complexity, from single elements to their complicated sets, constitutes undoubtedly a fascinating

[15] This probability is so small, that the time of the existence of the Universe is much to short to enable a realization of at least one such case in its history.

puzzle of the present physics. This problem will be dealt in more detail in the next chapter, devoted to thermodynamic evolution.

The problem of the arrow of time gains some clarification by the fact that even the time known from dynamic theories (as opposed to thermodynamic ones), i.e. quantum mechanics and relativity theory, is not ideally symmetrical, as it is usually assumed. It means that even such a simple system as an atom of hydrogen, composed of a proton and an electron, would behave in a slightly different manner, once the arrow of time is reversed. Moreover, everything indicates that the disruption of the symmetry of time is strictly related to the disruption of the matter-antimatter symmetry. It has been proved that the existence of one of those symmetries implies automatically the existence of the other. This connection is so deep that antiparticles are frequently treated as particles moving against the flow of time! As we remember, the disruption of the matter-antimatter symmetry was responsible for the existence of galaxies, stars and ourselves. As it appears, the same reason which enables the existence of matter seems to form the basis for the passage of time. The last conclusion should not, however, be overestimated, for the connection between thermodynamic arrow of time and the above-described disruption of the symmetry of dynamic laws seems to remain completely mysterious.

The situation is even more interesting in fact, because there exists an additional third asymmetry, strictly related to the first two, namely, the asymmetry "world – its mirror reflection". It consists simply in this that if the left and the right side of the Universe exchange their places (which would result in a mirror reflection of the Universe in relation to a certain plane), the properties of the Universe would change. This conclusion – shocking at first glance – has already found its confirmation in some studies on the radioactive decay, for more particles were emitted in one direction than in the opposite one. The differences between our world and its mirror reflection (until now existing only in theory) would not perhaps be so great, as it was described by Lewis Carroll in the book *Through the Looking-Glass, and what Alice found there* (the second part of *Alice's Adventures in Wonderland*), and certainly not so picturesque. We have to agree, however, that the realisation of an idea we would be inclined to regard as just a fairy tale for children raises the question if there exists in physics any limit fixed by the so-called common sense, which will forever remain forbidden to cross.

The three presented asymmetries (we know nothing about others) constitute a certain inseparable entity, expressed in that fact that the properties of the world will change, once mater and antimatter change places, or the right side of the world takes the place of the left one and vice versa, **or** the direction of time is reversed. However, if all these three operations are performed **together** (simultaneously), it will result in a world that does not differ by anything from the real world! Therefore, the correlation of the asymmetries appears to indicate some superior symmetry at some deeper level of reality. A triangle with three unequal sides lying on a plane can be used as a distant analogy. Such a triangle lacks any symmetry. If one performs on the figure one of the following operations, namely: mirror reflection with respect to a vertical axis, mirror reflection with respect to a horizontal axis or rotation around the intersection of those axes by 180 degrees, one will obtain a different triangle (in different position). On the other hand, if all the three operations are performed **together**, it will change nothing. The triangle will remain in its place (this can be checked easily). Therefore, while each operation taken separately changes the properties of the triangle (its position), then, the entire **group** does not produce any effect. A group of exactly this type is constituted by the three presented asymmetries of the Universe.

By way of summing up the section and providing at the same time a brief answer to the question posed in the title of the chapter, one can claim that the existence of matter (and of the flow of time, in a certain mysterious sense) is possible due to a very small imperfection (with respect to our aesthetic expectations) in the three symmetries of physical laws. Without this "defect", the Universe would exist as a timeless formation filled exclusively with radiation. This apparent imperfection, however, if examined more thoroughly, may turn out to be a manifestation of some higher perfection. To learn if it is so, one has to wait for further development of physics, and for the long-awaited "theory of everything" in particular (if it is possible for such theory to be formulated).

The content of the present chapter is briefly recapitulated in Table 2.

Tab. 2. Summary of the considerations concerning the reasons of the existence of matter.

Why does matter exist?

Early, hot Universe: two opposite reactions occurred with an identical rate:

$$\text{radiation} \xrightarrow{\text{creation (synthesis)}} \text{particle} + \text{antiparticle}$$
$$\text{particle} + \text{antiparticle} \xrightarrow{\text{annihilation (disappearence)}} \text{radiation}$$

Matter and antimatter originated in the same amount. Further evolution:

expansion and cooling of the Univese
↓
domination of annihilation over creation
↓
practically entire matter and antimatter vanishes

---

small superiority on matter over antimatter (1:10000000000)
↓
small part of mater undergoes 'salvation'

Reason: **insignificant difference in properties of matter and antimatter (matter transforms slightly slower in antimater than inversely)**

♣ **Existence of matter: the consequence of the break of three connected with each other symmetries of physical laws.**

1. asymmetry matter - antimatter: replacement of all particles with their antiparticles changes insignificantly properties of a system.
2. asymmetry space - its mirror reflection: replacement of right side with left side changes properties of the world.
3. asymmetry of time: reversal of run (arrow) of time changes insignificantly properties of particles.

♠ **We owe the existence of galaxies, stars, planets and ourselves to a "defect" in a perfect symmetry of our Universe.**

* * *

Cosmology offers us a very fascinating picture of the Universe as a whole, and particularly of its origination and evolution. At the same time, this picture seems to be very incomplete. We still do not know for sure how the Universe was formed, where its further development tends to, neither the existence of what we have in mind when speaking about the existence of the Universe. For centuries, the material, tangible essence of the "objective" external world was opposed to the fleeting, subjective and abstract "spiritual substance". We have been accustomed that, although man can be led astray by exuberant products of imagination and speculation detached from reality, the material world surrounding us will always remain the immovable bedrock of reason and cognition. Now it appears that the essence of matter (as well as of energy, space, time *etc.*) is equally ethereal, fleeting and mysterious as the essence of our psyche.

In fact, there is nothing strange in this. After all, we are not given access to the external world "in itself", in its immanent essence (although we frequently have such subjective feeling), but only to some sets of concepts, constituting a reflection, representation of some aspects of this world in our psyche. Therefore, what we call matter, time and space in physics is undoubtedly and to a great extent a collective construct of human minds, conceptual in its nature and belonging to Popper's third world. What differentiates the above concepts from entirely unrestricted creations of fancy, or from a great collective fiction or culture, is a certain correspondence between the "semantic structure" of these concepts and the "real structure" of the phenomena in the external world, established by common experience and (particularly) the methodology of sciences. This "adherence" of the "conceptual network" to the "things in themselves" cannot in principle be ideal. It is restricted by the structure of the neural network in the human brain, inherited in the course of biological evolution. Therefore, we are not universal cognitive machines with unlimited possibilities, as many scientists and philosophers have been inclined to believe (explicitly or implicitly) for centuries.

Scientific knowledge is not completely autonomous and is not created *de novo*. It is formed on the basis of the already-possessed knowledge, first of all the knowledge gained in common experience. However, nothing guarantees an adequacy of our system of concepts – formed during everyday activities, such as counting apples, gauging distances by means of a measuring tape, or measuring time with the aid of a watch – for the description of elementary particles or the Universe as a whole. The further away from the common experience – in the direction of small or great scales of distance, high densities and temperatures – the more exotic seem to be the laws of physics that are obligatory in these "regions" of reality. It should be noticed that even concepts so fundamental for our cognition as space, time, matter and causality do not necessarily come directly from the external world, for they are **also** derivatives of a certain concrete manner of integrating and ordering sensual stimuli in our brain, formed in the process of biological evolution for the purpose of hunting for mammoths, rather than for developing science and philosophy. Nonetheless, we trust that the mesh of our "conceptual network" will keep adhering better and better to the structure of the world, as a spider web sticks more and more tightly to a stone sculpture. The beauty of science consists also in the fact that we may never be able to expect all aspects of the world to become explained completely, and therefore the spider may always have something to do.

Various physical theories describing different aspects of the Universe constitute in a sense and are analogous to different projections of a certain spatial solid body on different planes. For example, the picture of a vertically-standing egg on a vertical plane is represented by an (egg-like) ellipse, while its projection on a horizontal plane is a simple circle. The projections of more complex solid bodies contained in spaces of more than three dimensions can be much more varied. The science composed of particular scientific theories attempts to re-create the picture of the entire solid body (the world) on the basis of only a few projections at its disposal, some of them being only superficial sketches. In time, science creates new projections, improves the already existing ones or discovers than what was regarded to be separate elements constitutes pieces of one greater, coherent picture. The final destination is to co-ordinate and unify all the pictures-theories into one entity and to obtain in this way an insight into the essence of the entire figure-world. This can seem impossible, either because of infinite complication of the world, or due to the limitations of our cognitive apparatus. Does it entail cognitive pessimism? Probably, it does not, for it is not the goal that is most important but the way leading to it.

The conditionings and limitations of our cognition (including scientific cognition) imposed by the basic structure of meanings in our brains (a result of biological evolution) will be discussed in the third part of the present book, devoted to the evolution of the conceptual network.

THERMODYNAMIC EVOLUTION

Essence of thermodynamic evolution

Terms such as entropy, degree of order, information and thermodynamics have already appeared on the pages of this book in the discussion of physical evolution in general, and the evolution of the Universe in particular. Nevertheless, the previous chapter focused especially on the evolution of the Universe, mutual interrelations between space, time, matter and energy as well as their dependence on a particular moment of the history of our Universe, while the rules of mutual transformations of energy forms, and also the related problems concerning process irreversibility and ordered structure formation – so important for understanding, for instance, of the physical basis of the phenomenon of life -- were treated marginally. The present chapter is an attempt to make up for these arrears. It outlines briefly the aspect of physical evolution of immense significance, namely the evolution in the domain of thermodynamics.

Many among the simplest aspects of the evolution of the Universe are by no means so obvious, as one could infer from the previous chapter. One should address at this point the following – far from trivial – question: why is it so, that the evolution of both the Universe as a whole and of its component objects proceeds just in this, and not in the reverse direction? This question takes a different form for an open (infinite) universe and a closed (finite, possessing positive space curvature) universe. An open Universe expands from the state of great (infinite) density, which leads to a decrease of this density in time due to ‘stretching’ of the space, whereas the opposite phenomenon – namely the contraction of space and collapse of the Universe from a stage of a very rarefied distribution of matter to a very dense terminal stage – does not take place. The models of the Universe based on general relativity – symmetrical with respect to time – do not give an *a priori* answer to the question why the first, and not the second case takes place (the relevant equations say only that the Universe is unstable, and therefore it should either expand or contract, whereas the choice of one of these possibilities is based entirely on experimental data).

There arises an even more important question: why does the evolution of matter distribution – from the initial stage characterised by a very homogenous matter (and energy) distribution towards formation of matter conglomerates in different scales of magnitude (galaxy clusters, galaxies, stars and planets) – take place in the expanding Universe? Black holes are the final (if we leave out the so called Hawking radiation unaccounted for), most condensed phase of matter transformations. They can further on merge with other black holes to form still larger and larger objects.

In the case of a closed Universe, the phase of expansion will be followed by the phase of contraction, which, in principle, could be an exact reversal of the expansion phase, finding its epilogue in the final singularity. Does it mean that not only matter density and temperature will increase toward their initial values, but also galaxies, stars, planets and living organisms will start to evolve in the backward direction, till they reach (just before the Big Crunch) the state of homogenous matter (and energy) distribution, identical to the state just after the Big Bang?

The last question can be reformulated in the following way: why do we observe the clear trend towards formation of different kinds of objects, such as galaxies, stars and planets, together with various dynamic structures developing on their surfaces (namely, atmospheric phenomena and living organisms)? Why do not we have to do rather with a universal decay of such objects and structures? Cosmological theories describe changes in time of the curvature and size of space and the density of matter and, for instance, radiation energy averaged for great areas (the total amount of energy occurring in different forms, e.g. in the form of matter, does not change and probably equals zero). Without additional assumptions, the general relativity theory and quantum mechanics are unable to specify the kind of changes in the spatial distribution of matter (and energy), and thus, the macroscopic forms the matter starts to adopt. At this point we come back to the problem of directionality and irreversibility of physical evolution, sketched generally in the first chapter of this book. As mentioned before, the physical theory called thermodynamics provides the missing parameter – the arrow of time, whose absence is felt in the domain of dynamic theories.

Where do these specific properties of thermodynamics – absent and not deducible directly from the laws of dynamics (mechanics) – come from? After all, both dynamics (quantum mechanics, theory of relativity) and thermodynamics refer to the same physical world, characterised by a set of universal and (we deeply believe) consistent features. Both describe, for instance, the behaviour of elementary particles, atoms and chemical molecules or systems and objects composed of them. How does it happen that such basic properties of the world as the arrow of time, irreversibility or ordered structures appear in thermodynamics, while they are absent from dynamics; moreover, the language itself, the mathematical structure of dynamics, is not suited to make sense while speaking about this sort of terms (they are so alien to dynamics, as colours to a blind)?

The above problem is affiliated with another, much better known problem, namely the question of how the phenomenon of life – which after all is in a sense nothing more than a set of particular atoms, behaving in a strict agreement with the laws of physics – emerges from the physical level and what distinguishes it from this level? Or, to be more exact, where does the property of **purposefulness** of living organisms – directed to realise a certain sort of tasks (survival and reproduction) – come from? How the process of biological evolution, **inexpressible** in purely physical terminology, can lead to origination of systems so sophisticated structurally and functionally? After all, neither purposefulness nor tendency to formation of specific, functional structures, is in any way **imprinted** into single atoms! We believe, however, that all these properties **emerge** in some way from a certain kind of behaviour of some systems of atoms. What does the **specificity** of this behaviour have to consist in, to form such macroscopically ordered and purposeful objects as living organisms?

Struggling with different interpretations of reductionism in biology (rejecting at the same time the primitive version of reductionism, expressed in the statement that man is nothing but a set of atoms), we frequently forget that a similar problem appears already within physics itself at the point where dynamics and thermodynamics meet. *Nota bene*, the way the thermodynamic level emerges from the dynamic level can throw a new light on the essence of the phenomenon of life (as it is in principle a thermodynamic phenomenon as well), for it constitutes an intermediate stage between single atoms and a particular complex biological organism. Thus, thermodynamics bridges to some extent the gap between classical physics and biology. I will deal with the essence of the biological level in the second part of this book (while the problem of the emergence of the psychological level will be discussed in the third part). Now, I would like to return to specific features of thermodynamics, distinguishing it from dynamics.

Although both theories describe the same objective world, thermodynamics differs fundamentally from dynamics, because its conceptual (and mathematical) structure is meant from the beginning to account for different **aspects** of this world. The differences in the "later" properties of dynamics and thermodynamics (resulting from the accepted assumptions) find their source in different **starting points** of both theories. Now I will try to make the above statement more specific and to show how these two different "points of view" yield two pictures of the world difficult to reconcile with each other (but this is, in my opinion, a problem of our system of concepts and not of the objective reality).

Dynamics focuses on the behaviour of single objects, or on interaction between one object and another. For example, quantum mechanics describes the evolution of the spatial distribution of the wavefunction attributed to a given electron as well as the interaction between a negatively charged electron and a positively charged proton in the atom of hydrogen. Newtonian dynamics and relativity theory may refer to a movement of a stone in the gravitational field of the Earth, or to a relative movement of the Earth and the Moon resulting from the gravitational forces exerted reciprocally by them. Dynamics works best, when we apply it to single objects or to pairs of objects – already the behaviour of three bodies of comparable masses in the gravitational field causes great mathematical problems (the three bodies problem cannot be resolved strictly in the analytical way).

On the other hand, the starting point of thermodynamics is an interest in the behaviour of **sets** of (a great number of) atoms (elementary particles, chemical molecules), rather than in the properties of single atoms (particles). We deal with such sets – containing **many** elements (in fact, unimaginable amount of elements of the order of, for example, 10^{20}), and not with single elements (e.g. atoms) – when treating macroscopic objects known from everyday experience, such as a stone, gas in a container or our planet. And just such objects are referred to by the terms of the arrow of time and irreversibility (compare the example of a cap which can break down spontaneously, but will never assemble itself from the pieces) as well as of complexity and organisation (it is enough to mention living organisms here).

In such a system, consisting of a great number of elements, it becomes of course completely impossible to investigate the behaviour of each of the elements separately. The quantity of the necessary measurements or calculations reaches such an astronomical number that it is not worth to take all of them into account[16]. Fortunately, the exact knowledge of the state (position in space and velocity vector) of each atom (molecule) of a system under consideration appears to be to a large extent **redundant**, at least from the macroscopic point of view.

It is so, because various opposite reactions of particular elements **cancel** each other **out** after an **averaging**, and therefore do not manifest themselves at the macroscopic level. Take a

[16] Anyway, exact measurements and calculations would not be possible because of the indeterminism (and the uncertainty principle) of quantum mechanics.

closed vessel filled with gas, where particular gas molecules move in different directions with different (frequently high) velocities. On average, however, the same amount of molecules moves with a given velocity in a given direction as in the opposite direction. Therefore, the average velocity vector will amount exactly to zero, which seems to be obvious, because the gas as a whole does not move, as it is limited by the walls of the vessel. We may also be interested not so much in the directions (vectors) of molecule velocities, but rather in the average value of this velocity and in the average kinetic energy of a molecule proportional to it. Here, we are also able to calculate the average value of this energy, and it will not be, of course, equal to zero (unless the gas has the temperature of the absolute zero, but in such conditions the gas would solidify). Therefore, **statistical approach** is the essence of thermodynamics. It consists in quoting the average values of some parameter(s) for the entire set of elements, instead of the entire set of parameter values for particular elements. Additionally, it is possible to determine the distribution of a given parameter value in the system under consideration, expressed, for instance, as a percentage of gas molecules having a given velocity (to be more exact: a certain narrow range of velocities), while another percentage of gas molecules moves with another velocity.

On the basis of the average values of such microscopic parameters describing single molecules, we can determine the value of a certain **macroscopic parameter** related to the entire system, while referring it to particular elements of the system does not make any sense. **Temperature** – proportional to an average kinetic energy of molecules (microscopic parameter) – is an example of such a macroscopic, statistical, "phenomenological" parameter. It seems to be obvious that in a physical system a certain temperature can be attributed to gas, but not to a single molecule, as in the case of a language where meaning is ascribed to a word or sentence, but not to a single letter. Apart from temperature, gas is characterised by some other parameters, such as volume or pressure (an average force exerted on the vessel walls by gas molecules hitting them incessantly). These three parameters describe in a complete manner (in some important sense) the state in which the gas finds itself at the present moment (if we assume that the gas is in the state of equilibrium). Therefore, the thermodynamic description of a system composed of many elements with an aid of macroscopic parameters seems to be very economical in comparison with the dynamic description. In the latter case, it is necessary to quote three co-ordinates denoting the position in space plus three co-ordinates of the velocity vector for each of, say, 10^{20} molecules contained in a given gas sample. This description conciseness results, of course, from the statistical approach, i.e. from ignoring individual behaviour of particular elements of the system and satisfying oneself with determination of only the average values of microscopic parameters[17].

Focusing attention on systems composed of many elements – which is the core of thermodynamics – is the condition that makes it possible to formulate such terms as "order" or "entropy" (i.e. a measure of the degree of disorder) within this physical theory. Obviously, if we deal with a single object (element), it cannot be in any way "ordered." On the other hand, in the case of a system composed of many objects, their distribution in space itself can be chaotic (and therefore not ordered) or can form a certain structure. A spatial structure can be accompanied by a temporal (functional) structure, comprising a certain pattern of changes in spatial distribution of elements in time. Thus, thermodynamics enables one to introduce the concepts of structure, (dis)order and complexity, absent at the dynamic level (or rather, these concepts start to make sense at all only in the framework of thermodynamics).

From the point of view of dynamics, and of quantum mechanics in particular, an atom of carbon entering (as one of the substances of some organic compound) into the composition of a highly organised structure (let it be a human brain) is identical with an atom of carbon in CO_2 molecule – an element of homogenous gas (carbon dioxide) which has achieved maximal entropy (degree of disorder) in a given set of conditions. Dynamics states that the movement of this atom and its possible chemical bonds are subject (in both cases) to such-and-such laws of nature, but **nothing more**. Dynamics is not able to go up to the higher level and state anything about the rules governing the behaviour of systems composed of many atoms. This inability does not come from a provisional imperfection of present dynamic theories, chiefly quantum mechanics, which will be replaced in the future with better theories, allowing us to overcome the above limitations. In my opinion, these limitations are of fundamental nature and result from the essential assumptions of the dynamic approach (paradigm). To pass from single objects to systems composed of many elements, it is necessary to supplement dynamics with a new system of concepts, absent from the framework of dynamics and impossible to derive from it (similarly, it is not possible to describe

[17] In a similar manner, economy is a statistical science. For example, it is not interested (contrary to psychology), whether entangled recesses of consciousness and sub-consciousness of a given person will lead him/her to purchase the goods under consideration, but only what per cent of people will decide to purchase goods due to financial and other conditioning factors. In this sense, while psychology can be regarded as "dynamics" of a human being as a creature gifted with mental life, sociology and economy are "thermodynamics" of human societies.

the essence of life exclusively by means of the terminology taken from the physical level of reality).

To be sure, even a single object, such as an atom, can have – in a sense – a certain kind of structure, expressing itself e.g. in the fact that an atom is composed of its nucleus and electrons. However, this is not the sort of structure considered by thermodynamics. The "structure" of an atom is **necessary**, as it results unequivocally from the laws of physics. On the other hand, any possible structure is by no means necessary in thermodynamic systems – it can exist, but it may just as well not exist (or disappear). The probability of a spontaneous origination can be attributed to each thermodynamic structure; and the smaller the probability, the more organised the structure is. The most probable structures – for instance, a gas evenly distributed in a container – are characterised by the lowest degree of order and their entropy acquires maximal value.

As the concept of a structure may seem inadequate in the case of homogenous gas distribution in a box, corresponding in fact to complete lack of structure, a more general concept of **macroscopic state** has been introduced. This term denotes a certain macroscopically ("phenomenologically") distinguished kind of spatial ordering of system elements. For example, it is possible to obtain from a given set of atoms both a homogenous mixture and a human organism. Both cases correspond to different macroscopic states of the same set of elements. Besides, it is possible potentially to imagine a great amount of various other macroscopic states, such as segregation of all chemical elements into separate vials. One encounters a similar situation in the case of gas in a box – apart from the homogenous distribution of gas, there exists a theoretical possibility that all the gas molecules would be located only in one half of the box. A great number of other potential macroscopic states is possible here as well as.

The fact that **different** (in reality, frequently just unimaginably different) **probabilities** are attributed to various macroscopic states is the crucial element for the properties of thermodynamics resulting from the starting points of this theory. It means that the chance of a spontaneous origination of various macroscopic states as a result of a random translocation of elements in a given system is very different. No one needs to be persuaded that there is a much greater probability for some set of chaotically moving atoms to form a homogenous mixture of atoms than a human body. The same happens with gas in a box – its molecules will tend to adopt an even distribution in the box as an outcome of their random collisions. Matching velocities of molecules and directions of their movement so well that all of them would concentrate in one half of the box has got relative probability (in reference to the homogenous distribution of gas molecules in a box) smaller than that of throwing exclusively tails during 10^{20} (100 000 000 000 000 000 000) throws of a coin (assuming the number of molecules to exceed 10^{20}).

How does thermodynamics cope with determination of probabilities of different macroscopic states? Well, here the amount of different **microscopic states** corresponding to a given macroscopic state acquires fundamental importance. A microscopic state is a such-and-not-other, unique state of all the elements of the system under consideration. In the case of gas, it means that the position and velocity vector (i.e. the value of velocity plus its direction) of **all** the molecules entering into the composition of this gas are exactly determined. Different macroscopic states represent different **classes** of microscopic states – many microscopic states correspond to one macroscopic state. It seems obvious that very different disorderly distributions of gas molecules will appear to be roughly identical and will correspond to the same macroscopic state, called homogenous distribution. Much smaller, but still enormous amount of microscopic states can be attributed to the macroscopic state where "all gas molecules concentrate in one half of the box" – it is possible to imagine many combinations of positions and velocities which would yield such a state.

What does it all have to do with the probabilities of macroscopic states? A lot. It appears that the probability of an occurrence of a given macroscopic state remains in close relation to the number of microscopic states corresponding to this macroscopic state – the more microscopic states, the greater the probability. All microscopic states have in principle got the same probability (such that the sum of the probabilities of all possible microscopic states adds up to 1). Moreover, there occurs a continuous transition from one microscopic state into another, due to incessant movement of molecules. The next microscopic state may belong either to the same class of microscopic states (the same macroscopic state) as the previous one, or to another class (another macroscopic state). During its random roaming from one microscopic state to another, the probability that the system will pass from one class of microscopic states (macroscopic state) to another class depends on the relative numerical force (number of elements) of these classes. It is obvious that the transition from a macroscopic state with a relatively small number of microscopic states corresponding to it to a macroscopic state "containing" a great number of microscopic states will be much more likely than the transition in opposite direction. In such a case we simply say that the second macroscopic state is much more probable than the first. Thus we see, how the number of microscopic states combining into a given macroscopic state determines the probability of the latter.

Let us consider an example of four integer numbers belonging to the range between 1 and 100. Let us assume that each of these numbers changes in a completely random way into any number from this range. A microscopic state of this system will be represented by a present set of four particular numbers. Every moment the system will pass from one microscopic state to another. Now, let us distinguish the macroscopic state X, in which all four numbers are identical, from the macroscopic state Y, corresponding to any other situation (where all four numbers are not identical). The probability of transition from state X to state Y is of course much greater, than the probability of transition in the opposite direction. The arrangements of four identical numbers will therefore happen relatively rarely. This example differs from real systems mainly in the fact that the probability of a spontaneous manifestation (appearance) of the distinguished macroscopic state X is only very low, but not astronomically low[18].

Let us look at one more analogy. Let our system consist of a tray with a certain number of small balls laying on it. A microscopic state of this system corresponds to a specified combination of positions of **all** the balls on the tray (shifting slightly just one ball will already originate **another** microscopic state). The tray is being shaken and therefore our system is continuously passing from one microscopic state to another. Now, let us define two macroscopic states: 1: "all balls forming a rectangle covering the left half of the tray" and 2: "balls scattered randomly on the tray." It is obvious that the macroscopic state 2 is much more probable. It is so, because a much greater number of microscopic states, i.e. of different combinations of ball positions, corresponds to this macroscopic state. In state 1, the balls may occupy interchangeably one of a few defined "positions" in various parts of the rectangle, e.g. on its edge, in its corner or in its interior, but this macroscopic state comprises only a relatively small amount of microscopic states, because of the limited number of combinations of possible ball locations in these positions. On the other hand, the amount of potential combinations of ball positions in the second (chaotic) macroscopic state is practically unlimited. Therefore, the system will achieve state 2 with a much greater probability

[18] Statistically, it is possible that the water in a pan heated up on a fire will freeze (the situation has a non-zero probability). However, the probability of such an event is so small, that the time of the existence of the Universe is much too short to enable the occurrence of even one such "miracle."

than the state 1. If state 1 is represented on a target as a single point, while the rest of the target field corresponds to state 2, it is obvious that there would exist only a very minute chance of hitting the point (state 1) during a chaotic (blindfold) shooting, and a given bullet would almost certainly reach the remaining area of the target (state 2) (we assume that the target is large enough to allow all bullets to reach it). At this point, it becomes clear why (coming to a more realistic example) the even dispersion of gas in a box is much more probable than the concentration of all gas molecules in one half of the box.

Every system possesses a natural tendency towards the evolution from a macroscopic state of a small probability to a much more probable state. This means that if the initial state of a system has got a small probability, the system will inevitably tend towards a state of a greater probability. On the other hand, if we "launch" the system from the most probable state, nothing will happen – it will stay in the same state as before. The evolution from a less probable (more ordered) state to a more probable (less ordered) state is therefore **irreversible**.

At this point, something should begin to dawn upon us; there should emerge something suspiciously familiar that we have already met on previous pages of this book. But let us keep first things first. Let us consider the problem of irreversibility of dynamic systems evolution in the following example. Let us take a vessel with two indents in the bottom – one small and shallow indent A and one large and deep indent B. Fig. 1.5. presents such a vessel. There is a single small ball in the vessel, and the entire vessel is being shaken. The small ball can either jump from depression B to depression A or – inversely – fall from depression A into depression B. Both of these transitions have got a specified probability, such that the probability of the A⇨B transition (pAB) significantly exceeds the probability of the B⇨A transition (pBA). This is probably obvious – a falling of the ball into indent A requires a really unusual chance, consisting in a very specific selection of the ball's velocity and direction of its movement, while practically every considerable movement will cause a jump of the ball from indent A to indent B. Thus, the system possesses an immanent tendency to evolve from state A to state B.

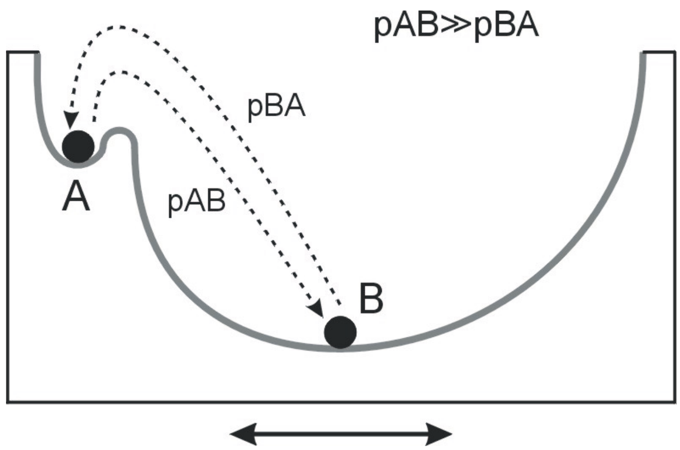

Fig. 1.5. A shaken vessel with two indents in the bottom: A and B. A small ball situated in each of these indents represents a different macroscopic state. Due to shaking of the vessel, the ball can pass from state A to state B or vice versa. The probability of the first transition (pAB) is much greater than the probability of the reverse transition (pBA). Therefore, the system exhibits a strong tendency towards the evolution from the macroscopic state A to the macroscopic state B.

Now, a single ball represents – differently than in the previous analogy with a tray – a state of the **entire** system, and not only of one element (in the example with a tray, the state of the system was the sum of positions of all balls-elements), while the bottom of the vessel determines the set of possible ball positions, and therefore of possible states of the system (let us remember that the ball never stops in the lowest point of any of the depressions, because the vessel is being continuously shaken). Of course, there exist numerous microscopic states (i.e. the possible ball positions on the vessel bottom in a given moment of time) in this case. On the other hand, we deal with only two classes of microscopic states, i.e. two macroscopic states. As we remember, all microscopic states can be regarded as equivalent, which means that the same probability is attributed to them. On the other hand, the probability of macroscopic states corresponds to a relative amount of microscopic states they comprise. It is clearly seen in Fig. 1.5 that the macroscopic state B is much ampler (contains considerably more possible microscopic states) than the distinguished state A, comprising only a few microscopic states. State A is in this sense more ordered, organised (possesses a smaller entropy which is, as we remember, the measure of the degree of disorder), i.e. it gives much more **information** about the localisation of the ball (if we already know that the ball is in the small indent A, it is much easier to find it than in the case when one possesses information that it is anywhere in the big indent B). Taking into account the overall probability, the ball almost certainly resides in indent B, and therefore the message that the ball is actually located there carries very little information. The macroscopic state B is therefore characterised by a high disorder of ball position (it can occupy one of the numerous microscopic states) and, consequently, by high entropy.

As it has already been said, the described system exhibits irreversible tendency to evolve from state A to state B, or to remain in state B if this state has been already achieved. We have seen that state B, as more probable, is also characterised by a smaller amount of information, lower degree of order (organisation) and higher entropy than state A (mutual relationships between these parameters will be discussed in detail later on). In this way, we have demonstrated step by step that a decrease in order and information as well as an increase in entropy must occur in each (isolated) system, until the values of these parameters stabilise at a constant, minimal (in the case of entropy – maximal) level possible in given conditions.

Thus, proceeding from the "starting points" (assumptions) of thermodynamics, we have succeeded in deriving its most fundamental, macroscopic properties that turned out to be inevitable results of an attempt to describe systems composed of **many** elements. Firstly, we came to the famous **second law of thermodynamics**, proclaiming that the entropy in isolated systems increases in time, until it reaches its maximal value (the system has to be isolated in order not to exchange matter and energy, and therefore entropy, with the environment). This is equivalent to a universal tendency to an increase of the degree of disorder. Energy tends to be dissipated – particularly during transformation of one energy form into another – in the most disorderly form, namely in that of heat energy.

The emergence of **the arrow of time** is the second, extremely important consequence of the statistical nature of thermodynamics, attempting to follow only average behaviour of elements in a system containing many elements. The arrow of time constitutes, in a sense, an alternative facet of the second law of thermodynamics – it corresponds to this direction in time in which the increase of entropy proceeds. In other words, the universal tendency of the degree of organisation to decrease not only extorts the unidirectional flow of time, but is (in a sense) even identical with this passage of time. Time flows in the direction in the space-time of the Universe in which entropy increases. A decrease of order with time implies that the degree of order reached a maximal value at the beginning of the Universe. I will return to this issue later on.

The second law of thermodynamics implies also **irreversibility** of all processes that are accompanied by a change of entropy. It means that for such processes, involving an entropy increase, we will not be able to observe in natural conditions any reverse processes, opposite to them. A cup can break into pieces, but it cannot spontaneously assemble itself from such pieces. Gas in a box tends to reach as homogenous (chaotic) distribution of molecules in space as possible. Generally speaking, every physical system tends to reach the state of **thermodynamic equilibrium**, in which entropy acquires the maximal value, and nothing happens once the value is reached.

Finally, thermodynamics provides such terms as "order," "structure", and "information" with a specific and at the same time very general (sometimes even too vague) meaning. Let us come back once again to our box with gas, to elaborate a bit on their sense. Let us distinguish again two macroscopic states: state A, in which all gas molecules are concentrated in the left half of the box (for example, immediately after the removal of a division separating the half of the box filled with gas from the half containing vacuum), and state B, representing homogenous distribution of molecules in the box. We already know that the destination state of this system is state B, as it is much more probable and therefore characterised by higher entropy. State A has slender chances to occur in a spontaneous way and therefore it can be achieved only by some **external** action, for example the above mentioned removal of a barrier separating two halves of the box (one of which has been emptied earlier). It happens so, because the macroscopic state A carries **information** responsible for the spatial ordering of gas molecules, and therefore for the kind of "structure" – understood in thermodynamic terms. The information cannot enter into the system from nowhere.

Molecules in state A are ordered – in the common sense of the term. Similarly, there is an order in a flat, when clothes are in a wardrobe, food products in a refrigerator and litter in a waste basket. In other words, we have to do with a spatial segregation of particular objects. Of course, a gathering of all gas molecules in one half of the box does not serve any practical purpose (in opposition to the arrangement of things in a flat), but distinctions of this kind are completely unimportant from the physical point of view.

It is perhaps more difficult to accept intuitively the above-formulated statement that state A has already a some kind of a well defined **structure**. Let us consider another example. A rectangle is covered with tiny squares fitting exactly to each other in, say, 1000 rows and 500 columns. We have therefore 500000 squares. Each square can be either white or black. The rectangle filled up with squares forms a certain pattern, or picture, like emulsion grains on a black-and-white photograph. Let us assume that an equal number of white and black squares is randomly scattered on the surface of the rectangle. In this case, one can perceive a homogenous, grey picture devoid, of course, of any structure. On the other hand, if we arrange the squares in such manner, that they represent a black-and-white reproduction of the painting "Mona Lisa," we will face a very complex two-dimensional structure (just as a sculpture corresponds to a three-dimensional structure). This structure results from a mutual relative arrangement of white and black squares in a "special," "meaningful," "defined" manner, that can easily be characterised by a sentence like: "this is a human face." A rectangle with black squares gathered in its upper half and white ones in the bottom half will also represent a structure, although a much less sophisticated one. This structure can be shortly described as "a rectangle divided into a black half and white half." In a sense, it represents an intermediate case located somewhere between Mona Lisa and a chaotic mixture of white and black squares. If it was painted by an abstractionist, it could even bear some germs of meaning. Anyway, it would certainly be less monotonous than the homogenous grey background and therefore it would deserve the name of a structure. Similarly, gas concentrated in one half of the box is a certain kind of structure, as opposed to gas homogeneously distributed throughout the entire box.

Both the concept "degree of order of a system" and the term "degree of structure complexity" can be defined and expressed through the amount of information contained in a macroscopic state corresponding to a given structure. I said earlier that the greater amount information a given macroscopic state has, the smaller is the number of microscopic states corresponding to it, and therefore the smaller is the probability of a spontaneous origination of this macroscopic state due to accidental arrangement of the elements of the system under consideration. It is obvious in the above presented analogy that white and black squares will arrange most probably in a grey mixture; the probability for the squares to be segregated into two equal parts, black and white, will be much smaller; while a spontaneous "painting itself" of "Mona Lisa" can be counted as almost completely precluded. It leads us to the conclusion that the masterpiece of Leonardo da Vinci contains the greatest amount of information, while the chaotic mixture of squares – the smallest amount.

There is, however, an alternative way of measuring of the amount of information associated with a given macroscopic state – a much more adequate one, in my opinion (at least for biological and cultural systems). According to this proposition, **the amount of information contained in a macroscopic state is proportional to the length of the shortest possible description of this state**. In other words, the more information contains a macroscopic state, the longer is the description necessary to show, how this state has been distinguished and what it consists in.

Let us apply the above rule to our rectangle. We will introduce the co-ordinates of each small square, determining the column (co-ordinate x) and row (co-ordinate y) of its location. Therefore, the position of a given square will be indicated unequivocally by a pair of numbers in the form (x,y), where x can adopt a value between 1 and 500, while y – between 1 and 1000; co-ordinates of a square may take the form of, for instance, (123,698). The **full** description of every possible macroscopic state of the system under consideration can be obtained in the following manner: "the squares having the following co-ordinates ... (here enters the enumeration) are black" (it is assumed that the rest of squares is white). This is the **longest possible** description of any macroscopic state (equivalent to the description of a single microscopic state), particularly in the case of an equal number of white and black squares. However, many macroscopic states can be characterised (or distinguished) in a much more economical way. The **shortest** possible description – as it suffices to use just the words "chaotic distribution" – can be devised to denote the macroscopic state representing "grey background," the case where white and black squares are scattered randomly and the entropy (degree of disorder) of the system reaches its maximal value. Slightly more information must be provided to describe the rectangle divided into two halves: black and white. It can take the following form: "all the squares with the value of co-ordinate y belonging to the range between 1 and 500 are black" (it is assumed that the squares with co-ordinate y between 501 and 1000 are white) or "all black squares are situated in the upper half" (but it is necessary to define earlier the "upper half"). Thus, the description of this macroscopic state is longer than the description of the "chaotic" macroscopic state, and therefore the macroscopic state "one half white, one half black" is more orderly and has a higher level of structure complexity. On the other hand, it would be necessary to define the co-ordinates of (almost) all black squares to characterise fully the macroscopic state "reproduction of Mona Lisa," and therefore the length of the description would approach the maximal possible length.

The discussed macroscopic state would contain (in a given particular system) the greatest possible amount of information, the most complex structure and the highest degree of order[19]. For exactly the same reason, we can consider a man as a structure much more complex (of much lower entropy) than a chaotic mixture of atoms contained in human body. We can see therefore, that the criterion of "the length of the shortest description" provides an instrument allowing us to estimate – at least intersubjectively – the degree of complexity of a structure and the degree of organisation (and therefore information content) of an examined macroscopic state. This criterion, however, is far from fully objective. A Martian would not certainly see the arrangement of squares representing the reproduction of Mona Lisa to contain a significantly greater amount of information than a purely random (chaotic) distribution of squares[20]. Biological structures as well (human organism included) can be considered as "especially distinguished" only from the point of view of the functionality requirements imposed by the purposes of the biological evolution.

It follows from the above discussion, that thermodynamics allowed – through the description of sets of elements – to derive such properties of the physical world as the tendency towards entropy increase, the arrow of time as well as the irreversibility of processes. It provided the possibility of speaking about the properties of physical systems slipping out of view of the dynamic paradigm, namely, structure complexity, degree of order as well as amount of possessed information. Unfortunately, as I have just said, it was necessary to pay a price for it, which took

[19]This statement is a simplification, since the positions of black squares (black points on a reproduction) are not completely independent of each other. Some points influence to some extent the positions of other points, in order to preserve e.g. the continuity of the oval of the face. For this and many other reasons, finding out the shortest description of a complicated macroscopic state would be very difficult to realize in practice. The estimation of mutual correlation of different point positions seems to be very troublesome. This, however, is only a technical problem, in no way invalidating the general principle.

[20]On the other hand, we can investigate, how complicated is the set of rules able to generate a given system (for example the Mona Lisa picture or a given living organism). For instance, the Mandelbrot set (very kin to fractals), which is apparently very (and in a sense infinitely) complex, can be generated by one very simple mathematical procedure. Therefore, it contains very little information. Inversely, to construct a brain of an adult human, we would probably need an enormously long list of instructions. In this case, the shortest possible description of the system would be very long. This is a quasi-objective way to estimate system complexities. However, the problem consists in how many microscopic states are contained in a given macroscopic state, i.e. where the "boundaries" of this macroscopic state are to be found. If we remove one atom from the Mona Lisa picture (human organism) or replace it with another atom, is it still the same Mona Lisa picture (or human organism) or not? What about ten, a hundred, or a million of atoms?

the form of certain **subjectivity** of thermodynamics, as the **macroscopically distinguished state** is an inseparable element of its conceptual construction. As we remember, this is just the tendency to transition from less probable macroscopic states to more probable ones that forms the basis of the second law of thermodynamics, arrow of time and the property of irreversibility. It is a macroscopic state that is characterised by a specified structure, ordering and informational content. However, the very separation of a certain macroscopic state – i.e. joining a group of microscopic states into a distinguished class – seems to be in principle a subjective procedure (intersubjective at best). The length of the shortest possible description indispensable to characterise completely a macroscopic state is, in my opinion, the best accessible measure of amount of information defining the essence of a macroscopic state (at least in the case of biological, mental and cultural systems). This length, however, is related to the state of our mind, to its knowledge and categories in which it sees the world. I think that the contamination with subjectivity of the very idea of macroscopic states is its fundamental feature. "Objectively," there exist only microscopic states and they have got equal probability and are equivalent in every other respect. A particular microscopic state, represented at some moment by all the atoms contained by a human organism, is of equally small probability as any other microscopic state of the same atoms that form a chaotic mixture. In both cases, the definition of a given microscopic state requires information about positions and velocity vectors of all atoms contained in the given system. A man is a much more orderly macroscopic state than a chaotic mixture of atoms, because the number of possible microscopic states yielding a homogenous mixture is greater by a trurly astronomical factor than the number of microscopic states that can be classified as the macroscopic state called "a human being." As it is impossible to formulate any objective criteria of such classification, we have to (in my opinion) satisfy ourselves with the immanent subjectivity of thermodynamics – at least in relation to biological and cultural phenomena[21] – or to change our

When we do not deal with single microscopic states, we have to deal with their classes, but distinguishing (defining) these classes is very arbitrary and therefore subjective.

[21]All this does not mean that there is no objective measure of entropy or information: the box with gas molecules concentrated in one half has certainly lower entropy (contains more information) than the same box with gas molecules distributed homogeneously. There can be defined in a strict mathematical way the correlation between spatial situations of particular atoms – this correlation is lowest for the random distribution of atoms. However, exactly the same amount of entropy (information) can be attributed to a human organism and to a homogenous mixture of all the atoms building this organism, concentrated in a very small volume in a box (which corresponds to concentration of gas molecules). Such a mixture of atoms, although located in a very specific (one

concept of objectivity and accept as fact that some properties and aspects of the world, absent from the purely dynamic level, emerge on the thermodynamic level (and later: on biological and psychological ones). It would be equivalent to attributing a more objective status to such aspects of the world as ordering, complexity, biological "purposes" and mental life. It would mean that any fundamental difference between subjectivity and objectivity would be, in principle, negation discarded. Personally speaking, I am an advocate of the second alternative, as I think that connecting objectivity only with the (sub)atomic level of reality has already become slightly too archaic, all the more so that quantum mechanics also meets very serious problems with objectivity just at this level. I will deal with the problem of blurring the sharp distinction between the objective and subjective spheres of the world in the third part of this book, devoted to psychical evolution.

We can conclude that in the case of biological (and mental) systems, thermodynamics offers us a good measure of the **quantity**, but not the **quality** (specificity) of information, distinguishing the living phenomena from physical objects. The same set of atoms, that forms a human body, could be arranged in such a way that those atoms would form a macroscopic state having a similar or greater amount of information as a human organism, but that will would be completely "senseless" from the biological point of view (for example a sufficiently complex system of convection currents or vortexes in water). After all, the protons, neutrons and electrons, presently forming your or my body, just after the Big Bang formed a mixture that contained more information (had smaller entropy) than they do now – as elements of a human organism (I will discuss this problem in the next chapter). Therefore, we need a measure, or a definition, of the **quality** of biological (and also psychological) information. In the second part of this book, I propose a cybernetic paradigm for the description of the phenomenon of life, in the frame of which the quality of information can be attributed to the amount (or "density") of negative feedbacks constituting a network of regulatory mechanisms, responsible for the functioning of living organisms. This network of regulatory mechanism is equivalent to the functional identity of a living individual, serving its survival and reproduction. Generally speaking, I believe that the

---

of many possible) places, instead of being dispersed evenly, is nothing special or meaningful for us, unlike a human organism. This suggests that the information connected specifically to biological phenomena is something different than purely thermodynamic information. In other words, what thermodynamics fails to treat is not the quantity, but the (e.g. biological) quality of quality of "biological information" originates from the self-directing of a living organism identity on propagation of just this identity, "made up" from these regulatory mechanisms. In the third part, I suggest that the quality of mental (psychological, self-conscious) information comes, analogously, from the relation of a directing on itself of the "cognitive center" in our brains.

Thermodynamic evolution of the Universe

As we remember from the chapter devoted to the (dynamic) evolution of the Universe, the matter contained in it – initially in the form of a homogeneous mixture of elementary particles – started to create gradually conglomerations in different scales of magnitude. There occurred the process of separating clusters of galaxies, galaxies themselves, and finally – subsequent generations of stars. With a high contribution of heavier chemical elements, synthesised in the nuclei of stars, and thrown afterwards out into the cosmic space during explosions of Supernovae, there began to form planets and their moons, as well as smaller rocky (and icy) fragments: asteroids, comets and meteorites. The initially homogeneous distribution of matter underwent therefore a change into an outstandingly heterogeneous distribution. Moreover, astronomical objects formed as a result of those transformations – galaxies, stars and planets – underwent further internal differentiation, which allowed them to obtain a specified internal structure. Galaxies are composed of a central part and spiral arms, stars possess a nucleus, a mantle and a crown, the planet Earth, apart from the two-layer nucleus and a fluid mantle, possesses also a planet crust and an atmosphere, both characterised by a complex structure and complex dynamics (additionally, most of the Earth's surface is covered by oceans, forming hydrosphere). In the case of a planet, each of its layers possesses a distinct chemical element composition, so that heavier elements exhibit a tendency to be located nearer the center (of gravity) of the planet. On our planet, and probably on many other planets in the Universe as well, there additionally came into being highly organised living organisms, whose degree of complication increased in time, in the course of the process of biological evolution. Finally, in some cases (at least in one case) life yielded rational creatures (as ourselves) who use a highly sophisticated brain (or a corresponding organ in the case of aliens, if they exist) to think. Therefore, it seems incontestable that there

---

information. In the second part of this book I will propose how such a quality of biological information can be defined.

exists a clear tendency of matter to achieve a highly heterogeneous distribution, as well as to form different kinds of orderly structures. Apart from the expansion of space, this property constitutes a fundamental feature of the hitherto evolution of the Universe.

Wait, wait! Just a while ago we said something else, quite the opposite, while considering the evolution of thermodynamic systems, among which one undoubtedly counts the Universe composed of many (it would be difficult to deny this!) atoms. In accordance with the second law of thermodynamics, the entropy – and therefore the degree of disorder – has to grow, whereas the amount of information and the complication of structure related to it – to **decrease**. We were proving that the spatial distribution of either gas molecules, or balls on a tray, should tend towards a homogenous distribution in a system. The mosaic forming the Mona Lisa picture should scatter into a chaotic mixture of squares. However, something exactly opposite has happened. Several billions of years after the moment when all atoms formed a homogeneous soup, there emerged the Mona Lisa picture from this soup – with some assistance of a certain Leonardo da Vinci, *nota bene* also a descendent of the soup! Have we therefore invested so much trouble, deriving logically, step by step, the most important laws of thermodynamics, to state at the end that these laws have nothing to do with the real world, that was to be inevitably subjected to them? Does the reality we live in defy the most elementary logic?

Not at all. The above-quoted "deviations" from the second law of thermodynamics – in the form of a spontaneous formation of large-scale matter conglomerations (astronomical objects) and the origination of highly-organised living organisms – only appear to be deviations. What is interesting, they do so for entirely different reasons. It means that the concentration of clouds of gas and dust into heavenly bodies apparently "circumvents" the general tendency towards universal increase of entropy in an entirely different manner than the process of origination and evolution of living creatures does. While living organisms "feed on" negative entropy (negentropy) of their environment, astronomical objects are completely self-sufficient with the respect to the formation of their own structure "against" the rule of the growth of chaos. The present subchapter is aimed to explain how galaxies, stars and planets manage the second law of thermodynamics, while in the next subchapter I will show how thermodynamics admits the origination of living organisms.

The **gravitational force** is the key to understanding the thermodynamics of the Universe as a whole, and of the evolution of matter distribution (directly related to it). Each material object with a mass exerts this force and at the same time is subjected to it. As it is widely known, the gravitational interaction manifests itself as mutual attraction of bodies. Due to this property, matter exhibits a natural tendency to gather in one point. The fact that for a long time the **entire** matter has not concentrated in one place – to form something like a huge black hole or a terminal singularity – is due to the expansion of the Universe. There occurs here an incessant rivalry between the movement of material objects away from each other due to the expansion of the Universe and bringing them nearer caused by gravitation. For today, the gravity has won at smaller distances, which brought about the formation of planets, stars, galaxies and clusters of galaxies. On the other hand, the impetus granted to matter (and space) in the time of the Big Bang – thanks to which clusters of galaxies still move away from each other – still predominates at distances of greater orders of magnitude. However, gravitation generally favours a heterogeneous distribution of matter in space.

On the other hand, mutual gravitational interaction between gas molecules in the box with gas is so insignificant that it can easily be neglected. The predominant role is played in gas by chaotic movements and collisions of molecules, resulting from their kinetic energy. Just because of an accidental movement direction of particular molecules, gas tends to enter into the state of random distribution of molecules in space. On the other hand, in the case of objects of astronomical scale of magnitude, their movement is by no means accidental – its direction is imposed mainly by the gravitational force. As a result, astronomical objects have a tendency to approach the common center of gravity. This explains why homogeneous clouds of gas and dust concentrated into galaxies, stars and planets, but the opposite process – as in the box filled with gas – does not occur, namely the evolution towards a homogeneous distribution of matter. The difference results from the size of the considered systems. On the cosmic scale, implying enormous masses, the gravitational force (proportional to mass) predominates over accidental thermal movements of molecules. After all, even the box with gas does not fly away into the cosmic space, but remains in an immediate contact with the surface of our planet because of the gravitational attraction towards the Earth. However, on a small scale, gravitation allows gas molecules to move as they want, since their mutual gravitational attraction is very weak due to their small mass (of course, atmospheric gases are kept near the Earth's surface by gravitational force exerted by our planet).

The difference between the direction of the evolution of matter distribution in the box with gas and in Cosmos is presented figuratively in Fig. 1.6. In the first case, the situation corresponds to the behaviour of balls on a constantly shaken tray. The shaking is to supply kinetic energy to balls. Even if we concentrate the balls in one place, they will quickly disperse chaotically on the surface of the tray as a result of accidental collisions. On the other hand, big masses in astronomical scales of distance behave like balls on a rubber membrane. They do not scatter chaotically; on the contrary – they gather around their common center of gravity.

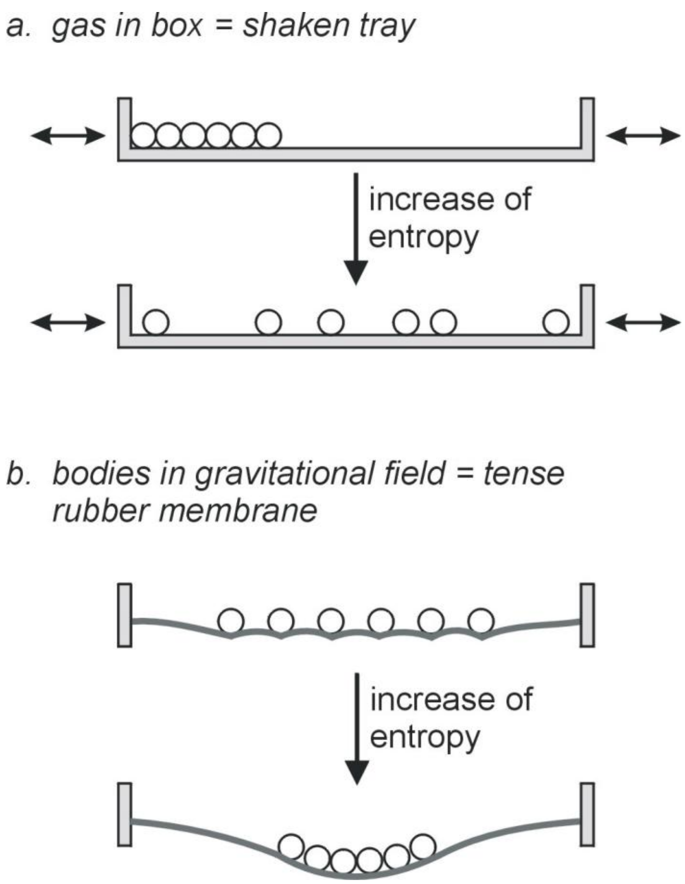

Fig. 1.6. A shaken tray with balls on it as an analogy of the evolution of the distribution of matter in a box with gas. Lower: balls on a rubber membrane illustrating the evolution of the distribution of mass (matter) in the astronomical scale. In the first case, the balls exhibit a tendency to a chaotic distribution, in the second case – inversely – they concentrate in one place.

Well now, we already know why astronomical objects, unlike gas in a box, have a tendency to evolve from the state of homogeneous matter distribution towards its concentration in different kinds of structures. Yet, does it mean that the second law of thermodynamics is "reversed" on the cosmic scale and that, instead, the principles of entropy decrease and of increase of the degree of order are now in force? Is the "thermodynamic arrow of time" turned in the opposite direction than in the microscale when we pass to large scale distances and masses? By no means. Origination of planets, stars and galaxies from the initially (almost) homogeneous distribution of gas and dust in space is **also** an expression of the universal tendency towards entropy increase and decrease of ordering. "No," somebody could say, "nobody will force me to accept such an obvious absurdity. After all, what happens in Cosmos is an exact reversal of what occurs in a box with gas. If a sphere is formed from all gas molecules in the center of the box, such a 'mini-planet' will instantly disperse over the entire box." Well, the devil of it is that the situation in the box is not at all a strict reversal of the situation in Cosmos. Again, what is responsible for this lack of symmetry is gravitation.

Since the Einstein's publication of the general relativity theory, we have known that gravitational force is equivalent to **curvature of space**. This thing done by a body with a mass when it exerts its gravitational attraction consists in curving of a "flat"[22] three-dimensional space into a certain kind of a "funnel," where other massive objects can fall[23]. The surface of a "normal" funnel is a two-dimensional formation, while a three-dimensional funnel can be imagined as a certain kind of "condensation" of space around a body with a mass. Similarly, a two-dimensional funnel corresponds in a sense to a "condensation" of a plane. If we look on such a funnel from above or – in a more mathematical language – perform a projection of the funnel onto a horizontal plane, then a certain arbitrary unit of length (let it have the value 1) on the plane of projection will

[22]At least in the scales of magnitude smaller than, say, one billion of light years. As we remember, the space of the Universe as a whole can be curved positively or negatively. However, in smaller scales, this curving is negligible, as the curving of the terrestrial globe in the scale of a football field is.

[23]We do not see this curving because the course of visible rays, moving with an enormous speed (the speed of light!) undergoes only a slight deflection in a gravitational funnel (although we known of the effect of a "gravitational lens," where we see a double, or even a quadruple image of a distant galaxy, caused by curving towards themselves of the courses of light rays coming from this galaxy by a gravitational field of a star lying much nearer, between the Earth and the considered galaxy). The effect of curving of space "felt" by much slower objects attributed with mass is much stronger.

correspond a longer (greater than 1) segment on the surface of a funnel. The closer to the center (outlet) of the funnel we are, the longer the segment becomes. The situation is illustrated in Fig. 1.7. Let us now arrange balls on the plane of projection in such a manner that they form one cluster. It turns out that this situation corresponds to a **homogeneous** distribution of balls on the surface of the funnel! On the other hand, a homogeneous distribution of balls on the plane of projection would mean just a rarefaction of balls in the funnel! If we therefore look on the problem from the point of view of a funnel, then a concentration of balls on the plane of projection corresponds to their homogeneous distribution. This is also illustrated in Fig. 1.7.

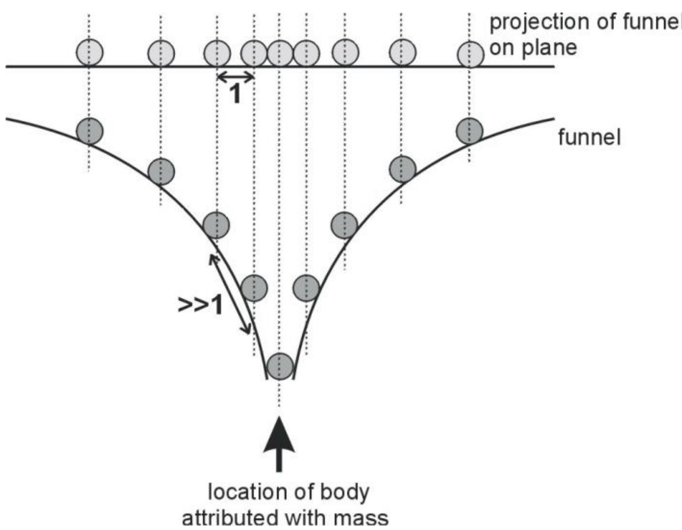

Fig. 1.7. A section through a two-dimensional funnel, symbolising curving of space around a body attributed with mass and the projection of this funnel on a horizontal plane, reflecting the "apparent" situation registered by our eyes via visible rays. The closer to the center (outlet) of the funnel (the steeper its walls) we are, the greater length on the surface of the funnel corresponds to an arbitrary unit of length equal to 1 on the plane of projection. From the point of view of the plane of projection, the surface of the funnel (two-dimensional space) undergoes "condensation" near its center. Homogeneous distribution of balls on the surface of the funnel makes an impression of concentration of balls on the plane of projection (and therefore for our eyes), and of their highly heterogeneous distribution in space.

The situation is analogous in the real three-dimensional space. While our eyes – through the mediation of light rays – see rather the projection of a three-dimensional funnel on this space, material objects attributed with mass "see" just the funnel itself, and not its projection, and behave accordingly. If we take into account the curving or "condensation" of space around the bodies acting gravitationally, then it turns out that this condensation of matter in one place corresponds exactly to a "homogeneous" distribution of matter. Therefore, the evolution of the Universe tends to such a state and the second law of thermodynamics appears to be maintained! The formation of clusters of matter on the astronomical scale turns out to be exactly the same process as achieving a homogeneous distribution by gas in a box[24].

I remind that gravitation – equivalent to curving the space to form three-dimensional funnels around bodies possessing mass (gravitational attraction corresponds to "falling" of other bodies into such a funnel) – is the culprit "guilty" of the reversal of situation with respect to small-scale (not astronomical) systems (e.g. the box with gas). When two masses meet, their funnels fuse forming a greater funnel, which even more efficiently "lures" into itself other bodies with their funnels. On the other hand, space "seems" to be much more flat for electromagnetic rays (including the visible rays that help us see the world) – due to their great speed (the speed of light) and zero resting mass – than for much slower bodies possessing a mass and therefore manifesting a tendency to fall into gravitational funnels. If we could see "with the eyes" of gravitating bodies, we would see a much more homogeneous distribution of matter in the Universe. What is more important, however, the evolution of matter distribution from a homogeneous state to still denser and denser as well as more and more massive concentrations would look for us as a transition from a highly heterogeneous distribution (with "holes" of rarefied matter in the places corresponding to funnels) towards a homogeneous distribution of matter.

Let us add that, from the point of view of a funnel, the "homogeneous" distribution corresponds to the greater "apparent" condensation, the nearer to the center of a funnel we are. It is not surprising therefore that such astronomical objects as the Earth or the Sun possess in their centres a nucleus of a relatively large density (specific weight), while the "lightest" layers, such as terrestrial atmosphere or Sun's crown, situate themselves outside (in the outermost position). To understand why it is so, it is necessary to have a look at Fig. 1.7. The nearer the center of a funnel, the steeper its walls seem to be, and therefore the greater "condensation" characterises the surface of a funnel (in the case of a three-dimensional funnel – its "space") in relation to the projection of the funnel on a horizontal plane. Thus, matter distribution must be denser on the plane of projection to give a homogeneous distribution of mass on the surface of a funnel.

[24] In the case of gas, the kinetic energy of movements of gas molecules overcomes a very week gravitational attraction between particular gas molecules.

Taking into account the curving of space caused by gravitational attraction of bodies with mass, we state that the distribution of matter in the Universe evolves towards a less-orderly state, with less amount of information and larger entropy. This results, shortly speaking, from the fact that the "real" distribution of matter in curved space becomes more homogeneous in time, in spite of an "apparent" tendency of matter to format different kinds of clusters.

* * *

Since we have already made sure that, instead of denying the second law of thermodynamics, the evolution of the matter distribution in the Universe yielding concentrations of matter in various clusters is just an **expression** of this law, we will see now how the tendency of entropy to increase was realised on the astronomic scale during the history of our Universe.

As we remember, in the initial period of its existence the Universe was (almost) **homogeneously** filled with matter and radiation. The phenomenon of inflation – related to an enormous and very quick expansion of space immediately after the Big Bang – led to a "flattening out" of all eventual divergencies from the homogeneous distribution of matter. In the high temperature of the Universe at that time, elementary particles were colliding incessantly with each other, absorbing and emitting quanta of radiation. Accidental collisions maintained homogeneous distribution of matter in space (chaotic movement of particles with huge kinetic energy predominated over the force of gravitation). Temperature and density, however, were constantly diminishing as the Universe expanded. Electromagnetic radiation as well was undergoing the process of rarefaction and cooling off. Protons and electrons joined to form atoms of hydrogen, and collisions between them as well as interactions with radiation practically ceased. Due to swelling of space, the average distances between atoms increased. Inherited from the earlier phase of the Universe evolution, the more or less homogeneous distribution of matter still persisted. The Universe would have undoubtedly remained homogeneously filled with hydrogen atoms until today, and their density would have fallen in the meantime to a few atoms per cubic meter and would have continued to fall, if the gravitational force had not remained on the stage of the drama.

Due to accidental collisions, some atoms gathered closer some other atoms and moved away from the remaining ones. These differences were rather small at the beginning and the distribution of matter still remained almost homogeneous. However, the "almost" was decisive for the later formation of stars, galaxies and their clusters. If some number of atoms created by chance a hardly separable cluster where particular atoms were located slightly closer to each other than the average distance between atoms in the entire space, then the mutual gravitational attraction between them could gain a slight advantage over the tendency to move away from each other caused by the expansion of the Universe. As a result, the initially tiny heterogeneities in matter distribution began to grow. The particles of matter within the limits of a "cluster-to-be" were more and more slowly moving away from each other and particular atoms started to approach each other. On the other hand, the space between conglomerations was being emptied, as the matter remaining there was swept out by the gravitational attraction of the nearest cluster. This process occurred simultaneously in very different scales. It means that huge areas of space witnessed the formation of clouds of hydrogen that were the beginning of galaxy clusters. Inside the clouds, matter concentrated into "galaxies-to-be," witnessing in turn the differentiation of yet smaller conglomerations – the "ancestors" of stars[25]. During billions years, the formation of stars, galaxies and their clusters gave them their "final" shape. In the present Universe, although the clusters of galaxies move away from each other increasing therefore the area of space (almost) devoid of matter, the distances between galaxies in clusters and between stars in galaxies is maintained on a more or less constant level.

The origination of all heavier chemical elements is also due to the second principle of thermodynamics acting in the gravitational field. Mainly hydrogen and certain amount of helium (constituting about 25 % of the weight of primary matter) appeared in space as a result of the Big Bang. The first generations of stars in galaxies formed exclusively from these two elements. The giant condensations of matter in the centres of stars (centres of huge gravitational funnels), however, were caused by the very tendency towards an increase of entropy that induced matter to create conglomerations in the gravitational field.

As a result of great pressures and temperatures prevailing in the conditions of so high matter density, thermonuclear synthesis was initiated, by which protons and neutrons joined into still greater and greater conglomerations (nuclei of heavier chemical elements), which was

[25]Some heterogeneities in the background radiation – reflecting very small deviations from a homogeneous distribution of matter in the early Universe in the largest scales of distance – have been recently discovered by the satellite COBE.

accompanied by a release of great amounts of energy[26]. From the human point of view, the above process fulfils three very important "functions." Firstly, it constitutes the source of electromagnetic radiation (visible radiation included), emitted by stars (just for this reason stars shine). And it is the light radiation emitted by our star (the Sun) that drives a majority of thermodynamic processes on the surface of the Earth, including the majority of phenomena related to life. Secondly, the pressure of radiation formed due to nuclear synthesis – and therefore, indirectly, to gravitation – restrains stars from collapsing further on, caused by the very gravitational force. As a result, a sort of balance settles between gravitation and the pressure of radiation. This dynamic equilibrium determines star dimensions. Finally, nuclei of all the chemical elements in the Universe heavier than hydrogen and helium (e.g. carbon, oxygen, silicon and iron, but also tens of other elements) originated in the thermonuclear synthesis. Our planet and we ourselves are built of those elements.

To serve as building blocks for planets and organisms, the heavier elements had to get out first from star interiors. This is not possible in the case of small stars, such as the Sun or smaller, that – having transformed all hydrogen and helium in their interiors into heavier chemical elements – simply die out slowly (passing earlier through the stage of a red giant, white dwarf and brown dwarf). However, more massive stars have a much more stormy life. At a certain moment, their core collapses suddenly, creating a neutron star or a black hole, while the external layers explode equally suddenly in all directions as a Supernova. It is exactly during Supernova explosions that the heavier elements are taken by the external layers, thrown out into the cosmic space. There, they enrich the clouds of inter-stellar gas, initially composed mainly of hydrogen (with some admixture of helium). The next generation of stars arises from such gas-dust clouds. Formation of planets with a solid crust (such as the Earth) follows. Life originates on some of them on the basis of elements synthesised in the interiors of massive stars.

Let us return, however, to the "ordering" role of gravitation. Gravitation not only causes an origination of conglomerations of matter (it has already been discussed) but also imposes on them the internal structure. The general tendency prevails for mass to form the greatest condensation in the center of a conglomeration (for the reasons mentioned before). There are more

[26] This energy is released as a result of formation of "bonds" between nucleons (protons and neutrons) due to the operation of strong nuclear forces. To separate the nucleons, it is necessary to supply a great amount of energy. It is therefore logical that in a reverse process (bounding of nucleons) this energy is released.

stars per volume unit in galaxy interiors than on their outskirts. Furthermore, there are reasons to believe that the centres of at least some galaxies are occupied by black holes of gigantic masses, representing possibly the most condensed form of mass (matter). In stars, the greatest density of matter (and also the highest pressure and temperature) is also found in their centres, where thermonuclear reactions occur. As to our planet, when it was originating, a metallic nucleus "smelted" in its center (as in a foundry furnace). Over this core – due to "gravitational separation" – the remaining matter gradually sorted out into: the mantle, crust, hydrosphere and atmosphere. Thus, we see that gravitation not only orders matter in space, concentrating it into different clusters, but also confers a defined structure to those conglomerations. Let us remember, however, that all of this occurs in agreement with the second principle of thermodynamics, and therefore both the "ordering" and the "structure" are here in a sense illusory. In reality, there occurs a decay of structure and order in this case. Or better – what we perceive as (apparent) gravitational order corresponds to something fundamentally different, or even simply opposite to the (true) thermodynamic order.

The process of growing mass concentration has been occurring continuously and will still proceed in time. It tends to closing the entire matter, or at least a possibly large part of it, in the so-called **black holes**. I have already said that at least some of them constitute the final stage of the evolution of massive stars. It still remains, however, to be explained what a black hole is in the first place.

If a sufficiently large mass is concentrated in a sufficiently small volume, then the gravitational field becomes so strong in the immediate neighbourhood of this mass that even light (more broadly – electromagnetic radiation), moving with the greatest speed possible in the Cosmos (the speed of light), is too slow to get out from the reach of attraction of this mass. It cannot therefore move away sufficiently far in space to be received by an observer's eyes. As nothing in the Universe is able to move faster than light (the special relativity theory is based on this assumption), **nothing**, no signal (of whatever nature it would be) can leave (escape from) a sufficiently steep gravitational funnel created by sufficiently dense matter concentration (the so-called Hawking radiation, to be treated in a moment, is an apparent exception). Everything that falls into such a funnel, be it matter or radiation, will never get out from it. It will simply disappear, as in a hole. By definition, a black body is a body that does not emit any visible radiation. No wander, therefore, that massive objects characterised by the properties described above have been named "black holes."

I remarked earlier that black holes represent the most condensed form of matter (mass) in the Universe. How much condensed? Infinitely, according to the present physical theories (principally the general relativity theory). For, there appears a singularity in the center of a black hole (similar to the previously-discussed initial singularity, or rather – because of the direction of the arrow of time – to the terminal singularity in a closed universe), with zero dimensions and infinite density.

Black holes are also the form of matter of the greatest (maximal possible) entropy. As I said, a situation where all objects with mass meet at one point – corresponding to their common gravitational center – is the top of their "thermodynamic dreams." Black holes guarantee their fulfilment. No wander, therefore, that black holes constitute the final stage of the evolution (transformation) of the forms adopted by matter.

But this is not the end yet. The entropy of a black hole depends on its size – the greater a black hole, the higher its entropy. Due to this fact, further growth of entropy can develop in two ways. Either a black hole sucks into its interior the surrounding matter as a result of gravitational attraction, increasing in this way its dimensions and mass, or two black holes join into one, with a similar result. It is required by the second law of thermodynamics, proclaiming inevitable tendency toward entropy increase. All this makes us face an extraordinarily "alien" vision of the far future of our Universe: (almost) entire matter existing in the form of black holes, joining into still greater and greater black holes. In an infinite universe, the latter process can also proceed infinitely. The gravitational attraction will restrain the expansion on growing scales of distance, so that the matter from vaster and vaster areas of the Universe will finally concentrate into one black hole. The process will proceed endlessly, especially that the rate of the increase of the maximal possible entropy[27] will exceed the rate of an increase of the actual entropy (expressed in the magnitude of black holes) due to the increase of the dimensions of the Universe. The difference between the entropy already achieved and the entropy possible to achieve will therefore increase continuously. However, this is not the type of thermodynamic gradients that could be responsible for driving the processes of life, for instance. Indeed, it is difficult to imagine life in the world of a

[27] The expansion of space increases the amount of possible microscopic states (in this case -- permissible positions in space), which causes an increase of the maximal possible entropy.

terrifyingly empty space and rarely distributed, extremely massive black holes. But this may just be a problem of lack of imagination.

In a closed Universe (with finite dimensions and time of existence) the **entire** matter will finally concentrate into the terminal singularity, with zero dimensions and infinite density. The present expansion will be stopped at some moment, and matter – whose growing portion will appear in the form of black holes – will finally gather at one point. In a sense, the terminal singularity can be regarded as something like an ultimate, final black hole with maximal entropy, in which the entire matter concentrates and entire space "curls up" into a point.

We remember that the initial singularity, preceding immediately the Big Bang (or, if one prefers, constituting its starting point) was characterised by similar properties. Does therefore the terminal singularity constitute simply the reversal in time of the initial singularity? If so, it would mean that, at the end, the Universe will come back to its starting point. Such a state of things would be also equivalent to the return of entropy to its initial value. How is this related to the second law of thermodynamics, proclaiming the tendency to entropy increase, and how is this possible if, as we have seen, entropy increases all the time during the evolution of the Universe?

Well, the terminal singularity is not at all a mirror reflection in time of the initial singularity, for, its "structure" (understood in the sense allowing one to talk about the structure of a point!) is completely different. Matter is distributed very evenly in the "nearest neighbourhood" of the initial singularity. It forms a homogeneous "soup" of elementary particles and radiation. The curvature of space is practically the same everywhere. On the other hand, the nearest neighbourhood of the terminal singularity is very heterogeneous, differentiated into massive black holes and empty space separating them. The curvature of space exhibits significant deviations from homogeneity as well, due to huge gravitational funnels of black holes. Future theories (e.g. quantum gravity), that will remove the unpleasant properties of the initial and terminal singularities – such as zero dimensions and infinite density – will probably allow us to make more sense when discussing the structure of the singularities themselves.

Let us remember that, allowing for the gravitational force, the above mentioned differences between the initial and terminal singularity mean that the entropy of the former singularity was low, while the entropy of the latter – high. And therefore, during the history of a closed Universe, there occurs a dissipation of order and the related information that is equivalent to an increase of entropy. Once again, the second principle of thermodynamics is saved.

