Part II

CONTROLLING INFORMATION

To give an account of the how and why anything is to trace it to its material, to its essential characteristics, and to its provoking cause; for in investigating the genesis of a thing men are chiefly concerned with the nature of what emerges from the process, with the impulse that initiated the process, and with what was already there to undergo the process from the start.

[*ARISTOTLE Phys.*, 198a32–35]

But how are we to understand "potentiality" here? Not in the sense in which we say that the potentiality of the statue exists in the bronze; for that implies that the whole of the bronze may actually become the statue, whereas it is not so with an illimitable potentiality, since it can never become an unlimited actuality.

[*ARISTOTLE Phys.*, 206a18–21]

Rhythm results wherever there is a conflict of forces not in equilibrium. If the antagonist forces at any point are balanced, there is rest; and in the absence of motion there can of course be no rhythm. But if instead of a balance there is an excess of force in one direction; if, as necessarily follows, motion is set up in that direction; then for the motion to continue uniformly in that direction, the moving matter must, notwithstanding its unceasing change of place, present unchanging relations to the sources of force by which its motion is produced and opposed. This however is impossible. Every further transfer through space, by altering the ratio between the forces concerned, must prevent uniformity of movement. And if the movement cannot be uniform, then (save where it is destroyed, or rather transformed, as by the collision of two bodies travelling through space in a straight line towards each other) the only alternative is rhythm.

[*SPENCER* 1860–62, p. 228]

6
Complexity: A Necessary Condition

This chapter raises some fundamental questions about the concept of *function* in biology. Especially in the last chapter we have seen that there are functions of the brain that cannot be reduced to a pure information-processing [Sec. 4.6] or an information-acquiring [Sec. 5.4], particularly when goal-directed actions establishing motor equivalence classes are involved [Sec. 5.3]. The first part of the book, dealing with information-acquiring, began with quantum mechanics [Sec. 1.1]: Now, we shall see how to integrate information and complexity when managing this problem. In this way, we shall also discover some interesting connections between quantum mechanics and complexity.

As a matter of fact, it is difficult to understand the way the brain works without considering some form of *reference* to external objects in which the goal level of biological systems is considered [Subsecs. 5.3.2–5.3.3]. We shall see that the main problem is a lack of appropriate consideration of the biological dimension. Obviously, the brain is a very special biological system, although, as I have explained [Sec. 3.1], it has been considered as the prototype of biological information-processing and -acquiring, like a sort of device or prebiotic physical system. Then, in the main part of this chapter, we shall try to focus on some necessary conditions of living beings, namely self-organization and complexity.

6.1 Information, Reference, and Functions

6.1.1 An Inversion in the Information "Flow"

Here, I shall show why it is necessary to go much further than the common conceptualization of information. We have already considered the difficulty of starting with a pure information-processing view of biological systems and information-*acquiring* in general. Now, the basic knowledge that we have acquired thus far allows us to deal with this very problem.

As is well known, the classical theory of information acquisition was more precisely a theory of information communication, thus preserving only a part of the original cybernetic program [Subsecs. 2.3.1 and 3.2.2]. In a controlled communication context, what is important is *how much* information the receiver r acquires about the source s: It concerns a certain weighted *average* of the conditional probability of all signals—having the form (2.3)—that can be transmitted from s to r, and therefore has nothing to do with single events or objects (whose meaning is, for this reason, irrelevant). In other words, the quantity of information embodied in a signal is only incidentally related to the reference of that signal. The most important quantities at play are the amount of information generated by a particular state of affairs (a source) and the amount of information carried by a particular signal about that state of affairs.

Such a view has been corrected but not radically changed through our general explanation of information-acquiring [Subsec. 2.3.2]. This explanation does not have a statistical nature but it still cannot solve our problem as such. Indeed, the most important differences we have introduced relative to the classical (communication-theory) understanding of information acquiring are that:

(1) There is a certain dependency (through coupling) of the selecting device on the input processor that allows such a selection to be a selection of (a part of) the input information.
(2) Selection occurs at the end of the information-acquiring process and not at the start, for the initial processor is only a source of variety.
(3) Only the selection at the end determines an irreversible process of information-acquiring.

Point (1) is also partly acknowledged by classical information theory as long as the channel is taken into account. However, in that context, classical information theory does not sufficiently account for the fact that this dependency is fundamental for the subsequent act of selection of the output information. This also means that often the problem is not understood in its conceptual generality: The issue at stake is the necessary dependence of any selection on a coupling and not on the transmission, which is a much more specific problem (quantum mechanics shows that there can be indeed information sharing without information transmission [Subsecs. 2.2.2–2.2.3]). Now, it is also clear that this selection act is *about* the input information. However, this aboutness cannot be understood, at this level, in terms of a *referential* act. It is rather a pure physical process in which a system is able to acquire (and eventually store) information about another system even without any kind of referentiality or intentionality in that sense [Sec. 5.4]. Here, this physical acquisition of information *could* eventually be used by some agency, so that we can speak of a form of potential information [Subsec. 2.2.2]. Let me give an example: If, in one laboratory, the position of a particle is accidentally measured (say, due to some electric fluctuation that switches on a measuring device while nobody is in the laboratory) and eventually even stored, some scholar thereafter having access to the storage device can recover this potential information, and tell something about the particle. However, we cannot say here that the selecting devices (detectors) or the storing ones were purposeful or goal-directed to the particle. The same thing is what spontaneously happens in nature everyday with quantum systems [Subsecs. 1.3.3 and 2.2.4]. The situation is totally different for living beings. Organisms must informationally *control* their environment to survive, which means that they cannot acquire information by chance or wait for spontaneous physical processes. They must actively search for the signals that allow them to control situations that are fundamental for their own survival. In the previous chapter we have indeed seen that when actions and motor aspects are involved we have a mapping from the agent to the world and not vice versa as it happens in information acquiring. Therefore, information-acquiring, even in the expanded form I have proposed, is no longer sufficient.

Given this proviso, the model for information-acquiring I have proposed in Subsecs. 2.3.1–2.3.2 has a considerable advantage relative to the classical one, namely that of representing *a necessary condition* of a goal-directed information control. Indeed, there can be no information control on some source if there is no selection at the end of the process, assuming the input is a source of variety and not already a defined message. If this is the case, it makes perfect sense, under suitable conditions, to apply the final selection act again to the unknown input information in order to compare the first and the second outputs (i.e. producing interpretations in a wide sense) and in this way try to extract more information about the input. If the input information (in relative absence of fluctuation or noise, i.e. in normal conditions) coincided from the start with the selected message (as assumed in the classical communication theory), this necessity would not arise at all. I am not speaking here of the weighted average of the conditional probability of all signals that can be transmitted from the source to the receiver, or of the higher or smaller fidelity of the

transmission. These are quantitative issues that would be largely insufficient to explain how *a new class* of physical systems has arisen in the history of out planet, that is, organisms as biological, complex, systems able to control information.[1] The problem is about a yes-or-no alternative and concerns the fact that the unknown input information could have a *totally different relevance* for the self-maintenance of the system, e.g., being, disruptive and not survival-promoting. Thus, in a situation in which there is incertitude about the source, to be able to have a goal-directed action on this source (i.e. performing information control on it) will be crucial for the emergence and survival of more complex systems. In other words, I am trying to derive the possibility of a goal-directed behavior from the necessity of ensuring, in determined conditions, that the final selection act is really adequate to the unknown input information. Summing up, the problem of the organism is *not* to get some information from a certain source (an operation that any physical system can do), but, following von Helmholtz[2] [Subsec. 4.4.5], to operate an inverse (Bayesian) inference: To estimate the factors that may have given rise to a certain sensation or perception once this sensation or perception has occurred.[3]

Therefore, the problem of the reference of the information dealt with by biological systems is crucial for their operations and functions. Often, the formulation of proposals for dealing with this problem was strictly dependent on an examination of psychological and mental processes, that is, on aspects that are not completely adequate for this context of inquiry. However, these considerations are also relevant for our investigation and I shall try to reformulate them in a language that is more appropriate to the present context of discussion.

6.1.2 A Semantic Computationalism

One of the most difficult concepts in science and philosophy and certainly one of the less understood is that of *function*. It is a ubiquitous concept, and there probably does not exist a single field in biology or cognitive science in which it is not extensively used. We essentially have three options for dealing with this problem: To assume that there is a mapping between structural properties of the system and the functional level, to consider functions as detached from the material substrate of which or on which they are functions (functionalism), and to formulate the hypothesis that functions somehow emerge from the lower level of reality. The option that I adopt, as explained in Ch. 2, is the third one. Here, organisms are understood as emergent complex systems controlling environmental information and therefore deploying many different vital functions. Before I develop this point of view, let me briefly examine the other two options.

To speak of a system's function and goal means to go further than both the physical level and the pure computational level. The necessity of taking into account the problem of the reference of a certain message (and even somehow a semantics for information-processing), and therefore to go beyond a pure syntactic assumption [Subsec. 3.2.1] was clear not only to the fathers of cybernetics but also to some of the leading thinkers of classical computation, for instance Zenon Pylyshyn[4] [see also Subsec. 3.7.2]. Pylyshyn was well aware that mental activity and at least certain types of human behavior are determined by representations. According to him, however, we can also distinguish syntactic and semantic aspects in a computer. Concepts are represented in the machine as symbolic expressions and programs, while it is the physical realization of these representations that determines the machine's behavior.

[1] See also [*ROSEN* 2000, pp. 288–96]. [2] [*VON HELMHOLTZ* 1867, pp. 586–611].
[3] [YUILLE/KERSTEN 2006] [SUMMERFIELD/KOECHLIN 2008].
[4] [PYLYSHYN 1980]. See also [*PYLYSHYN* 1984].

These rules and concepts are:

1. Expressed in terms of syntactic operations over symbolic expressions,
2. These expressions are then "interpreted" by the inbuilt functional properties of the physical device.

Of course, the machine does not interpret the symbols—as letters, for example: They are only formal patterns that cause the machine to function in some particular way. For this reason, according to Pylyshyn, all semantic distinctions must be mirrored by syntactic distinctions and such features must in turn be reflected in functional differences in the operation of the device. Jerry Fodor also felt the necessity to correct pure (syntactic) computationalism with some form of representationalism.[5]

The fact remains that this standpoint does not seem to account for what is fundamental in the purposeful action of living organisms. When this purposeful action is reduced by Pylyshyn to semantics, and this in turn to symbols, it seems a way to bypass the problem. Theorists like Pylyshyn cannot explain what holds together the representational and the computational. The heart of the problem is that computers do not operate on symbols with semantic content. Whatever meaning, truth, or reference (and therefore symbolic value or content) programs have is *derivative*, tracing back to interpretations imposed by the *programmers and users* of the system.[6] The semantic computationalism thesis is that computers, like human brains, manipulate symbols.[7] This is exactly what computers cannot do. Computers function the way they do only because there are humans that have programmed them to do so and other humans that use them in order to obtain certain results. The programs are truth-preserving in the same sense in which equivalent rules would preserve truth when applied by logicians to symbols on paper or a blackboard. It is true that classical computers can be seen as second-order machines: Given the formal specification, i.e. the program, of a first-order machine, they will "become" that machine. This, however, does not solve the problem of the relation between structure and behaviors: Here, the greatest problem is represented by the fact that, in general, we cannot derive behaviors from structure, nor structure from behaviors.[8] Moreover, computers do not seem to be the best model of the brain, since a (classical) computer is essentially a passive device, while the brain is essentially a dynamic device, for producing and controlling certain outputs as shown in Ch. 5. Finally, it is difficult to conceive the brain in terms of a problem–solving device (which would act as a pure information-processing device for solving any type of problem), since the main difficulty with the general problem-solver paradigm [Subsec. 3.2.1] is that there is not a single state-space in which any solution may be found.[9] However, in order to be a general problem-solver a system must be able to reduce any type of problem to a single space. The conclusion is simply that living beings are not organized in the way that classical computationalism, assumed them to be,[10] and this also holds for its semantic or representational version.

6.1.3 Dretske's Proposal

I shall consider now a proposal by Fred Dretske that is very much related to semantic computationalism. According to Dretske,[11] a signal should not only carry enough information [Subsec. 6.1.2] but also the *right* information, i.e. it must carry the information quantity generated by the system

[5][FODOR 1980]. For this reason, Fodor was criticized by Stich [*STICH* 1983]. [6][SAYRE 1986].
[7][NEWELL *et al.* 1958]. [8][LANGTON 1989b]. [9][*BODEN* 1988, pp. 151–4]. [10][BEER 1995a].
[11][*DRETSKE* 1981, DRETSKE 1983].

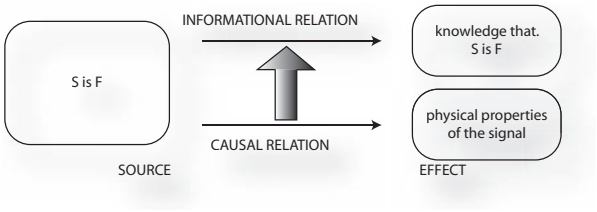

Fig. 6.1 How, according to Dretske's proposal, the *physical* properties of the signal determine the correct informational reference of the message.

s having the property F (and not, say, the property G) *and* s must actually possess F. Therefore, equivocation must be zero.[12] Now, according to Dretske, knowledge is caused by the information that s has the property F if and only if those physical properties of the signal by virtue of which it carries this information are the ones that are causally efficacious in the production of the knowledge; in other words, the knowledge that s has the property F is caused by the information that s in fact has the property F which in turns derives from a causal relation. Here, the indexical relation of the signal to the property F of the system s is grounded on the causal connection between the physical properties of the signal and the property F of s [Fig. 6.1].

The reader will recognize in this case a more refined treatment of a similar argument already used by Pylyshyn and discussed in the previous subsection (information mirrors some structural properties of a signal). For example, if my home's doorbell rings three times, I know that it is my friend John. Here the information-carrying physical property of the signal is the temporal pattern (which is connected with a specific causal action: John pushing three times on the bell) and not the amplitude or the pitch. However, this also conveys a specific information to me: My friend John is at the door. If a fly is frightened away, according to Dretske, this is a physical consequence and *not the effect* of information because the fly would be disturbed by any sequence of rings or knocks. However, I remark that if there were someone else in my place, he or she would not understand that it is John at the door, which shows that, together with the information-carrying physical properties of the signal, I need a previous agreement with my friend. In other words, the physical properties of the signal alone do not ensure the right transmission of information. In fact, when I say "This is my friend John," this is already an interpretation and not a simple exchange of information. Conversely, the fly also receives some information (and without agreement): A pitched ring can sound as an alarm signal that something has happened, and this not only according to many people but also to some animals.

In conclusion, it seems to me that Dretske's proposal says both too much and too little and therefore misses the problem of the referent by trying to establish a direct relationship between the codified feature of information (the structure of the signal) and the way in which one receives this as a meaningful piece of information about the source that produced the signal (which is a referential relation). However, a signal as such tells nothing about its source, nor about any other referent if there is not either (in the human case) a previous understanding of, or (for organisms) information control on, some *additional conditions* allowing that result. The fundamental reason for these difficulties is the fact that any informational sources is *not the cause* of information reception (or selection) by itself [Subsec. 2.2.1].

[12]On this point see [ARBIB 1983] [ARMSTRONG 1983a]

6.1.4 Functionalism

As we have seen [Secs. 3.1–3.2], one of the first answers in cognitive science to the problem we are dealing with in this section was functionalism, the true background of the treatment of the brain as an information-acquiring and -processing device. It is strongly based on the distinction between hardware and software, where the assumption is that the same software can run on different hardwares, such as brain functionalities being instantiated either in the brain or in computers.[13] Actually, the very same physical state recurs in a computer under circumstances in which very different processes, operating in quite different domains of interpretation, are being executed. This is the reason for providing a functional description beyond the physical or even a pure structural one. It is a fact that all the different schools of cognitive sciences agreed to a certain extent with functionalism.[14] Also, connectionism inherited this perspective [Subsec. 3.8.1].

It is important to understand from the start that functionalism cannot be very helpful for our task. Indeed, the set of possible configurations in the real world that can give rise to a given biological or cognitive function is in general very limited and there is therefore a certain link between structure and function, even if this link cannot be understood too strictly, given that there is always a certain underdetermination both ways (a single structure can give rise to several functions, a function can be realized in different structures[15]). We have already met this problem in examining Hubel and Wiesel's theory of vision [Sec. 4.3]. For instance, to fly you need a certain wing-like structure. So, the structure has a certain relevance here. Instead applying a strict functionalist point of view to certain configurations or structures can lead to some paradoxes. For instance, Ned Block[16] showed that there are systems without mental states to which, according to functionalism, one should nevertheless attribute mental states—for example, China's population could be understood as a giant neural network. It is possible that functionalism makes a confusion between information coding, which can indeed be performed in an arbitrary number of different bases [Subsec. 2.2.2], and the issue of the system's characteristics that can give rise to a certain function. Finally, I have also mentioned that in the brain there is ultimately only hardware and no software [Subsec. 3.3.1]; an issue that shall be investigated in this part of the book.

6.1.5 A Closer Look at Our Problem

Therefore, there are two aspects of the brain that seem difficult to reconcile: (a) The fact that the brain resembles an information-processing or -acquiring machine and (b) the fact that a living organism behaves typically in a purposive and adaptive way implementing specific functions, as already understood by the fathers of cybernetics[17] [Subsec. 3.2.2]. Cybernetics was well aware that the notion of purpose is central for exploring the ways biological systems work. Again, I emphasize here the tension between cybernetics and cognitive sciences' functionalism. In particular, the difficulty is not with reflex behavior but with learned behavior, which is not inborn and is not genetically determined in detail. When organisms learn, not only does their behavior change, but it usually changes for the better (to a more survival-promoting form). What is the neuronal correlate of learning? Association may be a first degree of learning [Sec. 3.8]. However, it is difficult to lead any learning activity from pure association: The correctness of the organism's response is not found in the process or acquisition itself but in the relations that it bears to *other* processes, even in very elementary conditioning experiments. For instance, when a dog is conditioned to salivate, this could not happen without the neural activity that determines this process together

[13][PUTNAM 1967, *PUTNAM* 1981, *PUTNAM* 1988]. [14]See also [*DUPUY* 1994].
[15][*PIAGET* 1967, pp. 144–47]. [16][BLOCK 1978]. [17][*ASHBY* 1952].

with the other ones that are connected to the activity of eating. Moreover, we have seen that perception is always connected with some expectancy [Sec. 4.1]. How can we account for this in a pure information-acquiring model?

It seems to me that the only solution is to consider the brain as a biological organ, and try to understand how higher cognitive functions have arisen starting with more modest but still extraordinary biological functions. Artificial Intelligence, both in its computationalist (syntax only or at best semantics mapped to syntax) and representationalist-functionalist versions possibly misconstrued the nature of the plastic behavior displayed by organisms. Even the low cockroach displays robust (apt to resist the environment), flexible, practical intelligence that computers lack.[18] The cockroach is indeed much more than the simple "sense-predator and initiate-random-run" command.[19] Two main features common to all biological systems are to be carefully considered here:

- Biological systems are *softly assembled* [Subsecs. 4.4.4–5.3.3]. This means that most subsystems are not crucial for the sustenance of the whole and the whole organization emerges through appropriate local interactions. As we shall see, this is an important feature of complex systems. In this way, individual variability in biological systems—which reflects such a trade-off—is a consequence of their being soft assembled and cannot be considered as noise or an obstacle but as an essential feature which makes them able to respond to environmental fluctuations.[20]
- However, there is also a manipulation of the external world that obeys a *hierarchy of goals and levels* (cognitive competencies are also acquired and rooted in manipulations of the external world) [Subsec. 5.3.2–5.3.3]. This second feature neither comes solely from, nor depends solely on, the complex nature of biological systems, even if complexity already implies, at a structural level, a certain hierarchy, as we shall see in the following. In other words, complexity is a necessary but not sufficient condition for having biological systems: It is an important property of biological systems but it does not exhaust their reality.

The combination of these two features is really unique. If only the first were important, suitable connections of complex patterns would suffice to explain life, as the connectionist approach maintains. Indeed, as we shall see, the weakest point of neural networks (as well as of connectionism[21]) is that artificial networks are mostly trained while life is self-training, i.e. it has the capacity of self-correcting. The latter capacity would not be allowed by a life conceived only as an assembly of parts or production of patterns. In the history of science and philosophy this problem has been deeply considered, especially in relation to the way the mind is related to the external world. If our mind or our brain were only a collection of images and descriptions without goals and purposeful acts, it would be like a screen, where different images come and go without any relationship among them.[22] This is precisely the original point of view of associationism, which we have seen to be the father of connectionism and distributed computation [Subsec. 3.8.1]. Indeed, Hume says[23]:

> The mind is a kind of theatre, where several perceptions successively make their appearance; pass, re-pass, glide away, and mingle in an infinite variety of postures and situations. [...] The comparison of the theatre must not mislead us. They are the successive perceptions only, that constitute the mind; nor have we the most distant notion of the place, where these scenes are represented, or of the materials, of which it is compos'd.