Already in the preceding chapter devoted to the dynamic evolution of our Universe, we came to the conclusion that the origination of the Universe was equivalent (at least in a certain aspect) to the appearance of **information** concerning distribution of matter and space in the gravitational field. This conclusion returns now with more intensity in the context of thermodynamics[28]. Bodies with a mass have a natural tendency to fall mutually into their gravitational funnels. A homogeneous distribution of matter in space is therefore equally improbable as a (spontaneously originated) distribution of boulders on a ground where each boulder rests on top of a very steep land projection in the shape of an aiguille. It is obvious that erosion and wind will cause the boulders inevitably to roll down into the ravines separating aiguilles. This is more or less what happens during the evolution of the Universe, but, additionally, particular boulders are the source of ravines (gravitational funnels) for other boulders and these ravines fuse into larger depressions in time, joining their strengths to attract other boulders. Undoubtedly, the macroscopic state "each boulder on the top of an aiguille" is distinguished among all other macroscopic states. This state carries a huge amount of information – it looks as if somebody has arranged the situation on purpose. Let us remember that the sudden expansion of the Universe (especially in the so-called phase of inflation) contributed in an important manner to the "flattening" of the possible initial heterogeneities in matter distribution. It has also been preventing (until today) the entire matter from collapsing into one point. Solving the problem of where the expansion comes from remains one of the greatest challenges of cosmology.

In the above considerations, emphasis was lied on the gravitational force. Three forces, however, are responsible for the general tendency of matter to form different kinds of conglomerations. First among them are strong nuclear forces that made quarks join into protons and neutrons, and were also responsible for the following formation of atomic nucleuses of different chemical elements out of protons and neutrons. Then comes the electrical attraction that led to creation of atoms by joining negatively charged electrons to positively charged atomic nucleuses. Finally, the force of gravity caused a formation of conglomerations of atoms and black holes. The "transformation" of matter into black holes as well as joining and enlarging of the last ones will probably proceed further on in time.

[28] In my opinion, information related to the separation of matter and space is an intrinsic, immanent aspect of the real world, ignored by dynamics (mechanics). Without an incorporation of it into our theories, we will not be able to explain the beginning of the Universe in a satisfactory way.

As a matter of fact, black holes can "evaporate" due to the so-called "Hawking radiation." A spontaneous creation (permitted by quantum mechanics) of pairs particle-antiparticle – e.g. two photons, two quanta of electromagnetic radiation – occurs at the border of the horizon of events of a black hole (i.e. the area wherefrom no signals can leave, light included). In a strong gravitational force of a black hole, one of them may have positive energy, while the other – negative energy (as we remember, masses moving apart from each other in space is equivalent to a negative energy of gravitation; to be sure, photons do not have a resting mass, but they possess a determined energy equivalent to mass). The particle with positive energy is emitted into the cosmic space (hence the Hawking radiation), while the particle with a negative energy falls into the black hole, decreasing its positive energy and therefore mass (the Einsteinian equivalence of energy and mass). In this manner, the black hole diminishes in time – its mass (energy) is carried into space by the photon with positive energy (the sum of its energy and the negative energy of its partner falling into the black hole is of course zero).

However, the above-described process is extremely inefficient. Large black holes dispose of their mass very slowly, and the greater a black hole, the more slowly it "evaporates." The growth of black holes due to absorption of matter and joining of two holes into one compensates more than efficiently the waste of mass caused by the Hawking radiation. Therefore, this radiation seems to exert no significant influence on the above-presented scenario of the evolution of the Universe.

I have already mentioned that it is difficult to imagine life existing in the late stages of the evolution of the Universe, even if it is infinite and will exist forever. Although, from the formal point of view, the "thermal death" of the Universe will not take place in the sense that there will occur an increase of thermodynamic gradients (the span between maximal and actual entropy) due to expansion of space, the almost empty space with occasionally scattered black holes does not constitute a promising scenery for biological processes. Life is possible only in a certain span of the history of the Universe, when matter, still on its way to condensation into black holes, undergoes the evolution from the stage of hydrogen (and helium) to the stage of heavier elements during the process of thermonuclear synthesis in the interiors of stars. The above process is crucial for the existence of life for two reasons. Firstly, it delivers the building material – in the form of the heavier elements thrown out into space during explosions of Supernovae – for planets where life can develop, and also for life itself. Secondly, it leads to an emission of electromagnetic radiation that includes the vital visible part of the spectrum (this is just the shining of the Sun and stars), functioning as the original driving force for all life processes. Therefore, we have to accept the fact that the phenomenon of life constitutes only a temporary ephemera in the history of the Universe, regardless of its ultimate future. As it will turn out, life is (from the physical point of view) only one of the ways of dissipation of energy, or of an increase of entropy. At the moment when a certain defined **kind** of entropy increase – consisting in the thermonuclear synthesis in the interiors of stars – terminates, there will be simply no place for living processes in the Universe. Of course, the stage where conditions favouring origination and existence of life prevail will not end suddenly – life will gradually become less and less frequent, and its origination less and less probable.

Thermodynamic evolution of life

The second principle of thermodynamics, concerning inevitable entropy increase in time, have been already mentioned many times on the pages of this book. A logical conclusion can be drawn that, if there exists the second law of thermodynamics, there must also exist the first one. What does the **first** principle of thermodynamics – consigned to silence until now – refer to? Well, it states simply that the overall amount of energy remains **constant** in an isolated system (that does not exchange matter and energy with environment). It means that energy can neither originate from nothing nor vanish without trace. On the other hand, it can undergo transformation – and this is very important – from one form into another. As we know, energy may appear in many forms, as kinetic, potential, gravitational, thermal energy and so on. It should be remembered that matter also constitutes one of the forms of energy, in accordance with the famous equation of Einstein – the equivalence of mass (inseparable from matter) and energy. The first principle of thermodynamics allows therefore a transition of energy from one form into another, under the condition that the overall amount of energy remains unchanged.

On the other hand, the first principle says nothing about the direction of such transformations. It does not specify the forms of energy more willing to undergo transformation into other forms, neither the kind of energy preferred as the final stage of the whole sequence of transformations (evolution of forms) of energy. It its at this point that the second law of thermodynamics enters on stage. After a small re-formulation, it proclaims that energy has a tendency to transformation from a more organised (useful) form into a less ordered form (possessing a smaller ability to perform a useful work). In the scales of magnitude where the life processes occur (sufficiently smaller from astronomical orders of magnitude for the force of gravity to loose its crucial role), thermal energy is its most disordered form. Therefore, due to the second principle of thermodynamics, energy becomes finally dissipated into the form of heat.

Both principles of thermodynamics have important implications for physical aspects of the functioning of living organisms. The first principle teaches us that life does in no way **produce** energy, contrary to the frequent yet simplified usage that allows one to talk, for instance, about "production" of energy in the form of ATP. Life only **transforms** energy from one form into another. The energy stored in the form of ATP comes from the energy released during oxidation of nutritional substances (e.g. glucose), and its merit consists in universal accessibility for different processes in a cell that needing energy to take place. Speaking more generally, life constitutes only one of the processes participating in the dissipation of energy (transformation of its form with low entropy into a form with high entropy) on the surface of the terrestrial globe. This last process has in turn only a minute share in the overall energy transformations in the Universe.

The surface of the Earth receives significant amounts of only one kind of energy from only one source (if we leave out small amounts of energy still released from the interior of our planet). This is of course the electromagnetic radiation, particularly the visible radiation (light), emitted by the Sun. As it receives energy, the Earth must also get rid of it somehow, so that the amount of energy stored in a given moment on the surface of the Earth could remain more or less unchanged – if the surface of the Earth only absorbed energy without dispersing it, the temperature of the planet would increase infinitely. Practically the only way in which our planet disposes of energy consists in emitting infrared (heat) radiation into the cosmic space. As I have mentioned, the energetic balance of the surface of the Earth is exactly zero – the same amounts are absorbed and emitted. It is crucial, however, that it is visible radiation that is absorbed, while it is infrared radiation that is emitted. Such situation is possible thanks to the fact that the sky over our heads is characterised by a high degree of heterogeneity: a small area on it is occupied by a hot spot of the Sun, while the rest corresponds to the cold emptiness of the Cosmos. If the entire sky possessed the temperature of the Sun, the surface of the Earth would soon acquire the same temperature, reaching the state of thermal equilibrium with the sky, the state with maximal possible entropy in which all spontaneous processes, such as life, would become impossible. Fortunately, this does not happen. Cosmic space is much colder than the surfaces of stars, and the Earth (having an intermediate temperature) constitutes only a "stop" for the electromagnetic radiation emitted by our star into space, a stop which is exceptional in this sense that it changes one kind of radiation into another.

The visible radiation that reaches us from the Sun originates in transformation of another form of energy, namely the kinetic energy of gas molecules in the interior of the Sun. The latter, in turn, finds its source in the energy released during the process of binding two protons and two neutrons in a nucleus of helium in thermonuclear synthesis. Nucleons (protons and neutrons) in a nucleus are bound by a strong nuclear force. Nucleons moved apart in space (like masses exerting the gravitational force) possess a potential energy attributed with a negative sign. On the other hand, the potential energy of nucleons gathered in one place amounts zero. Therefore, joining nucleons into a nucleus involves a transition from negative energy to zero energy, or simply a **growth** of the energy of the system[29]. It is this energy that is released in thermonuclear synthesis and transformed into the kinetic energy of atoms and molecules. Thermonuclear reactions, in turn, are possible thanks to gravitation that squeezes together (compresses) protons and neutrons in the interiors of stars, so that strong nuclear forces operating in very small distances can become significant. Gravitation appears to be the initial "detonator" of the entire series of transformations of one energy form into another described above.

Let us remember that this chain comprises transformation of energy released in the process – occurring in the interiors of stars – of binding nucleons into a nucleus of helium (and then of heavier elements as well) into the kinetic energy of atoms, and next, into visible radiation[30], as well as transformation of visible radiation into infrared (heat) radiation – occurring on the surface of the Earth (the heat radiation undergoes afterwards dissipation in the Cosmos). The last process is of course most interesting for us, as it is the driving force for life phenomena. The transformation of visible radiation into infrared radiation means (the second principle of thermodynamics returns like a boomerang) that the latter possesses higher entropy. In accordance

[29] According to the Einsteinian equivalence of energy and mass, the nucleus of helium, composed of two protons and two neutrons, has a mass smaller than the sum of masses of free protons and neutrons. The difference in mass corresponds exactly to the amount of the released energy.

with quantum mechanics, the waves of electromagnetic radiation propagate in "packets" (portions) called quanta. The wavelength of visible radiation is smaller than the wavelength of infrared radiation, and therefore (again, due to quantum mechanics) the energy of quanta of visible radiation is larger. According to the first law of thermodynamics (the principle of energy conservation), the overall energy of radiation absorbed and emitted by the surface of the Earth must be the same. And therefore, the number of high-energy quanta of visible radiation absorbed by our planet is smaller than the number of low-energy quanta emitted in the form of infrared radiation. The energy of infrared radiation is therefore characterised by a greater dispersion is space (it is scattered over a greater number of places, less ordered), which means, of course, greater entropy.

Taking into account everything said above, from the point of view of thermodynamics of the Cosmos, dissipation of the energy of visible radiation into the form of the energy of infrared radiation, and therefore production of entropy, is the main "role" of the Earth's surface. Many processes are responsible for the realisation of this "task." The simplest of them is absorption and emission of radiation by bare rock (including desert sands). The surface of rocks warms up during a day by absorbing electromagnetic radiation of the (mainly) visible range of spectrum. Afterwards, both during a day and (chiefly) in the night, rocks give heat away (and undergo chilling) by emitting infrared radiation. This is presented in Fig. 1.8a. The transformation of visible radiation into heat radiation occurs here in two stages. First, the energy of visible radiation transforms into kinetic energy of the molecules of rocks, and then, this energy transforms into energy of infrared radiation.

[30] I ommit ultraviolet (UV) radiation on purpose, since a great part of it is absorbed by outer layers of the terrestrial atmosphere, while the remaining part of this radiation reaching the Earth's surface is not utilized by living organisms.

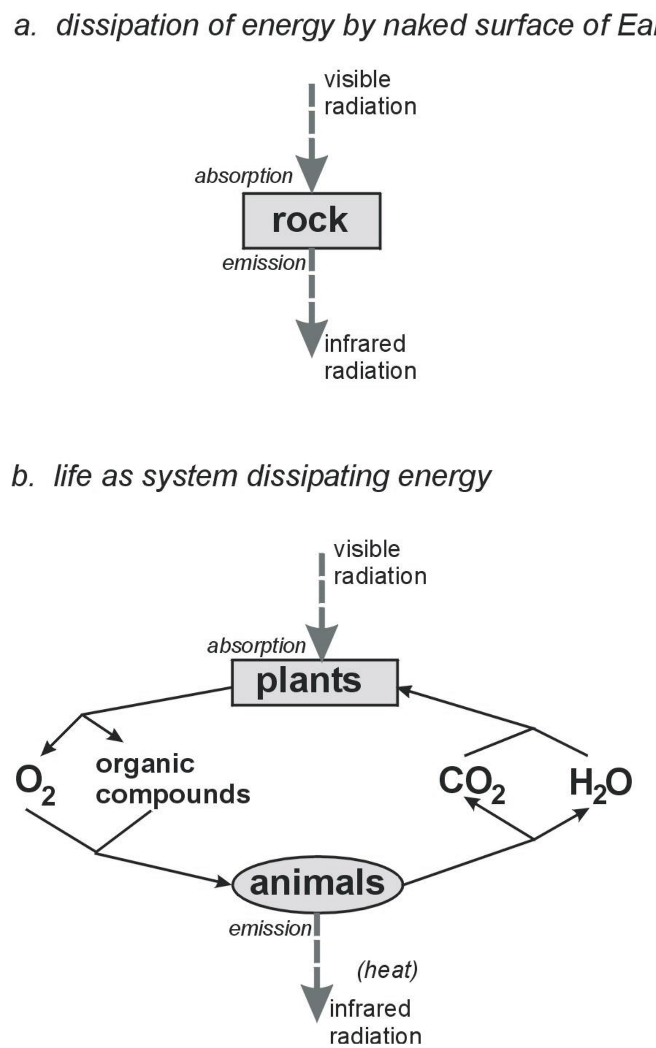

Fig. 1.8. Two possible ways of dissipation of energy by the surface of the Earth. Generally speaking, the dissipation of energy consists here in absorption of a small amount of high-energy quanta of visible radiation and emission of a large amount of low-energy quanta of infrared (heat) radiation. A series of processes occurs here, e.g. simple absorption and emission of radiation by rock (a), circulation of gases and liquids in atmospheric and hydrologic phenomena (not shown) as well as the circulation of matter in living systems (b).

The situation looks slightly more complex, in the case of warming up of ocean surface. The transformation of one kind of radiation into another sets here in motion a sequence of hydrological and atmospheric processes connected with a cyclic circulation of water. The surface of oceans (this concerns as well smaller reservoirs of water, e.g. seas and lakes) gives away the heat received from the Sun due to absorption of visible radiation by evaporating a certain amount of water molecules in the form of steam. Light rays transfer their energy to water molecules in oceans and increase the velocity of their movement (and therefore their kinetic energy), which enables steam to move up into higher layers of the atmosphere against the force of gravity, and thus to gather potential energy. There, the steam cools off, getting rid of a part of its energy by emitting infrared radiation that is sent into the cosmic space. As a result, steam forms clouds, where it undergoes condensation allowing it to fall down on the surface of the Earth in the form of rain. Thus, it transforms its potential energy into the kinetic energy of the drops of rain. The part of rainfall falling on land forms streams and rivers flowing back to oceans. The kinetic energy of flowing water – serving, for instance, to drive turbines producing electric current in water power plants (but also responsible for water erosion) – is nothing more than transformed energy of the radiation of the Sun. When the water finally finds its way to an ocean, it can again undergo evaporation and the cycle closes. Therefore, the transformation of visible radiation into infrared radiation drives here a certain cyclic process and forms a certain dynamic structure. Warming up and cooling off of air masses causes additionally such atmospheric phenomena as barometric fronts, winds and cyclones. The transformation (dissipation) of energy in the atmosphere and hydrosphere does not, therefore, bring about a completely chaotic behaviour of the molecules of water and air (though, atmospheric and hydrologic phenomena can be considered in a certain important sense as chaotic ones – compare the "butterfly effect" discussed before). On the contrary, one can see certain regularities, and therefore an ordering, in the behaviour of atmospheric currents. They have nothing to do with the structures formed by gravitation – the gravitational interaction between different parts of atmosphere is insignificant. Again, there happens something apparently contradictory to the spirit of the second law of thermodynamics – the degree of order grows instead of diminishing. Let us keep the problem in mind, leaving it to be solved later, and let us proceed further.

Biosphere – the complex of all living organisms – is yet another system dissipating energy on the surface of the Earth. The phenomenon of transformation of some forms of energy into other forms is here, like in the case of atmosphere and hydrosphere, a multi-stage process, closed into a cyclic sequence of transformations. This property is characteristic in general of the systems participating in the production of entropy (dissipation of energy) – compare the structures formed by convective currents that has also already been mentioned (and will be discussed in a moment). But let us keep first things first. The visible radiation is absorbed by plants and its energy is transformed in the process of photosynthesis into chemical energy of bonds between atoms in organic compounds. The primary product of photosynthesis is glucose, originating from carbon dioxide and water, while oxygen is released. The organic compounds synthesised by plants serve in turn as a building material and source of energy for animals. The latter combust nutritional substances with participation of oxygen, while carbon dioxide and water are created, and the whole cycle closes. The produced energy (stored in the form of ATP) is used by animals to perform different kinds of work: chemical work (synthesis of different compounds), mechanical work (movement), electrical work (formation of electric potential across cellular membranes) *etc*. Each of the processes involves certain energetic losses (this is valid for the metabolic transformations in plants as well), due to which a part of energy is dissipated into environment in the form of heat and finally undergoes emission into the cosmic space in the form of infrared radiation. This is presented in Fig. 1.8b. Additionally, saprophytes decompose dead bodies of living organisms, which also leads to production of heat. Therefore, life does exactly the same that naked rock as well as the atmo- and hydrosphere do – it transforms the visible radiation reaching the surface of the Earth into the infrared radiation dissipated into the emptiness of the Universe[31]. As in the case of atmospheric phenomena, this involves formation of ordered structures – it is not necessary to convince anybody that living organisms are characterised by a highly organised structure. After all, the entire biosphere is a complicated dynamic structure, organised not only in space, but first of all in time. The formation of such structures involves undoubtedly a drop in entropy. This is therefore the proper moment to address the question of how the phenomenon of life (and atmospheric phenomena as well) is related to the second principle of thermodynamics pronouncing the inevitable tendency of entropy to increase in a system.

Let us return for a moment to the already-mentioned dynamic structures related to energy dissipation, such as convective currents in a liquid. If we place a liquid, for example water, in a vessel warmed up from the bottom, then a certain gradient of temperatures will form between the

[31] It should be mentioned that the biosphere dissipates energy (transforms visible radiation into infrared radiation) much more effectively than naked rocks do. This is because, instead of absorbing visible rays and emitting the energy in the form of heat radiation, rocks reflect majority of visible rays that reach them. On the other hand, plants absorb a more significant part of the visible quanta they receive and transform them into the energy of chemical bonds, which is finally dissipated as heat. For this reason, the areas covered with plants look darker on satellite photographs than deserts or rocky mountains. In this sense, the existence of life is favoured for thermodynamic reasons, because life is **darker** than the naked surface of our planet, and therefore dissipates energy (produces entropy) more effectively.

bottom of the liquid layer (the bottom of the vessel) and its upper surface (it will contact air in an open vessel). The bottom limit of the water layer will be simply warmer than the upper limit, and the difference of temperatures will achieve the higher value, the more intensively the vessel is warmed up. The layer of liquid will of course transfer heat from its bottom (warmer) to the upper (colder) surface. However, the manner in which it will occur depends on the gradient (difference) of temperatures between both limits of the liquid layer.

For small gradients of temperature, heat is transferred by way of **conduction**. The liquid molecules situated at the bottom of the layer take over the kinetic energy from the bottom of the vessel, which increases the velocity of their movement. Due to their collisions, the energy is transferred to the molecules situated higher, that in turn transfer it further up, to the upper limit of the liquid, where the heat is given away, for example by evaporation. This is illustrated in the upper scheme of Fig. 1.9. The movement of molecules in the layer of liquid is here entirely **chaotic** (the movement of a given molecule is in no way correlated with the movements of other molecules), although the lower the molecules are situated, the quicker they move (the higher the temperature of the liquid).

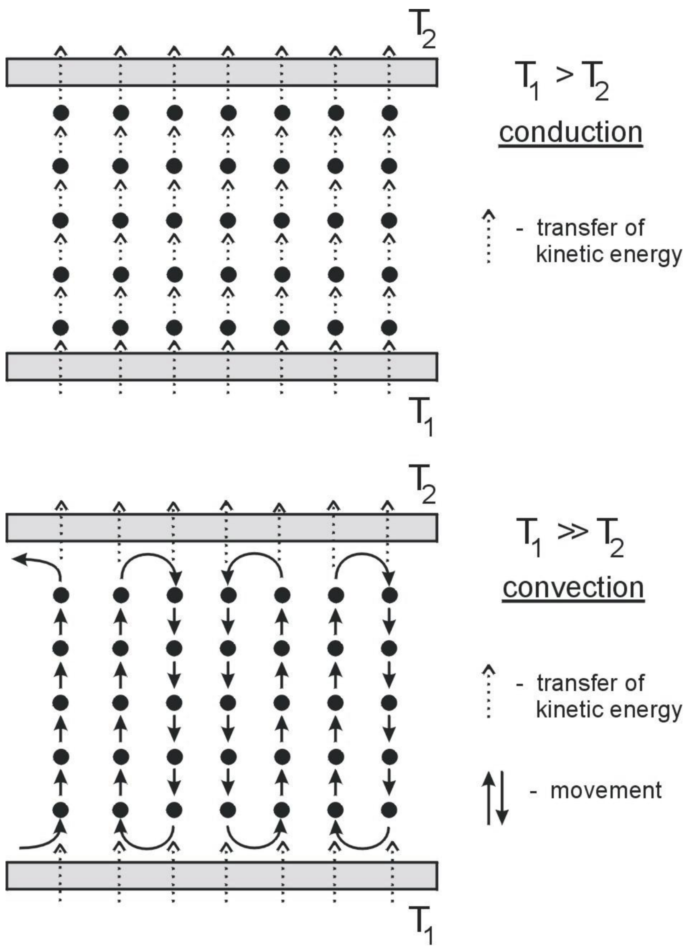

Fig. 1.9. The flow of heat through a layer of a liquid in the form of conduction and convection. Heat flows from the warmer bottom boundary of the layer with temperature T_1 to the colder upper limit of the layer with temperature T_2. If the difference of temperatures between the bottom and the top limit of the layer is small, the heat transfer takes the form of conduction, consisting in transferring of kinetic energy by molecules of liquid moving with greater speed (black balls situated lower) to slower molecules situated higher. In the case where the temperature gradient exceeds a certain critical value, there appear convective currents: a group of molecules of liquid, having taken over kinetic energy from the bottom limit of the layer, moves together as a warm upward current, to fall down in the form of a cold downward current, after giving away energy to the upper boundary of the layer.

However, the moment the difference of temperatures between the upper and lower limit of the layer exceeds a certain critical value, the conductance becomes an excessively inefficient mechanism, levelling the gradient of temperature, due to which there appears a completely new type of behaviour. The movement of molecules ceases to be chaotic, and it becomes **organised** into macroscopic structures called convective currents (and the entire process is called

convection). The molecules of liquid that "warmed up" and took heat over from the bottom limit of the layer, instead of transferring the kinetic energy to the molecules situated higher up, gather into an organised upward current. The movements of particular molecules become therefore mutually correlated. When they reach the upper limit of the layer, the molecules get rid of the gathered "surplus" of the kinetic energy by transferring it to the upper boundary of the liquid (to the air), and afterwards they move down as a downward current. There, due to the contact with the warmer bottom boundary of the liquid phase, their kinetic energy increases again and the cycle closes. The mechanism of convection is illustrated in the bottom scheme of Fig. 1.9. Convection allows for a significantly faster heat transfer than conduction does, and therefore it appears at higher thermodynamic gradients, where a significant displacement of a system from the state of equilibrium forces the tendency to a quick return to this state (i.e. to an intensive production of entropy).

The tendency of thermodynamic gradients(in this case – the gradient of temperature) towards equilibration, that is to an increase of entropy (entropy is maximal when the entire system has identical temperature), or to dissipation of energy (creation of thermodynamic gradients requires supply of an organised form of energy, and therefore decay of the gradient is equivalent to dissipation of energy), **forces** an origination of organised macroscopic structures, when the value of this gradient exceeds a certain critical value. We call them **dissipative structures**, because they are directly related to dissipation of energy. The dissipative structures (convective currents included) will of course disappear the moment equilibrium of temperatures is reached and entropy reaches its maximum. It happens so, because they are dynamic phenomena, driven by an increase of entropy, that is by dissipation of energy. Thus, we can see – although it may sound paradoxical – that the universal trend toward decay of an ordering can bring about a transitory order. In other words, the decay of order occurs in an orderly manner. A great contribution to theoretical explanation of the genesis of dissipative structures is due to Ilya Prigogine.

Similar dissipative structures form on the surface of our planet, although their scale is much greater. The before mentioned cyclic circulation of water in the atmosphere and hydrosphere is a good example. The analogy is very close: visible radiation absorbed by the surface of our planet corresponds to a higher temperature (is emitted by hotter bodies), while infrared radiation sent further into the Cosmos – to lower one. Let us remember that few high-energetic "visible" quanta of electromagnetic radiation are dispersed by the Earth in the form of many low-energetic "heat" quanta. It is the increase of entropy involved in this dispersion that brings about all kinds of "organised" atmospheric and hydrologic phenomena.

This is also the driving force of all life processes, conditioning the existence of the biosphere. From the point of view of thermodynamics, life is only one more (although undoubtedly a very complex) dissipative structure on the surface of our planet. Life is by its nature a circular process, involving a cyclic transformation of matter and energy[32], kept up by the thermodynamic gradient between the hot surface of the Sun (about 6000 Kelvins) and the cold space of the Cosmos (its "temperature" corresponds to the temperature of the background radiation, that is to about 3 Kelvins). If, for some reasons, the Sun went out and the differences of temperature vanished, life would not be possible.

All of this, however, does not explain how the origination of ordered dissipative structures, such as convective currents and living organisms, can be reconciled with the second law of thermodynamics proclaiming the inevitable decrease of order in time. If this principle is universally compulsory, then why the phenomenon of life does not comply with it? The answer appears to be trivially simple. Living organisms also respect the law of entropy increase in time. However, this law is related only to **isolated systems**, that do not exchange matter, energy and information (and therefore entropy) with their surroundings (the entire Universe is undoubtedly such a system). On the other hand, living creatures constitute **open systems**, i.e. a continuous stream of matter and energy flows through them. A living organism emits exactly the same amount of matter and energy, as it absorbs (a small part of the absorbed matter can be used as the building material for a growing organism, but this matter will be returned at the moment of death). However, due to metabolic transformations (of matter and energy), both matter and energy leave a living organism in a form completely different from the one in which they are absorbed.

[32] Many circular processes, driven by a flow-thorough and dissipation of energy, occur also on a smaller scale, e.g. in the cells of organisms. In the second part, I discuss the circulation of protons across the inner mitochondrial membrane, that is the process of pumping them outside – driven by energy received from electrons taken over from respiratory substrates – as well as their return inside, coupled with gathering and storing of energy in the form of ATP. Another example can be found in the continuous synthesis of ATP and the inverse transformation – the hydrolysis of ATP (ADP is formed) in reactions requiring energy for their occurrence. Many cyclic processes of this kind occur in the cell: here belong the circulation of ions (e.g. Na^+ and K^+) across the cellular membrane and the cyclic transformations of many metabolites (e.g. in the Krebs cycle or Calvin cycle). Life is similar to a certain extent to a complex of vortices (turbulences) in a liquid, or just to the convective currents appearing at a sufficiently high temperature gradient.

What an organism gives off to the environment possesses much higher entropy than what was earlier consumed. Atoms of carbon in a molecule of glucose in food are dislocated (as it has already been discussed) in a much more orderly manner than atoms of carbon in carbon dioxide, produced as a result of oxidation of glucose. The structure of glucose (the information related to it), vanishing during combustion of this sugar, becomes in a (quantitative, not qualitative) sense transformed into the information related to the structure of an organism. Therefore, while the entropy of an organism actually diminishes during its growth and development, the entropy of the environment increases even more. As a result, the entropy of the **whole system "organism + its environment"** increases in time as well! Thus, living things can be regarded as islands of order on a sea of disorder, and they contribute actively in increasing the overall disorder faster and faster.

Thermodynamic dissipative structures, such as living organisms, have a completely different nature than the gravitational structures, i.e. the heavenly bodies discussed before. While the latter constitute a direct expression of the universal tendency towards an increase of entropy, living organisms are a **local** deviation from this tendency. Feeding on the low entropy of their environment, they bring about an effective increase of this entropy, which allows them to decrease – only in a local and transitory manner – the entropy of their own bodies.

Summing up, thermodynamics does not exclude the phenomenon of life. On the contrary, the thermodynamic evolution of a system in a state far from equilibrium virtually imposes the formation of dissipative structures – living organisms included – actively participating in efficient production of entropy. The phenomenon of life can appear only at a certain stage of the thermodynamic evolution of the Universe – the one where conditions favouring origination and existence of life prevail. No matter what else life could be, it is also a structure effectively dissipating energy. This conclusion will close the first part of this book. In the second part, I will deal with other aspects of life, as well as the problems of what is its essence and what is decisive for the specificity of the biological evolution.

EVOLUTION OF LIFE

CLASSICAL FORMULATION OF BIOLOGICAL EVOLUTION

Life has existed on the Earth for about three billion eight hundred million years. This unimaginably long period constitutes roughly about one fourth of the age of the Universe, estimated at more or less fifteen billion years[33]. This time witnessed the generation of the entire presently observed and immeasurable abundance of living forms. All of them have originated in the process of biological evolution from simpler forms. Moving back in time, one reaches finally the moment when the germs of life originated spontaneously from inanimate matter. As the origination and development of life was a continuous natural process, it is necessary – in order to investigate the phenomenon of the spontaneous birth of life and its evolution – to determine exactly what the phenomenon of life consists in as well as what are its essential features, distinguishing it from inanimate physicochemical processes.

There are many characteristic attributes that are ascribed to the presently known living forms. The set, as rule, allows this phenomenon to be distinguished from the phenomena of the inanimate nature in a pretty univocal manner. This is due mainly to two reasons. Firstly, in spite of the enormous diversity of forms, the most fundamental principles of construction and function of all the organisms existing presently on the Earth are astonishingly similar (in fact: identical). Secondly, even the simplest present living forms exhibit a significantly higher degree of complexity than the first organisms. The complexity itself seems therefore to be a good determinant of life. However, a simple enumeration of the properties characterising the organisms on the Earth known to us carries certain dangers with it. On the one hand, one can easily ascribe the status of inseparable attributes of life to its accidental features, characteristic for the **present terrestrial life**, but not necessarily for what we would like to consider as the **life in general**. On the other hand, one should be careful not to mistake the properties representative for **all** living organisms for the features attributed only to **complex** biological systems. In other words, we want our criteria for dividing phenomena (objects) into living and inanimate to be so deep and universal, that it would be possible to apply them to the forms of life existing on other planets, to the first living organisms on the Earth, as well as to some presently known phenomena that arise justified doubts if they should be classified as living systems or not. In even simpler terms, the point is to reach the very essence of life, its most immanent characteristics.

Bearing in mind this far-reaching aim, I will try first to enumerate the most fundamental properties associated with a great majority of relatively very complex living systems known to us on the Earth. Then, I will perform a selection, separating the properties that can be regarded as immanent components of the essence of life from its secondary properties, less important for the understanding of this phenomenon. Finally – in the subsequent chapter devoted to a cybernetic formulation of the definition of life – I will try to give abstract, formal characteristics of living systems, reduced to an absolutely minimal quantity of necessary elements.

Here are the main properties attributed to living organisms existing presently on the Earth, presented in a possibly short form:

- they are built of **organic compounds**,
- they possess a complex, hierarchical **structure**,
- they exhibit a complex and hierarchical system of **functions**,
- hey involve a transformation of matter and energy, that is **metabolism**,
- they are characterised by an ability to **grow** and **move**,
- they **reproduce**, that is produce descendant organisms (systems),
- they **inherit genetic information**, that is transfer their properties to progeny,
- they undergo **evolution**, that is a transformation of one form into another.

Below, I will discuss these properties in more detail, analysing to what extent each of them is really specific exclusively for living organisms.

All living forms known to us are built of **organic compounds**, chiefly of nucleic acids and proteins, but also of a great amount of others, including carbohydrates (sugars), lipids *etc.* (organic compounds can be defined simply as compounds of carbon, reduced to a great extent, that is containing hydrogen; they may also contain other elements, the most important of which are oxygen, nitrogen and phosphorus). The specific structure and function of living organisms is mainly due to specific organic compounds. Among them, nucleic acids serve for recording genetic information and determining all (principal) properties of organisms. They also participate in the process of reading this information, that is in a synthesis – in a proper place, time and quantity – of appropriate proteins. Proteins constitute one of the most important structural and functional

[33] Different estimations of the age of the Universe give values between 10 and 20 billion years.

components of a cell, while a part of them, called enzymes, are responsible for the performance of a huge amount of biochemical reactions occurring in this cell. Lipids (fats) serve as the fundamental building material of cellular membranes. Finally, sugars can fulfil both the role of a building material (e.g. cellulose or simple sugars bound to proteins) and the function of storing a reserve of substances (monosugars, bisugars, starch), that are oxidised, if need be, to produce energy (ATP). An enormous amount of other organic compounds fulfils a great variety of functions. The fact that some determined kinds of organic compounds (among all the possible ones) play the basic role in the formation of structure and function certainly constitutes an inseparable attribute of the forms of life presently existing on the Earth. Although such objects as viroids – located at the border between life and "inanimate matter" – constitute simply a thread of a ribonucleic acid (RNA), their reproduction, or any living functions, cannot occur outside a host cell, containing the full set of organic compounds necessary for life. Presently, the terrestrial life is inseparably linked to the organic building material, particularly proteins and nucleic acids.

However, this does not need to apply to life in general, whose various forms may exist in different places of the Cosmos. It seems to be reasonable to assume that the essence of life consists in a determined kind of structure and function, rather than in a concrete sort of building material. In other words, it is not the chemical components of living organisms that are of the fundamental significance. Rather the kind of the interactions between these elements is crucial (which does not mean that some sorts of chemical compounds are not particularly predestined to be a building material of structurally and functionally complicated systems). Living beings are undoubtedly complex systems. However, the basic problem related to the essence of life focuses in the **kind** of complexity in question, rather than in the **substance** the systems are composed of. Organic compounds can originate outside living organisms, both as a result of natural physicochemical processes and in laboratories. They have been found even in some meteorites (particularly in the so-called carbonic chondrites). Therefore, these compounds are not inseparable from life processes. On the other hand, in principle, there is no reason why each kind of life should have to be based on exactly the same types of organic compounds that form the basis of the life on the Earth.

However, there exists another aspect of the matter that prevents us from concluding rashly that living systems, in principle, can be built of anything, as long as they are able to realise a determined complex of functions. Among the chemical elements known to us, only carbon forms chemical compounds (namely just organic compounds) complex enough to become components of highly complex living structures. Only organic compounds are characterised by a sufficient diversity, that allows some of them to be "adapted" to fulfil so specialised functions as recording of genetic information, catalytic activity, isolation from the environment and so on. A complex system – characterised by a high variety of mutual interrelations between different elements – cannot be composed of only one or a few kinds of these elements. For example, nobody will built a TV set of only diodes and resistors.

It is true that ultimately both a living organism and a television set are composed of several kinds of atoms (or, reaching even more deeply, from three elementary particles: protons, electrons and neutrons). This would suggest that complicated systems (undoubtedly a TV set is such a system) do not have to be composed of molecules of organic compounds. However, in a TV set – an object built with human hands – it is not single molecules that constitute the fundamental functional unit, but significantly greater elements such as wires or diodes, based on the crystals of semi-conductors. Such elements cannot originate spontaneously, like organic molecules, and their shape and function are not the result of the type of the [selected] building material. A copper wire connecting two distant sub-systems represents one of infinite possibilities of the spatial arrangement of x atoms of copper entering into its composition, and the chance for such a shape to form accidentally (not to mention connecting properly elements within a given system) is practically equal to zero.

Therefore, for life to be able to originate spontaneously, it must be composed of molecules representing one of the few 'necessary' – because determined by the nature of chemical bonds – configurations of atoms entering into their composition. The structure and function of these molecules must also result from physicochemical laws, e.g. from the spatial distribution of nuclei and electron coats. Only in organic compounds, are these atoms able to form sufficiently diverse combinations of connections. It seems, therefore, that life is "destined" to be based on organic compounds of carbon, although not necessarily exactly those we know as the components of life on our planet.

The concepts concerning, for instance, life based on silicon are highly speculative. It is possible to create complicated devices from this chemical element (and also from many others), but they seem to have nothing to do with life. The functioning of a TV set, to return once again to this example, is based on a man-designed shape of its elements (kinescope, resistors, capacitors and so on) as well as on the system of mutual connections between them. Both the structure of particular elements and the scheme of their interconnections have been intentionally imposed in the technological process of their production. A TV set could not originate spontaneously, in the absence of human intellect, since nothing in nature is able to "plan" its functional parts – they are not in any way 'necessary' (univocally determined), as they do not constitute a simple consequence of physicochemical laws.

On the other hand, the structure and properties of various organic compounds are determined by physicochemical laws (e.g. quantum mechanics, specifying the shape of atomic and molecular orbitals), conditioning specific properties of their molecules. These compounds can therefore originate without the participation of intellect. Living organisms – being complex agglomerations of organic compounds – in a sense projected their complicated structure and function by themselves, in the process of biological evolution. It was possible, because biological evolution could "start off" from relatively simple complexes of these compounds, formed spontaneously by physical and chemical forces (the origin of life will be discussed in one of the subsequent chapters). A spontaneous evolution of a TV set or another system based, for instance, on silicon would not have anything to start from, since physicochemical processes themselves do not lead to production of sufficiently complex and diversified silicon compounds, able to fulfil the function of progenitors of diodes or capacitators, and especially to arrange these elements into a coherent entity. The same concerns all other physical systems known to us, except organic compounds.

However, it should be stressed once more that, even if it is in fact true that organic compounds form the exclusive basis for life to originate (if it is to be a spontaneous process), then, they do not have to be necessarily of exactly the same kind as the compounds entering into the composition of terrestrial organisms. Anyway, the above restrictions refer only to the life formed in a spontaneous way, for it is possible to imagine organisms projected and produced in an artificial manner by rational beings. Such organisms could be composed of various materials, such as silicon, metals or plastic (I will return to the problem of artificial life further on).

The second property, inseparably connected with the phenomenon of life seems to be found in the complex **structure**, organised in a hierarchical manner, where systems from a lower level constitute components of systems at a higher level. Therefore, what we consider as a simple building unit at a higher level, an element of some superior system, constitutes itself a composed system at a lower level, containing elements of a yet lower order (particular elements are attributed with a certain degree of autonomy). Cells, for example, form tissues and organs of higher organisms, while, on the other hand, they are themselves composed of a nucleus, lisosomes, mitochondria, Golgi apparatus and other organelles. The hierarchical structure of living organisms contains many levels. Among the lowest, we can distinguish the level of atoms of various chemical elements, the level of simple chemical compounds and moieties, and further on, the level of organic macromolecules, such as proteins, nucleic acids, carbohydrates and fats. They are the building blocks of complexes of macromolecules, for example protein-lipid membranes, ribosomes or chromosomes, which, in turn, enter into the composition of various organelles (mitochondria, chloroplasts, cellular nucleus and so on). The next level is the level of entire cells, the components of which are constituted by the above organelles. This is followed by the level of tissues, formed from one or several types of cells (e.g. epidermal, muscular, neural cells), the level of organs composed of complexes of adequately spatially organised tissues, the level of systems of organs (e.g. alimentary canal divided into particular segments, equipped with glands producing digestive enzymes and so on), an finally – the level of entire organisms. The most complex organisms – including man and other vertebrates – contain all of the levels quoted above. On the other hand, the simplest living forms, for example bacteria, have stopped on the level of a single cell, loosing additionally on the way the level of cellular organelles (the lack of organelles like nucleus, mitochondria and lisosomes constitutes one of the fundamental features distinguishing non-nuclear organisms – *Prokaryota* (where bacteria and cyanobacteria belong) – from nuclear organisms – *Eukaryota* (including protozoa, algae, fungi, plants and animals).

Differentiation of particular structural levels constitutes undoubtedly a convenient way of describing the structure of living organisms. However, it must be emphasised that these levels have been singled out in a somewhat arbitrary way, and that it is not possible to determine an absolutely clear-cut division between them. As an example, one can take a plasmodium, that is a cell containing many nuclei. In the case of the alga *Caulerpa*, the entire plant, which is tens of centimetres long, divided into various "organs", including "leaves" and a "stem", is just one big plasmodium. It seems that we meet here an intermediate level between the level of a cell, on one hand, and the level of a tissue, or even of an organ, of the other hand. A similar relativity of the division into levels can be found for other levels as well.

Of course, possession of a structure certainly does not constitute a property attributed exclusively to living organisms. A structure (understood as a certain spatial differentiation, presence of different components) is possessed by a huge majority of physical systems, both static (e.g. crystals) and dynamic (e.g. convective currents). A complex structure is found in the scale of our entire planet as well as stars and galaxies. The Earth constitutes at the same time a perfect example of an inanimate object with a strongly pronounced hierarchical structure. The whole terrestrial globe is composed of atmosphere, hydrosphere (oceans and fresh waters), planetary crust, a liquid mantle appearing underneath it, as well as a centrally situated metallic nucleus. The terrestrial crust, for example, is composed of many kinds of rocks, frequently arranged in relation to each other in a very complex manner (resulting from the dynamic tectonics of the crust of our planet which, for instance, folds rocks, creates sedimentary rocks, brings about outflows of lava which is related to formation of basalts and so on). Particular rocks also possess a complex structure. They contain many different minerals that can assume the form of greater or smaller crystals. Of course, in the ultimate instance, all of them are composed of chemical compounds and atoms. Thus, one may distinguish many levels in the structure of our planet. This structure itself is hierarchical, as in the case of living organisms. It is possible to try to prove that the structure of living organisms exhibits nonetheless a higher degree complexity and regularity of repeatable substructures (e.g. cells) than the structure of physical objects. However, as there is not objective measure of complexity that would allow one to compare objects as different as a planet and a bacterium, an estimation of complexity can be very arbitrary at best (I suppose that many people would rather tend to grant a higher degree of complexity to a planet). Moreover, first organisms were most probably significantly simpler than present forms. Finally, complex structure and function are also characteristic for such man-made products as a computer, robot or TV set, the classification of which as living beings would rise our resolute objections. Therefore, it is certainly not a possession of a (hierarchical) structure alone that constitutes the principal determinant of the phenomenon of life.

Another property of living organisms is that they fulfil various **functions**, which are organised in a hierarchical manner, just as the elements of their structure. Among the several distinguishable levels, one can be indicate here the biochemical (metabolic) level, the cellular level as well as the physiological level, comprising the functioning of tissues, organs and systems of organs. Complexes of functions at a lower level form partially autonomic blocks that jointly perform some function at a higher level. For example, a complex set of biochemical reactions participating in the process of protein synthesis belongs to one of many blocks, enabling the functioning of cells, which requires an entire set of proteins, among them enzymes. Among such proteins in muscle cells, for instance, one finds actin and myosin that are responsible for contraction of muscle fibres and constitute the main component of their tissue. Another complex of biochemical reactions indispensable for the contraction of muscle fibres is the energetic block responsible for production of energy in the form of ATP (various blocks are of course mutually connected horizontally – the energetic block delivers energy (ATP) to the block of protein synthesis, while the block of protein synthesis produces enzymes entering into composition of the energetic block). The muscle contraction can fulfil different tasks at the physiological level: it can maintain a constant circulation of blood in the blood vascular system and thus provide all cells of the organism with oxygen and nutritional substances, as in the case of the heart; it can serve locomotion (skeletal muscles) or drive peristaltic movements of guts, causing a gradual flow of nutritional substances and, later, their non-digested reminders, through subsequent segments of the alimentary canal (smooth muscles). Similarly, an event at the biochemical level, such as synthesis of a given hormone, and subsequently at the cellular level (secretion of the hormone), serves the function of preserving homeostasis at the physiological level (e.g. control of the level of sugar in blood by glucagon and insulin). All these actions have their common superior purpose, which is survival of an entire organism and its reproduction.

A certain kind of purposefulness seems to be an important characteristic trait of the functioning of living organisms. This purposefulness is expressed, among others things, in a tendency towards the preservation of an organism as an intact entity in spite of changeable environmental conditions. A tendency of this type is sometimes referred to by the name of homeostasis. Taking up and assimilation of food – the source of building substances and energy – is a good example. Another manifestation of the fact that actions of living organisms are an expression of a tendency toward a decidedly determined aim is found in the production of descendent systems, possessing the properties of parental systems. These superior functions (survival and reproduction) govern an entire system of inferior functions on all levels of hierarchy, all of them characterised by purposefulness as well. A great majority of the processes proceeding in living organisms do not simply occur, but serve a purpose (retaining of homeostasis and production of progeny – from the point of view of the superior purpose).

What about the "functions" of non-biological systems? Physicochemical phenomena at the basis of a living system functioning are principally the same processes (in terms of their category) as those that occur in inanimate nature. In both cases, the same laws rule the behaviour of atoms and their complexes – no qualitatively new factor, no "vital force" decides about the essence of life. In inanimate systems, particularly in man-made devices, it is also possible to find a hierarchical ordering of different functions (which, in turn, appeared in first organisms in a rudimentary form). The only property specifically distinguishing the complex of functions related to life could be the above discussed purposefulness. However, this concept is to a great extent subjective (as many other concepts, such as complexity, referring to macroscopic systems). For one may just as well state that the purpose of feeding is survival, and that the purpose of gravitation is that stones should fall down on the Earth, and planets should revolve around the Sun. This questions the usefulness of the purposeful character of a function as a good indicator differentiating the specificity of life. What is more important, however, a purposefulness – very similar to the one found in living organisms – is met in such man-made products as robots or computers, designed to fulfil strictly determined tasks. Additionally, the functioning of these devices has a hierarchical structure, which strengthens even more their similarity to living systems. The operation of particular subsystems (frequently complex), such as the microprocessor or diode, participates in and co-creates the work of the whole, aimed to realise the purpose imposed by the constructors of the device. In conclusion, it is not in function as such where the essence of the phenomenon of life should be looked for.

The functioning of a living system on the biochemical level manifests itself in the form of metabolism, that is **transformation of matter and energy** flowing through the system. From their environment, organisms take up different substances, of which – after appropriate conversions – they build their bodies. They also draw energy from various sources, and transform it afterwards into some useful form (e.g. energy of chemical bonds in ATP), which can be used for performing different kinds of work, e.g. chemical work (synthesis of various chemical compounds), mechanical work (contraction of muscle, movement of vibraculum), electrical work (transduction of neural impulses) and so on. A part of this energy is dissipated in the form of heat. Autotrophic organisms, mainly plants and some micro-organisms, synthesise the components of their bodies from carbon dioxide, water and other inorganic components, while they utilise the Sun radiation or various inorganic compounds as the source of energy. On the other hand, organic compounds produced by other organisms serve as the source of both matter and energy for heterotrophic organisms. Generally speaking, transformation of matter and energy is a process in which matter and energy leave a living organism in a form completely different from that in which they were taken up. A significant fraction of energy is finally dissipated in the form of heat. As to matter, heterotrophic organisms transform organic substances (e.g. sugars, proteins, fats) and oxygen into carbon dioxide and water (a part of food, once transformed, servers as a building material, while not digested reminders are excreted), while an inverse transformation proceeds in autotrophs – water and carbon dioxide serve for synthesising carbohydrates and other organic compounds (oxygen is released). The metabolism in living organisms is a multi-step process, which involves a great number of chemical compounds and biocatalysts (enzymes).

Transformation of matter and energy occurs commonly in physical systems as well. Crystallisation of rocks under the influence of high pressures and temperatures or releasing of various gases to the atmosphere during volcano eruptions can serve as examples of the former. This last process constitutes at the same time an intermediate stage of process of transforming the energy liberated during disintegration (decay) of radioactive isotopes into heat energy, followed by the transport of this energy onto the surface of our planet, where the heat energy finally radiates into the Cosmos in the form of infrared waves. A similar process involves the radiation emitted by the Sun, which reaches the surface of the Earth. A relatively small number of high-energy quanta of electromagnetic radiation in the visible range of spectrum is absorbed by the surface of our planet, to be afterwards radiated in the form of low-energy quanta of infrared (heat) radiation. From the point of view of thermodynamics, life is nothing more than another process participating in this transformation – the direct (autotrophs) or indirect (heterotrophs) source of energy for the totality of life on the Earth is the Sun (the visible radiation emitted by it); this energy is finally released (radiated) in the form of heat (the infrared radiation related to it). The process of life as a certain form of transformation of matter and energy can, therefore, be characterised adequately in physical (thermodynamic) terms. Incidentally, perhaps the most drastic example of "transformation of matter and energy" in physical systems is thermonuclear synthesis in the interiors of stars. The lighter elements serve here as the substrate for the synthesis of heavier elements. Additionally, a small part of matter (mass) undergoes transformation into enormous amounts of energy (kinetic as well as radiative energy).

Generally speaking, living forms are also capable of **growth** as well as of **active movement**. Growth is observed both in bacteria cells, until they undergo division on two descendent cells, and in a man, during the period of his individual development. As to movement, it can mean translocation of entire organisms (running of an antelope, creeping of an earth-worm, swimming of an euglena driven by a vibraculum), or a change of location of its different parts in relation to each other (e.g. beating of the heart, movement of cytoplasm in a cell). A given system moves actively if the reason of the movement lies inside of the system, and is not a result of a passive reaction on influences of the environment (like for example a rolling of a poked stone down a slope).

The world of inanimate objects offers us a broad range of examples of both growth and movement. Crystals in a saturated solution of salt increase their dimensions; so do stalactites and stalagmites; the terrestrial globe was formed in a result of joining of rocky pieces. Planets revolve around the Sun and rivers flow down to the sea. To be sure, this is not active movement, since it is caused by the force of gravitation, external in relation to the considered objects. However, the movement of the surface of the terrestrial globe – that is the drift of continental plates, caused by convective currents in the mantle of our planet, which are in turn driven by the heat released in the process of radioactive decay of isotopes – should certainly be regarded as active (if we consider our entire planet as a system). Anyway, one does not have to search far away to find an excellent example of an inanimate object moving in an active way, for one finds it in an ordinary car.

Sometimes the ability to **receive of stimuli (perception)** from an environment and to **react** adequately to them is quoted as an attribute of life. Perception of a stimulus can assume as different forms as a picture of a mouse on the retina of an eye of a buzzard, on one hand, or stimulation of the production of enzymes responsible for lactose decomposition in bacterium by the presence of this sugar in the environment, on the other hand. The adequate reaction would consist in catching the mouse or transformation of lactose and its utilisation as an additional source of energy. However, it is enough to look at the thermostat-freezer system in a refrigerator – responsible for recording an increase in temperature and for the reaction in the form of its lowering – to reject the hypothesis living systems are exceptional with respect to perceiving stimuli and reacting to them. Another example one finds in a computer, "reacting" to various combinations of signals introduced via a keyboard by a complicated system of operations. If this analogy of perception seems extremely far fetched, one should consider the fact, that there have already been constructed devices ("perceptrons") able to recognise different patterns, for example the letters of alphabet. Moreover, the concept of perception is not sufficiently sharp. In a sense, the movement of a stone down a slope constitutes a reaction to a "perceived" stimulus – poking this stone by someone.

Life seems to exhibit three properties – strongly related to each other and constituting, to some extent, different faces of the same thing – completely absent from inanimate systems. Life is composed of **individuals**, which **propagate** and **evolve**. As I will prove further on, these three concepts (a living individual, propagation, biological evolution) cannot exist separately and they mutually co-define each other. For it can just as well be said that biological evolution is a change of properties in subsequent generations of reproducing living individuals, or that a living individual is a creature capable of reproduction (production of descendent individuals and transfer of its properties to them = heredity) as well as of evolution (gradual change of these properties in subsequent generations). Now I will try to specify how the terms of propagation (connected with heredity), biological evolution and living individual should be understood.

The first essential property of living systems is the capability of auto-copying (**propagation**), that is of production of descendent systems. This rush toward a quantitative increment is expressed in the production of new objects, similar to the parental one, and not in spreading of some organic mass constituting one entity. Life is therefore composed of particular individuals, separate organisms. Taking up substances (and energy) necessary for this from their environment they produce individuals similar to themselves in the process of reproduction. The "reaction" of producing new individuals by already existing ones can be presented in the following way:

X parental individuals + building substances → Y descendent individuals

where the following condition is (at least potentially) fulfilled: Y > X. Since in an appropriately long interval of time the number of produced descendent organisms is always greater than the number of exist parental organisms, the quantity of living organisms exhibits a natural tendency to increase, limited only by the resources of the environment. The "production cycle" of descendent individuals by parental individuals is presented in Fig. 2.1.

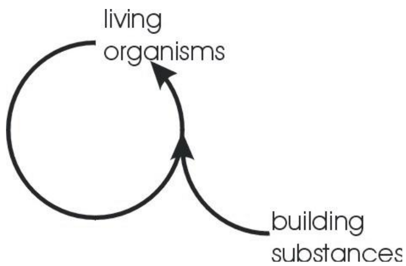

Fig.2.1. The cycle of reproduction of living organisms. The existing organisms produce descendent organisms from building substances taken up from an environment.

The above presented process of reproduction is not unique for the phenomenon of life. Chemistry knows the so-called **autocatalytic** reactions, where (simple, monomeric) molecules of a certain compound (let us call it A) catalyse the formation of further molecules of this compound from molecules of another compound (X). A reaction of this type has the following form:

$$\mathbf{A + X \rightarrow 2A}$$

and, from the formal point of view, it is strictly analogous to the above described "reaction of synthesis" of new organisms by already existing organisms from substances drawn from the environment. Fig. 2.2, presenting an autocatalytic cycle of simple chemical molecules, is analogous to Fig. 2.1. Therefore, propagation – understood as a production by a type of systems of objects similar to them (auto-copying) – does not constitute a property attributed only to living organisms. Another example of inanimate objects capable of "reproduction" (and also of taking up substances from an environment and of growth) are coacervates which I discuss in more detail in one of the next chapters. Also (free) neutrons undergo the process of auto-copying in a nuclear chain reaction, in which hitting a nucleus of a radioactive isotope of a chemical element (e.g. uranium) by one "parental" neutron leads – if this nucleus is split into two lighter nucleuses – to production of three "descendent" neutrons.

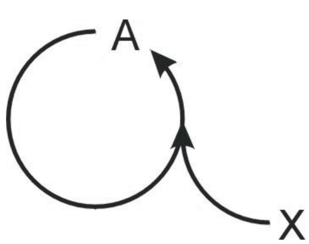

Fig.2.2. The autocatalytic cycle of compound A. Compound A catalyses formation of new molecules of compound A from compound X.

However, the reproduction of living individuals differs by one very important property from the "reproduction" of the physicochemical objects discussed above. For in living organisms, one observes the process of inheriting the properties, that is the transfer of their **identity** to descendent systems. A given neutron or a molecule of a particular simple, monomeric chemical compound cannot transfer their identity to "descendent" neutrons or molecules, because they do not possess such identity. Neutrons and molecules of a particular **simple** compound are as they must be, considering elementary physical and chemical laws. Their structure does not carry any **information**, because this is the only possible structure. Having at disposal the laws of physics (not necessarily those known until now), one can (at least in principle) determine the properties of any neutron or any molecule of compound A, since they are '**necessary'** and no additional information is needed here. Additionally, considering their ideal sameness, no **competition** occurs between particular neutrons or molecules of a simple chemical compound. For none of them are "better" or "worse" than others in the efficiency of auto-copying. As will be shown further on, it is the competition between different individuals – adopted to a different degree to survive and produce progeny – that is decisive for the essence of biological evolution and life.

A completely different situation is observed in the case of living individuals, or at least of such organic compounds, fundamental for life, as proteins and nucleic acids. Particular molecules of proteins and nucleic acids can differ in the sequence of aminoacids and nucleotides, respectively, while the dissimilarity of particular living individuals can adopt innumerable forms, conditioned by various combinations of their properties. The number of possible combinations of sequences of aminoacids in a protein chain containing 100 aminoacids is virtually astronomical (20^{100}, because there exist 20 kinds of aminoacids), but still incomparably smaller than the amount of various organisms which could potentially function on the Earth. *Nota bene*, all hereditary properties of living organisms are just a derivative of the sequence of monomers in the above mentioned macromolecules, especially in nucleic acids. Particular living individuals possess a very strongly expressed identity – thus, there is much to be inherited. The structure of a horse does not result, in a necessary manner, from physical or chemical laws applied to the atoms of which it is composed.

To describe this structure completely, it is necessary to deliver additionally a huge amount of **information**, which includes the information about the structure of the horse's genetic record, manifesting itself in the sequence of nucleotides in DNA. But this would not suffice yet. The isolated DNA of a horse alone will never develop into a horse: it also needs proteins, "taking care" of its protection and reproduction as well as remembering (aminoacyl-tRNA synthase) the genetic code, the whole fertilised egg cell (zygote), ensuring proper conditions for the action of proteins, finally a mare able to carry the embryo until the moment of birth, and then, to feed and protect the colt. The inherited and evolving biological identity of a given individual of horse is therefore an entire complicated system. The genetic information recorded in DNA[34] is only one of its elements. The evolutionary identity (complex of hereditary individual properties) of two horses is always slightly dissimilar. The identities of a horse and of a hear differ much more drastically. Various identities can, under given environmental conditions, have a different degree of efficiency in ensuring the survival of an individual and in producing offspring, and therefore in transferring and eventual spreading (propagation) of itself. In organisms reproducing sexually, of course, only a half of the identity of each parent is transferred onto a given descendent. However, there is no reason for identity – understood as a complex of properties – to be indivisible.

In conclusion, an individual is not only a system producing systems similar to itself, but also a system possessing identity (individual hereditary properties) and transferring this identity to its progeny. The identity is equivalent to the information about a not indispensable and therefore, in a sense, arbitrary structure of a given system. The more possibilities of choice, the more information is carried in a choice of one of them. A simple chemical molecule catalysing formation of further identical molecules has "at choice" only one possibility concerning its own structure, and therefore, it possesses zero amount of information. Such a molecule cannot have an individuality, and thus it cannot have identity. However, already organic polymeric macromolecules – such as nucleic acids and proteins which (as we will see later) gave the origin

[34]The temporal continuity of genetic information recorded in DNA alone is insufficient to ensure a production of next generations of living organisms. It also requires the temporal continuity of the entire system responsible for DNA replication, and for reproduction itself (the last process involves also a reading of genetic information, that is synthesis of proteins). Therefore, this is not the genetic information (set of genes) alone that corresponds to the identity of a living individual, but the structure of the whole capable of autocopying a system which is identical with this individual. This is not a complex of Dawkins's selfish genes that evolves, but an entire reproducible system.I will discuss this problem in more detail later on.

to life – possess a resolute identity determined by the sequence of their monomers (nucleotides or aminoacids). Living individuals – being very complex systems – have identities pronounced even more strongly.

The identity transferred from generation to generation of propagating living individuals undergoes **evolution**. If one follows the history of life down to the moment of its origin, one notices a gradual transformation of some forms into others and a divergent evolution of different forms coming from a common ancestor, until the process produces the enormous diversity of presently observed biological species. These changes can lead to an increase in complexity of living forms (although this does not happen always), and therefore to still stronger and stronger pronouncement of their individuality (identity). After all, the first living organisms had a form of some primitive proto-cells, while man – one of the presently existing organisms formed through evolution – possesses the most complex of all systems in the Universe known to us, namely the human brain. Apart from complexity, living organisms are characterised by excellent adaptation to the environments they live in. However, it should be clearly emphasised that this rush towards complexity and adaptation is not a result of any preimposed instructions, but constitutes a derivative of basic mechanisms of the process of evolution, operating "from-bottom-to-top," so to say.

Natural selection is the main of the mechanisms underlaying biological evolution. This process results directly from the fact of reproduction of living individuals as well as from heredity, that is transfer of individual's own identity (complex of individual properties) to its progeny, as well as from the physical nature of living organisms and their environment. If we start from these "premises", the natural selection, and therefore evolution as well, become in fact inevitable. As I have discussed before, reproduction and inheriting of properties constitute an inseparable attribute of life. Life therefore, once it has originated, must evolve. Natural selection constitutes a logical consequence of the very essence of living organisms. The ability to evolve may therefore be indicated as an inseparable component of the definition of these organisms. By drawing logical conclusions from the premises at our disposal, let us therefore investigate where this inseparable interrelation between life and evolution comes from.

The first premise states that, as we remember, organisms propagate and therefore produce descendent organisms that inherit the properties (identity) of their parents. In the simplest case – of a bacteria, for example – a cell of bacterium undergoes division into two cells. These cells divide further, yielding 4, 8, 16, 32... cells. In optimal conditions – i.e. if the amount of food and place is appropriate, if the environment is not polluted with the products of metabolism, and so on – the amount of bacteria would increase in a geometrical progression, tending to infinity. After a relatively short time, the progeny of one bacterium dividing every 20 minutes would constitute a huge sphere expanding with the speed of light. A similar reasoning applies to all other organisms. The famous Fibbonacci's sequence is a good example. It describes the amount of progeny produced in subsequent years by a pair of rabbits (and then by this progeny). All sequences of this kind increase exponentially, tending to infinity.

For bacteria reproducing asexually (by division), this process is presented in Fig. 2.3. Each individual bacterium divides into two descendent individuals which survive and divide further on. In subsequent generations we have 1, 2, 4, 8 ... individuals (cells). Their quantity increases unlimited to infinity. In the case of sexually reproducing individuals an analogous scheme would be slightly more complex, although this does not have any consequences for the course of our reasoning. Therefore, subsequent schemes will also be limited to the consideration of asexually propagating organisms.

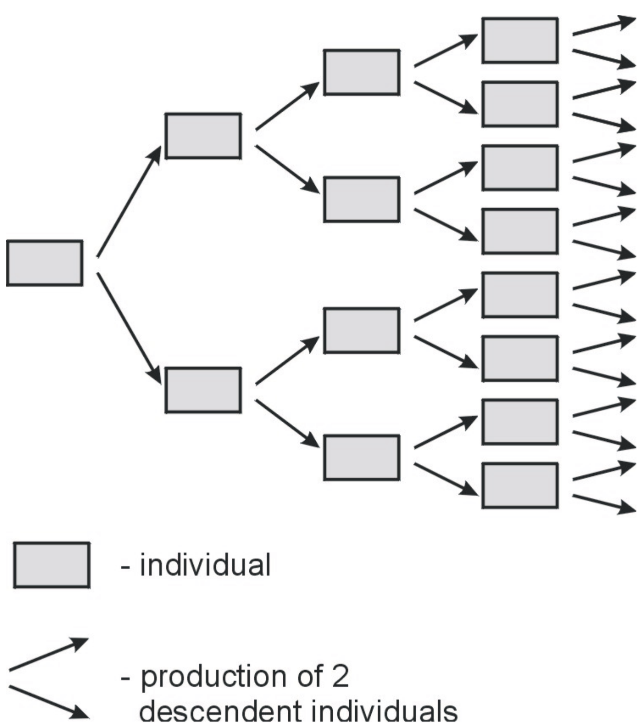

Fig.2.3. Reproduction of living organisms if there were no environmental limitations. Each individual (represented by a rectangle) produces two descendent individuals. The number of individuals increases in a geometric progression, tending to infinity.

The fact of propagation of living organisms, therefore, leads inevitably to the conclusion that the numerical force of a population of any organism would grow to infinity in ideal conditions. However, such an unlimited propagation of organisms never occurs in the real world, which is due to a **limited capacity of an environment**. Each organism occupies some place in space and takes up from the surroundings matter and energy in a proper form, necessary for it to live. The amount of both place and building-energetic resources on the surface of our planet is of course limited. Additionally, all organisms live in some concrete environments, specific for them. Their capacity is frequently much smaller in comparison with the capacity of the entire terrestrial globe. Each environment allows for the existence of only a limited number of living organisms. Once this number is reached – and the capacity of an environment exhausted – a part of the produced progeny must die. Only the part permitted by the resources of an environment can survive and further reproduce. This results from the premise recognising the physical realities of the world.

Fig. 2.4 describes such a situation. The capacity of an environment is here limited to four individuals (this artificially small number is to allow for a transparent schematic presentation; in real situations, it can amount to thousands or millions). If we start from one individual, then the capacity of the environment is saturated already after the third generation. Because the number of individuals cannot exceed 4, each time only one (on average) of two descendent individuals survives, while the other must perish without progeny. Therefore, each of the individuals produces now (statistically) one surviving descendent individual, which is illustrated by single arrows (in the case of sexual reproduction, two parents produce on average two descendants able to survive and further reproduce). Generally, a half of the individuals originated in each generation must die. If some organism, unlike bacteria, would produce more than two descendent individuals, then an appropriately greater part of this progeny would have to die. The principal question for biological evolution is **which** individuals survive and reproduce, and **which** of them die without offspring (Fig. 2.4).

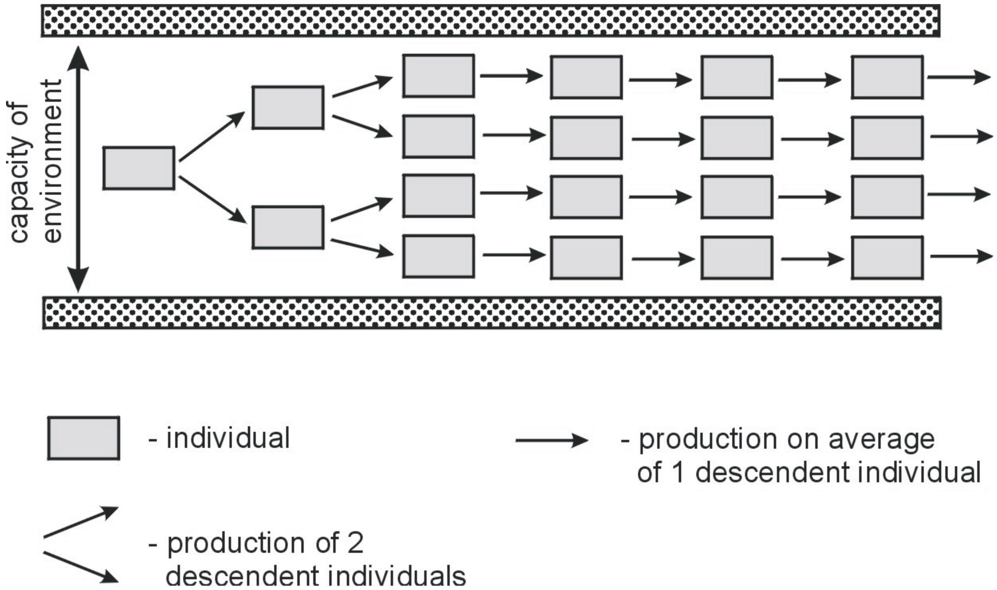

Fig.2.4. Reproduction of individuals when the capacity of environment is limited. Once this capacity is exhausted (artificially low in the scheme, equal to 4 individuals), only such number of the progeny can survive that is necessary to substitute parents. In the case of bacteria cells dividing into two descendent cells, statistically only one of them will survive and divide. The other one will die before reproduction. Of course, a chance can cause that there will survive and divide two descendent cells of one bacterium, while both descendent cells of other bacterium will die. However, all bacteria cells have in principle identical chances for leaving the average of one descendent cell.

And here the genetically-conditioned individuality (**identity**) of particular individuals enters the stage. It was assumed in Fig. 2.4 that all individuals are **identical** (and therefore that they do not possess identity), due to which each one has an equal chance to leave the same number of descendants, namely one. Pure chance decides that sometimes there will survive both descendants of one individual, while none descendants of another individual.

However, particular individuals in a population of some real species are never identical genetically. They are **genetically variable**. Descendent individuals can differ from parental individuals, while individuals of the same generation differ one from another. Therefore, it will depend to a great extent on individuals themselves, which of them will survive and propagate, and which will die without progeny, although, undoubtedly, chance will also play some role here.

Genetic variability (and therefore possession of identity by particular individuals) is **inevitable** and results from the **physical nature of the carrier of genetic information**, that is **DNA**. To build its own structure, a living organism needs some plan, describing this structure, as well as a certain algorithm, that is a set of instructions of how such a plan (record) should be interpreted (read). The sequence of nucleotides in DNA (or else in RNA, in the case of some viruses) is the universal manner of recording this plan for all life on the Earth, while a complicated system of a reading of genetic information – "translating" the genetic record (via the language of the genetic code, regulation of expression of genes and so on) into the structure and function of an entire organism – is the interpreter of this plan. In order to produce a descendent organism possibly similar to itself (i.e. of a possibly close identity), a parental organism must transfer this plan of structure and function to the descendent organism, which requires creation of an exact copy of its own plan. It is, therefore, necessary for it to rewrite its own sequence of nucleotides in DNA onto such a sequence in its twin descendent copies. Copying of genetic information consists in separating the double helix of DNA into two single threads, as well as in adding a complementary thread to each of the two. This complementarity is produced by spatial matching of nitric bases (entering into composition of nucleotides) laying in pairs opposite to each another: A-T and C-G (A, T, C, and G are abbreviations of the names of four nitric bases entering into composition of DNA). For a descendent individual to constitute a perfectly faithful copy of its parent (at least potentially, i.e. after a neglecting the influences of an environment and random processes), the genetic record would have to be copied with **infinite** accuracy[35]. Such an accuracy, however, is in principle impossible for purely physical reasons. Copying of the genetic record cannot be perfect, due to chaotic thermal movements of molecules, always present in temperatures higher than the absolute zero. They disturb the process of DNA copying (replication) and can lead to the insertion of an improper nucleotide (containing different nitric base than it is required by the principle of spatial complementarity to the opposite base in the maternal thread of DNA) into the newly synthesised thread. In this way, the sequence of nucleotides in the descendent thread of DNA, and therefore the genetic information co-conditioning the identity of an organism, will be changed. This change will be transferred to all threads of DNA synthesised from this modified thread. Such change is called **mutation**. If a mutation occurs in the genome of a unicellular organism, or in the genome of a cell of a generative (sexual) line in a multi-cellular organism, then

[35]This concerns of course only asexually reproducing organisms. In the case of sexual reproduction, each of the parents has to enter into a genetic compromise with the other, if both contribute more or less a half of the genetic record of their descendent. Its plan of structure (its identity) is therefore, in a sense, a resultant of the plans of structure (identities) of parental individuals.

it will be **inherited**, just as the remaining, unchanged part of genetic information. A mutation in some gene means an origination of a new **allele** of this gene. An individual with a mutation will differ genetically from the not-mutated individuals. Many mutations exert an important influence on the phenotype of an individual, that is on its final form (while genotype is a recipe for an organism, phenotype constitutes a ready product, conditioned also – apart from genotype – by influences of an environment). Due to such mutations an individual will differ from others in phenotypic respect as well. It should be emphasised that mutations are accidental and non-directional, which means that they originate fortuitously and do not tend towards a change of genetic information in any determined direction.

The above described mutations – also called punctual mutations and consisting in the replacement of one nucleotide for another, or in deletion or insertion of a nucleotide – lead to the origination of new alleles of a given gene, that is of its slightly changed versions situated in the same place (*locus*) in genome as the original gene. However, the origination of completely new genes is also possible. It occurs by way of duplication of a primary gene. Two identical copies – located one beside the other or occupying totally different places (*loci*) in genome **–** originate as a result of this mutation. One of these copies, as in the previous case, codes some protein fulfilling a certain function in an organism. On the other hand, the second copy can undergo different punctual mutations and thus cumulate genetic changes in itself. This does not threaten with a loss of functionality of the produced protein because this protein is still synthesised by the first, unchanged copy of the gene. In this way, the second copy would became a new gene coding a new protein. Such a new protein could incidentally appear to be able to fulfil in the organism some useful function, differing from the original one, and thus, it increases the efficiency and degree of complexity of this organism. New genes can also originate as a result of "shuffling" (arranging in different combinations) of exons (coding fragments of eukaryotic genes, separated by non-coding fragments – introns) responsible for coding particular domains of proteins (fragments of protein molecules fulfilling particular functions, e.g. enzymatic catalysis, anchoring in a membrane, binding ions of metals and so on). It is also possible, especially in micro-organisms, for genes to undergo a horizontal transfer, that is a transfer between two organisms, frequently unrelated with each other.

Recombination is an additional source of genetic variability of individuals in a population, secondary in relation to mutations. It can consist in mixing of genetic information in sexual reproduction due to the fact that a descendant obtains a half of its chromosomes from each of its parents, while these chromosomes undergo earlier a random segregation during their allocation to an egg and a spermatozoon. This increases significantly the number of combinations of various alleles of particular genes situated in different chromosomes, which originated in different individuals in a population. On the other hand, in organisms with the asexual type of reproduction, in order for two mutations to meet, they would have to originate within the same clone, which is much less probable.

Apart from this, genetic information can also be "shuffled" through an exchange of parts of chromosomes between two homologous chromosomes in diploid organisms. This phenomenon is called crossing-over. Here arises the possibility of an arrangement of the alleles of genes in the same chromosome into different combinations. Both the sexual reproduction and crossing-over are an efficient source of recombinative genetic variability. However, recombination is called a secondary generator of variability, since it does nothing more, but mixing and shuffling of the elements within the range of genetic variability created earlier by mutations.

The joint effect of mutations and recombinations produces genetic differentiation of individuals within a population. Let us return to Fig. 2.4 and see what will happen if a mutation occurs in the genome of one of the individuals. We distinguish three types of mutations, according to the kind of the effect exerted on an individual – their carrier. They are neutral, harmful and profitable mutations. **Neutral** mutations bring neither advantage nor disadvantage. This is the case when – due to the property of degeneration of the genetic code (some aminoacids are coded by several different triplets of nucleotides) – a substitution of a given nucleotide in a codon (a triplet of nucleotides coding an aminoacid) with another nucleotide does not change the coded aminoacid. In other cases, even if another aminoacid is coded as a result of a mutation, this fact can have no influence on the structure and function of an entire protein, especially if the aminoacid in question is situated in a protein chain outside the important regions of this protein, e.g. active centre of an enzyme. Finally, a mutation may cause some perceptible phenotypic effect, which, however, has no influence on the functioning of an organism – for example a change in the distribution of spots on the shell of a snail.

Generally speaking, neutral mutations do not influence the chances of survival (in a given environment) of their carrier-individuals, or their ability to leave progeny. They reproduce as effectively as individuals without a mutation and therefore their offspring will exist in a population together with the offspring of the non-mutated individuals. This is illustrated in Fig. 2.5. As it can be seen, a neutral mutation changes nothing. A mutated individual produces descendent individuals with an identical efficiency as it did before the mutation occurred, and therefore it has statistically the same chances to transfer a mutated gene to the next generation, as other individuals have to transfer non-mutated genes. A gathering (cumulation) of neutral mutations can lead to a change of the genetic composition of a population, which may (but does not have to) appear advantageous in new conditions. However, neutral mutations alone do not lead by themselves to an increase of adaptation or complexity of organisms.

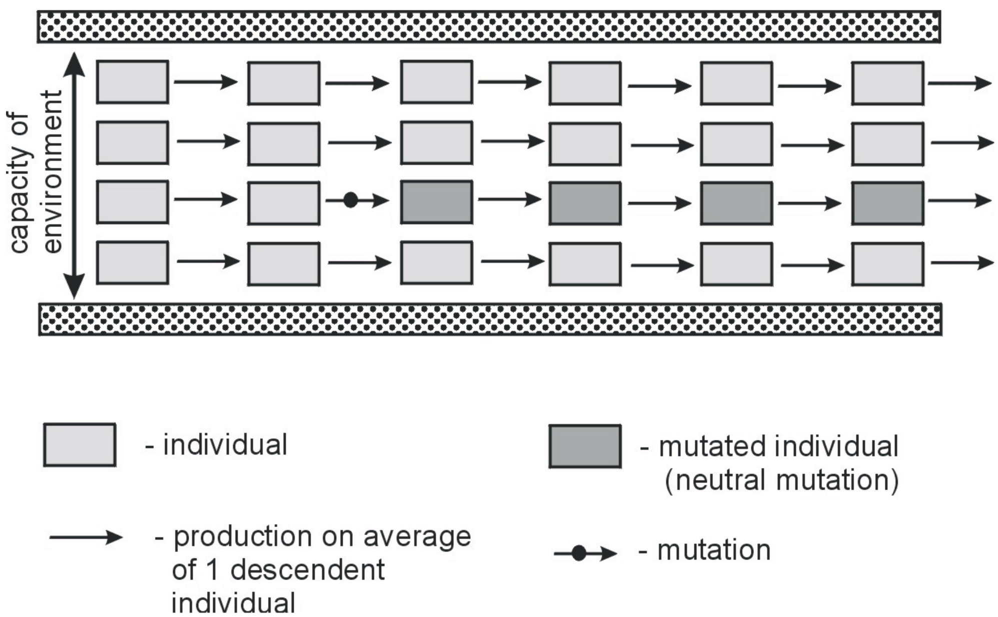

Fig.2.5. The effect of a neutral mutation. A neutral mutation (dot) causes the appearance of a mutated individual, which has statistically the same chances of survival and production of progeny as non-mutated individuals. A neutral mutation will be neither eliminated nor propagated in a population.

Many mutations are **harmful**, and therefore influence negatively the chances of survival and production of progeny. In the extreme case of lethal mutations, they are followed by an inevitable death of organisms possessing them, before reproductive ability is reached. A mutation in a gene coding some enzyme indispensable for metabolism, that brings about a loss of catalytic abilities of this enzyme, can serve an example. Frequently, the effect of a harmful mutation is less dramatic. It decreases only insignificantly the efficiency in survival and reproduction. In both cases however, the mutated individuals will sooner or later (in the case of a lethal mutation – instantly) be eliminated from the population by non-mutated individuals that are more effective in survival and propagation. This situation is presented in Fig. 2.6. We see that the place of a potential descendant of an individual with a lethal mutation is taken over by a descendant of another, non-mutated individual, filling up the "gap" (in a complete filling of the capacity of an environment) formed by the death of an individual bearing this mutation.

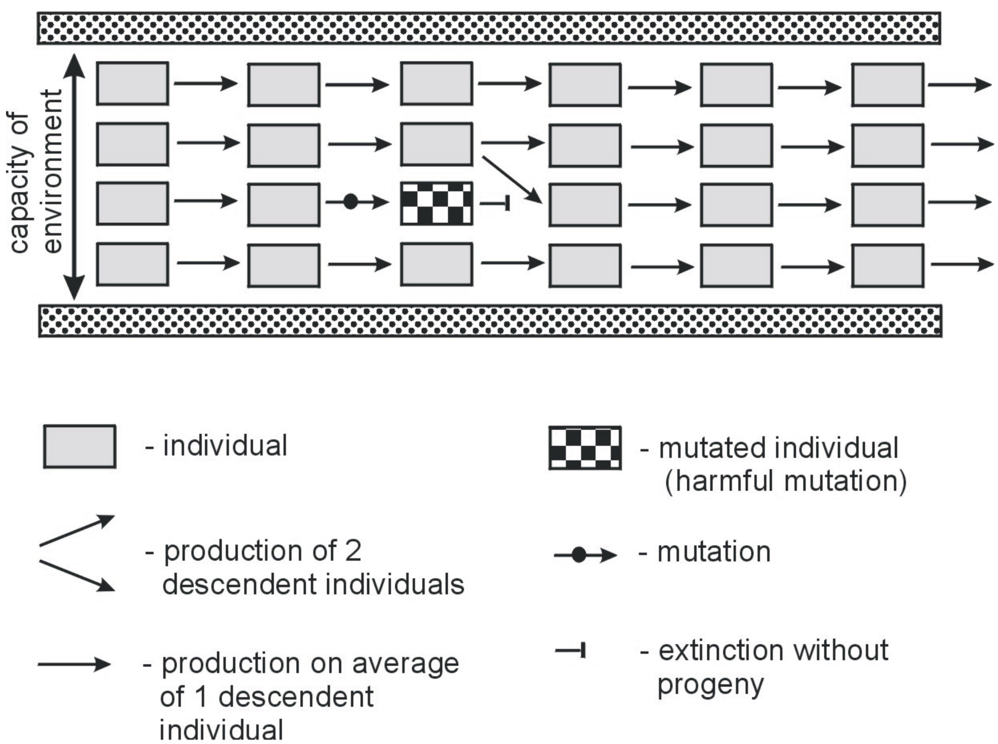

Fig.2.6. The effect of a harmful mutation. A harmful mutation (dot) diminishes the chances of survival or/and reproduction, and therefore the progeny of mutated individuals, also possessing this mutation, will sooner or later be ousted by the progeny of non-mutated individuals. A harmful mutation will be eliminated from a population.

Few mutations can turn out to be **profitable**, and therefore to give to a mutated individual an advantage over non-mutated individuals in the domains of the effectiveness in survival and/or production of progeny. If a mutation results in the appearance of an enzyme, permitting a decomposition of a nutritional substance present in the environment which could not be assimilated before, then a mutated individual will gain an obvious advantage over the individuals devoid of this enzyme. It will be easier for such and individual to avoid starvation, grow faster and produce more progeny. Its progeny will gradually oust (eliminate) the progeny of other individuals, until the entire population will be composed of individuals possessing this enzyme. This case is illustrated in Fig. 2.7. Here, two generations after the appearance of an individual with a profitable mutation, the entire capacity of an environment becomes occupied by individuals that are carriers of this mutation (a profitable mutation can additionally lead to an increase of the capacity of an environment – more individuals will be able to survive and reproduce due to the accessibility of a new nutritional substance). A profitable mutation produces an organism better adopted to its environment. The situation described above can be repeated and a subsequent profitable mutation can become consolidated in the population. Genetic changes of individuals in subsequent generations will be of cumulative character. Therefore, after a sufficiently long period of time, descendent individuals will already differ significantly from their ancestors. Finally, these differences will be so significant that a new species will originate. The above process is called **evolution**.

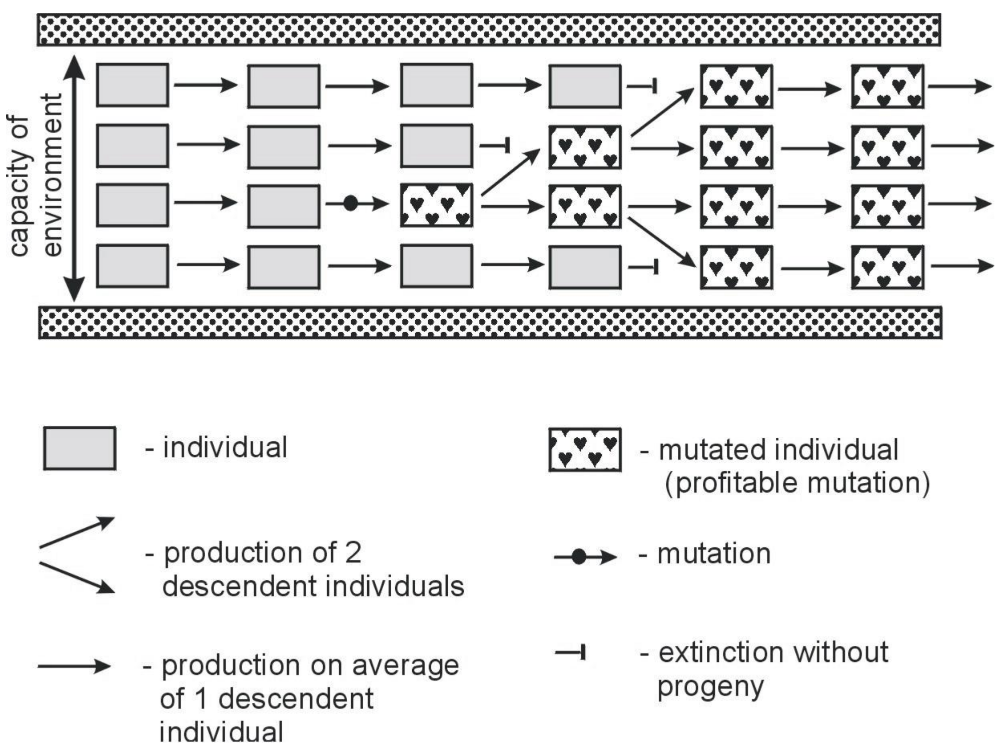

Fig.2.7. The effect of a profitable mutation. A profitable mutation (dot) will increase the chances of survival or/and reproduction of individuals possessing this mutation in comparison with individuals without a mutation. The progeny of mutated individuals will gradually

substitute non-mutated individuals and eventually all individuals in a population will become carriers of the profitable mutation.

Natural selection is the described mechanism of evolution. It expresses itself in the survival of individuals that are best adopted to the present conditions of an environment, due to the accumulation of profitable mutations. The competition for the resources of an environment can take place between individuals of the same species, as well as between individuals of different, frequently hardly related species (e.g. an ousting of herbivorous Australian marsupials by rabbits and sheep brought from Europe). Mutations which are profitable in one environment can be harmful in another environment. Therefore, it is the environment that constitutes the determinant factor of the adaptation of organisms. Dinosaurs were perfectly adopted to live in the tropics of the Mesozoic era, but unable – contrary to primitive mammals – to survive the fall of a meteorite (as well as the subsequent cooling off of climate and cutting down of sunlight by clouds of dust) on the turn of the cretaceous period and the Tertiary. On the other hand, some mutations, that improve the general efficiency of an organism, can increase its chances of survival in many different environments. Mammals can serve here as an example. They have taken into possession a huge number of environments on the Earth (lands, seas, air, tropics, polar zone, deserts and so on).

In conclusion, the general scheme of the natural selection is based on four principal premises:

- living individuals **propagate**, which produces much more descendent individuals than can survive;
- living individuals **inherit** properties (that is identity) of parental individuals;
- **accidental variability** prevails among individuals; it is caused by non-directional mutations (and recombinations);
- only the individuals survive that are **best adopted** to an existence and propagation in a given environment.

Natural selection constitutes in fact an inevitable consequence of the existence of living organisms (reproduction and inheriting of identity), as well as of the realities of the physical world in which they live (limited capacity of environment, inevitability of errors in copying of genetic information (mutations), properties of an environment which organisms have to adopt themselves to). This confirms the proposition that evolution constitutes an inseparable feature of the phenomenon of life ("composed" of propagating individuals), and therefore, that living organisms must evolve. Because mutations in DNA will never cease, life will always have a broad spectrum of offers concerning further ways of development to choose. Requirements of functionality and conditions of an environment will decide, by means of the natural selection, which of these accidental offers will be accepted (very few will). Next, the current structure of an organism will have the decisive influence on the set of offers proposed in the future (e.g. flowers of some plant can change their colour from yellow to red only because something like flowers came into being at some moment).

It should be stressed that natural selection does not realise any plan sketched in advance. Natural selection appears in systems possessing certain above-discussed properties and is as purposeless a phenomenon as the growth of crystals in over-saturated solution of salt, or the force of gravity responsible for falling of stones and revolution of planets around the Sun. It is undoubtedly an astonishing thing that such a simple mechanism can lead to such complexity and abundance of living forms. This is exactly what is decisive for the great importance and intellectual challenge of the Darwinian theory of evolution. However, in some fashion complex objects can already be produced by simple physical processes. It is enough to recall "flowers" "painted" by frost on window panes. Both the phenomenon of evolution and appearance of ice flowers do not have any purpose or sense – the phenomena simply take place. It should be repeated once more that natural selection (and evolution) constitutes an inevitable consequence of the very existence of living organisms.

Let us analyse a simple example. Let us imagine a small island inhabited by a population of a beetle. Most beetles have their crusts coloured in green, but there sometimes takes place a mutation resulting in the appearance of red spots on the green background. This mutation is not profitable due to energetic costs of production of the red pigment, and therefore the frequency of appearance of the allele responsible for spottiness is maintained in the genetic pool of the population at a low level. Let us assume now that insectivorous birds, feeding on our beetles, among other creatures, migrate from land to the island. If the red spots frighten birds away or cause a resemblance of the spotty beetle individuals to something inedible, then, the production of spots will become very advantageous (and gain a high positive selective value). The frequency of the allele conditioning this property in the genetic pool of the population will increase significantly, until finally this allele will totally oust (eliminate) the alternative allele, responsible for homogeneous green colouring. So, what has happened? The genetic composition of the population (as well as phenotypic properties of its individuals) has undergone a change, and therefore there **evolution took place**. The described sequence of events is presented in Fig. 2.8. While in the absence of birds, the mutation causing spottiness was disadvantageous, and therefore eliminated from the population, the natural selection started to favour it after an appearance of birds, due to which this mutation (the relevant allele) spread in the genomic pool.

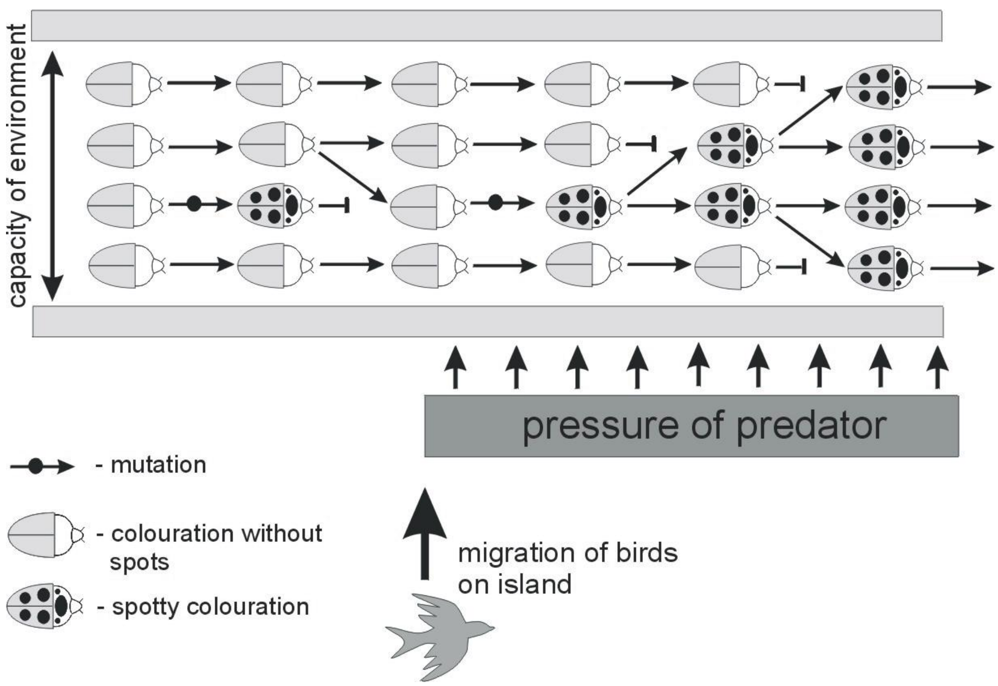

Fig.2.8. The evolution of beetles on an island. In the absence of predators, the mutation of spottiness is eliminated as a harmful mutation (unnecessary energetic expenses on production of a pigment). Once birds feeding on beetles migrate onto the island, the spotted coloration becomes profitable (frightening birds) and becomes fixed in the population.

The above example shows, that there is nothing mysterious or incomprehensible in the main mechanism responsible for biological evolution, namely in natural selection. Moreover, this example proves that the occurrence of biological evolution does not constitute a problem requiring any consideration. Evolution must occur. It is a consequence of the existence of living organisms.

The so-called industrial melanism[36] is a real example – similar to the case described above – where evolution "is caught in the very act".

* * *

Another aspect of life, equally essential for it as copying of an individual's identity (reproduction) and evolution, consists in the fact that life does not constitute a continuous mass of matter (like the rational ocean in the novel by Stanislaw Lem *Solaris*), but is divided into particular **individuals**, entering into its composition. This property seems to be absolutely universal in the terrestrial biosphere. Although in some particular cases (to be discussed in a moment) we could have serious problems as to **what** should be in fact regarded as an individual, I will try to demonstrate in the next chapter that the concept "individual" (in the evolutionary sense) can be defined quite objectively within the framework of cybernetic terminology. Moreover, I will attempt to prove that an individual is an absolutely indispensable component of the logical construction describing the phenomenon of life and evolution in cybernetic categories.

At present, I would like to discuss cursorily some aspects of the traditional, intuitive meaning of the concept "individual". The **structural aspect** is perhaps the most evident. As life is composed of individuals, this aspect is nothing else than the above discussed structure of living organisms. Each individual is not only in some way separated from its environment, but also differs from this environment by its internal structure. If not the second criterion, the properties of an individual would have to be attributed to a stone, which is after all an object separated spatially, yet it possesses an internal structure identical with the structure of the mother rock. The internal structure of an individual is connected, among others, with a possession of some barrier, separating it from the surroundings. The structural isolation of an individual from its environment is also accompanied by a different chemical composition and possession of hierarchical

[36]Its most known example is the following phenomenon: Before the period of industry development, a light form of a certain species of moth predominated in England. Its colour made it invisible for predators on trunks of trees covered with lichens, while dark forms, much better visible, were very rare. Later, air pollution caused by industry brought about extinction of lichens and the dark form propagated, while the occurence of the light form, in turn, became very sporadic.

substructures of a higher order. If one were to take, say, amoeba as a standard example of an individual, then it fulfils the above criteria in a satisfactory way. An amoeba constitutes a spatially separated (although changeable in shape) lump of cytoplasm, surrounded by the cell membrane. Its (bio)chemical composition differs drastically from the composition of its environment, no matter if it is water, soil or another kind of an environment. Even parasitic amoebas have slightly different set of proteins and other organic compounds than the organism of its host – this is decisive for the separateness of a given amoeba as a biological species. Finally, an amoeba possesses various structures of a higher order, such as macromolecules of proteins and nucleic acids, their complexes (ribosomes, chromosomes), cellular organelle (nucleus, mitochondria) and so on. In multi-cellular organisms, of course, the list is additionally supplemented by the levels of tissues, organs and systems of organs. The internal structure of an individual (living organism) has therefore, as I have mentioned before, a hierarchical character.

An individual is also separated from its environment **functionally**. This means that the number of relations (interactions) between the elements of an individual exceeds the number of interactions between an organism and its environment. If we separate arbitrarily an area of a rock, then the number of relations between its (also arbitrarily differentiated) "elements" is exactly the same as the amount of "interactions" between such elements occurring inside and outside the separated area of the rock. In the case of amoeba, inversely, the relative number of its internal relations exceeds significantly the number of external relations (between an amoeba and an environment). This is due to two reasons. The first one is a great functional differentiation (complication) of the interior of the amoeba (and therefore not only quantity, but also quality of functions differentiates an amoeba from its environment). The second one consists in cutting off most of the interactions (disturbances) with the environment by a structural and functional barrier (the cellular membrane is its important element). Only those interactions are admitted that have an important informative value, useful for survival and production of progeny, as biochemical or sensual signals do (it is difficult to distinguish them on the level of amoeba). In accordance with the already said, the functional "structure" of an organism has a hierarchical character, as its spatial structure does. It also exhibits some kind of purposefulness, expressing itself in the tendency to preserve the constancy of internal conditions, regardless the changes in external conditions (we call this homeostasis). This is realised by a complex system of regulatory mechanisms.

The discussed aspects (structural and functional) of an individual, as well as their sub-aspects, are strictly interconnected with each other and are mutually complementary. Therefore, they could be joined into one whole and labelled together as the structural-functional approach.

The classical paradigm of approaching a biological individual as a certain structural and functional entity, however, rises some problems. The first is that the classical definition refers also to systems which certainly are not living organisms. A robot and a computer fulfil, in principle, all the criteria and aspects discussed above. In the simplest dissipative structures, such as convective currents (e.g. Benard's cells), one can also perceive some ordering of structure and function, although it is present here in a very rudimentary form. It is not clear, however, if first organisms (individuals) were much more complex. Therefore, the structural-functional definition of an individual does not correspond well to the reference of the term "individual" in biological-evolutionary terms, and it does not constitute a sufficiently good criterion for distinguishing living individuals from inanimate systems.

Even if we limit the use of the structural-functional paradigm to the systems which we traditionally consider as living systems, it will still remain not very clear in relation to the problem what in fact should be regarded as an individual.

The first example is constituted by colonial coelencerates. In non-colonial species (e.g. hydra or *Actiniaria*), an individual constitutes a well-determined and separated whole, composed of a long, two-layer sack surrounding the absorptive-digestive cavity, with one opening to the outside – the mouth hole, surrounded by a set of tentacles. However, in many colonial coelencerates, individuals are only partially separated from the body of a colony. They have a common absorptive-digestive cavity, while their mouth wholes surrounded by sets of tentacles are separate, and so are the gonads. The structural-functional paradigm is not able to answer univocally if an individual in this case is constituted by a entire colony, or only by its partially separated and self-dependent "quasi-individuals".

Another doubtful case is found in myxomycete – primitive organisms exhibiting some properties intermediate between fungi and protozoa. For most of their life cycle, they exist as single amoebas. However, if nutritional substances are lacking, amoebas gather into a multi-cellular formation called plasmodium, serving for formation of a sporangium (which is composed of a stalk, sporangium walls and spores filling it) as well as for a sowing of spores. Some amoebas built the stalk and walls of sporangium, while others produce spores. It is obvious at the same time that only the latter amoebas – which transform themselves into spores – will have the possibility of transferring their genetic material to the next generations. The possibility is lacking for those that form the stalk and walls of sporangium. What is therefore an individual here? Amoebas that constitute structurally separated and self-dependent units during most of the life cycle? Or the plasmodium, that realises common reproductive functions for a short time? Or maybe amoebas cede in some moment their "individuality" to the plasmodium?

In social insects, both sexual forms as well as asexual casts at their service – workers, soldiers and so on – are generally treated as individuals. They fulfil well both structural and functional criteria. However, the entire colony constitutes also a certain entity in the functional and, to a less extent, structural sense, especially if it is considered together with its nest. One finds here a separation from the external world, hierarchical structure as well as a rich network of various interactions, abundant in different regulatory mechanisms. Therefore, the question arises whether a colony also should be considered as an individual. But then, what about particular insects? Can an individual be composed of individuals? This already seems to be an absurd. Following the path, kidneys, heart, liver as well as all particular cells entering into composition of a body would have to be regarded as individuals. On the other hand, why should not one treat the entire biosphere as an individual? After all, it is also an entity well separated both structurally and functionally. The classical definition of an individual seems to be completely defenceless here.

Viruses – composed of a thread of nucleic acid surrounded by a protein cover – and viroids – whose structure is limited to a naked thread of ribonucleic acid – fulfil the structural-functional criteria of an individual only partially. Undoubtedly, they are (at least viruses) structurally separated from the surroundings, while one can try to see hierarchical aspects in their structure (again – at least in viruses). On the other hand, they cannot, in principle, **function** outside the cell of their host. All their living functions, the whole network of biochemical transformations realised by them is totally dependent on the metabolism of the host. Therefore, viruses and viroids are not autonomous functionally. Moreover, viroids do not have any barrier, isolating them from their environment, while the protein cover of viruses can be regard as such. Therefore, again, the classical approach does not offer any conclusive criteria for awarding (or not) of the status of a (living) individual to the considered objects.

Most parasites are much more complex creatures than viroids or viruses. They also realise complex of functions, of which at least a part can be regarded as autonomic functions, i.e.

independent of the organism of their host. As a rule, there are no problems with counting them among living individuals (although, on the other hand, why should not one treat the whole system parasite-host as an individual?). However, there exists a group of objects possessing some properties of parasites, whose status is not so clear. Here belong tumours (cancers), prions and parasitic DNA.

Tumours, especially cancers, can be regarded, in a sense, as "would-be" parasites, the existence of which ends together with the death of the organism where they originate. The rate of growth and frequency of divisions of particular (types of) cells in multi-cellular organisms is, as a rule, subject to strict control. A cancer originates when a cell (or group of cells) breaks free, for some reasons, from this control and starts to propagate with the maximal velocity permitted by the accessibility of nutritional substances, oxygen and so on. This results in a growth of the mass of cancer cells in the form of a tumour, which pushes aside the neighbouring tissues and can lead to their damage. If a cancer is malignant, tumour cells have the ability to transfer via the blood system from the place of origination to other places in the organism, and to form next tumours there. This leads to quick death of the organism, which is of course equivalent to the extermination of cancer cells.

As a rule, a cancer (tumour) cell is not regarded as an individual. But, in fact, why? As far as the complication of structure and function is concerned, it is certainly more complex than all bacteria, the "individuality" of which could hardly be denied. The act of tumour-like transformation frees such a cell from the control of the organism of origin, and allows it, from the moment on, to live "its own life", as each individual does. The host organism can be treated simply as the environment of a cancer cell. The last differs from a "normal" parasite by the lack of ability to infect other host organisms. However, this is not a property which enters into the structural-functional definition of an individual[37]! Moreover, there is a type of a cancer appearing in dogs, called venereal sarcoma, which transfers itself from one dog organism to another by way of sexual intercourse. Therefore, the cells of at least this particular cancer constitute a parasite in the strict sense.

[37]It cannot be excluded that the mankind will also cause a total extermination of its environment, that is the terriestrial biosphere, an will become extinct if he does not find another environment (another planet). However, it would not enter anybody's head to deny the "individualness" of particular people.

Let us analyse another example. Prions are pathogenic factors of proteinic nature. They cause terminal diseases of brain in animals and man. Their effect on changes in behaviour is illustrated by the names of these diseases, such as "the disease of crazy cows" tormenting cattle, or kuru, called also "laughing death", widespread among cannibals on New Guinea. Many years of trials aimed at determining the character of the infectious factor led to the surprising conclusion that this factor is constituted by single molecules of a protein, normally produced by the organism of an animal (both ill and in good health) which have been called prions. It turned out that the molecules of the prion protein can occur in two conformations (spatial arrangements): "normal" and "infectious" one. Only the latter causes a disease. Moreover, it has been proven that the "infectious" form can stimulate the molecules of the normal form to transform into infectious molecules. Therefore, a single molecule of the infectious form which gets into an organism (brain) of an individual in good health can lead to its death, by transforming all its normal prion molecules into infectious molecules. Summing up, prions constitute the simplest infectious factor known to us, as it is simply a changed form of a protein which is normally synthesised by an organism.

Viewed in the structural-functional perspective, prions are further removed from being individuals than viroids. Their structure is extremely simple, and it is difficult to find in them any autonomous functions (realising their "own" purposes) – in spite of the fact, that they have constant contact with the metabolism of the "host" – unless we consider transformation of normal molecules into infectious molecules as a rudimentary form of a function. However, prions exhibit a property, lacking in tumours, that allows them to be classified as parasites, and therefore individuals. They can transfer from one organism to another, and therefore, they are able to cause infection. This should be viewed as another signal of difficulties faced in the attempts at a clear definition of an individual.

"Parasitic" DNA, such as some repetitive sequences (transposons) and chromosomes B, constitute an interesting example of parasites of the genome. Most probably, they come from fragments of genetic information which became independent in the past, and since then they live and evolve "on their own account", aimed to achieve their own evolutionary success (replication and spreading of their copies) rather than the benefit of their host. They resemble tumours in this respect. Parasitic DNA never leaves the genome of its host. It is transferred together with its generative cells. Its "customs" will be discussed in more detail in the context of a cybernetic concept of an individual. Here, it is enough to say that the classical approach has little to say in the case of parasitic DNA to answer the question: "to be or not to be" an individual. One could argue that the connection of parasitic DNA (due to its genesis and function) with the organisms on which it parasites is so close, and its structure (especially single threads of the parasitic DNA) is so simplified, that it cannot be awarded the status of an individual. However, this argument refers, at best, to our intuition and taste.

Finally – the case of symbiosis. Does the body of a lichen – composed of two co-existing components: alga and fungus – constitute one individual, or maybe both components should be regarded as individuals? In the phenomenon of symbiosis, coexisting organisms (to mutual advantage) are always in some way related structurally and functionally. This interrelation may range from very loose and fleeting connections to complete mutual life-and-death dependence, depending on a concrete case of symbiosis. It is not clear therefore if being an individual should also be considered as gradual, or the system of two co-operating partners should be considered as an individual.

The above examples show that the structural-functional definition of an individual is not able to indicate clear and unambiguous attributes of being an individual. What is worse, it cannot join together, clearly and directly, the phenomenon of "individuality" with the phenomena of life and evolution, although these three things seem to be inseparable. This is why the classical paradigm does not seem to be especially heuristically fruitful. Its usefulness in explaining such problems as the specificity of life or subject of evolution is small. The classical formulation does not throw a new light on old problems, neither does it allow for posing new ones, as it acts principally *ex post*, which means further specification and codification of the already existing intuitive understanding of an individual. All this does not mean that the classical formulation is completely useless. Quite the contrary. It can be very useful, but only as a supplement of some deeper and more fundamental formulation. In the next chapter, I will propose a possible formulation of this type.

Above, I enumerated many selected aspects, associated directly with the phenomenon of life. Their list could certainly be continued further on. Afterwards, I performed a selection of only these aspects, that constitute an unfailing criterion allowing one to distinguish between living organisms and inanimate objects. As a result, we obtained three mutually co-defining concepts: an individual, reproduction (connected with inheriting of identity) and evolution. Living organisms are systems producing descendent systems and transferring to them the complexes of their properties, their identity, which undergoes evolution in time, from generation to generation.

In my opinion, the essence of life consists in an **identity** of living individuals **directed on itself**. This identity includes the entirety of structure and regulatory mechanisms aimed at survival and reproduction of the individuals being its carriers, and therefore at continuing the existence of this identity in time. The identity of living individuals exists only to maintain existence, protect, copy and propagate nothing else, but just this identity. Additionally, identity evolves in time, and therefore the aim is not so much to preserve the identity strictly, but rather to preserve the continuity of changes of this identity in time. This permanent continuation, preserving in later stages the memory of earlier stages (in the form of current structure and function), makes those identities – passing smoothly from one into another – into a certain whole. If the last statement seems difficult to comprehend, please recall man who, after all, is someone completely different in his old age, compared to himself in his childhood. His body and especially psyche undergoes evolution during the life time. In spite of this, he is still – in a very important sense – the same man. It could be concluded that the memory about earlier stages present in later stages is what links the subsequent "stages" of a given man (apart from the "memory of structure", the principal role is of course played in man by the psychological memory).

In conclusion, life can be defined from different (complementary) points of view, using such terms as individual, reproduction, evolution or finally – identity. However, there arises a basic question: isn't it so, that the mentioned aspects constitute some "excessive" and macroscopic manner of describing the phenomenon of life, the essence of which lies at some deeper level? Or, in other words: is it possible to separate, single out some simpler core, the essence of the phenomenon of life, in relation to which the mentioned aspects are either secondary manifestations, or macroscopic categories, being a derivative both of the phenomena of the real world and of our minds? I will deal with this problem in more detail in the next chapter, devoted to the cybernetic paradigm of description of life and evolution.

CYBERNETIC FORMULATION OF BIOLOGICAL EVOLUTION

None of the above-described paradigms used for characterising the phenomenon of life reaches, in my opinion, the very heart of the essence of the phenomenon of life. Neither does it the classical (biological) paradigm, based on the criterion of structure, function, reproduction and evolution, nor is it achieved by the thermodynamic paradigm, treating life as a highly complicated dissipative structure. The thermodynamic paradigm is too general, and therefore description of dissipative structures in terms of quantity of information is not able to characterise the specificity, value and the kind of information related to life. The biological paradigm, inversely, is to excessive. It indicates many properties attributed to life, but does not state clearly what life really is in general. Both paradigms have been worked out for one particular case of life, namely for the present life on the Earth which, generally speaking, differs very clearly from inanimate nature. However, there are objects on our planet that generate serious problems concerning their classification as living individuals, as it has already been and will be discussed. Much more serious problems of this kind may arise during exploration of other planets. The recent discovery of supposed traces of life on Mars leaves room for possible divergence of interpretations, and we deal here not with life itself, but only with its possible remains in the form of organic compounds and some fossil structures resembling bacteria in shape. In this case, we can therefore take cover under the aegis of ignorance and use it to explain our hesitation as to the status of the discovered traces. However, let us imagine that we get hold of some strange, completely alien and incomprehensible creature or object, and we are to classify it as a living being or an inanimate object, and suddenly it turns out that we lack conclusive categories or reference points which we could borrow from the terrestrial life, and which would enable us to make an univocal decision (this problem was excellently illustrated by Stanislaw Lem in such novels as *Solaris* or *The Invincible*). A criterion differentiating life from inanimate nature would also be very useful in considerations focused on the origins of life on our planet. It would have crucial significance for answering the questions: at what moment did life emerge from the suspension of 'dead' organic compounds, as well as: what was it for life to come into being?

In the present chapter, I would like to propose a cybernetic paradigm for characterising the essence of life, which, in my opinion, enhances important progress towards the solution of the above-presented problem, although it might not indicate an ideal and absolute criterion for the "animating" a system, because such a criterion is probably impossible to be found. Firstly, although life certainly cannot be defined exhaustively as e.g. a dissipative structure connected with production of entropy, nevertheless life undoubtedly **must also** be a dissipative structure, regardless of other properties. Thermodynamics tells us therefore that life cannot be found in systems remaining in the state near thermodynamic equilibrium (e.g. in cold interstellar dust or in a Cosmos with a negative curvature of space, in a couple of hundreds billions years) and we may not neglect this fact under any circumstances while aiming at a full characterisation of the phenomenon of life. Therefore, apart from the 'core' of the essence of life, the following conditions necessary for the existence of this core are also important: displacement of a system from the state of thermodynamic equilibrium, presence of a proper type of chemical compounds, organic compounds of carbon in particular (for example, it is not possible to imagine life in the form of gas – except in science-fiction novels – because gaseous substances lack sufficient wealth of structure of molecules) and so on.

Secondly, the phenomenon of life is a certain macroscopic category, which is "located on the side of our mind", but does not necessarily exist 'objectively' (whatever this could mean) in the real world. In the face of this, an attempt to discover an unequivocal and absolute criterion differentiating life from inanimate matter may also be, to some extent, nothing more than a tendency of the human mind to classify the continuous spectrum of phenomena, i.e. the world, into discrete categories. Being aware of these threats and limitations, I will try however to formulate cybernetic characteristics of the minimal semantic core corresponding to the phenomenon of life, which, in my opinion, reflects best what we would be inclined to identify with life.

Cybernetics is a science on regulation, first of all on regulatory mechanisms in complex systems realising some purpose ('purposeful’ systems, such as living organisms and many devices built by man), whatever we would understand by such purpose. The fundamental concepts of cybernetics are: **negative feedback** and **positive feedback**. The negative feedback constitutes the basic regulatory mechanism. Its task is to maintain a value of a certain parameter (this could be temperature, rate of reaction, concentration of compound or anything else) at an assigned (fixed) level – any disturbance causing a deviation of the parameter value from the assigned one in any direction induces actions leading to a change of the parameter value in the **opposite** direction (hence the name 'negative’), and therefore to compensation of the effect of this deviation. The positive feedback, on the contrary, tends to a change of a parameter value in the direction

consistent (hence – 'positive') with the direction in which the deviation from the assigned (fixed) value (which is frequently simply zero) took place. Therefore, this feedback causes a still greater and greater increase (or fall) of the value of a given parameter. While in the negative feedback, the parameter value tends to the assigned value, oscillating around it; then in the positive feedback, the parameter value escapes from the assigned value the faster, the further from this value the parameter already is. Fig. 2.9 presents schematically a system without any feedback, compared to a system containing a negative feedback and a system containing a positive feedback. Below, I will discuss in more detail both kinds of feedback and indicate various examples.

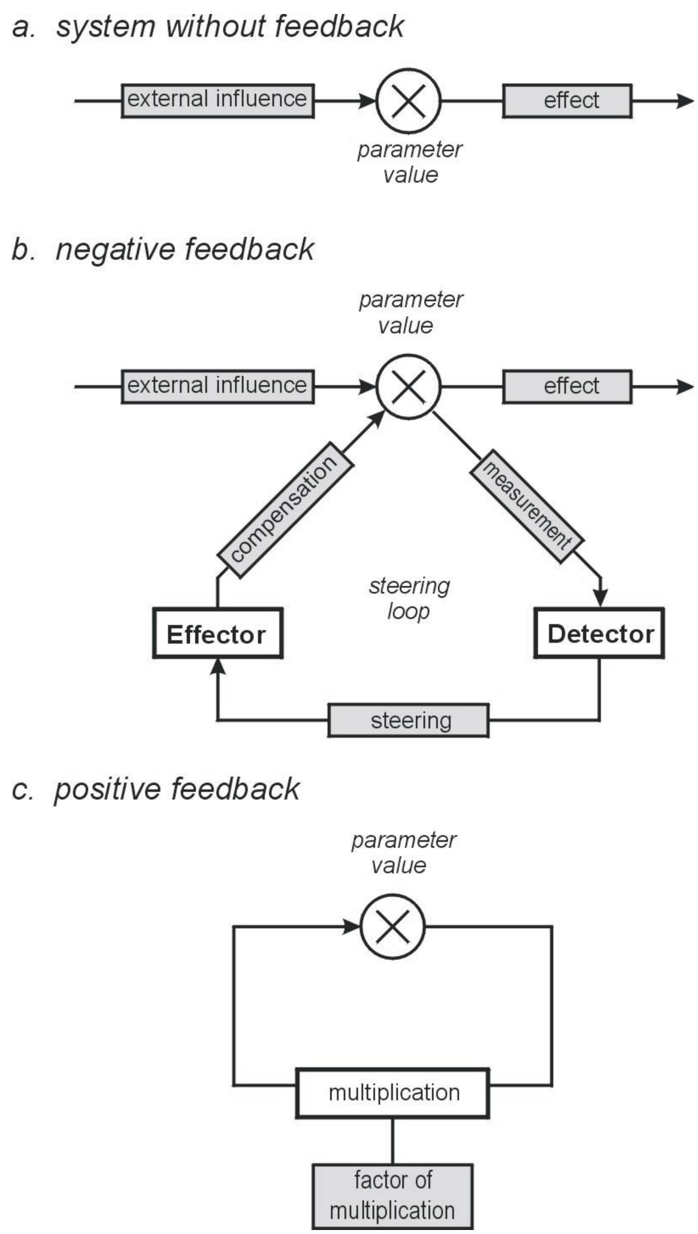

Fig. 2.9. Negative feedback and positive feedback. (a) system without feedback, reacting passively to influences of an environment; (b) system with a negative feedback, regulating (maintaining on a constant level) a value of some parameter and compensating influences of an environment; (c) system with a positive feedback, increasing in a given period of time a value of some parameter by a certain constant factor (multiplication of the current value) – the greater the value of a parameter already is, the faster this value increases.

Negative feedback constitutes the central concept of the entire cybernetics. In the colloquial language, it can be simply referred to as "regulatory mechanism". However, as cybernetics defines what we understand as regulation in a much more precise and formalised manner than the colloquial language does, we will keep the term "negative feedback".

Fig. 2.9a presents a system without any regulation. The **value** of a certain considered **parameter** (let it be, say, temperature) is here the result of a passive response of the system in question to an **influence of environment**. If we put a box to into a warm room, the temperature inside it will quickly reach the temperature of the room. If we construct the walls of the box to insulate thermally possibly well, then the process of equilibration of temperatures will be slowed down, but the final effect will be the same. If there is food (for example meat) in the box, then the **effect** of the room temperature prevailing in the box will manifest itself in spoiling this food quickly. This effect will be disadvantageous in relation to the **purpose**, determined by man, namely, to preserve the freshness of the stored food for as long as possible. Therefore, in a system without regulation, the value of a given parameter reacts passively to the interactions with the environment, and, as a result of a "harmful" influence of the surroundings, may exert a disadvantageous effect (considering a certain predetermined purpose).

The **regulation** of a given parameter value consists in maintaining this value close to the **assigned optimal value**, which produces the best effect with respect to some purpose assumed in advance. In our case, such an assigned value is the temperature sufficiently low to prevent the spoiling of food for a long time and, at the same time, high enough to avoid a too high expense of energy for its maintenance. The negative feedback is the mechanism allowing to maintain a parameter (temperature) at the assigned (optimal) level. As it is the negative feedback that constitutes the basis of a thermostat operation in the refrigerator, which, as we know, is a device used to store food due to its ability to maintain a constant (with slight oscillations) low temperature. Let us analyse the action of a refrigerator in relation to Fig. 2.9b, presenting a system with negative feedback.

A refrigerator standing in a warm room, in spite of thermal insulation, has a natural tendency to equalise the temperature of its interior with the temperature of the room (the influence of the surroundings). The result of this would be a temperature increase in the refrigerator interior (and therefore an increase of the parameter above the assigned value) and, as a consequence, spoiling of food in it. However, the parameter value (temperature) in the refrigerator is continuously **measured** by an appropriate **detector**, in this case a device called thermostat. When the detector finds out there is a parameter value increase above the assigned value, it sends a

steering signal, which sets in motion an appropriate **effector** (here: a cooling device), which counteracts the deviation of the parameter value from the assigned value caused by external influences, and therefore **compensates** the disturbances brought about by these influences. In the discussed case, it means such reduction of temperature that it returns to the assigned optimal value. When this happens, the cooling device is turned off.

Generally, the greater the deviation of a parameter value from the assigned level, the stronger the compensation induced by the negative feedback to **diminish** this deviation (hence the name "negative"). The whole discussed cause-result sequence – namely: increase of a parameter value above the assigned level–measurement–detector–steering–effector–compensation–decrease of the parameter value back to the assigned level – is called the **steering loop**. It results in maintaining, in spite of external influences, such a value of a given parameter which brings about a profitable, desired effect.

One could ask, however, if the cooling device could not chill a refrigerator at a constant rate, compensating accurately the rate of its warming up by an exchange of heat with the surroundings? Unfortunately, it could not. We created "comfortable" conditions for the refrigerator in our example, assuming a constant (although high) temperature in the room. In most situations, external conditions are never ideally unchangeable – they undergo continuous, unceasing fluctuations. Let us imagine fluctuations of the temperature of the surroundings of our refrigerator with the window in the room kept open during the entire year. Then, the external temperature would change from time to time, and therefore the rate of a warming of the interior of the refrigerator – linearly proportional to the difference between the outside and inside temperatures – would also change. In this case, the steering loop appears to be absolutely necessary.

The assigned parameter value does not need to be constant. It can change depending on conditions. For example, the optimal temperature of the body of a bat is different in summer and in winter, during hibernation. It is important for the optimal parameter value to serve the superior purpose of the entire system, in this case the survival and reproduction of a bat.

Let us pass now to the positive feedback. It exhibits, as the name suggests, an action with contrary direction to the one of the negative feedback. While the latter causes an "attraction" of the value of a given parameter to a certain (finite) assigned value, the positive feedback "pushes away" the value of a given parameter from the assigned value the more strongly, the further away from the assigned value the parameter moves. The parameter value tends therefore quickly (exponentially) to infinity, and the obtained effect can be called "anti-regulatory" or "explosive".

Positive feedback takes place in the case of the previously mentioned nuclear chain reaction, as well as in that of a propagation of a colony of bacteria. Each free neutron or bacterium are able to produce new neutrons (bacteria). The more neutrons (bacteria) there are at a given moment of time, the faster the production of new free neutrons (bacteria) proceeds. This is due to the fact that each neutron absorbed by a nucleus of a radioactive isotope of uranium (U^{235}) leads to the emission of three new neutrons, while each bacterium cell divides into two descendent cells. Therefore, the amount of free neutrons (bacteria) undergoes tripling (doubling) in the period of time between the emission of a neutron and splitting of a subsequent nucleus of uranium (or the period between divisions of bacteria cells). Generally speaking, in the positive feedback a non-zero value of some parameter causes an increase of itself, which proceeds the faster, the greater this value is at a given moment. Another example of the kind of feedback in question is found in the already discussed autocatalytic reaction, in which a molecule of a chemical compound catalyses the synthesis of molecules identical with itself. It can also be found in an avalanche, where a jostled stone on a hill slope sets next stones in motion, so that a great mass of stones soon goes down the hill. Frequently, processes due to positive feedback are referred to by the name "chain reaction".

Let us, however, come back to the negative feedback. It is the basic regulatory mechanism in such goal-directed systems (i.e. realising some purpose) as living organisms or devices built by man[38]. In the case of man-made products, negative feedback can manifest itself in a correction of

[38]This mechanism functions also in a sense in various naturall inanimate systems. However, a problem arises here, concerning the discrimination between the negative feedback and the state of dynamic equilibrium which is accidentally disturbed by turbulent processes. Let us take any star as an example. The radius of such a star is, as I mentioned before, the resultant of the operation of two forces: the force of gravity tending to contract the whole matter of a star into a small area, as well as the pressure of radiation 'distending' the matter outwards. Because heat in the interior of a star is tranferred by convective currents, its surface resembles to some extent the surface of boiling water. Some fragments (bubbles, protuberances) move away temporarily from the surface of a star, exceeding thus the 'assigned' distance from its center, determined by the radius of a star. There, however, the pressure of radiation is not sufficient to counterbalance the gravitational force and the matter that has been thrown out comes back to the star surface. We do not treat this phenomenon as a negative feedback because human brain does not distinguish the decrease of the strenght of radiation with distance (faster than the decrease of the strenght of gravitation) as a 'special', purposefull regulatory signal. However, in principle, the described case does not differ in any important manner from the 'typical' negative feedback -- these differences are derivative of the

deviations from the assigned course of a flight of a rocket, in maintaining a constant pressure in the steam boiler, in optimisation of tasks performed by some types of robots, or in the above-discussed function of a thermostat. At present, however, we will be naturally interested in the regulatory mechanisms based on negative feedback that occur in living organisms.

Typical examples of the negative feedback can easily be found at the biochemical level. This feedback occurs, for instance, in the case of metabolic pathways of synthesis of many aminoacids. Such a pathway is illustrated in Fig. 2.10. A given aminoacid constitutes the final product P of this pathway. The whole pathway consists in a transformation of subsequent metabolites in next metabolites by subsequent enzymatic reactions: A is transformed into B, B into C, afterwards into D, and finally it comes to the synthesis of P. While metabolite A is a precursor not only of aminoacid P, but also of many other compounds, metabolites B, C and D occur exclusively in the pathway of synthesis of P. The task of negative feedback here consists in maintaining a relatively constant concentration of P, so that, on one hand, it does not undergo an excessive accumulation in a cell, and, on the other hand, it is present in a sufficiently high concentration to fulfil the demand of the reactions which utilise it. The factor disturbing the concentration of P is the changeable intensity of different reactions for which this aminoacid is a substrate, e.g. protein synthesis. If a sudden increase (decrease) of the rate of the synthesis of proteins (that is usage of P) were not followed by the "tuning" of the rate of production of P to the rate of its consumption, the concentration of P would fall (rise) drastically in a short time. This would be disadvantageous for the reasons mentioned above. Therefore, the concentration of P acts as a negative-feedback signal, inhibiting the first reaction – that is the transformation of A into B – which leads (exclusively) to its own production. In the conditions where the aminoacid P accumulates in a sufficient amount, the rate of this reaction is inhibited (which also prevents excessive accumulation of B, C and D). On the other hand, if the concentration of P drops due to intensive consumption, the transformation of A into B becomes unblocked and the rate of metabolite flow through the pathway increases, which counteracts the decrease in the concentration of P. In the quoted above example, the allosteric center of the enzyme catalysing the A $\rightarrow$ B reaction, the center that binds P, is the detector performing the measurement of the

---

'valuation' imposed by our cognitive apparatus. It is also important that the maintaince of a constant radius of a star can hardly be identified with any sensible purpose – from the human point of view. I will say more about the 'fact-creating' role of our mind in the chapter devoted to the evolution of the conceptual network.

concentration of P. The steering process proceeds by a conformational change (of the spatial arrangement) in the molecule of the enzyme resulting from binding P to the allosteric center, which causes a decrease of the activity of the catalytic center; this, in turn, fulfils the role of an effector compensating – via a change in the rate of the synthesis of P – fluctuations in the concentration of this aminoacid.

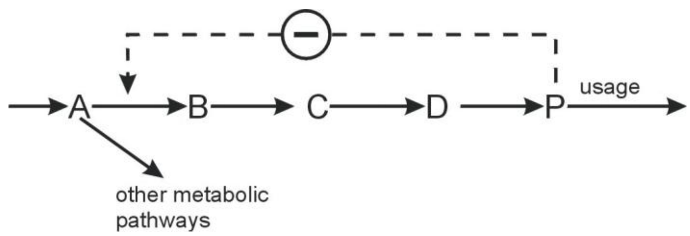

Fig. 2.10. Negative feedback in a metabolic pathway leading to synthesis of aminoacid P. Higher concentrations of P inhibit the first reaction, leading directly to synthesis of this aminoacid, thus preventing excessive accumulation of P, as well as of intermediate metabolites (B, C and D).

The regulation of the rate of energy production in response to energy demand varying in time (its utilisation in various processes inside a cell, such as protein synthesis, transport of ions through the cellular membrane, or a mutual translocation of actin and myosin filaments during muscle contraction) is another example of a negative feedback on the biochemical level. The chemical compound called ATP is the universal "carrier" of energy in the cell. It is synthesised (in animals) mainly in cellular organelles called mitochondria in the process of oxidative phosphorylation. ATP is produced from another compound – ADP. When those different kinds of work (protein synthesis, muscle contraction, movement of a vibraculum) are performed, ATP "gives back" its energy and is transformed into ADP, which is transferred to mitochondria, where it again serves as a substrate for synthesis of ATP. The rate of ATP utilisation by a cell varies. It is relatively small during rest, while during muscle contraction, for example, it becomes large. It is necessary to adjust the rate of ATP production to the rate of its usage. Mitochondria have to "learn" in some way how fast they are to supply this compound (otherwise, the entire ATP would be exhausted quickly, or it would accumulate in amounts that are lethal for a cell). ADP concentration plays the role of the signal informing these organelles about the current energetic demand. During more intense work, the process of ATP conversion into ADP is quicker, and therefore the concentration of the latter compound grows. This activates mitochondria (strictly speaking, appropriate mitochondrial enzymes participating in energy production) to more intensive synthesis of ATP, which causes an increase of its concentration, while the concentration of ADP diminishes. This mechanism results in at least partial compensation of changes in the concentration of ATP and ADP during periods of increased demand for energy as well as an ensuring a constant inflow of ATP at an appropriate rate[39].

Both above presented examples of regulation at the biochemical level concerned modification of activities of enzymes already present in a cell. However, the appropriate enzymatic proteins are frequently synthesised only when they are needed. This prevents unnecessary expense in the form of energy and aminoacids used for production of enzymes that have nothing to do at a given moment of time. Such a mechanism operates, for instance, in the case of enzymes catalysing the pathway of decomposition of lactose in bacteria. The synthesis of these enzymes occurs only if lactose is present in the surroundings. This sugar unblocks the expression of genes encoding the indicated enzymes. Thus, the regulation takes place at the genetic level (the details of the mechanism of this regulation is described by the famous theory of operon). The regulation here is aimed at optimisation of two different parameters, depending on situation. In the presence of lactose, the negative feedback predominates. It secures a sufficient amount of building material as well as respiratory substrate for energy production (the function of both is fulfilled, among others, by lactose). If this sugar is lacking, the above purpose must of course be realised in another way (by assimilation of other compounds from the surroundings). Then, there comes into prominence the necessity of inhibiting unnecessary utilisation of energy and building material in the form of aminoacids for the synthesis of enzymatic proteins. All this takes place within the framework delineated by a superior purpose, consisting in ensuring as great inflow of energy and building material for a bacterium cell as possible, which allows it to grow and divide fast (the purpose may also be defined as maintaining a constant concentration of nutritional substances in a cell during the maximal rate of growth and divisions). This superior purpose governs two different "assigned" values of concentrations of enzymes belonging to the pathway of decomposition of lactose: zero concentration when the sugar is absent, or a

[39]The described mechanism is in fact more complex – e.g. calcium ions may at certain moments stimulate directly the production and consumption of ATP.

concentration ensuring the optimal rate of degradation of lactose if it is present. The assigned value of a parameter is therefore conditional here – it depends on circumstances.

As we are discussing regulatory mechanisms in bacteria, it is worth mentioning substances called **alarmons**. Alarmons appear the moment a certain class of organic compounds are lacking in a cell. For example, ppGpp is the signal of aminoacid hunger, while cAMP occurs in the case of carbon hunger (lack of sugars). Alarmons stimulate the synthesis of various enzymes, participating in the production of this type of organic compounds that are lacking at the moment. Their significance as steering signals in negative feedbacks aimed to maintain a sufficiently high concentration of different compounds is obvious. Production of cilia – serving bacteria cells for locomotion – is another regulatory mechanism acting in the case of a lack of nutritional substances. Cilia allow active seeking of organic compounds in the surroundings, which is why they are also an element of the steering loop aimed to maintain a constant concentration of building and energetic substances in a cell.

The frequency of divisions in bacteria constitutes a simple derivative of the growth rate (doubling of mass) of bacterial cells. However, the moment a bacterium undergoes division – i.e. a cell exceeds the critical size, which constitutes the signal for division – is determined most probably by the proportion of the amount of cytoplasm to the amount of DNA in a bacterial chromosome. In the process of division, the amount of DNA undergoes duplication, while the amount of cytoplasm does not change. Therefore the ratio of cytoplasm to DNA is reduced by half, to increase gradually afterwards during the growth of descendent cells, until the next division. In this negative feedback, the cytoplasm/DNA ratio is maintained within some determined boundaries (its value changes only by half). This is important because DNA would not be able to synthesise the great amount of proteins needed by excessive cytoplasm, while insufficient amount of cytoplasm would not be able to produce enough energy and building material needed for duplication (doubling the quantity) of DNA during a cell division. Let us remember that the whole cytoplasm is in a sense only a carrier, *vehiculum*, for genetic information recorded in DNA, and it serves propagation of this information (on the other hand, however, this information is nothing else than a description of the structure and function of the entire cell, that is cytoplasm, first of all; I will come back to this vicious circle, so characteristic for the essence of life).

In bacteria, an entire network of negative feedbacks – constituting the functional essence of those organisms (as systems possessing purposeful regulatory mechanisms) – is realised first of all at the biochemical and genetic level. The scheme of this network is recorded in a DNA thread, while the interpretation of this record is carried out by a whole complex of enzymes (with the aminoacyl-tRNA synthase in the forefront), translating (the key is here the genetic code) the instructions expressed in the sequence of nucleotides in DNA into the sequence of aminoacids in proteins, which determine the structure and function of those proteins.

How does this record "know" which proteins and in what temporal sequence should be produced to fulfil optimally the function of survival and multiplication of this record itself? The genetic information is "blind" – it does not "see" in any way the physical world or the rules governing it. The plan of the network of negative feedbacks recorded in genetic information originated in the process of evolution by trial and error. Only those loops of negative feedbacks, among all the accidentally originated ones, were consolidated into DNA (as a record of the structure of proteins as well as of the pattern of expression of genes determining the time and intensity of synthesis of various proteins) that appeared advantageous for survival, possibly fast growth and reproduction. The remaining loops underwent elimination. This situation resembles to some extent a blind man, who does not see a room, but has learnt to move in it gropingly. As the "picture" of the room finds some reflection in the set of rules which are used by the blind man during translocation from one place to another, so the system of negative feedbacks recorded in the genome of each organism constitutes a "reflection" of the external world in which this organism lives. However, this reflection concerns mainly the world aspects of important significance for survival. The top achievement of evolution – the human brain connected to the entire complex of sense organs and reconstructing the picture of the world – is also subject to these limitations. The implications of this fact for the nature and universality of our cognition in science and philosophy will be discussed in the third part of this book.

After this digression, however, we should return to the network of negative feedbacks and take a look at a couple of examples found in higher organisms. In multicellular organisms, there appear higher levels of structural and functional hierarchy: the level of tissues, level of organs, level of systems of organs and finally – level of the whole multicellular organism. Each of these levels possesses its own system of negative feedbacks, and they are also organised hierarchically – negative feedbacks at a lower level "work" to realise feedbacks at a higher level. Let us consider concentration of sugar (glucose) in blood. It is controlled by a complex of two antagonistic hormones: insulin diminishes this concentration, while glucagon increases it. This is carried out in the process of depositing glucose in various tissues in the form of glycogen or (after transformation) fats, or in the reverse process of releasing glucose from tissues into blood. The more or less constant level of sugar in blood is important because an excessive lowering of its concentration would lead to damage (or simply necrosis) of the tissues which, like brain, use glucose as the only source of energy, while an excessive amount of the sugar leads to a diabetic coma. The glands producing the two hormones are able to detect the level of sugar in blood and to react adequately, by increasing or decreasing hormone production intensity. Participation in the process of two steering signals (hormones) acting antagonistically ensures the that regulation is more precise and fitted to conditions than it would be in the case of only one signal present.

The level of sugar in blood undergoes various fluctuations, because different tissues utilise it with an intensity which is changeable in time. Apart from this, there is much more glucose after a meal than after a period of fasting. The balance between the processes of releasing glucose (and fats as well) from and depositing it in storage tissues must be adjusted to the current needs. This is not enough, however. It is clear that the reserves of nutritional substances in an animal body are not unlimited and will start to get exhausted after a longer period of starvation. Therefore, there arises a need of their completing, that is simply of a taking up food from outside. Therefore, a lowered level of glucose in blood stimulates the center of hunger located in the hypothalamus. This changes the state of activity of a whole animal which – instead of resting blissfully, as after an abundant meal – increases its behavioural activity, undertaking intensive functions aimed at finding food. Once hunger is successfully satisfied, the activity of an animal abates again. This subsequent negative feedback acts at the level of an entire animal.

Let us return however to insulin and glucagon. Their production and release are dependent on the realisation of other negative feedbacks, operating at lower levels – the genetic and biochemical ones. First of all, both hormones are proteins and therefore their heightened synthesis requires genetic regulation – activation of appropriate genes. Secondly, the synthesis of proteins is connected with energy utilisation, that is with decomposition of ATP to ADP. Therefore, it becomes necessary to increase the rate of production of ATP by mitochondria. The negative feedback described a few paragraphs earlier is responsible for this. We see therefore that negative feedbacks operating at different levels of hierarchy are closely interrelated. The mutual relations in the network of negative feedbacks in living organisms are extraordinarily abundant and complex. Above, I discussed only a couple of simple examples. One should be aware, however, that each of the feedbacks depends directly on tens and hundreds of other feedbacks. For instance, active search for food requires a precise co-ordination of movements, which, in turn, requires appropriate regulation of tension of particular muscles, changing in time very quickly. It is also necessary for an animal (e.g. carnivorous animal) to correct "on the run" the current trajectory of its movement on the basis of signals coming from sense organs (receptors) (e.g. a view of a prey). In many cases, it is even difficult to define in a simple way **what** in fact constitutes the parameter, the value of which is to be optimised. Such a parameter constitutes frequently a quality much more complex than temperature, for instance, and its assigned value can change in time as an answer to a change of need determined by values of other parameters. All this, however, only insignificantly weakens the heuristic power lurking in the cybernetic paradigm. In the present chapter, I will try to show that the application of this paradigm allows a much better understanding of the essence of living organisms.

In conclusion, the essence of the functioning of living organisms consists in direct or indirect interconnection of all negative feedbacks (mechanisms optimising a parameter value). As each is important for survival of the entire organism, therefore, it creates an indispensable "environment" for other feedbacks. For each two negative feedbacks, it is possible to detect a vicious circle in which one of the feedbacks conditions the existence of the other and vice versa. Let us consider the example of the two already discussed feedbacks: 1. decrease of the level of glucose in blood → stimulation of the center of hunger in the brain → searching for and taking up food → restoration of the previous level of sugar in blood; 2. increase of utilisation of ATP and decrease in its concentration → increase of ADP concentration → stimulation of ATP production by mitochondria → increase of ATP concentration.

The first negative feedback ensures a sufficient amount of glucose, which, among other functions, is a substrate for energy production (ATP). The second negative feedback ensures production of a sufficient amount of ATP to enable the transport of sodium ions (Na^{+}) and potassium ions (K^{+}) across the cellular membrane of appendages (axons and dendrites) of neural cells, and therefore conduction of neural impulses required for active search for food. Not only both negative feedbacks cannot function one without the other (and without a great number of other feedbacks). They also mutually impart **sense** one to another (again – in connection with all other negative feedbacks present in an organism). This statement is true in relation to any considered pair of feedbacks. This mutual co-conditioning of negative feedbacks, making their "meaning" (purposefulness) relative to other feedbacks, constitutes a deep-level property of living organisms. We will also meet it later (the third part of this book) in the conceptual network, constituting (in my opinion) the "substance" of our psyche and (self)consciousness. This will be one of the contributions to the claim that the psychical level emerged from the biological level according to a similar principle that governed emergence of the biological level from the physical level.

Taking things generally, I propose that the most general scheme of the functioning of living organisms can be formally described **exclusively** in the form of the network of negative feedbacks (or, more widely, regulatory mechanisms), without reference to the manner of their material realisation (building material of particular elements, physical basis of different processes or detailed constructive and functional solutions). Additionally, this system of negative feedbacks, constituting the basis of the (purposeful) function of living organisms, is organised in a hierarchical manner. As a rule, the regulation at a lower level of hierarchy serves as an element of a higher level (although frequently the situation is reverse, as in the discussed vicious circle of the regulation of the level of glucose and ATP). The system of negative feedbacks constitutes one, internally interconnected and co-conditioning entity. However, a question arises what in fact constitutes the **superior purpose** of **all** these regulatory mechanisms. What does the **whole** system of negative feedbacks possessed by living organisms serve? As I stated before (this topic will be discussed further on), the goal is survival and production of progeny, that is maintenance and propagation in time of one's own **identity**, which corresponds in the cybernetic formulation to just **such-and-not-another system of negative feedbacks**. The parameter whose value undergoes optimisation in the operation of the whole network of negative feedbacks constituting a living individual is therefore the identity (complex of properties) of an individual, its survival in time and propagation. The system of negative feedbacks creating a given organism is **auto-directed** on itself – its only purpose is its own copying (multiplication). However, as I showed in the example of bacteria (this applies to any other living organism), reproduction of organisms constitutes a leading example of a positive feedback, in which a non-zero value of a parameter tends to increase, and does so the faster, the greater this value is at a given moment.

Finally, the appropriate moment has come to define (this is my own proposition) life (a living individual) in cybernetic terms as a **system of negative feedbacks subordinate to (being at service of) a superior positive feedback.** A supplementary criterion is the condition that **all** sub-units and negative feedbacks entering into the composition of a system in question are aimed **exclusively** at the realisation of the principal positive feedback[40]. I claim that all living system are necessarily subject to the above definition, and all systems subject to this definition should be recognised as alive (or, if somebody prefers, belonging to the class of life-like systems). In one word, the proposed definition, corresponding to the very essence of life, constitutes the **smallest conceptual system, necessary and at the same time sufficient for an unequivocal determination what life is**. It can therefore be used to specify univocally which objects and processes on other planets can be regarded as alive, from which moment we may speak about life in the process of its spontaneous genesis on the Earth about 4 billions years ago, and what constitutes a living individual among the discussed above, problematic phenomena known to us presently on our planet.

Various aims realised by each living organism can be divided into **subordinate purposes**, expressed by the network of regulatory negative feedbacks, as well as the **superior purpose**, manifesting itself in the existence of a positive feedback.

We saw, in the example of bacteria, where this superior positive feedback comes from. Living systems have a natural tendency for reproduction and multiplication of their copies, when attainable environmental resources allow this. The sequences describing the amount of individuals in subsequent time intervals (if environment capacity is not limited) grow exponentially and tend to infinity. It has been calculated that, but for a limited amount of food and environment poisoning with metabolism products, the progeny of a single bacterium would in a relatively short time become a huge sphere expanding with the speed of light. The same applies to a pair of sloths

[40]This last condition excludes the classification as alive of such systems as cultures, states and economical companies, equipped with a network of negative feedbacks and tending to expand at the cost of other, similar systems. However, the above systems are composed of human units, and each of them realises also (if not primarily) its own private biological, psychical and cultural purposes, and not only the purposes of the superior system. Moreover, such systems as cultures, states and economical firms realise their potential of expansion rather by increasing their own range of operation, than by producing decsendent systems with (almost) the same identity. Therefore, we cannot speak here about individuals, which are so important for the cybernetic definition of life.

(perhaps it would be necessary to wait for the described effect a little bit longer), although sexual practices of these animals rarely produce associations with the speed of light.

In reality, as it was emphasised above, the capacity of an environment is of course always limited. This concerns other quoted cases of negative feedback as well, e.g. the "environment" of an avalanche of stones (stone-spangled slope) or of a nuclear chain reaction (a sphere of a radioactive isotope of uranium). In these cases, a chain reaction finally dies out because environmental resources (namely: unstable stones on a slope and nuclei of a radioactive isotope of uranium) get exhausted. On the other hand, life still preserves its potential to a violent chain reaction even if it has exhausted the capacity of its environment at a given moment. The proofs of this statement are abundant. This **potential of expansion** is manifested not only in the case when some organisms take possession of a certain area, where do not face competition of other organisms or such competition is weak (e.g. a sudden increase of the population of rabbits brought to Australia). The spectacular incidents of this kind happen rather rarely. Normally, the expansion potential of organisms (carriers of a determined identity) expresses itself in a strong pressure to expel other organisms (having a different identity; they can be individuals of the same species or of very far related species) from the limited capacity of the environment. In stable ecosystems, the other organisms exert as a rule an expulsive pressure of the same or similar strength, but in reversed direction. In fact, the rivalry is between identities of particular individuals (complexes of properties or systems of negative feedbacks), tending to control the entire capacity of an environment at the cost of other identities. Tiny differences in the strength of the pressure exerted by particular identities lead to a gradual expulsion of some identities by others, and therefore to evolution.

It should be stressed once more that the identity of living organisms –the parameter "optimised" (maintained and propagated) by the system of negative feedbacks constituting a living individual – is something much more complex than a simple, elementary parameter, such as temperature, for example. An identity comprises not only temporary properties of an adult individual, but also the whole living cycle from a zygote to the adult stage, including also the ability to produce progeny. In sexually reproducing organisms, the identity of an individual constitutes a resultant or combination of identities (complexes of properties) of its parents (the phenomenon of recombination of properties). However, even in organisms reproducing asexually, the identity undergoes changes as a result of mutations. Therefore, the question is not that of preserving a particular identity, but rather that of preserving the **continuity of changes of this identity** in time. The parameter called identity is therefore undoubtedly a complex parameter – its proper understanding has a decisive influence on the understanding of the phenomenon of life.

In principle, each property of an individual, e.g. the pattern of its coloration, can be regarded as an element or result of a negative feedback. In this last case, different concentrations of a dye are "assigned" to different places on the surface of its body. Whether the assigned pattern of coloration is "good", and therefore allows to hide from predators, co-decides about the efficiency of a given system of negative feedbacks, and therefore identity.

Identities compete with each other to control (occupy) the capacity of an environment and better identities oust worse identities. This is illustrated in Fig. 2.11, which uses similar convention as the figures presenting the idea of natural selection. We begin here with four identities (1, 2, 3, 4) that compete with each other, "pushing around" as far as possible in the accessible life space. As identity 3 is the strongest (best adopted), and therefore has the greatest potential for expansion, it gradually ousts the remaining identities and takes over their places. The weakest identity 2 disappears first and is followed by identities 4 and 1. The entire life space (capacity of environment) becomes occupied by identity 3. However, there soon appears a small modification (mutation) of this identity, which gives origin to identity 5, a little bit better than its predecessor. Before identity 5 ousts totally identity 3, there appears (again as a result of a mutation of identity 3) a still better identity 6, which ousts both the identity 3 and identity 5, temporarily filling the whole capacity of the environment. Because the process described above has no end, we may speak only about a temporary victory of some identity, until there appears a still better identity, that is of greater potential for expansion (stronger positive feedback), maintained by a more effective system of negative feedbacks.

Contrary to living organisms, inanimate systems equipped with a positive feedback – such as an avalanche or a lump of a radioactive uranium isotope – do not possess any negative feedbacks, maintaining their potential for expansion. Once all stones rolled down the slope, an avalanche comes to a standstill, and no further neutrons are produced, once all nuclei of a radioactive isotope are split. Therefore, neither stones in motion nor free neutrons have any mechanisms oriented purposefully at ensuring their survival and production of descendent "systems" (stones in movement or free neutrons). Therefore, while both living organisms and some physical objects have a certain superior purpose in the form of (at least a potential ability of) expansion, only the former possess a subordinate purpose in the form of a network of regulatory mechanisms based on negative feedback, responsible for maintaining individuality (identity) of living organisms and their potential for expansion.

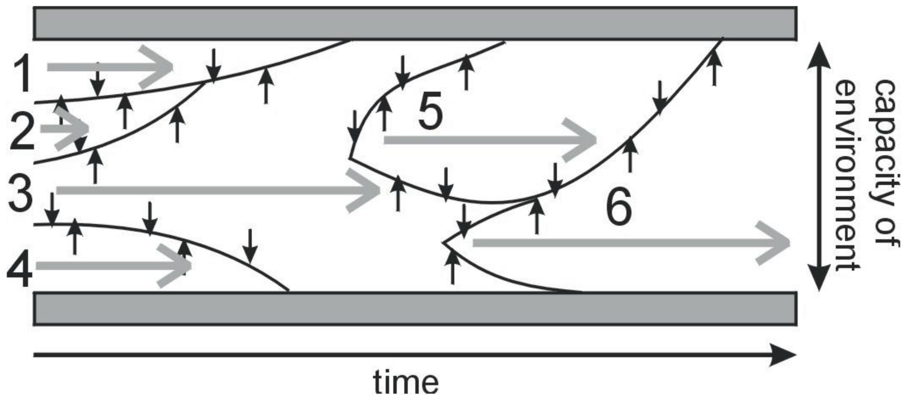

Fig. 2.11. Competition of six identities (complexes of individual properties) for the accessible capacity of an environment. Initially, identity 3 wins with identities 1, 2 and 4, but later it is ousted by identities 5 and 6 – its more successful "mutants". The large grey arrows designate the "impetus" of particular identities towards survival in time, while small black arrows designate the "pressure" exerted by a given identity on other identities, in order to oust them from the accessible capacity of an environment.