A pantomiming or theatrical being could never survive in a real world. The point is: If the brain or the mind were like this, an organism could never understand that two or more images can

[18] [*CLARK* 1997, pp. 1–3, 43–61]. [19] [RITZMANN 1993]. [20] [*CLARK* 1997, p. 81].
[21] For instance see [*CHURCHLAND* 1995]. [22] [AULETTA 2003a]. [23] [*HUME* 1739–40, p. 253].

be related to the *same* referent and not be images of different referents (they are indeed only images, which continuously alternate or replace one another [Subsec. 4.4.1]), and therefore could never correct its error.[24] In my opinion, organisms use the ability to give rise to new reactions by associating new schemata to an already known referent or applying old schemata to different and new referents. Therefore, the problem of any organism is not representation (or perception) in itself, neither behavior as such, but rather *error correction*[25]: As a matter of fact, the necessity of information control arises from the need to correct the way in which an information source has been dealt with [Subsec. 6.1.1], a problem demanding a totally different approach to be solved in a satisfactory way. The theme of error correction or suppression will emerge repeatedly in what follows at several levels. For example, suppression of prediction errors is formally related to minimizing surprise [Subsec. 2.3.3] by sampling environmental inputs to maintain homeostasis. At a high level of sensorimotor and cognitive functions, it appears as predictive coding and in the context of reward prediction error and reinforcement learning. Here and in the following few chapters I shall first focus on complexity. The consideration of behavior will grow out of such an enquiry.

6.2 Matter Structures and Combinatorics

In order to overcome the previous difficulties, it is crucial to consider how living organisms emerge from physical reality. Indeed, only an elementary consideration of biological systems can really ground our understanding of the relations between cognition and biology. Organisms are complex systems but their basic structures and interactions are of a molecular and chemical type. Therefore, we must preliminarily deal with this issue here so as to bridge the gap between our brief account of quantum physics in the first part of the book and an analysis of complexity in this and in the next chapters. In particular, to have a short look at molecular physics and basic chemistry, as well as at complexity theory, is crucial in order to correctly understand the specific way organisms build their order through their metabolic processes. We shall subsequently discover that this metabolic aspect is not independent from a cognitive one.

As remarked, the fundamental atomic and molecular building blocks of our world do not necessarily present the codification structure that is typical of quantum information [Subsec. 2.4.1]. Here, entropic and thermodynamic considerations are fundamental. Apart from some relevant considerations about combinatorics at the beginning of the next subsection, the following pages should be understood as a very quick reminder of these issues for the reader who has studied these disciplines in the past or as a short introduction to them for those who do not know but would like to acquire some basic notions (indeed they can turn out to be useful in what follows). Otherwise, the present section as well as part of the next one could easily be skipped.

6.2.1 Atoms

When several quantum systems interact, it is possible that they become localized, losing in this way the typical nonlocal aspects (features) that they have at a basic level. At that level, there is no univocal and general space–time structure [Sec. 1.3]. On the contrary, when they become localized its is reasonable to assume that space and time emerge. It is also quite reasonable to assume that

[24] This is a well-known problem in AI, but it is in general taken as an unavoidable fact and not one that is essentially in conflict with the way organisms deal with their environment [*HAWKINS/BLAKESLEE* 2004, pp. 63–4].

[25] [SHAPIRO 2007].

space emerges as a sort of hypersurface (in a kind of holographic process) enveloping one or more localized quantum systems. In this case, space (as well as time and gravitation) would emerge out of the information that quantum-mechanical systems represent[26] [Sec. 2.2]. The fact that this information is then projected on this hypersurface is a further consequence of the general principle that we can have access to information only conditional on some effects at the threshold between the original event (enclosed by that surface) and some other system [Subsec. 2.3.1]. It is also reasonable that in such new conditions new kinds of interconnections emerge that often no longer represent codified information or a mapping to codified information, even if they originate from centers of codified information. This is finally due to the randomness of these multiple interactions. In this case, physical parameters like mass and energy of the involved particles play a central role in determining the constraints that allow for new kinds of interconnections that no longer have a direct informational value. In other words, quantum-mechanical information is the source of molecular order but the latter is a true emergent order of reality [Subsec. 2.4.2]. To be specific, electrons, protons, and neutrons as well as atoms themselves to a certain extent can still be understood as a finite set of elementary units that (with a certain proviso) are also alternative. However, they can be combined [Fig. 2.7] but *not lineary* combined in the same sense in which informational units do (indeed the birth of gravitation due to massive particles has produced non-linear effects in the history of our universe). This is evident when considering the process of the solar fusion of hydrogen atoms for constituting helium, in which a large amount of energy is released. Another interesting problem is when a hydrogen molecule (out of two hydrogen atoms) is constituted. Here, the electron's motion generates a charge cloud that overcomes the nuclei's repulsion, thus allowing for the formation of the molecule. In general, we cannot say that a molecule is formed by atoms but rather by elemental centers.[27] All this means that there is no syntax ruling these combinatorics, otherwise we are forced to admit that the term *syntax* only means "lawful combination," which is a very poor meaning. Finally, all these combinatorics show no alternative codes. This means that there are no alternative ways for a hydrogen atom to be built. This in turn implies that there is no code at all and that we cannot consider atoms and molecules as a set of instructions for doing something else. However, the above-mentioned characters still allow atoms to be *used* as single-code information codifiers (they are in either the excited or the ground level, a circumstance used in quantum computation), even if molecules are increasingly difficult to use for that task. Molecules could perhaps be understood as noncodified combinations out of single-code binary nodes (atoms). Moreover, any physical process (from atoms to black holes) still presents a certain entropy, which allows the possibility that codified information is recreated elsewhere in appropriate conditions.

When an electron (which represents an electrical negative charge) and a proton (which represents an electric positive charge) interact to constitute an atom, the energy acquires discontinuous values [Subsec. 1.2.4]. This means that electrons occupy certain orbital levels in a discontinuous manner. If $E_0/2 = 13.6$ eV (eV meaning electronvolt[28]) is the energy of the ground (lowest or first) level, the energy of each n level ($n = 1, 2, 3, \ldots$) is given by the formula[29]

[26][VERLINDE 2011]. Unfortunately, the author follows thermodynamical considerations instead of quantum-mechanical ones, but the main insight is very fruitful.

[27][EARLEY 2006].

[28]It is equal to the amount of kinetic energy gained by a single unbound electron when it accelerates through an electric potential difference of one volt.

[29][*AULETTA et al.* 2009, Ch. 11].

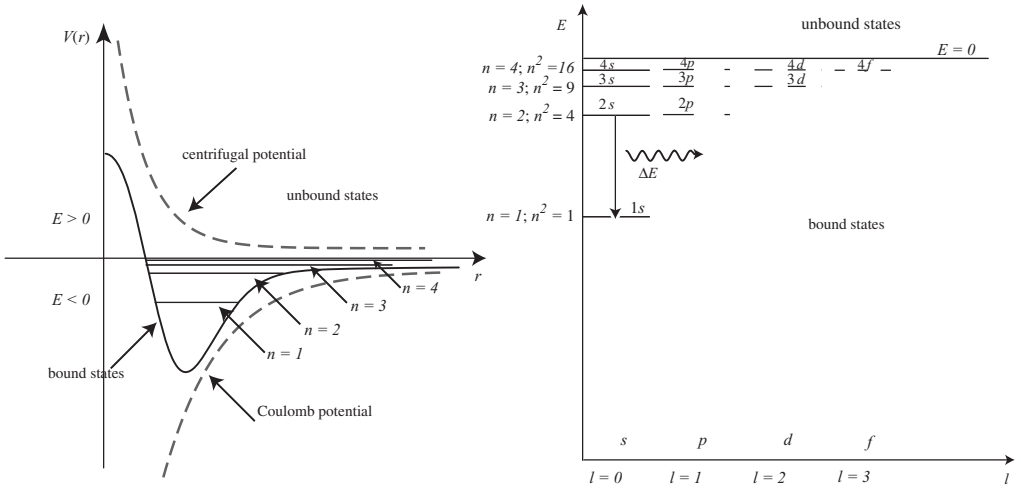

Fig. 6.2 On the left, the electromagnetic potential as a function of the distance r (on the abscissa) of the electrons from the nucleus. With energy below zero we have trapped (bound) states for the electrons, i.e. possible orbitals at specific energy levels.

On the right is a schematic representation of the same populated energy levels (only the first 4 shown) as a function of the two quantum numbers n (indicating the orbital shell) and l (l is related to the orbital momentum of the electrons). The quantity n^2 indicates the number of suborbitals for each shell, which are 1 for any s wave, 3 for any p wave, 5 for any d wave, 7 for any f wave. Therefore, level $n = 1$ has energy $-E_0/2$ and number of orbitals $n^2 = 1$. The level $n = 2$ has energy $-E_0/8$ and number of orbitals $n^2 = 4 = 1 + 3$. The level $n = 3$ has energy $-E_0/18$ and number of orbitals $n^2 = 9 = 1 + 3 + 5$. The level $n = 4$ has energy $-E_0/32$ and number of orbitals $n^2 = 16 = 1 + 3 + 5 + 7$. Each orbital "position" can be occupied by two electrons with opposite magnetic polarization. The arrow shows the transition of an electron from the $n = 2, l = 0$ state to the $n = 1, l = 0$ state, with emission of a photon of energy $\Delta E = E(n = 2) - E(n = 1) = -E_0/8 + E_0/2 = 3E_0/8$.

$$E_n = -\frac{1}{2n^2} E_0, \tag{6.1}$$

which, when going further and further away from the central nuclear charge represented by the proton, tends to a continuous distribution (absence of interaction, when $E = 0$) [Fig. 6.2]. This can easily be understood by recalling that an electron behaves like a nonclassical wave [Subsec. 1.2.2]. This means that it oscillates at certain frequencies and that the more energy the electron has, the higher the frequency at which it swings (like an ordinary guitar string that can be played with different strengths). This means that the electron needs a certain precise amount of energy in order to fit into a certain orbital shell (only an integer multiple of the oscillations is stable) and that a little bit more or less would not "close" the curve, with the consequence of a self-destructive interference [Fig. 6.3].

I make use here of a one-dimensional wave function $\psi(x)$ of the position x which is another way to describe the state of a quantum system and corresponds to the state vector $|\psi\rangle$. It can be mathematically thought of as a scalar product $\langle x \,|\, \psi\rangle$ between this vector and the eigenvector $|x\rangle$ of the position [Subsecs. 1.2.3 and 1.2.7]. When the number l grows, we have an increasing

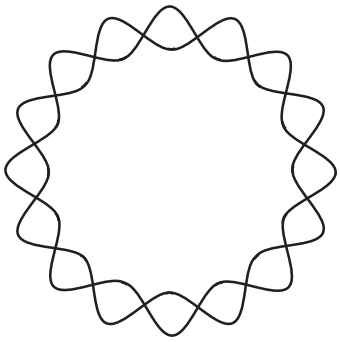

Fig. 6.3 The oscillation of an electron needs to be a precise frequency to close the curve at a certain orbital.

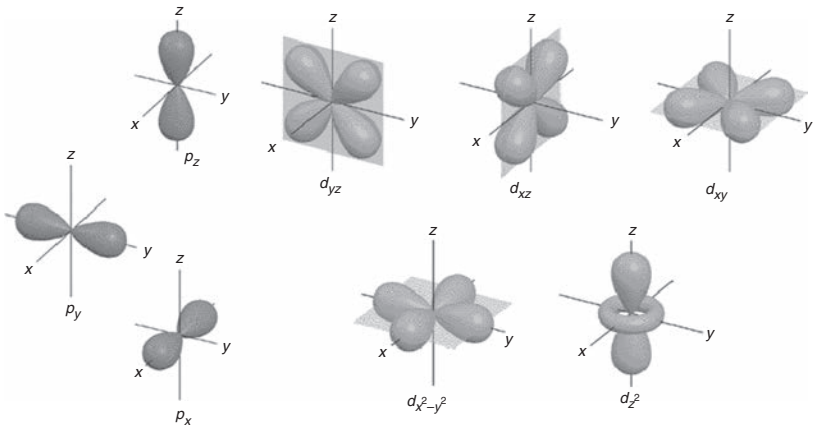

Fig. 6.4 Representation of p-states (the three states on the left) and d-states (the five states on the right). Adapted from http://www.chem.ufl.edu/itl/2045_s00/lectures/lec_10.html.

number of components for each l as well as for the suborbitals of each orbital shell ($n = 1, 2, 3, \ldots$). The components are called s, p, d, f states. The s states show a spherical symmetry, while p states show a double bulk about the x, y, and z axes, one bulk with a positive sign, the other one with a negative sign, corresponding to the positive or negative phase of the wave function [Fig. 1.3, and Fig. 6.4]. Indeed, it is very important to understand that electrons cannot occupy the same orbital level in the same state (the Pauli exclusion principle) and therefore need to have opposite directions (up or down) relative to a fixed axis of magnetic polarity called spin. The spherical symmetry of the s states is a consequence of quantum mechanics and especially of the spatial symmetry of the wave function. Also the other symmetries of molecules are a consequence of quantum-mechanical distributions of atoms and electrons. Taking into account the limitation imposed by the exclusion principle, we can have two electrons in the first shell (denoted by $1s$), 8 electrons in the second shell (2 electrons for $2s$ and two times the three axes = 6 for $2p$), 18 electrons for the third shell (2 for $3s$, 6 for $3p$, and 10 for $3d$), 32 for the fourth shell (2 for $4s$, 6 for $4p$, 10 for $4d$, and 14 for $4f$), and so on. The order of occupation is

$$1s\ 2s\ 2p\ 3s\ 3p\ 4s\ 3d\ 4p\ 5s\ 4d\ 5p\ 6s\ 5d\ 4f\ 6p\ 7s\ 6d\ 5f$$

After the latter level ($5f$), it becomes increasingly difficult to get stable elements due to the weakness of the electromagnetic force binding electrons to protons when the former are increasingly distant from the nucleus, approaching the limit ($E = 0$) between bound and unbound states.

Let me now give some examples.[30] The carbon atom C (number of electrons $Z = 6$) is described by $1s^2 2s^2 2p^2$, where $1s^2$ means two electrons in the orbital $1s$, and so on. However, the last two electrons should occupy two different $2p$ orbitals because in the mean they repel each other less than if they were in the same orbital. We can assume, in general, that an atom in its ground state adopts a configuration with the greatest number of unpaired electrons. Therefore, C can be thought of as $1s^2 2s^2 2p_x^1 2p_y^1$. Nitrogen N ($Z = 7$) can be thought of as $1s^2 2s^2 2p_x^1 2p_y^1 2p_z^1$ while oxygen O ($Z = 8$) as $1s^2 2s^2 2p_x^2 2p_y^1 2p_z^1$.

6.2.2 The Chemical Elements

This summary may be useful for the following examination. I limit myself to some essential features. Let us consider that all chemical elements are organized in 4 major blocks [Fig. 6.5]:

- The s block (H, He, and the first two groups on the left),
- The p block (groups 13–18),
- The d block (groups 3–12, together with La and Ac),
- The f block (elements 58-71 and 90-103).

By increasing the group number, we have a progressive reduction of the atomic radius and an increase both in ionization energy and electronegativity. It is important to stress that all the macroscopic properties of the classical elements are emergent on the basis of the previously described quantum laws [Subsec. 2.4.2].

The *alkali metals* are very reactive metals that do not occur freely in nature. These metals have only one electron in their outer shell. Therefore, they are ready to lose that one electron in ionic bonding with other elements. Among the alkali metals are: Lithium (Li), sodium (Na), and potassium (K).

The *alkaline earth elements* are metallic elements that have an oxidation number of +2, making them very reactive. Because of their reactivity, they are not found to be free in nature. The alkaline earth elements are Beryllium (Be), magnesium (Mg), calcium (Ca), strontium (Sr), barium (Ba), and radium (Ra).

The 38 elements of the third group are called *transition metals*. As with all metals, the transition elements are both ductile and malleable, and conduct electricity and heat. The interesting thing about transition metals is that their valence electrons, or the electrons they use to combine with other elements, are present in more than one shell. This is the reason why they often exhibit several common oxidation states. Among the transition metals are: Chromium (Cr), manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu), zinc (Zn), silver (Ag), platinum (Pt), gold (Au), mercury (Hg).

Posttransitional metals are unlike the transitional elements in that they do not exhibit variable oxidation states, and their valence electrons are only present in their outer shell. All of these elements are solid, have a relatively high density, and are opaque. The 7 known elements classified here are: Aluminum (Al), gallium (Ga), indium (In), tin (TI), thallium (Sn), lead (Pb), bismuth (Bi).

[30] [*ATKINS/DE PAULA* 2006, pp. 340–1].

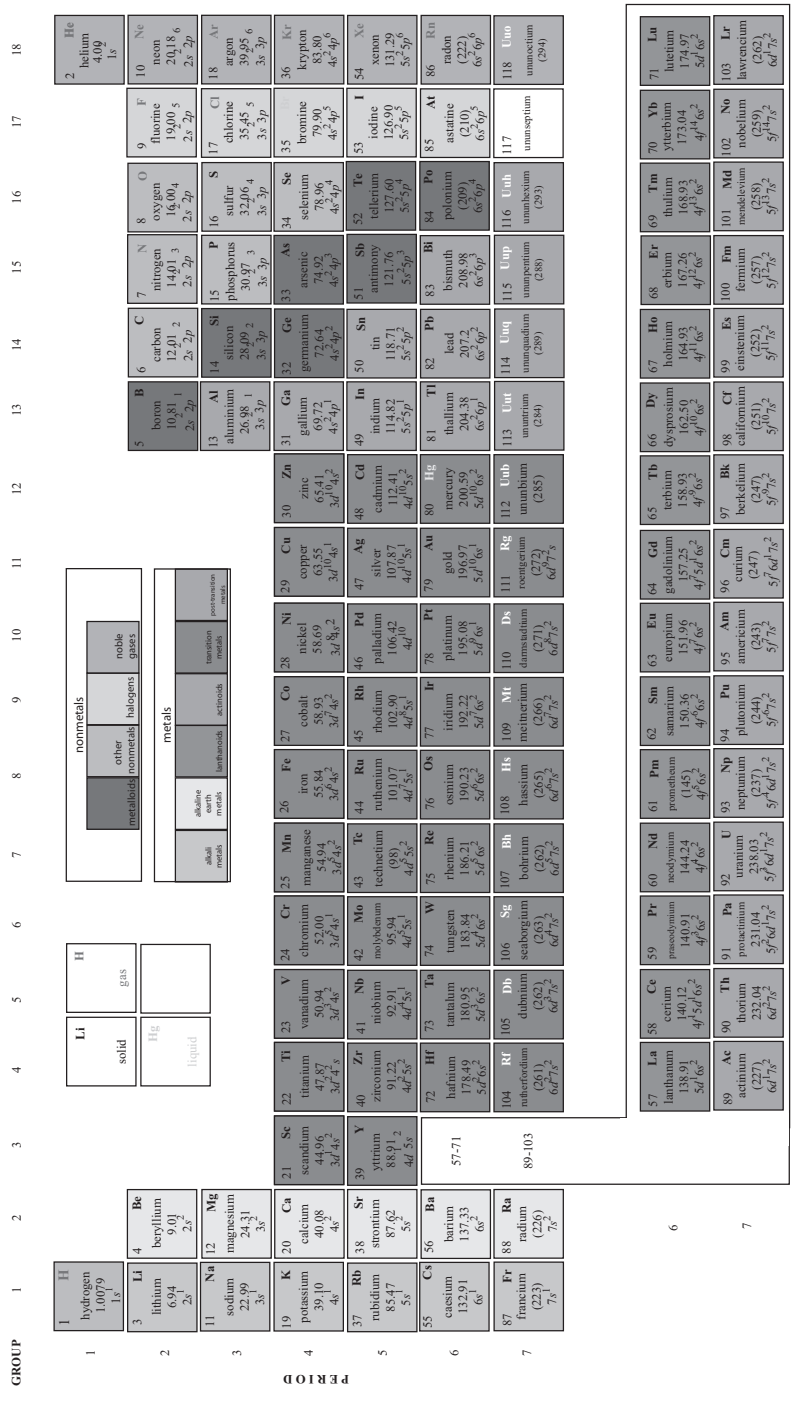

Fig. 6.5 Periodic table of elements. The atomic number of the element (number of protons), symbol, name, atomic weight, and relevant periodic orbitals are shown. (The figure is reproduced in color in the color plate section.)

Metalloids are the elements found along the stairstep line that distinguishes metals from nonmetals. The only exception to this is aluminum, which is classified under the previous group. Metalloids have properties of both metals and nonmetals. Among the metalloids are boron (B), silicon (Si), germanium (Ge), arsenic (As).

Other nonmetals are not able to conduct electricity or heat very well. As opposed to metals, nonmetallic elements are very brittle. Nonmetals are: Hydrogen (H), carbon (C), nitrogen (N), oxygen (O), phosphorus (P), sulfur (S), selenium (Se).

The *halogens* consist of five nonmetallic elements. The term "halogen" means "salt-former" and compounds containing halogens are called "salts." All halogens have 7 electrons in their outer shells, giving them an oxidation number of −1. The halogens are: Fluorine (F), chlorine (Cl), bromine (Br), iodine (I), astatine (At).

The six *noble gases* were considered to be inert gases until the 1960s, because their oxidation number of 0 prevents them from forming compounds readily. All noble gases have the maximum number of electrons possible in their outer shell (2 for Helium, 8 for all others), making them stable. The six noble gases are: Helium (He), neon (Ne), argon (Ar), krypton (Kr), xenon (Xe), radon (Rn). The remnant elements are the so-called rare earth elements.