Many devices constructed by man (robots or thermostats) are equipped with such a subordinate purpose (system of negative feedbacks). However, they do not possess any superior purpose (in cybernetic terms), and even if they do, it is a purpose of their users. Devices do not take care of themselves, their own survival and production of identical descendent devices. Therefore, they do not constitute entirely autonomous systems (individuals), directed exclusively on themselves, but are elements of some superior system, the purposes of which (determined by man) they realise. If we wanted to treat a living organism as a certain device, we should ask: What purpose, **apart from** the organism itself, does this organism serve? What useful task has it been planned to execute? However, such a purpose or task simply cannot be found. From the point of view of man, we come to the paradoxical conclusion that living organisms serve no purpose, apart from themselves. A thermostat in a refrigerator regulates temperature to prevent spoiling of food for man – this is understandable. But a thermostat which has originated spontaneously and regulates temperature just for fun – isn't it strange? Of course, it is their own existence that auto-copying living organisms maintain. This resembles a thermostat built of pieces of ice, that maintains low temperature to keep itself from melting and, additionally, to produce thermostats similar to itself[41]. This already seems to be pure nonsense. However, this is just what the essence of life and evolution consists in. These phenomena, like stones falling down in the gravitational field, serve nothing and do not possess any purpose. They simply exist. Therefore, we have to abandon our anthropocentric point of view and the associated habits of evaluation in order to understand well the sense of evolution.

The thermodynamic aspect of the essence and evolution of life was described in one of the previous chapters devoted to thermodynamic evolution. Thermodynamics tells us that life, whatever else it would be, is **also** a dissipative structure participating in effective production of entropy on the surface of the terrestrial globe. As a certain determined structure, life is attributed with some degree of ordering, and therefore a **quantity of information** connected with it. However, thermodynamics says us **nothing** to answer the questions what the **type of ordering specific for life** consists in, how life differs from other dissipative structures (after all, convective currents or turbulences in a flowing liquid also possess some quantity of information), and therefore what the specific **quality (value) of information** is characteristic for the phenomenon of life.

Therefore, I would like to propose here a formulation of the idea of the quality of information related to life within the framework of a formal science, namely cybernetics. Above, I have prepared the background for this by defining living organisms as systems of subordinate negative feedbacks serving the superior positive feedback: the (rush to) expansion. In this context, the quality of biological information is equivalent to the **purposefulness** of negative feedbacks, manifesting itself in the realisation of the main "task" of living individuals (their identities), namely propagation. I propose hereby the following simple formula describing the quality of biological information, or, more strictly, the **amount of the biologically relevant (purposeful) information**:

---

[41]This example is intentionally grotesque. However, one could imagine an artificial life created by man or other thinking beings. This would be something like a robot producing robots similar to itself from substances (and energy) taken up from an environment. They would have to provide these descendent robots with a plan instructing how to built next similar robots. The record of this plan could undergo small accidental changes (mutations) due to imperfections of the physical carrier of this record. Such 'individuals' could compete between each other for the accessible resources of an environment as well as evolve. This – named 'dead evolution' by Stanislaw Lem in his novel "The Invincible" – constitutes in fact an evolution of artificial life.

$I_B = \log_2(n*2)$,

where: I_B denotes the amount of the biologically relevant information (in bits), n stands for the number of negative feedbacks present in a given organism, and 2 in the brackets corresponds to the minimal number of distinct states recognised by a given negative feedback. For example, the thermostat in a refrigerator is able to distinguish only between two states: 1. the present temperature greater than the assigned temperature, and 2. the present temperature not greater than the assigned temperature. It is also able to carry out only two types of reactions, respectively: 1. freezing, and 2. doing nothing. The thermostat is equipped with only one negative feedback, recognising only two distinct states, and therefore, according to the above equation, we deal in this simplest possible case with the quantity of cybernetic information equal to $\log_2(2)$, that is one bit. This definition of the biologically-relevant (purposely) information can be easily extended for the case where paricular negative feedbacks distinguish (and adequately respond to) a greater nunber of states (parameter values).

However, the information related to thetmostat is not of course biologically relevant information, since the negative feedback in the thermostat does not serve any positive feedback of its own, analogous to the tendency to expand observed in living creatures. Nevertheless, the valuable information (with respect to the superior positive feedback) "contained" in living organisms can be (at least in principle) easily presented in this way. The amount of the biologically related information is undoubtedly huge, even in the simplest living organisms. Moreover, the more complex a given organism is, the more biologically relevant information it possesses (the amount of negative feedbacks constituting its identity is greater). Therefore, the cybernetic formulation of the essence of life can offer a useful criterion not only for the problem of what life is, but also for the problem **how much** life is present in a given case.

Apart form the maintenance of its own potential for expansion, one should indicate an additional aspect of the system on negative feedbacks, namely the fact that such a system becomes an **individual**. I already discussed this term in the previous subchapter as a term co-defining the phenomenon of life together with the terms: reproduction and evolution. Within the cybernetic terminology framework used here, to be an individual means to possess a certain autonomous purpose, expressing itself by the above-described combination of positive and negative feedbacks (a system of negative feedbacks subordinated to a positive feedback – potential for expansion). This purpose assumes the minimum plan: to survive, and the maximum plan: to spread its own system of negative feedbacks (that is its properties, its individual identity) in a possibly great number of copies similar to the original. This purpose should be considered as completely devoid of reason and, in fact, it seems to be totally nonsensical from the anthropomorphic point of view. For what sense could an unrestrained multiplication and spreading "without purpose" have? However, a living individual has two alternatives to chose from (the "choice" consists simply in the occurrence of accidental mutations, changing the system of instructions recorded in the genome): either it accepts this purpose and exerts as strong expansive pressure as it can afford (even not so much to "gain living space," but first of all not to loose it to others, who also strive forward with all their might; compare the model of the Red Queen[42]), or it is immediately eliminated from the game. In the latter case, its properties do not pass on to its progeny and vanish, and the entire capacity of the environment is filled by individuals who inherited from their parents the highest potential of expansion. There is no room to "deliberate" on the sense of such an enterprise (this is, of course, a metaphor). From the anthropocentric point of view, such a thoughtless realisation of a purpose – a plan imposed simply by a system of cybernetic mechanisms – seems at best to be sterile, and decidedly repellent in the worst case.

An individual, in the sense of the term preferred here, is in principle identical with the subject of evolution. To avoid confusion between the traditional understanding of an individual (mainly structural in its essence) and the evolutionary-cybernetic definition proposed in this work, I will introduce the term "**evoluon**" to refer to the latter. I will use this term interchangeably with the term "cybernetic individual" and in opposition to the term "structural individual". We will see in a while that a cybernetic individual (evoluon) is not always identical with a structural individual. For the sake of the present considerations, I will shortly discuss one case to visualise the abstractive definition of the cybernetic individual in an example.

In the light of what was said above, an evoluon is a self-copying system, the exclusive purpose of which is constituted by self-multiplication, and not only a part of such a system. This property does not necessarily refer to a structural individual. This is discernible best in the case of a colony of ants, composed of a reproducing queen and asexual castes, workers and soldiers,

[42]This model, born in the frame of population genetics, describes the situation in which only an as fast evolving as possible ensures a 'keeping in the same places', which means not loosing the competition with the rivals which also evolve possibly fast, that is not being pushed out from the accessible capacity of an environment. The name comes from the book by Lewis Carroll

taking care of the queen and protecting its progeny. According to the traditional (mainly structural) paradigm of an individual, an ant worker **is an individual** (at least by analogy to the individuals capable of reproduction in the case of non-social insects; it possesses a head, limbs, abdomen and the whole rest, except a functional reproductive system). From the cybernetic point of view, however, a worker of ants **is not an individual**, because it serves a superior purpose of some greater entity – in this case the whole ant colony – and is not self-sufficient in the realisation of this purpose (a worker is connected with other members of the colony by a network of behavioural negative feedbacks). On the other hand, the entire colony **is a cybernetic individual (evoluon)**. It possesses a reach network of negative feedbacks, regulating (for instance, by means of pheromones) the values of various parameters, e.g. the numerical force of particular casts (including sexual ones). The queen cannot be considered as an evoluon either, for it constitutes only a part of the system of negative feedbacks enabling the reproduction of the whole colony (production of descendent colonies). The queen does not realise its own potential of expansion while **depending only on itself**. Similarly, e.g. in the case of a horse, the **whole** organism is a cybernetic individual, and not only its somatic (corresponding to workers) or reproductive (corresponding to a queen) cells, that, after all, exhibit a relatively large structural and functional autonomy. (Let us remark that a horse comes from unicellular organisms and is, from the formal point of view, a strictly integrated **colony** of particular cells). In unicellular organisms, evoluon applies to a whole cell, and not to its genetic apparatus or particular genes ("selfish DNA"). For, only the entire system of negative feedbacks maintaining its potential for expansion is a cybernetic individual. Summing up, a sterile ant worker cannot in any case be regarded as a cybernetic individual (evoluon), because this worker realises the biological-evolutionary purposes of an entire colony, and not its own purposes.

I realise that the above definition of a cybernetic individual (evoluon) may seem contradictory to the intuitive, common understanding of the term "individual". In this understanding, the criterion of "individuality" is first of all a structural-functional criterion, indicating structural-functional separation of some system. However, there arises a question, if such a paradigm is sufficient and if it is heuristically fruitful. Both brain and liver in our body, as well as the entire biosphere, constitute examples of systems clearly separated in the structural and

---

Throughy the Looking-Glass (the second part of *Alice's Adventures in Wonderland*), where the Red Queen only running with the greates possible speed could keep in the same place.

functional sense. The same can be said about above discussed cells. However, they are certainly not individuals. For, they do not realise **self-dependently** any purposes of their own. Organs in a body are subordinated to the interests of the whole organism, while no purposes of the biosphere can be formulated within the (non-subjective) framework of cybernetic terminology. For the same reasons, neither a robot, nor a computer, or a thermostat is an evoluon. Because, they do not realise any autonomous superior purposes of their own, but only the purposes imposed on them by man.

The case of multicellular organisms (the example of a horse) indicates that the intuitive concept of an individual is founded on human categories of seeing the world, directed at **spatial structures**. Now, I would like to develop this topic in more detail. The body of a horse is composed of billions of cells, fulfilling various functions, integrated by mutual interactions into one whole. However, it is known that far ancestors of mammals (and generally of all tissue organisms) were single cells, each of which constituted an independent evoluon. Later, as a result of incomplete separation in the process of division, cells began to form multicellular aggregations, differentiating in the course of time into reproductive cells and cells fulfilling other functions. A horse is such a multicellular aggregate. However, nobody uses the term "individuals" to refer to the cells entering into its composition – only the entire organism of a horse is an individual, because it is this entity that realises its own potential for expansion. After all, just as a horse constitutes an aggregation of many cells, the colony of social insects is an independent complex of particular insects, realising its own superior purpose and not the purposes of its members. Particular insects in a colony are interconnected by an abundant network of negative feedbacks, comprising an exchange of food, regulation of the numerical force of particular castes, maintenance of constant temperature and humidity in the nest, as well as many other purposes. A tenacious defender of the common concept of an individual (stressing the structural aspect) could draw our attention to the huge difference between a horse and a colony of social insects (to horse's advantage) in the degree of system organisation as well as of integration of its particular elements. This argument misses the point, however. Let us consider a primitive multicellular organism, for example *Volvox*. Although its cells are already differentiated into reproductive and somatic cells (as insects in a colony), the number of these cells as well as the degree of their functional integration certainly do not equal the wealth and complexity of the biggest colonies of ants and termites. Nobody, however, says that not the entire *Volvox* is an individual, but its building parts (cells), although both kinds of systems (*Volvox* and a colony of insects) seem to be very similar from the formal point of view.

Why, therefore, *Volvox* is considered to be an individual, while a colony of insects is usually denied this status? The reason is quite trivial. The organism of *Volvox* is a complex of cells integrated mainly **structurally**, while the **functional** component predominates in the interrelations of insects in a colony. Human brain integrates stimuli from the external world (first of all visual stimuli) in such a manner that it is easier for the brain to see structural relations, while functional connections are less pronounced. Therefore, it is easier for us to consider a robot as an individual than to view a colony of insects in this way. The structural aspects, so important for the way we view the world, seem to be largely accidental for the biological evolution. As we will see below, systems whose spatial organisation differs so much as in the cases of a thread of DNA, a single cell, an elephant or a colony of insects can be cybernetic individuals (evoluons). This results from the fact that the functional aspect – in the cybernetic formulation – provides the main (in my opinion) criterion that distinguishes an individual in the context of biological evolution. According to this formulation, an individual is a system of negative feedbacks, realising its own superior purpose – expansion of its identity, regardless of how these feedbacks manifest themselves in the material-spatial aspect. For the above reasons, I consider the commonly accepted (structural) concept of an individual as misleading and heuristically sterile, and propose to introduce instead the evolutionary-cybernetic paradigm, which identifies a (cybernetic) individual with what I have named evoluon.

In the previous chapter, I described some cases beyond the reach of the traditional paradigm, for it does not supply any decisive criterion for an univocal determination of what is an individual (and to what extent) and what is not. What is worse, it does not offer any objective measure or point of reference that would allow to order the discussion of the above topic. In the present chapter, I would like to show what the cybernetic approach proposed by me can offer in this respect. I have already discussed its application to social insects. I will discuss other cases later, while now I will deal with myxophyta and colonial coelenterates.

In myxophyta, it is possible to treat as an individual either particular amoebas or a plasmodium formed to produce spores, if nourishing substances are lacking in the environment. I think that in this case the genetic identity of amoebas in a given population imposes the choice of the system identified with a cybernetic individual, although it may seem strange at the first look. If there is an identity of this type, a plasmodium – or the population of amoebas on a given area which can be treated as an "individual in dispersion" – should be recognised as a cybernetic individual (the fact that different relations occur between amoebas during formation of a sporangium, so that the entity they compose constitutes one system of negative feedbacks, decides about the classification of the group of amoebas as an individual). It is probable that the entire local population comes from one spore and then the situation seems to be clear.

What happens, however, in a reversed situation where a plasmodium is formed of a few genetically different groups of amoebas? Then, the plasmodium can be treated as some kind of an intra-species reproductive symbiosis (let us remember that formation of a sporangium and dissemination of spores is the purpose of a plasmodium) of a number of cybernetics individuals (groups of genetically identical amoebas). In this case, there appears an interesting possibility of a reproductive parasitism. As I mentioned before, only some of the amoebas transfer their genetic material to spores, while the rest builds a stem and walls of a sporangium. A symbiosis would make sense only if each of the individuals (genetically identical groups of amoebas) participates to a comparable extent both in production of spores and in building of a sporangium. However, if a group of amoebas constituting one individual were able to outwit other groups in such a way that it would dominate the production of spores, while the other groups would be pushed out to somatic functions (formation of sporangium), then such a strategy would gain an evolutionary success[43]. This would result in an origination of a parasitic population (species), exploiting another population (species) in order to realise its own reproductive functions. A similar mechanism is encountered in the case of social parasitism in ants, where a parasitic species (closely related to its host) is devoid of the caste of workers and exploits a colony of its host for production of progeny, eliminating partially or totally the sexual forms (queens) from this colony. I think that the possibility of sexual parasitism in myxophyta indicates a new and interesting direction of future studies. Genetic differentiation of amoebas in a plasmodium seems to be the key issue.

The problem of genetic identity occurs also in colonial coelenterates. A colony of coelenterates, originating by gemmation of some polyps from other polyps, differs in fact little from a clone of bacteria reproducing by division, or from a population of a partenogeneticly

[43]However, this depends on many factors, for example on the probability of meeting a "host" by such a parasite.

reproducing plant on a meadow. Of course a meadow cannot claimed to be an individual, because particular plants do not co-operate (there is no network of negative feedbacks between them) in propagating their common genetic identity. The matter is less simple in colonial coelenterates, because in this case there is some degree of co-operation. Therefore, some aspects of "individuality" have been ceded here to the entire colony. The fact, however, that a single element of a colony of many colonial coelenterates – i.e. a polyp with a mouth orifice and a set of tentacles – possesses gonads is an argument in favour of recognising it, at least partially, as an individual. It is genetically independent and works on its own account within evolution. This is particularly manifested in the case where particular polyps differ genetically as a result of a somatic mutation, and therefore, they are evolutionary competitors[44]. The same applies to plants in a meadow. Even if they all are genetically identical, nothing connects them functionally and each of them will follow its own (evolutionary) way from the moment of the first mutation.

In a colony of coelenterates, some aspects of an individual seem to be attributed either to particular polyps, or to the whole colony. The network of negative feedbacks, which is also spread between different polyps – although it is much denser in each particular polyp – has a crucial significance in this case. On the other hand, genetically identical specimens of some species, not connected with each other by any negative feedbacks, should be considered as separate individuals, because each of them has, due to inevitable mutations, a possibility to follow its own evolutionary way and, if it does not become extinct without progeny, it will certainly go this way.

In the above case, the conceptual tools of the cybernetic paradigm do not allow one to indicate sharp division lines between individuals and non-individuals, because such a differentiation is impossible here. However, such a demarcation seems to be attainable in most other cases (those already discussed and those to be discussed below). Where it cannot be realised, the concept of an evoluon provides a convenient instrument measuring the "degree of individuality"[45]. In general, the main role of the cybernetic formulation is to create a certain

[44]The situation is even more complicated, because many colonial coelenterates exhibit a polymorphism of polyps: different polyps are specialised structurally and functionally to perform different tasks: nourishment, defence or reproduction. Although a nourishing polyp without gonades can be regarded as an individual in the structural sense, it is certainly not an individual (evoluon) in the cybernetic (evolutionary) sense.

[45] The ‘degree of individuality’ (d) may be easily defined (this is my own proposition) in the following way:

general point of reference, which allows to speak in a sensible, universal and possibly objective way about the phenomena of life and evolution. This paradigm offers a certain criterion which, if applied to different phenomena, allows one to order and classify them, or at least relate them to other phenomena in a fertile way. Finally, which is perhaps most important, it helps to form an intuitive opinion on the essence of life.

I will return for a moment to the problem of the subject of evolution, that is the fundamental evolving unit. This status was attributed to single genes (Dawkins), auto-copying systems or their groups. I will not describe here the arguments supporting each of these opinions. I mentioned earlier that, within the approach presented here, the system of negative feedbacks maintaining the superior positive feedback, that is simply an evoluon, becomes, automatically in a sense, the subject of evolution. Formulating the matter more strictly, one should not speak about an evolution of individuals, but rather about an evolution of their identity (complex of properties). For it is identity that is associated with something divisible and gradable, able to undergo continuous changes. Therefore, this formulation does not generate greater semantic troubles in the case of e.g. sexual reproduction, in which not a part of a parental individual, but just a part of its identity is transferred to progeny. This small shift of significant accents changes nothing in the logical construction of the cybernetic paradigm, while it seems to fit better our common conceptual tradition.

The case of "suicidal plasmids"[46] living in bacteria testifies best that it is identity (and therefore a certain biological **information**, as discussed above) and not any particular carrier of this identity in the form of an individual, that constitutes the subject of evolution. Those plasmids contain two genes: one conditions resistance against a determined kind of toxin (colchicine), the second is responsible for production of this toxin. The protein encoded by the first gene is

---

$$d = \frac{I_{Bi}}{I_B}$$

where I_B is the overall biologically relevant information (as defined above) of a considered ‘candidate’ for an individual, taking into account both negative feedbacks serving only the ‘candidate’ and negative feedbacks serving the whole system a part of which the ‘candidate’ is (e.g. a colony or the entirety constituted by symbionts), while I_{Bi} stands for the biologically relevant information related only to the ‘candidate’ for an individual and not to the whole system (serving only its own purposes and not the purposes of the whole system, e.g. colony). Of course, d = 1 in the case of a fully self-depencent individual, while 0 < d < 1 in the case of ‘semi-individuals’ in a colony or in the case of symbiosis.

[46]I have taken this example from the book *Problems of Biology* by John Maynard Smith.

synthesised non-stop, while the production of colchicine is triggered when the condensation of bacteria exceeds a certain level. The production of colchicine leads to death of the bacterium cell producing the toxin (even the gene of resistance for colchicine cannot help in this case, which is due to a great concentration of the toxin in the bacterium cell which produces it), and therefore also of the plasmids contained in this cell, as well as of all other bacteria **not** containing the discussed plasmids. This facilitates significantly the survival of the remaining bacteria possessing these plasmids. Therefore, by committing a suicide, a given copy of a plasmid promotes the survival of other copies possessing the same identity. This is an example of a genetic "altruism" slightly similar to the one encountered in the case of asexual castes of social insects.

Let us come back, however, to the analysis of particular phenomena in the perspective provided by the cybernetic paradigm to characterise life and evolution. As I mentioned, cybernetic properties of a living individual are not exhibited, for instance, either by stones in an avalanche or by robots. To be sure, the former are able to transfer their "individual" properties – i.e. in this case the movement first of all – to the surroundings (other stones). However, they do not possess any purposeful mechanisms to maintain the potential for expansion, and at the same time, as a consequence, they are not attributed with a well defined property of being an individual (they do not possess an identity and a network of negative feedbacks related to it). A complicated robot, on the other hand, has a rich complex of negative feedbacks. This device, however, has not got any possibility to use this complex for production and support of its own potential of expansion. On the contrary, all its superior purposes are determined by man. Therefore, if I were to decide which of the above examples is closer to the essence of life, then (contrary, perhaps, to most people) I would point to an avalanche, where the stones possess at least their own superior "purpose" (which can be identified with movement), although in a rudimentary form, and are self-supporting in the "realisation" of this purpose.

I have recalled the case of an avalanche of stones on purpose. For we already know an example of quasi-life, which in the formal (cybernetic) respect is practically identical with a stony avalanche. These are the before-mentioned prions, that is infective protein molecules, causing some diseases of brain in animals and people. As I said above, these molecules may appear in two forms: normal and infective, differing one from the other by a conformation (spatial arrangement) of the protein chain (the metamorphosis of Dr Jeckyll into Mr Hyde consists here in the prosaic transition of an α-helix into a β-structure, similar to an accordion). Normal molecules occur in a healthy brain and fulfil there some (yet unknown) function. Infective molecules stimulate (most probably through an ordinary contact) a transformation of normal molecules into infective molecules, that can transform still other molecules, which finally leads to the accumulation of the infective form and to death of the organism (it is not known exactly how – either the infective form or the lack of the normal form may be harmful). This sequence of events can be initiated by delivering at least one external molecule in the infective form, or by a spontaneous transformation of a normal molecule into an infective variant. Prions do not have any complex of negative feedbacks to maintain their potential for expansion. The phenomenon of transformation of the normal into the infective form is a simple physicochemical process – it does not involve even the simplest network of mutual relations of separate elements. The analogy with an avalanche of stones is therefore strict, for there one also encounters a stone moved by somebody or spontaneously separated from a rock due to erosion, that transforms other stones from the motionless form into the moving form and a chain reaction takes place. An infective molecule of prion, like a moving stone, has not got a system of purposeful mechanisms directed at reproducing itself in the environment (which is constituted by normal molecules of prion and still unmoved stones, respectively). Elementary physical phenomena or chance, whatever one prefers, and not even a substitute of a purposeful action, decide whether a next prion molecule (stone) undergoes transformation (moves). Therefore, prions are not alive.

However, we will come to different conclusions when the cybernetic paradigm is applied to viroids, i.e. naked molecules of RNA living of plants. They are able to get into the cells of their hosts and induce an intensive production of their own copies, which disturbs the functioning of a cell (deprives it of energy and building substances), leading to its death. To be sure, viroids have not their own metabolism, as viruses. Therefore, one could argue if they possess their own system of negative feedbacks. This would be a futile discussion. For us, it is important that viroids are able to enter into the system of negative feedbacks of a cell of a host and to change its course by modifying the optimised **parameter**. This optimised parameter is simply a (possibly great) number of copies of a given viroid, while the regulation of such parameters as the concentration of ATP, nutritional substances and so on becomes neglected, which finally leads to death of a cell and release of viroids (naked, circular molecules of RNA) which can next infect other cells. The whole structure of a viroid, composed by the sequence of nucleotides in a thread of RNA, is **adopted** to penetrate into cells of a host and to shift their metabolism to enhance the proliferation of the viroid molecules. This aim becomes preferable at the cost of synthesising nucleic acids of the cells and, generally, of any other processes. The identity (sequence of nucleotides) of a viroid may evolve, increasing its infective effectiveness, as well as enabling it to infect other species of plants.

All living functions of a viroid are performed by enzymes of its host, at the cost of its energy and building substances. This "ceding" on the host of a part of the network of negative feedbacks needed by a parasite is typical for the system parasite-host. In principle, this is what constitutes the essence of parasitism. However, in viroids, this cession has adopted an extreme form – in fact, all the regulatory mechanisms are transferred to the host, while the task of a viroid is only to enter into the system of negative feedbacks of the host and modify it to realise its own reproduction. Since viroids are evidently directed "intentionally" to perform this task, they should be recognised from the cybernetic point of view as evoluons.

The whole sequence of events caused by viroids is strictly **purposeful** and leads to multiplication of viroid molecules, as well as to infection of next host cells or individuals. Such a phenomenon is not encountered in the case of prions. In the normal form, prions most probably fulfil some useful function in the host organism (they are encoded in its DNA, after all), while the harmfulness of their infective form is simply **accidental**, an evolutionary trap of the organism producing them, something like auto-allergy[47] (but this is a rather far fetched analogy). We have therefore no reasons to think (and this seems impossible for fundamental reasons) that the structure of prions has been improved by evolution for the purpose of increasing their infectiveness. As they do not constitute a subject for evolution – they are not able to transfer the information about their own structure (sequence of aminoacids) to their "progeny", and therefore to preserve continuity of their identity in time. This information is not in themselves, but in a cellular nucleus, on which they have no influence. In fact, not prions undergo replication (the number of them is constant in time), but only their **infective conformation** (at the cost of the normal conformation). Therefore, prions are by no means evoluons, contrary to viroids. Nevertheless, the comparison of prions and viroids (both phenomena represent a single chemical molecule) significantly decreases the precipice separating, as it could seem, life and an avalanche.

[47] In the case of autoalergy, the immunological system of a given animal produces antibodies against the own proteins of this animal what, as in the case of prions, constitutes an example of an action of a certain part of an organism for a harm of the whole.

If the status of evoluons is awarded to viroids, it is certainly granted to viruses as well, for their units, virions, are molecules of RNA or DNA encapsulated in a protein cover, protecting them and facilitating the process of infection. The molecule of a nucleic acid in a virion is much greater than the naked thread of RNA of a viroid. It encodes a few proteins, including the proteins of the cover and the proteins participating in copying the DNA or RNA thread of the virus in a host cell, as well as taking part in the synthesis of the proteins of the virus's cover. Viruses already exhibit germs of their own network of negative feedbacks. Bacteriophages inject actively their nucleic acid thread into a bacterium cell, while many viruses encode their own enzymes, optimising the velocity of multiplication of descendent virions in a host cell. From the point of view of the cybernetic paradigm, however, viruses should generally be put in the near neighbourhood of viroids.

Another interesting case – in a sense intermediate between prions, on the one hand, and viroids and viruses, on the other – is constituted by neoplasms (tumours), especially malignant tumours (cancers). The abnormal transformation of a cell can be treated in cybernetic categories as abandoning by this cell of the realisation of the superior purpose of the whole organism and imposing its own superior purpose, that is simply its own expansion. A neoplasm cell does not know that its life will finish with the death of the whole organism (after all, this is not always true – many tumorous lines are presently cultivated for scientific purposes), as dinosaurs did not know that they would become extinct as a result of a fall of an asteroid. Both cases differ of course by their time scale – and by not much more. We frequently consider cancer as a completely nonsensical phenomenon, in opposition to life. However, this is a delusion of human brain, caused by our tendency to anthropomorphism. A tumour in not less (and not more) sensible than life in general (more strictly: the difference is only quantitative and not qualitative). Tumour cells simply do not possess appropriate mechanisms to secure their survival after the death of the organism they take into possession. However, in the "environment" of an organism, they manage not worse than bacteria on a medium. The fact that tumour cells do not have time for further evolution and self-improvement is caused by a limited capacity of their "environment". Therefore, they are not able to produce early enough a sufficiently complex network of negative feedbacks (they do not lack a positive feedback!). In other words, tumour cells have insufficient time to adopt to their broadly understood ecological niche. Let us notice that the evolution of "normal" parasites tends to decrease their harmfulness for a host – because, it does not pay to kill the organism they feed on. Of course, a tumour differs from a normal parasite by its inability to infect another host[48]. On the other hand, during its short period of existence, a cancerous cell frequently produces at least some **purposeful mechanisms** (negative feedbacks), resulting for example in insensibility to the immunological system or other defensive reactions of the organism, adaptation of metabolism to a lowered oxygen concentration, large rate of growth and divisionns, ability to migration and so on. Thus, they ensure survival and development in its environment (that is the host organism). Therefore, such a cell is a cybernetic individual (evoluon) from the formal (cybernetic) point of view. Within the cybernetic paradigm, neoplasms are, in a sense, located somewhere in-between prions and viroids (and viruses). Tumours possess a network of negative feedbacks ensuring their expansion in a temporary (but this is a problem of an arbitrary scale of time – the existence of life on the terrestrial globe is also temporary if counted in billions years), limited environment of a living organism. They are also characterised by some origins (germs) of evolution (e.g. the pressure of the immunological system can favour the best-adopted tumour cells). If they were allowed to live long enough, they would certainly start to evolve normally. This situates tumours higher than prions, that do not possess any system of negative feedbacks. Tumours, however, cannot survive and propagate in a broader and more persistent environment, such as a population of individuals of a host organism, conversely to viroids and viruses, able to transfer from one individual to another. This proves a greater degree of animation of the last ones, although viruses and viroids have a much less complex structure than tumour cells. What counts here, however, is the cybernetic criterion.

The last[49], incredibly interesting phenomenon I would like to discuss is that of parasitic DNA. It stands on the border of life and inanimate processes (but more on the side of life). One of

---

[48]However, there exists at least one example, already discussed, where a cancer has become a typical, 'normal' parasite. In the case of the venereal sarcoma in dogs, cancer cells can infect other host organism, transferring themselves from dog to dog during sexual contact.

[49]I do not discuss here plasmids in detail. They are small, circular molecules of DNA, several copies of which are encountered in bacteria cells. They are able to transfer themselves from one cell to another, although their status seems to be peculiar. Plasmids can transfer genes responsible e.g. for immunity to antibiotics or for bacteria sex, and their reproduction is independent of the multiplicaton of DNA in a bacterial chromosome and division cycle of the cell. Plasmids are profitable for its host, and it controlls the number of their copies. Additionally, plasmids can incorporate themselves into bacteria DNA, and afterwards leave it, while carrying various genes with themselves. Because plasmids can be transferred from one bacteria cell to another, this leads to an exchange of genetic information between bacteria. All of this indicates that plasmids constitute, in a sense, a substitute of sexual reproduction, leading to a recombination of genetic

its forms is constituted by the repetitive DNA occurring in the genome of the majority of higher organisms. A large part of or even the entire DNA entering into the composition of the genome of different organisms is the so-called non-repetitive (unique) DNA. This name comes from the fact that the sequences of nucleotides occurring in any place of this DNA do not occur in any other place of the genome of a given organism[50] (of course, the shortest, few-nucleotide sequences, which naturally must repeat, should be excluded here). Many parts of DNA, first of all the fragments of chromosomes encoding particular proteins, that is simply genes[51], belong to the non-repetitive DNA. On the other hand, all higher organisms (including man) possess also some amount (frequently very great) of the repetitive DNA, where nucleotide sequences occur in a certain number of copies. The relative amount of the non-repetitive DNA in a genome varies in different organisms, being particularly high in vascular plants and caudate amphibians (*Urodela*). We classify this DNA into three groups, comprising a small, medium or large number of copies of repeatable sequences. A part of the repetitive DNA plays undoubtedly an important role in the cell functioning, as it regulates the expression of genetic information and controls the processes of growth and differentiation of an organism. However, there is evidence suggesting that the remaining part does not serve in any way the interests of an organism. In such a case, there instantly arises a question, why an organism does not simply get rid of it. The copying of such DNA requires additional energetic expenditures, and a greater volume of the chromosomes containing this DNA makes the process of karyokinesis (division of the cellular nucleus) more difficult. Removal of the repetitive sequences would undoubtedly be advantageous for a cell an the biological evolution should have done so a long time ago. Why, therefore, didn't it happen?

---

information and increase in the pool of this information in the entire population of bacteria. On the other hand, some autonomy of their reproduction as well as the lack of constant connection with a given cell (the possibility of 'infecting' other cells) suggests that they can realize some purposes of their own and that they have their own potential for expansion. In this formulation, they would be an untypical kind of symbionts, exhibiting some individual properties. At the present, it is difficult to determine univocally their status.

[50]Of course, diploid organisms (possessing two copies of the genetic record), to which a great majority of higher organisms belong (including man), have two copies of non-repetitive sequences in DNA.

[51]Few genes occur in a greater number of copies, as for example the genes encoding histons or rRNA.

The only possible explanation[52] can be reduced to the hypothesis that the parasitic DNA has some mechanisms that allow it to avoid being removed from the genome of a given cell and to transfer to the genomes of descendent cells. Therefore, it possesses a system of negative feedbacks maintaining its existence (a possibly large number of copies of parasitic repetitive DNA is here the regulated parameter; it should not be too great, however, not to kill the carrier-cell; contrary to viroids, the repetitive DNA cannot infect other cells and therefore must undergo propagation together with the multiplication of the genome during cell divisions). In this respect, it exhibits properties of a parasite, that does not allow to be thrown out from a host organism either, and exists at its cost. On the other hand, the matter of the potential for expansion, that is the positive feedback, seems to be less clear. The repetitive sequences could multiply without any limitations, but this would lead to the death of a whole organism, and therefore of the sequences. The fact, that they multiply only to the extent that secures normal functioning of the organism, certifies that they have achieved a higher level than tumours which are less "anticipating". What does therefore their potential for expansion consist in? In the case where the genomes of all individuals in a population contain the maximal admissible (for an organism to function undisturbed) amount of the repetitive DNA, it can be said that the environment for the parasitic repetitive sequences is saturated and therefore no expansion is possible. On the other hand, the potential for expansion can manifest itself in competition between different kinds of repetitive DNA, in the tendency to oust rivals from the genome. The situation is identical as in the case (shown in Fig. 2.11) of competition between different identities of "normal" individuals for a limited capacity of an environment. At this point, the positive feedback is already manifest.

We can also speculate what will happen when individuals without parasitic DNA (or with a lowered amount of it) appear in a population (as a result of a mutation or emigration from another population). The parasitic DNA seems to be connected with diploid organisms reproducing sexually. Such organisms possess two copies of the genetic record, one obtained form the mother and the other from the father. In this case, if a given organism obtained a complete set of chromosomes without parasitic DNA from one of its parents, then these chromosomes would probably become infected by repetitive sequences coming from the other parent. The non-functional (for the organism) repetitive DNA would be, therefore, sort of a parasite of

[52]Under the assumption that the so-called parasitic DNA does not fulfil any function unknown to us, and does not brings profits for the whole organism.

chromosomes, able to transfer from one chromosome to another, but never leaving not only the organism (or its reproductive cells), but even the cellular nucleus of the host. The last property differs parasitic DNA from, for example, viroids that transfer from one individual to another and live of entire cells rather than exclusively on chromosomes. For this reason, the parasitic repetitive sequences of DNA can be situated slightly below viroids on the "inanimate-living" scale, but already on the side of the living. As we can see, a pretty big crowd forms in our precipice existing seemingly between the phenomenon of life and inanimate nature. Nevertheless, the cybernetic paradigm determines a sharp border between the inanimate and the alive.

The so-called supernumerary chromosomes B are another example of the parasitic DNA. They have been found in hundreds of species of plants and animals. These chromosomes are not necessary for a normal functioning of the organisms containing them (this is shown by the simple fact that individuals without chromosomes B develop equally well, or even better, than their carriers) and do not bring any visible profits (with few exceptions, doubtful in any case). Chromosomes B simply use their "carrier" to multiply (during chromosome copying when a cell divides) and spread, which does not remain without influence on its energetic balance (the copying of the genetic information is an energy-consuming process). It seems that organisms try to get rid of their chromosomes B, for example by loosing them during cellular nucleus division, and that chromosomes B are able to counter-act such attempts. Chromosomes B are therefore most probably parasites of the genome of their host, like some repetitive sequences, although they differ from repetitive sequences in their structure (a chromosome is a much more complex object than a short DNA thread). As they exhibit potential for expansion and possess a supporting network of negative feedbacks (remaining in a cell and copying during its division can be regarded here as the "regulated parameter value," and this is simply realisation of the positive feedback: the potential for expansion, while the attempts on the part of their carrier to get rid of chromosomes B can be regarded as the "disturbing influence of an environment"), which testifies unfailingly to the fact that they are cybernetic individuals (evoluons).

The most spectacular example of parasitic DNA I am aware of is the cytochrome B parasiting on the genome of the vesp *Nasonia vitripennis*. The number of sets of chromosomes decides about the sex in this insect – males are haploid (single set of chromosomes), while females are diploid (double set of chromosomes). The determination of sex in progeny is regulated by fertilisation of only a part of (haploid) egg cells by (haploid) spermatozoons. Females develop from fertilised cells, while males develop from non-fertilised cells.

The chromosome B is transferred in spermatozoons of *Nasonia vitripennis*, together with normal chromosomes coming from a male. After fertilisation, the chromosome B induces a removal of the whole set of the father's chromosomes, and therefore an extermination of the genome it was a part of. The zygote, therefore, becomes haploid (there remains in it a single set of chromosomes coming from a mother and, of course, the chromosome B) and develops into a normal male which also produces spermatozoons carrying chromosomes B. In the next fertilisation, they lead to extermination of the genome of their next host. The selfishness of the chromosome B in the discussed vesp exceeds significantly the selfishness of "normal" chromosomes B. While the latter only insignificantly decrease the vitality of the genome they live off, the chromosome B in *Nasonia vitripennis* manifests total ingratitude towards its temporary host by eliminating its whole genome. The discussed chromosome B is therefore unquestionably a subject of evolution possessing its own purposes and perfectly adopted to their realisation in its own "environment". In other words, it is an individual in the cybernetic meaning of this word (evoluon). Therefore, while one could still have doubts if "normal" chromosomes B are actually parasites, they do not seem to arise in the described case.

Parasitic DNA (both repetitive sequences of DNA and chromosomes B) is – beside tumours – another example of something that could be named "life in life". This phenomenon would consist in the fact that a certain part of a living system (organism) – an element of its network of negative feedbacks – serving initially the realisation of the superior purpose of (the identity of) this organism, that is its expansion in the external environment, starts to tend to its own expansion in the "environment" of this organism, which is due to an accidental mutation. Such a possibility is allowed for by the hierarchical structure of a biological organism. On each level of organisation (tissues, cells, organelles, macromolecules), particular building elements (e.g. cells) exhibit a high degree of autonomy, necessary for their functions to be performed effectively. The autonomy of cells is also due to historical reasons. Once (in the evolutionary scale of time), cells led an independent life (as unicellular organisms), while multicellular organisms originated, when dividing cells stopped separating after the process of division. As self-dependent objects, cells had to possess a full set of mechanisms allowing them to survive and expand. Many of these mechanisms remained still efficient in cells functioning within the bodies of multicellular organisms, and their subordination to the function of the whole organism requires many controlling mechanisms. These mechanisms however are not unfailing. An abnormal transformation into a tumour consists, in a sense, in gaining full autonomy by a cell (group of cells), which it lost hundreds millions years ago. On the other hand, in a quest for free-living ancestors of the parasitic repetitive DNA, one would need perhaps to go back as far as to the Eigen's hypercycles or the "RNA world" (discussed in the next chapter). Neither can one exclude that viroids and at least a part of viruses originated by such an "autonomisation" of structures, that earlier functioned as parts of more complex systems. Alternative hypothesis assumes that these objects are a result of an extreme simplification in the process of evolution of cells of the simplest prokarionts (such as intracellular parasites – rickettsias and mycoplasms that, as parasites, had also enough time to loose much of their autonomy).

The above-discussed examples concern the case where secondary systems – with their own purposes and potential for expansion, i.e. independent individuals – become differentiated within the framework of an initially homogeneous system, a single individual from the cybernetic point of view. However, a reverse phenomenon is also possible, namely, joining of independent individuals into a certain whole, which can lead in extreme cases to formation of one individual, one subject of evolution. A typical example of this situation is found in symbiosis, consisting in such co-existence of two organisms (or, more broadly, species) that both partners derive profit from this co-existence. Lichens are a leading example of symbiosis. Their body is constituted by a combination of cells of a fungus and an alga. Neither of the elements can survive separately in natural conditions. However, both components retained autonomous systems of reproduction and can exist independently in laboratory conditions. One could ask in this context: how many individuals do we encounter here? One or two? (for the sake of simplicity, I treat the alga component and the fungus component as some entities, although it would probably be better to consider particular cells as individuals in the case of alga). The answer may be correlated with circumstances, which leads to the claim that this is one species in the natural environment, while two species – in the laboratory environment. It is also possible to recognise only natural conditions (whatever this could mean) as representative and relevant. Let us, however, consider a case, where both partners can exist separately in nature, but then their development is slightly slower than in co-existence, and as a rule they loose the "fight for existence" with other organisms. Again – how many individuals (or, if somebody prefers, species) do we have here?

The intensity of symbiosis, that is mutual dependence of individuals, can be graduated – from a very loose, facultative relations, giving only small profits, up to complete dependence. Mitochondria – cellular organelles of eukaryotic organisms responsible for oxidative phosphorylation can be quoted as an example of the latter.

Mitochondria are almost certainly descendants of once free-living bacteria which entered into symbiosis with ancestors of unicellular eukaryotic organisms, perhaps similar to modern amoebas. We can imagine two ways to this co-existence, crowned with a great evolutionary success. Firstly, ur-mitochondria could have been initially parasites of ur-eukarionts, but later lost their pathogenic character, which allowed them to gain decidedly greater profits from peaceful co-existence. It is also possible that ur-eukarionts hunted for bacteria, digesting them within their cytoplasm. An accidental mutation which defected a digestive enzyme could have cleared the way for symbiosis. Whatever happened, it resulted in co-existence of two previously independent cybernetic individuals (subjects of evolution), which now adopts so extreme form that particular partners not only cannot exist separately, but also the majority of mitochondrial proteins is presently encoded in the cellular nucleus, where most genes of former bacteria have been transferred. Only residual amounts of DNA have remained in mitochondria until today. It still encodes a few proteins and possesses its own translatory apparatus. It is a relict of the "independent" past. It could seem therefore, that we may have no doubts – at least in such an extreme case – that the present eukaryotic organisms constitute single individuals.

However, this is not completely obvious either. Mitochondria do not originate in an eukaryotic cell from nothing. They are always a result of a division of parental mitochondria. These organelles retained their own genetic information (although it is residual) as well as the system responsible for its expression (much more similar to such a system in bacteria than in their own "hosts"). Therefore, mutations can occur in their DNA, as in any other DNA, and they may, for example, cause changes in the structure of the encoded proteins. If such a mutation leads to the occurrence in a given mitochondrium of a handicap of the function of its translatory apparatus (e.g. ribosomes), then such a defective mitochondium will reproduce worse and leave fewer descendent mitochondria than other competing mitochondria. A reverse situation takes place when a given mitochondrium gains advantage over other mitochondria due to accidental mutation (this is much more difficult to achieve). It is also possible for genomic conflicts to occur between mitochondria coming from different organisms (e.g. parents). Therefore, it is unquestionable that mitochondria still undergo some residual evolution, although only in the "environment" of eukaryotic cells, and therefore they are still individuals, although in a rudimentary sense. However, as the number of mitochondria is subject to strict control by a cell and a huge majority of the purposes of a single mitochondrium agree with the purposes of the whole organism, we can claim that, in this case, an almost complete evolutionary "fusion" of two individuals into one individual has been achieved[53].

The entire above-presented discussion was to show that the principle of the excluded middle is not obligatory in the alternative: to be or not to be an individual – it is possible to be an individual only partially. In the case of the phenomenon of symbiosis, there exists an entire spectrum of intermediate states between two independent subjects of evolution (evoluons) and one individual (the density of the network of negative feedbacks between partners in symbiosis can serve here as a measure). The cybernetic paradigm serves as a convenient reference point, although not always does it allow one to take a conclusive decision whether a given system is one evoluon or constitutes a complex of two (or more) evoluons. The problem consists in the possibility of partial interconnection between the networks of negative feedbacks of two systems. In other words, being an individual is not a discrete (0 or 1) fact – it can be graduated.

Let us, however, pay attention to the fact that such uncertainties appear only in the situations where we deal with the question **how many** individuals are present in a given system – one, two or more. I claim, however, that the cybernetic paradigm is able to offer an unequivocal answer to the question if in a given case we deal with any (at least one) living system (evoluon) **at all**, and therefore, if there is a subordination of a network of negative feedbacks to the superior positive feedback. Thus, the cybernetic criterion retains its validity in relation to objects of doubtful status, existing on other planets, and also in relation to the question which moment in the process of spontaneous origination of life on the Earth can be indicated as the beginning of the existence of life.

Finally, if everything said above is referred to human societies, the problem arises what to do with people who do not want to or cannot reproduce for cultural (celibacy) or biological (inborn genetic defects, operation, accident) reasons. Are they not individuals? This seems

[53]It would be very interesting to discover parasitic mitochondria which, instead of producing energy for a cell, take care only of their own reproduction at the cost of 'loyal' mitochondria (and the whole cell). However, they would have to do this with moderation, not to eliminate the normal mitochondria that keep the cell alive.

completely nonsensical. However, they are not individuals from the biological point of view, that is cybernetic individuals (evoluons). On the other hand, contrary to other animals, man has also other superior purposes, apart from the expansion of his genes, namely "expansion" of his convictions, beliefs, moral opinions, knowledge about the world, emotions[54], in one word: expansion of his psyche in the entire conceptual network of culture (I will define the conceptual network in the last part of this book). All of this constitutes, in my opinion, the essence of man as a **cultural individual**. The word "expansion" should not be identified here with rapacity, or understood pejoratively. After all, a priest teaching about the love of one's neighbour also works for the expansion (dissemination) of his opinions within culture. Each act of participation in the life of a society is a projection of one's own personality into the system of cultural meanings[55]. This proves only that being an individual in the psychical-cultural sense is (happens) to be more important for man than the realisation of himself as an individual in a purely biological-evolutionary sense. It should be pointed out at this occasion that these purposes to a large extent agree with each other. After all, parents transfer to their children not only their genes, but usually their convictions, ethical rules, etc. as well. On the other hand, the decision to have a smaller number of children that can be better educated and prepared for life (and not only to survival) testifies already to a significant contribution of cultural purposes.

The existence of a cultural (and psychical) individual, with its own purposes and mechanisms of their realisation which are not reducible to biological purposes, has been overlooked by socio-biology. Socio-biology has shown us how incredibly far the biological (especially evolutionary) conditioning of our nature reaches into our behaviour, motivations, way of thinking and structure of our societies. However, the success of the socio-biological programme of studies in the explanation of many aspects of human societies prompted some people to postulate that our **every** act can be satisfactorily explained on the basis of biological mechanisms. This is, in my opinion, an obvious absurd. Analogously, biophysics and biochemistry show how deeply the properties of living organisms are conditioned by their physicochemical building material. However, this does not contradict the claim that "something new" emerges when we pass from the physical level to the biological level. Similarly, the evolutionary source of many aspects

[54]That is mems according to the Dawkins' terminology.

[55]Only few personalities have a possibility of 'expansion' in a greater scale (as, for example, a personality of a famous writer), because the environment is already saturated with great amount of cultural contents and ideas.

of human behaviour does not imply that each phenomenon at the psychical and social level can be explained within the framework of biological terminology. What is crucial for the essence of living organisms is the axis: adaptation to an environment – lack of this adaptation, which is absent at the physical level. Similarly, the axis: pleasure – pleasantneess, which is absent on the biological level, constitutes the main motivation of human acts at the psychical (as well as social and cultural) level. Simply speaking, people frequently make things which serve their pleasure, but work against their reproductive success (for example, by applying contraceptives). In my opinion, socio-biology is not able to say anything heuristically fruitful about such phenomena.

Socio-biology has two more limitations. Firstly, a lot of (most) cultural contents is transferred through the cultural channel (from generation to generation, but also horizontally, within one generation) and not through the biological channel (genetically). Secondly, even many of the behaviour-related properties conditioned genetically (e.g. short sight or sexual coldness) are conditioned by neutral or harmful mutations (from the biological point of view), compensated by different cultural mechanisms and products. Summing up shortly, although socio-biology helps very much in understanding numerous aspects of human behaviour, its extreme form unnecessarily ignores (or even denies) the existence of the psychical-social-cultural level of our reality. I will develop this topic in more detail in the third part of this book, devoted to the evolution of (self)consciousness.

THE ORIGIN OF LIFE

At present, living organisms are characterised by an enormous variety of forms, as well as by a huge degree of complication of structure and function. There is surely no need to prove that an amoeba, an oak and a bat differ significantly. On the other hand, even simple cells of prokaryotic organisms (bacteria, cyanobacteria) possess a highly organised scheme of structure and complex metabolism. The extremely simplified forms, standing on the border of life, such as viroids and viruses, cannot reproduce outside a cell of a host and constitute living forms only as a part of its metabolism which is redirected by the parasite. Therefore, we still maintain that even the simplest living entity, able to exist independently in an inanimate environment is very complex now.

These remarks concerning variability and complexity of life are not formulated without a certain purpose. One could imagine that it is possible to find out the essence of life simply by extracting the properties common for all living organisms. The great diversity of these organisms at the macroscopic level suggests that one should look for similarities at the biochemical level, all the more so, because e.g. bacteria in fact do not reach above this level. And here, a real surprise is waiting for us – the fundamental principles of the functioning of metabolism, together with copying and reading of genetic information, are practically **identical** in all organisms known to us! This identity concerns the carrier of the genetic information (DNA), fundamental elements of the machinery of a cell (enzymes built of proteins), universal carrier of energy (ATP), the barrier surrounding a cell (protein-lipid membrane). The manner of copying DNA, the highly complex (containing mRNA, tRNA, aminoacyl-tRNA synthases and composed of two sub-units ribosomes) apparatus for reading genetic information (synthesis of proteins), finally, the genetic code, establishing the correspondence between the sequence of aminoacids in proteins and the sequence of nucleotides in DNA (and mRNA) – all are the same. A high degree of similarity of many metabolic pathways, including the enzymes occurring in them as well as converted metabolites – similarity, for example, between the pathways responsible for energy production, including the chain of electron transport connected with protein-lipid membranes – also testifies to high uniformity of the existing living organisms. I will not continue this enumeration – the above examples should suffice to support the thesis that claims identity for the basic principles of functioning of all forms of life known at present on the Earth.

As we can see, the problem in finding the common denominator for all life forms on the Earth consists in the fact, that this denominator is so vast. What seems to be the most fundamental scheme of living organisms turns out to exhibit astonishingly great degree of complexity. By the way, this testifies to the origination of all presently living organisms from only one ancestor, a quite highly organised one. It is logically necessary that it possessed **at least** this set of properties which characterises today all living organisms (very few exceptions, such as small deviations from the universality of the genetic code, can be neglected as secondary or unimportant). The extraction of common properties from all presently living organisms does not seem therefore to be particularly fruitful for reaching the heart of the essence of life. As it is doubtful if such a high degree of complexity is utterly **necessary** for life in general, or else it is possible to imagine much simpler living forms, fully functional and able to develop further in spite of their simplicity. If so, what are the decisive criteria that a given system must meet to be a living entity? In the previous chapter, I treated this problem from the point of view of the cybernetic paradigm. In spite of all its heuristic fertility, it is undoubtedly an abstract formulation, and therefore slightly detached from the real material world. For it seems an important thing not only what life is from the cybernetic point of view, but also how general cybernetic instructions are realised in the simplest possible forms of life on the Earth.

The above question leads us inevitably to the problem of the origin of life on our planet. It was certainly the only moment when life, emerging from inanimate matter, adopted as simple forms as possible. If we were able to find out **what** in those simplest forms decides whether they are animate or not, we would get closer to understanding the sense and essence of life. In such considerations, not the least important role is played by the curiosity aiming at finding out **what** initial stages life passed (or could pass) trough to achieve the present level of complication and diversity. The roots of the "success" of the process of evolution should be looked for mainly in the predispositions to further development which characterised (accidentally or by necessity) the first organisms on our planet.

In some circles, there still prevails the opinion that life is a too complex phenomenon to be able to originate spontaneously. Even the simplest prokaryotic cells living presently are enormously complex objects, composed of thousands of proteins and other compounds, possessing genetic information in the form of the DNA double helix which contains millions of nucleotide pairs, equipped with a complex and appropriately organised metabolism, ensuring transformation of matter and energy, finally – possessing a spatial structure. A spontaneous origination of such a highly organised system from inanimate matter seems to be completely improbable. Its probability is comparable to writing of *Hamlet* by a chimpanzee hitting the keys of a typewriter at random, or assembling of a ready-to-fly Boeing jet by a wind blowing over a scrap metal store (an example quoted by Hoyle). This argument is still quite frequently used against the possibility of a spontaneous origination of life.

However, the above argument should be considered as an entirely pointless. To explain this, I will use the example of a lottery game called Toto-Lotek[56], where one wins the main prize if he manages to guess six out of forty nine possible numbers. The probability of indicating in one "shot" the right six accidental numbers selected from forty nine numbers is equal to about 1 : 14000000, that is one to fourteen million. As it is necessary to guess right any of the six winning numbers out of forty nine numbers, and **at the same time** one of the remaining five out of 48, one of the remaining four out of 47 and so on, so that all six numbers are indicated right in one shot. The probability of such an event is equal to $6/49 \times 5/48 \times 4/47 \times 3/46 \times 2/45 \times 1/44$, which gives approximately just the above-mentioned one to fourteen millions. Therefore, one would have to "shoot" about seven million times to have 50 % chance for the main prize. Such a slender chance discourages many people from playing Toto-Lotek. It should be added at this point that the chance of a "lucky guess" in the "lottery of life", that is e.g. a spontaneous origination at one moment of a bacterium cell form inanimate matter (mixture of organic compounds) is yet far smaller, in fact so astronomically small that it is not worth being taken into consideration at all. Therefore, was life really not able to originate spontaneously?

The error in the above reasoning consists in the assumption that the process of origination of life resembled the strategy of playing Toto-Lotek. However, this was not the case – the whole present chapter is devoted to showing this point. Let us consider another strategy. Let us assume that if at a given shot we guess right at least one number out of the six, then we are informed about this and we leave it as good in the next shots, trying to guess the remaining numbers. After some time, we have already guessed two, three or four numbers, until we reach finally all six numbers. How many shots would we need on average if playing in this manner? The chance of guessing right of one of the six numbers from forty nine numbers is equal to c.a. 1 : 8, and therefore we have the chance of 50 % to guess right any number in c.a. 4 shots. When we have

[56]It is a lottery game popular in Poland.

already learned the first number, there remain five numbers out of forty eight (therefore now we cross out 5 numbers from 48). The probability of guessing the second number right amounts approximately to 1 : 10, and therefore we need on average 5 shots to perform this guessing right with the chance of 50 %. Analogously, guessing the subsequent numbers right requires respectively about 6, 8, 11 and 22 shots (this can be easily checked). With the probability 50 %, we will be successful to guess right all six numbers after about 56 shots (4 + 5 + 6 + 8 + 11 + 22). This is an incredibly small number in comparison with seven million (7000000) shots needed for guessing all six numbers right if the previous strategy is followed. Therefore, from among the two described strategies, the second one is much more effective. Moreover, life originated and evolved following exactly this method.

In other words, life could originate because it originated **step by step**. On each stage, it underwent various accidental changes – "shots" – and the external conditions (environment) decided which of those shots went home, that is which accidental changes appeared to be profitable for a given primitive organism, ensuring its survival and possibly production of similar systems. This resulted in a gradual accumulation of advantageous changes, that is in remembering the numbers already guessed correctly. The structure and function of even the simplest cell did not need to originate at once. On the contrary – it was being formed gradually during millions of years of evolution by the accumulation of many profitable changes, each of which improved this cell and frequently increased its degree of complication. It comes from the above, that life simply does not play Toto-Lotek.

Anyway, the process of origination is this aspect of the phenomenon of life which still remains the least known. However, a great majority of scientists shares the opinion that life was born spontaneously on the Earth almost 4 billions years ago, that is slightly more than a half billion years after the formation of the planet. The Earth formed about 4.5 billions years ago from dust and gas thrown into the cosmic space during the explosion of a Supernova star (our planet owes its composition to this star, for all chemical elements in the Universe, apart from hydrogen and helium, are formed in the interiors of stars). The force of gravitation brought about a concentration of dust particles and rocky pieces around the common center of gravity. As the mass of rock and dust shrank and thickened, the pressure and temperature prevailing in it increased, which led to melting of its building material. At this stage, the Earth resembled a sphere of melted magma. Heavier elements, especially iron, had a tendency to gather in its middle, which led to the formation of the metallic nucleus, while lighter elements, e.g. silicon, formed the outer layers of our planet. Its gravitational attraction gathered small rocky pieces from the neighbouring space, which caused both an increase of the mass of the planet, and further growth of temperature on its surface, due to unceasing collisions with those pieces of rock. However, the terrestrial globe started finally to cool down, radiating heat energy out into the Cosmos. Its surface changed the state from liquid to solid, which is equivalent to the formation of the terrestrial crust. During all this time, the melted surface of our planet and many volcanoes, once the terrestrial crust was formed, released various gases which formed a primitive atmosphere. After further cooling of the Earth's surface, one of the components of this atmosphere, namely steam, began to undergo condensation which is how oceans originated. At this moment, a proper ground for the origination of life had been already prepared.

At the moment of origination of life, the conditions on the terrestrial globe were different than these prevailing presently. The oceans exhibited a lower level of salinity. The atmosphere – which had a moderately reductive character – contained carbon dioxide, nitrogen, steam, hydrogen, ammonium and methane. On the other hand, there was no oxygen and, as a consequence, no ozone layer. The ultraviolet radiation lethal for life, stopped by this layer today, could reach the surface of our planet without any obstacles. The volcanic and tectonic activity (e.g. earthquakes) had a much greater intensity; atmospheric discharges were probably more frequent as well. And of course the terrestrial biosphere did not exist yet.

It seems paradoxical that those conditions, decidedly unfavourable for the present life, enabled at that moment the origination of this life. First of all, the gases entering into the composition of the atmosphere constituted the substrate for an abiogenetic (not caused by living organisms) synthesis of simple organic compounds. Generally speaking, organic compounds can be defined as highly reduced (containing hydrogen) compounds of carbon, that can also contain oxygen, nitrogen, phosphorus, sulphur as well as other elements – in smaller amounts. In the present atmosphere, organic compounds could not originate since it has a strongly oxidative character (a lot of oxygen accompanied by the lack of hydrogen and its compounds, if we neglect certain amount of steam). On the other hand, the primeval atmosphere in the initial period of the existence of the Earth contained carbon and nitrogen in a reduced form (as methane and ammonium), while oxygen was bound with carbon in carbon dioxide. Therefore, it had all the most important elements necessary for the origination of organic compounds – the basis of life – as well as appropriate, moderately reductive properties.

To enter into chemical reactions leading to origination of simple chemical compounds, the gases entering into the composition of the atmosphere at that time needed some source of energy that could be transformed into the energy of appropriate chemical bounds. Most probably, this energy was supplied by ultraviolet radiation and atmospheric discharges, and possibly by other sources such as explosions of volcanoes or decay of radioactive isotopes. The organic compounds, that originated as a result of these factors, fell down into oceans and underwent dilution in water. This solution, probably containing at least some aminoacids and nucleotides, simple sugars, hydrocarbons as well as a couple of other compounds, was named in various picturesque ways, e.g. as a "primeval broth" or as the "Urey's soup". The above mentioned monomers underwent spontaneous polymerisation, forming e.g. short polypeptide (protein) chains as well as threads of nucleic acids. The concentration of organic compounds in the seas at the time was most probably rather small. However, presumably, they could frequently undergo significant condensation, for example by a periodic evaporation of small water pools or by adsorption on different minerals. This enabled organic compounds to undergo condensation to the extent allowing them to enter into mutual interactions. Some of these interactions could lead to origination of structures of a certain degree of organisation. Organic compounds were relatively durable at that period, because there were lacking the two main factors leading today to their quick decomposition, namely oxygen and living organisms.

Summing up, it can be said that what is lethal to life today, enabled the origination of this life at the early stages of our planet's history. The factors enabling the origination of life were, first of all, the atmosphere containing methane and ammonium, lack of oxygen and the ozone layer protecting against UV radiation, great volcanic activity and at last also lack of the life itself. This explains why life does not originate spontaneously today – for the same reasons (lack of appropriate conditions) for which we could not have lived on the Earth over three billions years ago. Therefore, what was favourable for life at the very beginning and promoted its origination, does not have to be favourable for life at the present time. Lack of oxygen, ammonium and ultraviolet radiation in full abundance – indeed, going out to take a breath of "fresh air" must have meant something completely different for our furthest ancestors than it does for us.

Until now, everything looks rather clear and simple (although slightly too general). We may announce statements similar to those above without an excessive risk of exposing ourselves to the objection that we in fact do not quite know what we are talking about. However, a fundamental question arises: what comes next?

Here, honestly speaking, we are doomed to speculations. We may never learn the details of the process of origination of life on the Earth. However, this does not mean that nothing can be said on this subject. For, knowing what life consists in its mature phase, we can, *ex post*, infer some properties which it should have possessed at its dawn. However, one should allow for the margin of error always involved in this kind of extrapolations, and treat them with proper caution.

Having made all the necessary reservations, I can at last risk formulating the thesis that (in my opinion) the origination of life consisted in a more or less simultaneous (in the geological scale of time, i.e. in the order of thousands and millions of years) formation of its three aspects: **genetic** aspect, **structural** aspect and **energetic** aspect[57].

Perhaps, the genetic aspect associates most with the above-proposed cybernetic definition of life as a complex of negative feedbacks, at the service of a positive feedback, and therefore I will discuss this aspect first. However, before looking back, I will remind shortly what it means – in the presently existing representatives of the living world – when we speak about recording genetic information, copying and reading it, that is "translation" of such a record into the structure and function of an organism.

In a huge majority of organisms, the genetic record takes the form of a sequence of nucleotides in a double helix of DNA (deoxyribonucleic acid), composed of two threads wrapped around each other. Only in viroids and some viruses the carrier of genetic information takes the form of a single thread of DNA or RNA (ribonucleic acid). Four kinds of nucleotides enter into the composition of DNA. They differ by their nitric bases. These bases are: adenine, thymine, cytosine and guanine (in shortening A, T, C, G). Apart from one of the above bases, all nucleotides contain a sugar deoxyribose as well as a phosphate group. In RNA, T is replaced by U (uracil) similar to it, while deoxyribose is replaced with ribose. In the double helix of DNA, a

[57]The aspects I distinguish, as well as their number, are of course a little bit arbitrary, since one could prefer to emphasize other aspects, e.g. enzymatic aspect. Additionally, all aspects of life are not separate and independent – they are strictly related to each other and partially overlapping. Indeed, different scientists distinguish different kinds and numbers of such aspects. These

given nucleotide (concretely – its nitric base) of one thread must spatially correspond to an appropriate nucleotide in the other thread. Such spatially matched pairs are created by: A with T and C with G. Therefore, the sequence ATGCT on one thread will correspond the sequence TACGA on the other thread. In this sense, each of the threads in the double helix of DNA constitutes, in a sense, a negative of the other thread.

This manner of recording genetic information suggests automatically the mechanism of its copying (Watson and Crick have been awarded the Nobel Prize for proposing the double-helix model). It suffices to separate both threads and apply the principle of spatial matching of nucleotides to provide each of them with its negative. Such a negative can be created by adding to the original thread a new thread, formed of spatially-matched nucleotides, complementary with respect to the nucleotides in the old thread. The enzyme responsible for this process in the currently-existing organisms is called DNA polymerase. As a result of this operation, two helixes are formed, identical (if the process of copying was accurate) with the original helix. This is exactly the way in which living organisms act when they transfer genetic material of one dividing cell to two descendent cells.

It is not enough for an organism to copy genetic information. It is also necessary for it to read the data. The reading of genetic information can be reduced, in principle, to production of certain proteins in an appropriate amount, place and time sequence. For proteins – as enzymes and carriers (molecules transferring different substances through cellular membranes) – constitute the whole machinery of a living cell, responsible for the synthesis of proteins themselves, of other building compounds (sugars, lipids), for formation of cellular structures, production of energy and connected with this driving of various cellular processes, and at last – for copying genetic information. Other proteins can also serve as a building material (structural proteins), function as regulators of the process of reading genetic information, take the function of hormones or digestive enzymes. DNA therefore determines the structure and function of an organism through different proteins, for the production of which it is the matrix.

Continuous immediate reading of the original copy of genetic information could lead to its various damages (it is enough to recall what happens with a too frequently used map). Therefore, it is purposeful to keep such a plan away from the rest of metabolism, best – behind the hood of

---

different propositions are not mutually exclusive: they just deal with the problem from slightly different angles, the choice of which depends on personal preferences of a given author.

the nuclear membrane, and to use formerly prepared secondary copies for protein synthesis. An additional advantage of such a solution is constituted by the possibility of controlling the number of such copies depending of the needs of a cell. Apart from this, there is no need to copy the whole DNA at once – it suffices to multiply these segments (called genes) which are responsible for encoding the structure of determined proteins (one gene encodes one protein), that are needed at a given moment.

The role of such additional copies is realised by a certain sort of RNA called mRNA (messenger RNA). The principle of copying sequences of nucleotides from DNA onto RNA (this process is called transcription) is the same as the principle of copying (replication) of DNA – each molecule of mRNA constitutes a negative of some segment of one of DNA threads (before the synthesis of a given mRNA, a segment of a double helix of DNA must of course be locally unwound, separated into two DNA threads, so that one of them could serve as a matrix). One difference –a purely formal one anyway – is that U replaces here T in the formation of a pair with A. mRNA, contrarily to DNA, occurs in the one-thread form.

As I mentioned, both DNA and RNA are built of four nucleotides arranged in various sequences. On the other hand, proteins (synthesised according to the mRNA matrix) are composed of 20 aminoacids, also arranged in a different sequence and forming one chain. The sequence of aminoacids in a protein chain determines exclusively and unequivocally (with few exceptions) the final structure and function of a protein[58]. And just this sequence is recorded in the sequence of nucleotides in DNA (and, of course, in its working copy, that is mRNA).

For objective reasons (there are only 4 nucleotides, while there are as many as 20 aminoacids) it cannot be that one aminoacid corresponds to one nucleotide in the record of the sequence of aminoacids by the sequence of nucleotides. In reality, a sequence of exactly **three** nucleotides corresponds to each aminoacid. The so-called **genetic code** determines which triplets of nucleotides (**codons**) correspond to particular aminoacids. For example, the sequence (codon) GCU corresponds to the aminoacid alanine, while AAG – to lysine (the sequences of triplets in the genetic code correspond to sequences in mRNA – appropriate triplets in the maternal thread of DNA are of course complementary). Because the number of combinations of four kinds of nucleotides arranged in a three-element sequence (64 different triplets) is greater than the number

[58]Different aminoacids are characterised by different spatial structure and different chemical properties.

of aminoacids (20), several triplets (codons) correspond to most aminoacids. Additionally, three triplets (UAA, UAG, UGA) do not encode any aminoacid, but constitute the signal that reading of a given fragment of mRNA is terminated, and thus indicate where is the end of the record of the sequence of aminoacids in the protein encoded by the gene the copy of which is a given segment of mRNA[59].

However, there arises the question how the correspondence between different aminoacids and the appropriate triplets is set up. During the process of copying DNA into DNA or into RNA, the reason was purely physicochemical – the spatial matching of pairs of nucleotides (nitric bases). However, there does not exist any spatial correspondence, any natural matching between the structure of aminoacids, on one hand, and the nucleotide triplets, on the other hand. Therefore, living organisms had to create a substitute of such adequacy. To join two pipes of a different diameter (one of them can have at the same time a round section (profile), while the other – a square section), one can use special couplers with two different ends, with their shape and size corresponding to the sections of the pipes being joined. Living organisms act similarly, but they use **two** couplers to join a given triplet (codon) with the corresponding aminoacid. These are: a protein (enzyme) called the aminoacyl-tRNA synthase (I will call it further the aa-tRNA synthase) and the second kind of RNA, namely tRNA (transporter RNA), having the shape of a relatively short chain, winded to form a clover leaf. Different kinds of both aa-tRNA synthase and tRNA occur in the same number as the triplets encoding aminoacids[60]. A given aa-tRNA synthase recognises the appropriate aminoacid with one of its ends, while one of the ends of a proper tRNA molecule – with its other end (I use here the word "end" in the functional sense and not in the structural sense – in a protein, such an end can be in fact a catalytic pocket, and in tRNA the functional ends not necessarily overlap with the ends of the chain). The second end of tRNA, in its middle loop (that is in the middle clover leaf), possesses a triplet of nucleotides (the so-called **anticodon**) complementary to (that is being a negative of) the triplet encoding a given aminoacid, that is a codon (it is perhaps not necessary to add that each of the kinds of aa-tRNA synthase and tRNA has a slightly different spatial structure of each of their ends, which is responsible for their specificity). Therefore, the (informational) joint between the sequence of triplets in DNA and the sequence of aminoacids in proteins undergoes closing (completing) at this point. The whole

[59]These are the so-called nonsensical triplets or codons.
[60]In reality, there are less of them due to the so-called Crick's tolerance principle.

succession takes therefore the following form: a triplet in DNA – a triplet in mRNA – end 2 (a triplet) in tRNA – end 1 in tRNA – end 2 in aa-tRNA synthase – end 1 in aa-tRNA synthase – aminoacid[61]. This succession is presented in Fig. 2.12. The whole process seems undoubtedly to be rather complex.

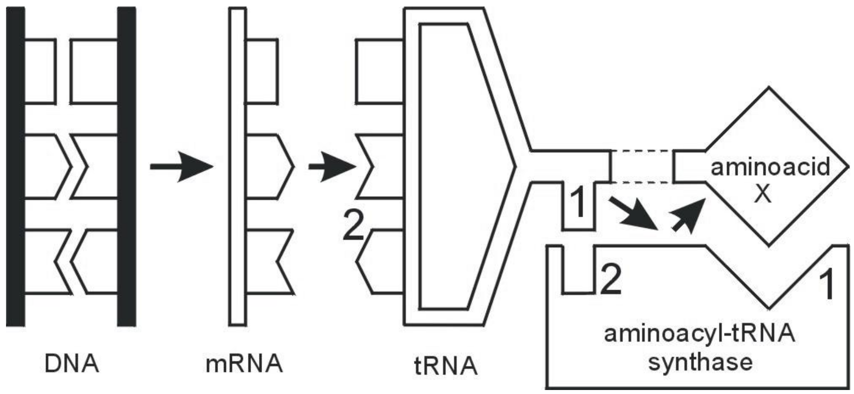

Fig. 2.12. "Chain" of transfer of genetic information. The sequence of nucleotide triplets in DNA is "translated", via the codon triplets in mRNA, anti-codon triplets in tRNA and appropriate aminoacyl-tRNA synthases, into the sequence of aminoacids in a protein. 1, 2, "ends" 1 and 2 of tRNA and aa-tRNA synthase.

Let us see now how the above-described encoding system behaves in operation. When a given gene (segment of DNA) receives (through the mediation of some regulatory protein) a signal about a demand for the protein encoded by the gene, the gene is transcribed (rewritten) onto a segment of complementary mRNA (the enzyme called RNA polymerase is responsible for this). This process occurs in the cellular nucleus. On the other hand, molecules of tRNA with proper aminoacids attached to their end 1 (the process of a selective joining of aminoacids to appropriate tRNAs is catalysed by adequate aa-tRNA synthases) are already waiting in the cytoplasm. The synthesis of proteins takes place in large protein complexes, ribosomes, containing the third kind of RNA, called rRNA. A thread of mRNA is being moved through the centre of a ribosome (or, what is equivalent, a ribosome moves along this thread). Subsequent triplets of nucleotides (codons) recorded in this thread are complemented with subsequent complexes tRNA-aminoacid, which is done through the mediation of the complementary triplets (anticodons) on end 2 of tRNA.

[61]This entire succession never exists as a physical object at a given moment of time – the subsequent stages of the transfer of genetic information follow one another in time. In this sense,

Aminoacids "belonging" to the subsequent tRNAs become sequentially joined in a chain (while disconnected tRNA molecules are released and can bind next aminoacid molecules) – the sequence of aminoacids in this chain corresponds therefore to the sequence of tRNA molecules with triplets (anticodons) complementary to the triplets (codons) in mRNA. Once the process reaches a nonsensical codon, that does not correspond to any tRNA, protein synthesis is terminated and a ready chain of connected aminoacids is released.

Although significantly complicated, the above description is in fact very simplified. In reality, each of the discussed stages is composed of a succession of sub-stages, catalysed by complexes of various enzymes. Moreover, it is subject to a variety of regulative processes, aimed at linking all particular processes in one sensible functional whole, serving biological interest of a living organism. A huge amount of regulatory proteins participates in this, and the final effect of their operation is a structure and function of a given organism, from a bacterium cell to such a complex object as human body.

In order to consider the origination of life on the Earth, I propose to reverse the arrow of time and look at the process of gradual simplification of subsequent elements (aspects) of living organisms, i.e. the process of their diminishing specialisation and differentiation in the reversed course of evolution, which leads us to the point where life "dissolves" into inanimate matter. Such a sequence of transformations, which I will call anti-evolution, could not occur in the direction consistent with the normal flow of time (and not against its flow) for thermodynamic reasons, but also because the logical structure of fundamental mechanisms of evolution, such as the natural selection, makes them asymmetrical with respect to time. The antievolution should therefore be treated as a thought experiment, without attributing any real existence to the process. With this reservation in mind, we will try to investigate how anti-evolution could realise its task, that is reduction of the complexity of modern living organisms, their transformation into a suspension of organic compounds in the "primeval soup" in such a way that all intermediate stages would be imaginable and fully functional.

Returning to the genetic aspect of the origination of life on the Earth, one should ask if the whole above-described apparatus of copying and reading of genetic information could originate with all its complexity spontaneously, at one moment, from an accidental mixture of aminoacids and nucleotides as well as of other simple organic compounds. The question is of course rhetorical

---

the analogy with the succession of couplers between two pipes cannot be treated too strictly.

and the answer – negative. However, it provokes a subsequent question: if (and how) the system of transferring and reading genetic information was able to form in stages, step by step. The problem is far from trivial, since it concerns a functionality or simply a possibility of existence of intermediate (transit) stages.

Let us assume that we are building a plane from a car (to make the analogy exact, we have to do this while driving!), by adding subsequent parts and waiting until it takes off. Can a plane with one wing be an intermediate stage in our endeavour? Of course, not! We will not only be unable to take off, but we will crash against the first tree at the side of the road. Therefore, to be successful in our undertaking, we must add both wings and tail (and also a propeller, appropriate engine and a lot of other devices, e.g. a steering system) at the same time, in one step. This means a sudden, jump-like increase of the complexity of our device. An adequate question in the context of the origination of life on our planet is therefore clear – if a formation of a highly complex system (now I am discussing the genetic aspect, but the same applies equally well to other aspects) at one moment seems improbable, then, is it possible to think of any **functional** transitional stages, that would allow to achieve the same effect in many steps? If not, then life could not originate spontaneously, on the basis of physicochemical laws known to us, and it would become necessary to find another explanation of its genesis[62].

In the case of the system of copying and reading of genetic information, intermediate stages would mean either a simplification of the system elements, or a reduction of the number of these elements, that is removal of some of them, and most probably, both these things at the same time. The possibility of simplifying particular elements does not probably rouse any special doubts. Anyway, I relieved in a sense the process of anti-evolution of this task, reducing significantly in my description the whole complexity of the functioning of the genetic system in modern organisms. Therefore, I will not deal with this problem any more. As I mentioned before, it does not seem to be possible for the described system, even in its simplest version, to originate spontaneously, by a pure chance and in one step. It remains to suppose that, in its dawn, the system of a copying and reading of genetic information contained less elements than it does at present.

[62]In fact, the argument claiming, on one hand, an impossibility of spontaneous origination of life in its full complexity, and, on the other hand, non-existence of imaginable intermediate stages, is still quite commonly used by opponents of a spontaneous biogenesis and the process of evolution in general.

I will remind shortly once more the discussed chain of transfer of genetic information: DNA – mRNA – tRNA – aa-tRNA synthase – protein. At the first glance, it seems that removing any of the elements (links) of the chain would lead to a complete loss of sensible integrity of the whole chain, to a disintegration of its function. Elimination of DNA would be equivalent to depriving the above system of a universal carrier of genetic information. The discussed "couplers" (tRNA, aa-tRNA synthase, and also mRNA) would not have in this situation anything to translate into the sequence of aminoacids in a protein. Lack of proteins would mean absence of enzymes, and, what follows, of the whole metabolism, being the basis of the functioning of living organisms. In reality, the production of proteins constitutes the purpose of the whole system, and therefore its absence leads automatically to the loss of any sense of the remaining rest. On the other hand, a removal of any of the couplers mediating in the reading of the genetic record would cause a break in the continuity of the stream of transfer of information from DNA to proteins. If tRNA is lacking, for example, the aa-tRNA synthase would lack an instrument for recognising the appropriate codon triplet on mRNA – the structure of its end 2 fits to end 1 of an adequate tRNA, and not to the "structure" of a codon. After all, it is enough to imagine that one of the couplers between pipes is removed (we assume that there are no two identical couplers as to the shape and diameter of their ends) to understand what the problem consists in. If such a coupler has two different ends (and this constitutes the essence of a coupler), then the free ends of the pipes this coupler is designed to join will not fit each another.

However, one can try to imagine possible manners of getting around the above problem (which does not mean that the obtained conclusions have to correspond exactly to what once actually took place). I will assume therefore that the genetic system could function even when some of its elements are removed. After all, our existence testifies to the fact that the process of anti-evolution would contrive some way of reducing this system. Now I will try to analyse two such possibilities that, in my opinion, actually took place during the origination of life.

Firstly: is DNA actually fully indispensable for recording genetic information? No, and this is not a speculation – viroids and some viruses use RNA for this purpose. Therefore, one could speculate that the first organisms could have used what we call today the mRNA for two purposes: as a matrix for protein synthesis and as a carrier of genetic information. Only later, living organisms learned to use DNA for the second purpose, for the reasons quoted a few paragraphs above: because such a record is characterised by greater durability and it provides the possibility of regulating gene expression at the level of transcription ("rewriting" sequences of nucleotides from DNA to mRNA). The until-now existing enzyme called the reverse transcriptase that catalyses rewriting of sequences of nucleotides in the reverse direction, that is from RNA to DNA, can be a "living fossil" from those immemorial times. We do not know if this really took place, but **in principle** it was not necessary for DNA to be a component of the first genetic systems. This shortens our chain by one link.

I would also like to speculate (this is my own proposition) on the aa-tRNA synthase. Today, its absence would make impossible the recognition and binding of adequate aminoacids by tRNA. However, this did not need to be always the case. tRNA has a relatively complex spatial structure, and once it could have possessed a fragment directly recognising adequate aminoacids by the principle of spatial matching. A progenitor of the aa-tRNA synthase, on the other hand (let us remember that we deal with an enzyme), could appear slightly later and initially occur only in one form, common to all kinds of tRNA. In such a case, its role would consist only in catalysing, speeding up the spontaneously occurring binding between a concrete aminoacid and a given tRNA, but not establishing the correspondence between appropriate kinds of these compounds. Only later, when different types of the aa-tRNA synthase originated and took over the function of a selective recognition of the corresponding molecules of aminoacids and tRNA, the tRNA itself could have lost this function. The fragment on tRNA corresponding structurally to a chosen aminoacid was removed by evolution as already unnecessary, while the fragment recognised selectively by a proper aa-tRNA synthase was exposed. Therefore, in my opinion, there is again no need to assume that the aa-tRNA synthase constituted a component of genetic systems from the very beginning of their existence – it is possible to imagine a functioning of these systems without its participation.

A reduction of the amount of basic elements of the genetic system of modern organisms is, therefore, admissible, at least in principle (from the logical point of view). However, we may never learn which way evolution used for this purpose (or which way anti-evolution would use). Nevertheless, the argument, wielded by some antagonists of evolution and spontaneous biogenesis, that it is highly improbable for such a complex object as life to originate spontaneously, seems to be invalid. Life simply did not have to start its existence from a stage as highly organised as we know at present.

After reducing the number of components of a primitive genetic system – by means of a specific kind of Occam's razor principle (do not multiply the elements necessary for the origination of life beyond necessity) – we are left with the following system: mRNA – tRNA – protein. Further simplification of this system, if it is still to remain a self-copying and evolving living system, seems to be impossible. We have here, in my opinion, the absolute minimum: the sequence of aminoacids in a protein, which decides about the functional usefulness of this protein; the record of this sequence in RNA (pre-mRNA), which can be transferred to descendent organisms; as well as the "key" to the genetic code, translating the second of these sequences to the first one, that is appearing in the number of kinds corresponding to the number of codon triplets – tRNA. This is complemented with one simple protein facilitating the copying of RNA and another protein joining aminoacids in a protein chain (in the simplest version, different tRNAs with added aminoacids could – in my opinion – attach themselves to a mRNA thread in the places of proper codons, and aminoacids were afterwards joined in a chain according to their spatial ordering). Such a system seems already to be able to form spontaneously, by physicochemical processes occurring for millions (in fact – hundreds of millions) years in a mixture of simple organic compounds dissolved in practically all waters on the Earth at that time. After all, at the beginning, this system could have been significantly simpler, that is, it could have contained less aminoacids than presently, and it could have possessed not triplets, but dyads of nucleotides encoding one aminoacid, separated by single nucleotides not participating in the coding process. This hypothesis is supported by the fact that, in the modern genetic code, an exchange of a nucleotide in the third place in a codon triplet frequently does not lead to a change of the aminoacid encoded. Because the amount of possible triplets where each position can be occupied by one of the four nucleotides (the number of combinations is equal to $4^3 = 64$) is much greater than the number of aminoacids (20), some of aminoacids are coded by two, four or even six triplets (codons)! Therefore, one could easily imagine that at the beginning, instead of 64 triplets, there were 16 (4^2) dyads (separated by single non-coding nucleotides), and the same or slightly smaller number of encoded aminoacids. Therefore, the genetic code did not have to originate immediately in its final form. Such a possibility could constitute one of possible explanations of some exceptions from the universality of the genetic code, that is the fact that in some cases (very few, anyway), the same triplets encode different aminoacids.

Anyway, the genetic system of the simplest living organisms was (in my opinion) most probably composed of proteins and nucleic acids (here belong DNA and RNA). Now, I will try to answer the question why both these compounds, and not for example only proteins or only nucleic acids, entered into the composition of the first self-copying systems.

I mentioned above that the superior purpose of life is to maintain the potential of expansion, being the result of a positive feedback. This is a manifestation of the fact that living organisms reproduce, that is multiply the copies of themselves. However, considering the origination of life, we deal with chemical systems which are not yet alive. The phenomenon of autocatalysis, (compare Fig. 2.2) in which some siple monomeric molecule catalyses the production of (copies of) itself, is an already mentioned example of a positive feedback in chemistry. Therefore, the term "autocatalysis" is frequently used while speaking about copying of simple systems composed of proteins and/or nucleic acids.

Why, therefore, there originated autocatalytic systems composed of proteins and nucleic acids and not only of one of these components? Proteins, to which all enzymes belong, are as a rule effective catalysts. Even artificial protein chains with an accidental sequence of aminoacids can frequently speed up significantly the rate of different reactions, although their specificity and effectiveness is of course much smaller than the specificity and effectiveness of "professional" enzymes. In the "primeval soup", there could have originated (and probably did so many times) accidentally, by random joining of different aminoacids, short protein chains exhibiting some catalytic activity in the direction of joining free aminoacids into short protein chains. Many of the newly formed chains were able to catalyse the origination of other such chains, those in turn... – we already know this scenario. The described auto-catalytic cycle (leading to the origination of copies of a molecule) is a typical example of the positive feedback. Could the cycle as such constitute the beginning of life?

Everything seems to indicate that it could not. Why? Because the described beginning of the evolution of such a cycle is at the same time the end of its evolution – this cycle will never be able to produce anything more than protein chins with an **accidental** sequence of aminoacids. Even if a certain protein chain – built as a result of random incorporation of such and not other aminoacids, in such and not another sequence – appears to be much better, more efficient (for example in the rate of catalysis) than other chains, it will not transfer this property to its "progeny", that is to the protein chains synthesised by it which will be composed of aminoacids arranged in a random sequence, just like the descendent chains of other protein chains. Therefore, proteins in an autocatalytic cycle do not have the possibility of transferring their identity to descendent molecules.

The matters take on a different aspect in the case of autocatalytic cycles of nucleic acids. As we remember, due to spatial matching of the nucleotides situated opposite one to another in the maternal thread and descendent thread, the latter is a negative of the former. A negative of a negative gives of course a positive, and therefore the progeny of the descendent thread will be identical with the primeval (original) thread. Using a more familiar terminology, a daughter in the world of threads of nucleic acids is complementary to a mother, while a grand daughter is identical with a grandmother. It results from the above that threads of nucleic acids can transfer their identity to descendent threads, with a faithfulness limited only by the accuracy of the copying process. In this respect, nucleic acids seem therefore to be an ideal material for evolution and origination of life.

However, at a closer look, the situation does not seem so good. Copying of nucleic acids by adding a complementary thread can occur spontaneously, without participation of enzymes, in a solution containing free nucleotides. Yet this process is slow, of low efficiency, and it allows only relatively short chains to be copied. Outside a laboratory where ideal conditions are created for such reactions, self-reproducing threads of nucleic acids could be destroyed by various physical and chemical factors. There would take place a kind of race between the production of new threads in an auto-catalytic cycle and the disintegration of the existing threads under the influence of environmental factors. Judging from what we know, nucleic acids would loose this race – their degradation would be faster than the origination of new threads[63]. An additional, very important factor is constituted by the limited accuracy with which chains of nucleic acids undergo spontaneous multiplication. This leads to the conclusion that only a small part of the "progeny" of even short chains will be identical with the maternal chain. In the case of longer chains, the chance for a sequence of nucleotides to be copied exactly is slender. It appears therefore that neither nucleic acids alone, nor proteins alone can transfer their **identity** in a sufficiently effective way to descendent molecules.

[63]Additionally, if the thread being copied were too long, the descendent thread could dissociate away from it before copying is terminated.

Therefore, in my opinion, auto-catalytic cycles of nucleic acids (as analogous protein cycles) are also unable to produce life and to undergo the process of evolution. On the other hand, such an ability is evinced by **hypercycles** (Eigen's term) of proteins and nucleic acids. In an auto-catalytic cycle, a protein catalyses a formation of a protein, and a nucleic acid catalyses a production of a nucleic acid. In a hypercycle, on the other hand, there takes place a specific co-operation: a protein catalyses also the copying of a thread of nucleic acid in which it is encoded itself, while the thread of nucleic acid serves also as the matrix for production of this protein! Therefore, this, probably the first case of "symbiosis" (please, treat this term as a metaphor) in the history of life on the Earth, yielded a system able of self-copying and evolution, and therefore of living. Nucleic acids contributed to this system the possibility of preserving and multiplying its "individual" identity, while proteins – effective executive mechanisms for the realisation of this purpose (catalytic abilities). Entire hypercycles gained the ability of effective copying and transferring their identity to descendent hypercycles.

While it can be accepted that the moment of the origin of life is identical with the appearance of the first hypercycle (which I will try to discuss in a moment within the framework of the cybernetic paradigm), this does not mean that nothing existed before. The recently discovered molecules of RNA with some catalytic properties (so-called **ribozymes**) have lead to the supposition that auto-catalytic cycles of RNA preceded hypercycles of proteins and nucleic acids in the history of life. This hypothesis has recently became very fashionable, as it frequently happens with hypotheses formulated on the basis of a freshly made scientific discovery. However, I think that this fashion will be gone soon. Firstly, the catalytic activity of ribozymes is rather ineffective and limited to a very narrow set of reactions. Secondly, it still remains a mystery how proteins become involved in the co-operation with RNA. However, if auto-catalytic cycles of RNA really preceded hypercycles of proteins and nucleic acids (which remains highly problematic), then they were successfully substituted by the latter, because proteins are much more effective catalysts than RNA.

One could equally well suggest (and I am personally disposed to prefer rather this point of view, proposing it as an alternative possibility) that the above-mentioned protein cycles constituted an earlier stage. Proteins with an accidental sequence of aminoacids produced in the cycles could certainly catalyse many various reactions, and not only (not even mainly) joining of free aminoacids, that is production of next protein chains with a random sequence. I will use the term "**polycatalon**"[64] to refer to such a multi-functional complex of proteins with an accidental sequence of aminoacids, catalysing many different reactions, which, in my opinion, constituted an important stage in the origination of life. Polymerisation of nucleotides present in a solution into threads of nucleic acids complementary to already existing threads could appear to be one of those "extra" catalytic functions. This would be an important intermediate step on the way towards formation of hypercycles. Once they are created, a polycatalon could also be the starting stage for any kinds of (advanced) catalytic activities, when the functions of particular proteins had undergone further improvement and specialisation in the course of evolution.

The scenarios mentioned above do not need necessarily to exclude each other. It is perhaps most reasonable to assume that auto-catalytic cycles of both proteins and nucleic acids existed independently, before they became joined in hypercycles of proteins and nucleic acids. I propose that primitive protein chains from the very beginning supported catalytically reproducing threads of RNA (before they began to be coded by these threads), coexisting with them within objects similar to coacervates (discussed below bubbles composed of organic compounds, able to grow and divide). RNA (and also free nucleotides) could gather for purely physical reasons (better dissolvability in coacervates than in water) in the interior of coacervates, where the environment is conducive to reproduction of RNA due to non-specific catalytic activity of proteins. In such a case, auto-catalytic cycles of proteins and nucleic acids coexisted independently within coacervates, until proteins started to be encoded by nucleic acids, i.e. until hypercycles were formed. I think that coacervates not only constituted the ancestor of the first cell in the structural aspect, but also have led to establishing a closer "acquaintance" between proteins and nucleic acids. The possibility of a purely physical role (as a spatially organising factor) of objects similar to coacervates in the formation of hypercycles seems for me to be underestimated. I will mention it during the discussion of the genesis of the structural aspect of life. Anyway, the origination of hypercycles did not need to happen at once, in one step, which supports again the thesis that life emerged gradually from inanimate matter.

---

[64]An experimental premise suggesting the possibility of a spontaneous formation of protein chaines with an accidental sequence of aminoacids and unspecific enzymatic activity is constituted by the so-called protenoids, obtained by Fox through the exposition of aminoacids to high temperatures. Dissolved in water, protenoids form aggregations called microspheres. Microspheres exhibit the ability to bud descendent microspheres and to carry out relatively varied enzymatic activity.

While discussing the origination of hypercycles, one should remember the necessity of existence in the genetic apparatus of not only the record of genetic information (nucleic acids) and the metabolic machinery (enzymes) coded by it, but also of the coupler between them, that is the tRNA discussed before. For the minimal possible and effective auto-catalytic system, able to inherit its identity must be composed of **three**, and not only **two** parts. The third part, that is the mentioned coupler, should contain (determine)) a kind of correspondence – established in a conventional (that is primarily accidental, arbitrary) manner – between the information in the genetic record and the information in the structure of "executive factors", that is, respectively – the sequence of nucleotides and the sequence of aminoacids. The attachment of different codon triplets to particular aminoacids, today irrevocably universal in the whole living world (with some, very few exceptions), was an accidental phenomenon at the moment of its formation, since there is no correspondence between the spatial structure of a given aminoacid and spatial structure of the codon triplet corresponding to it[65]. Therefore, only the continuity of existence of the information about the structure of particular kinds of tRNA (it is presently encoded in DNA) can ensure the "sensibility" of genetic information. Therefore, if the correspondence between different sequences of nucleotides in codon triplets and particular aminoacids is called the genetic code, then the mentioned coupler, tRNA, is the key to this code. Without this key, any "communication" between proteins and nucleic acids (potentially coding the proteins) would be impossible and they would have to exist as separate systems. The necessity of origination of this coupler between pre-mRNA and proteins makes even more probable the hypothesis proposed by me before a while, the one concerning the coexistence of auto-catalytic cycles of proteins and nucleic acids within a limited volume of the interior of coacervates (where favourable physicochemical conditions prevailed) **before** hypercycles came into being. Thus, it became possible for RNA and proteins (and also for nucleotides and aminoacids) to gather in an appropriately high concentration, while they were protected against destructive influences of the environment, and a sufficiently long time for

[65]Similarly, there is no reason for the name of a horse to be "a horse" (after all, this name is completely different in other languages), since this word does not reflect in any way, for somebody not knowing the English language, the shape or colour of a horse, or any other "objective" property. In this sense, both ethnic languages and genetic languages are conventional. However, while we know many ethnic languages, there is only one genetic language. This is not because it is the only possible genetic language, but since other genetic languages, which most probably originated in the times of the origin of life, simply lost the competition with the now-predominating genetic language and become extinct.

origination of hypercycles was assured. Coacervates composed of intensively reproducing proteins and nucleic acids (hypercycles allow such effective reproduction) grew faster and underwent divisions more frequently, which provided them with selective superiority over other coacervates in the rivalry for nucleotides and aminoacids dissolved in the "primeval soup". And these were already the germs of life and evolution.

One, incredibly fascinating conclusion results from the combination of the above-discussed arbitrariness and universality of the genetic code: all the organisms living presently on the Earth come from one ancestor. This should be understood in the most literal way: if we would take any individual of earthworm, ladybug and alga and followed the sequence of its ancestors, going back by millions and billions years, and would do the same with the sequence of our own ancestors (say, in the male line), then, these sequences would meet at a certain moment and go further in the past as one sequence. The common ancestor – completely dissimilar to either man or earthworm – would be at the meeting point of these two sequences. Of course, there would be different common ancestors for man and earthworm, on the one hand, and for man and alga, on the other hand – the evolutionary ways of man and alga diverged earlier, no doubt. The common ancestor of all living organisms was probably something significantly more complex than a mere hypercycle – the biochemical similarity of all living organisms testifies to this.

But where in fact does this conclusion concerning common descent come from? Maybe, in spite of all, the ancestors of to-day, say, bacteria, plants and animals, could originate independently from inanimate matter, while their biochemical similarity results from some fundamental physicochemical phenomena which gave birth to all these organisms? Well, such a possibility should be decidedly excluded. As I mentioned, the genetic code is universal, an therefore common (with very few small modifications) for all organisms on the terrestrial globe. It is also conventional ("arbitrary"), its form does not result from any physicochemical reasons or predispositions. It is contingent, resulting from pure chance. This "convention" is identical with the continuity of existence in time of (the information about the structure of) the key to this code, that is several tens of different molecules of tRNA (as well as of aa-tRNA synthase) or, in other words, the continuity of the correspondence between different aminoacids and proper codon triplets. The amount of possible combinations of the correspondence between different codons and different aminoacids, that is simply the number of possible genetic codes, is a truly astronomical number. It should be expected that all such codes (or at least most of them) would have a similar functional efficiency (similarly, it would not change the efficiency of the English language, if the notation of letter "a" were exchanged in it with the notation of letter "e"). The genetic code known to us is therefore in no way privileged in relation to all other possible genetic codes, which had the same chance to originate. On the other hand, the probability of a two-fold or multi-fold independent origination of exactly the same code by pure chance seems to be extremely low. This leads to the conclusion that the whole life descends from one ancestor that possessed the genetic code in its present form. This does not mean that different forms of life, having different genetic codes, could not originate many times in the earliest periods of the existence of the Earth. However, even if this happened, only one form has survived to the present times.

I have already mentioned above that, in my opinion, life emerged from inanimate matter at the moment of origination of first hypercycles. Now, I would like to show how this conclusion ensues from the cybernetic definition of the essence of life, presented in the previous subchapter. I remind that according to this definition, life is a system of negative feedbacks sustaining the superior positive feedback, expressing itself as the potential of expansion of the system identity. Why therefore do hypercycles **already** fulfil the above cybernetic criterion, while auto-catalytic cycles of proteins and nucleic acids do not meet it **yet**?

Firstly, as it has been already discussed, proteins or nucleic acids alone cannot inherit efficiently their identity, because they do not possess any regulatory mechanisms, that is negative feedbacks, allowing them to "optimise" the **main parameter**, which is identity. The situation is different in hypercycles: proteins maintain the identity of nucleic acids in subsequent generations by channelling their accurate copying, while nucleic acids maintain the identity of proteins by coding them, due to which the identity of the whole system remains preserved. Environmental influences that disturb the regulated parameter (identity) by causing errors in exact copying of RNA threads were compensated by **selection** and **elimination** (resulting from the former) of hypercycles with harmful mutations in RNA. In such hypercycles, RNA with a changed sequence of nucleotides coded a defective protein (or proteins), much less efficient in effective copying of nucleic acid threads and/or synthesis of itself on the basis of the sequence of nucleotides in RNA. The growth and reproduction of a coacervate with such hypercycles was slower than in the case of other, not defected coacervates, and therefore the former was loosing in competition with latter for free aminoacids and nucleotides present in the environment[66]. Expressing this in cybernetic terms, the steering loop of a negative feedback took the following form: deviation (harmful mutation) from the assigned identity (optimal in current conditions) → elimination of the defective identity by natural selection → restoration of the assigned (optimal) identity of the individuals (in this case coacervates) in a considered population. In this sense, the natural selection is, in my opinion, the historically first and remains the most important negative feedback in living organisms, serving the superior positive feedback – maintenance and propagation of the identity of individuals.

I also propose that the natural selection of hypercycles took place at the level of coacervates and not, as in the original Eigen's proposition, at the level of single hypercycles. Of course, as a result of a much less probable, profitable mutation (an accidental change in a complex system can much more easily spoil it than improve it), there would appear yet more effective hypercycle, with still higher potential of expansion of its own identity. Therefore, this **first negative feedback** optimising the survival and propagation of a given identity (the continuity of changes of this identity in time, to be strict) **was inseparably connected with natural selection, that is with evolution**[67]. The steering loop optimised (maintained) here the best identity among the available identities. As I mentioned before, not single hypercycles, but whole coacervates, able to contain several, mutually supporting hypercycles, were (in my opinion) the subject of this evolution. The protein of one of such hypercycles could, for example, catalyse more effective copying of a thread of RNA, while the protein of another hypercycle – the process of coding and synthesis of proteins. I think the coacervates containing such "super-hypercycles" (let me introduce one more term) certainly can be regarded as the first living organisms, for they were able to evolve further on. As a spontaneous origination of coacervates containing auto-catalytic cycles of proteins and nucleic

[66]These substances could come from the previously-discussed abiogenetic synthesis of organic compounds or from desintegration of some coacervates. In my opinion, the selective pressure concerned not only the rate of growth and divisions, but also the persistence of coacervates containing such and not other hypercycles.

[67]Only afterwards, there appeared "typical" negative feedbacks, acting within a given system and not between systems. For example, enzymes repairing damages in DNA as well as replacing improper nucleotides (not mateched with their counterparts in the complementary thread in DNA), inserted to a thread of DNA due to a mutation, with proper nucleotides (coupled with nucleotides in the opposite thread of DNA) – these enzymes helped to preserve the identity of the genetic record and therefore of the whole organism.

acids[68] (as a result of simple physicochemical processes) seems to be at least a probable process (and possibly an inevitable one in the conditions prevailing on the Earth 4 billions years ago), the origination of life required **only** for some germs of the genetic code to develop. Afterwards, everything went on smoothly – the originated systems gained the ability to reproduce (multiply their identity) and evolve, and therefore they gained the ability of gradual **auto-complication,** proceeding from generation to generation. The development of such complex organisms as e.g. a horse became already only a problem of time. Indeed, it is a wander that so **little** was needed for life to appear on our planet.

I think that together with the origination of life, there emerged in a sense a new, biological level of reality. The determinant distinguishing it from the physical level is the kind of purposes or senses constituting the biological level, defined as a certain (already characterised here many times) complex of cybernetic mechanisms. Three important properties, absent from inanimate physical objects, characterise biological cybernetic individuals, that is evoluons.

Firstly, the identity of such individuals is auto-directed on itself, on its survival and expansion. The system of negative feedbacks constituting this identity has the sole purpose of continuing its own existence, which is realised by production of systems of regulatory mechanisms similar to itself. Secondly, the "functional substance" of living organisms is constituted by a **relative network of interrelations** between particular negative feedbacks, where the sense, the meaning of a given feedback is determined by its context, i.e. other feedbacks it co-operates with. In the last instance, the aim of each negative feedback is related to the entire network of feedbacks (regulatory mechanisms)[69]. The conditioning of the meaning of a given feedback occurs therefore

---

[68]In the beginning, only proteins served catalytically nucleic acids, which did not repay (return) yet with a codying of proteins.

[69]For example, the regulation of the level of glucose in blood discussed above is not a purpose in itself, but constitutes the basis for other negative feedbacks. Glucose is one of the main respiratory substrates and therefore its presence in an appropriate amount enables the negative feedback connected with the oxidative production of ATP, aimed at the production of energy with the rate adjusted to the present demand for energy. The maintaince of a more or less constant concentration of glucose in blood would not have any sense if we do not take into account this task (among many others). There are many other functions of glucose and each of them is connected to many negative feedbacks, optimising the values of various parameters. The main purpose of all these feedbacks is replication of DNA and its transfer to offspring. If we heat up a lump of sugar in a physical environment, it will burn away at the greatest possible rate. On the other hand, living organisms optimise the rate of oxidation of sugars in order to maximise their reproductive success.

by way of **connotation** (co-defining). Thirdly, the discussed network of relations between negative feedbacks constitutes a certain "theory" on the surrounding world "worked out" by a living organism. This theory – that is a **picture of the world** imprinted into the network of regulatory mechanisms responsible for the response of an organism to different properties of this world – is not abstractive, but purely pragmatic in its essence. Therefore, it does not reach in any way the "essence" of the world, for it is concerned only with effective functioning in the world, ensuring the realisation of the superior purpose, that is survival and production of progeny. Such "theories" (corresponding to different identities) are formulated randomly (as a result of mutations), and current conditions of the environment segregate them into "adequate" and "erroneous".

We will meet the above properties, that is the auto-directing on itself, the network of relations meaning through connotation as well as the creation of an "instrumental" picture of the world, in the last part of this book, when I discuss the emergence of the psychic level (self-consciousness) from the biological level. It will appear that, from the formal point of view, the origin of psyche is strictly analogous to the origin of life and consists in the appearance of the properties enumerated above on another level, in a sense.

Summing up, in the above-presented hypothetical history of development of the genetic apparatus all the way back until the moment of the birth of life, the stress was put on the principle "not everything at the same time", due to which the origination of the present degree of complexity proceeded in stages (step by step). It is also important that each of these stages was fully functional and contained in itself a possibility of a transition to the next stage (that is to a higher level of organisation of a system). This implies that physical and chemical laws known to us are sufficient, at least in principle, to explain the origination of life and we do not have to resort to non-orthodox explanations. This does not contradict the statement that some new level of organisation, which cannot be satisfactorily interpreted in the terminology of physical laws, appeared together with the origination of life. For the appearance of a negative feedback is inseparably connected with the appearance of a certain purpose – as we know, this term does not mean anything in a purely physical description of the world. Therefore, the system of physical concepts is not sufficient to characterise completely the essence of life. It is also necessary to introduce at least cybernetic terms, describing mechanisms and systems directed on the realisation of a certain purpose (tendential systems).

The concept of purpose or sense seems to be totally absent at the level of physical phenomena. The behaviour of both atoms and heavenly bodies is governed by well-defined physical laws and theories, with quantum mechanics and general relativity in the forefront. Neither the diffraction and interference of an electron when it passes through a double slit, nor the revolutions of the planets around the Sun have any sense or purpose. This applies equally to the physicochemical processes, e.g. movement of atoms or formation of chemical bounds, occurring in the "animate matter". Therefore, how does it happen that, when we take all physicochemical processes constituting a given organism, gather them together and take them as a certain macroscopic entity, they suddenly acquire traits of a clear functionality, directed on a certain purpose? How can purposefulness emerge from aimlessness? The physical paradigm of description of the world cannot answer this question.

To be sure, thermodynamics can describe life as a non-equilibrium process, tending to a possibly effective dissipation of energy and adopting "for this purpose" the macroscopic form of a dissipative structure. This throws undoubtedly a new light on the phenomenon of life, but it certainly does not reach the very heart of its essence. Thermodynamic description – although it emphasises the aspect of information and macroscopic organisation and undoubtedly constitutes a step in the right direction[70] – should be seen as decidedly excessively general.

Sense, i.e. purposeful character of actions, emerges at the level of auto-reproducing systems attributed with identity. Or, rather, it appears in our picture of these systems. Human reason associates (frequently unconsciously) purposeful behaviour with the existence of negative feedbacks: living organisms, robots or a thermostat in a refrigerator – all behave purposefully. "Meaningfulness" or "purposefulness" are certain categories of our seeing of the world, like time, space, quantity or causality. Therefore, while one should avoid any absolutisation of these concepts, they are instrumentally convenient terms which are used by our evolution-created brain to comprehend and "assimilate" the totality of the sensations received by sense organs. The cybernetic approach fulfils here an important role in this sense that it allows to define the concept of purposefulness in a possibly objective and strict way.

---

[70]On the other hand, thermodynamic systems having the form of dissipative structures, even as simple as a whirl of water in a sink (and living system even more so), cannot be described in the language of the classical (that is dynamic) physics. Within dynamics, such concepts as information, macroscopic organisation or arrow of time have no meaning. This constitutes a clear

Finishing this short discussion on the topic of biological "senses" (I will return to the topic of senses characteristic for different levels of reality at the end of this book), I would like to stress clearly one thing. What I call the evolutionary purpose or sense of a living organism should not be understood teleologically. For evolution does not possess any externally imposed purposes or directions of development, and certainly it was not its purpose to create man. What I call the purpose of evolution is its internal property, resulting from the cybernetic mechanisms that form its foundation. Therefore, evolution is not like a man who is to reach some concrete point in some area of countryside, with a map and compass in his hand, directing himself towards a tree at the horizon. Evolution is rather like a crowd of people going through an infinite labyrinth, where most corridors end blindly, possess trapdoors or are mined, and only few corridors lead safely further on. Each of the men throws a coin at each branching to decide which corridor to choose. No wander therefore, that most of the men enter blind branches and are eventually killed by mines or trapdoors, and only the luckiest ones continue further their wandering. And even they are not certain if the next corridor they choose will not turn out to be a lethal trap. Therefore, while a man with a compass and a map has the purpose of reaching a particular point in the countryside, the only purpose for a man in the labyrinth is simply: **"to go further on"**, and his destination (if there is one) is determined mostly by chance (and the way already covered by him). The amount of possible ways to choose from is practically unlimited and only very few of them will be exploited. Therefore, the probability for just the species *Homo sapiens*, with all its features, to originate in the process of biological evolution was practically equal to zero (if biological evolution were to begin once again, it would almost certainly not lead to the origination of man). On the other hand, there was undoubtedly a finite probability that at last some rational beings would originate on our planet.

* * *

The structural aspect of the origin of life is the second (after the genetic) one, which I would like to distinguish and emphasise. In the simplest case, it consists in a separation of a living organism (individual) from the environment surrounding it. Hypercycles by themselves did not

---

signal that the physical paradigm of description of the world can be sensibly applied only to some of its aspects.

possess any spatially determined structure. Proteins and nucleic acids entering into their composition were not constantly bound together. In water solutions (as in any other solutions), there takes place the process of diffusion of substances dissolved (or suspended) in them. Therefore, some enzyme, just after termination of its synthesis, could flow away from the maternal genetic system (hypercycle) and not only be irrevocably lost, but also perform a useful work for another, competitive system. Apart from this, the concentration of various simple organic compounds (the building material of more complex compounds) in the "primeval soup" was probably not too great, and therefore there arose a need to gather them in a possibly small volume, in the nearest neighbourhood of proteins and nucleic acids constituting the hypercycle. As I argued before, a prolonged physical contact was most probably indispensable for joining auto-catalytic cycles of proteins and nucleic acids into hypercycles, that is for a primeval genetic code to get formed. Additionally, a clear separation of the interior of an organism from its surroundings would help to eliminate unprofitable influences (disturbances) of its surroundings. Finally, some clear physical division between an individual and its environment would be useful to separate well such a system as an individual. For all these reasons (as well as perhaps for several other reasons), it is indispensable for newly originating organisms to form an envelope around them (or at least a clear "phase border" between the interior and the exterior).

In all living organisms known to us today, such a cover, delimiting the borders of what is called the cell, is the double protein-lipid membrane. Double, because it is composed of two layers of lipids. Molecules of lipids are built of a polar, hydrophilic head (willingly contacting with water) as well as of a hydrophobic tail (avoiding water). In the lipid bilayer, the heads are turned outward, in the direction of the water phase (the heads of one layer contact with the external environment and the heads of the other layer – with the interior of a cell), while tails are turned inwards, in the direction of the surface where both layers join. The cross-section of the bilayer looks therefore as follows: head-tail-tail-head. Molecules of various proteins, being able to perform different movements within the membrane, are dipped in the lipid layers. The movements are: diffusion in the plane of the membrane, rotary motion, passing from one lipid layer to the other an so on. The above-presented modern opinion on the general structure of the protein-lipid membrane bears the name of the "liquid mosaic". Such a structure characterises not only the membrane surrounding a cell, but also membranes of various intracellular structures, such as a nucleus, mitochondria, chloroplasts, endoplasmic reticulum *etc*.

As the framework of the protein-lipid membranes described above – decisive for their continuity and integrity – is formed by lipids, it seems probable (at least in my opinion) that double membranes built of lipids (or rather of chemical compounds in some way analogous to them, which means composed of molecules equipped with a polar head and hydrophobic tail) constituted the cover of first organisms. Such bubbles, surrounded with a double lipid membrane (called liposomes) can be artificially produced today by means of different physical methods. These methods (for example, applying ultrasounds to a suspension of lipids in water, or pushing such a suspension under pressure through millipores) do not resemble the conditions prevailing on our planet soon after its formation (or at present). This is one of the reasons to suppose (and this is what I am going to propose) that the progenitors of lipids, building first cellular membranes, were slightly different from the modern lipids.

The second reason is the fact that lipid membranes are fully impermeable for many simple organic compounds, which first organisms had to take up from environment as building material and, perhaps, as source of energy. The membranes of modern cells possess a broad set of specific protein carriers, carrying selectively substances to and from a cell. First living organisms, for obvious reasons, did not have such carriers yet. As long hydrophobic tails constitute the main barrier for the flow of most substances through a lipid membrane, they were probably – in my opinion – both shorter and less hydrophobic in pre-membranes. Only later, when there appeared protein carriers, those hypothetical ancestors of lipids could be supplanted by more hermetic, modern lipids. The present protein-lipid membranes have a very complex structure and fulfil manifold functions. Here belong isolation from the external environment, transport of different organic substances and ions, providing anchoring places for different enzymes, transduction of hormonal signals and neural stimulation as well as production of energy (to be discussed in a moment). However, as in the case of the apparatus for copying and reading of genetic information, the entire complexity of structure and function could not originate suddenly from nothing, in one step – it is a result of a long and complex process of evolution. At the beginning, there probably existed simple, homogeneous bubbles, formed spontaneously as a result of some physical process. They could already contain most primitive auto-copying systems, organised spatially in the form of objects similar to coacervates (to be discussed in a moment).

Finally, the third reason indicating that the composition of primitive biological membranes was different than the present one is the fact that bubbles built of modern lipids do not have the ability of a independent up-take of free lipids from the surroundings and their incorporation into the bubbles themselves. As a consequence, such bubbles are not able to grow, neither to divide, after reaching a certain critical size. These properties would be very desirable for first organisms, since they had not yet (which seems obvious) any efficient mechanisms of control of the growth and division of their lipid-like covers. However, is it possible that bubbles built of any spontaneously formed organic substances and suspended in water would exhibit the above properties? We do not know. However, Oparin's coacervates (already mentioned many times) suggest strongly that such a possibility should be seriously taken into account.

If some organic substances (e.g. Arabic rubber, various proteins) are suspended in water, the molecules of these compounds will manifest a tendency to concentrate in small droplets. Such droplets discovered by Oparin are given the name "coacervates". Coacervates absorb molecules of organic substances that still remain in the suspension, which leads to an increase of their dimensions. If the size of a droplet exceeds a certain limit, determined by different physical processes, the coacervate will undergo a division into two descendent droplets. These droplets grow further, until they reach the dimensions indispensable for the next division. For an external observer, this phenomenon resembles very much the behaviour of living cells.

Undoubtedly, the analogy of Oparin's coacervates with living systems is rather superficial. Coacervates are a simple, although undoubtedly fascinating physical phenomenon, which is far less complex than living organisms. However, in my opinion, ignoring them could turn out to be slightly premature, for, if the first auto-copying systems accidentally adopted the structural properties analogous to the ones possessed by coacervates, this would have facilitated greatly the "life-start" of such systems. For, in my opinion, they could put off the problem of creating a (complicated) enzymatic system steering the growth and division of the cover. In the meantime, they could deal with more urgent needs, first of all the creation of an effective genetic apparatus. This is very important due to the previously formulated principle: "not everything at the same time".

At the earliest stage, coacervates (I consequently use this term in a broad sense, comprising all coacervate-like objects, and not only the narrow class of objects described by Oparin) probably did not possess any lipid-like cover. They were separated (in my opinion) from the environment simply by the phase limit between their interior, built of proteins, nucleic acids and maybe other compounds, and the water surrounding them. I think that such a primitive barrier could already fulfil some of the enumerated functions of a more "professional" membrane. The environment inside a coacervate could be more stable than the external environment, due to limited diffusion, for example. As a result of a better solubility inside coacervates than in water, the interior of a coacervate could gather monomers, serving as the building material of proteins and nucleic acids, namely aminoacids and nucleotides. These compounds were being continuously used for synthesis of the mentioned macromolecules, which resulted in their further uptake by a coacervate. Because the building material of such coacervates was to a large extent – according to my proposition – a polycatalon (complex of proteins of an accidental sequence of aminoacids and broad gamut of catalytic activity), a catalytic environment prevailed in their interior, that favoured replication of both proteins and nucleic acids. This led to an increase in the dimensions of coacervates and, as a consequence, to their divisions. These coacervates that managed to form hypercycles – causing a more efficient multiplication of proteins and nucleic acids – could grow and replicate faster, winning the competition with other coacervates. As they inherited the information about the structure of proteins and nucleic acids (RNA), these coacervates could also undergo improvement in the process of evolution. The predispositions of the coacervate-like objects I propose to be the first "carriers" of life are summarised in the Fig. 2.13. If the coacervates of that time were mainly composed of proteins and nucleic acids (Fox's microspheres, bubbles composed of poly-aminoacid chains synthesised outside living organisms, testify that such objects could be stable), then (this is what I suggest) coacervates were simply the **structural manifestation** of the existence of hypercycles. This emphasises the extremely tight relation which most probably (in my opinion) bound the emergence of the genetic aspect and structural aspect of life.

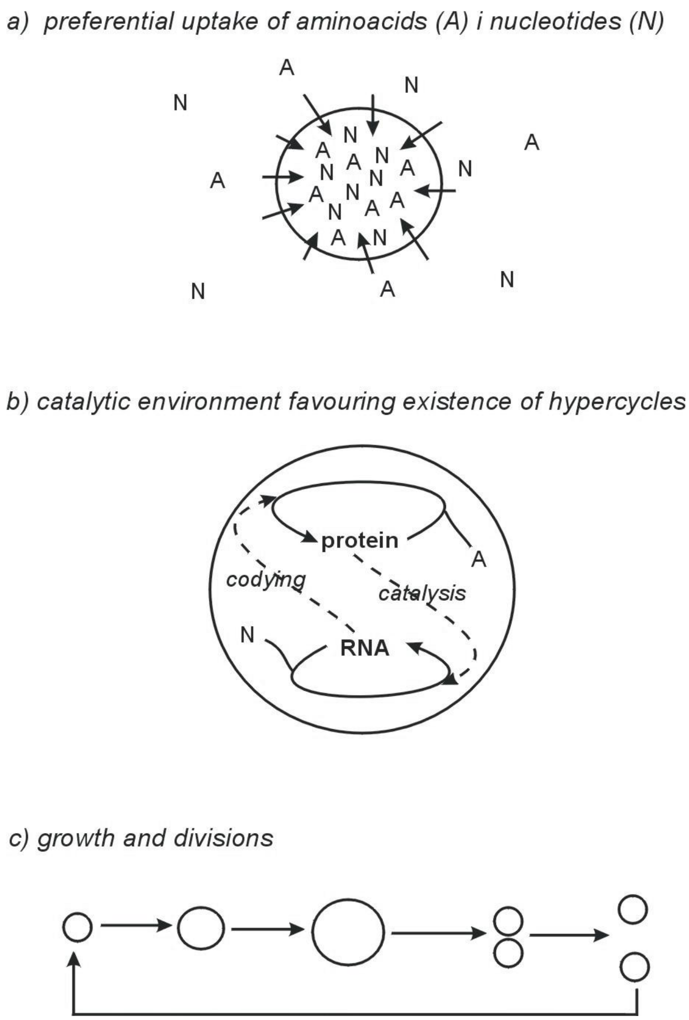

Fig. 2.13. Predispositions of coacervate-like objects to become the first "carriers" of life. Coacervates absorbed preferentially aminoacids and nucleotides from the surroundings, physically concentrated proteins and nucleic acids in a small space (volume), enabling formation of hypercycles, as well as grew and divided, which allowed for a natural selection of coacervates containing better hypercycles.

* * *

The energetic aspect is the third important aspect of the origination and development of life on our planet that I distinguish in this book. As thermodynamics shows, dissipation of energy that leads to production of entropy constitutes an indispensable condition for the existence of dissipative structures (to which the phenomenon of life belongs). I explained in the chapter devoted to physical evolution where the energy which drives all living processes on our planet comes from. Shortly speaking, the Sun[71] is the direct source of this energy for the whole biosphere. While now only green plants, transforming a part of the solar energy into the energy of chemical bounds in organic compounds, are able to assimilate this energy, at the beginning the whole life, in fact, drew energy directly from the Sun. For just the ultraviolet radiation emitted by our star was probably the main factor producing simple organic compounds from the mixture of atmospheric gases (another factor taken into consideration, namely lightnings, like all atmospheric phenomena, also draw their energy, indirectly from the Sun). However, the origination of first self-replicating systems created the need not only of the flow of energy, but of its transformation (transduction) in a channelled manner, proceeding in certain determined ways, so that dissipation of energy and production of entropy would not be a "purpose in itself", but serve as a driving force for living processes. Thermodynamics does not describe detailed mechanisms of the transformation of energy by dissipative structures. It only formulates certain conditions that are indispensable for the existence of these structures. From this point of view, life does not differ in any important way from convective currents or a whirl formed by water in a sink. Therefore, while non-equilibrium thermodynamics is heuristically interesting as a bridge built between physics and biology, then it seems to be decidedly too general for the description of energetic transformations in living organisms. Therefore, before I will pass to the energetic aspect of biogenesis, I will remind shortly how modern organisms manage the transformation and utilisation of energy.

Without risking great oversimplification, it can be said that the chemical compound called ATP (adenosine triphosphate) is the universal "carrier of energy" in a living cell. Its universality comes from the fact that (almost) all processes producing easily accessible energy converge in the synthesis of ATP, while (almost) all processes requiring energy for their occurrence draw it from decomposition (hydrolysis) of this compound. In short, reactions delivering and utilising (easily accessible form of) energy are mutually interconnected via a common accumulator of this energy, namely ATP.

Among processes producing energy that is easily accessible to living organisms, one can list aerobic respiration and anaerobic respiration (fermentation) in animals, fungi and numerous bacteria; photosynthesis in green plants, cyanobacteria and some other bacteria; and "pseudo-

[71]A marginal source of energy for living organisms can also be found in decay of radioactive isotopes in the interior of the Earth.

photosynthesis" in some halobacteria. On the other hand, among the processes that utilise the energy (i.e. are driven by hydrolysis of ATP) one can enumerate: synthesis of proteins, polycarbohydrates (polysugars), fats an so on; transport of organic substances and ions through biological membranes; contraction of muscle, conduction of neural impulses, movement of chromosomes during a cell division, movements of cytoplasm in general, movements of cilia and vibracula, secretion of different substances from the cell, bioluminescence (production of light) as well as many other phenomena.

ATP is a nucleotide containing three phosphate groups (further on, I will discuss its relation to the nucleotide A, described earlier, while discussing the structure of nucleic acids). It is formed as a result of adding the phosphate group to ADP (adenine diphosphate, containing two phosphate groups, which can be easily calculated). This process requires a supply of energy (which comes from the visible radiation emitted by the Sun or from oxidation of respiratory substrates)[72]. On the other hand, the reverse reaction, i.e. hydrolysis (decomposition) of ATP into ADP and inorganic phosphate (phosphate group) releases large amounts of energy. This energy can be dissipated in the form of heat, or transferred to other process (e.g. chemical reactions, by coupling this process with hydrolysis of ATP) that need energy to proceed. Therefore, if we treat the system ATP + ADP as some kind of a battery (accumulator) of a living cell, the synthesis of ATP from ADP will correspond to loading this battery, while hydrolysis of ATP coupled to performance of some work (chemical, mechanical, osmotic work and so on) will correspond to its unloading. While discussing the phenomenon of negative feedback, I discussed the way a cell ensures the rate of ATP supply (synthesis), adjusted to the current demand for this compound.

ATP can be synthesised by transferring a phosphate group from some intermediate metabolite, that forms a transitory stage of the conversion of an organic compound taken up as nourishment. This manner of ATP production is called **substrate phosphorylation**. It takes place in glycolysis, for example, that is the metabolic pathway of glucose decomposition (and of other simple sugars), being a part of both aerobic respiration and fermentation. However, while substrate phosphorylation constitutes the only mechanism of ATP production in the latter process, in aerobic respiration it plays only a minor role, and it does not occur at all in photosynthesis. A completely different kind of phosphorylation takes place in these cases instead of substrate

[72]The amount of energy 'contained' in ATP depends on the ratio of concentrations of ATP, ADP and inorganic phosphate.

phosphorylation. It is called **oxidative phosphorylation** in the case of aerobic respiration and **photosynthetic phosphorylation** in the case of photosynthesis. The general principle of energy production in the process of "pseudo-photosynthesis" in halobacteria is also very similar. Because all these processes do not differ with respect to the essence of the basic mechanisms, I will designate them with the common name "**membrane phosphorylation**".

The existence of the membrane phosphorylation was proposed in the 60s by Peter Mitchell within the framework of the so-called **chemiosmotic theory**. This is one of the most elegant and fruitful theories concerning the biochemical level[73], so no wander that its inventor was awarded the Nobel Prize. According to the chemiosmotic theory, the intermediary in the synthesis of ATP is found in the gradient (difference of concentration) of protons across an appropriate protein-lipid membrane, connected in a given case with energy production. This can be the inner mitochondrial membrane or cellular membrane of bacteria in the case of aerobic respiration, the membrane of thylakoids in chloroplasts for the process of photosynthesis, or the purple cellular membrane in the case of "pseudo-photosynthesis" in halobacteria.

Protons must be pumped from one side of such a membrane to the other side, in order to form a gradient of protons, that is a difference in their concentration on both sides of a membrane (and electric potential, due to transfer of a positive charge possessed by a proton). This purpose is realised by a certain kind of proteins, called proton pumps. In mitochondria and bacteria, protons are pumped to the outside of mitochondria and bacteria, respectively, while in the case of chloroplasts – to the inside of the grana of thylakoids. In the case of aerobic respiration, proton pumps – being parts of the respiratory electron transport chain (containing also cytochromes, among other elements), anchored in the inner mitochondrial membrane – are driven by a flow of electrons through the chain. These electrons are transferred from decomposed nutritive substances (sugars, fats, proteins) into the respiratory chain through the mediation of compounds called NAD and FAD. Electrons (and hydrogen ions) are finally transmitted by the respiratory chain to oxygen, which leads to the formation of water. The whole electron flow sequence takes the following form: nutritious substances → NAD (or FAD) → respiratory chain → oxygen. The protons are pumped, as I mentioned, at the stage of the respiratory chain. As the proton gradient serves synthesis ATP (more about it in a moment), it is clear how the organic compounds taken

[73]In line with a few concepts that formed the foundations of molecular genetics, such as the model of double helix, genetic code or the theory of operon.

up with the nourishment and oxygen breathed in by lungs (or analogous organs) serve the process of energy production (and why they are indispensable for life processes in most heterotrophic organisms, including man). Hetetotrophic organisms simply oxidise organic substances with oxygen. However, this combustion proceeds along very complex routes and a significant part of released energy is not dissipated as heat, but drives the process of proton pumping (and afterwards synthesis of ATP).

The process of pumping of protons in photosynthesis (in its light phase) is analogous. Here, the flow of electrons through the (photosynthetic) electron transport chain drives proton pumps, carrying protons across the membrane of thylakoids. However, two aspects of photosynthesis constitute a reversal of aerobic respiration. Firstly, electrons flow in the opposite direction, i.e. from water (two protons and two electrons become separated from oxygen) to NADP (a compound very similar to NAD). The energy needed to break a molecule of water (free oxygen is released in this process) and transfer electrons to the photosynthetic chain is delivered by visible radiation, the absorption of which is mediated by chlorophyll – the green pigment giving the colour to plants. Secondly, protons are pumped inside the grana of thylakoids, so that the gradient of protons gains opposite direction than in mitochondria (greater concentration of protons inside than outside).

The process of proton pumping is significantly simpler in halobacteria than in the case of mitochondria (as well as aerobically-respiring bacteria) and chloroplasts. The only proton pomp functioning here (in the previous cases there were two or three kinds of them, arranged into a respiratory chain or a photosynthetic chain) is formed by a protein called bacteriorhodopsin and driven directly by light radiation. Absorption of a quantum of radiation (photon) leads to a conformational change (of spatial arrangement) in a molecule of bacteriorhodopsin and to binding of a proton from the internal side of the cell membrane, while a return to the previous spatial arrangement is connected with a transfer of a proton into the outside of the cell. A shot from a cross-bow can serve here as an analogy. The energy supplied from the outside is needed for drawing a cross-bow, just as an absorption of a quantum of radiation is needed to produce a conformational change in a molecule of bacteriorhodopsin. After releasing the trigger, the cross-bow returns violently to its initial state, which is connected with shooting an arrow. Similarly, during a quick relaxation of tension in its molecule, bacteriorhodopsin is able to perform work, which takes the form of a transfer of a proton to the other side of the membrane, against the gradient of protons concentration. Proton pumps operating in the respiratory and photosynthetic chains work according to a similar principle. The main difference consists in the fact, that the energy resulting in a conformational change of a pump is delivered in the form of electrons, taken over from the previous component of the chain, while a liberation of this energy (pumping some number of protons – this number is different for different pumps) is connected with a transfer of these electrons further on, to the next component of the chain. In this way, electrons flowing through the respiratory chain, for example, pass through the sequence of three proton pumps and lead to triple pumping of protons.

However, irrespective of the shape taken by the mechanisms of proton gradient formation, the final effect is the same in all cases described above – a difference of the concentration of protons (as well as of electric charges) on both sides of a membrane responsible for conversion (transformation) of energy. Now, the second phase of the membrane phosphorylation proceeds, namely spontaneous return of protons through the membrane to this side of the membrane, where the concentration of protons is smaller, coupled with a synthesis of ATP from ADP and phosphate. Each kind of atoms, ions and molecules possesses a natural tendency, described by laws of thermodynamics, to a flow from the compartment where their concentration is greater to the compartment with a lower concentration. Additionally, in the case of charged ions, their movement across a membrane is favoured by a (properly directed) electrical potential (positively charged protons are attracted by a negatively charged compartment, where their concentration is smaller). Pumping of protons "upstream" the gradient of concentration and electric potential requires energy. On the other hand, their return in the opposite direction is connected with a release of energy, which can be used for ATP synthesis. This reaction is catalysed by the enzyme called ATP synthase. In mitochondria, it is situated in the inner mitochondrial membrane (similarly as the respiratory chain) and its molecules have the shape of "mushrooms" with their "caps" directed towards the inside of mitochondria. The ATP synthase possesses a channel perpendicular to the surface of the membrane, through which protons can come back to the interior of mitochondria, which is accompanied by ATP synthesis. Because the inner mitochondrial membrane is hardly permeable for protons, the main possibility for them to return to the mitochondrial matrix is provided by the ATP synthase channels. The mechanism of membrane phosphorylation has a similar form in chloroplasts, aerobically respiring bacteria and halobacteria, but in chloroplasts, again, everything is reversed. The "caps" of molecules of the

ATP synthase are oriented outwards the membrane of thylakoids, which is the direction of the back flow of protons. The general idea of the membrane phosphorylation is presented in Fig. 2.14 ("**H^+**" in bold designates the compartment where the concentration of protons is greater).

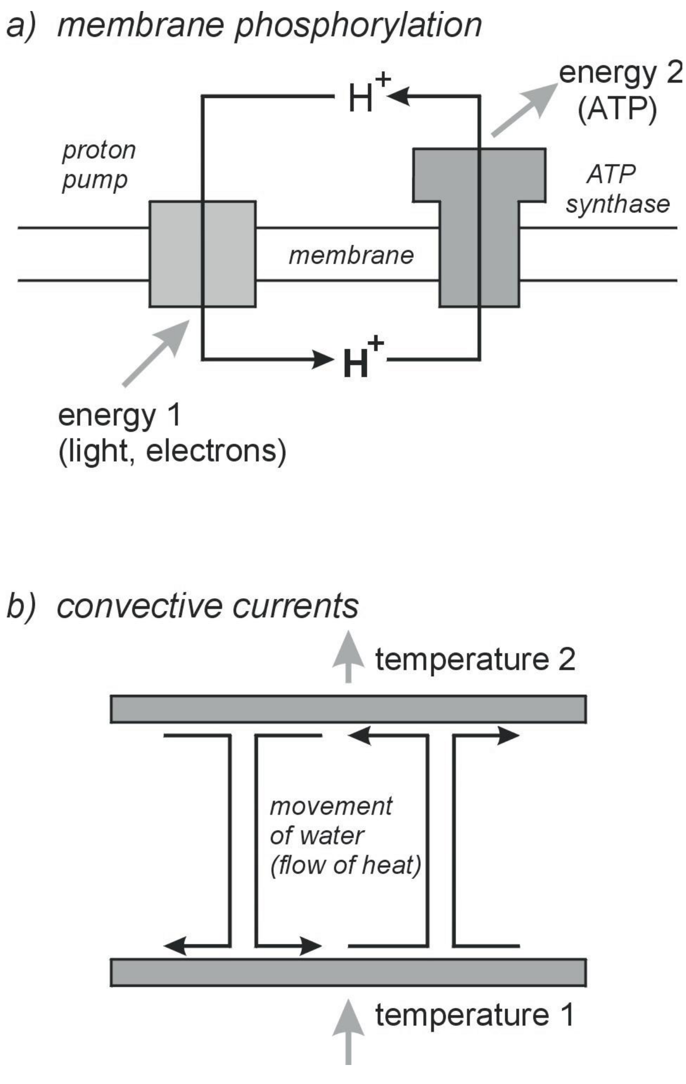

Fig. 2.14. General scheme of membrane phosphorylation and its formal similarity to a simple dissipative structure – convective currents. As the flow of heat between two plates of different temperatures forces upward and downward convective currents in the water between them, so the transformation of energy from one form into another form in living organisms causes "upward" and "downward" currents of protons (hydrogen ions) from one side of a membrane connected with energy production to the other side of this membrane.

The whole mechanism of membrane phosphorylation can be presented by the following analogy. Let us take a box filled with air and divide it into two halves. We can built into this division a small air pomp, driven by a small combustion engine, pumping molecules of air from the left chamber to the right chamber. This operation will produce underpressure in the left chamber, overpressure in the right chamber, and a gas concentration gradient across the division. If we make a small hole in the division, the air will come back with a hiss from the right chamber to the left one. We can, in turn, place a little windmill connected with a dynamo on the way of the flowing air and thus produce electric energy. In the above example, the chemical energy released during combustion of petrol becomes transformed into the energy of air pressure (strictly speaking – a difference of pressures), an this, in turn, into electric energy.

Similar transformations of energy take place in membrane phosphorylation. The chemical energy of respiratory substrates or the radiation energy of photons is transformed here into the electro-osmotic energy of the proton gradient (that is energy connected with the difference of proton concentration[74]), and then into chemical energy "accumulated" in ATP. In the above analogy, the air pump supplied with energy by a combustion engine corresponds of course to the proton pumps of the respiratory chain or photosynthetic chain (driven by the flow of electrons) or to bacteriorhodopsin (utilising radiation energy directly); the gradient of air pressure is the proton gradient; the windmill with a dynamo corresponds to the ATP synthase; while the electrical energy is an analogue of the chemical energy of ATP.

At this moment, I would like to make a small digression and mention an extremely important feature of the energetics of living organisms. In heterotrophic organisms, for example (animals in the first place), the energy which is transformed into the energy of the gradient of protons, and then into the energy accumulated in ATP molecule, comes primarily from the energy of chemical bounds in a molecule of a respiratory substrate, e.g. glucose or fatty acid. Only "high-energy" electrons, transferred from such a substrate onto the respiratory chain, serve the purpose of pumping protons against the gradient of proton concentration. However, the energy in the form of ATP is indispensable for the first stages of the chemical transformation of respiratory substrates (and also for taking them up from an environment). Therefore, some amount of ATP has to be **already** available for **further** ATP production. This amount is needed to activate the

[74]It can be compared with the potential energy of a stone lifted above the ground level: "discharging" the energy (the stone falls down) can produce work, e.g. in the form of heat.

first reactions leading to production of this compound. This constitutes another example of the previously-discussed vicious circle, so characteristic for the essence of living organisms. In this particular case, in the scientific jargon one ought to speak about the non-linearity of living systems (there is a minimal threshold concentration of ATP, below which the biochemical pathways leading to ATP synthesis cannot be activated and a cell dies).

Let us come back to the leading topic of this fragment of the book. (Some) protein-lipid membranes play the crucial role in the production (transformation) of energy in living organisms. While discussing the structural aspect of the origination of life on the Earth, I indicated that it was necessary for first auto-copying systems to possess a pseudo-lipid cover. It would be difficult not to relate these facts one with the other. All the more because a huge majority of the processes laying at the basis of life cannot proceed spontaneously. A supply of energy is needed for them to occur. Therefore, it seems to me that the originating life had to "think" quite early about gaining possession of some system of production of this energy.

How could such a system originate? Let us look once again at the simplest system of transformation of energy in the case of membrane phosphorylation, namely the one in halobacteria (see also Fig. 2.14). There are two components ensuring continuous circulation of protons, namely bacteriorhodopsin, pumping protons outside a cell under the influence of light radiation, and ATP synthase, coupling the return of protons to the interior of a cell with synthesis of ATP. (In the above system, protons circulate in a closed circuit, dissipating energy in this way – which is strikingly similar to the upward and downward currents in convective movement; this is illustrated in Fig. 2.14. This observation makes once more the point that life is simply a very complex dissipative structure, whose "purpose" – from the thermodynamic point of view – consists merely in effective dissipation of the energy of solar radiation that reaches the surface of the Earth).

Now I would like to propose a description of possible spontaneous origination of the above described system of production of energy by the mechanism of membrane phosphorylation. A removal of one of the two basic elements constituting this system (that is either proton pump(s) or ATP synthase) takes all the sense from the whole system. On the other hand, a simultaneous origination of both components as a result of a coincidence seems to be improbable. Therefore, one should formulate again the question, whether there exists any functional transitional stage which could be an intermediate phase in the process of origination of the simplest energetic system?

In my opinion, one can imagine such as stage, containing only one of the two mentioned components. As both components are necessary to fulfil the function of energy production, that hypothetical intermediate stage must have been connected with other purposes. Fortunately, we do not have to rely here exclusively on our imagination. In my opinion, the process which could be the required "missing link" exists until now. This is the transport of organic substances and ions through protein-lipid membranes.

Modern bacteria and mitochondria use for this purpose nothing else but just the proton gradient (called also the protonmotive force) across the surrounding membrane. Transport of different substances occurs now through the mediation of protein carriers, specific for each substance. In my opinion, at the dawn of life, various chemical compounds must have been simply passing through a pseudo-lipid membrane. The proton gradient can favour uptake of particular substances from the surroundings in two ways. Positively charged molecules (cations) have a tendency to gather in the negatively charged (due to pumping out of protons) interior of a bubble surrounded by a membrane. On the other hand, electrically neutral or negatively charged compounds will undergo a selective transfer into the bubble interior, if their transition across the membrane is in some way coupled with translocation of protons. Presently, specialised protein carriers are responsible for that coupling. At the dawn of life, this could have been some simpler mechanism.

Additionally, weak acids (to which many simple organic compounds belong) tend to dissociate (release an H^+ ion) in an alkaline environment, prevailing inside a bubble, while in an approximately neutral environment (e.g. sea water), they exhibit a greater preference for binding this ion. If weak acids can cross through protein-lipid membranes (as it takes place in modern organisms) only in the neutral (not dissociated) form, then, once they get into the bubble, they are in a kind of trap – they dissociate and cannot get out of it. This leads of course to accumulation of such a compound in a bubble.

Therefore, I think that the formation of the proton gradient, even if it did not serve ATP synthesis at first, could have been profitable for first organisms. Organic substances, they built their "bodies" of, were probably available in the oceans in a pretty big dilution, which increased additionally as they were used by originating life. Therefore, it become an urgent need to extract these various compounds efficiently from the water. The proton (or, more generally, ion) gradient is a perfect tool for this purpose. Only one element – proton pump – suffices for the gradient to come into being. It is difficult to say what could have been used by first organisms as the source of energy for pumping of protons, although, some analogue of bacteriorhodopsin that uses light radiation (and is encountered in modern halobacteria) is a probable candidate, due to the simplicity of its structure and function. The energetic function of the proton gradient, that is ATP production, could have emerged later on, together with the origination of some prototype of the ATP synthase. In this approach (proposed by me), production of energy would be secondary with respect to the transport function, which needed only the presence of a proton pump.

The question still remains, what primeval function could ATP serve in a cell. For it is also hardly probable that auto-copying systems simultaneously learned to produce ATP and to use this compound as a source of energy. However, one should realise that ATP is identical with the nucleotide A entering into the composition of RNA. In the process of polymerisation of the latter, two phosphate groups are torn off from ATP, while the rest of the molecule, containing a remaining phosphate group, becomes incorporated into an RNA thread. By analogy to ATP, there also exist compounds called UTP, CTP and GTP, constituting "precursors" of the "bricks" that build an RNA thread, namely the formerly discussed U, C and G. What is interesting, UTP, CTP and GTP also fulfil today the role of carriers of energy (like ATP), although they serve a significantly smaller number of reactions than ATP and therefore they are not so universal. UTP, CTP and GTP are presently mostly synthesised from UDP, CDP and GDP through the transfer of a phosphate group from ATP (ADP is formed). In my opinion, however, a primitive ATP synthase was initially unable to distinguish between different types of nucleotides and utilised the proton gradient to "load" each of them with energy. The energy for binding particular nucleotides into a DNA thread is provided by hydrolysis of two phosphate groups from their loaded form XTP (where "X" stands for A, U, C or G). Therefore, before a thread of a nucleic acid is synthesised, energy has to be delivered to nucleotides, via the synthesis of XTP from XDP or XMP (the last contains only one phosphate group). This, in my opinion, suggests a possibility that primeval organisms learned to transform energy and accumulate it in the form of ATP more or less at the time when they acquired the apparatus for recording genetic information. In such a case, the function of energy accumulation would increase significantly the efficiency of copying genetic information in primitive auto-copying systems[75]. Here we see the main feature of the origination of life I would like to emphasise. I propose that the genetic, structural and energetic aspects of biogenesis (origination of life) are mutually interrelated and that each stimulated the origination of the others. The common denominator of the structural and energetic aspects is found in the membrane closed into a bubble. The energetic aspect is connected with the genetic one through the use of common XTP molecules[76], while a spatial (structural) arrangement of proteins and nucleic acids within coacervates could be (in my opinion) an indispensable condition for a primitive genetic apparatus to come into being.

The energetic aspect of life, as the structural one (and contrarily to the genetic aspect), is not **directly** related to the cybernetic essence of life. While the structural aspect created a physical barrier separating a primeval genetic apparatus (reducible in fact to hypercycles) from an environment, the energetic aspect propelled the whole system, ensuring its efficient functioning as a dissipative structure. Anyway, the energetic aspect remained almost from the beginning under the control of the genetic apparatus (proton pumps are after all proteins coded by nucleic acids), which directed the dissipation of energy along appropriate paths, thus incorporating the energetic aspect into the network of negative feedbacks of this primeval auto-copying system. This constitutes a premise supporting the claim of superiority of the genetic aspect in relation to the remaining two aspects, that constitute merely particular cases of a material realisation of general cybernetic instructions.

* * *

Now I would like to recapitulate briefly everything said above, that is to show what the initial stages of the evolution of life on the Earth could look like. This scenario comprises some commonly-accepted ideas, some concepts chosen from several existing alternative opinions, and certain propositions of my own, as has been indicated in the text.

---

[75]The fact that very few modern organisms use a carrier of energy simpler than ATP, namely pyrophosphate, that is two joined phosphate groups, seems to testify against such a hypothesis. However, it is not certain if this is an originally primitive property, or a secondarily simplified property.

[76] It is worth noting that the nucleotide A also enters into the composition of other metabolites important for the metabolism of a cell, such as NAD, NADP, FAD or coenzyme A. This proves its primitive and universal character.

As the first stage, one could consider an object similar to a coacervate (perhaps already surrounded with a pseudo-lipid membrane, as a result of elementary physical phenomena), containing (or built of) short auto-copying protein chains with an accidental sequence of aminoacids. Such a set of simple proteins, which I called polycatalon, was able to catalyse various chemical reactions, including autocatalysis. Simultaneously, threads of nucleic acids (RNA) replicated in coacervates exploiting universal "enzymatic" properties of polycatalon. A coacervate grew and underwent divisions, as a result of synthesis of proteins and RNA from aminoacids and nucleotides absorbed from an environment. Both kinds of replication (auto-copying of bubbles and of chains of proteins and threads of nucleic acids occurring in these bubbles) were not (yet) mutually co-ordinated and proceeded independently. If some of the proteins in the polycatalon were able to act as primitive proton pumps, this led to the formation of proton gradient and more efficient uptake of substances from an environment. At this stage, the system was still purely physical and did not exhibit the fundamental properties of a living organism.

Other proteins in a polycatalon could have been able to transform XDP into XTP, drawing energy for this process from dissipation of proton gradient. On the one hand, this presumably led to a formation of an energy production system, while, on the other hand, it enhanced efficient synthesis of nucleic acids (probably RNA). In such a case, both processes – efficient reproduction of nucleic acids and the subsequent formation of the system allowing genetic information to be encoded and expressed, regarded today to be crucial for the existence and evolution of life – could have been primarily by-products of the appearance of a primitive ATP synthase (I want to remind that this is my suggestion). The process of energy production in the form of ATP (and also CTP, GTP and UTP) could enter into the biogenesis process through the backdoor, due to an "a-typical" use of the proton gradient, that initially resulted simply in transport of substances via the membrane of a bubble. This in turn, would enable (also "by chance") the appearance of "energetically loaded" monomers entering into the composition of RNA, and therefore, origination of auto-catalytic cycles of proteins and nucleic acids, and of hypercycles afterwards. One should remember that the origination of life seems to be a purposeful process only *ex post*. In its essence, it was a phenomenon resulting from normal physicochemical processes, devoid of any purpose.

Finally, I think that the crucial turning point in the course of biogenesis consisted in the appearance of primeval tRNA, coupling auto-catalytic cycles of proteins and nucleic acids in the above-discussed hypercycles (although, it is probable that auto-catalytic cycles of nucleic acids profited earlier from catalytic aid of some proteins, entering into the composition of the polycatalon). If one would necessarily want to fix a single moment of the origination of life (the cybernetic paradigm can be used as the criterion), then, the above event seems to be the best candidate. Presumably the primeval genetic code possessed a significantly simpler form from its present form. It is probable (again – at least in my opinion) that codons contained initially only two (and not three) coding nucleotides, separated with a single non-coding nucleotide. It seems also to be probable that the primeval genetic code recognised not single particular aminoacids, but rather groups of aminoacids with similar physicochemical properties. There are several such groups: hydrophobic, hydrophilic, aromatic, alkaline, acidic. Even at present, in highly specialised proteins, an exchange of one aminoacid for another from the same group frequently does not influence significantly the properties of a protein. In primitive proteins, different aminoacids from the same group were perhaps used interchangeably. Initially, five kinds of tRNA – recognising the relevant groups on the basis of their physicochemical properties – sufficed to code five groups of aminoacids. A spontaneous origination of such a code does not seem to be so extremely improbable, when it is recognised that life had all the waters on the terrestrial globe as well as tens (or even hundreds) of millions of years to test different variants. After this decisive step, the evolution of life became a self-propelling process, and the main progress consisted in differentiation and specialisation of proteins in the polycatalon, which was equivalent to enriching and improving the genetic record. This constituted an example of a positive feedback (not related to the positive feedback resulting from auto-copying): more efficient enzymes enabled better coding and reading of the genetic information, while this, in turn, allowed for production of a greater amount of more efficient and specialised proteins, i.e. first of all enzymes.

The main role of the genetic apparatus was, therefore, to enable differentiation and specialisation of particular proteins in the polycatalon, through ensuring transfer of information about their structure and function from generation to generation. Since then, all accidental but profitable changes in the aminoacid sequence in a protein chain were not lost after the "death" of a given chain (every organic compound undergoes eventually decomposition). They were accumulated as changes in the sequence of nucleotides in a nucleic acid. As the copying of a twin thread of nucleic acid proceeds with (quite) a considerable accuracy, and it does so much faster than decomposition of this compound – which is due to the operation of enzymes (even the nonspecialised ones that entered in the polycatalon) – the information concerning the structure of proteins is retained and can evolve. In this sense, one can claim (as I propose) that polycatalon developed the apparatus of copying and reading genetic information in order to make its own evolution and specialisation[77] possible, and thus, it became a hypercycle.