As said, any orbital (energy level) of an atom can be occupied at most by two electrons with opposite spins. The most elementary atom in nature is hydrogen (H), with the atomic number 1 (1 electron, 1 proton, no neutrons), followed by helium (He), with atomic number 2 (2 electrons at the same energy level, 2 protons, 2 neutrons). Lithium (Li) has the atomic number 3 (3 electrons disposed within two orbitals, 3 protons and 4 neutrons). Berillium (Be) has the atomic number 4 (4 electrons in two orbitals, 4 protons, 5 neutrons), boron (B) has the atomic number 5 (5 electrons in 3 levels, 5, 6). Carbon (C) has the atomic number 6 (6, 6, 6), nitrogen (N) a.n. 7 (7, 7, 7), oxygen (O) a.n. 8 (8, 8, 8), fluorine (F) a.n. 9 (9, 9, 10), neon (Ne) a.n. 10 (10, 10, 10). The second layer of elements (from sodium (Na) to argon (Ar)) have a.ns. from 11 to 18. The third layer goes from potassium (K) to krypton (Kr), with a.ns. from 19 to 36; the 4th layer from rubidium (Rb) to xenon (Xe), with a.ns. from 37 to 54.

6.2.3 Molecules and Chemical Compounds

It is also very important to understand that the macroscopic properties of chemical compounds cannot be derived from more elementary laws and therefore represent a true *emergence* [Subsec. 2.4.2]. Indeed, the average relative locations of the particles in a liquid are expressed by the radial distribution function $g(r)$ defined in such a way that $g(r)r^2dr$ gives the probability that a molecule will be found in the range dr at a distance r from a reference molecule. However, again in accordance with my understanding of emergence, these distributions can only be simulated by knowing the specific macroscopic characters of the fluid in question.[31]

Further strong evidence for emergence is phase transition (from a liquid to a solid state, for instance). Although phase transition is well defined in general terms (the phase that minimizes the free energy is selected), the specific parameters describing a phase transition of a concrete chemical (like water) cannot be guessed *a priori* since experimental data are necessary, and therefore needs to be simulated.[32] The fact is that water is not constituted by H_2O molecules but rather by dynamic aggregates $H_{2n}O_n$, where n is often much larger than unity.[33] Again, general laws do not determine single results.

[31][*ATKINS/DE PAULA* 2006, pp. 606–07]. [32][*ATKINS/DE PAULA* 2006, p. 177]. [33][EARLEY 2006].

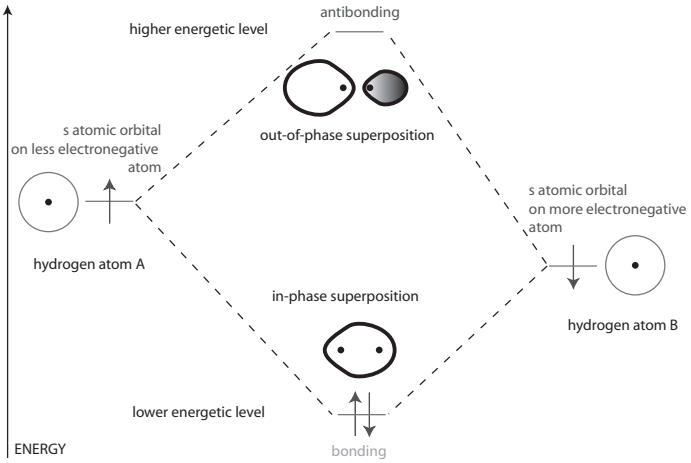

Fig. 6.6 The bonding and antibonding between two hydrogen atoms, A (on the left) and B (on the right). Shadowed and unshadowed regions mean position probabilities for electrons being in opposite phases. The elements like O or F are the more electronegative atoms which contribute more to the bonding orbital (the two opposite spins of the electrons are indicated with double vertical arrows below), while atoms like C are the less electronegative ones, that contribute more to the antibonding orbital. The case in which the electrons are shared equally by the two atoms occupying the same intermediate energetic level, portrays a pure *covalent bond*. When the difference between the two electronegativities is too large, we find that a filled orbital on the anion [see footnote 40, p. 76] has the same energy level as the atomic orbital on one of the atoms and the empty orbital on the cation has the same energy level as the atomic orbital on the other atom. In this case an *ionic bond* is established (like between metals and nonmetals). Here, an intermediate case is shown.

Let us consider how some molecules are formed. In so-called valence–bond theory,[34] the wave function of the molecular hydrogen describing atoms A and B can be written as the superposition

$$\psi(1,2) = A(1)B(2) + A(2)B(1), \tag{6.2}$$

where 1 and 2 designate the two electrons.[35] This means that both electrons are shared by the two atoms. This sharing, combining and establishing interdependencies have not the same nature as information combining.

The so-called LCAO approximation helps us to understand that electrons accumulate in regions where atomic orbitals interfere constructively[36] [Fig. 1.3]. In this case, we have bonding, which we can write down as the probability $\psi^2 = A^2 + B^2 + 2AB$ for any two atoms A and B. These positively interfering regions (called internuclear) allow for orbital shrinkage that improves electron–proton interaction more than it is decreased [Fig. 6.6]. They are regions of lower energy. We have antibonding when $\psi^2 = A^2 + B^2 - 2AB$, an orbital that, when occupied, contributes to a reduction of the cohesion of the two atoms because the electrons are found everywhere but between the two nuclei, exposing the latter in this way [Subsec. 6.2.1] and determining a repulsive reaction between the atoms.

[34] [*AULETTA et al.* 2009, Ch. 12].
[35] [*ATKINS/DE PAULA* 2006, pp. 364–71].
[36] [*CLAYDEN et al.* 2001, pp. 95–105].

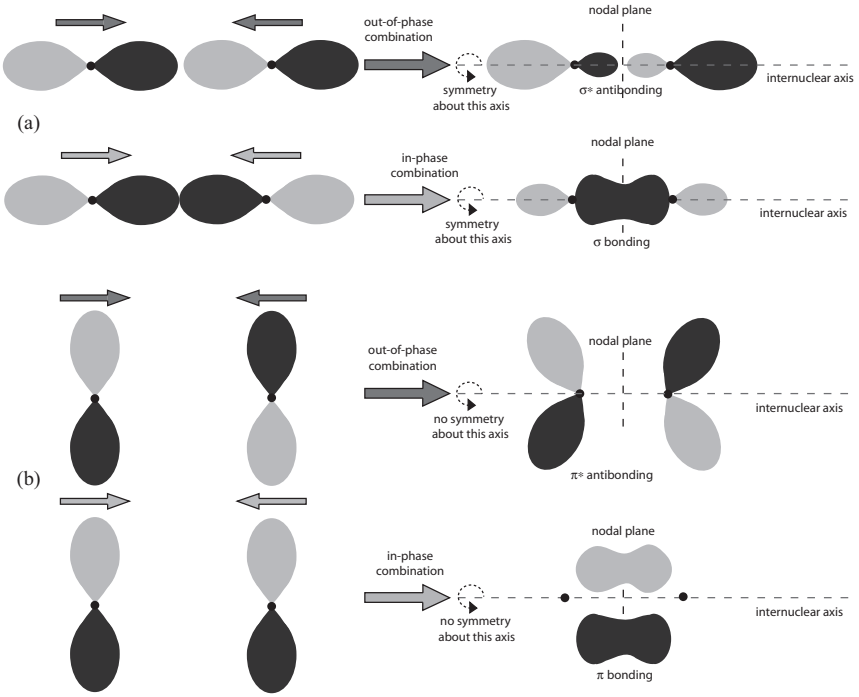

Fig. 6.7 (a) Constitution of antibonding (σ^*, first row) and bonding (σ, second row) combinations along the symmetrical axis (the two shades mean position probabilities of electrons being in different phases). (b) Constitution of antibonding (π^*, third row) and bonding (π, fourth row) combinations, which are orthogonal to the symmetrical axis (this means along the p_x or p_y axes).

When applied to the nitrogen molecule N_2 and taking z as the internuclear axis (connecting the two nuclei), we imagine each $2p_z$ orbital pointing towards a $2p_z$ orbital of the other atom. Here a cylindrical symmetry is established along the internuclear axis (the so-called σ bond). The other $2p$ orbitals ($2p_x$ and $2p_y$), instead, establish a side-by-side bond through the nodal plane (a so-called π bond) [Fig. 6.7]. Let us also consider oxygen, whose electron configuration is $1s^2 2s^2 2p_x^2 2p_y^1 2p_z^1$. When giving rise to water, the unpaired electrons in the O$2p$ orbitals can each be paired with an electron in H$1s$ orbital. Since the $2p_y$ and $2p_z$ orbitals lie at 90° to each other, the two constituted σ bonds also lie at 90° to each other [Fig. 6.8].

Carbon's covalent bond shows how the molecular bonds cannot be derived simply from quantum theory: It is again an instance of emergence. The ground state configuration of C is $1s^2 2s^2 2p_x^1 2p_y^1$, which suggests that carbon can only form two bonds and not four bonds, which is indeed what actually happens [Fig. 6.9]. The problem is circumvented by postulating promotion, i.e. the excitation of an electron (of level $2s$) to an orbital of higher energy ($2p_z$). In this case we have $1s^2 2s^1 2p_x^1 2p_y^1 2p_z^1$. Again quantum mechanics helps us to understand that here four hybrid (superposition) orbitals are established:

$$\psi_1 = s + p_x + p_y + p_z, \;\; \psi_2 = s - p_x - p_y + p_z, \tag{6.3a}$$

$$\psi_3 = s - p_x + p_y - p_z, \;\; \psi_4 = s + p_x - p_y - p_z. \tag{6.3b}$$

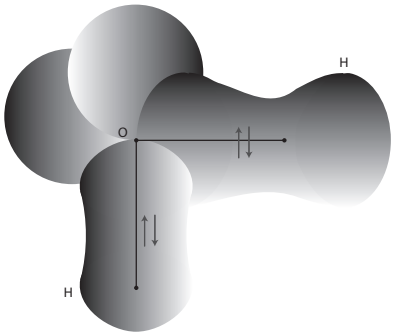

Fig. 6.8 Distributions of electrons in the bonds between oxygen and hydrogen atoms to constitute water molecules. The double arrows mean the two opposite directions of the magnetic polarization of the two electrons (spin up and down) occupying the same orbital.

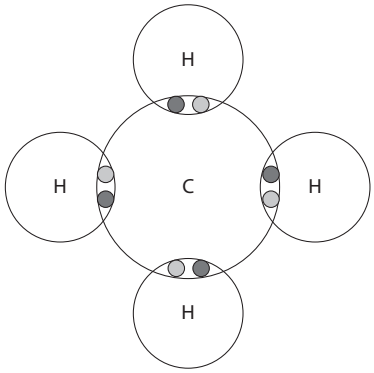

Fig. 6.9 Example of covalent bonds: Carbon, by sharing 4 pairs of electrons with 4 hydrogen atoms, gives rise to methane. The hydrogen atoms' electrons are shown in dark gray, and the carbon's electrons are shown in light gray.

Each hybrid orbital consists of a large lobe pointing in the direction of a corner of a regular tetrahedron (the triangular pyramid), the axes making angles of 109.47° [Fig. 6.10]. We see that global order and local interactions go complementarily together according to quantum-mechanical principles but with new and surprising results.

All heteronuclear diatomic molecules are polar, i.e. with a permanent electric dipole moment: This is the measure of the polarity of a system of electric charges (like electrons and protons) with a displacement vector pointing from the negative charge to the positive charge. Molecular symmetry is of the greatest importance for the issue of molecular polarity, even more than the problem of whether or not the atoms constituting the molecules belong to the same element. Indeed, ozone is homonuclear but polar, since the central O atom is different from the other two (having two bonds instead of one) and the dipole moments are not cancelled[37] [Fig. 6.11].

[37][*ATKINS/DE PAULA* 2006, pp. 621–35].

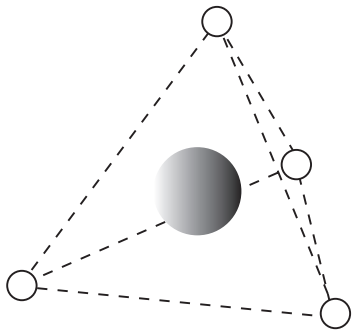

Fig. 6.10 The methane structure: The carbon atom occupies the center of the pyramidal structure, the four hydrogen atoms occupy the corners.

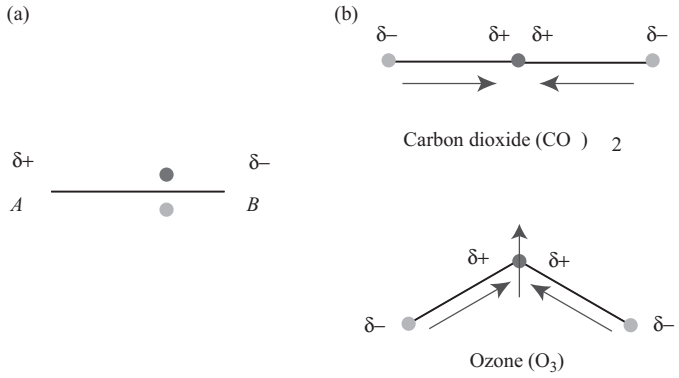

Fig. 6.11 (a) If an atom B is slightly more electronegative than an atom A, the atom B will attract the electron pair rather more than A does [Fig. 6.6]. This means that the B end of the bond has more than its fair share of electron density and so becomes slightly negative. At the same time, the A end becomes slightly positive. The symbol δ_- means slightly negative while δ_+ means slightly positive. (b) An example of a neutral molecule (carbon dioxide) and of a polar one (ozone).

Interaction between molecules, especially considering electric dipoles and ions, gives rise to a large variety of multipoles. When a hydrocarbon molecule (composed of hydrogen and carbon atoms, like benzene) is surrounded by water, the H_2O molecules form a clathrate cage isolating it from its fluid environment. Hydrocarbon molecule coalescence in water is entropy-favored.

Let us now consider some organic compounds.[38] Ethene is formed by adding two carbon atoms and 4 hydrogen atoms in the way shown in Fig. 6.12. The involved carbon atomic orbitals are $2s, 2p_x, 2p_y, 2p_z$. We can consider a hybrid form of $2s, 2p_x$ and $2p_y$ for each carbon atom to form the σ bond between the carbon atoms and the 4 σ bonds between each carbon atom and two hydrogen atoms. We see again the relevance of global quantum features for the symmetry properties of the molecules. However, the specificity of molecules (especially of organic molecules) relative to

[38] [*CLAYDEN et al.* 2001, pp. 105–110].

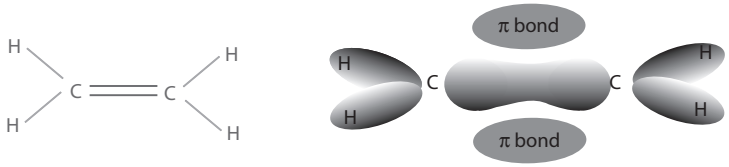

Fig. 6.12 Ethene. The two bonds with the hydrogen atom on the left form an angle of 117.8° (and the same for the two bonds on the right). All σ bonds are shown in gray scale, while the π bond in dark gray. The combination of the C–C σ bond and of the π bond constitutes the double connection between the two carbon atoms.

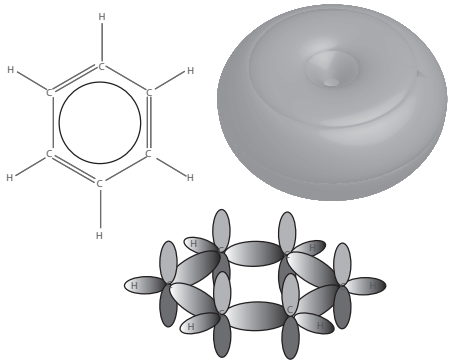

Fig. 6.13 The ring constituted by benzene. The σ bonds are shown in grayscale while the two phases of the π bonds are shown in pale purple and green (the colors refer to the color plate). Actually, all the π bonds (in green and pale red) constitute a single delocalized (quantum-mechanical) system shown in the circle inside the hexagon and as the tridimensional green structure on the top right. (This figure is reproduced in color in the color plate section.)

elementary quantum systems like electrons is that they are *integrated systems* in which both local interactions and long-ranging correlations are dynamically intertwined to give rise to a relative stable structure. Indeed, another interesting example is provided by benzene [Fig. 6.13].[39] The alternation between double and single bonds in the C-C ring is called conjugation. Actually, one of the ways in which chemical transformation is done is through rotation, which can be done only around a single bond. As we shall see, organisms also are integrated systems; however they integrate not only physical processes but also functions.

6.2.4 Chemical Reactions

Let us now consider some chemical reactions, paying particular attention to organic ones. All chemical reactions have some fundamental properties[40] in common:

- *Stoichiometry*: Stoichiometry rests upon the law of the conservation of mass, the law of definite proportions (i.e., the law of constant composition), and the law of multiple proportions. It is described by integral numbers counting the molecules that interact and form as a consequence

[39][*CLAYDEN et al.* 2001, pp. 154–55, 174–76, 549]. [40][*PALSSON* 2006, p. 15].

of the chemical reaction. Therefore, as previously remarked, it expresses a key property of information coding (*without* being information coding): Combinatorics. It is a property that is independent from environmental conditions (pressure, temperature, etc.).

- *Relative rates*: All reactions, and especially biological ones, are thermodynamic, and therefore the relative rates are dependent on environmental conditions like temperature and pressure.
- *Absolute rates*: The absolute rates of the reactions inside a cell are highly manipulable and indeed are continuously manipulated through enzymes and other catalysts. This is the continuous aspect of chemical reactions.

Any reaction starts with reactants and ends with products (the reverse transformation is also sometimes possible).

For starting a chemical reaction, we need to overcome a certain energy barrier (activation energy). Indeed, in general, molecules repel each other.[41] During a reaction the energy difference between starting reactants and products is very important. This is given by the difference in the Gibbs free energy $G = h - TS$, where h is the enthalpy and S the thermodynamic entropy [Appendix to Ch. 2]. This equation is the thermodynamic analogue of Eq. (2.23) in the form

$$h = G + TS, \tag{6.4}$$

showing an ordered and disordered part, and therefore confirming again that the quantities involved here are of thermodynamic and not informational type—even if information can be extracted from any system having a minimal degree of order [Subsec. 2.3.4], i.e. not being in a maximal-entropy state or a zero-entropy state (as occurs for quantum systems [Subsec. 2.3.3]). At a constant temperature and pressure the maximum additional work that is not due to expansion is given by the change in Gibbs energy: $dw_{\text{add,max}} = \Delta G$. The Gibbs free energy G always decreases when temperature is raised (at constant pressure and composition) and decreases most sharply when the entropy of the system is large; G always increases when the pressure of the system is increased (at constant temperature and composition) and is more sensitive to pressure when the volume of the system is large. At constant temperature and pressure, chemical reactions are spontaneous if

$$\Delta G = \Delta h - T\Delta S \leq 0. \tag{6.5}$$

If the energetic gap is not too big, then both the forward and backward reaction may occur. If ΔG is positive, reactants will be favored at equilibrium; while, on the contrary, if ΔG is negative, products are favored:

- In the case where ΔG is negative, we have an *exothermic reaction*, which releases heat (and energy), and therefore the products have less stored energy than the reactants. In other words, the enthalpy of the products is less than that of the reactants. In this way, we have available (free) energy for giving rise to endothermic reactions.
- In the case where ΔG is positive, we have an *endothermic reaction*, typical of bond making, where high-energy electrons are needed to form with other atoms chemical compounds, which absorbs heat (and energy), and therefore the products have more stored energy than the reactants. In other words, the enthalpy of the products is more than that of the reactants.

I stress that since the sign of the Gibbs free energy is negative or positive depending on the thermodynamic bond making or breaking, we see here a form of thermodynamic interconnection

[41] [*CLAYDEN et al.* 2001, pp. 113–22 and 307–31].

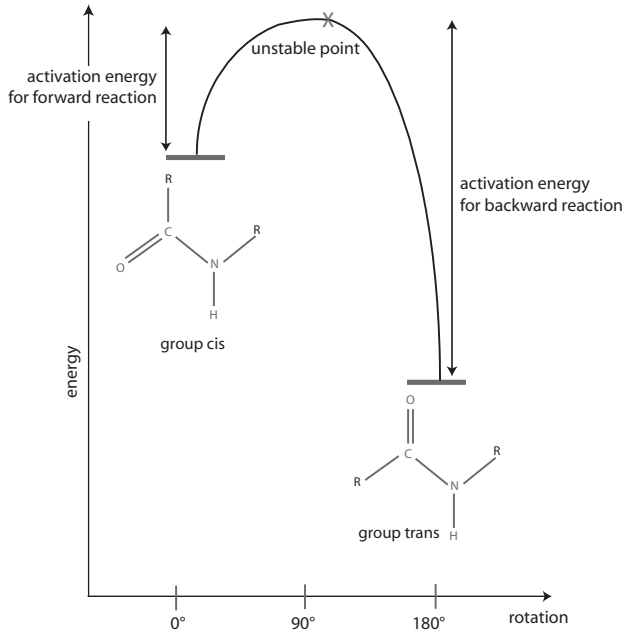

Fig. 6.14 The 180° rotation about the single bond CN (R indicates any here irrelevant, molecular component). The intermediate state at 90° rotation is less stable. The energy barriers are general aspects valid for any chemical reaction.

between molecules and compounds that is not of the informational type—like mutual information, the ordered part of Eq. (2.23).