It is frequently assumed (since the publication of "The Selfish Gene" by Richard Dawkins) that the subject of evolution is nucleic acid (presently mainly DNA), that exploits the entire enzymatic apparatus and the whole system (organism) – its "carrier" – to multiply and realise expansion of itself ("selfish DNA"). However, one could equally well maintain that life can be in fact reduced to a complex of enzymes that (primarily as the polycatalon) "produced" the genetic record to copy themselves more effectively, for enzymes catalyse the formation and/or uptake from an environment not only of themselves or their building substances, but also of all other components of living organisms, responsible for the entire structure and all functions of these organisms. After all, DNA is only a passive record of the structure of enzymes, while enzymes constitute the functional machinery of the cell, able – among others – to replicate their own record. Such a point of view finds an additional corroboration if hypercycles of proteins and nucleic acids were actually preceded in time by protein cycles (polycatalons). Undoubtedly, the genetic information – recorded in nucleic acids (at present, mainly in DNA) and encoding the sequence of aminoacids determining the structure of all proteins in an organism – constitutes an immensely important element that allows the identity of this organism to be preserved and multiplied. However, proteins are not only the product of the system of reading genetic information. As enzymes, they fulfil a very important function in catalysis of particular stages of this process. For example, DNA polymerase is responsible for replication, RNA polymerase for transcription, while aa-tRNA synthase and ribosome proteins – for translation. A lot of other proteins fulfil various roles, aimed at keeping the entire system (including genetic record) alive and producing its copies. For, during evolution, a system of negative feedbacks has been built up to support the genetic apparatus, and proteins play crucial role in the functioning of the system. Therefore, both proteins and nucleic acids participate in recording, copying and reading genetic information. One ought to remember that the sequence of nucleotides in genes has no sense without the genetic code, establishing the correspondence between this sequence and the sequence

[77]Let us remember, however, that there exists an alternative scenario of the course of events, namely the "world of RNA", that is initial origination of auto-copying threads of RNA with some

of aminoacids in the protein coded by a given gene. Genetic information is not only the information recorded in the sequence of nucleotides in DNA, but also the information about the manner of reading this information by the system primarily built of different proteins. Similarly, the information contained in this book is not only a determined sequence of letters, but also the knowledge of English language as well as of some fundamental concepts in biology provided by the Reader. Therefore, in my opinion, nucleic acids or particular genes are not the subject of evolution. Its subject is constituted by entire **auto-copying systems** (evoluons), composed of nucleic acids and proteins – in the simplest case.

* * *

Now I will pass on to enzymes and the role in the evolution of life played by the process of their "emergence" from particular proteins of the polycatalon. Enzymes, as biocatalysts, are responsible for carrying out most of the transformations occurring in a cell at the biochemical level, that is metabolism. What does the process of enzymatic catalysis consist in[78]? Let us imagine a cube of sugar placed on a table in a closed room. Molecules of sugar will react with oxygen contained in the air, which will eventually lead to complete oxidation of sugar. This process will produce carbon dioxide and water and release a portion of energy in the form of heat. However, this transformation will proceed very slowly, probably for years. We can speed it up significantly by warming up the cube of sugar – it will start to burn, which will terminate the process of oxidation in a few minutes, releasing a great amount of heat (the same amount of heat, although spread over years, will be released in spontaneous oxidation, described before a moment).

Why did the initial heating of the lump of sugar speed up the reaction so much? Each reaction to proceed needs the so-called energy of activation. This is so, simplifying the matter somewhat, because the reacting molecules must collide with a certain minimal velocity to break the old chemical bonds and/or to form new ones. In the indoor temperature, only a very small

---

catalytic properties.

[78]Enzymatic catalysis is based on the same principle as inorganic catalysis, but differes from the latter by, among others, efficiency (enzymes as a rule speed up the catalysed reactions much more than inorganic catalysts do) and specificity (enzymes catalyse only one particular reaction or a narrow group of similar reactions).

number of molecules acquires a sufficient kinetic energy, as a result of accidental collisions with other molecules. Therefore, a spontaneous oxidation of sugar proceeds very slowly in these conditions. An initial warming of a cube of sugar supplies sufficient amount of kinetic energy (which is after all proportional to temperature) for many molecules to gain appropriate velocity. The reaction (burning of sugar) produces large amount of heat, so that further warming up is not necessary – the reaction sustains itself (after exceeding the so-called flash point).

Therefore, a reaction can be speeded up by increasing temperature. Living organisms, however, cannot use this method for at least two reasons. Firstly, high temperature accompanying combustion would destroy the structure of a living cell. Secondly, a cell is not interested in the production of brilliant fireworks with flames and in releasing great amounts of heat, but rather in accumulation of the produced energy (or at least a part of it) in some useful form. Heterotrophic organisms (man included) carry out exactly the same reaction as described above, that is oxidation of sugars (e.g. glucose) with participation of oxygen to produce carbon dioxide and water. However, they do this (roughly) in indoor temperature, and the transformation proceeds in many steps, each of them catalysed by a separate enzyme. Some of these steps (membrane phosphorylation, substrate phosphorylation in glycolysis) lead to ATP synthesis (universal carrier of energy). However, there still remains the question how enzymes accelerate the rate of reactions.

I mentioned earlier that, for a reaction to happen it is necessary that the kinetic energy of molecules which are to react should be at least equal to the energy of activation. As it is not possible to accelerate the movement of molecules (increase temperature), and therefore to increase the kinetic energy to an appropriate level, it is necessary to lower the energy of activation. This is what enzymes do. By an appropriate shape of their active center and dislocation of electric charges in it, enzymes influence some chemical bonds in a substrate molecule in such a way that they become much easier to break, which significantly accelerates the rate of reaction.

A rubber ball falling out of a glass will serve as an analogy for a chemical reaction. Such a ball cannot fall out of the glass (we assume that the glass is positioned vertically, bottom down), unless we shake the glass. The shaking corresponds to collisions of molecules (above the temperature of the absolute zero molecules have always some kinetic energy). We can increase the intensity of shaking to quicken (make more probable) the desired result – falling of the ball out of the glass. This corresponds to increased temperature. However, we can also lower the edge of the glass, that is the energy of activation. It is the strategy applied by enzymes.

Enzymes accelerating chemical reactions increase significantly the intensity of a huge amount of processes in living organisms. Thanks to their specificity, they catalyse only those reactions which are necessary for a cell. The possibility of inheriting all changes in their structure and function is of crucial significance, and it opens the way to quick improvement and specialisation of the enzymatic system – essential improvement, compared to a weakly differentiated polycatalon. This resulted in improving all functions performed initially by polycatalon, such as synthesis of proteins and nucleic acids, pumping of protons, synthesis of ATP, as well as in adopting new tasks, such as transport across membranes, cell division and thousands of others. This is one more reason why the formation of the genetic apparatus can be regarded as the moment at which life, having forced its way through a sequence of "narrow passages" on the way to its origin, had at last an open way to colonise the surface of our planet.

* * *

This will finish the description of a possible course of biogenesis. I would like to stress once more that the above-presented process of origination of life on the Earth constitutes to a large extent a mental speculation, partially invented by me. It is consistent with the known scientific facts and is based on them, but it reaches beyond them in many cases. It has turned out many times that the imagination of the nature significantly exceeds human imagination. Did actually the birth of life have such a course as presented above? In details – almost certainly not. In a general outline – maybe[79]. This is not the most important thing at this point. My aim was not to report strict a certain historical phenomenon, but to show that, at least in principle, life did not have to originate at once as a highly complex system, such as we know it today. On the contrary, it is possible to imagine its development proceeding in stages, step by step. A strong intuitive objection arises, when one faces the necessity for highly organised life to originate from inanimate matter in one discrete step, which is consistent with the principle "everything or nothing". Such a process of origination is seen as an extremely improbable event. On the other hand, if we agree the

[79]In fact, there have been proposed various concepts differing, at leasst in details, from the story presented by me, for example, the mentioned "RNA world" or an origin of first living systems on the "matrix" of different minerals. These concepts can constitute either an alternative possibility or a supplementary approach. Certainly, we will never learn all the details of the process of origination of life.

"alive" is separated from the "inanimate" by a sequence of intermediate states, than we can imagine a slow, gradual development, an emergence of life from inanimate matter, where particular stages, passing more or less continuously one into another, are fully functional and able to evolve further on. Using the analogy sketched above, I intended to show that, against intuitive expectations, it is possible to reconstruct in such a way a car running at full speed by adding single new elements, and that not only it will not crush against the nearest tree, but eventually will fly up in the sky. After all, the evolutionism many a time struggled with the problem of "impossible-to-think" intermediate stages as totally non-functional. As far as I know, there were eventually proposed probable solutions of most "paradoxes" of the type. As to the rest, well, the imagination of the nature seems to be still much richer than human imagination.

It seems therefore that life could originate spontaneously. The above history of the birth of life (speculative at many points) was presented as an attempt at obtaining a picturesque illustration of these aspects of life which seem to be most important for understanding of its essence and properties, among which one finds the possibility or even inevitability of evolution. The cybernetic definition of life formulated in the previous chapter says nothing about the problem how some general instructions were realised in a given case of the life on the Earth. The presented reasoning was to visualise that, although the genetic aspect is very closely related to the cybernetic paradigm, the originating life had also to take care of the realisation of other aspects necessary for its coming into being, including the structural aspect and energetic aspect. Initially constituting simple physicochemical processes (e.g. growth and reproduction of coacervate-like bubbles), these aspects, after passing under control of the genetic aspect (this control is performed through appropriate structure and function of adequate enzymatic proteins), became part of a complex network of negative feedbacks, maintaining the survival and reproduction of the whole system.

Thus we arrive to the stage when there appear first living organisms. And at this point we will finish our considerations concerning biological evolution. The reason is simple. Although it may sound paradoxically, from the point of view of understanding of the most important fundaments of the essence of life, nothing interesting happened during the following three billion years. Please, do not understand me wrongly. Of course, life formed in this period an enormous diversity of forms. There originated *Prokaryota* (bacteria, cyanobacteria) and *Eukaryota* (nuclear organisms). There appeared plants and animals, both unicellular and multicellular. Plenty different plans were produced, whose structure gradually underwent differentiation and complication. Life left oceans and conquered land, took possession of air, extended its reach from tropics to high mountains and the polar zone. Finally, the species *Homo sapiens* originated and developed a psyche conscious of its own existence and able to create a culture. The abundance of the forms of life is undoubtedly impressive. However, as I discussed earlier, the fundamental mechanisms of functioning of all living organisms known to us are in fact identical. The difference between various forms can, in principle, be reduced to coding of different proteins by their DNA, among them chiefly regulatory proteins.

The regulatory proteins are responsible for controlling expression of genetic information, that means for deciding which genes will be activated, when and to what extent, which will be followed by a synthesis of the appropriate proteins. The entire individual development of a man, from a fertilised egg cell to the adult stage, is a result of the action of an extraordinarily complex system of regulatory proteins (and, indirectly, a result of spatial gradients of the concentration of the so-called morphogenetic factors, formed as a result of their operation). Nevertheless, such proteins are already found in bacteria. Their operation is described in this case by the theory of the operon, for the formulation of which Jacob and Monod received the Nobel Prize. The regulation is reduced here to production of enzymes serving decomposition of different nutritive substances in response to the appearance of these substances in an environment. Molecules of the sugar called lactose, to use the already discussed example, inactivate a regulatory protein called the repressor, which in its active form inhibits the synthesis of the mRNA encoding three proteins responsible for transport and conversion of lactose. This results in the synthesis of these proteins and decomposition of this sugar.

In *Eukaryota*, and particularly in multicellular organisms, the system of the functioning of regulatory proteins is decidedly much more complex. It determines the course of the embryonic development, and thus, the degree and kind of complexity of the adult organism. Generally speaking, the velocity of increasing the structural complexity in the process of evolution was proportional to the current level of complication of this structure. By analogy to the economical statement that "money make money", it can be said that in the biological evolution "structure makes structure". Indeed, e.g. the structural threshold necessary for development of multicellular forms was exceeded relatively recently, less than one billion years ago. Since then, we observe a violent evolution of tissue organisms, plants and animals, an increase of their level of organisation with the speed of a chain reaction. But this is a completely different story. One should not forget that the whole abundance and complexity of the living world is a resultant of some fundamental mechanisms constituting the essence of life, formed over three and a half billion years ago in the process of biogenesis. In the last instance, we are still a little bit more complex and improved hypercycles.

EVOLUTION OF CONCEPTUAL NETWORK

Finally, we have come to the third of the greatest evolutions in our reality – the psychic evolution. The first two parts of this book investigated the emergence of biological evolution from the physical level, and the physical evolution – from nonentity. Continuing the quest for the uniformity of the world, which constitutes the main axis of this book, I will try to show in the present part how the psychic evolution leading to our self-consciousness emerged from the biological (neurophysiological) level of the functioning of our brains. I will try to show the enormous extent to which our perception and interpretation of reality (expressed to a large extent in our philosophy) is conditioned by the mechanisms of integrating and processing sensual stimuli in the brain, formed in the course of individual and evolutionary development. Eventually, we will see how the linear reductionist sequence – leading form the physical level via the biological level to the psychic level – closes up into a loop, and therefore how, in a sense, (our picture of) the physical level emerges form the way in which our minds work.

The first part of this book devoted to physical evolution I devoted mainly to popularisation of the existing scientific facts, theories and ideas, although I gave my own interpretations to some problems. While describing biological evolution in the second part of this book, I presented and popularised a widely accepted scientific background, but I also introduced many propositions of my own, indicating this fact in appropriate places in the text. The present part, dealing with the evolution of psyche and (self)consciousness, contains mainly my own concepts, although, of course, they are based on the known scientific facts. Therefore, the Reader should be cautious and avoid treating the ideas expressed in this part as commonly accepted opinions. As far as it is possible, I try to separate well-grounded knowledge from subjective interpretations and speculations.

I start from the very fact of the existence of the third, psychic level of our reality, that is the sphere of our subjective experience, and afterwards I try to derive this level and its apparently "autonomous" properties from the functioning of our brains at the biological (neurophysiological) level. I assume that the psychic level differs from the biological level in a similar way as the biological level differs from the physical level. It means that in both cases a higher level emerges in some way from a co-operation of elements at a lower level, and no qualitatively distinct being, like a vital force or non-material spirit, is necessary to explain this. On the other hand, a higher level cannot be described within the terminological framework developed for a lower level.

Taking into account the present state of knowledge, the present part of the book, although based on scientific facts, must be regarded as a semi-philosophical work[80]. Therefore, as it happened many times in the history of mutual relations between science and philosophy, it will probably appear that many detailed opinions proposed here will be verified by the future science as simplified, naive or even false. Further development of neurophysiology and theory of consciousness will probably formulate problems and propose conclusions which we do not even suspect – having at our disposal the present conceptual apparatus. Therefore, particular solutions presented in this part of the book should not be taken too literally; more important are general propositions and conclusions. I wrote this part with the conviction that its fundamental idea – concerning the connotative meaning of concepts as well as the evolution of conceptual network in the ontogenetic and phylogenetic development – is valid. If I am wrong, I hope that the present argument will at least contribute to a clear formulation of the problem.

I will start from the assumption of the material background of psyche. I accept that it is a derivative of the physiological activity of the human nervous system. I also assume (which I will try to justify) that concepts and their complexes – forming a conceptual network[81] – are the content of consciousness (or, rather, that the content of consciousness can be best described as a network of concepts). Finally, I will investigate a presumable evolution of this network in the individual (ontogenetic) and evolutionary (phylogenetic) development of man.

In the below considerations, I provide some terms, such as "concept", "sense" or "denotation", with a slightly different meaning than the one established in the philosophical tradition. At the same time, I avoid defining of them in an extemporaneous and concise manner, as I believe that the entire book constitutes the best context determining the sense of these terms. Generally speaking, I prefer something which can be named a "definition by usage". For a long rung, this seems to me to be a more effective strategy. I realise that such an approach can result in certain inconveniences in reading. Nevertheless, I hope that things will gain clarity when the

[80] I discuss a broader philosophical context connected with the topic dealt with here in the book "Absolute – imaginary reference", which has not been translated into English yet.

[81] As I will discuss later, I do not believe in anything like "qualia" or simple, elementary, atomic sensations and impressions.

mentioned terms will become entangled in various relations with other terms (concepts), which will provide them with better semantic "localisation".

At this point, I would like to clarify shortly one problem (I will come back to it later). In my opinion, there does not exists anything like consciousness alone, without self-consciousness (it does not even make any sense). I other words, I believe that consciousness appears together with self-consciousness and constitutes a part, or an aspect of it. For, to be conscious, one must be a subject, and to be a subject, one has to be aware of his own existence, and therefore, one has to be self-conscious. In the other case, we would be forced to attribute consciousness to any purposeful system (equipped with a set of negative feedbacks), able to "perceive" some properties of the external world, such as a bacterium, or even a thermostat in a refrigerator[82]. This would be an obvious absurd. To be conscious, one has to be conscious that he is conscious – this is in fact what the "self-" in the term "self-consciousness" means. There are also other reasons for the presented opinion[83] – I will discuss them later in more detail. Now I only want to remark that I will use the term "consciousness" in the above meaning. To emphasise the inseparable (in my opinion) dependence of consciousness on self-consciousness, I will frequently use the term "(self)consciousness" to denote consciousness.

[82]Only purposeful systems, equipped with negative feedbacks, are able to "perceive" some properties of the external world, just as a thermostat "perceives" temperature and a bacterium is "aware" of the presence of sugar in its environment.

[83] Half-jokingly: in Polish, we have in fact one name for consciousness and self-consciousness, and therefore, according to the Shapiro and Warth theory (speaking about the influence of ethnic languages on the representention of the world existing in different cultures and minds), we may have predilection to treat these two term as one thing.

WHAT IS CONCEPTUAL NETWORK?

In the present chapter, I will use the common understanding of "concept," which I will precise however in some aspects. I understand a "concept" as a certain unit of a meaning or sense. And because everything that reaches our consciousness and constitutes its content **must** mean something, all of this is either a concept or a complex of concepts. Therefore, I treat a "concept" as a "unit" of consciousness, of thought. In the face of this, my opinion can be called a conception of "conceptual thinking", in opposition to Wittgenstein's "language thinking". All kinds of beings which we can perceive, think, imagine, and finally grasp mentally in reality, dream or a mystical, religious or narcotic trance – all of them correspond to a concept. A tree, a centaur, Julius Caesar, a category, an angel, love, justice – as well as the whole multitude of "objects" so misty and undetermined that we do not find linguistic names for them – all of them are concepts. Concepts are not only "individual beings", as for example linguistic names in a sentence, but also more complex objects, corresponding to entire sentences, conceptions and ideas. There are no beings that are not concepts (in our consciousness). In the presented here, broad understanding of concept, every sense, even the most vague one, is a concept. Therefore, while speaking about consciousness or psyche (more strictly – about their content), we refer, in a sense by definition, to complexes of concepts.

Such a complex of concepts, mutually interconnected by certain determined relations and exhibiting some specific properties, a complex that fills our consciousness, I will call the **conceptual network**. Concepts in conceptual network are not separated by sharp borderlines: some of them pass smoothly into other. Like in the case of hills in a landscape – although there "exist" particular hills, there are no sharp limits between them. Therefore, conceptual network is a continuous formation (contrary to, for example, language, which is composed of discrete names). Two properties decide about the identity of a given concept. One consists in its determination (specification), the degree of separation of the sense of the given concept, the "intensity" of its meaning in the "semantic field". This property decides how clearly and intensely a given concept appears to our consciousness, how obvious and univocal it seems to be. The second property differentiating concepts is the set of connections between a given concept and other concepts in the network, i.e. the question: to which concepts it is related semantically (and in which way), which other concepts "define" this concept. This property decides about what a given concept **means**.

Because the meaning of a concept is determined only and exclusively by its semantic context, reference to other concepts, we deal here with a meaning by **connotation** (contrary to a meaning by denotation, where the meaning would be realised by a direct correspondence between concepts and objects designated by them). The concepts "defining" a given concept are in turn themselves defined by other concepts, and so on; eventually all concepts are defined by all concepts. The whole conceptual network constitutes a proper semantic context for a given concept. Below, I will show how the connotativity of concepts within conceptual network results from the structure of the neural network in the human brain. It should be mentioned additionally that the first of the above-listed properties is secondary with respect to the second one, namely the determination (specification), the semantic intensity of a given concept is a derivative of the number of relations to other concepts, i.e. of the abundance of its connotative connections in the conceptual network.

If the "dimensions" of the semantic space are represented in the form of "significative axes" (fixing a sort of a Cartesian system of co-ordinates), then, concepts are formed by a polarisation along these axes in relation (reference) to other concepts. Examples of simple significative axes are: hot – cold, light – dark, large – small, good – bad. It should be remarked that significative axes, like concepts, are continuous objects, and continuity operates both "along" an axis and "between axes", at a separation point of one axis from another. Additionally, the type of polarisation (opposed ends of axes) in a given case is also determined by concepts. Here, the connotativity of conceptual network becomes manifest as well (concepts are determined by significative axes, while significative axes – by concepts). Concepts differentiated with an aid of only one significative axis have their own clear "anti-theses", as in the case of the pair of concepts: fast – slow. Concepts determined by a great number of such axes do not have, as a rule, a univocal anti-thesis, that means a concept described by the opposite extreme on **each** of those axes – such a combination of meanings simply does not determine any real or at least imaginable object. It would be difficult, for example, to think a thing being **in each respect** the opposite of the elephant (this would have to be a thing which is "small", without a trunk, inanimate, not grey and so on, *ad infinitum*). The existence of significative axes does not, therefore, imply that each concept has its opposite. However, it still remains valid that that senses appear at the moment of "layering" or polarisation of a "semantic vacuum" along significative axes.

What is the relation between concepts and sensations? Are sensations concepts? In my conception, sensations are equivalent to an "activation" of some concepts in the conceptual network. This activation can be caused by stimuli from the external world and enhanced by the phenomenon of attention. However, concepts can also be activated by the processes of thinking, remembering, dreaming, that is generally by endogenous autonomous activity of the brain. The "activated" concepts constitute the current content of our (self)consciousness. Incoming sensations are related to the existing concepts and are possibly incorporated (after adequate processing) as new concepts. This is how sensations are "interpreted" and "understood". The sensory stimuli from the external world activate some concepts much stronger than autonomous processes in the brain, and therefore we perceive sensations as much more clear and "real" than thoughts, dreams and recollections. There is nothing like "pure sensations" or "qualia". The interpretation by the conceptual network of signals coming from receptors determines entirely the "content" and "meaning" of incoming sensations.

Discrete names in a language correspond to these concepts in the conceptual network, that are clear, univocal, and determined (specified) to the largest extent. Concepts that are misty, scarcely palpable and difficult to interpret, do not possess their counterparts in the linguistic layer. Apart from this, language itself, an object apparently so autonomous and discrete, is also – at its deeper level – built of concepts, which lie at the base of linguistic names. This should be understood in such a way that both a symbol itself (e.g. a sequence of letters entering into the composition of a word), its meaning (a set of senses designated by this symbol) and the relation of correspondence between them (that means between a symbol and its meaning) are limited (built) only and exclusively to (of) the conceptual network. Conceptual network is therefore a structure more general and more primeval than language. For this reason, concepts and not linguistic names and sentences are the content of our consciousness. Language originally served the purpose of communication between two psyches, two conceptual networks. The appearance of language subsequently enhanced the very process of thinking itself, since language is a perfect instrument allowing one to operate (manipulate) the whole conceptual network (yet, it is at the same time a part of this network). That is probably the source of some conceptions of "linguistic thinking". Anyway, it would not perhaps be possible for higher forms of thinking and (self)consciousness, being a realm of man to originate without language. However, this does not change the fact that concepts are the fundamental "substance" of psychic processes.

The difference between the widely accepted opinion on the mutual relation between language and concepts and the conception presented in this chapter is presented in Fig. 3.1. In the common version, language is a certain system of discrete signs (characters) which denote various (also discrete) aspects ("facts") of the world. Concepts, or meanings, constitute the semantic "lining" of particular linguistic names, determining their sense. Language is here superior in relation to concepts, which play only the role of a part of the whole linguistic system. In the conception proposed by me, this fundamental scheme does not change as a matter of fact. However, the sphere of concepts grows here to comprise everything that lies on the "subjective side of reality", and therefore not only the ordinarily understood meaning of linguistic names, but also other, misty concepts that do not possess names, as well as the picture of the facts of the world created in our mind, and finally the whole structure of language together with its discrete names and the relation of denotation. Moreover, the "concept" itself is also a concept, and therefore exhibits the property of self-applicability. Simplifying a little, the apparently "sharp" (discrete) linguistic signs, as well as the seemingly univocal relation of denotation are – at a deeper (fundamental) level – simply built of naturally continuous concepts. Different fragments of conceptual network are everything we are able to reach directly with our mind. In this formulation, conceptual network appears to be a superior object, much more general than language. The presented scheme emphasises the fact that the relation between the facts of the world and their representations in the mind (conceptual network) is not a simple two-element relation, a one-to-one correspondence, but a result of a gradual, initially random adjustment of the mesh of conceptual network to different aspects of the objective reality, occurring in the course of biological evolution and individual development.

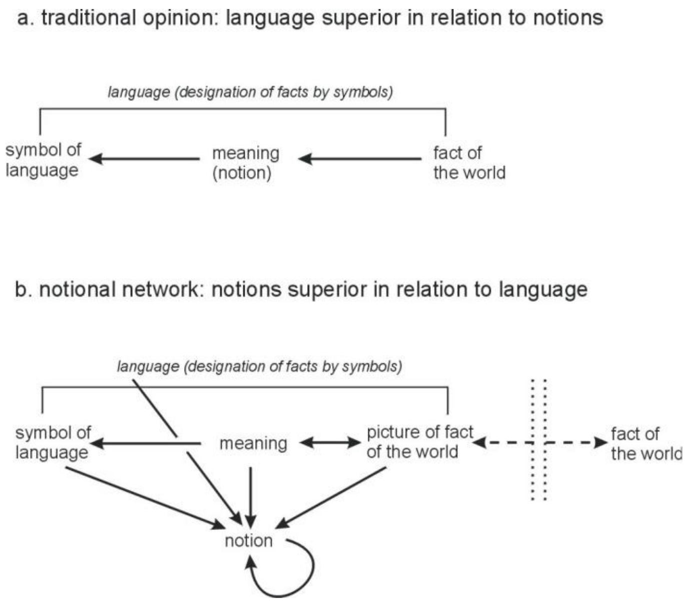

Fig. 3.1. Comparison of the traditional opinion on the relation between concepts and language with the opinion represented in the present work. (a) Traditionally, language is regarded as a complex of signs denoting facts of the world. Meanings – which can be identified with concepts – are hidden under names. (b) In the present work it has been assumed that concepts are the fundamental "component" of the subjective side of the world. They form the building blocks of our picture of the world, language with its elements (names and the related meanings) as well as any other meanings that are accessible to our mind, including what is colloquially called concepts. The relation (based on biological functionality) establishing correspondence between facts of the world and their picture in the mind is in a sense bi-directional, since the already possessed picture of the world co-conditiones (together with the objective reality) its own development.

However, at this point I looked slightly too far ahead. Therefore, let us go back to the determination of the "phenomenological" essence of the conceptual network, without entering deep into the conclusions resulting form its biological genesis. Of course, the attention of consciousness is at a given moment focused only on a very small fragment of conceptual network possessed by us. Nevertheless, the understanding of concepts used at a moment is based on the entire network. Conceptual network is not a homogeneous creation. It contains better or worse differentiated and internally coherent areas, such as scientific disciplines, religion or finally the sphere of common (ordinary) concepts. We know very well, at least from introspection, that we did not obtain our present conceptual network – with the whole abundance of knowledge, information and ability of thinking – at the moment of our birth; on the contrary, at the very beginning we had almost no concepts at our disposal. Afterwards, our lifetime witnesses growth of our knowledge about the world, development of our view of life, increase of our understanding of different aspects of reality, products of culture, science and art, change and enrichment of the meaning of "the same" linguistic names. It is obvious therefore that our conceptual network had to evolve in the individual (ontogenetic) development. We know that the species *Homo sapiens* originated in the course of biological evolution from animal ancestors. The long sequence of our progenitors contains apes and monkeys (some people talk about having only a "common ancestor with monkeys", which is obvious hypocrisy; if this ancestor lived presently, it would certainly be classified among monkeys), crossopterygians as well as – to reach yet earlier times – some undetermined unicellular organisms, to mention just a few. While one could detect some germs of (self)consciousness in monkeys, then it would be difficult to say this about protozoa. Therefore, consciousness, together with the underlying conceptual network, had to undergo a development during biological evolution. Finally, while (self)consciousness created culture, science and religion, then, it was itself to a great degree formed by them. Therefore, the development of conceptual network had to proceed along a parallel route with the development of civilisation.

Analysing the essence, structure and development of the conceptual network, we analyse what psyche and (self)consciousness is. Analysing the neural network, we learn something about the conceptual network. These two assumptions constitute the basis for the further part of my reasoning.

To avoid any misunderstanding, I have to make the following reservation: I do not treat conceptual network as an object which is in any way absolute or "objective", as philosophy does in relation to, for example, spirit and matter. I treat it only as a convenient **model**, allowing to describe the interesting aspects of the discussed problem.

NEURAL NETWORK AS BASIS OF CONCEPTUAL NETWORK

Where does the connotative structure of conceptual network come from? As I remarked earlier, I assume material background of (self)consciousness. I accept, therefore, that the psychic sphere is in some way a derivative of the physiological activity of the brain. The functional unit of the brain is constituted by a neural cell (neurone), which together with other neural cells form a neural network. The activity of a single cell within this network consists in perceiving stimuli (impulses) from other cells through appendages called dendrites, in processing these stimuli with participation of information already possessed by a cell (memory), accompanied by possible transfer of a transformed impulse to other cells through an appendage called axon. The axon of one neural cell is connected to dendrites of other neural cells through synapses. It is an obvious thing that different neural cells, fulfilling different roles, entering into the composition of various centres in the brain, must differ one from another. Therefore, what do these (functional) differences consist in? We can distinguish two separate categories.

The first category consists of internal differences. Each neural cell has a characteristic logical function, which transforms the stimuli reaching the cell into signals which are sent further on (to yet other cells). Speaking more simply, each cell makes a different use of a given set of received impulses. One cell enhances a given combination of signals, sending a strong signal at the output, while another cell does not react to the same input combination by sending any impulses. As the input set (combination) I understand the set of stimulated dendrites and the time sequence of stimulation, while by the response at the output refers to that whether an impulse is sent further by the axon and in what manner (I mean here a time modulation). The manner some signal at the output is attributed to a combination of signals at the input constitutes the specificity of a given neural cell (it should be stressed that the response of a particular neurone, and chiefly a complex of neurones, to a given set of stimuli is not strictly determined; it is possible only to indicate probabilities of some reactions).

Let us consider the idea of a specific transformation of stimuli by a neural cell in a very simple example. Let us assume that our cell has only two inputs (dendrites) A and B, one output (axon) C, and that these inputs and outputs can assume only one of two states: 1 (signal present) and 0 (lack of signal). It should be remarked that the identification of dendrites and axons with inputs and outputs is a certain simplification; one should rather speak about synapses situated on dendrites and the axon. In this (very simplified) case, our cell can execute one of a few logical functions. Let it be: if not A and B then C; in other words, a signal on the output (axon) C will appear then, and only then, when the dendrite (input) B is stimulated and the dendrite (input) A is not stimulated. Another cell can realise a different logical function, e.g. if A or B then C. Of course, the quantity of possible functions transforming signals from inputs into the signal at the output increases dramatically, together with an increase in the number of inputs (dendrites and synapses localised on them) and in the variety of their states (intensities of activation). The kind of the function realised can be changed under the influence of past experiences (stimulations). We call this memory. Memory concerns both particular neurones and their complexes. It can be related to a formation of new connections or to a selection of the existing connections (synapses) between dendrites and axons, to changes in the "weights" of synaptic connections[84], or (presumably) to changes in the information recorded in proteins and nucleic acids in the body of a cell. Its physiological (or maybe biochemical) basis is not fully understood yet.

Does the logical function realised by a given neural cell determine univocally its role and significance within the whole neural network? Does it fully determine the specificity of this cell? And if not, then, is this property at least the decisive factor? The answer to both questions is negative. The connections via dendrites and axons with other neurones (complexes of neurones) constitute the primary factor, determining the "meaning" of a given neurone (or a complex of neurones) in the functional "context" of the neural network. It should be emphasised that there can be many neural cells which realise a given concrete logical function. These cells can fulfil different roles in different centres, participate in various psychic processes. On the other hand, there exist (probably) only one cell in the whole brain, connected to exactly these-and-not-other cells (cell complexes) in the exactly such-and-not-another manner (in the functional sense). We see therefore that internal differences between neurones (in the character of the logical functions realised by them) are completely insufficient to determine the specificity and individuality of their role in the brain[85]. External differences are at least equally important. They concern the functional

---

[84]The greater the weight of a given synaptic connection is, the easier it is to for a stimulation of this synapse to cause the stimulation of the whole neural cell and, in consequence, transmission of a signal further on by its axon.

[85]E.g. a neuron (or a simple network of neurons) responsible for "seeing" a movement (compare the below example with a frog) transfers an impulse further if before a while it did not receive signals from a receptor (photosensitive cell), and now it receives such a signal or inversely, whereas it does not react to an uninterrupted sequence of signals or to absence of signals.

surroundings of a given cell in the neural network, its significative context, that is this which other neurones this cell is connected to and in what way. The logical function of such a cell taken out from its surroundings looses its sense and becomes an empty structure which, as it is able to carry any real content, carries no content at all. The information "contained" in a given neurone as an element of the neural network is therefore the sum of the information about its internal "logical structure" and the information about its functional "localisation" within the neural network (and therefore, in fact, about the structure of the **entire** neural network). The first aspect is connected with local memory; the second – with the global memory, which may be partially a dynamic memory, consisting in a circulation of some signals in neural circuits and in concrete values of the weights of particular synaptic connections. Only this first kind of information and memory can be preserved after destruction of the structure of neural network, and therefore possibly recorded and transferred genetically. Therefore, only the structure of single neurones and relatively very simple neural networks can be inherited. The structure of more complex networks, primarily those related to human psyche, must therefore **evolve** during the individual development through the acquisition of new experiences.

A simple example will illustrate the idea of relativity of the meaning of the internal logical function of a given neural cell in relation to its "semantic environment". As I mentioned before, a neural cell is connected to other neurones through its dendrites (inputs) and axon (output). Moreover, it can receive impulses from receptors, which react to stimuli from an environment (e.g. chemical, auditory, visual stimuli), as well as send signals to effectors (muscles, glands *etc.*). Let us imagine a fictitious water organism of several cells, containing only one neural cell which realises our logical function: if not A and B then C. Such an organism is presented in Fig. 3.2. This organism can feed on organic substances present in water, while in their absence, it produces the building material and energy in the process of photosynthesis, for which light is needed. Let us assume that input A of our neural cell is connected with a receptor sensitive to the presence of organic substances in the water, input B with a receptor reacting to light, while output C sends impulses to a vibraculum – a movement organ which is able to drive the whole organism in the

However, if we take such a neuron out of its appropriate place in the neural network and put it in a centre responsible for feeling pain as well as for related unconditional reflex (a withdrawal of a hand after touching of a hot object, for example), this would lead to lamentable results. After putting a hand into fire, we would feel pain for a fraction of a second and after this – nothing at all. Having a neural system working in such a way, we would be eliminated by evolution.

direction of a source of light. The organism in question would feed on organic substances when they are accessible, while in their absence, it would flow in the direction of a source of light. The functional sense of such an action is clear and obvious. However, this sense consists not only in the form of the logical function realised by the neural cell, but first of all in its significative context, in an attribution of a concrete meaning to its inputs and outputs. In a complicated neural network, this meaning is determined by the fact which neurones, and in what manner, a given neurone is connected to. The function of these "neighbouring" neurones is determined, in turn, by their own "significative context", and so on, until the considered network of connections spreads to the system, comprising the whole neural system as well as all receptors and effectors connected to it. Only this system fixes, in the ultimate instance, the adequate functional context, its reference system. Any change of this context changes to some extent the functional identity of a given cell. If in our example of the water organism we substitute the presence of gold in the environment for A, the presence of a predator for B, and going into its direction for C, then our poor creature, in the absence of gold in water (which is a rather typical case) would go into the mouth of the first predator it meets. It is difficult to claim such a behaviour to be purposeful and profitable.

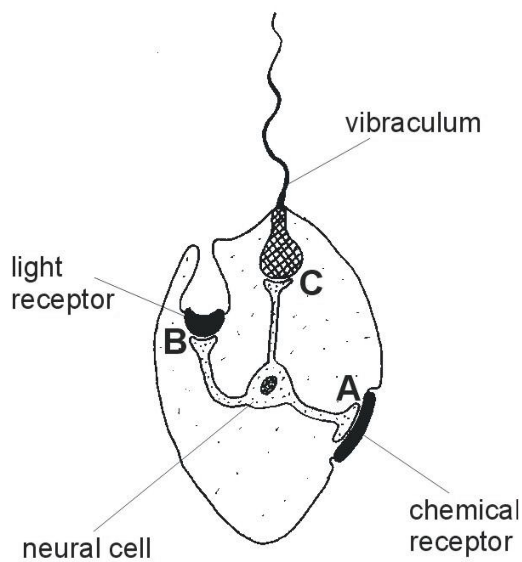

Fig. 3.2. Imagined aquatic organism with only one nervous cell. Input A (dendrite) is connected to a chemical receptor of nutritional substances dissolved in water, input (dendrite) B is connected to a receptor of light, while output C (axon) – to a mobile vibraculum. The nervous cell realises the logical function: "if not A and B then C". As a result of this, in the absence of

nutritional substances and in the presence of light the organism goes in the direction of the light source.

This analogy applies perfectly to more complex neural networks. It leads to the conclusion that the structure of complex neural networks cannot be inherited. For the genetic record can concern a functional structure of a neurone, but not the complex network of connections which is mostly acquired during individual development. The amount of information contained in the brain of an adult human exceeds most probably the amount of its genetic information. This does not mean that nothing is inherited in the structure of the neural network. In primitive organisms, almost their entire neural network together with its "significative context" is determined in the genome.

In higher organisms, including man, the following functions are probably encoded genetically: the manner of integration of the data coming from receptors (e.g. their ordering in functional structures which we interpret afterwards, *ex post*, as spatial, temporal or causal relations), organisation of the centres of the brain responsible for vegetative "functions" (feeling cold, hunger *etc.*), and also **predispositions** of neural cells and neural networks composed of them to a selective formation and removal of connections, and also to acquisition of memory. However, the neural network responsible for a complicated system of behaviours, being the background of memory and (self)consciousness, is formed during an individual life in the process of learning.

What are sensations according to the neural-network conception underlying the conceptual network? In my opinion, the sensation of blue is "blue" because the neural "structure" (the group of neurones connected functionally in an appropriate way) corresponding (equivalent) to the concept of blue is connected (in fact this is much more complicated) to receptors in the retina sensitive to the "blue" range of the spectrum of visible radiation. During our life, this "structure" learns that a part of perceived objects sends signals to the receptors of "blue" and our conceptual network learned that "blue" is the common property of the objects which stimulate the receptors of "blue" (and only these). Therefore, the "blue" in the conceptual networks co-defines all blue objects, while all blue objects co-define the blue (again – meaning by connotation). We would not understand or even "see" the "blue" without all these blue objects we come across in our life and without their incorporation (as representations) into our conceptual network. Therefore, the blue is by no means a "simple", independent, autonomous being. What we really receive (through our eyes) is electromagnetic radiation of a certain concrete wavelength. We do not perceive the "blue" itself (because nothing like this exists), but only the very core of the concept of blue, when the neural sub-network underlying the concept of blue is activated by signals from the "blue" receptors (and also by impulses generated internally within the brain in dreams, thoughts or remembrances).

The very essence of the perception of blue is determined exclusively by the neural context of the concept of blue. We do not experience the "blue" because "blue" receptors send different signals than "red" receptors, but because the neural structure corresponding to the "blue" was stimulated during our whole life by blue objects, "acting" via "blue" receptors. Let us imagine that we cut the neural connections between the "blue" receptors and the "blue" centre in the brain of some patient, and connect the "blue" centre with "red" receptors instead. Then, we show to our poor patient a red rose. What will he see? A **blue** rose. This is because the subjective "blue" is not what stimulates "blue" receptors, but what activates the "blue" centre in the brain. What is blue, then? Blue is only the common property of all objects which – during the development of the brain – stimulated "blue" receptors, the property which was represented in the brain in the form of the "blue" centre. The "blues" of different humans differ slightly one form another. The "blues" of a human and a fly differ much more (a fly has different receptors and much smaller neural network). Why do we experience the "blue", and a camera does not? Because a camera has not got the concept of blue, defined by billions of other concepts, which can be activated by "blue" stimuli, and therefore is not able to interpret and "understand" these stimuli as just "blue" ones.

We see from all the above considerations that conceptual network is inseparably related to the neural network. A comparison of conceptual network and the neural network leads to one important conclusion. In both cases, the meaning, the sense of an element of the network (a concept or a neural cell, or a group of such cells) occurs by way of connotation, that is by reference to other elements. The general structure of both networks is therefore analogous. This seems to be in some measure understandable, because conceptual network is a "result" of the activity of the neural network. Nevertheless, this statement has far reaching implications, concerning the basis and nature of human consciousness (psyche). The structure of the neural network, imposing to some extent the general structure of the conceptual network, significantly restricts the range of possibilities concerning the essence of (self)consciousness. While recognising the general similarity between conceptual network and neural network, one should at the same time avoid far-fetched analogies. In particular, it would be naive to suppose that one single neural cell (or at least a simple complex of such cells) corresponds to one concept. A concept is probably a derivative of a complicated functional structure of the neural network, which does not have to be strictly localised spatially within the brain. A concept is a dynamic structure, changing under the influence of different received impulses. The structures at the basis of different concepts partially interpenetrate each other. The details will be probably explained in the future by science (neurophysiology and the theory of consciousness). Fortunately, they are not crucial for our further reasoning.

There arises a question, however, namely: how can conceptual network be a continuous object, if the underlying neural network is composed of discrete neurones. The answer is a huge number of neurones building even relatively simple neural networks, as for example in the visual cortex. The complication of the networks responsible for the processes of thinking and consciousness must be yet much greater. As spots on a photograph are considered as continuous objects, although they are composed of discrete grains, because these grains are by a few orders of magnitude smaller than the dimensions of spots, so the discreteness of particular neurones is lost at the level of the functioning of the brain and psyche as a whole.

EVOLUTION OF CONCEPTUAL NETWORK IN INDIVIDUAL DEVELOPMENT

It is time to pass on to the evolution of conceptual network in the individual (ontogenetic) development of man. As I mentioned above, this evolution is parallel to the evolution of the neural network. However, we will not be interested now in the details of the latter. I will confine myself to the fact that this evolution consists in a formation of new neural connections (appendages and synapses), decay of a part of existing connections, change of the "weight" of existing synaptic connections, biochemical record of memory, origination of impulses circulating in neural circuits, selection of existing circuits and perhaps in other processes, yet unknown to us. All these phenomena occur in the brain, the general plan of the structure of which as well as predispositions to the above enumerated changes are encoded genetically.

At the moment of birth, a man possesses almost no conceptual network. However, he has germs for its development, centres of crystallisation of this network (if they already constitute some simple conceptual network is to a large extent a conventional problem). They are primitive "significative axes", determined by purely biological meanings or "values". Such significative axes are for example: "warmth-cold", "satiety-hunger", "feeling of safety – lack of it"[86]. They come from the fact that we have appropriate centres in our brains that receive signals from adequate receptors (and other centres) and inform about the state of affairs the primeval "evaluating system" – the centre of pleasure[87], stimulating this centre or not, depending on the situation (the fulfilment of basic biological needs[88]). The above-mentioned axes differentiate stimuli arriving

---

[86]This "feeling of safety" in a new-born child has very little to do with an obiective assessment of a situation. It reduces to the need of perception of a "factor enssuring safety", e.g. the voice of a mother. For a baby, the feeling of safety is satisfied by each stimulus associated with a satisfaction of hunger, warmth and so on. The ability to associate such, otherwise neutral, stimulus with the satisfaction of biological needs is ensured just by the primitive significative axis "feeling of safety". In our ancestors, this had much more significant implications for a new-born child, for example a defence against predators, parasites and so on.

[87]Which can be identified more or less with the reward system in the brain, in which the neurotransmitter dopamine plays the main role.

[88]In adult humans, the reward system can be stimulated also by a satisfying of non-biological needs, for example cognitive, ethical or aesthetical needs, although they could nevertheless have biological roots (in fact, this is quite trivial, since we evolved from purely biological creatures). This testifies to the thesis that humans have achieved a level which is higher than the biological

from the surroundings, segregate them into some categories, being the germs of first concepts. The segregation and differentiation occur **in relation to** already existing significative axes. The originating concepts form in turn the basis for new semantic axes.

The existing significative axes constitute a sieve (selector) and interpreter for stimuli coming from receptors[89]. Sensations received by sensory organs are only a bunch of unordered signals, which is a complete nonsense, when an interpretative key or algorithm is not applied to it. A great majority of animals have encoded genetically many, if not most of such algorithms decoding sensory stimuli. In man, these genetic predispositions are reduced to the mentioned semantic germ, the centre of crystallisation of senses: a few significative axes which are afterwards developed during experience accumulation. Origination of concepts and their attribution to the phenomena of the external world occur by induction, via a multiple coincidence of similar sets of stimuli. If a given system of stimuli causes many times the same effect, and enters into the same interactions with the existing conceptual network, then, it becomes incorporated into this network as a new concept. Therefore, the sensations from the external world are the "substrate" for newly built concepts. These sensations are initially confronted only with the primary significative axes, and then, with the already possessed conceptual network, to which they are "referred". In this way, the meaning of new concepts is determined and defined by the already existing complex of meanings contained in the conceptual network.

As conceptual network develops, there originate new, secondary significative axes on the basis of the already existing concepts. Subsequent complexes of stimuli are located as new concepts in the network in relation to the existing concepts. In this way, projections (images) of individual, "real" objects (aspects, facts) of the external world are created in the conceptual network. I will call such concepts, directly corresponding to simple "facts" of reality, the primary concepts. The general, abstract secondary concepts are created (in a sense at a higher level) similarly like the primary concepts, as a result of a "perception" by the mind of many similar sets of concepts which are particular in relation to them (these can be both primary concepts as well as secondary concepts, lower in the "hierarchy of generality").

level, namely the psychic and cultural level. Therefore, I do not agree with the protagonists of an extreme form of sociobiology.

[89]These signals undergo a preliminary integration by appropriate centres in the brain. However, the integrated stimuli have to be "understood" by "higher" centres, being the basis of (self)consciousness.

Let us investigate the initial stage of the evolution of conceptual network in a new-born child. As I mentioned, it has at its disposal the simplest significative axes of a purely biological meaning. Now, if the voice, smell and touch of the mother is associated by an induction (multiple coincidence of stimuli) with warmth, satisfying of hunger and feeling of safety (for example, due to the fact that the first regularly heard voice in its life is associated by a new-born child with a feeling of safety), then, such a combination of auditory, gustatory (milk), aromatic, tactile stimuli will become a germ of the concept "mother".

As experiences accumulate, the concept of a mother is enriched with new elements, supplemented by visual stimuli, related to other concepts, made more precise. Let us accentuate the fact that "the same" concept "mother" differs in a new-born child significantly from this concept in an adult human. The "centre of crystallisation" of first concepts are impulses: satiety, safety, cognitive impulse. The last one makes a germ of the concept "toy" from an attractive (e.g. colourful) object yielding to manipulation. In a new-born child, conceptual network is very poor – those few concepts fill its entire world. At the very beginning, there are maybe only two concepts. Simplifying a bit, let us call them: "mother" and "lack of mother". Afterwards, still new and new stimuli are perceived, these concepts undergo differentiation into many derived concepts, for example "mother" divides into "true mother", "nanny", "father" and so on, as the baby starts to distinguish particular persons satisfying its hunger and need of safety. This sounds a little bit grotesque, because the use of linguistic names to describe these hardly-formed concepts is not appropriate. Since, as I mentioned earlier, linguistic names correspond normally to concepts which are best determined, attributed with the highest intensity of the "semantic field". Apart from this, while discussing the psyche of a new-born child, we in a sense go introspectively back to times when we did not yet possess language (this is rather obvious, because language is also a form of the conceptual network, and a quite complex one).

The problem whether the development of the psyche of a new-born child looks exactly as it was presented above is not important. A strict description is not possible at least because of the very weak congruence between originating germs of concepts and the sphere of language. Therefore, the point here is rather to present the general principle, and not to give an accurate account of facts. The following are the general properties of conceptual network development in question: "stratification" of the existing concepts into better-precised concepts, location of new concepts in conceptual network by induction, multiple perceptions of sensations or more detailed concepts, and finally, specification of already existing concepts. These processes are strictly connected with each other. They in fact constitute various manifestations of the same process. As the meaning of concepts within conceptual network is realised by the relation of connotation, each concept in this network is "defined", more or less directly, by all other concepts. The incorporation of new concepts to the network results therefore in better specification of the already existing concepts. The "stratification" of hitherto uniform concepts proceeds as a result of the appearance of new semantic axes. The appearance of such axes is equivalent to the appearance of new meanings, and therefore new concepts. Already existing concepts are referred to these new axes, which results in their better specification (determination).

The set of reactions of a baby, its repertoire of behaviour is as poor as its representation of the world determined by the possessed network of concepts. Simplifying the matter, it can be said that the system called "a baby" is able to manifest one of two states at the output (please, forgive me this cybernetisation of a defenceless suckling). Either it does nothing, when all its needs are satisfied, or reacts with a scream, when it feels hunger, is wet or feels the mother's absence and the corresponding feeling of lack of safety. The choice of one of these states depends on the stimulation of the primeval "evaluating factor", i.e. the centre of pleasure (reward system in the brain). The evolutions of the degree of complication of the perceptual system, of the picture of the world (conceptual network) as well as of the system of behaviours, proceed along parallel lines during ontogenetic development. In particular, both the ability of perceiving various aspects of the world and of reacting adequately requires the possibility of their recognition, interpretation and "understanding" (both in the biological and psychological sense) based on the possessed network of meanings (conceptual network).

The first main trait of the evolution of conceptual network in the ontogenetic development of man is certainly its quantitative increase, connected with the origination of new concepts and specification of the already existing ones. The development of conceptual network does not start, since it cannot, from a complete semantic emptiness, because some meanings have to fix the sense of the newly created concepts. However, the few inborn, biological significative axes constitute the germ of psyche which is so extremely small in comparison with the enormity of conceptual network of an adult man that one can practically say that the latter starts its evolution from nothing at the moment of birth.

I regard the moment of birth as the beginning of the development of the conceptual network, and therefore of the origination of psyche and (self)consciousness, because it is connected with a violent exposition of an embryo to signals coming from the external world. An embryo in the mother's uterus, connected to her circulation of blood by a umbilical cord and wrapped in the placenta, practically does not receive any stimuli[90] from the environment, and therefore is devoid of the "substrate" for the formation of concepts. Its environment is extremely unchangeable. It ensures constant temperature, unceasing inflow of nutritional substances and outflow of needless products of metabolism. Therefore, while it is difficult to determine when an embryo becomes a man in the biological sense, it begins to gain the attributes of humanity in psychic aspect only after its birth. On the other hand, the "humanity" of an embryo can be fixed by its ability to produce psyche. However, the same can be said about a spermatozoon and egg cell before they join into a zygote.

I would like here to separate clearly conceptual network which – in my opinion – is practically absent from a new-born child (although, as I mentioned, the problem is partially conventional), or formed only to a small degree, from the inborn "significative axes" and the mechanisms which integrate stimuli into spatial, temporal, causal, *etc*., structures. The former is already a representation of the external world, although very poor, while the latter is a set of predispositions, a frame for the formation of the picture of this world. The difference is slightly similar to the difference between the genetic information of some organism, together with the mechanisms responsible for the formation of spatial and functional structures of this organism (genetic code, translation, transcription, regulation of embryogenesis), and a ready organism itself, which originates due to interaction of the genetic record with the physical environment, for the genetic information "in itself" means nothing – it has a sense only in relation to the physical reality from which (and in which) this organism originates. Genes can contain encoded sequences of enzymes which catalyse reactions leading to the synthesis of chlorophyll, but the fact that this compound is green and enables the process of photosynthesis is a property of the physical reality.

The contribution of the environmental factor is much greater in the formation of conceptual network than during the individual development of a biological organism, whereas the influence of

---

[90]Some exceptions seem to be changes in the level of hormones and nutritional substances in the blood or the beating of the heart of the mother. However, the degree of diversity and complication of these "stimuli" is so small in comparison with the richness of sensations received after the birth, that the former can be neglected without the risk of excessive simplification.

inborn mechanisms – much less determined. A plant, having at its disposal a certain limited set of simple chemical substances and facing a set of physical conditions, will always produce chlorophyll, when at the same time the whole diversity of the remaining factors remains unimportant. With the optimal temperature, composition of soil, amount of light and so on, a plant will develop fully. If some of these factors diverge from the optimum, the plant will grow dwarf-like or simply die. This is the limit of its flexibility. On the other hand, the quantity and diversity of properties of an environment (as various sensorial stimuli) that are important for a developing conceptual network of a man is much greater, which increases immensely the abundance of possible ways of formation of this network.

In a sense, the external world (at least some of its aspects) is in some way "imprinted" both in the mechanisms of reading genetic information and embryogenesis, as well as in the cognitive mechanisms of our brains. The process of imprinting was performed by biological evolution, "responsible", after all, for the possibly most efficient functioning of living organisms in the physical reality. However, such an inborn record must still undergo realisation. In the process of embryogenesis, this means building of appropriate structures of an organism from physical matter, while in the case of the development of the conceptual network, it means formation of this network from the "substance" of incoming sensations. As I said before, the analogy is not accurate, because the diversity of the accessible physical elements of an environment is rather limited, while the number of possible sensations – enormous. Of course, sensory stimuli are also of physical character, but their perception at the neuro-psychical level is completely different (because of the degree of complexity of the structure of sensations) from the "perception" of physicochemical properties of the surroundings at the biochemical and physical levels of functioning of living organisms.

However, the general rule is the same: just as an organism originates only after "confrontation" of the genetic record with the physical reality and builds its body from the components of this reality, so conceptual network forms its "substance" from incoming sensations and for this purpose it uses some algorithms in the form of inborn mechanisms of stimulus integration. Anyway, only the simplest and most general instructions for the development of conceptual network are inborn (as primary significative axes). The rest originates gradually as a part of this network, determining the frame for the directions and character of its further development. Generally, sensations are absolutely necessary for conceptual network to originate and develop, since it is "built of" these sensations to a great degree.

The appearance of some "surplus" in relation to a purely "visual" presentation of the world is the second aspect of the ontogenetic development of the conceptual network. While primary concepts are simple derivatives of processing the stimuli reaching our sensory organs, then, the secondary (general, abstractive) concepts originate on the basis of primary concepts or other secondary concepts and therefore their connection with the world is more indirect, that means, at least potentially, less restricted. Therefore, secondary concepts can be separated from the real world (e.g., the concept "angel"), although they do not have to. This surplus in conceptual network is inseparably connected to processes of thinking, that is to the autonomous dynamics, the activity of the neural system. Nothing else but this activity – related to the integration, association and co-ordination of data from receptors, memory records and signals sent to effectors – leads to the origination of secondary concepts.

If we acquire practically the entire conceptual network during our individual development, it has to be formed in the process of learning. One can learn in many ways: by observation, by trial and error, by imitation, finally through the mediation of linguistic transmission, oral or written. The accessibility of particular ways of learning to a man depends on the degree of development of its conceptual network. A baby, who possesses only its germinal form, is capable only of passive observation of the surrounding reality. Elements of this reality gradually acquire different meaning for the baby, depending on their connection with the fulfilment of purely biological needs (hunger, safety *etc.*). Having acquired some orientation in the external world, it is possible to experiment with it, for example to move oneself in it or move various objects. This helps to develop the spatial representation of the world and to learn its properties. The learning by trial-and-error method is in fact a method of active perception, where a baby "observes" its own, initially purely accidental actions, and the reaction of the world to them. Gaining knowledge by imitation of parents or other adult people already requires some understanding of the properties of the world, and therefore, it requires operating secondary concepts.

Mastering of a system of artificial symbols, that is an ethnic language, increases greatly the efficiency of the processes of learning. The language, itself a part of conceptual network in a human mind, extraordinarily improves the use of this network. Linguistic names help to "identify" concepts. The rules of syntax are useful in their appropriate ordering. Language, as a common social convention, allows for mutual "translation" of conceptual networks of different individuals. New information, concepts, are located in the existing semantic and syntactic structures of a language (and thus, of conceptual network underlying it), and therefore there is no reason to learn them by trial and error. This accelerates marvellously the process of learning and development of the conceptual network. An additional advantage of language is of course the fact that it allows to transfer information at a distance, both in time and in space.

A given language not only stimulates the development of conceptual network and the picture of the world contained in it, but also, shapes this picture in some manner through its structure. The structure of the world is unique, while each language has a slightly different structure. The structure of a language orders in a given concrete way the conceptual network, and therefore the representation of the world formed within it. There are known languages that do not know verbs in our meaning, and therefore the meaning of temporality in their picture of the world is different than in our picture. Languages that have names only for two colours or three numbers in an obvious way condition the seeing of the world. In still other languages, the names of objects are complexes of properties attributed to them, for example, the names for both a hand and a tree have a segment standing for a ramification. A different structure means a different logic of a language and conceptual network, as well as of the world seen through their prism. Therefore, the ultimate shape of conceptual network depends to a large extent on its ethnic language of origin.

Linguistic names are credited with a greater meaning than they have in reality. Hypostases are the most flagrant example; however, every linguistic name is in principle a hypostasis. Linguistic names (due to their discrete nature) frequently determine and separate sharply aspects of reality, that are not so well separated and determined at all. This primarily concerns general terms, but also names of individual beings. Each human, regarded as an individual being, changes in time, and therefore is somebody different at a given moment than a minute, week or twenty years earlier. There occur physiological transformations in his body, experiences accumulate in his brain, he translocates in space (that "the same" man preserves its identity in different places in space is not obvious at all!). Due to this, the attributing to him an unequivocal name constitutes a misuse. In other words, the name "John" applied to a given concrete man is not at all a name of an individual being, but a general name, a category which refers to the set of "states" of a person called John in different temporal moments, even after an amputation of a limb, etc. At the same time, attribution and classification of these states by the category "John", and first of all, the very fact that the category is distinguished, is a property of our mind. After all, one could emphasise the occurrence at one moment in time of "temporary states" of different persons, and create from them a fundamental category: "temporary human community", while at the same time secondary significance will be attributed to the property of continuity of states of one person in subsequent moments of time. Therefore, it is completely arbitrary that we regard John as an individual being[91].

Such a state of affairs is due to the mechanisms in our brain which integrate sensations in such-and-not-another way. Physicists, describing the behaviour of a set of particles (atoms, molecules), frequently use the so-called phase-space rather than the ordinary three-dimensional space. The phase-space for one particle is six-dimensional: three dimensions determine space co-ordinates of a particle, while the remaining three dimensions – the co-ordinates of the vector of its impetus. For a system of many particles, the number of dimensions describing this system in the phase-space is equal to the number of particles multiplied by six. Such a space is unimaginable for the human brain (like, for example, the curved four-dimensional space-time!). We predict its properties using the appropriate mathematical formalism. However, physicists did not invent the phase-space to make ordinary mortals feel stupid, but because it is much more **useful** for some purposes than the ordinary three-dimensional space. Let us remember that usefulness was the main reason for which evolution shaped our seeing of the world in categories of space and time. Therefore, there is in principle no obstacle to suppose that some unknown creatures perceive the world in the categories of the phase-space. This means that the phase-space would be for them a primary, "visual" manner of seeing reality, while they would have to make a considerable effort of mind and mathematical skills to familiarise with the Euclidean space, as we familiarise ourselves with the phase-space. The "world" of such creatures (this means: its representation in conceptual network of those creatures) would be completely different from our "world".

The so-called "facts" of the world, and also the general way of seeing them, are a derivative of our cognitive apparatus (language being its part). The fact that even the most fundamental

---

[91] The fact that some categories seem natural to us, while others seem so grotesque and contradictory to the common sense, results from certain cognitive structures in our brains formed during biological evolution. Evolution consolidated such structures that ordered the sensations received by us in categories useful for survival of the species. In social relations, for example, such a category is the concept of an individual. This name functions very well in relation to people. However, when we apply it to a colony of coelencerates, composed of individuals partially separated one from another, it looses its clearity.

properties of human seeing of reality do not have to be obvious at all can be easily seen when man is compared with animals. Let us take the faculty of vision. Perception of an image is the fundamental property of our seeing. If we look, for example, at the picture *Mona Lisa*, then we see in it a pattern of some colourful spots, forming a portrait of a certain female in black, and, first of all, the most famous and puzzling smile in the world. Of course, the fascination with a smile of a woman in mourning has first of all cultural roots (related primarily to the Mediterranean culture). The fact that in the picture we perceive a person at all, results from the fact that we are humans – a Martian would not probably be able to read even this information. However, at least one thing seems unquestionable, namely, that there "objectively" exists in the portrait an arrangement of colourful spots which can be received by each organism, having at its disposal the faculty of vision (such organisms could eventually possess an organ of sight with a smaller resolution, which would make these spots less clear and distinct).

Let us look now at *Mona Lisa* with the eyes of a frog. A frog perceives first of all movement, and therefore a portrait hanging on a wall would remain to it a homogeneous light (or dark) background. On the other hand, a fly flying by against the background of the picture, that would be hardly noticeable for us, would release a sudden reaction in the brain of a frog as the only object perceived! The discussed amphibian has an eye which is built similarly to our eye. It is sensitive to a similar range of electromagnetic radiation, focuses a picture on the retina with an aid of a lens (this similarity does not have to take place at all in the case of creatures from other planets – it certainly does not take place in the case of insects[92]). Therefore, the same spatial distribution of light signals causes in both cases a similar spatial distribution of stimulation of photosensitive cells on the bottom of the eyeball. The reason that a man and a frog perceive something completely different lies in a completely different manner of integration by their brains of signals coming from the retina. Simplifying the matter greatly, human brain has a greater tendency to produce pictures on the basis of the intensity of stimuli received from different regions of the retina. On the other hand, the brain of a frog is prepared to record the derivative of the

[92]Insects receive the range of colours from ultraviolet to orange, while man – from violet to red. They have a complex eye, built of simple eyes (ommatidia), completely dissimilar from our eye. Finally, they see the polarisation of light, imperceptible for us.

intensity of these stimuli in time[93]. In other words, it receives mainly not the intensity of stimulation of some group of photosensitive cells, but rather a **change** of this intensity, that is the fact that a signal was arriving from a given receptor before a while, but it is not arriving now, or inversely. This comes from the fact that the brain does not necessarily see the same what the eye sees. Signals coming from the retina can be stopped or integrated and transformed in various ways. This applies of course to other senses as well. The mechanisms of this processing are so deeply imprinted in the most fundamental categories of our thinking and seeing of the world, that we willingly regard them as the properties of this world itself.

Even in the case of "normal" seeing, that is of seeing an image, the grouping of some complexes of stimuli into "facts" is a derivative of the integration of these stimuli by our mind. We can imagine ourselves creatures whose description of the world does not possess categories corresponding to our concepts of space, time and causal relations. These creatures could be no so well adopted to live in the real world and to explore it. However, we cannot exclude the possibility that they would do equally well, or even better. Their logic would differ from ours, and even "reality" would mean for them something completely different. We are not able to find even a single, most elementary fact of our mind, about which we could claim that it is an objective property of the world. Of course, some convergence, similarity of concepts to "real" facts would have to be imposed by biological evolution as a condition of purely instrumental efficiency of living in the physical reality. But similarity does not mean identity.

Let us return now to the linguistic aspect of our seeing of the world. Normally, the facts of the world as well as linguistic names are assumed to have an independent and fully autonomous existence (which finds its culmination in the philosophy of early Wittgenstein). According to the conception presented here, facts are constituted by our language (as well as by underlying cognitive structures of our brain) from the "chaos", or rather from "nonsense" of sensations that reach us. Acknowledging a "John" to be a fact depends on the manner of integration of these sensations. Facts are products of our mind, and not objectively existing categories of the world. Therefore, the language uses names which have not got well-determined designates. This opinion stands in obvious opposition to the philosophy of the early Wittgenstein, where the only possible

[93]The difference between a frog and a man is not at all as drastic as this simplified description could suggest. A man also has a certain preference for movement perception, especially in dusk and at the margins of the field of vision.

and necessary structure of names designates the set of autonomous facts (existing outside our mind).

The "fact-creating" role of language determines to a large extent the essence of our whole culture, science, religion (the last is perhaps the most obvious case), the sphere of common sense. It shapes the form and content of human thinking. Science differs from philosophy or religion, for instance, by its methodology, which ensures congruence between linguistic (and conceptual) structures and the structures of the world. However, this congruence is far from perfect and this differs the real science from its idealisations. The error of hypostasis consists here in a naive belief in a **real** and "sharp" existence of such objects described by science as the orbital of an electron, gravitational force or biological species, being general concepts. The same also applies to the planet Earth and a given, concrete individual of alga or canary. The qualification "real" in this case does not mean an existence totally outside of our psyche, which I by no means intend to negate, but an existence in the forms, categories or even values produced by the human mind. At the same time, I do not share the original opinion of Immanuel Kant that "things in themselves" are completely non-cognizable. If our categories had nothing in common with the real structures of the world, we would not survive as a biological species, and we would not achieve so spectacular success in natural sciences.

On the other hand, however, the early Wittgenstein and many modern epistemological optimists treat the world as a puzzle. The elements of this puzzle, if fitted appropriately one to another (for which the rules of logic serve), will give a complete representation of the real world on the obverse. If each of these elements were labelled with a corresponding name, than, once the puzzle is finished, we would obtain on its reverse a complete linguistic description of the world. The concept of the "absolute" essence of linguistic names (assuming a designation of autonomously existing facts) leads to such a conclusion. The concept of conceptual network presented here proposes a completely different approach, according to which we see the world as faint shapes behind a hazy screen. The degree of haziness can be diminished by developing the conceptual apparatus and by increasing its congruence with the real world, which is behind the glass. This is realised primarily by the methodology of natural sciences. For the screen to become fully transparent, it would be necessary to built a conceptual apparatus with an infinite degree of complication, which is of course impossible. For the purpose of describing the shapes behind the screen, it could be helpful to divide its surface into small squares of different colours. On the other hand, the same operation performed with a use of small triangles may appear to be much less effective. However, to conclude from this that the world is composed of small squares would be premature, to say the least. I think that the analogy is clear – a description of the world by means of language, so as a description of the picture behind the screen by means of small squares, has, as every approximation, limits of its applicability. Maybe the main purpose of this part of the book is to present these limits.

Another analogy to the projection (representation) of the world by conceptual network can be found in a spider web covering some sculpture (say, Venus de Milo). The denser the web and the more closely it adheres to the sculpture, the more adequate is the "description" of the sculpture in categories of the threads of the spider web and their interconnections with other threads. However, no matter how dense this web, it will never be able to explain the stone the sculpture is built of. Therefore, although we can create still better and better **models** of the world in our conceptual network, attempts at finding its "essence" will be futile for the simple reason that the "substance" of conceptual network is different from the "substance" of the external world[94].

According to philosophers who prefer linguistic thinking, *modo* early Wittgenstein, the known is separated from the not-known by a clear boundary – the known is the already ready part of our knowledge, grasped in linguistic structures and therefore true for ever, while the not-known is the whole rest, waiting to be grasped in this way. On the other hand, according to the conception of conceptual network proposed by me, nothing is fully true or false in our picture of the world. Cognition is a continuous process which consists in refining (or rather: co-refining) of concepts, and in developing distant areas of conceptual network that can present in a new light, or even change completely, these elements of our knowledge which we regarded as unshaken and obvious. Therefore, the process of cognition occurs here gradually and never leads to a completely certain, absolute knowledge. Even logic, which could seem to be an absolute and autonomous gauge of the unity of at least the most basic properties of our mind (language) on one hand, and the objective reality – on the other hand, appears to be nothing more but an evolutionary-grounded mechanism by which our brain treats stimuli from the environment.

[94]This does not imply a return to the classical dychotomy of spirit and matter. Conceptual network is a derivative of the functioning of matter, although additional concepts are necessary to describe it, for example cybernetic ones; after all, the same applies to the phenomenon of life, which is in a sense something completely different than the dead, physical matter.

Let us come back to the evolution of conceptual network during individual development. If a new-born child does not have much of a conceptual network, it is obvious that it does not posses (self)consciousness, which is attributed to an adult man. Therefore, (self)consciousness must appear during individual development. At the same time, its origination is not a single act. It must develop gradually, together with the evolution of psyche and conceptual network. Within the framework of the conception of the conceptual network, this is equivalent to a creation within this network not only of a representation of the external world, but also of a picture of itself. This means, in a sense, that the network enters at a certain meta-level and looks from above, i.e. perceives its own existence. In unconscious psyche, the processes of thinking occur at the level of the conceptual network, and their "field of vision" comprises only the representation of the external world in this network. Together with the appearance of self-consciousness, this field starts to comprise also conceptual network related to the "cognitive apparatus" itself. This is the origin of the differentiation into "I", that is the area of conceptual network which knows both the world (its picture in the conceptual network) and itself, as well as the "world", that is the area of conceptual network which is known, but itself does not constitute a part of the cognitive apparatus. This is the source of the extreme opposition in our psyche of the mind and the external world. They appear to be categorically identical only at the level of the conceptual network, for both are parts of this network (concepts or complexes of concepts, depending on the approach).

The formation of (self)consciousness is strictly connected with quantitative development of the conceptual network, origination of secondary concepts, processes of learning and a mastering of language. (Self)consciousness emerges gradually and one cannot determine the moment of its origin in a non-arbitrary way. It is of course possible to use different tests, like recognition by a child of its own face in a looking-glass. However, such tests are able to probe only a few aspects of (self)consciousness, and the criteria of classification of different reactions as conscious or unconscious are arbitrary and non-verifiable (e.g. recognition of one's own mirror reflection has been found in some birds, that are not suspected of having (self)consciousness).

Above, I described four main aspects of the evolution of conceptual network in the individual development of the man: quantitative growth, origins of surplus, learning as well as production of (self)consciousness. I distinguish similar aspects below, while discussing biological evolution of the conceptual network. This suggests that, in a sense, the ontogenetic development of conceptual network is a repetition of its phylogenetic development, similarly as in the case of the development of morphological properties. However, I do not think that, apart from showing some convergences, this concept could appear to be heuristically fertile (as the theory of recapitulation did not appear to be so), for most of our ancestors had a major part of conceptual network encoded genetically, while a human baby starts the development of this network almost from scratch.

The process of formation of conceptual network described above proceeds in a connotative way. New concepts are incorporated on the basis of the already existing concepts (significative axes). The meaning of concepts corresponding to objects of the world is also determined by their relations to other concepts. The projection (representation) of the world in conceptual network in a denotative way would correspond to direct attribution of concrete concepts to some facts of reality. However, it is not possible to imagine any physical link between a material phenomenon and its representation in consciousness (mind), any mechanism which could establish a univocal relation of representation. For a man does not know (since he cannot) what complexes of signals he will meet during his life.

For the signals from the surroundings to appear in consciousness as senses, a key which translates them is needed. Such a key, at least in man, is not inborn (however, even some inborn mechanisms, like e.g. the spatial and temporal integration of sensory data occurring in man and, to a larger extent, in animals, originated in the process of learning, although it happened during evolutionary, and not individual, development). As I mentioned, this key is formed according to the principle of induction, by multiple perception of similar complexes of stimuli. For a designate of a concept to be "in advance" connected with this concept by a denotative link, concepts for all phenomena which a man could experience or think would have to be encoded in the genotype, and learning would consist simply in the selection of these concepts. This is an obvious absurd. Therefore, contrary to the impression based on the seemingly unequivocal attribution of linguistic names to the facts of the world, the correspondence between the world and its picture in human mind is far form univocal and obvious.

This led, at the moment when man turned towards philosophy, to the creation of what I call an "evolutionary trap of consciousness". Our (self)consciousness (mind) was formed by the processes of evolution in the "objectively" existing world in such a manner (the only possible manner) that it is not able even to prove with absolute certainty the existence of this world. For it is not possible determine in a sensible way the truth value of a "hypothesis" concerning the existence of something like the external reality (this is why idealistic monism has originated in philosophy). Each analysis of the concept "the world" carried out with consequence leads to absurd conclusions, and implies that this concept is empty, devoid of its designate – the "external world" cannot be in any way distinguished from a certain area of our conceptual network. Nevertheless, the "hypothesis of the external world" is very useful in the operational meaning, that is in categorisation and manipulation of the entire spectrum of phenomena which are accessible to us. Therefore, this hypothesis is characterised by practical usefulness. We can talk about absolute existence of the world only in this instrumental sense (anyway, the same concerns absolute, independent existence of our mind). Similarly, the question of the existence or non-existence of general (universal) beings is reduced to various (unconscious) interpretations of the manner of formation and the essence of secondary concepts. The unresolvable (or rather apparent) character of philosophical problems is a direct derivative of our neurophysiology.

I have already mentioned that what constituted the primary "evaluating factor", the "drive of cognitive activity" was the centre of pleasure in our brain (which can be roughly identified with the dopamine-related reward system). As in the case of a baby this statement refers to the simplest, biological significative axes, it rather does not lead to controversies. How does the situation look like in an adult man, in a fully developed conceptual network? Well, in this case any activity and evaluation is still related, directly or indirectly, to the stimulation of the centre of pleasure. It seems to be characteristic for man that this centre can be stimulated not only by stimuli fulfilling biological needs (satisfying hunger or sexual impulse), but also by "stimuli" of cognitive, artistic and, what is more important, ethical nature. Therefore, one of the main traits of humanity at the neurophysiological level is the ability to draw satisfaction from science, music, fine arts, but also from altruism, kind-heartedness and disinterestedness. A translation of ethics into the language of neurophysiology should not shock anybody – after all, it does not cease to be ethics because of this. At the same time, in order to avoid misunderstandings, I would like to point out that the commonly known and strictly localised (spatially and functionally) centre of pleasure (or more generally the reward system) does not necessarily exhaust all functions of what I call the functional (motivational) centre of pleasure[95].

As it results from what was said until now, conceptual network is a derivative of two things. The first of them, inborn, is the groundwork for its development in the form of genetically

[95]I apologize for this reservation, but it is justified by our (or at list my) ignorance.

encoded structures of the brain, possessing determined possibilities and predispositions. The second is a complex of sensations and experiences faced during lifetime. These two conditioning factors are supplemented by chance, able to direct the development of conceptual network in an at least partially unpredictable way. There are no two individuals with identical brains existing in the world, due to both genetic and phenotypic variability. Moreover, life experiences of different people differ drastically. These differences are enhanced by the considerably stochastic nature of neurophysiological processes. Therefore, it seems to be obvious that there are no two identical conceptual networks in the world. There does not exists an absolute identity of senses (meanings), valid for all psyches, or an ideal translatability of senses between particular psyches. Therefore, although it is possible to understand well a given man at the level of linguistic communication, the understanding will never become complete, since the same words and sentences refer to slightly different concepts in different people. Commonness of the language does not, therefore, prove commonness of comprehension. These differences are of course (from the statistical point of view) greater between people speaking different languages, brought up and educated in different cultures.

Therefore, there is nothing like ideal confirmability of concepts which underlie the same linguistic names in different people. This applies to even as "sharp" concepts as the concept of number "two". At the first glance, this statement seems to be rather farfetched. Therefore, I will try to analyse it in more detail.

I will start by reminding that the concept of number "two" was not known to us at the moment of birth, but we came to its gradual understanding and specification during the development of the conceptual network. We learned that it is an abstraction of some repeatable property of objects of the external world (e.g. two stones, two apples), which are not at all sharp concepts. We did not acquire the understanding of the concept of number "2" immediately and in one step during our individual development. In my opinion, a child associating different concepts and names creates in its brain different "hypotheses". Some of them become later "enhanced" by the repetition of facts that fit to them, while others are eliminated as they do not find confirmation in further experience. If we show two apples and two roses to a child and indicate them by saying "two", then it can produce different associations. Either it will link the word "two" with the number of shown objects or, for example, with their red colour. If the name "two" coincides further on with two objects which can have different colours, or with any number of objects having red colour, one of these "hypotheses" will be confirmed, while the other one is abolished. Normally, we consequently present two objects to a child and provide the name "two", so that it comes **gradually** to some, but certainly not "complete", comprehension of this number.

Remembering different situations where the name "two" is pronounced and an (unconscious) attempt at abstracting a common property for these situations is a partially alternative, and partially supplementary possibility in relation to the creation of *ad hoc* hypotheses. As experiences accumulate, the degree of understanding of number two by a child, the level of abstractness of this concept grows. At some stage, a child can for example associate this concept with two similar objects in space, but not with two events following one another in time. One can imagine a culture which would use a different system of names and concepts for counting phenomena in time and in space (this undoubtedly would not remain without an influence on the comprehension of the world by this culture). At a slightly different stage, a child would apply the concept of number "two" to very similar objects, for example two cats, but not two different objects (if the child is shown a hen and a horse in a picture, it will not perceive them as **two** animals)[96]. As we see, the specification, understanding, "abstractness" of the concept of number "two" is gradual.

The same concept for a mathematician and for a child differs only quantitatively. For a child, this concept is evidently not sharp. If it were perfectly sharp in an adult man, this would mean a necessity of reaching a qualitative jump – a sudden and complete understanding of number two at a certain moment of individual life. However, it is not possible to find out such a moment; on the contrary, we can observe many small steps in a progressing process of "understanding" (this does not contradict the fact that there exist moments of revelation when comprehension and specification of a given concept advances very much – an occurrence of gradual, evolutionary changes does not mean that they must proceed all the time in at the same pace). Even the apparently so sharp concept of number "two" can be further specified and sharpened – mathematics provides the adequate evidence to support this claim. Number "two" can be further defined in various ways, for example as a set of sets possessing the property of "duality" (containing two elements), as was proposed by Frege, or as an abstraction of a double use of a function in the Church's "lambda calculus". Each determination of number "two" attributes

[96]I do not know if particular examples I quote here actually take place in reality. They only serve me to illustrate some general principle.

specified semantic connotations to it which are absent in other formulations. A "sharp", ideally specified concept of number "two" is only a delusion founded by our mind, having nothing in common with facts.

Taking matters strictly, a concept of a number cannot claim absolute sharpness, because numbers have in fact nothing to "count" in the real world, since singular beings do not exist in any ideal manner. Macroscopic objects do not exist, because, when we take them apart atom by atom, we are not able to state when our object ceases to be itself and becomes something completely different. Just as in the paradox of a bald man, it is not possible to determine how many hairs are to be plucked from the head of a man with a bushy shag to make him bald. In the case of two colourful spots, one red and one of a colour that differs infinitely little from red, we also cannot answer the simple question without restoring to an arbitrary decision: "how many red spots we deal with?": one, two, or maybe one with a fraction? On the other hand, microscopic objects (elementary particles) do not exist in an absolute way due to the uncertainty (indeterminacy) of quantum mechanics. A wavefunction which determines the probability of detecting an electron at a given point of space makes this particle "blurred", in a sense. There is a certain probability (although it is small) that our electron will be found not where the probability is greatest, but in a neighbouring galaxy. Additionally, due to quantum mechanics, all electrons are mutually equivalent and exchangeable, and behave as various copies, "manifestations" of one electron. And where is our absolutely sharp concept of a number when we are to answer the seemingly so prosaic question: "how many electrons are there in this room"?

The problem of the "sharpness" of concepts can be quite well illustrated by the analogy with a knife. Can a knife be absolutely, **infinitely** sharp? Of course – not, because finally, going down towards still smaller and smaller scales of magnitude, we meet the granularity of atoms (great majority of them being atoms of iron) as well as the quantum indeterminacy of their electron cover. This granularity cannot be "skipped" in any way – the edge of a knife cannot be sharper than a single row of atoms (in practice, we have not even approached this ideal). A knife may **seem** to us to be ideally sharp in practice, for our purposes, which does not mean, however, that it is absolutely sharp. When sufficiently magnified (which is possible by means of the scanning microscope) the blade of an unused razor resembles a landscape composed of undulating hills.

A similar situation takes place in the case of concepts. They cannot be "sharper" than it is allowed by the granularity (resolution) of conceptual network or, if somebody prefers, the neural network underlying it. In practice, mathematical (and logical) concepts are sufficiently sharp for many purposes. However, the fact, that there are still disputes in mathematics as to how strict (how formalised) should be a mathematical proof, testifies that these concepts still lack ideal sharpness. A good example can be found in the proof (in its first version) of the famous Fermat's theorem, quoted by A.J. Wiles, or the proof of the Kepler's hypothesis by W.-Y. Hsiang. In both cases, a part of mathematicians regarded the presented proof as sufficient, whereas other mathematicians saw gaps in it.

Further arguments supporting the opinion that concepts lack absolute sharpness are provided by attempts at modelling the phenomenon of learning by means of artificial neural networks. Simulations of the process of learning, analogous to the one occurring in a child, were performed on electronic systems which imitated (to some limited extent) the structure of our neural network. Learning could focus, for instance, on recognition of letters or pronunciation of English words. Also in this case, the "education" advanced **gradually**. In the course of learning (consisting in enhancing or weakening different connections in response to a correct or false identification of a letter), the per cent of instances of correct recognition of a given object increased. No clear point existed, however, before which the system did not recognise an object at all, while afterwards it always answered correctly – the transition was **continuous**. Similarly, a child comes gradually to the understanding of pictures and concepts, including the concept of number "two". Of course, anything like "absolute" understanding of a number is not possible – in the ultimate instance, a number should also be regarded as a manner of ordering the world by our mind, and not as an objective property of the world.

A significant regression in the degree and abstractness of the comprehension of numbers can be the result of damaging the brain in an adult man. This can happen in the case of an injury of the centre responsible for spatial orientation, situated in the temple lobe (Gerstmann centre). Most mental functions remain unchanged in a man with such a defect. He is still a normally intelligent individual. However, he is not able to perform even the simplest operations concerning numbers, e.g. adding up in his mind. Nevertheless, he is able to realise elementary calculations by means of material objects, for example fingers or apples. Therefore, it cannot be said that such a man does not understand at all the concept of number, but only that this concept is related in him to the "visual" properties of the surrounding space (let us remember that the centre of spatial orientation has been damaged). This constitutes a further premise supporting the thesis that our concepts are not as abstract as we would be inclined to wish them to be, and that their roots are in sensory stimulations. The localisation of the function of spatial orientation (the sense of location) and the understanding of numbers in the same centre confirms this conclusion.

What has happened to the "abstractive" part of the concept of number in the described case? The Gerstmann's centre is a fragment of the brain cortex, and therefore of the part of the brain responsible for association of stimuli, memory and thinking processes. As I proposed earlier, these processes lead to the origination of secondary concepts, that enrich the "sensory" substance of primary concepts with the mechanisms of integration and association of different impulses in the brain (I should add here – better late than never – that the difference between primary and secondary concepts is liquid and arbitrary). Therefore it is probable that the majority of secondary concepts acquired during individual development and concerning numbers and spatial orientation was destroyed in the described case as a result of the damage of the neural network underlying them. There only remained primary concepts, localised "deeper" in the structure of the brain (or maybe less centralised, "blurred" over a large area of the brain cortex). In a normal man, they undergo processing and interpretation by secondary concepts. However, in the last instance, our "abstract" understanding of number "two" is still dependent on and inseparably connected with the ability of counting on one's fingers!

Each concept means by connotation, and therefore, in fact, by reference to all other concepts. Because no two identical conceptual networks exist, there do not exist two **perfectly** identical meanings of the concept "number two" either, although the differences between these meanings can be so small that they are not recorded by our mind. This does not mean, of course, that there does not exist (whatever this could mean) a property of the external world which corresponds to our concept of number "two", but only that the "reflections" of this property in different minds are not ideally the same. If different persons are shown a photograph of one hill or of two distinct hills, then they will of course provide convergent answers for the question: "how many hills are there in the photograph?". However, if the photograph presents one hill with two peaks, separated with a small pass, then the provided answers will differ. Of course, one could argue that different persons differently comprehend the concept "hill", and not the concept "number two". However, as the latter concept originated (like all concepts) by induction, then it cannot be "pure", sharp and devoid of empirical "contaminations". Actually, the concept of number two applied to two discrete objects, for example two atoms, **is not exactly the same concept** as the concept of number two applied to continuous objects (e.g. hills), although the proneness of the human mind, and particularly of science, to shred the world into discrete categories with fixed labels of linguistic names thoroughly obmen of letters these differences.

Two different concepts of number two in different conceptual networks resemble two superimposed identical regular octagons, whose centres overlap, but one is turned in relation to the other in such a way that its angles are situated between the angles of the other one. Therefore, the area common to both octagons is much greater than the areas that belong only to one of them, but not to the other. If we treat the octagons as a metaphor of conceptual network "meshes", the common area represents the zone of identical understanding of a given concept. This explains why the understanding of number two is, in a huge number of cases, convergent for two different people. The greater the number of angles in such superimposed regular polygons, the smaller areas belong to only one of them (this area is greatest for triangles). If we took two polygons with a very great number of angles, we would treat them as two ideally overlapping circles, due to limited resolution of our eye.

Similarly, the "resolution" of our "introspection" can be too small to discern the difference between concepts of number two in different conceptual networks (a circle can stand in this analogy for a "real" property of the world, which underlies the concept of number two, while a polygon circumscribed around the circle can correspond to an approximate representation of this property in the conceptual network). This does not mean, however, that such differences do not exist.

EVOLUTION OF CONCEPTUAL NETWORK IN BIOLOGICAL EVOLUTION

Although we know what biological evolution consists in, it is frequently for us too abstract a process to fully understand its consequences. We know, for example, that life originated on the Earth some 3.8 billion years ago and its development during this period led from the simplest forms to the emergence of the species of rational beings – man. However, mind and (self)consciousness had also to originate for rational beings to come into being. Moreover, it had to do so from something that had not been consciousness yet. As biological evolution is a continuous process, germs of thinking had also to be formed gradually. In order to illustrate the fact of biological evolution and its consequences, I will use the metaphor of a Gallery of Ancestors. This is a long, straight corridor, where, in heavy oak frames, portraits of our ancestors in the masculine line (for instance) hang on the wall. At the beginning of the corridor, one sees himself, then his/her father, grandfather, great grandfather and so on. We walk along the corridor, looking at the portraits. Soon after our great grandfather, there appear some men with swords and in armour, then – in russet overcoats with spears, and then – dressed in animal furs and armoured with clubs, finally – completely naked creatures with stony pestles in a hand, and afterwards – without even these tools. After a few hours of our walk, our forefathers in the Gallery of Ancestors start slowly to be more and more hairy, they begin to stoop, foreheads and chins move back, their arms become longer and longer, while legs – shorter and more bent in knees. The size of their cerebral skull diminishes. At the same time, our subsequent ancestors more and more frequently pose to the portrait while hanging from a branch and not, as previously, standing on the ground. After some time, we see on a tree a creature of the size of a cat, with a long tail and prehensile limbs. Then we catch ourselves – maybe for the first time – reflecting that it was not so long ago that some elegantly dressed older man with a splendid moustache and a gold watch on a chain looked on us from a portrait. His father resembled him very much, his grandfather resembled his father, a mister with a club did not differ in fact by anything (apart from the clothes) from the mister with a spear, and so on. So where did this strange creature come from?

But it did not appear all of a sudden. Its immediate descendants resembled it very much. Our grandfather and this creature in front of our eyes differ drastically, but at the same time, a father was always similar to a son in our Gallery of Ancestors! Amazed with this contradictory and simultaneously evident character of our insight, we move on. As we have a long way to go, we mount on a bicycle. We move by portraits of small insectivorous mammals with a long snout. Then, we see mammal-like reptiles, furless creatures that lay eggs, labyrinthodonts resembling large salamanders, finally crossopterygians. Here we may experience another shock. It is difficult to believe that this fish with goggle eyes and a dull expression of the snout is our ancestor in a straight line. Tired and dejected, we get into a car. We pass by a long sequence of worms digging in mud, sedentary creatures filtering water – whose mouth gets finally overgrown and its function is taken over by the anus – till we reach still simpler and simpler unicellular organisms painted by means of the microscope. After a very long and tiring travel, we finally reach the end of the corridor. In the last portrait, we see an oil painting which presents some turbid mixture of proteins and nucleic acids. We read on the verdigris-covered plate below: Sir John Hypercycle. We would try in vain to find any traits of similarity with our great grandfather.

The wanderings along the Gallery of Ancestors was meant to visualise the fact that, although we differ from our forefathers very drastically, the process of emergence of our species in the course of biological evolution was continuous, and no sharp barrier separates the human from the non-human. This implies that the situation was similar in the case of the origination of (self)consciousness – one of the most important, if not the most important attribute of humanity.

Let us now turn back in our Gallery of Ancestors in order to pass it once again, in the opposite direction, this time to look not at the external appearance of our forefathers, but at their originating and slowly developing conceptual network, which finally becomes the background of what we call (self)consciousness. For the development of our spirit was a **continuous** process, as the evolution of the body was.

Conceptual network is a "result" of the activity of the neural system (network of neural cells). Therefore, we can talk about germs of conceptual network only in organisms that possess at least the simplest neural network. Such a network appears in coelencerates (e.g. hydra) in the form of a loose complex of primitive neural cells with "equal rights" (or, if somebody prefers, cells fulfilling their role), and in worms, where we find concentration of neurones in stems and ganglions, e.g. in the "head" region. The "conceptual networks" of these organisms are extremely simple. Their organisation is mostly inborn, and their exclusive function is to "re-code" (translate) stimuli coming from the environment into the behaviour of the animal. Therefore, they have a purely biological "meaning". For the discussed above fictitious organism with only one neural cell, this meaning is a mechanical ("reflectionless") realisation of the logical function: "if the organic substances are lacking in the water and you see light, go in its direction". Anyway, it would be a simplification to think that "meanings" of this kind are characteristic only for the neural network. After all, a bacterium that produces enzymes decomposing lactose in response to the appearance of this sugar in an environment, by activating (more strictly: removing the inhibition of) the genes responsible for the production of appropriate enzymatic proteins (the theory of operon), also responds adequately to the appearance of nutritious substances in its surroundings. Its reaction is a part of a "network" of biochemical and genetic mechanisms. Therefore, biological "meaning" can be identified with a generally purposeful behaviour of living organisms (at different levels: biochemical, physiological, neural, and so on). However, only the network of neural cells developed into a system complex and "excessive" enough to transform biological meaning, slowly and unnoticeably, into psychological meaning – the basis of our (self)consciousness. Therefore, one should look for the beginnings of conceptual network in primitive neural systems.

As I mentioned, the "significative axes" (e.g. satiety – hunger, satisfaction of the impulse to transfer one's own genes – lack of this satisfaction) are inborn in most of our ancestors, and their meaning – exclusively biological. Anyway, both the set of stimuli to which a given organism reacts, and the behavioural repertoire of responses are frequently extremely poor. Jacob von Uexkül used a tick as an example. This is not a representative example in the context discussed here, because the tick does not belong to our Gallery of Ancestors (it is a representative of a quite different branch of the evolutionary tree), and – as a parasite – its behaviour is very simplified. Nevertheless, one can suppose that – at some stage of evolution – our ancestors had at their disposal a behavioural repertoire of a comparable degree of complication.

The whole life of such a tick consists of climbing on a tree or bush to wait for a passing-by animal (frequently for several years in complete immobility) and fall down on it and find a proper place to attach itself and consume blood (and, of course, of meeting an individual of the opposite sex and transferring its genes to progeny). A tick needs a very limited set of stimuli to realise this task (mainly aromatic stimuli – a scent of butyric acid secreted by sweat glands of mammals, thermal stimuli and an elementary sensitivity to light). Its "world" – its conceptual network – is therefore very poor. For a tick, the concept "roe deer" is a combination of a characteristic scent, warmth and satisfaction of hunger. For us, "the same" concept is defined by millions of other concepts, concerning the appearance and behaviour of a roe deer, its taxonomic position, anatomical structure, physiology, biochemistry, cultural aspects (e.g. hunting) and so on. We know that a roe-deer is built of atoms, these in turn – of a nucleus and electrons, which are ruled by the laws of quantum mechanics… Therefore, in man, the semantic context for the concept "roe-deer", i.e. something that "defines" this concept consists of his entire conceptual network composed of a great number of concepts – nodes. In a tick, this network has only a couple of nodes. No wander therefore, that the "psyche" of a tick is completely incomparable to the human psyche.

A tick has to possess a certain minimal behavioural repertoire to find its host. Such problems do not concern internal parasites, such as a tapeworm, the eggs of which are unconsciously swallowed by a host. Their life activity, therefore, is reduced to anchoring in the alimentary canal of a host and production of eggs. It seems that speaking about conceptual network of a tapeworm borders on absurd. This animal represents the level at which our free-living (non-parasitic) ancestors were at the moment of the origination of their neural system. Since then, one can speak about the development of the conceptual network.

The co-evolution of the neural system and conceptual network proceeded on many planes (it had many aspects). First of all, it was a quantitative development which consisted in increasing the number of elements of both systems: neurones, receptors and effectors, on one hand; and concepts, on the other hand. Origination of new senses and development of the already existing ones favoured growth of the quantity and variability of data about the surroundings that were reaching an organism. Therefore, apart from the primeval chemical sense, there appeared other senses: sensibility to light, sound, mechanical stimuli. At the same time, the already existing receptors underwent gradual differentiation and improvement that resulted in the formation of so sensitive and efficient sensory organs as the eye of vertebrates, the echolocation system in bats, the smell receptors of some moths detecting single molecules of sex pheromones or the "heat eye" of a rattlesnake. A more developed neural system was needed to transform sensations coming from these receptors. This resulted in the creation of projection of the surroundings in the neural network that was both richer and more adequate to reality. Concepts corresponding to particular "objects" or aspects of the real world were becoming more precise, because a broader spectrum of sensory impressions (both with respect to their diversity and resolution) served to "define" them, and because these sensations were integrated by the neural system with greater efficiency. What we see as a fish with all the details of its appearance and behaviour, for some primitive water organism can just mean "a movement in the water" which is a signal of danger. An improved picture of the world was useful for organisms only to the extent it increased their chances for survival and production of progeny. The development of receptors and the neural system had therefore to proceed parallel to the evolution of effectors (increase of their number, diversity and efficiency of action). This led to the formation of still better developed locomotion which could secondarily fulfil also certain manipulative functions (e.g. prehensile upper limbs in monkeys, apes and man). This "operational" efficiency in registration of and adequate reactions to complex signals from the external world required organisms to create in their neural network a system – network of interrelations (dependencies) between its different elements – which corresponded in some way to the structure of the world. This was exactly the origin and the beginning of development of the conceptual network.

What was said above can be summed up in the following thesis: association of stimuli from an environment with reactions of an organism to these stimuli – which is initially efficient only instrumentally (i.e. related to the survival of the organism and not its "purely cognitive" abilities) – leads to a more or less adequate projection (reflection) within the neural network of the structure of the world. Representation of different aspects of the world – being the result of this projection – is always imperfect, and its form is to a large extent dependent on the physiology of sensory organs as well as on the neural network. The essence, "substance" or content of concepts comes as much from the structure of the world as from the structure of receptors and neural connections – from their manner of ordering of sensations. Man, acquiring knowledge of the world, frequently forgets that his thinking, conceptual network, and therefore categories in which he considers reality, reflect not only the "necessary" properties of the world, but also – to a large extent – accidental (evolution-shaped) mechanisms of the operation of our receptors and of processing the receptor-provided data by the neural system.

The increase of the "excessive" character of both systems was the second aspect of the parallel evolution of the neural system and the conceptual network. It was related to the fact that some their elements ceased to be responsible for direct "translation" of environmental stimuli into reactions of an organism. Instead, they took over the function of processing and integration of sensory impressions, locomotive co-ordination, accumulation of memory, and finally – association and co-ordination of various functions of the neural system. This resulted of course in the development of the conceptual network, which meant increase of the number of received stimuli and manifested reactions, while their integration supplied new dimensions in the semantic space, new significative axes differentiating (defining) concepts. The above-mentioned "surplus" of the neural-conceptual system created the possibility an "internal dynamics" of the neural system (its "autonomous activity") to originate and develop, which corresponded in the psychical sphere to germs of the processes of thinking. At the same time, the “sensory” givenness of the sense of many concepts decreased. In the organisms where conceptual network realises the function of simple translation of environmental stimuli into the behaviour of an organism, all concepts are immediate derivatives of the structure of signals transmitted from sensory receptors. They correspond therefore to particular facts of the external world, and the complication of their structure reflects the "resolution" of receptors and the degree of integration of data they transmit. I call such concepts "primary concepts". Primary concepts are therefore various "simple" objects and processes of the external world, seen through the prism of the senses of a given organism. Therefore, one can qualify them as accessible through direct “sensory” experience. However, the appearance of the above-mentioned areas of excess in the neural system (and conceptual network) made a part of concepts lose their accessibility through direct “sensory” experience. It means that the meaning, the form of these concepts is no longer a simple projection of facts of the world, co-shaped by the physiology of senses. The "structure" of these concepts – which I call "secondary concepts" – is formed on the basis of primary concepts, but it is also shaped by mechanisms of memory operation, integration and association of neurophysiological processes, and co-ordination of different parts of the neural system. Therefore, secondary concepts comprise all kinds of general and abstract concepts, not corresponding directly to simple facts of the real world. This does not mean that their structure has nothing to do with this world. Quite the contrary. After all, they are formed in the process of integration of primary concepts, and the manner of this integration is not completely accidental either, for it had to go through the sieve of natural selection. This selection spared only such types of integration which somehow reflected the structure of the world, allowing an organism to survive as well as acquire an effectiveness in a purely instrumental sense.

Therefore, secondary concepts are not, as they cannot be, completely separated from the world. They differ from primary concepts for they capture reality not only through the prism of the physiology of senses (and simple neural networks related to them), but also through the mechanisms of integration and association of stimuli, that is through the processes of thinking. Capturing the world in such categories as three-dimensional space, time or causal relations results therefore from the manner of integration of sensory data by our brain. Predispositions to capturing the world in this way (that is to such processing of sensory stimuli) are presumably inborn (although, of course, the concepts of space, time and causal relations are not inborn). The fact that some organisms are supposed to perceive the world in two dimensions testifies that it can be done differently. However, this does not mean that the manner of sensory data integration is completely arbitrary. The external world must have a structure which allows to the brain to order received stimuli within categories of space, time and causal relations. Therefore, "something" corresponds to these concepts in reality.

The above conclusions are to some extent analogous to the opinions of Immanuel Kant. This philosopher maintained that stimuli (sensations) reaching us from "things in themselves" (non-cognizable in principle) are captured by our mind in some a priori categories, like time, space or causal relations. However, these categories were of absolute character in the philosophy of Kant, while the structures of our thinking shaped by evolution are, of course, not absolute. Apart from this, Kant claims that "things in themselves" are completely non-cognizable. In the evolutionary approach, we are forced to recognise gradable character of the level of the cognition of the real world. This level is very low in organisms with a weakly developed and differentiated neural system. It increases together with the development of the neural network, receptors and effectors, reaching the maximum development in man. It means, the top in relation to the hitherto accomplished achievements of evolution. It is undoubtedly possible to go further. It is also possible to choose a slightly different way, developing different receptors (sensitive, for example, to infrared radiation or ultrasounds) and different mechanisms of integration of signals they transmit. This would lead to formation of a slightly different picture (representation) of the world in the conceptual network, yet the picture would correspond – better or worse – with this world. Man is able to compensate indirectly for the lack of some senses and inborn manners of seeing the world by creating appropriate "prostheses" through the development of science and appropriate devices. Due to this, he can further improve his manners of cognition of reality by gaining knowledge about it in ways that partially eliminate the necessity of direct sensory insight[97]. In this formulation, cognition of "things in themselves", although impossible to an absolute degree, can be enhanced relatively by developing conceptual network and continuous verification of the

[97]Since the primary "substrate" of all concepts are sensations coming from receptors, it is of course impossible to do wihtout sensory ocularness completely.

"correspondence" between its structure and the structure of the external world. Natural sciences are particularly useful for this purpose, because of their methodology.

The picture (representation) of the world created in conceptual network is, therefore, neither a perfect projection of this world, nor an accidental structure completely separated from the world. The development of conceptual network proceeding along parallel route to the evolution of the neural system led to an increase in the amount of concepts and to their better specification, and therefore, to still more and more faithful projection of still greater and greater number of aspects of reality. All concepts come in some way from sensations received from the external world. The connection of secondary concepts with the world, however, is much more indirect than the connection of primary concepts. What in the structure of conceptual network is not a projection of aspects of reality arises from the physiology of senses and from the mechanisms of the neural system functioning.

The status of mathematical concepts – concerning such "objects" as number, point, straight line as well as relations between them – still remains to be clarified in this formulation. One of approaches, referring to the Platonic tradition, holds an idealistic assumption of an existence of these concepts (objects, relations) beyond both the human mind and the material world. The role of a mathematician would be to discover the world-and-man-independent structure of mathematics. However, such a hypothesis seems to me to be excessive and unnecessary. A competitive hypothesis can be formulated within the conception of the conceptual network. Mathematical objects and relations between them would refer in this formulation to concepts within our mind representing some aspects of the real world, when they are abstracted from accidental properties. These concepts constitute building blocks which can be arranged in different combinations, which allows one to create different mathematical systems. Such systems can be created because the building blocks fit together in some combinations, while in other combinations they do not, yielding either an internal coherence of the system or a contradiction. Only few of these systems are used by us for describing the physical world, while other systems are "separated" from this world. But this separation is only apparent. It constitutes an introspective impression of our mind, because concepts of mathematics – its building blocks (both mathematical objects and relations between them) – come from the world. In the same way as we create "unreal" mathematics, we could – possessing appropriate knowledge – create "unreal" biologies by projecting non-existing living organisms built of already existing building blocks:

chemical elements and compounds, or even cells, tissues and organs. A different attribution of aminoacids to nucleotide triplets (codons) in the genetic code could serve as a simple example. Again, not all combinations of building blocks would lead to formation of a "operative" system, which is due to internal properties of the building blocks. Just as systems within mathematics can be verified by the demonstrating their correctness or internal contradiction, so the criterion for "evaluating" constructed organisms would be found in their functionality or disfunctionality. Therefore, there is no need to assume the existence of mathematics outside the structure of the world and our minds. By the way, the exactness, power and elegance of mathematics testifies that the contribution of objective properties of the world to the formation of mathematical concepts in our mind is huge, incomparable with any other field of human knowledge, while subjective "contamination" of these concepts by the structure of our senses and neural connections is relatively small (however, the share of this contamination is certainly different in different branches of mathematics). This contamination certainly cannot be completely eliminated (which is shown by Gödel's proof or the turtle and Achilles paradox). Fig. 3.3 compares the "Platonic" concept of mathematics and the point of view represented in this book.

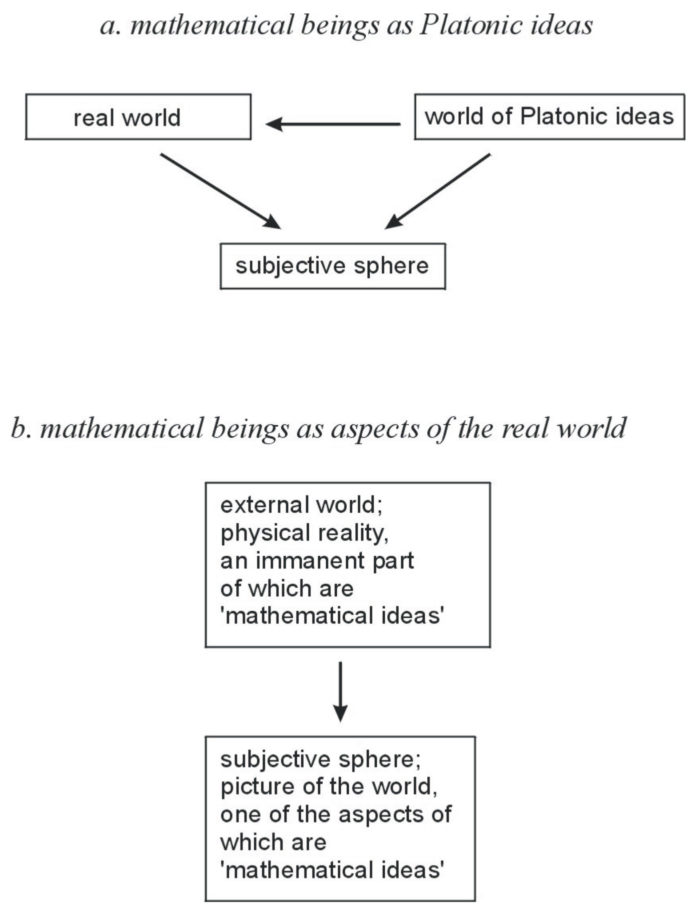

Fig. 3.3. Comparison of Platonian opinion on the reciprocal relation between the world of mathematical ideas, real world and the subjective sphere, with the respective opinions represented in the present work.

The excessive character of the nervous system with respect to the function of direct translation of the set of received stimuli into the behaviour of an animal cannot be separated from the third aspect of the common development of the nervous system and conceptual network, namely a growing part of behavioural repertoire as well as of integration and interpretation of sensory impressions was no longer inborn, but acquired in an individual experience. This was related to the development of the memory recording processes. These changes enhanced significantly the elasticity and dynamics of the conceptual network; its scheme was determined genetically in an increasingly smaller degree, while the process of learning acquired growing impact on their shaping. This was favoured by the care for progeny, when parents constituted a kind of a buffer, behavioural hood for a young individual, which compensated for any threats from an environment and supplied food as well as facilitated acquisition of experience until a young individual reached maturity. This allowed a young individual to be born with an "inborn" conceptual network composed of a simple system of "biological" significative axes, with a degree of complexity comparable to that of conceptual network of a tick or tapeworm, for it had time to develop its complexity.

Conceptual network formed in this way was completely different from the inborn, genetically-determined conceptual network. First of all, the former could be ampler and better differentiated than the latter. Firstly, only a scheme of a simple neural network can be recorded genetically due to connotative meaning of particular neurones in this network. Therefore, greater conceptual networks can develop only by way of learning. This proceeds by imitation of parents (this is another role fulfilled by them, besides creating a behavioural buffer between a young individual and the surroundings) or other adult individuals, or through the accumulation of experiences gained by the method of trial and error. Advantages of such learning are the following: quickness (the whole process is contained in an individual life), elasticity and excessiveness, allowing to react adequately to unpredictable situations that have never been faced yet. The evolution of inborn schemes of neural (and conceptual) networks is also the method of learning by trial and error, for, as a result of mutations, there originate various variants of neural (conceptual) networks, that "perceive" and integrate stimuli from the external world in a slightly different manner (creating slightly different pictures of this world), and thus, respond to them differently. Only the best among them, the most adequately adapted to reality are allowed to survive in the natural selection process. However, this way of learning is much slower (the acquiring of evolutionary memory is measured in geological time), less effective, and conceptual networks formed in the process are much less flexible and completely proof to correction during individual development. However, individuals of the same species can after all live in slightly different environments and a behavioural repertoire that is adequate in some conditions can be quite inadequate in other conditions. Therefore, inborn conceptual networks are considerably less well adapted to differentiated and quickly changing environments than networks shaped in an individual development. The absence of necessity to possess the structure of the whole neural network recorded genetically creates the possibility of much faster evolution of the nervous system, particularly the brain. Learning, as acquiring of experiences during individual development, also favours formation of thought processes as autonomous activities of the neural network. In an inborn, strictly determined neural network, there is no room for elastic association of various processes and decision-making, inseparably connected with the processes of thinking. A genetically programmed conceptual network is therefore poorer and stiffer, evolves slowly, is devoid of autonomous activity and reflects the real world less adequately than a conceptual network originated in the process of learning.

Finally, the fact that in a certain moment the cognitive apparatus ("the cognitive centre") in the brain became directed on itself led (in my opinion) to the origination of (self)consciousness, the emergence in conceptual network of the subject, able to follow (investigate) the processes occurring in this network. This should be, in my opinion, understood in the following way. A certain "cognitive centre" (this is a metaphor, meaning rather a functional, but not necessarily an anatomical centre) originated in the brain in the course of evolution. It associates integrated data incoming from receptors with records of memory, co-ordinates various functions of the central nervous system and makes decisions concerning stimulation of effectors. This centre, therefore, confronts signals from an environment with the acquired memory, knowledge and the picture (representation) of the world formed in the conceptual network. An important role is played in its operation by the autonomous activity of the neural network. Therefore, the processes comprising what we call psyche occur in this centre. If we separate it arbitrarily as a relatively well isolated system, then, the inputs of this system will be constituted by the receptors and centres processing signals coming from receptors. I think that the origination of (self)consciousness consisted in directing a part of the inputs of this system onto the system itself, that is on the self-recognition of the processes occurring in the mentioned centre. These processes started to process themselves as well, in the same way they formerly treated the data coming from receptors (or from memory records, as in the state of sleep). **Apart from a picture of the world, the above-mentioned "cognitive centre" created also within itself a picture (representation) of itself**[98]. This would therefore be a relation of self-applicability, analogous to some degree to a similar relation found in liar's paradox, Russell's antinomy of classes and Gödel's proof, or to the concept of concept

[98]This could enable the origination of language, in which some mental objects, namely linguistic names, are representations of other mental objects – concepts (which, in turn, are representations of the "real" objects of the world). I propose that creatures devoid of selfconsciousness can possess only concepts which represent real objects, while the beings conscious of themselves create (this is what selfconsciousness consists in) representations of representations of representations ...
and so on, *ad infinitum*.

presented by me, where "a concept" is also a concept. Taking into account the fact that the neural network is the background of the conceptual network, this last analogy does not seem to be accidental.

The neurophysiological process proposed above would therefore lead to the origination of the concept of the "*ego*" in the conceptual network. As every process of this kind, it was not (or at least it did not have to be) a discrete phenomenon (occurring in one step), since it is highly probable that directing of the "cognitive centre" input onto itself occurs gradually. For no single mutation could lead at once to the origination of a complex nervous structure underlying our self-consciousness. Some small non-continuous, "qualitative" change could have been necessary, but further processes had to occur by way of accumulating relatively slight mutations. Anyway, only some predisposition to produce self-consciousness may be encoded genetically in our brain. Self-consciousness itself is not encoded in this way. I think that if a man were isolated from the inflow of **any** afferent stimuli since birth, then, not only any self-consciousness, but even a psyche would not form in him (which follows from what was said in the previous chapter). For he would not be able to produce a more complex conceptual network, an elementary representation of the world, and therefore he could not be called a man in any psychological sense. Therefore, being born a man in the biological (carnal) respect does not involve being a man in the spiritual respect.

What was said above leads to the conclusion that (self)consciousness was born gradually in the process of evolution. Therefore, it is not possible to fix the moment of its origin – in a non-arbitrary manner – that is to point out our forefather in the Gallery of Ancestors who already possessed (self)consciousness, while his parents did not. If the formative process of (self)consciousness was continuous, one is tempted to ask the question if some of presently living animals (for example chimpanzees or dolphins) possess germs of (self)consciousness. Personally, I think that such germs are present in them (chimpanzees are able to learn very simple symbolic languages), and that the development of (self)consciousness was stopped at a low level in these animals due to the absence of a more complex system of communication (this relation is bi-directional, however; language stimulates the development of (self)consciousness, while (self)consciousness stimulates the development of language) and of such a perfect manipulatory tool as the human hand, enabling the realisation of more complex orders of the brain. Language – being itself a part of conceptual network – influences the possibility of developing this network in a fundamental way, for it facilitates significantly the process of operating with concepts within the conceptual network. Production of an ethnic language within conceptual network by the ancestors of man acted as a positive feedback, resulting in a chain-reaction development of this network, including the elements directed on its self-recognition. Since language is a part of culture, broadly understood, I will deal with this problem in the next chapter.

The origination of self-consciousness has, in my opinion, a certain common trait with the origination of life: both processes consisted in the appearance of the relation of self-application (directing a process onto itself). In the first self-recreating living system, the hypercycle, some protein(s) catalysed copying (replication) of RNA that coded this protein and was the matrix for its synthesis. Therefore, such a protein catalysed a reaction of self-recreation. (Self)consciousness, similarly, originated by self-reception of the "cognitive centre" in the brain, by directing of its inputs onto itself. In this sense, (self)consciousness is in relation to life what life is in relation to inanimate matter.

CULTURE AND EVOLUTION OF CONCEPTUAL NETWORK

In previous chapters I described the evolution of conceptual network in the ontogenetic and phylogenetic development of man. One more aspect of this evolution remains to be discussed. It is the cultural aspect. Culture accompanied the last stage of biological evolution and is present practically in the entire period of individual development of man. Therefore, it plays an enormous role as a factor co-shaping conceptual networks of society members, while it is, in turn, formed by these networks. Culture is a certain form of collective communication of people, both in space and in time, a binder which turns sets of individuals into complicated "organisms" of societies with a complex, hierarchical structure and multiple internal relations. Due to its distinct position in the formation of the conceptual network, culture requires to be treated separately.

One should start by determining what is culture in the context discussed here, and therefore, how it is related to the conceptual networks of people living in a given culture and creating it. Therefore, we are interested in the psychological aspect of culture (and not, for example, in its material aspects, i.e. the so-called products of culture). In this formulation, culture constitutes something like Popper's third world – a complex of scientific knowledge, ethical and aesthetic values, arbitrary conventions *etc.*, situated beyond the material world (the first world) and our psyche (the second world). Culture, therefore, would be attributed with autonomous existence, not reducible to psyches of individuals participating in it.

This concept is, unfortunately, an inevitable hyposthasis, since it represents a holistic, idealistic approach. It is possible to accept the opposite, reductionistic paradigm, announcing that culture does not exists outside human minds, and that it is a sum, average or resultant of the mental elements, of which culture is composed. Such an approach, however, by resigning from beings with universal status, hinders any communication concerning the complex of phenomena that surround us. It seems therefore that it is most reasonable to choose the standpoint called conventionalism, according to which culture is a conventional category, commonly understood concept accepted because of its usefulness in ordering, describing and interpreting the world of phenomena accessible for us.

It is in this sense, that I would like to propose a model of culture in its psychological aspect in the form of a conceptual network existing outside psyches of particular participants of the cultural sphere. Such a network would contain first of all an ethnic language used by a given culture, accompanied by its knowledge and scientific theories, masterpieces of literature, fine arts and music (existing in the form a complex of their perceptions realised by the members of the given culture), conventional symbols, system of values, and, finally, various religious beliefs. Of course the amount of information contained in conceptual network of culture exceeds significantly the capacity of human brain. This is the first reason why this network can exist only outside any consciousness, in the sphere of abstract beings. The second reason is the already-mentioned fact that there are no two persons with identical systems of meanings. Conceptual network of culture must therefore be composed of concepts that are more blurred than the concepts of particular individuals, who precise these concepts for their own use. Or rather, conceptual network of culture has to take into account the whole range of semantic systems of the members of this culture. On the other hand, these semantic systems do adhere to each other in some way and allow to be translated one into another, for they have developed in one culture. The ability to translate senses between different cultures depends on the degree of similarity and mutual assimilation of these cultures. Anyway, biological nature of man imposes some common traits on all cultures existing on the Earth, ensuring some degree of their similarity. One can imagine extraterrestrial civilisations that would not share with our civilisation even the biologically-conditioned meanings (for example civilisations where their members reproduce asexually). However, there will always remain a common ground in the form of purely scientific knowledge, because there is only one physical world, and the structures of conceptual maps of science in different civilisations must adhere to some degree to the structures of this world, and therefore they have also to adhere one to another.

Culture, defined above as a certain kind of conceptual network abstracted from the minds of its members, constitutes in modern societies an extraordinarily developed and differentiated object. However, as what we call culture appeared in our Gallery of Ancestors relatively recently, it had to evolve through much simpler stages from some germs which were not culture yet.

The origins of culture should be looked for in conceptual network formation during individual development – as opposed to the genetic transfer of conceptual network structure – which means here drawing information from parents and tribesmen. Imitation is the simplest example of this process. Young individuals learn in this way how to gain food and avoid danger, bring up progeny and respect hierarchical principles. Adult individuals can also assimilate new forms of behaviour from other individuals. It has been observed already in monkeys that if one of the members of a herd invents the method of washing grain in water, the whole herd does follows the example after certain time. This behaviour is of course transferred to next generations through young individuals as an already-established behavioural pattern of a given population. The complex of such behavioural activities can be enriched in time, as the population discovers still new possibilities of more efficient ways of gaining food or performing other, biologically important functions. This can consist in discovering simple tools, such as a stone used to break shells by a see-otter, a thorn used by some birds to pick insects out of slits in the bark, or the famous stick used by a chimpanzee to reach a banana located out of its reach. The enrichment may also proceed by finding new sources of food, removing parasites by way of mutual grooming and so on.

The above-mentioned kinds of behavioural activities – although frequently quite complex and leading to enrichment of conceptual network with new elements – are not the germs of culture yet. They fulfil functions that can be objectively explained with their biological significance. Thus, purposefulness determined outside conceptual network is their domain. These behavioural activities **mean** in an objective manner, at least in the context of cybernetic mechanisms established by biological evolution. These meanings lack subjectivity (or rather inter-subjectivity) and conventionality, i.e. the most immanent aspect of culture.

Let us imagine a herd of monkeys, where a loud scream means danger in the from of an approaching predator. Let us assume that monkeys can give out slightly higher-pitched or slightly lower-pitched voices, but that the pitch of the voice in a given case is accidental – monkeys sometimes scream in high pitched voice, and sometimes in low pitched voice, regardless of the type of predator approaching the herd. However, one of the monkeys (preferably high in the hierarchy of the herd, e.g. the leader of the herd) starts at some moment to react with a high-pitched voice to a panther and with a low-pitched voice to an eagle. As different strategies of escape are required with respect to different predators, such a distinction should be regarded as biologically purposeful. It is therefore probable that, after some time, all individuals in the herd will issue a warning against a panther in high-pitched voice and in low-pitched voice against a bird of prey, and that they will react adequately. Production of such a behavioural pattern is already a germ of culture.

One is tempted to ask: why should it be? What differentiates the above example from the previous examples, which we declined to include in the domain of culture-related phenomena?

After all, the behaviour serves in both cases the realisation of "objective" biological purposes. What is therefore the essence of culture? As I have already said – it is conventionality. It was accidental in the above example that a high-pitched warning voice was attributed to a panther, and a low-pitched voice – to a bird of prey. The correspondence could have been reversed and the biological purposefulness of this type of behaviour would not change. However, conceptual network of "culture" of the herd in question was enriched by two concepts, two meanings: "high-pitched voice – panther – climb onto a tree" and "low-pitched voice – bird of prey – escape deep into the crown". A high-pitched voice and low-pitched voice can be regarded as the most primitive names, meaning "danger from the ground" and "danger from the sky". Therefore, the example of formation of the germs of culture is also an example of origination of germs of language – in my opinion, their ways were inseparable almost from the very beginning.

Another example of formation of the beginnings of culture, devoid of linguistic "contaminations", is quoted by Stanislaw Lem. There, an ache in the right side of the herd leader's jaw makes it gnaw hunted birds with the left side of the jaw. This custom is taken over by the rest of the herd, acknowledging it since then as an obligatory custom. What had a biological sense in the case of the herd leader, who were avoiding stimulation of the aching jaw, looses this sense in other individuals. The purposefully grotesque character of this example allows it to present more emphatically the roots of culture. It is possible to invent numerous stories of the type. For example, tumbling of a sorcerer on the face caused by the fear of thunder could be received by the rest of his tribe as a worship of some supernatural being and thus give rise to a new religion. It is not important if the above descriptions of formation of primitive elements of culture are true (certainly they are very simplified). The point here is first of all to show origination of an intersubjective system of meanings devoid of any direct justification in biological functions, and thus, to show the emergence of a new level from the biological one, namely the cultural level that cannot be explained exclusively on the basis of the latter one.

As to the question why culture originated, one can say as much as in the case of the origination of life. Once certain conditions were fulfilled, culture had to originate as a result of development of psyche and social relations. This process was favoured by language that codified the conventional system of cultural meanings. Together with the appearance of self-consciousness, there arose the need of knowledge not only about the world, but also about oneself. This knowledge was constructed in an largely accidental manner from found elements – meanings, if the indifferent physical world expressed its silent agreement. Returning to the context of the present reasoning, however, I am trying to describe the mechanisms that led to the origination of conceptual network of culture. There is in fact one such a mechanism – if we remember various limitations imposed by the biological and evolutionary "medium" of culture – namely, chance. Chance and the active "material" that shapes culture, i.e. human psyche.

Conceptual network of culture does not start from nothing. Its development is based on a complex of purely biological meanings. However, although culture must take into account physical and biological realities of the world, it is not limited by anything within the borders fixed by them. Further development of culture is of course determined and "channelled" to a large extent by its history. This is understandable from the point of view of the conceptual network. New meanings are built in this network on the basis of the already existing concepts. Conceptual network therefore determines through its structure a certain limited range of development routes, while the selection of one of them is arbitrary and accidental. The evolution of conceptual network of culture is therefore a cumulative process, "remembering" its own history. The most paradoxical property of culture is the fact that fully conventional and accidental meanings become – once they have come into being – unshakeable and "true" in a absolute manner within a given culture. The long and inglorious history of religious wars testifies to the strength of meanings born from chance.

The accidental character and conventionality of meanings of culture explains the enormous quantity and diversity of cultures, that originated during the history of mankind, in spite of the unity of the physical and biological world we live in. We can see the divergence of cultural evolution in the variability of beliefs, philosophies, rites, customs, ethical systems, fine arts and languages. The more divergent two cultures are, the more untranslatable are their concepts and meanings – ways of seeing the world. These differences are enhanced by dissimilarity of ethnic languages, whose structures have also significant influence (Shapiro and Wharf's theory) on the formation of the ways of seeing the world (I agree with Quine's thesis concerning indeterminacy of translation of ethnic languages). Complete understanding of a culture by a man brought up in another culture is in principle impossible, considering the lack of mutual adherence of the conceptual networks of these cultures. One can of course try to "enter into the spirit" of the beliefs of a certain tribe, but if voluntary sacrifice of one's life for a local idol were the greatest privilege for these people, real understanding of this culture would be equivalent to great ecstasy at the moment of becoming "privileged" in their sense. Anyway, we would tend to suspect a man from the circle of the Mediterranean culture, willingly giving his life for some exotic godness, of mental illness rather than of understanding the discussed culture.

A simple derivative of the divergences in the evolution of cultures is the fact that the senses existing within separate cultures and related to religion, fine arts, social relations, language are purely intersubjective and disappear once one leaves a given culture. This is so, because the objective world is neutral in relation to cultures – it gives them full freedom, as long as they do not ignore the simplest facts of this world, like, for example, the necessity to nourish feed and to avoid predators as well as to protect oneself against cold in order to survive. A grater part (or even the majority) of a culture has therefore only such a connection with the external world as it must. Without this, a culture could not exist, since physics neither admits any ethical or aesthetic values, nor affirms any kind of divinity. However, there exists a cultural domain which opted for a significantly closer connection with the external world, and therefore resigned from the right to unlimited production of unrestricted and conventional meanings which become the exclusively valid and obligatory senses, once they have come into being. This domain is of course science.

Let us imagine a developing conceptual network, where new concepts are being created. Still other concepts develop on their basis. In this way, there come into being "conceptual chains," understood as sequences of primeval and derived concepts in the conceptual network. If the development of such chains is conditioned only by the already existing structure of meanings, the indeterminacy of their meanings – i.e. the impossibility to derive them from initial conditions – accumulates when we move along the chain. Sufficiently long chains – no matter how well determined concepts they are composed of – can "bend" and "deflect" optionally far away from a certain assigned line in the semantic space. The susceptibility to such "deflection" is the measure of indeterminacy of concepts (for example, it is relatively very small in mathematics). If several conceptual chains develop independently from some conceptual germ, for example from the "commonsensical" conceptual network, then, at a sufficient distance, these chains diverge sufficiently one from another, that is they fall into (semantic) contradiction with each other (different religions and philosophies can be quoted as an example).

The situation is different in natural sciences. Here, all newly-formed concepts are confronted with the external world without delay (or with a certain finite delay). For example, the theory of the Big Bang and the Steady-State Theory coexisted for some time, due to the absence of an experiment (or observation) which could falsify one of them. Thus, their choice reflected only the philosophical preferences of their inventors. However, the discovery of the background radiation judged in favour of the first theory. Therefore, the objective world fixes lines in the semantic space, which have to be followed by conceptual chains. Cosmic ether, phlogiston or vital force provide other examples of concepts that are already dead, once they have been falsified by natural sciences. Of course, mutual adherence of these lines and chains of concepts is never perfect. However, semantic indeterminacy does not accumulate here, together with the development of a conceptual chain. Instead, it remains more or less constant. Therefore, assuming that the external world is coherent and non-contradictory, it imposes coherence (of course only in an imperfect and approximate way) on the structure of conceptual network of science.

In the case of a broadly understood culture and the humanities (and also philosophies drawing on them), the situation is quite different. Indeterminacy, "blurry" character of concepts is large in these domains and the possibility of verifying their conceptual networks by confronting them with the objective world is very limited. Evolution of the conceptual networks of natural sciences which originated in different cultures must be convergent. The convergence is established by the adherence of their structures to the structure of the world. The methodology of exact sciences is able to compensate their deflections from this structure. Therefore, one faces here the mechanism of a negative feedback. The development of culture, on the contrary, consists in enhancing accidentally chosen direction to be followed. The structure already possessed by a culture determines the frames for its further evolution, while within these frames there rules chance. Therefore, one can speak here of a positive feedback. The further a given culture goes along a given route of development, the more decidedly it follows this route. Since the direction of development selected by a given culture is determined by chance, evolution of different cultures is divergent. The humanities, studying culture, create it at the same time. For natural sciences, the object of their studies exists independently. Culture and the humanities create by themselves the object of their studies. Although the biological constitution of man is their starting point, their object of interest reaches further, to what is added to this foundation. Discovery of structures is here indistinguishable from creation of structures. This creation is not limited in this sense that they do not have to adhere to anything apart from themselves. Therefore, completely freedom of choice prevails here. Culture, apart from the indispensable minimum, has no common points with the objective world. While we can compare science to a convolvulus, winding around a wooden pole (= reality), then, a complex of cultures resembles a bush bifurcating vigorously in all directions.

Two structures of physics developing independently (e.g. in different cultures or on other planets) could have different starting points (as they start from within a culture) and use different concepts. However, their structures would have to adhere, better or worse, to the structure of the objective world. As this world should be regarded as univocal, no matter how one understands this statement, they would also have to adhere one to the other, according to the principle of transitivity: if A adheres to C and B adheres to C then A adheres to B. The adherence of the structures of science to the structure of the world increases gradually in the course of the development of science. The more similar two independently originated structures of physics will therefore be, the better developed they are.

Therefore, the answer to the question if it is possible to construct two separate domains of physics and two domains of mathematics underlying them is both "yes" and "no". Yes, because they can start from different concepts, drown from the commonsensical language, initially use different assumptions and the stochastic processes that shape them may run along (to some degree) different routes. No, because structures built within them must adhere to the structure of the objective world, and therefore one to the other. This protects natural sciences against accidentality and optionality.

Evolution of conceptual network of culture possesses a symptomatic property, consisting in a clear tendency to absolutise linguistic names and their underlying meanings (concepts), as well as to attribute real existence to their designates. We see the naturally continuous world through the prism of language, i.e. in a sliced, discrete way. Names of language demarcate arbitrary categories within reality, "naturally" absent there. This leads inevitably to the creation of hyposthases and to attribution to the world of properties which are only and exclusively properties of language. The segregation of the totality of phenomena perceived by our senses (performing already a preliminary categorisation) into different compartments facilitates excellently our manipulation of the accessible information, and therefore the functioning and development of the conceptual network. This segregation, however, leads at the same time to distortion of our picture of the world.

The processes of autonomisation and "realisation" of words of language can well be seen in magic, religion and philosophy. Magic not only attributes some primitive necessity to the relation between names and their designates, completely ignoring its accidental character, but it gives words a performative power. The esoteric and necessary character of names manifests itself also in religion. Jehovah witnesses, for example, take pride in their knowledge of the true name of God. Finally, philosophy used to absolutise such names (and concepts) as spirit, matter, monads and so on. However, the world is not discrete like language, but continuous like the conceptual network. Therefore, recognition of the role of our language in creating hypostases has a fundamental significance for epistemology.

A great role is played here by science as well. Physics decomposed the concept of matter into field equations, functions of probability and other concepts, being as much real objects as products of our mind, so that matter has become in fact only an empty name. The process of decomposition of the concept of (self)consciousness (spirit) into other concepts performed by neurophysiology is also significant, though it has not advanced that far. The same applies to many other concepts and problems, with which philosophy struggled for centuries, and which were shown by science as empty or apparent. Therefore science, by escaping from the conventionality of culture, enables us to de-mythologise many aspects of our culture. At the same time, science suggests (in my opinion) that a more primeval and more adequate tool of knowing the world is found in conceptual network and not language. Language is such a "superstructure" added to conceptual network that allows a much easier, but also far from perfect, learning of and manipulation with different "elements" or "aspects" of this network.

Inaccurate "adherence" of language to the structure of the world can be seen at least in the multiplicity of ethnic languages, reflecting (representing) this world in many various ways. Accustomed to the European group of languages, we frequently ignore the enormous diversity of linguistic structures. They are so different that one could risk claiming that science in the form known to us could not originate at all in some of them (and therefore in the minds of the people speaking them!), or it would have a rudimentary character, adhering very weekly to the structure of the world. The fact that the Western science originated in the circle of the Mediterranean culture is probably a derivative of many factors, such as language and religion. It is also a matter of chance to some extent. However, there arises a question if such-and-not-another language could make it more probable for a given type of religion to originate, and help chance in this way. However it would be, one should not forget that language is not an autonomous object, but it constitutes a fragment of the conceptual network, reflecting at the same time the totality of culture. Therefore the present structure of language, science and culture is, in the last instance, a result of the evolution of conceptual network composed of a sequence of accidents and limitations.

As I mentioned, conceptual network of a given individual constitutes a derivative of many factors. Its shape is influenced by inborn neurophysiological predispositions, life experience as well as chance. The last two elements cannot be inherited genetically. Inheriting of the first one is also doubtful. It is a derivative of so great amount of genes, whose the co-operation yields so unpredictable results, that it is in fact difficult to speak about a somehow determined process of inheriting psyche. Therefore, the elements of one's (self)consciousness, and therefore of one's conceptual network, can be transferred (and propagated) in the genetic way only to a limited extent. A purely biological organism is "interested" only in the propagation of its genes. Man, however, has reached above the biological level, to the psychic and "psycho-social" level. Therefore, *Homo sapiens* realises not only purely biological purposes, but first of all psychic and cultural purposes, even if they stand in contradiction with the former ones. "Transfer" or "spreading" of one's own conceptual network, i.e. psyche, is one of such cultural purposes, by analogy to the transfer of one's genes. Therefore many people, having resigned from possessing their own progeny, sacrifice their lives to science, fine arts, religion or other kinds of social activity. Through broad social response, these fields constitute an effective way to "propagate" within culture of somebody's intelligence, personality or, in one word, the conceptual network. It is much more effective than the transfer of its elements by the genetic route. Because the identity of a man consists in the identity of its psyche, he can leave more after himself in the form of a scientific or artistic output than in the form of his genes. This does not mean that I reduce any intellectual and artistic motivation to the will to assure the perdurability of one's personality. What I say is that such a mechanism is able to overcome biological motivation, i.e. the drive to propagate one's own genes. This is the aspect of humanity which has not been taken into account by e.g. (an extreme form of) socio-biology. Man reached in his development above the purely biological level, which is utterly subject to the laws of natural selection, and climbed to the socio-psychical level (I use this name in opposition to the socio-biological level), which is ruled by its own laws. This is testified by the origination of ethics, which cannot be (in my opinion) completely explained on the ground of socio-biology (although socio-biology can point out the biological roots of ethics). Even the so-called reciprocal altruism (which is not altruism at all), frequently used in models of population genetics, assumes the existence of some minimal psyche which enables individuals to recognise other individuals in a herd. In my opinion, the socio-psychical level has created its own purposes and values which are independent, and frequently contradictory with the "purposes" of biological evolution. The central "directional mechanism" of the biological evolution is the following axis: fitness – lack of fitness. This axis is replaced in the psycho-cultural evolution with the axis: pleasure – lack of pleasure. Of course, the feeling of pleasure (underlined by the reward system in the brain) appeared in the initial stages of the evolution of psyche as a fitness-maximising mechanism. However, these two axes diverged to a large extent during evolution of humans (in highly-developt civilisations the natural selection is essentially weakened and leaves a broad space of manoeuvres for psychical and cultural mechanisms determining human behaviour) . The moment of this divergence can be regarded as the origination of a completely new, psychic level of reality. In my opinion, pleasure gained its autonomy as a psychic phenomenon together with the emergence of self-consciousness, as a result of directing the cognitive centre in our brains on itself. A different kind of evolution started at the moment when there appeared self-consciousness. The evolution, namely the cultural evolution, has nothing common (apart form obvious biological connections) with the natural evolution.

CONCEPTUAL NETWORK AND PHILOSOPHY

As I mentioned in the preface to this part, devoted to the conceptual network, I intended not only to popularise chosen achievements of modern science and to interpret them in my own way, but first of all to consider the philosophical implications of these achievements, i.e. look for philosophical conclusions that can be drown on their basis. For, there is no escape from philosophy. No matter whether one likes it or not, philosophy will always be situated between a relatively "certain" knowledge and nonentity and will constitute a certain "buffer" between well-separated areas of conceptual network and semantic emptiness. Philosophy rises questions that cannot be completely clarified by science, simply because these questions lie outside the (present) reach of its interest as well as of its methodology. In many respects, these are the most important questions concerning the nature of everything that exists and the status of our knowledge about this nature. Therefore, philosophy cannot be rooted out from the mind of man, a being interested in the world. This would be simply inhuman.

I sympathise with the famous Wittgenstein's statement that everything that can be said, can be said clearly, and if one cannot speak about something he should be silent. Unfortunately, this statement refers (in a negative sense) to a large part of the entire, many-centuries old, output of philosophical thought. Philosophy certainly has not fulfilled the hope placed in it. However, the postulate to turn away completely from philosophy on the way to satisfying our will to comprehend the world resembles the proposition to look for a lost thing under a lantern, because of better illumination. Such illumination, corresponding to the methodology of natural sciences, allowed them to reach the most spectacular successes in gaining knowledge about the world that have ever been available to mankind. What difference does it make, however, if we already know that what we always looked for in philosophy does not lie under the lantern? Is it not better, therefore, to roam blindly with a very slender hope for finding at least a trace of understanding, instead of basking in the lamp light, with full certainty that what is looked for will not be found? However, such blind search is regarded by most scientists (which is not devoid justification) as a waste of time, for if we even succeed sometimes in finding something, we will never be able to gain certainty in the darkness whether this is what we mean or simply an illusion of our mind.

Does not there exist another alternative? Maybe, by widening the circle of light still further through new scientific methods and discoveries, we will be finally able to discern in the shadows at least some distant objects, hidden in the darkness until now? And it may be that, dazzled with the light of the lantern, we do not notice that this has already happened?

Representatives of natural sciences hold as a rule in low esteem a predominant part of philosophy, especially its so abstractive branches as ontology and epistemology. For most of my friends, their contact with philosophy terminated after passing of the required examination during studies. At the present moment, they are too absorbed in their own fields to have time for anything else. Those few who can afford a broader look avoid philosophy like fire, regarding it to be as fruitless thing as magic or astrology. Agreeing with me in other fields, they are surprised with these interests of mine. And what is most important, a huge number of premises authorise them to adopt such a sceptical stance.

Despite its more than two thousand years of existence, Western philosophy neither reached a solution of any of its most fundamental problems, nor did it work out satisfactory criteria for estimating or selecting particular philosophical concepts. Its main achievement is the exhibition of its own limitations and apparentness of formulated questions and problems. Philosophy lacks methodology – such as sciences have at their disposal – and the concepts used by it are so blurred and undetermined that logic cannot be used in a sensible way to assess them. On the other hand, a huge amount not only of facts, discovered by science and of potentially enormous significance for philosophy, but also their significance and interrelation with other facts – all this remains disregarded, which is due to gaps in the knowledge of "professional" philosophers, who are, after all, not specialists in particular fields of natural sciences. Scandalising a little bit, one could therefore ask if it would not be better to take philosophy away from philosophers and to hand it over to representatives of natural sciences?

Of course, each man (including a natural scientist) undertaking philosophical problems automatically becomes a philosopher. Therefore, the question formulated above is partially a rhetorical one. But only partially. For, the point is the practiced kind of philosophy and the manner of carrying out the practice. One can continue the philosophical tradition of many centuries, where philosophy is treated as the queen of sciences and its practitioners remain in the permanent contact with the Absolute, which is why they are in the possession of Truths independent of our knowledge in other fields. However, such an approach seems to be futile and, at best, threatens with abandoning philosophy itself in favour of dealing with its history or fragmentary contributions to already existing conceptions. At worst, the approach leads to autistic separation from reality in the name of unrestricted visions, whose sole verification would be found in the deep conviction of their inventors as to their validity (leaving aside a more practical danger, namely that of driving to despair the people whom their desire to base their knowledge on any relatively reliable methodology would, in other circumstances, lead to turn to philosophy). A modest alternative would consist in attempting to explore the huge knowledge accumulated by natural sciences, particularly in the field of physics, neurophysiology, theory of evolution, molecular genetics, mathematics, cybernetics, informatics and theory of information. This would allow a philosopher to use many concepts in a way that takes into account the context of modern knowledge about the world, instead of using these concepts in the meaning formulated over two hundred years ago.

Spirit (psyche) an matter are such concepts, the meaning of which changed enormously as a result of development of science. These concepts form the basis of the two greatest, opposed systems in the history of philosophy: idealism and materialism. The first assumes that the material world is secondary in relation to (self)consciousness, or even that it is simply a product of mind. The second one, inversely, claims that matter is primary and assumes that consciousness is a derivative of biochemical and physiological processes occurring in human brain. The opposition of these systems is based on the opposition of the concepts of matter (external world) and mind (self-consciousness). However, the development of theoretical physics and neurophysiology resulted in enormous change in the meaning of these concepts. Even more, they have become "de-defined", decomposed into other concepts and therefore devoid of meaning, in fact.

In the commonsensical understanding (that constitutes the starting point for philosophy), matter is something tangible, possessing a location in space, weight, consistency, colour, dimensions – it is something that can be transformed in different ways, but not annihilated. All these attributes disappear in theoretical physics. The essence of so well known and clear concept becomes here very mysterious and exotic. Matter can vanish, transforming itself into energy, in accordance with the famous Einstein's formula: $E = mc^2$. In quantum mechanics, a spatial location of an elementary particle, for example an electron, can be determined only approximately, due to the Heisenberg's indeterminacy principle. Similarly, there is no sense to speak about the colour of an atom or electron at the moment, when they do not emit the "carrier of colour" – the electromagnetic wave (the so-called "colours" of quarks – building blocks of protons and neutrons – are only facetious names for certain property that has nothing to do with colour; the same applies to "flavours" of quarks). Particles of matter sometimes behave not like particles, but like waves. An electron in an atom is not in any concrete place (before it is recorded by an appropriate detector), but is as if "blurred" (delocalized). Spatial distribution of this blurring – corresponding to the probability of finding this electron in a given point of space – is described by the so-called wavefunction. However, we are not able to decide if this wavefunction is a material object or a construct of our mind, and therefore we cannot draw a sharp border between reality and the content of human consciousness! According to the general relativity theory, matter is not only the substance filling space, but it also shapes this space, causing its curving proportional to mass. If the general energy of the Universe is zero, then, space can be treated in a sense as "negative" matter! We are no longer able to recognise in this entanglement of absurd and curiosities our familiar concept of matter formed on the basis of experience of reality through our senses, the concept that formed the subject of philosophy for two thousand years. Such an apparently clear and obvious concept, after closer examination, escaped our grasp, leaving an empty name.

The case of the concept of (self)consciousness is similar, although here the process of its decomposition into simpler concepts has not advanced that far. We know already from neurophysiology that the groundwork of our psyche is formed by processes (e.g. polarisation and depolarisation of the cellular membrane of the appendages of nervous cells: dendrites and axons) occurring in the network of neural cells in our brain. Consciousness depends on its material background in important ways. A severe brain damage can result in complete change of personality, while cutting the *corpus callosum* connecting both brain hemispheres (the operation of callotomy applied several decades ago to cure epilepsy) leads to formation of two, almost entirely isolated centres of (self)consciousness in one brain. Neurophysiology has also advanced much in studying the mechanisms of perception and of processing sensory stimuli as well as in explaining the essence and the manner of memory records. Not much place has remained for "pure (self)consciousness" or the "spirit in its immanence". It is probable that it will finally be possible to describe the very "core" of the (self)consciousness – that is the feeling of one's own "self" – in biological (and cybernetic) terminology.

As we see, consciousness appears in neurophysiology to be a product of matter (that is the neural network in the brain), while matter (e.g. electrons) seems in physics to be still more and more shaped by mathematical constructs of our mind. Therefore, it may appear that there is no sense to choose between materialism and idealism if we cannot show any principal difference between matter and consciousness, all the more because these concepts seem to be empty names, corresponding to nothing univocal and absolute. If it happens so, then, the greatest problem of the "queen of sciences" will probably be resolved (and in fact acknowledged as nonsensical), not in the course of the "development" of philosophy lasting for over two thousand years, but thanks to flourishing of physics and biology in the 20-th century.

Immanuel Kant's attempt to decide whether the Universe had a beginning or if it existed eternally is another example of anachronisms in philosophers' considerations when they use concepts formed on the ground of colloquial language (and not on that of terms defined by natural sciences). Kant, on the basis of an analysis of the concept of time in its common meaning, came to the conclusion that each of these two possibilities is internally contradictory. Therefore, neither is it true that the Universe existed eternally, nor is it true that its existence started at a certain moment in the past. The absurdity of this conclusion constitutes a flagrant example of the futility of application of logic to philosophy, as well as of philosophers' excessive confidence in the power of their discipline. It appears, in fact, that philosophy is not at all the "queen of sciences", announcing the truths about the most important things, while the rest of sciences deals with down-to-earth problems, unworthy of the attention of the queen. To the contrary, the latter must wait submissively for the advance in natural sciences to draw on their new discoveries, concepts and paradigms. One of the inventors of the theory of black holes, Stephan Hawking, showed that there may be no sense in speaking about time before the beginning of the Universe. Time "emerges" in some region of space-time (close to what we call the Big Bang) from one of spatial dimensions. Therefore, Kant's considerations would appear simply nonsensical, because of his lack of knowledge as to the nature of time. Again, it was not philosophy that approached this problem in a new perspective.

Einstein's relativity theory is another flagrant example of outdistancing philosophy by science. This theory – that completely changed our opinions in the topic of time, space, matter, energy, origin of the world and so on – not only was not created by philosophers, but it was also rejected by some of them (e.g. by Bergson) for a long time, either due to commonsensical thinking or to misunderstanding of its mathematical formalism. Fortunately, the world effectively resists Bergson's intuitions, and the methodology of science has more objective research tools at its disposal than somebody's private conviction, no matter how deep it may bee.

There exists a certain important difference between science and philosophy. The development of sciences is cumulative, that means new knowledge does not replace completely the knowledge already possessed, but is rather added to it and constitutes its supplement and development. In philosophy, simplifying matters slightly, it is necessary to start again and again from the very beginning. If a representative of natural sciences tries to express his opinion on philosophical topics, he is frequently charged with incompetence, because, for instance, he is not well-acquainted with the subtleties of a philosophical system of a Mr X, or he does not know the fifty distinguished sorts of tautology. However, there arises a question: **does he need such knowledge useful for anything**? For philosophy, loosing contact with reality, begins to deal more and more with itself, with its own products and schemes established by its tradition. It encystes in an autistic isolation from the world. In order to attain anything new in philosophy, it seems reasonable to get rid of this entire ballast and start from the foundations, i.e. from the knowledge, concepts and paradigms offered by natural sciences.

Great possibilities in this respect seem to be hidden in neurophysiology. All people (including philosophers) think with the aid of their brains. However, did the inventors of the concept of linguistic thinking, eulogists of the universality of logic or philosophers dealing with epistemology undertake the effort to investigate what conclusions concerning the nature of human thinking result from what is already known about the functioning of the network of nervous cells, mechanisms of memory formation or perception and treatment of sensory stimuli? No, they did not, which testified by the products of their speculations. Analysing the manner of formation and the nature of meanings in human brain, one can say much about the essence of our thinking and understanding of the world, about the possibility of reaching the "absolute truth", or about the manner of existence of general (universal) and abstractive beings. It can be stated – in a not excessive simplification – that the entire philosophy can be reduced to neurophysiology. Anyway, neurophysiology can say more about purely philosophical problems concerning human mind than the traditional philosophy based on speculations. Therefore, if they by definition investigate human mind, then, **are philosophers competent in philosophy?**

Undoubtedly, the above-sketched picture is slightly simplified and one-sided. Many philosophers are knowledgeable in the field of sciences, but they deal mainly with philosophy of nature, mathematics or science. Speaking about philosophy and possibilities of its "salvation" by sciences, I meant first of all the two most important fields of philosophy: ontology and epistemology. Such branches of philosophy as philosophy of man, philosophy of culture, aesthetics and ethics are not, by their nature, subject to the methodology of sciences, and therefore they will never be able to count on "salvation" (and they are likely to be replaced in the future by some moderate version of socio-biology). Their sin, shared with the rest of philosophy, consists in speculative freedom, lack of verifiability of statements, creation of empty concepts (hyposthases), which can generally be reduced to simple lack of any reliable methodology. In physics, nobody is interested in private prejudices and views on life of a given scientists. It is important if his experiments are repeatable and his theories predict correctly the results of further experiments. On the other hand, a philosopher is not able to say anything beyond the internal melody of his soul. The inventors of the Stationary State theory (Hoyle, Bondi, Gold) – who did not like the theory of the Big Bang for philosophical reasons – would have probably defended firmly their position until today, but for the fact that their theory yielded predictions contradicting astronomical observations made later on. Einstein – probably the greatest genius in the history of science – believed that the Universe was unchangeable in time and therefore he introduced the so-called cosmological constant into the equations of the general relativity theory, which "saved" the Cosmos from expansion or contraction. Einstein removed the cosmological constant as unnecessary, however, after Hubble's discovery of the escape of galaxies, that implied expansion of the Universe. Similarly, he was not right (at least according to the currently predominant opinions) when he questioned now commonly accepted indeterminism of quantum mechanics (due to his philosophical preferences as well). These examples show that even the greatest genius can be wrong and make mistakes if he relies exclusively on his own belief and intuition and does not confront his opinions with the objective world. However, this does not bother philosophers who – as Hegel or Theilard de Chardin – create systems that surpass the greatest physical theories with their impetus, but whose content is an arbitrary mumbo jumbo.

Therefore, the answer to the question if philosophy can be saved is: yes – at least partially. However, philosophy has to renounce the sin of bumptiousness and ignorance, resign from visionary attempts at imposing on the world what it should be like. Instead of this, philosophy should try to find out what can be said about this world in a possibly certain way. For this purpose, philosophy must rely on the only field of human activity that possess a methodology indispensable for a relatively objective description of the world, namely on natural sciences. Otherwise, philosophy may expose itself as futile or comic, which is even worse.