Here, it is very important to distinguish between configuration and conformation of a molecule.[42] If a transformation from one molecule to another cannot be effected without breaking some bonds, we say that the two molecules have different *configurations*. However, the same chemical molecule can exist in a number of different *conformations*, given by rotating a part of the molecule about a certain axis [Subsec. 2.4.3]. This is typical for protein folding. For example, ethane can exist either in a staggered or in an eclipsed conformation. Ethane is similar to ethene [Fig. 6.12] but with four hydrogen atoms at each end. When looking at the molecule from one of its ends, if the four hydrogen atoms of this end can cover those of the other end, the conformation is eclipsed. Otherwise, it is staggered.

Let us now consider an example: The rotation about a C–N bond [Fig. 6.14].[43] The intermediate form is less stable here and therefore not commonly found in nature. However, the back transformation is still possible here. In order to increase the rate of a reaction (that depends on the activation energy) the chemistry of life makes use of enzymes,[44] which have the specific function to lower the energetic level of the intermediate state (therefore stabilizing it and keeping control of the reaction).

[42][*CLAYDEN et al.* 2001, pp. 384–5 and 448–52]. [43][*CLAYDEN et al.* 2001, pp. 305–7].
[44][*DIXON/WEBB* 1960]. See also [*OPARIN* 1957, pp. 363–73].

6.3 Self-Organization and Emergence of Complexity

In order to understand what a living system is and what contribution this understanding can give for deepening our analysis of cognitive processes, we must also grasp what the necessary conditions are, at the physical level, for the emergence of something such as life from the molecular and chemical level previously described. The wide area that stands between the physics and chemistry that we have so far considered on the one hand, and the biology on the other, is the study of complexity, where, in accordance with the fundamental thermodynamic laws, further new properties emerge that strongly contribute to some of the basic features of living organisms,[45] in particular to the arising of functions. To understand how this all comes together, an enlarged concept of dynamics [Subsec. 2.4.1] become very useful here. Let us first consider this problem in a qualitative way and more quantitatively in the next section.

6.3.1 Self-Organizing Systems

Complex systems constitute a very important part of our world [Subsec. 2.4.2]. Complex self-organizing systems instantiate a certain amount of order thanks to a local decrease of entropy: They are open systems downloading entropy into the environment.[46] Their emergence is made possible by the fact that quantum systems can locally increase their entropy, while on the other hand, the net balance of the entropy of our universe, following quantum-mechanical laws, can remain constant [Subsecs. 2.4.1–2.4.2]. I stress here that quantum mechanics has shown the necessity of introducing a theory of open systems [Subsec. 1.3.3].

Complex systems pertain to the class of self-organizing systems. The main difference is that complex systems show a hierarchy of levels of organization while this is not necessarily the case for any self-organizing system. It is also important here to carefully distinguish between self-organization and autopoiesis (or self-production). The former is only the ability to give rise to emergent structures thanks to feedback mechanisms and other features that we shall consider now. Self-production is the ability that *organisms* show to create and maintain their own structures. This point is relevant because, according to Ashby, there are two meanings of self-organization: (1) changing from unorganized to organized (self-organization, in my terminology), (2) changing from a less adaptive organization to a better one (self-production, in my terminology). I maintain with Ashby that this second form cannot be accomplished by a nonliving system.[47]

The most important feature of self-organized systems is that the interactions among the elements occur on the basis of purely local information, without any reference to the global patterns, which indeed *emerge* from these interactions,[48] as, for example, the waves produced by people standing up in a football stadium.[49] With *pattern* I understand a relationship among elements in a set such that, when an arrangement of a subset of these elements is specified, the probability of guessing the arrangement of the remainder generally increases with the size of the previous subset. In other words, the information that can be extracted from the remaining elements progressively diminishes (this means that a pattern is a signal although it does not necessarily represent codified information). Patterns are formed in open systems that are thermodynamically far from equilibrium, where there is a *competition* between various tendencies.[50] A stable state can "diffuse" here into an unstable state.[51] This is an important connection between quantum

[45] The magnificent structures that can arise in this way also satisfy our aesthetic sense [*BALL* 1999].
[46] Therefore, also for complex systems, the scientific study of their context is crucial [*MITCHELL* 2009, p. 13].
[47] [*ASHBY* 1952]. [48] [BONABEAU *et al.* 1999, pp. 8–14]. [49] [FARKAS *et al.* 2002].
[50] [TURING 1952]. [51] [BEN-JACOB/GARIK 1990, BEN-JACOB *et al.* 2000].

mechanics and self-organization: Indeed, also in quantum mechanics, every single system behaves in an unpredictible way but in the mean their behavior is subjected to laws, since these laws rule probabilities [Subsec. 1.2.8 and Sec. 1.3]. Recall also that in quantum mechanics locality is blind to globality [Subsec. 2.2.5]. We shall also meet this feature in distributed computation.

Self-organization has three basic ingredients:

- Multiple local interactions that allow for the incidence of random fluctuations or noise.[52] It is a typical bottom-up effect.
- Positive feedback, that is, amplification of input signals. Randomness is often crucial for finding new solutions, and fluctuations act as seeds from which structures may stem and grow. This can lead to the creation of spatial-temporal structures in an initially homogeneous medium.[53] In other words, to establish a certain amount of complex order, some form of disorder and symmetry-breaking is necessary. When far from equilibrium, self-organizing systems are forced to explore new spaces of possibility and new patterns (a feature that has an enormous importance for the way the genetic systems evolve). When there is amplification of patterns, positive feedback is wave-like and global.
- Negative feedback, that is, outputs damping input signals: It "calms down" or steers perturbations arising from below. Negative feedback counterbalances positive feedback and helps to stabilize collective patterns: It especially characterizes complex systems.[54]

Summing up, this means that self-organizing systems integrate oscillatory motion and structures in a new and tangled way.[55] Patterns play a fundamental role in life. For example, in the way bacterial colonies arise[56] [Fig. 6.15]. A self-organized system is often characterized by some further key properties:

- The possible coexistence of several stable states (attractors), a phenomenon called multistability. Which attractor the system will converge to depends on random initial events.
- The existence of bifurcations when some parameters are varied.

An *attractor* [Fig. 3.24] is a set of states (i.e., points in classical phase-space [Subsec. 1.2.5]) that is invariant under the system's dynamics. When a system is in a state located in the neighborhood of the attractor (i.e., in the area of the phase-space that is under the influence of the attractor, which is called its "basin") it asymptotically tends to reach the attractor during its time evolution. Therefore, a dynamical system may have multiple attractors, each with its own basin of attraction.

6.3.2 General Characteristics of Complexity

Four features are important here[57]:

[52] [BAK *et al.* 1988] [BEN-JACOB *et al.* 2000].

[53] Homogeneity is not at all a mark of high entropy, as is sometimes assumed. Quantum macroscopic systems like the Bose–Einstein condensate are very homogeneous due to the high correlation between the bosons but show very low entropy [Subsec. 2.3.3]. This confusion stems again from the tendency to connect entropy and heat exchange.

[54] [FREEMAN 2000a].

[55] [*SPENCER* 1860–62, pp. 294–5]. In classical mechanics motion and structure are not necessarily connected. One often speaks of the motion of a point (a unstructured system) and of static systems. We come back here to the issue of an expansion of the traditional concept of dynamics [Subsec. 2.4.1].

[56] [BEN-JACOB *et al.* 2000].

Fig. 6.15 Simulation of (the fractal development of) a colony of *Paenibacillus dendritiformis* bacteria in E. Ben-Jacob's laboratories. Adapted from http://www.popsci.com/scitech/gallery/2009-02/cannibal-bacteria-colonies.

(1) Complex systems present hierarchical structures[58] having (a) different levels of order and complexity, and (b) a relational web at each level of the structural hierarchy.
(2) They are top-down systems (upper-level variables and constraints influence or canalize the lower-level, efficient dynamics).
(3) They present recurrent basic structures at any level (called motifs in cellular networks).
(4) They show a certain plasticity and adaptive potentiality.

About Point (1), recall also that the units themselves of a structural whole, in our case a complex system, may be structured or exist as complex systems as well [Subsec. 2.2.5]: I have stated that this is a general principle of nature where there are structures. Here, a specification of a higher-level state determines a family of lower-level states, each of them able to give rise to the higher-level state. A higher-level state can influence a lower-level state (see Point (2)) in order to induce this to map to another lower-level state.

Also because of the soft assembly [Subsec. 6.1.5], hierarchical structures are modular [Sec. 3.6], that is, each level is partially shielded from the others. Since, however, modularity applies also to (relative) independent elements at the same level of a hierarchy, it is better here to use the term *information encapsulation*: Information encapsulation is the hiding of certain lower-level structures having informational value or content relative to a higher level of a hierarchy,[59] which represents a further complexification of the information accessibility principle [Subsec. 2.2.2]. A very common example of this is the way in which DNA is encapsulated (as we shall see, life indeed makes use of information). Modules and encapsulated unities have the specific advantage of representing discrete units. A clue to the reason why modules and information encapsulation are used by complex systems and evolve in biology can be found by looking at the way they are used in engineering.[60] Modules in engineering convey an advantage in situations where the design specifications change in time: New devices or software can easily be constructed from existing, well-tested modules. A nonmodular device, in which every component is optimally linked to every other component, is effectively frozen and cannot evolve or adapt to meet new optimization conditions. Similarly, modular biological networks may have an advantage over nonmodular networks in

[57][ELLIS 2005a]. [58][ELLIS 2004]. On hierarchies see [*SALTHE* 1985]. [59][*BOOCH et al.* 2007, pp. 50–8].

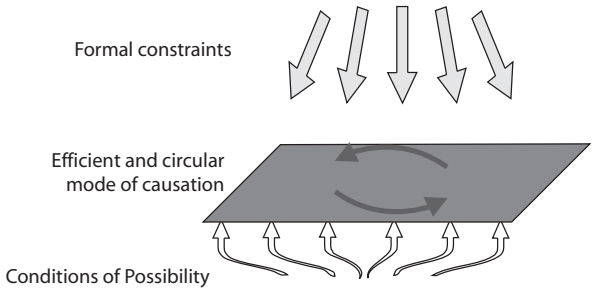

Fig. 6.16 Relations between different levels of a complex system.

real-life ecologies, which change over time: Modular networks can be readily reconfigured to adapt to new conditions. Furthermore, they are robust against impairment or the failing of components. The way in which they do this is by building degenerate structures and functional equivalence classes, as we shall see below.

Therefore, organisms show a complementarity between modularity (discontinuity) and connectedness (continuity),[61] which allows for the integration of different levels of organization (hierarchical organization) [Subsec. 2.4.4]. In this way, modularity and information encapsulation contribute to plasticity.[62] I would like to summarize these results by pointing out the general conditions that determine the way in which networks show hubs and modularize. According to web theory,[63] nonrandom networks show two basic principles: Continuous addition of new nodes and preferred linkages (in the mean, new nodes attach to nodes that are already best linked). This is the consequence, as already mentioned, of two more general principles that also rule quantum mechanics[64]: The continuous generation of variety (and therefore the intrinsic randomness of single events) and the establishment of preferred channels (channelization) due to the external conditions that generate regular patterns [Sec. 1.3].

As we shall see, it is here that we must find the root of the emerging *function* of a whole over the component parts. Indeed, as I have stressed [Subsec. 6.1.4], a function is neither fully dependent on a structure nor fully independent. It stems from a certain higher-level configuration able to canalize the operation of the parts in a proper way. Also, further elements are important here, and we shall deal with this problem extensively in the next chapters.

About Point (2), let me stress from the start how the relation of the different levels of a complex system should be understood. There cannot be a direct, efficient mode of causation on a certain level either from above, or from below.[65] The reason is very simple: Any level of reality is self-consistent relative to higher levels of organization and its specificity is irreducible to lower levels of organization, according to our previous examination of the issue of emergence [Subsec. 2.4.2 and Sec. 6.2]. For this reason, causal influence from below must be understood rather as *conditions of possibility* for further developments at the middle level [Fig. 6.16]. Causal effects from above, on the contrary, must be seen as *formal constraints* canalizing and restricting the space of possibilities of the middle level.[66] I wish to stress that both conditions of possibilities from below and formal

[60][ALON 2003]. [61][ULANOWICZ 2009b].
[62][*WEST-EBERHARD* 2003, pp. 12–13, 59–61, and 83–4] [ELLIS 2004].
[63][BARABÁSI/ALBERT 1999] [*BARRAT et al.* 2008, pp. 64–8].
[64]See also [BIANCONI/BARABÁSI 2001]. [65][AULETTA 2008d] [KIM 1999].
[66][*MITCHELL* 2009, pp. 41–4].

constraints from above cannot produce any effect at the middle level by themselves *without the dynamical contribution* of circular (that is, feedback–based) or efficient *causation at that level*.[67] For example, quantum-mechanical fluctuations may well have effects on molecules and chemical elements. However, without specific interconnections and interactions happening at this level, those fluctuations would be fully ineffective.

About Point (3), we shall see that self-replication of structures is fundamental for the transmission of life across generations.

About Point (4), according to Skyttner,[68] there is a suboptimization principle here: If each subsystem, regarded separately, is made to operate very efficiently (thus not being soft assembled [Subsec. 6.1.5]), the system as a whole will not operate with much efficiency. Moreover, in any complex decision network, the potential to act effectively is guaranteed by an adequate concatenation of information.

An important feature is also the following: Since order does not in itself guarantee complexity (a perfect correlated system like a quantum system shows order but no complexity), a system cannot increase its complexity by simply increasing the number of the interrelations between its components, but needs either to add a further dimension in its "space" or to enlarge the number of subsystems (by connecting itself to other systems). An example of the first strategy is represented by the secondary structure of proteins, in which a bidimensional complex is built by folding a linear sequence of unities by establishing hydrogen bonds among elements. An example of the second strategy is represented by the quaternary structure of proteins, a process through which two or more proteins join to give rise to more complex proteins.

It is possible that the class of complex systems is extensionally equivalent to that of biological systems. In other words, all natural complex systems are probably biological systems. However, complexity is not the only feature that characterizes the latter ones. Many other aspects are also important, as we shall see. In this sense, complexity is only a necessary (but not sufficient) condition of life, as the title of this chapter announces.

6.3.3 System Theory and Complexity

The problem of complexity can also be considered from a system-theory viewpoint [Subsecs. 2.4.4 and 3.2.2]. This is very important since we follow a system-biology approach. Three aspects are relevant here:

- System theory introduced into science the consideration of open systems.[69] Indeed, the steady state of a living organism is different from thermodynamic equilibrium [Subsec. 6.2.4] and the latter is not a necessary condition of the former. Stability is a property of the whole system and cannot be assigned to any part of it, because the presence of stability always implies some coordination of the actions between the parts.
- Given a negative feedback, a system's equilibrium state is invariant over a wide range of initial conditions. This is known as equifinality (convergence).[70] While in closed systems the final state is unequivocally determined by the initial conditions, in open systems the same final state (attractor [Fig. 3.24]) may be reached from different initial conditions and in different ways.

[67] This seems also to be the main tenet of the original proponent of downward causation processes [CAMPBELL 1974]. This helps also to overcome the biggest misunderstanding of ontological reductionism [Sec. 1.1]: Indeed, I know of no single example of a lower level of reality acting on more complex ones without additional dynamical processes at the latter levels.

[68] [*SKYTTNER* 2005, pp. 100–1]. [69] [VON BERTALANFFY 1955]. [70] See also [*MITCHELL* 2009, p. 48].

Convergence to some patterns is grounded in the fact that complex and chaotic systems can reach stable states from several points (the basin of attraction).[71]

- The amplification of fluctuations may give rise to bi- or multi-furcations along the system's time evolution, thus determining the possibility that it will tend toward different attractors according to the "chosen" branch. This is multifinality (divergence).

Unfortunately, equifinality led the German embryologist H. Driesch to support vitalism after observing this phenomenon in sea urchins, which can develop into the same mature form by starting from a complete ovum, from one half, or from the fusion product of two whole ova.[72]

6.4 Structure and Complexity

Having established the general conditions for complexity, let us move on to the interesting issue of whether or not there can be a quantitative measure for univocally distinguishing between more and less complex systems. In complexity an important role is played by mutual information, especially when organisms are involved [Subsec. 6.3.2]. However, as mentioned, mutual information alone [Subsec. 2.3.4 and especially Fig. 2.8] only guarantees order, not complexity. As we have seen, the initial disentanglement of quantum systems means an increase in entropy [Subsec. 2.2.4]. When we approach the level of molecular structure and the chemical elements, when more and more systems are added and classically connected, complexity grows [Sec. 6.2]. The growing complexity of the systems implies a further increase of the entropy but an even larger increase of the *maximal* entropy theoretically attainable by the system (since the number of different possible dispositions in which the compounding elements can be, grows much more than the elements' number). Therefore, the *gap* between the current entropy of the system (which is not maximal, due to some interrelation between the subsystems) and its possible maximal entropy also grows. This allows for further order and complexity [Subsec. 2.4.1], a mechanism that is widely used when new levels of complexity emerge at phylogenetic scale.

6.4.1 Stored and Structural Information

We have remarked in the previous section an important difference between self-organization and complexity: Self-organizing systems are only the result of thermodynamic principles ruling exchange of entropy in systems far from equilibrium. Complex systems are hierarchical systems showing information encapsulation. We have also mentioned that information codification plays an important role in organisms. A careful analysis of this point must be postponed: Here, I am not interested in the genesis of information codification in biological systems or in the analysis of how they deal with information codification. My more modest task is to show the relations between information and complexity.

We must find a measure of complexity that puts together two different aspects[73] [Subsec. 2.4.4]:

(1) The differentiation of the system in relative independent parts (in some cases, information encapsulation between different levels and modularity at the same level) and
(2) A global order (coherence of the whole).

[71] [VON BERTALANFFY 1962].

[72] [*DRIESCH* 1908]. Driesch's experiment also provided some of the first evidence for epigeny, that is, against a strict preformationist view (as well as against a mosaic view) of the organism [*VON UEXKÜLL* 1926, pp. 181 and 191–3].

[73] [AULETTA 2010].

Let us come back to the issue of stored information at an abstract level [Subsec. 2.3.4]. Let us consider again a system composed of two subsystems, say a and b (which could represent some string of information). Consider Fig. 2.9. For the sake of simplicity, both the systems a and b are assumed here to be in an initial state of maximal entropy $H(a)$ and $H(b)$ whose general form is (2.10), and $H(a,b)$ is the total entropy of a and b. Let us also consider an external environment with entropy $H(e)$, whose information our strings will store. According to Eq. (2.17), we can write

$$H(a|e) = H(a) - I(a:e) \quad \text{and} \quad H(b|e) = H(b) - I(b:e). \tag{6.6}$$

Analogously, we can write a conditional entropy of both the systems a and b on e, that is,

$$\begin{aligned} H(a,b|e) &= H(a,b) - I(a,b:e) \\ &= H(a|e) + H(b|e) - H(a:b|e), \end{aligned} \tag{6.7}$$

where an inspection of Fig. 6.17 can be very useful: $H(a,b|e)$ is the whole white region of $H(a)$ plus the whole white region of $H(b)$ plus the light gray region that they share. The quantity $I(a,b:e)$ represents the information shared by a and b on the one hand and e on the other and is given by:

$$I(a,b:e) = I(a:e) + I(b:e) - I(a:e:b), \tag{6.8}$$

where $I(a:e:b)$ (the central dark region in Fig. 6.17), in analogy with Eq. (2.17), is defined by

$$I(a:e:b) = I(a:e) - H(a:e|b), \tag{6.9}$$

and represents the overall information shared by a, b, and e. Note that the quantity $I(a:e:b)$ is symmetric, as expected by a true mutual information [Eq. (2.17)]. Indeed, we also have

$$I(a:e:b) = I(a:e) - H(a:e|b) \tag{6.10a}$$

$$= I(b:e) - H(b:e|a). \tag{6.10b}$$

Now, taking into account Eq. (6.7) and considering Eqs. (2.27)–(2.28), where I substitute $H(a,b|e)$ to H_M, we can write the stored information by strings a and b of some environmental information e as

$$\begin{aligned} I_s(a,b;e) &= I(a:e) + I(b:e) \\ &= H(a) - H(a|e) + H(b) - H(b|e) \\ &= H(a) + H(b) - H(a:b|e) - H(a,b|e), \end{aligned} \tag{6.11}$$

that shows some redundancy since $I(a:b:e)$ is considered two times. This is not bad, since we expect a certain redundancy when we store information.