I would like the present book to become at least a small contribution to the style of philosophising comprehended in this way, to building of a bridge between natural sciences and this part of philosophy which is motivated mainly by cognitive honesty and not psychical comfort. Therefore, I would like to treat philosophy as a way of satisfying normal human curiosity, and not as a psychotherapeutic method which is to release us from fear (not less human) of the nonsense of the Cosmos surrounding us. At the same time, I am aware of the danger accompanying such an enterprise. If one wants to touch both banks and overestimates his own strength or underestimates the distance between banks, it is easy for him to find himself in deep water. It is difficult to rely then that somebody will offer a helpful hand from either of the banks. Scientists can turn away from speculations that are futile in their opinion, while philosophers may feel disgusted by the presence of a barbarian in the sanctuary. I wrote this book with the hope that this will not happen.

* * *

I think that it is high time to leave this too-long introduction and finally state in the explicit way what philosophical inclinations of the present book – and particularly of its last part devoted to conceptual network – I regard to be so important that they constituted one of the main stimuli to write it. At the same time, I do not say that I insisted on avoiding any philosophical considerations on the above pages, in order to manifest my hidden intentions at the very end. To the contrary, everything I want to say has already appeared – more or less explicitly – on the pages of this book. Now, however, I intend to order and summarise various philosophical aspects, so that my aim would become possibly clear and comprehensive.

In this, probably the most inappropriate place, I will indulge in yet another digression. I will use the slightly outdated (as I mentioned above) opposition of spirit and matter. The enormous development of physics in the 20th century told the humanity more about what we intuitively understand as matter than the whole earlier development of mankind. I quoted just a few, out of many, examples to support this thesis. All of them illustrate a deep transformation of the concept of matter, but also of related concepts: space, time, causal relations, the beginning and nature of the Universe, in short, a transformation of our comprehension of the broadly understood physical reality. I think that this topic deserves a separate book which I could not write, as I do not possess sufficient knowledge. I tried to point out some problems in the first part of this book, devoted to physical evolution. However, what does this has to do with the present subject of consideration?

Well, in short, I tried to show in this work what modern science – especially the theory of evolution and neurophysiology – can say about (self)consciousness as well as about related matters of cognition and sense (physical, biological, psychic and cultural). For I am convinced that the conclusion coming from what we already know about the functioning of our nervous network, about its phylogenetic and ontogenetic evolution (we owe all this to a great development of the modern biology) have equal impact on the understanding of the nature of our mind, as the revolution in physics had on our comprehension of matter. Perhaps the knowledge about the essence of (self)consciousness is even more important, for we can "reach" what we usually call matter only through the prism of mind. Knowledge of the history of the development (in the process of biological evolution) of the mechanisms of perception, integration and association of environmental stimuli is of principal importance for the cognitive possibilities of man and it determines the degree of validation and the status of this cognition.

Additionally, I would like to say something about psychical senses, by referring them to physical and biological senses. The manner in which subsequent layers of reality emerged from lower layers, namely the biological layer from physical layer and the psychic-cultural layer from biological layer, this manner – fixing in a sense the nature of these layers – is able to say much about the essence of the meanings (purposes, senses) characteristic for particular layers and decisive for the fundamental differences between them. It is hard to overestimate this kind of knowledge in the formation of the picture of reality we live in.

* * *

The fact that the meaning of conceptual network is of connotative nature has extraordinary importance for understanding of the essence of human (self)consciousness as well as of the degree of validation of our cognition and the status of the world picture created with its aid. As we remember, connotative character of conceptual network results inevitably from the structure of its physical "carrier" (background) – the neural network. However, we can forget about that for a moment, for our mind has direct access only to the conceptual network. This will allow us to see –

without superfluous "contaminations" – how the world looks in the perspective of the conceptual network.

In fact, complexes of concepts (as I mentioned, I regard sensations as a result of a "stimulation" of particular concepts in the brain by external stimuli) are the only things that are really **given** to us (the only information we have). We can draw conclusions about the external world only on the basis of the picture of reality shaped within the network built of these concepts. The same statement applies to our percecption of consciousness, mathematics, logic, language, scientific theories, the entire area of cultural senses and, in general, to everything that can be "reached" or thought in any way. One should realise that, in spite of huge variability of the entire accessible to us – understood in the broadest sense – reality, at the level of the conceptual network it constitutes only a complex of categorically identical concepts.

In the connotative conceptual network, any senses and meanings exist only and exclusively **in relation** to other senses and meanings. The status of our opinions can only be **relative**, dependent on the context of the senses "surrounding" them. For concepts do not designate directly any "objects" existing outside conceptual network – the meaning of concepts realises itself only in relation to other concepts. Impossibility of announcing any truths which are in any way **absolute** – valid regardless of their reference, any senses existing **beyond** the area of conceptual network – is the very first conclusion for philosophy drown from the conception presented here. However, this in fact means a suicide of the traditionally understood philosophy.

Because, (almost) the entire history of philosophy consisted more or less in attributing absolute sense to some concept (it could be matter, spirit, monads, four elements, will or logic and language), in order to create an axiological bedrock for a description of the world. All ontological systems tended to separate a fragment (concept) of conceptual network and to promote this concept to the dignity of an Absolute, as well as to make it primary in relation to the rest of phenomena. However, while studying a philosophical system based on an initial absolutisation of some concept, the process of analysis leads finally to the concept itself, and then not only all attributes of its absolutisation, but also the very concept dissolve into semantic emptiness, leaving only an empty name. The above-discussed semantic deconstruction into primary factors of the concepts "spirit" and "matter" – which has already taken place within natural sciences – can serve as an example.

Therefore, the belief in absolute truths is a beautiful utopia founded by our mind. If one wanted to say consequently "the truth and only the truth", one would have to remain silent. However, this obviously does not satisfy us. We would like to say something about the world, even if this will have to be devoid of legitimacy, to some extent, i.e. not absolute and internally contradictory. There is nothing wrong in this, as long as we remember that the cognitive status of our statements is limited and relative. An advantage of the conception of conceptual network can be found in the fact that such a state of affairs can be easily recognised within this conception.

The need for the absolute knowledge is of course easily understandable for psychological reasons. A quest for an unshakeable bedrock of universal truths always accompanied human cognition. Unfortunately, this cognition led humanity to gradual dissolution of Absolutes. The history of science and the philosophy related to it was a sequence of gradual de-absolutisation of everything. Since its beginnings, modern science confronted every intellectualist with the necessity to choose between cognitive honesty and psychically comforting philosophy. Since the end of the Middle Ages, the two mentioned options were never reconcilable.

The medieval picture of the Universe favoured man vary much. The Earth was situated centrally. It received light and warmth from the Sun, while the Moon dispersed darkness of the night and the sphere of fixed stars showed the way to wanderers. Finally, the heavens behind it, i.e. the Lord's headquarters, whose watchful attention was directed at the centrally situated Earth and its inhabitants. A small, cosy Cosmos, created for the needs of man and favouring him in each respect. The central location of our planet was identified with the central importance of man in the Universe.

Copernican revolution shook this crystal-clear picture up. The vision regarded as the only possible and absolute in the Middle Ages was devastated. Man was forced for the first time to shift the weight of importance beyond himself, and to resign form some Absolute. Copernicus was followed by Kepler, who decided on the basis of observatory data to replace circular orbits of planets with elliptical ones, although the circle had been regarded to be a perfect figure since the antiquity.

Subsequent discoveries changed dramatically the medieval picture of the world. Let us take a look at the current image of our Universe. Billions light years of dire emptiness, rarely punctuated with stars and planets, meaning nothing and serving nobody. And man, a creature that originated by blind chance in the evolutionary lottery, living on a small planet revolving around one among billions of stars. Even the picturesque *finis mundi* has been replaced with gravitational collapse, thermal death of the Universe or transformation of the Sun into a red giant and burning the Earth to naked rock. Therefore, science replaced the only possible, perfect Universe at man's service with an optional and accidental Universe, filled with light ages of nonsensical emptiness.

Not only Cosmos has been deprived of the seal of Absolute by science. Physics questioned the semantic autonomy of many concepts taken from colloquial language, seemingly so obvious and "tangible". Newton shook the absolute character of space, Einstein – of time. Therefore, even the most fundamental elements of the world we live in do not constitute any longer a sufficiently solid bedrock for the intellect. They are neither certain nor unshakeable. This follows from the fact – to be discussed in a while – that the manner of integration of sensory stimuli by our brain appears to be the only gauge of their "absoluteness". Anyway, the same applies to other "absolute" categories and therefore science quite easily shakes their "unshakeability".

Let us go even further. Quantum mechanics negated the commonsensical concept of causal relations. Matter also underwent de-semantization in physics. Presently, it is not possible to fix the border between what is a "real" element of the world and what already constitutes a mathematical model, a "product" of our mind. Therefore, even the absolute character of the border between concept and its designate has been questioned. It is difficult to feel intellectually safe in the world devoid of any absolute support, where nothing is certain and everything is relative.

The advancing process of de-absolutisation has not omitted Man either. Instead of a being made in the likeness of the Creator, we have now a biological species, of which apes are the closest living relatives. (Self)consciousness has not resisted the destructive (for the Absolute) power of science either. Observation of drastic changes in psyche after a brain injury, great progress in investigations of neurophysiological mechanisms of perception of external stimuli, of association of neural impulses coming from different parts of the brain and of physiological background of memory – all these developments have left very little place for the spirit in its immanence, for an autonomous, independent and indivisible consciousness (mind). A few more examples can be listed: de-absolutisation of logic (Russell) and of formal mathematics (Gödel), negation of strict determinism (it was an Absolute for physics in the 18th and 19th centuries) and of the "absolute" phenomenon of life (transition from holism to reductionism).

Logical consequences of this approach seems to be the following: stripping all meanings in general off of legitimation in the absolute sense, denial of the existence of any truths that would be valid regardless their reference. In this formulation, even the Pascalian reed appears to be a delusion – no absolute knowledge can be given to us. Here is also where Occams's razor principle finds its ultimate fulfilment – there is no need to assume that anything exists (in the absolute way). That is all on the subject of absolute truth. Now I will pass on to further consequences of the structure of conceptual network proposed in this work.

First, there is a need to explain the reciprocal relation between conceptual network and language. Language is a "discrete" formation. It is composed of discrete names – symbolical labels of objects designated by them. On the other hand, concepts in the concept network merge one into another in a more or less gradual way (hills in a landscape can constitute some analogy here), and possible discontinuities are hidden as a rule at lower scales of magnitude, like grains in a photograph. Reciprocal correspondence between language and conceptual network occurs in such a way that linguistic names correspond to the areas of conceptual network (concepts) that are most intensely marked, separated, or precised by the conceptual context. Linguistic names are granted to sufficiently clear and "individual" concepts. On the other hand, a huge number of blurred and weekly-defined concepts – not more than just allusions to senses – do not find their counterparts in the sphere of language.

Language is secondary with respect to conceptual network, because both its names, "objects" designated by them, and the relation of designation itself are all built of concepts. Therefore, language is contained within conceptual network and forms its part. At the same time, language is an excellent tool operating within this network, allowing one to operate efficiently this network as a whole. If we compare concepts to hills in a landscape, then, language – constituting a kind of map with the names of hills marked on it – facilitates significantly orientation in the territory. However, concepts (and not correspondence to some external objects, as it may seem at the first glance) are decisive for the meaning of language names. Without the conceptual "lining," language would constitute only an empty structure of names ruled by some formal rules. In this sense, language seems to lack semantic autonomy. Linguistic names correspond to appropriate concepts only in an approximate and equivocal way. Therefore, although discrete structure of the language induced many philosophers to consider it as a universal tool of cognition, constructions composed of bricks of linguistic names are far from coherent and non-contradictory.

Connotative character of concepts leads inevitably to the conclusion that perfectly (infinitely) sharp concepts are impossible. On the other hand, such concepts are demanded by logic and mathematics. Therefore, these disciplines are not – contrary to frequently formulated opinions – ideally precise, strict and autonomous. Neither are truths pronounced by them. The classical logic attributes value 1 to truth and value 0 to falsehood. As ascribing the value of absolute truth (or absolute falsehood) is burdened in the context of conceptual network with internal contradiction, the veracity of statements can be determined here by any number from the range between 0 and 1, but never with zero or unity. Therefore, not only that there are no sentences which are true in an ideal and context-independent manner, which may be relatively easy to accept, but there are also no statements that are completely false (it is always possible to find an interpretation in which even a most nonsensical sentence can be attributed with some meaning). The world evades logic, because the latter is a form of language. Conceptual network cannot be described correctly by logic executing operations on names, due to its limitations resulting from inaccuracy of the reciprocal correspondence between conceptual network and language. Early Wittgenstein absolutised logic to such a degree that he ascribed to it the status of the structure of cognition, even more – the structure of the world. From the point of view presented here, logic – as it frequently operates on concepts (names) which are not completely precised – should be recognised as useless for philosophy, where indeterminacy of concepts is particularly large, and should be replaced with something like a conceptual vision, which should aim to fit the system of meanings in the receiver's mind.

Neither in mathematics – aspiring together with logic to absolute strictness – is it possible to avoid internal contradictions and paradoxes, due to connotativity of conceptual network. The problem of *continuum* can serve here as a leading example. It can be formulated in many different ways, e.g. in the form of the paradox of Achilles and the turtle. Here, I will express the above problem in the following form: why does multiplication of zero dimensions of points in a segment of a straight line by an infinite number of points returns at different time 3 or 5, in centimetres or other units? The problem is so old that the majority of mathematicians decided a long time ago to skip it by subsuming it to the category of commonsensical thinking. From my point of view, however, the problem of *continuum* constitutes a strong argument for the thesis that general schemes of seeing of the world that have been worked out by evolution are not completely coherent and consistent.

The concepts used by logic, mathematics and science are not ideally sharp because they come from concepts formed within **colloquial** language, appropriately defined and precised to adapt them for these disciplines. For obvious reasons, the process of precising cannot be continued infinitely (this would require infinite time and infinitely great number of operations). Therefore, there will always remain in every concept certain margin of significative non-univocality and blurriness (I discussed above the case of the concept of number 2 in this light).

Science is related to the external reality in such a way that it tries to cover this reality with a network of concepts which adheres possibly well to the structure of the world. Particular scientific disciplines can be seen as kind of conceptual maps, representing adequate aspects of reality. Two kinds of problems appear in the scientific representation. Firstly, adherence of conceptual-network structure to the structure of the world is always limited and approximate. Secondly, reciprocal translatability of one map into another map is far from perfect. There are many problems at the "meeting points" between maps. One can find many examples, for instance: the impossibility of co-ordination of quantum mechanics with general relativity (both theories cannot be true at the same time), incompatibility of thermodynamics with dynamics (the first possesses a clear arrow of time, while the second does not distinguish between the forward and backward directions in time, i.e. it is symmetrical in relation to them), problems with the unification of all four forces known in physics (electromagnetic, weak, strong and gravitational forces), or the incommensurability of biology and physics (it is not possible to express biological "senses" or "purposes" in the language of physics). A simple analogy for the description of the world by science can be found in the process of covering of the surface of a globe with certain number of flat maps. To describe the globe as accurately as possible (we assume that the adequate measure is the distance of the edges of a map from the surface of the globe, assuming that the centre of a map touches this surface) it is necessary to use a possibly large number of possibly small maps. However, this would be very inconvenient in practice. Apart from this, we would obtain a huge number of fragmentary descriptions that would not be completely compatible, instead of one or a few unified descriptions. The number of maps should be decreased, in order to obtain a generalised description. However, this is possible only at the cost of description accuracy (defined as the distance of the edges of a map from the surface of the globe). Apart from this, at least two maps – projections of the globe hemispheres, for example the northern and southern hemisphere – are necessary for a complete projection of the globe. Therefore, a complete unification of the description cannot be performed for principal reasons. This is due to different nature of the represented and the representing, i.e. of curved surface of the globe and flat surface of maps (or: of reality and conceptual network). For this reason, in my opinion, the Great Unification of All Sciences is an obvious utopia. Aspectual character will always remain an inseparable property of our perception of the world. While valuing science, one should not over-appreciate it, for one runs the risk of falling into the trap of absolutism.

The complex of properties of conceptual network implied by the character of the neural network underlying it is not limited to the property of connotativity. The neural "roots" of conceptual network make conceptual network appear to us as a dynamic and diachronic object. It undergoes continuous changes and development, both in the process of biological evolution and during individual development of particular persons. In both these processes, conceptual network starts its developmment practically from nothing.

A huge majority of the philosophy of cognition, of mind, of perception or language approached human mind in the synchronic perspective, as an established in advance and in principle unchangeable complex of instructions and categories governing the perception of the world. This leads to inevitable absolutisation of our cognitive apparatus, as it took place in the case of Kant's categories of pure intellect or in the case of Wittgenstein's linguistic thinking. However, the very realisation of the fact that conceptual network evolves during individual development of man suffices to shake this absolutism. It may remains a debatable problem whether a child who has not mastered language yet is already capable of thinking. However, the child possesses certainly some germs of thinking which are not separated by any qualitative difference from "real thinking", that will afterwards develop from these germs. We can regard linguistic thinking as some convenient and frequently used **form** of thinking (if we consider language as a certain tool, allowing one to operate efficient with concepts). However, we should not identify it with the essence of thinking **in general**.

A similar situation occurs in the case of Kant's categories. Some germs of these categories are probably inborn in man. However, this does in no way testify to their necessity and absolute character. For, these categories originated in the process of evolution as mechanisms of integration of sensory stimuli which are to serve the survival of evolving organisms in their environment. This does not imply – which probably seems obvious – the categories of our mind are perfect and universal as cognitive tools. They are conventional at least in this respect that it is possible to formulate the world in other categories, completely unattainable for **our** mind. This seems to be testified by the progress of modern physics, born in a still greater and greater pain and evidently tending towards increasing subjectivisation (quantum indeterminacy, the problem of the arrow of time or the degree of ordering of a system in thermodynamics, the anthropic principle). The above problems are most probably the result of this trivial fact that evolution created our mind to enable us to hunt for mammoths and not to develop philosophy or science. The development of culture (at least of many of its aspects) that took place afterwards was not planned by and to a significant degree indifferent to biological evolution, just as the pattern of colourful spots on a shell of a snail – if it is devoid of adaptive functions[99] – has no significance for evolution.

The "connotative" and diachronic function of the neural network, although it determines the general structure of the conceptual network, does not in itself suffice to explain completely the nature of the latter. If the network of neural cells is seen as a kind of form, then, there is still need for a content filling this form. The generall structure of neural network determines certain permissible range of possibilities of flow of neural impulses, but the number of actual combinations of impulses within this range is practically unlimited. The question arises: **what** does in fact cause the stimulation of certain and not other neural circuits and in such and not another sequence. In other words, the problem can be reduced to the essence of the primeval source of concepts, or to the "substance" on the basis of which they are built.

This primeval "substance" from which concepts are created is nothing else but stimuli (sensations) coming from sensory receptors. They decide (directly or through memory records) about concrete routes of neural impulses in the neural network. Sensations, e.g. pictures, constitute some form of stimulation of complexes of neurones with a characteristic spatio-temporal pattern. The structure of neural connections in a given part of the brain – e.g. in the visual cortex (and also their communication via the visual nerve with photosensitive cells in the eye) – ascribes a given complex of impulses to a visual sensation and not to a feeling of pain, for

[99]The origination of culture in general could be conditioned by evolutionary (socio-biological) reasons. However, many particular properties of a given culture are purely accidental (or conditioned by socio-psychical mechanisms). Similarly, the fact that a snail **does** possess spots on its shell can have some biological justification, while the given concrete pattern of these spots is simply indifferent from the point of wiev of natural selsction. In the case of human societies, this range of indifference leaves a free space for cultural evolution.

instance[100]. The structure of connections is also a determining factor deciding whether a given picture has mostly spatial character or whether it is focused on movement perception. In turn, a neural structure of a higher order – within which patterns of association of particular sensations are realised – underlies concepts, namely primary concepts. Integration on a still higher level is the source of secondary concepts. Speaking in general, the "substance" of all concepts is constituted by a derivative of sensations and their processing by "higher" centres of the brain (one should remember, however, that sensations constitute sets of neural impulses received from receptors that have been integrated by, mostly inborn, neural mechanisms).

Of course, we do not "perceive" consciously all the pictures (sounds, flavours) co-creating a concept which appears in our consciousness, just as we are not aware of all connotations of this concept with other concepts. To do this, we would have to grasp mentally our entire conceptual network, together with memory, which is of course impossible (due, at least, to its size). Thinking about a horse, we do not recall all horses we saw in our life, or all information about the anatomical structure of the horse, its taxonomic position, manner of breeding, horse races and so on, although it is the whole complex of these factors that conditions our understanding of the concept "horse".

I would like to treat the topic of the sensory nature of concepts in a little bit broader manner. The variety of different sensory stimuli in man is quite large (one should also remember the internal senses, "measuring" e.g. muscle tension). For the sake of simplicity, I will use in the following considerations mainly the example of sight – undoubtedly the most important sense of man[101]. In this case, some combination of stimulation of visual cells in the retina of the eye is the substance for newly-originating concepts. However, the complex of signals reaching the surface of retina cannot be in any case identified with the picture which appears in our consciousness. For these signals already undergo significant processing in the retina itself and, afterwards, in the receptor part of the brain cortex (visual cortex).

Stimulation of each visual cell by absorbed photons is completely equivalent to stimulation of another cell. All cells are identically predisposed "at the input" to form a picture, similarly as

---

[100]If one were to cut the auditory nerve going out from the brain and connect it with the nerve going out from the eye, then the portrait of Mona Lisa would be received in our consciousnecc as a cacophony of sounds.

emulsion grains in a photographic film. In both cases, the "stimulation" of a photosensitive element constitutes, as a simple physicochemical process, a direct consequence of absorption of a photon, and therefore the arrangement of stimulation points on the retina or the photographic plate (that is, what we can call the "primary picture") corresponds simply to spatial differentiation (in the plane of the retina or plate) of the stream of photons falling into the eye or a camera lens.

However, this equivalence of stimulation of particular visual cells is lost in the above-mentioned stimulus processing. It appears that the information coming form some visual cells becomes, *ex post*, more important (or has a different "meaning") than the information coming from other cells. Which of the receptor cells acquire greater "importance" at a given moment is not determined in advance. On the contrary, the importance of the stimulation of a given cell depends on the stimulation of the neighbouring cells (or its lack). It can change in time, depending again on the stimulation of neighbouring receptor cells. As it results from one of the below-quoted examples, the information concerning stimulation of a cell "situated" along the contour of some colourful spot becomes more distinguished or "privileged" than a signal from a cell "situated" within this spot (for the present, I understand here a colourful spot in a purely physical sense, as a beam of photons of the same wavelength falling on some area of the retina). This is related to the preference of the brain to extract **differences** in the stimulation of neighbouring cells (but also, in the case of movement perception, of differences in the stimulation of the same cell in neighbouring moments in time).

In the face of this, one can claim that the meaning of the data coming from a given visual cell depends on the "context" formed by the data coming from neighbouring cells. Therefore, it becomes clear that the discussed "evaluation" of information occurs not within a single receptor cell, but at the level of collecting and integration of stimuli from many receptor cells (the visual cortex can serve here as an elegant and relatively well-known example). Additionally, it seems obvious that the picture reaching the "deeper" (in the functional sense) layers of the brain cannot be equivalent to (identical with) the picture appearing in the plane of the receptor cells of the retina. While the picture on the retina corresponds to a photograph made with a camera, "faithfully" reflecting the external reality, further processing of the data can extract from it certain properties which are desirable for some reasons. For example, contrasts are stressed, objects in

---

[101]The reasoning and conclusions derived on the basis of the functioning of other senses would be similar. On the other hand, one should stress the fact that many or even all senses take part in the

the background are blurred, fragments of seen objects are completed with lacking elements. Some "patterns" of the picture on the retina are preferentially selected in the process of integration of sensory stimuli, while other patterns are treated as secondary ones. Such privileged integration strategies applied to sensory stimuli can comprise movement perception, perception of vertical or horizontal lines, higher sensitivity to certain ranges of the electromagnetic spectrum, and so on. These rather elementary mechanisms of sensory signal processing overlap with higher-order mechanisms, which exploit the accumulated memory records concerning past experiences, perform association with signals coming from other sensory organs, co-ordinate the function of effectors, and finally – are related to the autonomous activity of the brain, which corresponds roughly to the processes of thinking.

The influence of mechanisms integrating signals from receptors on the very nature of our thinking and cognition is, in my opinion, enormous and commonly underestimated. **The most fundamental categories of viewing the world come from the manner our brain processes stimuli from receptors**. (These categories are equivalent to some complexes of significative axes, differentiating the structure of our concepts). There are two basic aspects to be considered here. First of them is the relative spatial arrangement of receptor cells in sensory organs. The second is the integration of impulses coming from these cells. Our thinking in the categories of three-dimensional space, for example, is a direct derivative of the fact that two flat retinal pictures in our eyes, whose vision fields overlap almost completely, represent the same objects seen at slightly different angles (binocular seeing). Comparison of these pictures produces the depth effect, thus enabling our brain to "add" the third dimension to the two dimensions fixed by the plane of the retina. It can be supposed that animals (e.g. a cow) with vision fields that do not overlap at all see the world principally in two dimensions. On the other hand, we have not got the smallest idea what are the most elementary categories of seeing the world used by insects, possessing complex eyes, or by bats that obtain the picture of the surroundings mainly through echolocation. Most probably these categories escape completely our possibilities of comprehension. Therefore, our manner of seeing the world is certainly not the only possible one. It is based on a certain kind of structure of sensory organs as well as on mechanisms of signal integration – both fashioned by biological evolution. It is simply devoid of any sense to deal with ontology and epistemology, and also with philosophy of mind, perception and language, without

---

creation of many concepts.

knowing these mechanisms and the higher-level cerebral processes of association. Just as it had no sense to study the nature of matter and time before formation of modern physics. For the status of concepts in our psyche and the resultant cognitive limitations must constitute the ultimate starting point for any further reflection.

One of the fundaments of seeing reality **spatially** consists in joining the signals from (groups of) photosensitive cells stimulated by light in a similar way **into one object**, if these cells are situated **close to each other** in the plane of the retina and not scattered disorderly over its entire area. It means that a special informative value is attributed in this case to a stimulation signal coming from such a cell that is neighbouring with other stimulated cells. In spite of the fact that the primary picture on the retina is two-dimensional, its appropriate processing (e.g. taking into account of perspective or comparison of pictures coming from both eyes) carried out at a slightly higher level of integration makes our seeing of the world principally **three-dimensional**. We have good reasons to suppose that there exist organisms "perceiving" only two dimensions. Moreover, there is no principal reason that would forbid certain organisms to receive reality as a four-dimensional space-time or a six-dimensional phase-space. Therefore, our three-dimensional seeing of the world is not an obvious thing.

Movement perception – i.e. registration of changes in intensity of photosensitive cell (group of cells) stimulation rather than changes in the absolute value of this intensity – can be regarded as an important factor enabling us to perceive the **flow of time**. The importance of the information about stimulation of a given cell increases significantly if this cell was not stimulated before or, inversely, when signals from it suddenly cease to come after a period of persistent stimulation. Not only a receptor cell itself, but also some complex of neural cells is engaged in such "temporal memory". It seems that many primitive organisms receive first of all the current status of the surrounding environment (for example a presence of some substance in the sea water or soil) and react adequately, while they do not pay attention to sudden, but short-lasting changes of this status. Such organisms are handicapped, from our point of view, in their ability to perceive time. Other organisms, like the above-discussed frog, have firm predispositions towards seeing movement rather than picture. Therefore, it can be said that the world of this animal is in a sense rather temporal than spatial (the clarity of this example is a little bit disturbed by the fact that both temporal and spatial components of the "frog reality" seem to be much more reduced than their human counterparts; therefore, while speaking about temporally biased type of perception in a frog, we mean only a relative contribution of these components).

Another elementary mechanism of visual stimulus integration is the disposition to perceive **contours**, that means borders between two groups of photosensitive cells, when each group is stimulated in a different way. For obvious reasons, this enables us to distinguish **lines**. Therefore, preference for distinguishing contours (lines) is located completely on the side of stimulus integration mechanisms in our brain, and not on the side of the "objective" world, or even of the primary picture on the retina.

Cells in subsequent layers of the visual cortex, integrating signals coming from the retina, are arranged and connected in such a way that we have a strong inclination to perceive **straight lines**. (This is an important information for philosophers convinced about the existence of Platonic mathematical ideas). This is suggested by the studies on the visual cortex of a cat (different cells in the visual cortex are stimulated when straight lines with a different inclination are viewed by an eye). No wander therefore that the appearance of Euclidean geometry significantly preceded in time the appearance of its non-Euclidean counterparts. (I leave aside the fact that human mind **is** able to produce the concept of a straight line. Maybe some rational insects, with their complex eyes, would not be able to form such a concept, producing instead other concepts, incomprehensible to us. Thus, one can ask: are there two different worlds of Platonic ideas – one for us and another one for insects?).

Signal processing tends to separate **discrete objects** from a more or less continuous picture. This is the source of, on the one hand, our inclination to divide the world into sharply separated facts and categories, which finds expression at least in the discrete nature of language, while on the other hand – of our need to create the concept of a **number** which is indispensable to integrate further a picture composed of objects distinguished in such a way. (I see here clear implications for philosophy of language and mathematics). The roots of **causal relations**, or more broadly of **continuity in time**, can be seen as a result of joining **into one object** of a spot translocating in the vision field in such a way that it occupies **neighbouring locations in space** in **subsequent time moments** within the "picture" on the retina, and does not, for example, move chaotically over the entire field of vision. This is why a flying bird in subsequent time moments appears to us as one object, while this does not concern chaotic spots on a noisy TV screen with a spoiled aerial, or a group of birds at a given moment of time.

Therefore, there is probably no doubt (at least in my opinion) that distinguishing such categories as time, space, cause or a discrete object belongs completely to the domain of the manner of perceiving stimuli by our brain. This does not mean that nothing corresponds to these categories in the external world, but only that they are arbitrarily chosen (by evolution) approximations of some aspects of reality – primary and fundamental for our mind – that do not have to be "objectively" fundamental. Moreover, the kind of adherence of our world picture to the world itself will always remain to some extent a mystery, due to cognitive limitations imposed by the structure of the conceptual network. In this sense, the discussed categories lie rather on the side of our mind and not on the side of reality "in itself".

I would like to emphasise once again, that the mentioned manners of sensory stimulus processing are not concepts yet – for the reasons enumerated above – but some predispositions, semantic axes that determine the essence of originating concepts and directions of stratification of senses. Many of these mechanisms (and presumably all of them functioning at the lowest levels of integration) are inborn mechanisms, fashioned in the form of functional structures of the brain during the embryonic development, before any stimuli reach the brain (with a few exceptions). Concepts, on the other hand, are formed on the basis of the content of sensations "passed" through these integrating structures. Nevertheless, I realise that the above discrimination is partially a problem of definition choice, which is often directed by our own taste.

It seems to be understandable that many concepts are too abstract (this applies chiefly to secondary concepts) for their direct connection with the "sensory structure" of the substance of concepts to be easily tracked down. A greater role in the creation of these concepts is played by higher-order mechanisms of integration and association – the background of thinking as well as, more generally, of psyche – referring to memory and to the already possessed picture of the world. These two components, namely immediate processing of sensory stimuli and integration at a higher level, exhaust **completely** the neurophysiological background constitutive for concepts. The "sensory component" predominates in primary concepts, while the "associative component" – in secondary concepts. Beside them, however, there is (in my opinion) no place for any purely "spiritual" element, for anything at all (like e.g. qualia). **Concepts are (in neurophysiological terms) completely reducible to integration and association of signals occurring at different levels**. Their essence and sense come from a given concrete **manner** of integration imposed by evolution – first, biological evolution, and then, cultural evolution. On the other hand, it is not possible to explain completely the psychical form, "content" and **purpose** of concepts appearing in (self)consciousness by means of the conceptual apparatus of neurophysiology.

This situation seems to be analogous – to a large extent – to the emergence of biological purposes and senses from the physical level. Although functioning of living organisms is "nothing but" a certain manner of carrying out physicochemical processes, it is impossible to explain the essence of life and evolution referring exclusively to physical terminology. There is a need for additional conceptual apparatus of cybernetics and information processing, that does not follow in any necessary way from the physical description of the world.

All the above-presented examples of simple mechanisms integrating sensory stimuli seem to suggest unequivocally that the meaning of receptor cells is realised by connotation, as in the case of the meaning of neural cells. In particular, a given cell is not a visual cell (only) because it reacts to light, but first of all because the whole nervous system **treats** it as a visual cell. The nervous system is a big network, some endings of which – the ones that have "loose ends", not connected with other elements of the network (they are commonly called sensory receptors) – are stimulated in different ways, or rather they transform and transmit certain sets of external signals that reach them. In a sense, our psyche and consciousness can be reduced to a certain determined manner of processing these signals at different, chiefly the highest, levels of integration. Nevertheless, this is only the aspect of consciousness that can be seen from the biological level. It explains what is the **principle according to which** the neural network underlying psyche works. It is not able, however, to say anything about the **content** of psychical senses.

Having supplemented the connotative character of conceptual network with its, more or less directly, sensory sensation character, I would like to return now for a moment to philosophical consequences of the discussed concepts. All opinions and statements (including philosophical ones) can come into being in our consciousness only and exclusively as elements of conceptual network (their formulation within the syntax of language constitutes a secondary property, which is discussed above). Therefore, they cannot say anything that exceeds – either in **scope** or in **manner** – what is allowed by the **connotative** nature of conceptual network as well as by the sensory-associative origin of concepts.

Everything that was said above enhances additionally the thesis that the structure of conceptual network does not allow us to proclaim truths which are absolute in any understanding. As psychical senses can function only at the psychical level, they simply disappear once the functional structure of the neural network is destroyed. If there were no thinking beings in the Universe, there would not exist any psychical senses. Similarly, there is no basis allowing to speak about the purposes of biological evolution at the physical level, beyond the functional structure of living organisms. What is more important, the structure of conceptual network is completely different from the structure of the world it reflects (as the essence of a spider's web constitutes something completely different from the substance of the sculpture it covers), and their reciprocal adherence can be at best imperfect. This implies that psychical meanings – as they are a part of the conceptual network – are of completely different nature than anything that is **beyond** them, and therefore beyond the psyche (that is probably the source of the subjective inclination to separate "spirit" from matter). The external reality (in relation to the conceptual network) can be known only approximately for two principal reasons. Firstly, we cannot reflect the world with **infinite** accuracy, which is due at least to limited capacity of our mind. Secondly, the very "substance" of concepts will always remain different from the "substance" of the world. As the previous analogy shows, it is not possible to tighten infinitely the mesh of the spider's web in order to reflect the smallest details on the surface of the sculpture. Nor will the spider's web become **identical** with the surface of the sculpture. Without this identity, conceptual network can only pretend to be a certain **model** of reality, reflecting (approximately), as any model, only some aspects of what is modelled. For all the above reasons, a formulation of any statements that are to pronounce about absolute truths – fully autonomous and non-relativised to anything beyond themselves – is from the start flawed with a logical error. Nevertheless, a large part of philosophy keeps committing this error.

The problem of mutual relation between spirit ((self)consciousness) and matter (body) (the mind-body problem) can serve as an example of a problem artificially generated by philosophy. Interactionism treats them as two independent beings that exert influence one on another. However, certain arguments were formulated to show the impossibility of a bi-directional causal relation between these beings (i.e. both body → mind and mind → body), due to their total dissimilarity. The nature of evident examples showing the existence of a mutual influence would remain a mystery in the face of this, and the "dualism" of spirit and matter would lead to a contradiction. Parallelism resolves this question by postulating parallel occurrence of material and mental phenomena, while the theory of two aspects regards them as two different manifestations of some third being. However, the first of these concepts does not explain anything, while Occham's razor would be of use for the advocates of the second concept.

The point is that the whole problem was erroneously formulated from the beginning. We forget that philosophy led to opposing so drastically the mentioned beings, and give them an unquestionably absolute sense, which in fact should be recognised as equivalent to production of hyposthases. Consciousness (and matter as well) does not exist in the absolute way attributed to it by philosophers. It can be reduced, in a sense, to a **manifestation** of brain functioning, just as the roar of water is a manifestation of its falling in a waterfall. Causal relation – if this term makes any sense here – is in this case unidirectional and leads from brain to consciousness. Consciousness at best **accompanies** the neurophysiological brain activity and it is its "by-product". Similarly, while the statement that a waterfall constitutes the cause of a roaring sound seems to be relatively reasonable, then, the influence of the roar on the waterfall appears to be a complete nonsense. The opinion presented here has much in common with the philosophical conception called epiphenomenalism. This is suggested by the fact that here, to explain the **grounds** of the phenomenon of (self)consciousness, it suffices in principle to describe the manner in which impulses are conducted in neural circuits and the general cybernetic mechanisms. A strict analogy holds here with the phenomenon of life. Spirit in its immanence constitutes something analogous to the "vital force" of living organisms, which has been recognised by biology for a long time as an empty hyposthasis. While I am inclined to agree that psychical senses cannot be reduced to biological senses, then I do not see any need to distinguish (self)consciousness (mental sphere) as some absolute being.

The conception of the mental sphere presented here goes undoubtedly further than epiphenomenalism, because it shows **in what way** psyche and (self)consciousness can be a derivative of some neurophysiological mechanisms functioning in our brain. Psyche in such a sense both is and is not reducible to a complex of neurophysiological processes. In the same sense, the phenomenon of life both is and is not reducible to a complex of physical processes.

Each process in a living organism is a physical process. At the same time, a certain **complex** of such processes characteristic for living organisms escapes physical terminology. Here appear purposes and senses that cannot be attributed to inanimate matter. The question if these purposes and senses – and therefore the essence of life – constitute an "objective" element of the world or a "subjective" category of our mind is devoid of any sense. This question must remain without answer. Anyway, the border between objectivism and subjectivism is becoming more and more obliterated even in the most strict natural science, namely physics. For example, such "objective" concepts of thermodynamics as information, macroscopically distinguished state or the arrow of time become "subjective" from the point of view of "classical" dynamics. Even considering the elements of the physical world described by dynamics (understood broadly, together with quantum mechanics and relativity theory) as orbital, matter or force of gravitation, one can doubt if they should be attributed with "completely" real existence, or perhaps they should be considered as a convenient manner of categorisation by our mind of the whole accessible spectrum of phenomena. Indeterminism of quantum mechanics resulting from Heisenberg's indeterminacy principle is probably the best known example of subjectivism in physics.

The problem of the relation between the psychical level and the biological level is analogous to the dependence between the biological level and the physical level. **Each** psychical process – including the processes underlying (self)consciousness – is a biological (neurophysiological) process. However, a certain **complex** of neurophysiological processes, that can be reduced to a determined pattern of circulation of impulses in the neural network, acquires **new senses** – absent at the biological level. The scheme of connections in the neural network together with the mechanisms of its functioning determine the connotative, and therefore sensory-associative, essence of the conceptual network, i.e. the "content" of psyche. Of course, not every imaginable combination of neural impulses in the brain would underlie (self)consciousness, just as not every set of physical processes would lead to the formation of a living system. Therefore, the fundamental problem becomes **how** (self)consciousness does emerge from the biological level, and **what** kind of structure and function of neural circuits is responsible for the emergence. According to my conception presented here, the general answer for this question is the same in the case of the roots of (self)consciousness as in the case of the roots of life, namely the **relation of self-applicability**.

Generally speaking, a new kind of senses (purposes) appears (can appear) when some system starts to be **directed on itself**. In the case of biological individuals, this self-focusing is manifested as a tendency of a given identity to survive and propagate itself. In principle, the very essence of this identity consists in a drive towards auto-copying. Therefore, it becomes an autonomous sense or purpose ("in itself") to maintain the value of this "parameter", that means identity, at the "assigned" constant level. Let us remember that the mechanism ensuring some parameter value to be maintained at a constant level is called negative feedback. After all, it is not accidental that systems attributed with negative feedbacks are given in cybernetics the name: "tendential systems". In the case of life, the relation of self-applicability consists in the fact that the adequate "parameter" (which is the identity of a given individual) tends to maintain a constant value of itself.

On the other hand, we can speak about (self)consciousness from the moment when the (broadly understood) dynamic picture of the world within the brain – the picture that is to serve as a frame of reference, segregator and processing centre for perceived stimuli from the world – begins to "perceive" itself as well. At the same time, I treat the concept "picture of the world" in a purely instrumental (non-psychical) sense, that is as a certain adequately processed and integrated complex of experienced sensations, serving as a frame of reference and interpreter for future sensations (I do not attribute to this concept the status of any categorically distinct being). In this way, the dynamic picture of the world becomes "aware of its own existence". In my opinion, this resulted in the appearance of psychical senses, completely alien to the biological and physical level. Probably, I do not have to convince anybody about the existence of these senses. And this is in fact everything what can generally be, in my opinion, said about the essence of "spirit" or (self)consciousness.

Therefore, the problem of the relation between spirit and matter (the mind-body problem) resulted – according to what I said above – only and exclusively from a confusion of concepts, i.e. from granting them some optional and "absolutistic" meanings without taking into account what modern science already can (or will be able to in near future) say on this subject. A completely autonomous consciousness is a superfluous being in the sense of Occham's razor. However, this leads to another problem. Lack of causal influence of consciousness (spirit) on the body triggers automatically the problem of the existence of (conscious) free will. Even after a superficial analysis, however, we come to the conclusion that free will in the philosophical sense does not exist. More, it is an internally contradictory concept, to which nothing sensible can correspond. Unfortunately, production of artificial problems constitutes one of the main sins of philosophy – even more, it is an important element of the essence of philosophy in general. I think that conventional and apparent character of the majority of philosophical problems – manifested in production of hyposthases, in disputes concerning terminological ranges rather than senses corresponding to these terms, or in a purely optional transfer of concepts from their natural surroundings to other semantic environments where these concepts, still hidden under the same linguistic name, become in fact different concepts – allow me to take such position. The quite well known question if mathematics is created or discovered can serve as an example of the last phenomenon. I maintain that this problem vanishes the moment when we realise that the creation and discovery mean something different in mathematics than the creation of artistic masterpieces or discovery of new lands.

The paradigm of conceptual network can also show in a new light such problems as the problem of the existence of **universal beings**. Philosophers either granted these beings with the status of objects which are equally real as individual beings, or refused them any real existence, or viewed them as some conventions which allow us to describe reality in a convenient way. The problem, however, disappears in a completely natural and automatic way at the level of the conceptual network. Concepts corresponding to universal beings do not differ principally from other concepts and originate in a similar way. Of course, they correspond rather to secondary concepts than to primary ones, and therefore their essence is determined to a large extent by the manner of processing and association in the brain of signals on a high level of integration. Like all other concepts, however, they mean by connotation, and man acquires them gradually, during his individual development, by assimilating inductively repetitive complexes of sensory impressions and concepts. It seems to me futile and devoid of any sense to speak about a manner of "existence" of universal beings different from the existence of concepts corresponding to these beings. Anyway, it is not possible to demarcate strictly universal concepts and individual concepts. In the ultimate instance, the hierarchy of the universality of concepts gets completely flattened and all concepts become simply elements of the conceptual network invested with equal rights.

* * *

While speaking about the essence of consciousness, it is difficult not to mention, at least superficially, the recently fashionable problem of artificial intelligence (AI). This sphere – as each field of philosophy lacking firm grounding in facts – is full of insinuations and extreme simplifications. The principal problem consists here in the question whether an artificial electronic system processing information, a computer or robot, can possess psyche and (self)consciousness.

Searle gives a negative answer to this question. He refers to a thought experiment known under the name of the "Chinese room". The so-called Turing's test constitutes here the frame of reference. It consists in creating a situation where a man communicates through a system of questions and answers with something that he does not see. The man is to decide if his interlocutor is another man or a computer. Communication can occur, for example, through sheets of paper transferred from one room to another. If our man, while communicating with a computer, cannot unequivocally decide if he deals with another man or not, then this would testify that this computer possesses (self)consciousness.

Instead of a computer, Searle's "Chinese room" hosts a group of Chinese who do not know English and receive cards with questions in English. These men would supply answers to the questions by operating according to instructions in Chinese, that would determine a complex of purely mechanical rules of how to transform English questions into English answers (how to "translate" a sequence of words in a question into a sequence of words in the answer). If such a complex of rules were sufficiently rich, the group of Chinese would pass Turing's test (they would be recognised as a conscious system), although the men in the room would understand neither the received questions nor the given answers. This is to testify that a computer, operating after all by mechanical transformation of data according to supplied rules, cannot posses (self)consciousness.

The analogy with the "Chinese room" is undoubtedly misleading, because the introduction of a group of Chinese men – attributed after all with psyche (although it does not understand English) – instead of a complex of processors in a computer creates only an unnecessary complication that blurs the picture, but does not offer anything really new. Both the group of Chinese people that do not know English and microprocessors in a computer or neural cells in human brain are similar complexes of elements that transform information according to rules defined in advance and therefore do not understand "single-handedly" what they are doing. Both the computer and the brain function **only** and **exclusively** on the basis of the operation of such elements (unless we believe in spirit which differentiates man from a computer, but then any discussion terminates). Brain "copes" somehow with production of (self)consciousness on the basis of these elements, so why a computer could not be able to do the same?

The heart of the problem lies (in my opinion) in the degree, but first of all in the **kind** of complication of a system of functionally interconnected elements that is necessary to produce (self)consciousness. Searle has only shown that some kind of simple systems operating according to formal rules cannot possess psyche, but this conclusion seems to be so obvious, that it is trivial in fact.

Therefore, what distinguishes this kind of complication that underlies (self)consciousness in our brain, but is absent in present computers? Of course, it is not easy to give an exhaustive answer to this question at the present stage of knowledge. Probably, we even do not have a presentiment of a majority of conditions which would have to be fulfilled. Nevertheless, I would like to try to outline at least some aspects of the essence of intelligence (artificial included) on the basis of what I have already said on the subject of (self)consciousness.

I claimed that the property of self-applicability, i.e. auto-directing of the cognitive centre in human brain on itself, is the essence of the (self)consciousness. Therefore, this property should also apply to a "conscious" computer: some of its parts, transforming information obtained from the input and producing some kind of a picture (representation) of the external world on the basis of this information, would also have to collect signals coming from itself. Before such a picture can be directed in itself, however, it must be produced first. Therefore, a conscious computer should possess some set of receptors, receiving signals from the external world, "centres" that integrate these signals, processors of memory enabling accumulation of data, as well as effectors, i.e. executive devices that allow it to exploit the possessed picture of the world for some purposes. It would also be desirable for such computer to be equipped with the possibility of learning by trial and error, allowing it to test the adequacy of assimilated "behavioural" rules to the realities of the surroundings. Its memory should be of connotative, relational character and not of linear one (as in the case of the record on a computer disc or magnetic tape). This applies also to the structure of connections between the processors of such computer. A parallel and not sequential processing of information should be implemented. Therefore, the principle of operation of such computer should somewhat resemble the mechanism of functioning of neural-like networks. Finally, it would be also necessary to face such an electronic brain with the task of its own survival in complex conditions of the surroundings, in order to create in it an equivalent of the "self", preferably connected to the centre of reward (pleasure), on one hand, and to the centre of punishment (pain), on the other hand.

The presently existing computers "think" linguistically, in a discrete manner. Their functioning consists in operating with sequences of discrete symbols. On the other hand, the operation of the brain occurs principally at the conceptual layer, which is deeper in this case than the linguistic layer. Unlike linguistic names (symbols), concepts are continuous in the sense in which spots on a photograph are. If single grains on a photograph are so tiny that the resolution of sight is too small to detect them, we have an impression of continuity of a picture. In this analogy, neurones correspond to photographic grains. Although neurones themselves are discrete, we can treat the neural network as a continuous object at the macroscopic level composed of billions of neural cells. Concepts in our consciousness function at this macroscopic level (and not at the level of a single neurone), while linguistic names are found at a yet higher level, as they are secondary objects in relation to concepts. Therefore, linguistic names are not fundamental elements here, unlike in the present computers. To the contrary, they are complex objects serving as labels for underlying concepts that mean by way of connotation. Therefore, a "thinking" computer should operate first of all at the level of a continuous, connotative conceptual network and not at that of linear, discrete "sentences" of machine language. For language itself is only a system of symbols ordered due to some grammatical syntax. Without the conceptual lining, it is devoid of any immanent meaning (one that is not imposed from outside). No wander therefore, that the simple and discrete system of formal rules used by the Chinese in Searle's example did understand nothing.

Formulating the matter in yet another way: one basic element of structure is responsible for one elementary function in the (present) computer, while in the case of conceptual network, its most elementary operations occurs at a significantly higher level of integration than the function of a single neurone, and therefore this operation can be a "continuous" (analogical) process. Apart from this, we build a computer according to a plan defined in advance, attributing to it some functions in a denotative way. This device realises the imposed task of transforming some stimuli at the input (e.g. keyboard) into certain signals at the output (e.g. monitor or printer). As such, it remains necessarily non-psychical. The conceptual network, on the other hand, develops due to its own purposes – initially conditioned in a purely biological way – and without any detailed external plan. Its evolution consists in building up new meanings on the basis of already existing meanings, which is why it is doomed to be connotative.

Summarising very shortly, I think that the present computer cannot reach the level of (self)consciousness first of all due to the absence of: 1. A cognitive centre directed on itself, and 2. A connotative layer of concepts underlying the layer of discrete linguistic symbols.

*   *   *

For the purposes of this book, I divided reality into three main layers or levels: physical level, biological level and psychical-cultural level. However, I am decided to avoid any absolutisation of this division. Its task is to attempt at a heuristically fruitful description of the world, and not a multiplication of unnecessary autonomous beings, whose interrelations would remain completely mysterious (as in the case of the relation between spirit and matter). The introduced categorisation – as any of this kind – must necessarily be arbitrary to some extent and it must lack sharpness, all the more so, because I myself put particular emphasis on showing how particular levels emerge one from another. Despite these reservations, I would like to believe that I have a relatively objective criterion of distinguishing these levels at my disposal, allowing me to avoid complete arbitrariness and production of futile hyposthases. This criterion is constituted by the kind of senses (purposes) functioning at particular levels. For each of the enumerated layers is characterised by a specific set of senses which can be defined in a relatively strict manner, by means of cybernetic terminology, for example.

If we determine sense in cybernetic categories, the physical level (both the whole Universe and each of its elements) has no sense. At this level, there are no negative feedbacks, which we earlier recognised as an inseparable element of purposeful behaviour. However, the very fact of the existence of the Universe can be treated as the "sense" of the physical level. As I mentioned before, the Universe can be regarded as some form of existence of nonentity if we choose an appropriate point of reference. It seems that all components of our Universe, such as matter, energy and space, when added altogether, give exactly zero as the result. The positive energy equivalent to mass is equal to the negative potential energy related to gravitation. The negative energy of gravitation comes from the mutual distancing of different masses. On the other hand, a distancing is nothing else but the appearance of space between bodies attributed with mass. Therefore, in accordance with the principle of preservation of energy, the zero energy of nonentity could undergo splitting only into some amount of positive energy and an equivalent amount of negative energy. The origin of space constitutes therefore the other face of the origin of matter and inversely – neither space without matter can exist, nor matter without space can. However, if the Universe actually constitutes globally some form of nonentity, one should admit that it is a very complex one. In particular, the Universe contains an enormous amount of information about **the**

way of splitting nonentity. This determines in fact the shape of everything that exists: elementary particles, atoms, stars, galaxies, chemical compounds and living organisms. And this information can be identified with the sense of the physical layer of our world. This is undoubtedly very little from the human point of view. A sense manifested only in naked existence, lacking any purpose or "direction" . At the same time, we perform some semantic misuse even when we speak about "existence", because we suggest that the Universe lasts in time. But our Universe, if considered as four-dimensional space-time, does not last in time at all (time constitutes its part or aspect). It simply **is**. Anyway, in some region of space-time differentiated arbitrarily by our cognitive apparatus as the "beginning of the Universe", time is simply (if we agree with Hawking) yet another spatial dimension which acquires different properties and "emerges" as a time dimension – in the common understanding – only at some distance from this beginning. The Universe simply is and this constitutes its sense. At the same time, the question if it had to originate or if it might equally well not exist seems completely nonsensical. Therefore, Cosmos can in no way be regarded as an intentional object, that possesses any justifications, purposes or senses, apart from those mentioned above. The "true" purposes and senses appear only at higher levels of reality – the biological level and psychical-cultural level – and they cannot act "backward", on the physical layer. For the kind of senses existing within each of them is the only quite objective criterion for distinguishing the mentioned levels of reality.

In cybernetic categories, the sense of the biological layer of reality (life) can be determined as a network of negative feedbacks maintaining a superior positive feedback, namely the potential of expansion. Several elements are necessary for this sense to exist or find "expression". First, a biological "sensible" system should possess its own identity, i.e. it should differ from similar systems and, all the more so, from other "objects" in its surroundings. This is one of the reasons why we cannot treat simple self-reproducing (catalysing their own synthesis) chemical molecules as alive. They are all identical. As the second condition indispensable for the existence of (the sense of) life, one can point out the presence of mechanisms that allow this identity to preserve its continuity in time, by producing progeny that is (almost) identical with parental individuals (significant contribution of an individual's identity to the formation of the identity of progeny is at stake in the case of organisms reproducing sexually). This excludes polycatalons from the circle of living beings. Although they possess identity created by chance, they are not able to transfer this identity to descendent protein chains. As I mentioned in the second part of this book, the mechanisms leading to the preservation of identity (such as catalysis of reciprocal synthesis carried out by proteins and nucleic acids in hypercycles, which allows the sequence of aminoacids in proteins and the sequence of nucleotides in nucleic acids to be preserved in time and transferred to "progeny") are in principle identical with the principal negative feedback, responsible for maintaining a constant value of some parameter in time, the parameter being the identity itself. The network of negative feedbacks – the next property necessary for something to be classified as a living system – is something that we intuitively associate with what can be called purposeful or sensible action. In this sense, a computer, robot or even a thermostat in a refrigerator are purposeful. However, the sense of the enumerated systems is not autonomous, since they do not realise any purposes of their own. The purposefulness of these devices is determined only and exclusively by their usefulness for man. They do not fulfil the last criterion linked inseparably with the essence of a living system. They lack the positive feedback, expressing itself as the potential of expansion. On the other hand, while this last property can be found in prions and stones in an avalanche, both are disqualified as candidates for living beings by the lack of any mechanisms that would maintain their "identity", and this identity itself adopts in them only a rudimentary form, expressing itself only in movement of stones and in infective conformation of a protein molecule in the case of prions.

A system possessing identity and at the same time attributed with appropriate complex of positive and negative feedbacks serving its maintenance becomes automatically an individual, able to evolve (evoluon). The identity of an evoluon, because of the physical nature of its "carrier" (after all, each biological organism is also a physical object) can (must) undergo tiny changes (mutations) as a result of accidental events. This new identity, similarly as parental identity, will tend towards its own expansion and, if it turns out to be "better" (which will be decided by broadly understood environment), it can win the naturally-arising competition for propagation. This means replacement of individuals possessing the old identity with individuals attributed with the new identity. And this is just evolution after all.

The definition of life (a living individual) as a system of negative feedbacks subordinated to a positive feedback allows for two things. Firstly, it separates relatively clearly and objectively the phenomenon of life from inanimate world. Secondly, it joins together in a natural way the concepts of life, an individual and evolution, showing that they are different manifestations, aspects or faces of the same thing. In this formulation, life, an individual and evolution seem to be connected inseparably and to imply each other.

However, the most important role of this definition – in my opinion – consists in the fact that it illustrates directly the "sense" and "purpose" of the existence of living organisms. For we intuitively associate these organisms – as in the case of devices created by men – with a purposeful behaviour. While the devices serve our purposes, however, then whom do living beings serve? They serve nobody, or rather themselves, i.e. the expansion of their own identity. And nothing more. From our anthropocentric point of view, an irrational expansion seems to be a decidedly repulsive as the only purpose of the greatly complicated machinery of life. But this is not everything. It seems to be ironic that even the organisms existing millions and billions years ago which achieved the evolutionary success (which is reducible to the fact that their descendants inhabit the terrestrial globe at present) lost their identity in the process of evolution, for the presently living species, including *Homo sapiens*, differ very much from their far ancestors. Therefore, even the criterion of identity preservation seems to be too exacting. Its weakened version retains only continuity of changes of this identity in time. (This resembles the problem of the identity of human psyche. Although it changes during our whole life from birth to death, the **continuity** of these changes proves that we deal with the same person in the case of a given man. For analogous reasons, a man and a robot with the man's psyche copied into it would not be the same person, despite the fact that the "psyche" of such a robot would differ less from the current psyche of the original than the latter from the psyche of the same man ten years before). One cannot of course speak about complete extermination of an ancestor's identity, which would boil down to breaking the connection between this identity and the identities of its descendants. The course of evolution, that had led to the appearance of this ancestor, limited or "channelled" very much the range of choices or evolutionary routes available to the individual. This is equivalent to establishing some range of possible forms which can be adopted by descendants. On the other hand, this range is still very broad, and the similarity of a descendent to its ancestor is frequently imperceptible macroscopically (a similarity can be found a little bit easier at the biochemical and genetic level). It is hard to imagine maternal affection towards us in the eyes of some coelacanth fish from the Palaeozoic period, or to detect any traces of similarity with an unicellular amoeba – our still further ancestor. Considered through the prism of cultural senses, the sense of the phenomenon of life is very ephemeral, if not completely nonsensical at all.

Life looks similarly when viewed from the physical level. Here, biological organisms are simply complexes of physicochemical processes, that can be reduced to mutual interactions of atoms and molecules. There are no important differences between those interactions and the phenomena occurring in rocks, water, atmosphere or the interior of the Sun. The building materials – proteins, nucleic acids and other organic substances – certainly do not decide about the essence of life either. For, on the one hand, they can exist beyond living organisms and, on the other hand, one can imagine living beings built of somewhat different compounds[102]. Neither is the complex structure of living organisms more purposeful than other, simpler dissipative structures (including the already mentioned convective currents) that "serve" only to dissipate energy effectively, as gravitation "serves" to attract bodies attributed with mass. There is no teleology in this. Biological organisms cannot be in any way distinguished or defined in physical terms. To do so, one also needs at least cybernetic terms. Anyway, physics itself is not coherent and self-sufficient in this respect. The arrow of time appearing in thermodynamics cannot be derived within the classical dynamics (e.g. quantum mechanics or relativity theory). Neither is it possible to formulate the ideas of life, individual or biological evolution in the language of physics. Therefore, the phenomenon of life can be attributed with a "sense" only and exclusively in the context of cybernetic mechanisms functioning at the biological level and its essence is contained only in these mechanisms.

Psycho-cultural senses are of different nature than biological senses. They share one property, however. Just as biological senses cannot be expressed in the language of physics, the essence of psychic-cultural senses escapes biological terminology completely.

I will begin with psychical senses. While they are not identical with biological senses, the latter undoubtedly constituted the basis for the origination of psychical senses. For brain – the "carrier" of psyche – was formed in the process of evolution as an organ fulfilling important biological functions. It ensured – in co-operation with peripheral nervous system and the whole set of sensory organs and effectors – adequate reactions to environmental changes. However, this function is not realised exclusively by brain. It is also realised (in a much less complex manner)

[102]On the other hand, it is probable that those compounds would have to be organic compounds (containing carbon-carbon and carbon-hydrogen bonds), the only ones offerring a sufficiently rich network of connections to become a germ of a more complex structure and function. Nevertheless, this applies only to life originated in a spontaneous way. If we are ever successfull in constructing

even by the simplest, decentralised neural networks (where neural cells do not form complexes), such as the neural system of a hydra. Moreover, biochemical-genetic "networks" are also able to react to changes in their environment (and therefore to behave in a purposeful manner). For example, the appearance of a new sugar (e.g. lactose) activates in bacteria – if other carbon sources are lacking – the genes responsible for production of enzymes that catalyse the pathway of decomposition of this sugar (the already-mentioned theory of operon). The operation of the brain is obviously much more specialised and differentiated. The richness of the world picture formed within it exceeds greatly that of the "world picture" of a biochemical-genetic network, which can be reduced to adequate reactions to some aspects of reality, that are important for survival. Nevertheless, this does not change the fact that – from the cybernetic point of view – the brain and the bacterium are not separated by any clear categorial barrier. Both the brain (neural network) and the biochemical network (more strictly: biochemical-genetic network) in bacteria are certain complexes of regulatory mechanisms (including negative networks) and information transfer pathways, that realise in co-operation with other sub-units the superior purpose of the entire individual – expansion. The picture of the world in the brain – equipped with such perfect devices for perception and preliminary processing of sensory stimuli as the eye or the ear – can be incredibly complex and it is capable of taking into account a huge number of aspects of this world. But there is still no place for (self)consciousness, and therefore for new categories of senses. Improvement and complexity of the cognitive apparatus is not sufficient. Something else is needed.

I think that this "something" can, in principle, be reduced to focusing of the cognitive apparatus on itself. It means that not only a picture of the external world is formed within this apparatus, but also a picture of the apparatus itself. In other words, the picture of the world is projected into a certain part or area of this picture itself. This very simple cybernetic mechanism can be expressed in the language of logic as the relation of self-applicability. As a certain – although rather far-fetched – analogy can be found in a three-dimensional model of a city located on a stone plate in the market square of this city. Apart from the delineation of streets, location of squares and particular buildings, the model will also contain a tiny model of our model of the city.

---

robots that would be able to copy themselves and evolve, to realise their own purpose, i.e. uncontrollable expansion, then they will become undoubtedly living organisms.

If such a model were to be perfectly faithful, this sequence of models within models would be repeated *ad infinitum*.

As we know, brain is built of nervous cells, able to conduct electrical impulses. It also applies the functional "centre" in the brain responsible for higher cognitive functions. This centre receives (preliminarily processed) stimuli from sensory receptors and confronts them with different areas of memory, that is with the already possessed picture of the world. This can of course lead to further development of this picture. Now, if certain neural cells within the centre connect – for some reasons – their input appendages (dendrites) with output appendages (axons) of neurones belonging to the same centre (maybe even with their own axons), instead of connecting exclusively with axons of cells transmitting stimuli from receptors, then, something strange will happen. The picture of the cognitive centre will be represented in this centre itself, and in this sense this picture will become **conscious** of its own existence, its own identity (or it will simply become self-conscious). This projection will of course proceed continuously, with a frequency corresponding to the period of the circulation of a neural signal in a closed loop (this time of circulation would correspond to the shortest, "atomic" time span perceived by consciousness). The result will consist in "self-copying" of the identity of this picture. Therefore, I am inclined to identify this continuous projection of the picture of the world into itself with (self)consciousness.

Did (self)consciousness originate by chance? In a sense, certainly it did so – as (a) mutation(s) in regulatory genes that control the processes of brain development. However, it cannot be excluded that, once it originated, it turned out to be profitable for the organism of the carrier. Therefore, the development of (self)consciousness was prized for purely biological reasons, as ensuring evolutionary success to its "carrier" (I think that the emergence of consciousness was a gradual process and that its germs can be found in chimpanzees of dolphins). This, however, does not change the fact that the appearance of (self)consciousness was accompanied by the origination of a new category of senses, qualitatively different from biological senses. A biological individual became a psychical individual as well, and their purposes did not need to overlap since then. The psychical identity is not identical with biological identity. As each tends first of all to realise its own interests (preservation and propagation of itself), there can arise conflicts between them.

The simple mechanism of origination of (self)consciousness proposed by me is not a pure speculation, similar to those which I regard as a sin of traditional philosophy, for this hypothesis is in principle falsifiable. It implies that at least some neural circuits receive signals also from themselves in the centre in human brain that is responsible for the highest-level of association and integration of signals coming from sensory organs and from various areas of memory (the center of self-consciousness). This fact can be experimentally rejected or confirmed (indeed, there exists anatomical evidence that there can exist closed circuits between the brain cortex and thalamus or within the cortex itself). This may be a problem of distant future. For the present, one must be puzzled by the fact that the same relation of being directed on itself (self-applicability) did lead to the origination of categorically distinct senses at least once. After all, the self-reproduction of living organisms is nothing else but the already-mentioned relation. An identity, self-projecting and ensuring its own continuity in time – does not it come annoyingly close to the sense of self-consciousness?

The moment (self)consciousness comes into being, all the elements of the world picture formed within the brain acquire (in my opinion) a psychical sense, apart from the hitherto exclusively biological sense. From this moment on, the picture not only exists, but also exists in relation to something. (Self)consciousness – assuring the preservation of its own identity in time – is exactly the intentional factor that "evaluates" particular elements of the picture. The picture of the world becomes something more than a complex of purely instrumental meanings, allowing an individual to function in a possibly efficient manner in the surrounding reality. The picture acquires **sense**. This sense does not concern man as a biological being, whose only task is to survive and produce progeny, but as a psychical subject, not reducible simply to the system of "purposes" of biological evolution. This is corroborated by the fact that very early in their history primitive societies created religion, magic and art, that certainly meet psychical and not (only) biological needs of man.

As man is a social creature and a psychic individual, he also becomes a cultural individual. According to the conception I proposed earlier, a cultural individual is equivalent, by analogy to a biological individual, to a "self-copying (self-projecting)" identity that tends to expand and preserve the continuity of its own existence. Here, however, unlike at the biological level, it is not the biological identity, but the psychical identity that is in question. The role of man as a member of society and culture is based first of all on the fact that he preserves his identity. People in society are not mutually replaceable (like ant workers in an ant-hill) not because of genetic (and therefore biological) differences between them. After all, it is important in society that just a particular man and not his identical one-egg twin is the husband of a particular woman (this would not make any difference from the biological point of view). A psychical identity should therefore be regarded as something certainly different from the biological-evolutionary identity.

In the face of such a distinguished role of a psychical identity, its preservation in time and possible extension of its operation and impact range – i.e. dissemination of its elements among the members of a given society – becomes a thing of primary significance. This expansion can adopt various forms. In the simplest case, it is realised by contact with other members of society, education of children, pronouncement of one's own opinions and so on. Scientific, literary and artistic activities are its more sophisticated manifestations. Finally, in extreme cases, it expresses itself as a pursuit of fame or power. The projection of identities of particular members of a society onto this society as a whole leads to establishing a certain system of cultural senses.

As I argued above, cultural senses are in fact born from chance. However, the selection of senses to become consolidated is decided to a large extent by the impact of the expanding psyches that "carry" certain senses in relation to the impact of the expansion of psyches promoting "competitive" senses. This is something resembling a natural selection of cultural senses (mems in Dawkins' terminology), and therefore a process analogous to biological evolution, although occurring at a different level. Now, I will try to analyse this analogy in greater detail.

There is probably no doubt that particular psyches differ from each other. These differences can be reduced to different world pictures and complexes of cultural senses. Inborn genetic predispositions play rather insignificant role in an "acquiring" cultural senses. The possibility that an African born and educated in a Christian society will adopt the belief in Christ acquires a much greater probability than the alternative that he will come back to the religion of his ancestors, particularly if nobody informs him about the "existence" of African idols. Therefore, it seems obvious that a huge majority of cultural senses is inherited through the mediation of cultural transfer, and not via the biological channel. Due to differentiation of these senses, however, their transfer from generation to generation does not seem at all to be a determined matter. A great role in this process is undoubtedly played by parents, school and so on. However, chiefly in modern civilisation of the West, man is exposed to great diversity of propositions of cultural senses to choose from. Such a possibility existed in all cultures since the origin of civilisation, although its range was presumably much more limited.

For the present, I limited myself to senses taken over by a given man from other people. However, a men is – at least potentially – an effective generator of new senses. What in most members of society occurs at a small scale and concerns secondary things, decides about the exceptionality of the so-called outstanding persons: scientists, artists, men of letters or prophets of new religions.

All this testifies that cultural senses exhibit large diversity, can be "inherited", and new senses can be generated on the basis of old ones. For the present, one can find here an analogy with genetic information of particular individuals in a biological population, which also exhibits a differentiation, is inherited and undergoes mutations leading to origination of new variants of this information.

Only very few of the senses originating within culture are consolidated into it for longer periods of time. On the other hand, as culture develops, there undoubtedly takes place an evolution of the senses which constitute this culture. Such changes can proceed gradually or they can be characterised by the appearance of great turning-points, as it took place in the case of the Copernican revolution. All this leads to the conclusion that some senses are consolidated in some way, while other senses are sooner or later eliminated. The principal question here concerns the mechanisms or criteria decisive for the selection of these and not other senses at a particular stage of cultural evolution.

One should remember that, a psychical individual constitutes a "seat" of cultural senses, just as a biological individual is a "carrier" of genes in a population. The latter tends to propagate its genes in a population, while the former strives to transfer the senses recognised by him to other "participants" in culture, particularly to his own progeny (education is an element of this process). In this case, the efficiency in spreading the senses accepted by a given man (understood as an individual in a psychic sense) measured against the efficiency of dissemination of the senses proposed by other people is decisive for long-term success. This depends not only on sheer attractiveness of these senses for their potential receivers, but also on the accessibility of mechanisms enhancing the dissemination of senses in a purely instrumental way. While speaking about such mechanisms, I am referring to the naturally privileged position not only of parents in relation to children in this respect, but also of people having access to mass-media, politicians, the Church, and finally men of letters , artists and scientists in relation to common people.

Bearing in mind the above conditioning factors, we must address the following questions: Why do some senses consolidate and undergo propagation, while other senses are eliminated? Why is the evolution of senses in various cultures divergent (when these cultures develop in at least partial isolation)? What is the reason for establishing slightly different senses in different sub-cultures of a given culture? What is the reason for the transformation of senses in subsequent periods of development of a given culture? A great role is undoubtedly played here by a chance – as in the process of biological evolution – expressed in the generation of directionless variabilities. These, however, fall within the range determined by the hitherto history of evolution. This, however, refers mainly to production of new senses and not to selection of already existing sense. So?

By analogy to the natural selection, the mechanism in question can be found in something that can be described as the cultural selection, imposed by the "cultural environment" existing at a given moment. In a sense, the environment for cultural senses is constituted by the set of all human psyches. The direction of development of this culture as well as particular selection of senses to be assimilated (taken over from other cultures) or established (within the culture) is determined to a large extent by the already existing cultural pattern. Of course, there are limitations for potential directions of this development. They are imposed by economical, biological and physical realities of the world, among other factors. They also form the cultural environment to some extent. However, the selection of newly generated senses carried out through the sieve of already existing senses seems to be more interesting for us at this point. If "carriers" of the new senses possess greater potential for expansion than cultural individuals representing the old system of senses – either due to pure attractiveness of these senses or to the above-mentioned instrumental mechanisms – then, new senses will consolidated and propagated within culture. Inversely, if the old system of senses – having turned out to be "better" in a given cultural context – resists any innovations ("mutations"), then the *status quo* will be preserved.

One can also speak about natural selection applied to entire cultures, considering their instrumental efficiency (e.g. economic, military, ideological, artistic, culinary). By no means does this constitute an expression of the tendency towards biologisation of the whole culture. I would only like to point out some functional similarities of the biological and psychical-cultural levels, according to the spirit of Bertalanfy's systems theory.

* * *

Finally, it is time to summarise very shortly the present book. Its main axis is constituted by the postulated unity of three distinguished levels of reality: the physical level, biological level and psychical level. In my opinion, the division into these levels is mostly a property of our picture of the world, and not of the world in itself. I propose that each higher level emerges from its respective lower level (or, rather, we perceive that something apparently completely new emerges at subsequent levels) as a result of the relation of self-applicability (directing on oneself). The psychical level appears when the cognitive centre in the brain (and the picture of the world formed within it) becomes directed on itself. The biological level emerges when some system of regulatory mechanisms (mostly: negative feedbacks) begins to take care exclusively of its own survival and propagation (reproduction, equivalent to a positive feedback). It may be that the physical level (this is only my speculation though) comes into being when some mathematical algorithm begins to refer to itself (like some simple mathematical procedure applied to itself generates the enormously complex Mandelbrot's set). If my point of view is correct, this would lead to strange implications.

Why? Well, the just discussed relation of self-applicability underlies three (in fact much more) famous logical paradoxes: the liar's paradox, Russell's antinomy of classes and Gödel's proof. They are called paradoxes (or antinomies) because they (especially the latter two) have shaken the fundaments of logic and mathematics. Although they are constructed according to rigorous logical rules, they can be neither true nor false. The liar's paradox can be best expressed in the following way: "Hereby I am lying". The formulation of Russell's antinomy of classes is a little bit more complicated: "Is the class of all classes, that are not their own elements, its own element or is it not its own element?". Both answers lead to contradiction. Gödel demonstrated in his famous proof that the following nonsensical statement can be constructed in each sufficiently complex arithmetical system: "Hereby it is proven that this statement cannot be proven"[103]. Russel's antinomy and Gödel's proof were a great shock for many logicians and mathematicians,

[103] This is my a little bit simplified formulation of Gödel's proof.

since these paradoxes excluded for ever the possibility for logic and mathematics to be absolutely perfect, coherent and completed systems. There will always exist sentences (statements) within these systems that are neither true nor false.

As I mentioned, the relation of self-applicability which underlies the above paradoxes is also responsible (in my opinion) for the emergence (at least in our picture of the world) of the subsequent levels of reality. Therefore, it may be that our logic and mathematics do not reflect properly the most fundamental properties of the real world, and what seems to be a paradox for our idealistic logico-mathematical systems is in fact the most fundamental aspect of this world. If so, we all, as living beings possessing (self)consciousness, are children of paradox.