Always with reference to Fig. 6.17, the structure that the three strings a,b,e can give rise to could be defined as the information that a shares with b, b with e, and e with a, i.e.

$$\mathcal{S}(a,b,e) = I_s(a,b;e) + I(a:b), \tag{6.12}$$

or, if we like to avoid any redundancy here,

$$\mathcal{S}(a,b,e) = I(a:e) + I(b:e) + I(a:b) - 2I(a:b:e). \tag{6.13}$$

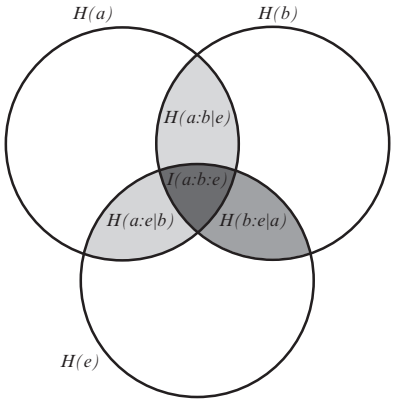

Fig. 6.17 The *stored information* $I_s(a, b; e)$ of the system $a + b$ relative to e: It is represented by the shaded regions [Eq. (6.11)]. *Structural information* of the system $a + b + e$: It is the whole (dark and light) gray region [Eq. (6.13)]. *Complexity* of the system $a + b + e$: It is the whole gray region minus the central dark region [Eq. (6.14)].

Obviously, other possible kinds of structures can also be defined, which is why this is only meant to be considered as an example. A more complex example involving five units (atoms) is represented by the tetrahedron shown in Fig. 6.10. The structures of matter (atoms, molecules, and so on) may be understood as order starting from qubits and ebits as the reservoir of information and order that is distributed in different ways in our universe[74] [Secs. 2.4 and 6.2]. When the number of interacting systems is greatly increased, we see a huge increase in their dependencies, but when there is a growth of complexity we see a proportional lowering of the overall mutual information.

6.4.2 Mathematical Measures of Complexity

We can use the previous equation for defining the complexity of the system $a + b + e$ as[75] [Fig. 6.17]

$$\begin{aligned} C(a, b, e) &= \mathcal{S}(a, b, e) - I(a : b : e) \\ &= [I(a : b) - I(a : b : e)] + [I(a : e) - I(a : b : e)] + [I(b : e) - I(a : b : e)] \\ &= H(a : e|b) + H(b : e|a) + H(a : b|e). \end{aligned} \tag{6.14}$$

This measure of complexity is therefore the whole structural information of the system $a + b + e$ minus their overall mutual information $I(a : b : e)$. It is interesting to note that both the structural information and the complexity between two sole elements, say a and b, collapse into their mutual information. This shows that the concepts of structural information and complexity require at least three elements [Subsec. 2.4.4]. Moreover, both complexity and stored information are concepts always relative to a certain environment or to a certain reference system. It is interesting to note that the difference between the maximal entropy attainable by a system and the entropy of its current state is due precisely to its structural order or, when there is codification, structural

[74] [FRAUTSCHI 1982, LAYZER 1977].

[75] See also [ADAMI/CERF 2000]. This can be considered to a certain extent to be a development of Kolmogorov's concept of complexity [KOLMOGOROV 1963].

information. Indeed, by again considering Fig. 6.17, it becomes immediately evident that by increasing the whole gray region, the whole surface given by the combination of the three circles goes down. Finally, we can understand that by lowering more and more the overall mutual information present in the structural information (6.13) we have a *natural transition to complex systems*.

The crucial point here is that mutual information, the quantity occurring several times in the second line of Eq. (6.14), *does not require* shared codified information but only that there is at least one local code in some of the involved systems and some overall combinatorics. Indeed, as Eq. (2.13) already demonstrates, it is solely concerned with the interdependencies among systems and their components. This allows that, together with thermodynamic interdependencies, complex informational-entropic dependencies are established that can then result in new, unexpected informational features, even if complex systems represent a weaker form of combinatorics than atomic and molecular combinatorics [Subsec. 6.2.1], since, in their case, we cannot speak of any number of elementary units. In spite of that, patterns can still be combined to give rise to other patterns as well as complex systems can give rise to other complex systems—this is evident when several organisms give rise to a whole ecosystem. As a matter of fact, only conditional entropies of a special kind are present in the last line of Eq. (6.14) which nevertheless allow for new forms of local information codification. Indeed, the case of organisms is such that there is always information nested somewhere. To understand this crucial point (that makes complexity far more interesting than molecular combinatorics), let us take a close look at the expressions of the form $H(a : e|b)$ occurring in the last line of Eq. (6.14). These quantities can be *equivalently* considered in two alternative ways:

- As the conditional entropy between the information shared by a and e, i.e. $I(a : e)$, on the one hand, and a third system b, on the other, that is, as $H(I(a : e)|b)$, or also
- As the information shared by a, on the one hand, and the conditional entropy $H(e|b)$ between e and b, on the other, that is, as $I(a : H(e|b))$.

It is easy to see that we have

$$H(I(a : e)|b) = I(a : e) - I(a : e : b), \tag{6.15a}$$

in accordance with Eq. (6.9). Instead, the expression

$$I(a : H(e|b)) = H(a) - H(a|H(e|b)) \tag{6.15b}$$

is more puzzling and difficult to understand. A look at Fig. 6.18(a) shows, however, that this corresponds to take the dark gray + middle gray region first, i.e. $H(e|b)$, and then the conditional entropy $H(a|H(e|b))$ (the whole light gray region). Finally, we subtract this from $H(a)$, which gives precisely the middle gray region. For expression (6.15a), we first take the middle gray + light gray region and then subtract the light gray region, as shown in Fig. 6.18(b). The relevance of this equivalence and of the relative expressions is that the latter show *both* conditional entropy and order as expressed by mutual information, and do so in accordance with an intuitive understanding of complexity as a mix of order and disorder.[76]

There are therefore several advantages of the measure (6.14) of complexity:

[76]I wish to recall that a similar but qualitative argument is presented in [*ULANOWICZ* 2009a, pp. 82–83]. My own derivation has been developed in full independence of any other known treatment.

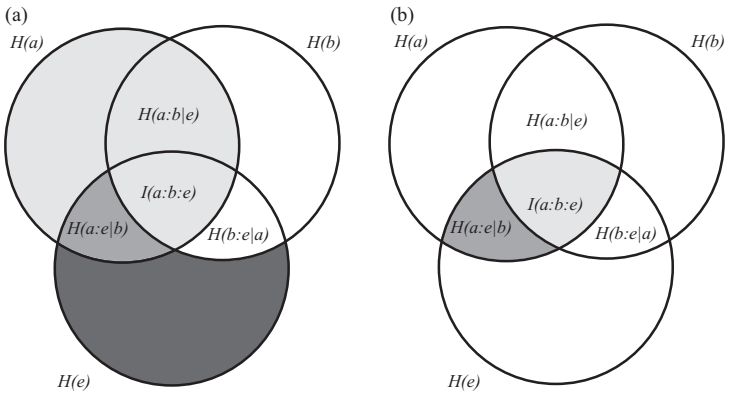

Fig. 6.18 The power of the formalism shown in Eq. (6.14): Any of the mixed terms on the last line can be considered as a combination of order and disorder. If we take the term $H(a : e|b)$, in particular, (a) shows that it can be interpreted as $I(a : H(e|b))$ while (b) shows that it can be equivalently interpreted as $H(I(a : e)|b)$. (This figure is reproduced in color in the color plate section.)

(1) It displays both coherence (order: expressed by the dependencies of the subsystems a, b, and e) and autonomy of the parts (expressed by the lack of mutual information $I(a : b : c)$). This means that the growth of complexity is not proportional to the growth of global order, as already anticipated.
(2) At a *global level*, it does not directly contain informational terms but only conditional-entropic ones. This means that complexity shows global patterns and functions that are independent of any explicit information coding, even in the case in which they stem from information codification.
(3) At a *local level* (the pairwise relations between the three subsystems) it does not require but allows for possible information exchange, sharing, and codification.

When the number of systems grows there are many more possibilities for partial interdependencies, which is the real quintessence of complexity. The quantities of the form $H(a : e|b)$ are very helpful for understanding such a process. Future research will show whether the formalism I am suggesting here can be useful to distinguish between several orders of complexity.

6.5 The Physical Nature of Complexity

6.5.1 Emergence of Complexity

Complexity represents, at a new level, the combination of the two aspects already found in quantum information [Subsecs. 2.2.4–2.2.5]: (a) Sudden points of crisis, which produce results that are somehow contingent, and (b) a wave-like emerging global behavior [Subsec. 6.3.2]. Let us now consider these aspects.

- The necessary physical condition of complexity is that the system is far from thermodynamic equilibrium[77] [Subsec. 6.3.1 and 6.3.3]. This allows for entropic fluxes from the system into the environment (and *vice versa*), in order that the system can locally decrease its entropy

[77][*PRIGOGINE* 1947] [*GLANSDORFF/PRIGOGINE* 1971][*NICOLIS/PRIGOGINE* 1977].

and build order (or, *vice versa*, to increase its entropy to break a too high homogeneity), a mechanism that is widely used by organisms in a controlled way. I have stated that order and complexity are two different concepts and that too much order diminishes the complexity of the system. Complexity consists in a symmetry-breaking such that the parts of a physical system become *spatially* (against the uniformity of a system in equilibrium) and *temporally* distinguished: In fact, the system "selects" one disposition among others in an unpredictable way depending on the specific conditions when the critical point is attained. This is a further application of the principle of universal selection [Subsec. 2.2.3] but with an important difference relative to quantum-mechanical or molecular systems: Here, the different selections cumulate so that the system shows a sort of protomemory and hence a quasihistorical dimension.[78]

- Therefore, when ascending the ladder of complexity in nature, the specific characters of quantum systems fail (here we no longer have pure random events), but the general characters of event-like aspects of nature remain [Subsec. 2.2.5]. In other words, phenomena like points of crisis, jumps, fluctuations, and the sudden breaking down of a system, are typical of many macroscopic systems and are manifestations of the principle of universal selection. We call this, often improperly, chance. One of the first scholars to have understood this point was E. Borel.[79] In particular, he was aware that the necessity of the global is not incompatible with the freedom of the local[80] [Subsec. 2.2.5].
- The main feature of complex order is that there are long-range *correlations* between distant parts of the system, as in quantum mechanics, where the whole cannot be reduced to the parts. We have also seen preferred linkages in networks [Subsec. 6.3.2]. As mentioned, these correlations are often not visible from lower levels of organization and the outputs are not fully determined by the inputs. Here, these correlations do not always depend on information codification: The mathematical equations describing the system are mostly nonlinear, while codified information needs to be linear (that is, built in an additive manner and reducible to a string of elementary unities, for instance as a sequence of zeros and ones in a columnar vector, like in quantum mechanics [Subsecs. 1.2.2 and 2.2.2]). Now, we find that in any possible condition whether with linear or nonlinear processes, whether at microscopic or macroscopic scale, nature always builds and rebuilds different forms of correlations, patterns, and structures. Thus, we are allowed to assume, together with the selection principle [Subsec. 2.2.3], a second very general principle of nature, a principle of universal correlation:

  Nature always establishes interdependencies and, when some are destroyed, many others are created anew.

  As a matter of fact, nature shows an overall tendency to disorder expressed in entropy growing (according to the second law of thermodynamics). Such a tendency grounds the irreversibility of the occurring dynamic processes. Therefore, the fact that we also observe continuous manifestations of order at all levels of our universe provokes our astonishment. How is it possible? Also, in quantum mechanics irreversible processes occur, paradigmatically when we perform a measurement. However, we may recall that a fundamental theorem of quantum mechanics tells us that for any irreversible process, we may find a larger context to embed it that is reversible as a whole [Subsec. 2.4.1]. Therefore, if the universe as a whole obeys quantum-mechanical laws (as I am inclined to think), we would expect that the tendency to disorder is continuously balanced

[78] [*ULANOWICZ* 2009a, pp. 68–69]. [79] [*BOREL* 1920].

[80] [*BOREL* 1920, p. 290]. He also assumed [p. 292] small deviations from a given law, showing an understanding of the general character of laws [Subsecs. 2.4.1–2.4.2].

by a tendency to order, so that the net result is precisely zero, as shown in Fig. 2.10. The conservation of both order and disorder in our universe is likely a far more general principle than the conservation of physical quantities like mass, energy, momentum. There are situations in which the conservation of energy can be violated (at least in very short times) as in the creation of virtual particles in a vacuum field. Obviously, the tendency to disorder is spontaneous and in this sense more fundamental. Indeed, according to statistical mechanics the possible disordered states of a system are much more than the ordered ones (for instance, there are likely infinite ways to break a cup, but only few to build it). However, my hypothesis is that every time such a tendency manifests itself, a *compensatory* tendency to order will also be produced elsewhere to preserve this net balance of our universe. This second tendency can be said to be less fundamental and not spontaneous, that is, forced by the first one. A consequence of these considerations is that the tendency to order will also be displayed through growing levels of complexification, without resorting to any form of vitalism. This implies that the irreversibility that we currently experience is due not only to disruption processes but much more to a certain directionality in the sense of growing complexification precisely thanks to the compensatory tendency to order even when (and because) entropy is conserved. These two tendencies to higher levels of local disorder and complex organization is indeed what we experience in our universe to be the rule from its early more uniform state. When a more complex reality emerges from a less complex one [Subsec. 2.4.2], the immediate result is that we have a larger number of components or factors that are integrated in the new system. Clear evidence for that is the higher number of elements and interrelations in a biomolecule like a protein relative to any abiotic molecule like water. In this way, getting the more complex from the less complex implies a growing number of constraints and therefore also a lower probability to find the ordered or stable configuration; for instance, the number of stable proteins is an incredibly small subset of all theoretically possible proteins. In other words, the number of possible disordered (and unstable) configurations grows exponentially with complexification. This implies that all biomolecules, for instance, share much more constraints among them than all abiotic ones do. This explains that along with the process of growing complexification there is also a sort of growing canalization (and therefore also a growth of control mechanisms), where this statement should not be taken in the sense that we have less and less variety, but precisely in the sense that we have more and more shared constraints.

- Another aspect of emerging complexity is multifunctionality: Different functions may be produced by the same elements and couplings. Multifunctionality is typical of synergy, the emergent property that results when the behavior of the whole systems is unpredicted by the patterns of their parts taken separately. The central feature here is the lack of one-to-one relationship between self-organized behavior and the structures that realize it [Subsec. 6.1.2], which is also the main difference between an organism and a mechanical object.
- A regime characterized by symmetry-breaking, multiple choice, and correlations is necessarily a dissipative structure.[81] As we shall see, complex systems can also be considered systems that are on the edge of chaos, and actually they share several properties with chaotic systems (systems whose trajectory is unpredictable).[82] For instance, both types of system show sensibility to the initial conditions. The sensibility of a system to past conditions is called hysteresis. This means that a complex system is such that its behavior depends crucially on its details.[83] Then, precisely as is the case for quantum systems, we can only have probabilistic predictions but not predictions

[81] [NICOLIS 1989]. [82] [*OTT* 1993]. [83] [PARISI 2006].

about single behaviors. That is, we need to deal with *classes* of behaviors and not with single behaviors[84] [Subsec. 1.3.1].

6.5.2 An Example of Nonlinearity

As I have mentioned, a nonlinear equation is such that the output is not given by the sum of the inputs. In general, the outputs themselves become parts of the inputs, so that a very easy form of nonlinear mapping is given by:

$$x_{n+1} = cx_n(1 - x_n), \tag{6.16}$$

where c is a constant. This map means that the n step that is itself the result of a previous $n-1$ step also enters as an element of the $n+1$ step. A very interesting example is represented by the so-called generalized baker's map,[85] through which an initial square of unit length $[0,1] \times [0,1]$ (whose x and y axes go from 0 to 1) is transformed according to

$$x_{n+1} = \begin{cases} c_a x_n, & \text{if } y_n < \alpha, \\ (1 - c_b) + c_b x_n, & \text{if } y_n > \alpha, \end{cases} \tag{6.17a}$$

$$y_{n+1} = \begin{cases} y_n/\alpha, & \text{if } y_n < \alpha, \\ (y_n - \alpha)/\beta, & \text{if } y_n > \alpha, \end{cases} \tag{6.17b}$$

where $\beta = 1 - \alpha$ and $c_a + c_b \leq 1$. The effect of the map is shown in Fig. 6.19. It is interesting to remark that after only two or five steps, as shown in Fig. 6.20, we already have the formation of strips following a certain organized pattern. In repeating this process over and over, we can obtain interesting results by starting with a simple bidimensional map.

6.5.3 Some Examples of Self-Organization and Complex Systems

One of the most studied examples of self-organization is constituted by the formation of Bénard cells.[86] Consider a thin layer of liquid between two large parallel plates [Fig. 6.21]. If the system is in equilibrium, with the liquid and the two plates at the same temperature, and the liquid is motionless, then the properties of the system are homogeneous. Suppose now that the upper plate is heated slowly. The heat will pass from the upper plate to the liquid and will be transferred through the liquid to its bottom layer by the process of thermal conduction. In thermal conduction there is no bulk motion of the liquid but rather a greater thermal motion of the molecules that causes the transfer of heat from the warmer layers to adjacent cooler layers. However, as the temperature of the upper layer is increased, a stage is reached (critical temperature) where the liquid overcomes its viscosity (the internal friction which opposes movement) and begins to undergo bulk motion. This results in a transport of heat by convection currents. The currents are not random but rather they lead to the formation of patterns, and often one first sees small convection cells (called Bénard cells).

Let us consider an example of emerging self-organization in chemistry. Two types of molecules (the reactants) can combine in order to produce two other types of molecules (the products).[87] The transformation has a constant relative rate (given by a certain pressure and temperature, considered

[84] [PARISI 1999]. [85] [*OTT* 1993, pp. 75–8].

[86] [NICOLIS 1989, pp. 318–19] [*BERGÉ et al.* 1984, pp. 83–91].

[87] The reader interested in self-organization and complexity within chemical reactions in biology can read the very good but technical textbook of Murray [*MURRAY* 1989, pp. 109–30] or [*BABLOYANTZ* 1986].

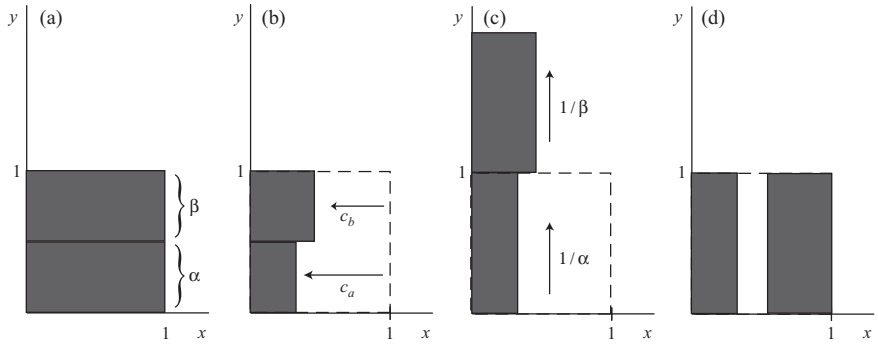

Fig. 6.19 The map shown in Eq. (6.17a). (a) We first divide the unit square into two parts, the top part, with $y > \alpha$, and the bottom part, with $y < \alpha$.

(b) Then, we compress the two pieces along the horizontal dimension (the upper piece by a factor c_b, as shown in the second line of Eq. (6.17a), the lower one by a factor c_a, as shown in the first line of Eq. (6.17a)).

(c) Next, we vertically stretch the upper piece by a factor $1/\beta$ and the lower piece by a factor $1/\alpha$ (in such a way that here both have unit length, as shown by the two lines of Eq. (6.17b)).

(d) Finally, we take the upper piece and we place it back in the unit square on the right side of the lower square, with its vertical edge coincident with the right vertical edge of the square. In this way we obtain two strips.

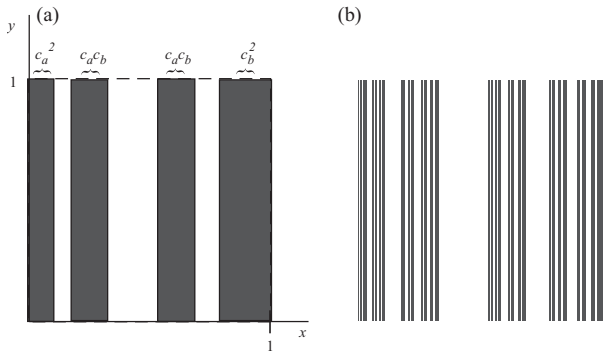

Fig. 6.20 The generalized baker's map after two steps (a) and after 5 steps (b): It is easy to verify that the number of strips is equal to 2^n for the n^{th} step.

here as constant). Though all reactants should in principle disappear in the reaction, it is never completely so, because the inverse transformation sometimes occurs [Subsec. 6.2.4]. Normally, we attain the chemical equilibrium when a fixed ratio is attained (it is called the equilibrium constant). At equilibrium, both transformations occur with exactly the same velocity (the so-called detailed balance). But if the system is open, then one can obtain that the concentrations of some stuff grows, and this is a stationary nonequilibrium state. In such a state, a mechanism can develop that is known as *autocatalysis* (the self-reproduction of organisms is a form of autocatalysis): For example, free radicals produce more free radicals, a behavior which can occur even if the system has no spatial inhomogeneities. Also autocatalysis is a typical nonlinear effect.

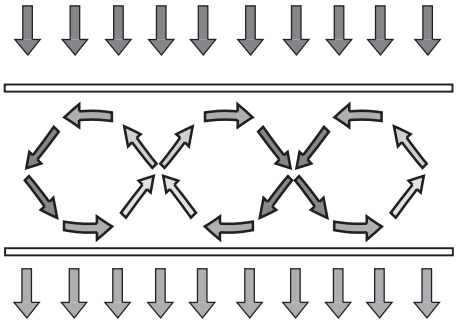

Fig. 6.21 Bénard cells between two metal layers. Each one runs either clockwise or counterclockwise (this depends on the long-ranging correlations). They assume a hexagonal form.

Fig. 6.22 Chemical waves produced during the Belousov–Zhabotinsky reaction. Adapted from http://www.flickr.com/photos/nonlin/4013035510/lightbox/.

One of the most interesting oscillatory behaviors (which has become the paradigm of oscillatory behavior as such) is the Belousov–Zhabotinsky (BZ) reaction [see the Appendix below].[88] The basic mechanism consists in the oxidation of malonic acid, in an acid medium, by bromate ions (BrO_3^-), and catalyzed by cerium, which has two states (Ce^{3+} and Ce^{4+}). Periodic oscillations are observed in the cerium ions. With other metal ion catalysts and appropriate dyes (for instance iron Fe^{2+} and Fe^{3+}) and phenanthroline, the regular periodic color shows the behavior of a chemical clock (alternance of blue and red population). Note that such an oscillation is completely different from that of a pendulum: This, if enhanced, shows a larger amplitude and a larger period (because of the invariance under time reversal), while a BZ reaction resets itself, as seen before, by showing an asymptotic stability, which is essential for *irreversibility* and *reproducibility* (a first manifestation of protomemory, as I have said) of events. Only long-range correlations guarantee that no destructive interference occurs, wiping the oscillatory behavior out. If the chemical system is not uniform, then regular spatial-temporal patterns are observed in the form of propagating wave fronts [Fig. 6.22].

[88][BELOUSOV 1951] [ZHABOTINSKII 1964, ZAIKIN/ZHABOTINSKII 1970]. See also [*TYSON* 1976] [*BERGÉ*] *et al.* 1984, pp. 91–7].

While the Bénard cells in physics show only a loss of transitional invariance, chemical reactions in the form of the BZ show chirality (when an object, for instance, shows preferred direction of rotation or of building helicoidal structures).

Many of these features, as mentioned, are also conserved in biological systems. Organisms, for instance, show a sensibility to initial conditions that is unknown to mechanical systems, in which the memory of the systems is represented by its current state. For this reason, a living being is a sort of a historical structure. Moreover, living beings at the cellular level show inhomogeneities that are self-reproducing.[89] In the formation of multicellular organizations we have a synthesis and emission of specific chemical substances by some cells, which is periodic, followed by the oriented movement of the other cells toward the regions of maximal concentration (this movement affected by the concentration of a substance is called chemotaxis), with the result that spatial-temporal wave patterns occur, and there is finally an amplification of the signal (in a feedback loop).

Life also shows chirality and violates symmetry: All proteins are made of L-amino acids (L stands for levogyre) and genetic material is made of D-sugars (D stands for dextrogyre).[90] It can be shown that if we have an open system, the equilibrium between the concentrations of L-molecules and D-molecules becomes unstable and the system spontaneously evolves into a state of broken symmetry. Random fluctuations cannot be dominant if the passage through the critical value happens slowly.

For each complex system there can be different optimal or stable regimes according to certain variables. This is again particularly relevant for life. For instance, a horse must pass after a certain threshold (given by oxygen consumption and other factors) from walk to trot and from trot to gallop[91] [Fig. 6.23]. These regimes are given by the task at hand. Also the physical structure plays a role. It can be observed that the legs swing at a pendulum rate. Longer legs have a longer period and for this reason a bigger animal does not go so fast relatively to a small one. Another important feature characterizing complex systems is relative coordination (with relative coordination it is not necessary that the phase between two rhythmic patterns is kept constant). For instance, two hands that tip together switch from being in phase to being in antiphase. This can be explained by taking into account the difference in frequency between the two patterns, which will influence the stable states or regimes.

6.5.4 Summary

In self-organizing processes, we initially have random starting patterns. However, in order to reach some stable order, one of them will rapidly dominate, so that we have here a true selection principle at a physical or chemical level. As in quantum mechanics, we have to distinguish here between global and local: We have competition at the microscopic, local, level among all different possible patterns, and cooperation at the global level. I have stressed that correlations are continuously established in nature at all scales and levels of complexity.

Complex order emerges as a compromise between positive and negative feedback, that is, at a chemical level, between the nonlinear chemical-like process, which through fluctuations sends innovating signals continuously but incoherently into the system (the system would finally be dominated by noise, i.e. becoming unable to reach the bifurcation points), and the processes that capture, relay, and stabilize them (and which on the other hand tends to full homogeneity). These systems are characterized therefore by circular forms of causality.[92]

[89][NICOLIS 1989, pp. 325–30]. [90][KONDEPUDI 1988]. [91][*KELSO* 1995, pp. 69–158].
[92][*KELSO* 1995, pp. 1–67]. See also [*HAKEN* 1976].

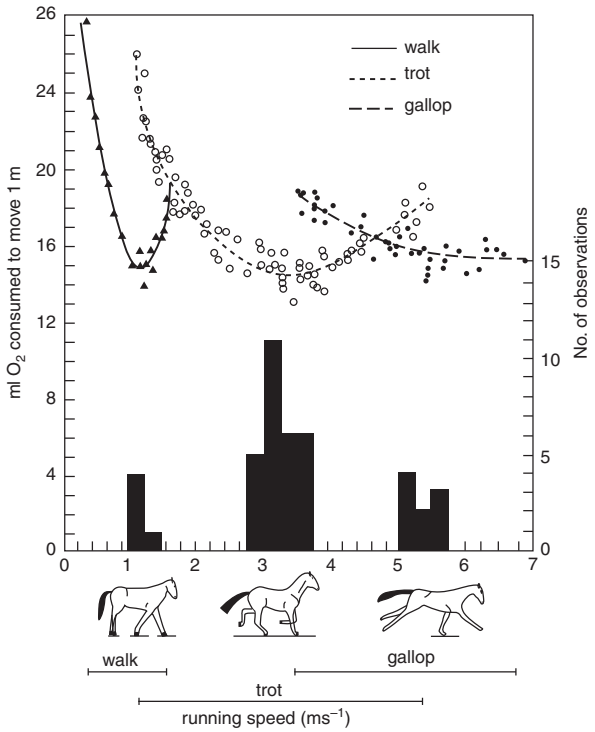

Fig. 6.23 Oxygen consumption per meter moved and preferred speeds. It is interesting to note that it can be lower at a gallop than at a walk. Adapted from [*KELSO* 1995, p. 71].

The relevant degrees of freedom are constituted here by collective variables that can be compressed into a few *order parameters*. These order parameters rule the specific configuration or conformation of a system. Parameters that lead the system to different patterns are the *control parameters*, that are quite unspecific in nature. Therefore, the knowledge of the system reduces to three features: Control parameters (boundary conditions), interacting elements, and emerging patterns (as expressed by order parameters).

6.6 Biology and Complexity

Let me first give a general account of the problem and then enter into some technical details.[93] In life there is a compromise between conflicting constraints. Life must emerge (for reasons that will be clarified below) as a whole, as a system. This shows that natural selection is indeed only one factor concurring to produce order: The other is self-organization. S. Kauffman has pointed out that living systems are sufficiently large chemical systems. When connecting unities, once the ratio of threads to unities oversteps the 0.5 value, most of the unities become connected; when it passes the value

[93] [*KAUFFMAN* 1993, pp. 287–404]. A good and easy account of this matter can be found in [*KAUFFMAN* 1995].

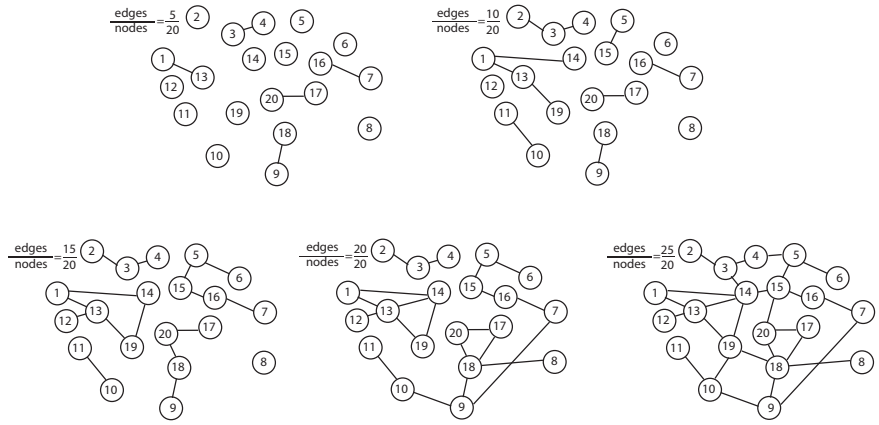

Fig. 6.24 Kauffman's connected web. Twenty nodes are connected by an increasing number of threads (edges). For a large number of nodes, when the ratio of threads to nodes passes the threshold of 0.5, most elements become connected; when it passes the threshold 1.0, a single giant network arises. There are two fundamental variables: How many inputs (K) each node receives and what the control rules are. As K augments, the length of the networks' state cycles increases much more. Adapted from [*KAUFFMAN* 1995, p. 55].

1.0, a single giant network will mostly arise[94] [Fig. 6.24]. In this way, autocatalysis [Subsec. 6.5.3] is almost inevitable and a true metabolism could arise here from an early protometabolic stage in which autocatalytic processes were established and combined.[95] The most important condition is that molecules reach a critical diversity (because, in this case, the ratio of reactions or threads to molecules or nodes becomes even higher). Self-reproduction may occur when the total system has reached such a high level of complexity that an ordered division into two daughter systems may occur. Alternatively, we may have a break in the system (a disordered breakdown of the system into several parts and not into daughter systems). An ordered division can more easily occur if the parental system replicates some structure in itself. We have already seen the tendency of complex systems to replicate structures [Subsec. 6.3.2]. However, to be able to split orderly, a system must also show *few* replicated configurations in itself, which in turn demands a hierarchical organization.

For an autocatalytic set to be ordered (and to represent a transition to a true metabolism) it must exhibit some homeostasis [96] [Fig. 6.25]: The tendency to stabilize some relevant internal order parameters against environmental fluctuations, allowing for survival in very different environments. In order to ensure its homeostasis, which requires the capacity to find and assimilate free energy and to discharge entropy into the environment [Subsec. 6.2.4], several work cycles are required.[97] In other words, life needs to be some form of agency, where, following Kauffman, I assume that the necessary condition for agency is the ability of an autocatalytic system to produce and reproduce itself by performing several thermodynamic work cycles, that is, cycles through which, according to thermodynamic laws for open systems, it is able to perform work and come back to its initial state. Therefore, an agent links exergonic (spontaneous chemical) reactions and endergonic (driven out

[94][*BARRAT et al.* 2008, pp. 5–8]. [95][WALDROP 1990]. [96][*CANNON* 1932].
[97][*KAUFFMAN* 2000].

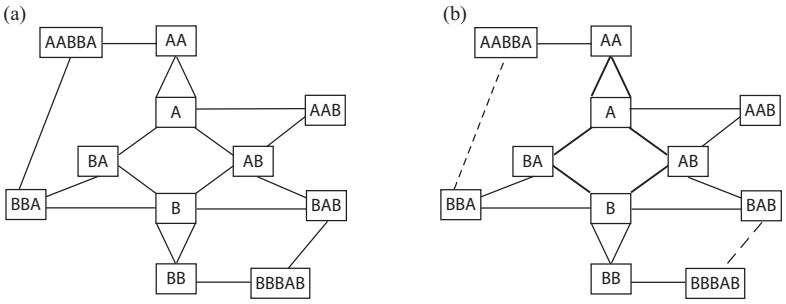

Fig. 6.25 (a) The schema shows a network of chemical reactions in which smaller molecules (A and B) are combined in larger molecules (AA, AB, etc.) which give rise to still larger molecules (BAB, BBA, BABB, etc.). Simultaneously, these larger molecules can be broken up again into their constituents.
(b) Here, there are some catalysts in order to speed up some reactions. The result is the selection of a pattern of heavy lines indicating a catalyzed subgraph. Inspired by [*KAUFFMAN* 1995, pp. 59–61].

of equilibrium by adding free energy) reactions thanks to some form of control. This links biology with cognition.

6.7 Concluding Remarks

The main results of this chapter can be summarized as follows:

- Information acquisition is not a sufficient paradigm for dealing with biological systems; the concept of information control and therefore of function (and complexity) must be introduced.
- Chemical bonds and structures do not depend on information codification but are ruled by entropic processes.
- Self-organizing systems arise thanks to amplification and feedback effects that are able to establish new patterns and global interdependencies given certain local events. The crucial point here is the exchange of entropy with the environment.
- Complex systems are a particular class of self-organizing system characterized by hierarchy and information encapsulation, a top-down causal process, recurrent structures, and adaptability.
- Hierarchical complex systems allow for functions to emerge.
- Mathematically, complexity can be defined as the sum of the joint conditional entropies of the parts of the system on any single other part in abstraction from whole-system coherence.
- Biological systems are complex systems. Organisms, in particular, can arise through autocatalytic chain reactions that arise spontaneously when a system grows in complexity.
- Self-replication in complex systems is the physical basis of self-replication in life.

Biological systems have two main roots: One is to be found in information codification, the other in complexity. In the next chapter we shall consider how organisms work by combining these two aspects. In Ch. 8 we shall finally learn the principles that rule such a combination.

For the development of the following examination the principle we have formulated above is extremely important. There are indeed good reasons for assuming that correlations are established and formed again at any level of complexity, so that even if some are broken, other ones are created anew.

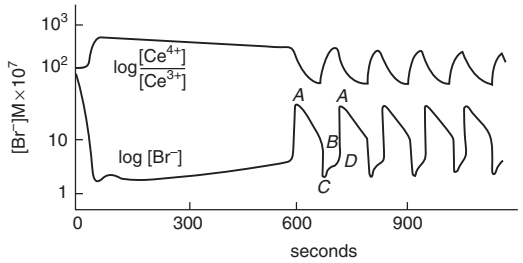

Fig. 6.26 Experimentally measured periodic limit cycle of temporal variation in the concentration in the ratio of the cerium metal ion concentration $[Ce^{4+}]/[Ce^{3+}]$ and the bromide ion concentration $[Br^-]$ in the BZ reaction. Adapted from [*MURRAY* 1989, p. 180].

Appendix: The Belousov–Zhabotinsky Reaction

Biochemical reactions mostly involve proteins called enzymes which are excellent catalysts that react selectively on compounds called substrates (for example, hemoglobin in red blood cells and oxygen). We have an *autocatalysis* when a chemical is involved in its own production, for example[98]

$$A + X \frac{k_1}{k_{-1}} 2X, \tag{6.18}$$

where $\rightleftharpoons$ denotes that the reaction is reversible and the k_j's are rate constants. The law of mass action tells us that the rate of reaction is proportional to the product of the concentration of the reactants. If A is maintained at a constant concentration a, the law of mass action gives the rate of reaction as

$$\frac{dx}{dt} = k_1 a x - k_{-1} x^2, \tag{6.19}$$

where x is the concentration of X and, as $t \to \infty$, $x(t)$ tends to a final nonzero steady state $x_S = k_1 a / k_{-1}$. This reaction exhibits a strong feedback with the "product" inhibiting the reaction rate. But if the reaction is defined by

$$A + X \rightleftharpoons k_1 k_{-1} 2X, \quad B + X \to k_2 C, \tag{6.20}$$

then X is used up for producing C and the system can be shown to exhibit a simple bifurcation, with the two steady states

$$x_S = \frac{1}{k_{-1}} \left(k_1 a - k_2 b \right), \tag{6.21}$$

according to the sign of the parameter $k_1 a - k_2 b$.

The basic mechanism of the Belousov–Zhabotinsky reaction consists of the oxidation of malonic acid, in an acid medium, by bromate ions, BrO_3^-, and catalyzed by cerium, which has two states, Ce^{3+} and Ce^{4+}.[99] Periodic oscillations are observed in cerium ions. Basically, the reaction can be separated into two parts, I and II, and the concentration of bromide ($[Br^-]$) determines the one dominant at any time [Fig. 6.26]. When the concentration $[Br^-]$ is high, I is dominant and during

[98] [*MURRAY* 1989, pp. 109–30]. [99] [*MURRAY* 1989, pp. 179–98, 702–12].

this stage Br^- is consumed: That is, we move firstly along the line AB, and then, once we have overcome the critical point B, we move quickly to the low level C [see again Fig. 6.26]. At this point, II takes over: During this stage the Ce^{3+}, which in part I was dominant, is now changed to Ce^{4+}, which reacts to produce Br^- again, so that its concentration increases starting Phase I again. All can be reduced to five key reactions, which can be written as

$$A + Y \overset{k_1}{\rightarrow} X + P,\ X + Y \overset{k_2}{\rightarrow} 2P, \tag{6.22a}$$

$$A + X \overset{k_3}{\rightarrow} 2X + 2Z,\ 2X \overset{k_4}{\rightarrow} A + P,\ Z \overset{k_5}{\rightarrow} fY, \tag{6.22b}$$

where the chemical elements are

$$X = HBrO_2,\ \ Y = Br^-,\ \ Z = Ce^{4+},\ \ A = BrO_3^-,\ \ P = HOBr, \tag{6.23}$$

the rate constants k_j's ($j = 1, \ldots, 5$) are known, and f is a stoichiometric factor (usually = 0.5) [Subsec. 6.2.2].

Reactions (6.22a) are roughly equivalent to process I, and reactions (6.22b) to process II. Using the law of mass action, we write the following equations for the involved chemicals' concentration

$$\frac{dx}{dt} = k_1 ay - k_2 xy + k_3 ax - k_4 x^2, \tag{6.24a}$$

$$\frac{dy}{dt} = -k_1 ay - k_2 xy + f k_5 z, \tag{6.24b}$$

$$\frac{dz}{dt} = 2k_3 ax - k_5 z, \tag{6.24c}$$

where each lower case represents the concentration of the related chemical written in upper case. If $f = 0.5$ and k_5 is very large, the last reaction in (6.22a) is very fast so that the third and the fifth reactions in (6.22a) collapse in a single reaction

$$A + X \rightarrow k_3 2X + Y, \tag{6.25}$$

which describes a bimolecular system that cannot oscillate. By a suitable nondimensionalization

$$\epsilon = \frac{k_5}{k_3 a}, \quad q = \frac{k_1 k_4}{k_2 k_3}, \quad \delta = \frac{k_4 k_5}{k_2 k_3 a}, \tag{6.26}$$

we could write Eqs. (6.24a) as follows

$$\epsilon \frac{dx}{dt} = qy - xy + x(1 - x), \tag{6.27a}$$

$$\delta \frac{dy}{dt} = -qy - xy + 2fz, \tag{6.27b}$$

$$\frac{dz}{dt} = x - z. \tag{6.27c}$$

To find out the nonnegative steady states (x_s, y_s, z_s), first we set the lhs of Eqs. (6.27a)

$$z_s = x_s, \tag{6.28a}$$

$$y_s = \frac{2fx_s}{q + x_s}, \tag{6.28b}$$

$$2x_s = (1 - 2f - q) + [(1 - 2f - q)^2 + 4q(1 + 2f)]^{\frac{1}{2}}. \tag{6.28c}$$

The other nonzero steady state is negative. Let us first write Eqs. (6.27a) in the vectorial form:

$$\frac{d}{dt}|r\rangle = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \epsilon^{-1}(qy - xy + x(1 - x)) \\ \delta^{-1}(-qy - xy + 2fz) \\ x - z \end{pmatrix} \tag{6.29}$$

Let us now introduce the stability matrix $\hat{S}$ with eigenvalues λ:

$$\hat{S} = \begin{bmatrix} \partial a/\partial x & \partial a/\partial y & \partial a/\partial z \\ \partial b/\partial x & \partial b/\partial y & \partial b/\partial z \\ \partial c/\partial x & \partial c/\partial y & \partial c/\partial z \end{bmatrix}. \tag{6.30}$$

Linearizing about $(0, 0, 0)$ we write down the determinant

$$|\hat{S} - \lambda\hat{I}| = \begin{vmatrix} \epsilon^{-1} - \lambda & \epsilon^{-1}q & 0 \\ 0 & -q\delta^{-1} - \lambda & 2f\delta^{-1} \\ 1 & 0 & -1 - \lambda \end{vmatrix} = 0, \tag{6.31}$$

where $\hat{I}$ is the identity matrix, which implies

$$\begin{aligned} 0 &= \left(\epsilon^{-1} - \lambda\right)\left[\left(-q\delta^{-1} - \lambda\right)(-1 - \lambda) - 2f\delta^{-1} \cdot 0\right] \\ &\quad -q\epsilon^{-1}\left[0 \cdot (-1 - \lambda) - 2f\delta^{-1} \cdot 1\right] + 0 \cdot \left[0 \cdot 0 + \left(q\delta^{-1} - \lambda\right) \cdot 1\right] \\ &= -\frac{\lambda}{\epsilon}\left(q\delta^{-1} + q\delta^{-1}\lambda + \lambda + \lambda^2\right) + \frac{q}{\epsilon}\delta^{-1}2f\delta^{-1} \\ &= \lambda^3 + \left(1 + q\delta^{-1} - \epsilon^{-1}\right)\lambda^2 - \left[\epsilon^{-1}\left(1 + q\delta^{-1}\right) - q\delta^{-1}\right]\lambda - \frac{q(1 + 2f)}{\epsilon\delta}. \end{aligned} \tag{6.32}$$

Since there is at least one positive root for $\lambda \geq 0$, the steady state $(0, 0, 0)$ is linearly unstable. Linearizing about (x_s, y_s, z_s) we find

$$|\hat{S} - \lambda\hat{I}| = \begin{vmatrix} \epsilon^{-1}(1 - 2x_s - y_s) - \lambda & \epsilon^{-1}(q - x_s) & 0 \\ -\delta^{-1}y_s & -\delta^{-1}(x_s + q) - \lambda & 2f\delta^{-1} \\ 1 & 0 & -1 - \lambda \end{vmatrix} = 0, \tag{6.33}$$

which implies

$$\lambda^3 + A\lambda^2 + B\lambda + C = 0, \tag{6.34}$$

where

$$A = 1 + \frac{q + x_s}{\delta} + \frac{E}{\epsilon}, \tag{6.35a}$$

$$E = 2x_s + y_s - 1 = \frac{x_s^2 + q(x_s + 2f)}{q + x_s} > 0, \tag{6.35b}$$

$$B = \frac{q + x_s}{\delta} + \frac{E}{\epsilon} + \frac{(q + x_s)E + y_s(q - x_s)}{\epsilon\delta}, \tag{6.35c}$$

$$C = \frac{(q + x_s)E - 2f(q - x_s) + y_s(q - x_s)}{\epsilon\delta} = \frac{x_s^2 + q(2f + 1)}{\epsilon\delta} > 0. \tag{6.35d}$$

Note that $A > 0$, since $E > 0$, and that $C > 0$; B can then be positive or negative. We follow Descartes's rule of signs: If, by ignoring coefficients which are zero, N is the number of sign changes in the sequence of coefficients of a characteristic polynomial, then there are at most N roots which are real and positive, and there are either N or $N - 2$ or $N - 4, \ldots$, real positive roots. Then, at least one eigenvalue λ is real and negative. Since $A > 0$ and $C > 0$, the remaining necessary and sufficient condition (all together are called Routh–Hurwitz conditions) for all the solutions λ to have negative real parts ($\Re(\lambda) < 0$) is

$$AB - C = \phi(\delta, f, \epsilon) > 0. \tag{6.36}$$

Therefore, for the steady state to be linearly unstable, δ, f, ϵ must lie in a domain in (δ, f, ϵ) space where $\phi(\delta, f, \epsilon) < 0$; the bifurcation point is $\phi(\delta, f, \epsilon) = 0$. If $B \gg 1$, the asymptotic solution of the characteristic polynomial is given by:

$$\lambda \sim -\frac{C}{B}, \ -\frac{A}{2} \pm \imath\sqrt{B}, \tag{6.37}$$

while if $B < 0$ and $|B| \gg 1$, then

$$\lambda \sim -\frac{C}{|B|}, \ \pm\sqrt{|B|}, \tag{6.38}$$

so that, for large positive B, the steady state is linearly stable, while, if B is large and negative, it is unstable. When the parameters are such that $B = C/A$ (the bifurcation point), we can solve for the roots, namely: $\lambda = -A, \ \pm\imath\sqrt{B}$. If $B = C/A - \omega$, with $0 < \omega \ll 1$, it can be seen by looking for asymptotic solutions in the form

$$\lambda = \pm\imath \left(\frac{C}{A}\right)^{1/2} + O(\omega) \tag{6.39}$$

that the $O(\omega)$ term has a positive real part. Thus, near the bifurcation surface in the unstable region, the steady state is unstable by growing oscillations. In the vicinity of the surface $\phi(\delta, f, \epsilon) = 0$, the system exhibits a small amplitude limit cycle solution with period

$$T = \frac{2\pi}{\left(\frac{C}{A}\right)^{\frac{1}{2}}}. \tag{6.40}$$

It can also be shown that for a given q, there are two critical f_c's such that, for the steady state to be unstable, f must lie in between. For small q we have:

$$\frac{1}{4} \simeq f_{c_1} < f < f_{c_2} \simeq \frac{1+\sqrt{2}}{2}. \tag{6.41}$$

We know that certain parts of the cycle are covered very quickly. A good approximation is given by the simple relaxation model

$$\epsilon \frac{dx}{dt} = y - f(x), \quad \frac{dy}{dt} = -x, \tag{6.42}$$

where $0 < \epsilon \ll 1$ and $f(x)$ is a continuous function such that $f(x) \to \pm\infty$ for $x \to \pm\infty$. The Belousov–Zhabotinsky reaction aptly shows the importance of oscillatory patterns in creating self-organizing systems.

7
General Features of Life

After some introductory considerations about complex systems in the last chapter, in the present one I shall briefly examine some basic elements of life. This means that several issues and problems are touched upon here in a very synthetic form. In the next chapter we shall see some important consequences of this inquiry, while in Chs. 9–11 these aspects and lessons will be deepened and more appropriately discussed in their proper context.

It is important to stress that it is not my aim (in the present chapter or in the following ones) to give a complete account of biology, nor of some of its main subfields. I shall rather examine this matter always from the specific point of view of dealing with information.

After a short recall of some main positions about the definition of life, the proper notion of a biological system is introduced. This notion implies the combination of a metabolism, of a genetic system, and of a selective system. The latter three subjects will be deepened in the folllowing sections.

7.1 A Discussion About Life: A Preliminary Look

According to Rodney Brooks,[1] we need to find an explanation of why life is so different from all artificial systems that attempt to simulate the behavior of living organisms. Several approaches are possible here,[2] and in general each one will stress different features of life, for instance self-organization, self-replication, self-reproduction, epigeny, or homeostasis. The most basic one was traditionally considered to be self-replication. There are historical reasons for this. In fact, during the 20$^{\text{th}}$ century biology became, and still largely is, centered on genetics, with a certain underestimation of ontogenetic and epigenetic aspects. This is also the reason why the first attempts at artificially simulating life were centered (e.g. by von Neumann[3]) on self-replication.

For this reason, I wish to stress some issues about heredity from the start. By *heredity* I understand, following W. Johannsen,[4] the presence of identical genes in ancestors and descendants. Now, in order to have heredity, we need self-production, i.e. the self-generation by the system of its own structures: At a very elementary level, as we have seen, a complex system can only split if it is able to reproduce some basic structures, which implies growth[5] (the system must increase its size) [Sec. 6.6]. For any complex system that builds structures thanks to a favorable thermodynamic exchange with the environment, it is indeed necessary either to grow or to somehow be destroyed:

[1] [BROOKS 2001]. [2] About life's different definitions see [*LAHAV* 1999, pp. 113–21].
[3] [*VON NEUMANN* 1952, *VON NEUMANN* 1966]. [4] Quoted in [*JABLONKA/LAMB* 1995, p. 16].
[5] An idea probably formulated the first time by E. Haeckel [*WEISMANN* 1889, I, p. 72].

To build order or to undergo disorder[6] [Subsec. 2.4.1] or death, i.e. the cessation of metabolism. The reason is that in this exchange process, a stable equilibrium point is an ideal case and as such is at most a transient situation. This is also the basis of the so-called *exploratory* behavior that is shown by *any* biological system.[7] Therefore, given suitable conditions, the organism will grow and split, and in this way finally also reproduce itself giving rise to some sort of duplicates. In the case of multicellular organisms, the organism is able to maintain itself up to the point at which the succession of many generations of cells in the same organism implies such a cumulation of errors and drawbacks that the senescence process begins. This is an additional pressure for self-reproduction, since it can be considered as a process for restoring the original freshness and integrity of the organism.[8]

When we speak of self-reproduction, it is important to distinguish here between limited and nonlimited heredity. Only the latter produces a continuous (and even evolutionarily increasing) variety of forms and is therefore characteristic of living beings. In order to be unlimited, heredity needs to be modular, i.e. there must be a finite code and a set of elementary unities,[9] according to the rules of information codification [Subsec. 2.2.2]. A characteristic error of traditional biology, for instance made by Weismann and his followers, is to think of heredity as a *material* transmission through the generations and not as a transmission of information independently from the individual material constituents,[10] which are indeed continuously replaced. Hereditary transmission as transmission of information is something that is very difficult to understand from a point of view centered on self-replication alone. At least in bacteria, self-splitting appears endless: For Weismann, death is not a primary necessity but is only secondarily acquired as an adaptation. This means that, according to Weismann, it is not contrary to the nature of life to be unlimited, but only that there are no special advantages this solution.[11] The difficulty is the following: The assumption of the existence of self-replication structures (DNA) does not suffice in order to have life, for one also needs, at the very least, the structure and function of proteins, that is, of *other* structures and functions that need to be specified. This is the basic "meaning" of the codified information in organisms, and it is not purely self-replication.

Summing up, life is characterized by both metabolic and informational features. Therefore, from the start, we are going beyond the idea that living beings can be reduced to replicators only (genotypes).[12] The triad replicator–interactor–lineage has also been proposed.[13] Replicators guarantee longevity, fidelity, and fecundity: They are entities that pass on their structure largely intact in successive replications. D. Hull[14] introduced the idea of interactors (phenotypes), due to the fact that the notion of replicator implies a passive behavior relative to the environment. This important point was understood for the first time by C. Waddington.[15] Interactors interact as cohesive wholes with their environment in such a way that these interactions cause replications to be differential. Moreover, Hull defined lineages as entities that can change through time as a result of both replication *and* interaction. According to Hull, it is also possible to speak of the

[6][*SPENCER* 1860–62, p. 245]. [7][*GERHART/KIRSCHNER* 1997, pp. 146–98].

[8][*LLOYD MORGAN* 1891, pp. 125–7].

[9][*MAYNARD SMITH/SZATHMÁRY* 1995] [*MAYNARD SMITH/SZATHMÁRY* 1999].

[10][*WEISMANN* 1893, p. xiii]. On this point see [*PIAGET* 1967, p. 97] [*WILLIAMS* 1992, p. 11] [*DE DUVE* 2002, p. 27].

[11][*WEISMANN* 1889, I, pp. 24–6 and 111]. This is also the reason why DNA is generally assumed to be immortal or to live at least 100 million years [*DAWKINS* 1976, pp. 25–35]. This statement is true only if it is taken to mean "the information coded by the DNA is long-living" [*ELDREDGE* 1985, pp. 136–7], and surely not if one takes DNA in its materiality [*MCGRATH* 2005, p. 39].

[12][*DAWKINS* 1976]. [13][*PLOTKIN* 1993, pp. 86–99]. [14][*HULL* 1988b, 407–12].

[15][WADDINTON 1961b, WADDINGTON 1968b].

triad replication–interaction–selection. This would represent a true evolutionary instantiation of our original informational triad, Processor–regulator–decider, which is already important at the quantum level and for any further process of information acquisition and exchange [Subsec. 2.3.2]. According to another tripartition, which is especially relevant for single cells, the phenotype is the seat of metabolism, the genotype the seat of heredity, of memory, and structural information, and the ribotype the seat of instructional information, and thus being the connection between the previous two aspects.[16]

7.2 Biological Systems

The concept of the biological system is wider than that of the organism.[17] There are indeed biological systems that are superordinate or subordinate to organisms. Examples of biological systems constituting the organism are the genetic system, the metabolic system, the protein, any cell of a multicellular organism, and the brain. Examples of biological systems of which organisms are a part are the environmental niche and the social group.

Following the previous suggestions, my main assumption[18] here is that [Subsec. 2.3.2]

Biological systems represent the integration of the three basic systems that are involved in *any* physical process of information-acquiring: The processor, the regulator, and the decider.

Nonliving self-organized systems do not have this capacity.[19] For instance, Bénard cells are depedent on input (hot plates) that represents both an entropic source and a signal driving the evolution of the system. In this way, the processor is outside (the control of) the system. The fundamental insight developed by Friston and coworkers[20] is precisely that biological systems are more than simply dissipative self-organizing systems, for the reason that they can negotiate a changing or nonstationary environment in a way that allows them to endure (to change in an adaptive sense) over substantial periods of time. This means that they avoid phase transitions that would otherwise change their physical structure. Biological systems seem therefore somehow autarchic.[21] However, any biological system is still dependent on its own external environment for regular (self-)maintenance. Any biological system also produces variability as a response to environmental challenges and tries to integrate both aspects inside itself. Therefore, such a system must avoid two opposite dangers: That of anarchy, in which any subsystem or component tries to separate from the system (it is a form of mechanical and entropic disruption) and that of a too great uniformity and regularity (order), which would make the biological system unable to resist the smallest environmental fluctuation [Subsecs. 2.4.1, 2.4.4 and 3.2.2, Secs. 6.4–6.5]. This is a general issue that shows (for reasons that will be explained below) that biological systems are *adaptive* but *never fully adapted*. This feature is present in any form and at any level of biological systems, being rooted in their typical kind of integration.

The only relatively stable solution to this general problem is represented by organisms, a special class of biological systems. As I have mentioned, an organism is made of biological (sub-)systems (which in turn can be constituted by other, smaller, biological systems in a hierarchical structure [Subsec. 6.3.2]) as well as being related to other biological systems, like environmental niches and social groups. Any subordinate and superordinate system can show a tendency to constitute itself

[16][*BARBIERI* 2003]. [17][*MATURANA* 1970, p. 11]. [18][AULETTA 2008b].
[19]This does not imply that organisms have a mind [BATESON 1971, p. 315].
[20][FRISTON *et al.* 2006, FRISTON/STEPHAN 2007]. [21][AULETTA 2008a]

as an organism. An example can be found in social insects, where the hive of bees tends to become a superorganism. However, such a strategy or tendency produces results that are far less stable than for any true organism. For instance, anarchy through uncontrolled worker production of males in honey bees is a counterstrategy against worker policing and an example of cheating invading cooperative areas.[22] Also, the so-called altruism of insect societies is rather enforced through strict policing by nestmates.[23]

The peculiarity of organisms relies on the fact that they show a specific way to integrate processor, regulator, and decider: I assume in the following that any organism consists of a genetic system, playing the role of the processor,[24] a metabolic system, playing the role of regulator, and, in the most basic example of organisms, the unicellular ones, a lipid membrane acting as a decider.[25] The metabolic system is not concerned with codifying information but rather it deals with the regulation of entropic fluxes between the organism and the environment and is therefore involved in all processes of growth and maintenance [Sec. 6.6]. I emphasize that metabolism is a specific regulatory activity [Subsec. 2.3.2], since regulation, as we shall see, concerns many aspects and subsystems of any organism: For instance, the immune system also has a regulatory activity provided by a specialized type of T cell.[26] Moreover, proteins act as the executive instance allowing and speeding all basic chemical reactions needed for performing the previous activities or for connecting these different subsystems [Subsec. 6.2.4].

Recently, evidence has been found for the above assumption.[27] The above characters may seem to many to be a pure matter of fact. As we shall see, they show instead a sort of necessity and is the first manifestation of the importance of constraints (the adaptive side of emergence [Subsec. 2.4.2]) when we deal with biological systems. It would therefore be appropriate to recall here the words of W. Wundt, who stressed that[28]

> If physiology is obliged, by the uniformity of interaction of physical forces throughout the universe, to accept the postulate that the processes of life have their ultimate basis in the general properties of matter, psychology finds it no less obligatory to assume, in this same matter, the universal substrate of natural phenomena, the presence of conditions which attain to expression as the psychical aspect of vital phenomena. But this latter statement must not mislead us. The latent life of inorganic matter must not be confused, as hylozoism confuses it, with real life [...]; nor must it be considered, with materialism, as a function of matter. The former interpretation is wrong, because it assumes the existence of vital phenomena at a point where not these phenomena themselves are given, but only the common ground upon which they rest and whereby they become possible; the second is wrong, because it posits a one-sided dependence, where in reality we find an interrelation of simultaneously presented but incommensurable processes. We employ the concept of material substance to denote the ground of all objective phenomena.

In the following few chapters I shall often consider the most elementary living organism: The bacterium. Here, the main elements are clearly distinguishable [Fig. 7.1]: The DNA is for coding. The metabolic system is able to control, through protein mechanisms and thanks to the membrane, entropic fluxes with the environment. In unicellular eukaryotes we have incredible progress insofar as these three systems are modularized (more sharply distinguished but also more integrated [Subsec. 2.4.4]): The genetic system is sharply confined inside the nucleus and the metabolic activity happens in the cytoplasm between the nucleus and the membrane (it is a concentric three-layer structure).[29] The nucleus can be considered a cell inside the cell. Also, transcription and translation

[22][RATNIEKS/VISSCHER 1989]. [BARRON *et al.* 2001]. [23][WENSELEERS/RATNIEKS 2006].
[24][*KOZA* 1992]. [25]See also [*DE DUVE* 2002, pp. 10–11]. [26][WALDMANN 2006].
[27][RASMUSSEN *et al.* 2004]. [28][*WUNDT* 1893, pp. 31–2]. [29][*DE DUVE* 2002, p. 42].

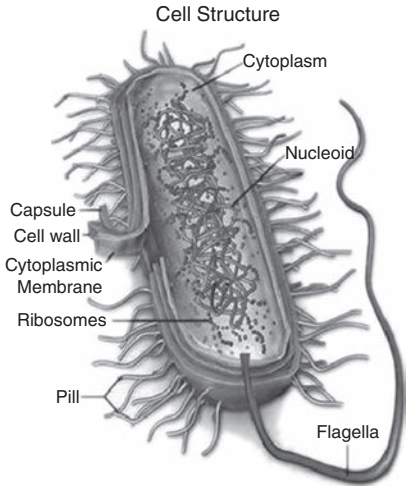

Fig. 7.1 The basic anatomy of a bacterium. We distinguish between the main elements: A membrane, a flagellum, the cytoplasm with the ribosomes, and the DNA. Adapted from wikipedia commons.

are temporally and spatially segregated. The reason is that in such a way a major control becomes possible. In the following, I shall also shift very often to the examination of eukaryotes due to the incredible number of insights they provide about elementary mechanisms of life.

7.3 The Metabolic System

7.3.1 Basic Chemistry of Life

Life has chosen carbon (C) as its basic chemistry because this element can form up to four covalent bonds with other atoms [Subsecs. 6.2.2–6.2.3]. I recall that a covalent bond is a form of chemical bonding that is characterized by the sharing of pairs of electrons between atoms, or between atoms and other covalent bonds [Fig. 6.9]. In this way, carbon can also form long chains and rings [Fig. 7.2], which are necessary for building the macromolecules that constitute an organism. Moreover, it can give rise to an interesting variety of macromolecules: This variety is necessary for life, allowing a relative freedom of metabolic process from pure chemistry.[30] Water (H_2O) is also very important to life and is called the universal solvent, since many substances dissolve in it. Moreover, since it is in a liquid state on most of the Earth's surface, it is also the ideal medium for allowing combinations of complex molecules.[31] Its molecule is formed by covalent bonds between the two hydrogen atoms and the oxygen atom [Fig. 6.8]. Since they are polarized, water molecules are interconnected through hydrogen bonds [Fig. 2.7]: A hydrogen bond is a dipole–dipole force between an electronegative atom, i.e. oxygen (O) or nitrogen (N), and the positively charged hydrogen atom that is already bonded to an electronegative atom[32] [Fig. 7.3]. Some polar molecules are hydrophilic,

[30] [*MONOD* 1970, Ch. 4].

[31] See [*DENTON* 1998, pp. 22–45] for an examination. However, I do not share the main conclusions of the author, as I shall explain in the last chapter of this book.

[32] [*BERG et al.* 2006, pp. 7–10].

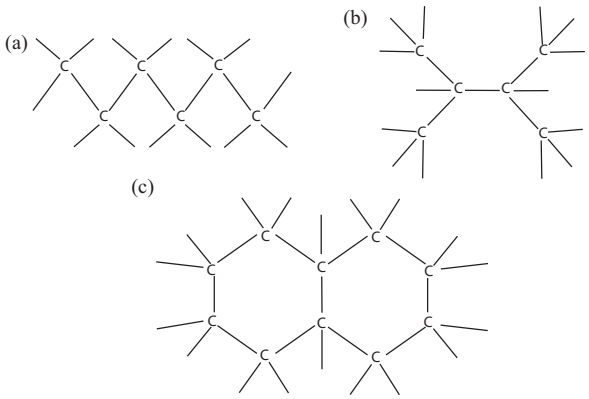

Fig. 7.2 Carbon (C) atoms can give rise to chains (a), branched trees (b), or rings, that is, honey-comb structures (c). Again an example of interesting structures [Subsec. 6.4.1].

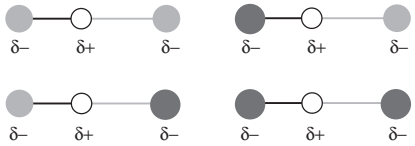

Fig. 7.3 Hydrogen bonds can form between the couple N-H (on the left column) or O-H (on the right column) and the electronegative atom N (top row) or O (bottom row), which gives rise to four possible bonds (that on the right of the bottom row establishes a water molecule). Color conventions: O in dark gray, H in white, N in light gray, hydrogen bond in gray.

while nonpolar molecules are hydrophobic (they interrupt hydrogen bonds). Fats are typical hydrophobic molecules. In water they form a so-called clathrate cage since their coalescence is entropy-favored.

Of course other chemical bonds are possible. Electrostatic interactions (electrovalent or ionic bonds) can often form between metal and nonmetal ions from the attraction between two oppositely charged ions [Fig. 6.6]. We can also speak of metallic bonding when there is the electromagnetic interaction between delocalized electrons, called conduction electrons, and the metallic nuclei within metals. Another type of noncovalent bond is represented by van der Waals interactions: The electric charge on an atom fluctuates with time and does not have a perfectly symmetric distribution. Then the atom and its neighboring atoms can attract each other until, at a shorter distance, the van der Waals force does not act in a repulsive way.

Weak bonds (like the hydrogen bond, ionic bonds, and van der Waals attractions) between different parts of the same macromolecule determine both the shape of the three-dimensional structure of macromolecular organic chains and how these structures interact with one another. A single noncovalent bond is too weak to oppose the thermal motions in water. The requirement that two atoms never overlap limits the possible bond angles. Diffusion (random walk), which enhances the possibility for the "right" molecules to meet, is the first step for molecular recognition. The types of motion are those allowed at a molecular level: Translations, vibrations, and rotations. But thermal motions not only bring molecules together but also pull them apart. There is an

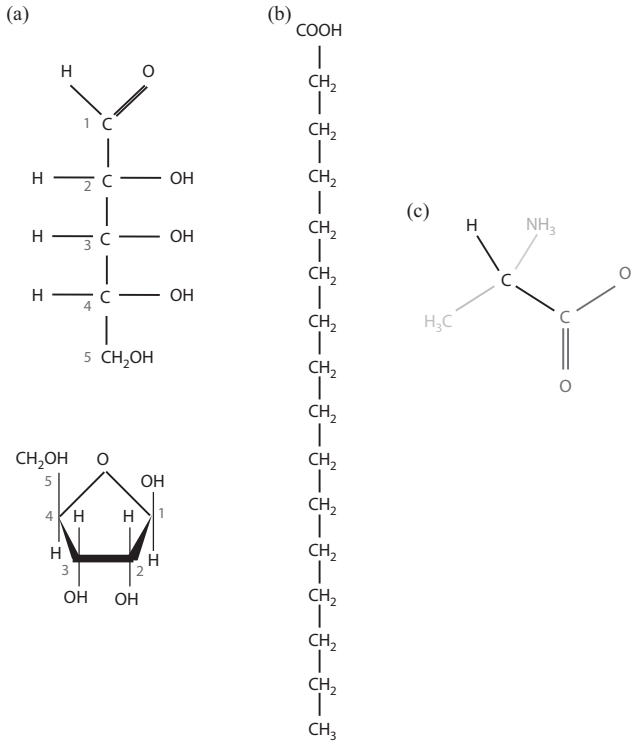

Fig. 7.4 Some examples from the main families of organic compounds. (a) Ribose is a member of the sugar family (its five carbon atoms are numerated): It forms ring-like structures. (b) Palmitic acid is a fatty acid. (c) Alanine (Ala) is an amino acid (of the subgroup of non-polar side chains). All amino acids consist of a central carbon atom (in black), an amino group (in middle gray), a hydrogen atom (in black), a carboxyl acid group (in dark gray), and side chain (in light gray).

equilibrium constant (called the affinity constant) at which the rates of formation and dissociation are equal. Organisms have developed very complex mechanisms for controlling such processes. For instance, because of the random factor in molecular interactions (due to quantum-mechanical effects, especially fluctuations), minor side-reactions occur occasionally, which means that a cell makes errors. As we shall see, cells have also developed repair mechanisms to correct these errors.

Organic molecules can be divided into four major families[33]: Sugars, fatty acids, amino acids [Fig. 7.4], and nucleotides.

(1) Sugars are the principal food compound of many cells, and therefore the main chemicals providing for metabolism (although they are also involved in recognition processes).
(2) Fatty acids are sources of food and also contribute to the cell membrane (they have a hydrophilic head and a hydrophobic tail), which in this way allows for the isolation of cells in water-based solutions, and therefore for information selection relative to the exterior

[33][*ALBERTS et al.* 1983, pp. 45–124].

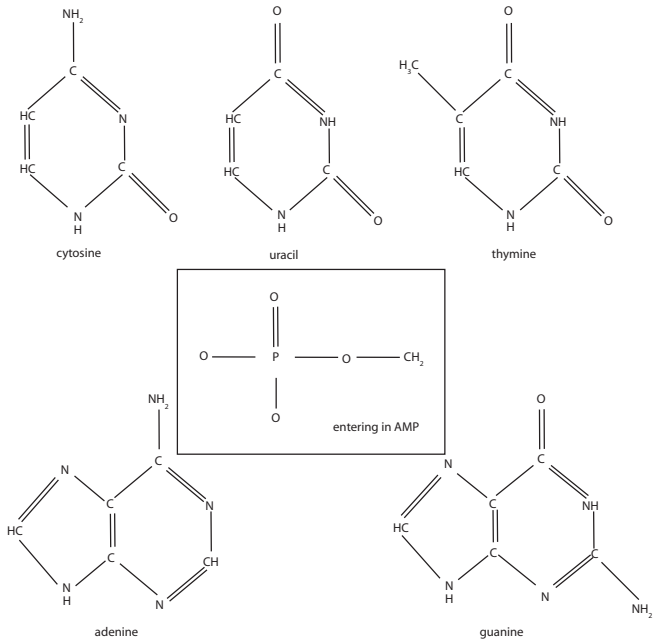

Fig. 7.5 The five bases of nucleotides. *Top*: The three bases (cytosine, uracil, and thymine) constituting the subgroup of pyrimidine (characterized by being a single ring with two nitrogen atoms). *Bottom*: The two bases (adenine and guanine) constituting the subgroup of purine (characterized by being a double ring with 2 + 2 nitrogen atoms). *Middle insert*: A monophosphate entering in the constitution of the nucleotide AMP (one of the four in a RNA molecule) is shown. Phosphates, together with bases and sugars, constitute nucleotides.

environment (an operation that is done by taking advantage of the action of certain proteins). Lipids are soluble in organic molecules.

(3) Nucleotides are universal carriers of genetic information and work as true information processors. They are constituted by the combination of a base, a phosphate, and a sugar [Fig. 7.5]. DNA (deoxyribonucleic acid) has cytosine (C), thymine (T), guanine (G), and adenine (A) as bases, and deoxyribose as the sugar forming a part of its backbone, while RNA (ribonucleic acid) has C, G, A, and uracil (U) instead of T, and ribose as a backbone sugar. Nucleotides are joined together by a phosphodiester linkage between 5′ and 3′ carbon atoms. The numbers refer to the numbering of carbon atoms in the ribose molecule [Fig. 7.4(a)]. In the backbone of DNA the 5′ carbon of a deoxyribose is linked to 3′ of the following one by a phosphate group, and so forth.

(4) Amino acids are constituents of proteins, each with a different side chain that determines the properties of the protein, attached to an α-carbon atom, an amino group (for instance, H_2N), and a carboxyl group (for instance, COOH). There are 20 amino acids. Proteins are the executive of the cell.

As mentioned, we may assume that life began as a molecule or a complex of molecules in which metabolic (autocatalytic) activity [Sec. 6.6] and informational instructions were tightly interwoven. However life may have begun, the crucial point is that the organism is a *systemic unity* in which

the three aspects of an organism—metabolism, genetic programming, and a selecting membrane [Secs. 7.1–7.2]—must be present in order that the whole can work. Thus, there is a certain gulf between the production of some basic constituents of life—like amino acids[34] or RNA—and a living organism, so that it is difficult to have life by chemistry alone[35] without protecting both the metabolism and information codification, as I shall explain in the following. As a matter of fact, evidence has accumulated about the possibility of an ancient RNA world: Experiments have been done in which an RNA-copying enzyme (having both metabolic and genetic features) has been produced. However, this polymerase (it has also been pointed out that RNA can also function as RNA ligase[36]) is not able to fully copy very long templates, since the rate of RNA progeny production does not exceed that of parental decomposition. This is the reason why RNA alone cannot evolve spontaneously.[37] In any case, the RNA, and the other organic molecular components of the cell, represent the sources of variety that have made life possible, but in a way that is still difficult to fully understand.

We come back here to the issue of emergence [Subsecs. 2.4.2, 6.3.2, and 6.5.1]. John Stuart Mill was well aware that there are cases to which a compositional conception of causality no longer applies.[38] With the birth of life, we have reached (through a selective pressure,which is the driving force of evolution) an integrated system of such a level of complexity that it displays new features and behaviors. Let me consider the same problem from another point of view. Recall[39] that in order to self-replicate a molecule must come back to its initial state (it undergoes a constant resetting). There are in general only three ways to come back to an initial state:

(1) By a reversible (physical) process,
(2) By recovering a lost structure, or
(3) By throwing away a structure or elements added during the process.

The first option is impossible for complex systems [Subsec. 6.5.1], the second one contradicts the fact that here the restoration of the initial state comes after a process that in general implies growth. Therefore, only the third option remains. As a matter of fact, DNA replication, the process through which two double strands of DNA are produced from a single double strand, is irreversible. Therefore, we must have an informational component able not only to *code* but also to *select* since irreversibility can only result from some information selection [Subsec. 2.2.1]. Moreover, this process must also be *regulated*, otherwise the different steps would not fit with each other. Therefore, this first basic molecule showing the characters of life (which we shall call a chemoton) must have shown the same basic systemic organization that we find now developed in the present organisms.

It is finally important to stress that all living beings have the same basic metabolism [Fig. 7.6], and this is organized as a network following a scale-free exponential structure.[40] This could be the result of an evolutionary convergence instead of a descent from a common ancestor, provided that life started with an ancestral community of primitive cells, as we shall see in the following.

7.3.2 The Metabolic Cycle

Metabolism is focused on energetic chemical reactions allowing order-building and making use of (free) energy[41] already acquired and stored, while discharging entropy into the environment [Subsec. 2.3.3]. Therefore, the whole process deals with *thermodynamic* entropy and is in accordance

[34][MILLER 1953]. [35][*FOX* 1988]. [36][ORGEL 1986]. [37][STROBEL 2001].
[38][*MILL* 1843, III, Ch. 6]. [39][ELITZUR 1994]. [40][JEONG *et al.* 2000].
[41]Called *calorique* by Lamarck [*LAMARCK* 1809, II, 15].

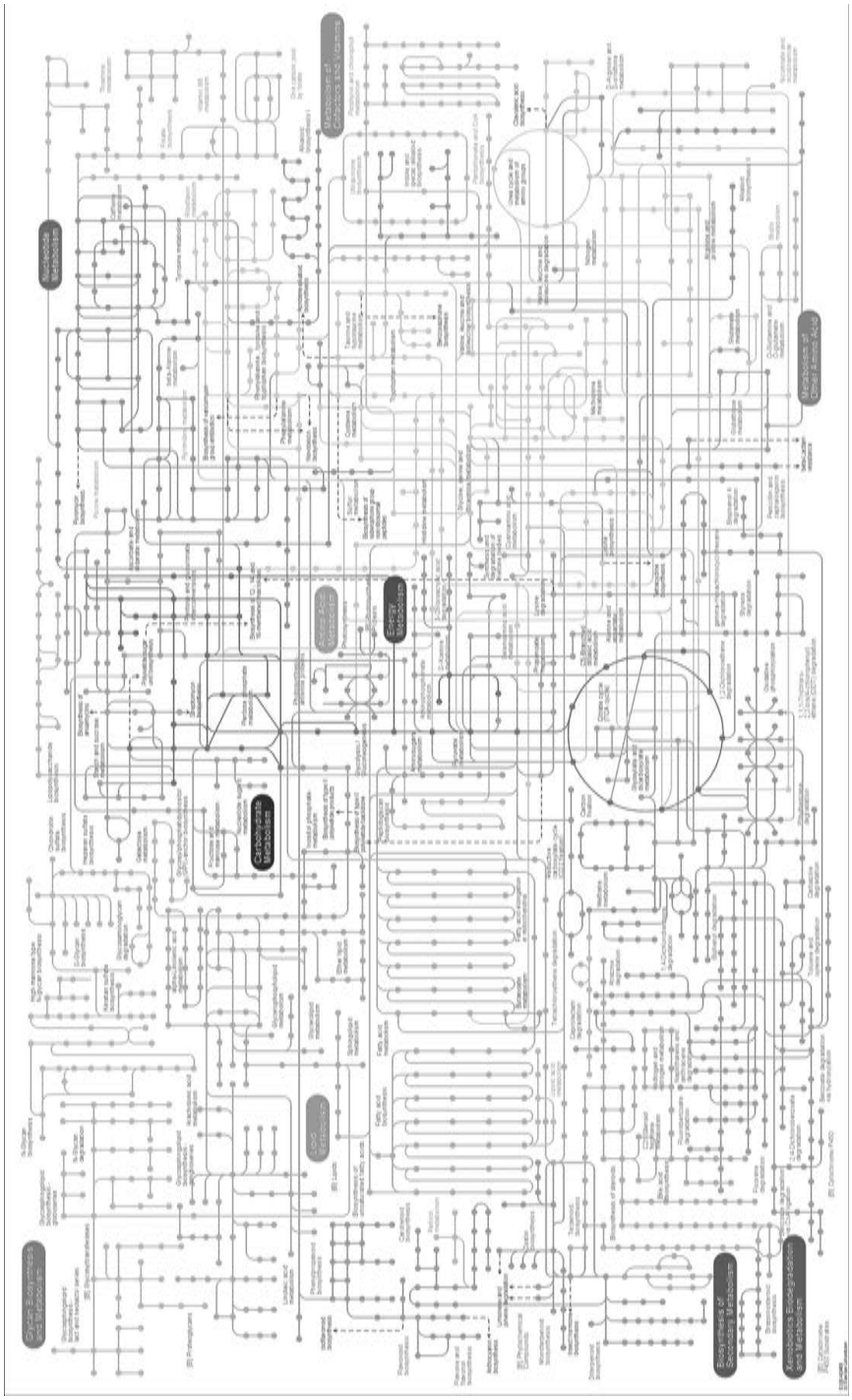

Fig. 7.6 The general structure of metabolic pathways. Any known metabolism can be framed here. Adapted from http://www.genome.jp/kegg/pathway/map/map01100.html. (The plate is reproduced in color in the color plate section.)

with the second law of thermodynamics. It is interesting to observe that quantum systems download information into the environment in order to acquire disorder [Subsec. 2.2.4], while biological systems download entropy into the environment so as to acquire order [Subsec. 2.4.1].

It is appropriate to recall here that Schrödinger had already pointed out that the metabolism is aimed at acquiring negative entropy, and therefore free energy or energy suitable for building order, either for the sake of self-preservation, or for the ability to do work or grow.[42] Therefore, Schrödinger stressed that living beings do not feed only to acquire new stuff, substitute their matter, or obtain energy (even if energy is important as mechanical energy). Besides, energy is acquired to replace the heat continually given off into the environment, since heat dispersion is a way to download entropy. In other words, Schrödinger pointed out that the structure or order, i.e. the formal side, is a much more important factor in metabolism than the material side. For this reason, he had also invoked the exigency of a new physics in order to understand life; namely, a physics not countermanding the old one but still representing a further generalization of it.

Metabolism can be incredibly complex [Fig. 7.6]: Even in tiny *E. coli* about 1,000 chemical reactions are already taking place. However, the basic reactions can be categorized into fundamental types, so that we can speak of a universal chart of metabolism.[43] The two main reactions are *catabolic* reactions (which are exergonic), for obtaining useful energy from previous fuel (represented by carbohydrates or fats), and *anabolic* reactions (which are endergonic), for building complex molecules like proteins (and therefore order) [Subsec. 6.2.4 and Sec. 6.6]:[44]

$$\text{fuel} \xrightarrow{\text{catabolism}} \text{CO}_2 + \text{H}_2\text{O} + \text{useful energy}, \tag{7.1a}$$

$$\text{useful energy} + \text{molecules} \xrightarrow{\text{anabolism}} \text{complex molecules}\,. \tag{7.1b}$$

The less complex molecules on the left of the second equation represent the necessary material *variety* for building higher order (the complex molecules on the right of the same equation). The main energy currency in the organism is represented by ATP [Fig. 7.7]. ATP is the immediate donor of free energy rather than a long-storage form of free energy. In order to make use of the energy stored in ATP, this chemical undergoes hydrolysis, in an exergonic (spontaneous) reaction, to release inorganic phosphate (P_i) and ADP (constituted by a diphosphate, adenine, and ribose) [Fig. 7.8], that is,

$$\text{ATP} + \text{H}_2\text{O} \rightleftharpoons \text{Pi} + \text{ADP}. \tag{7.2}$$

Hydrolysis is a chemical reaction during which one or more water molecules are split into hydrogen and hydroxide ions (a diatomic anion OH, consisting of oxygen and hydrogen atoms). The transformation from ATP to ADP liberates energy suitable for endergonic reactions, such as work, signal amplification in transduction, active transport, and chemical synthesis, which allows for the building of structures, for instance proteins, which combine to form the ontogenetic structure of the organism, i.e. macromolecules whose structure is relevant for their function. The inverse transformation (from ADP to ATP) is provided by photosynthesis or oxidation of fuel molecules. I recall that *oxidation* involves the loss of electrons by a molecule, atom, or ion, while *reduction* involves the gain of electrons by a molecule, atom, or ion.

[42][*SCHRÖDINGER* 1944]. [43][MOROWITZ *et al.* 2000] [*MOROWITZ* 1992, pp. 49–52].
[44][*BERG et al.* 2006, pp. 409–29].

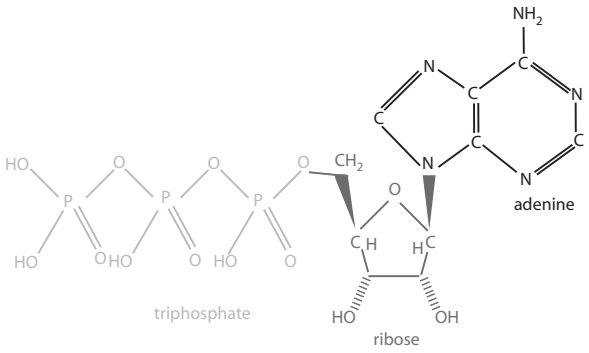

Fig. 7.7 Adenosine triphosphate (ATP), the energy short-term storing chemical, is a combination of adenine and ribose (which gives rise to adenosine) and of a triphosphate. The triangle denotes a bond pointing toward us out of the plane of the paper, while the hatched one goes away from us.

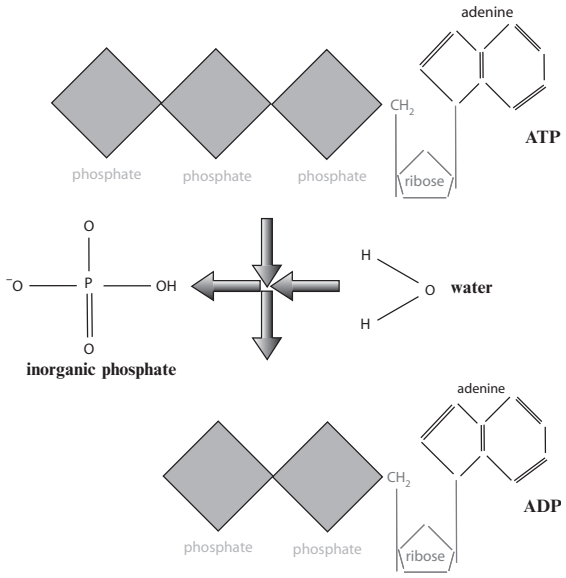

Fig. 7.8 The chemical reaction producing ADP from ATP.

Let us now consider some simple mechanisms for building ATP: By bringing high-energy electrons to the ground state, one can use the differential energy for building ATP [Fig. 7.9], a mechanism which is used in cellular respiration.[45] Another widely used mechanism (for instance, in chemiosmosis) is represented by the protonmotive force[46]: Two reversible proton (H^+) pumps are coupled, one driven by the transfer of electrons between two carriers and the other by ATP hydrolysis [Fig. 7.10]. The first pump transfers protons by making use of electrons that are first

[45][*DE DUVE* 2005, pp. 41–53]. [46][*DE DUVE* 1995, pp. 99–102] [*DE DUVE* 2005, pp. 133–48].

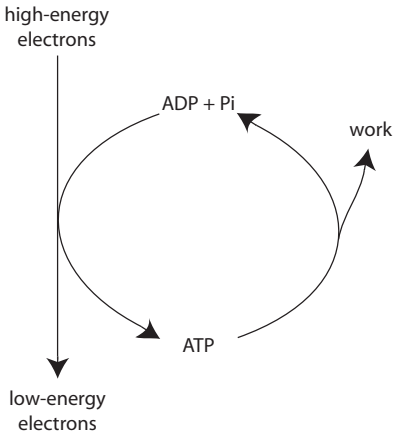

Fig. 7.9 The use of high-energy electrons for building ATP. Inspired by [*DE DUVE* 2005, p. 44].

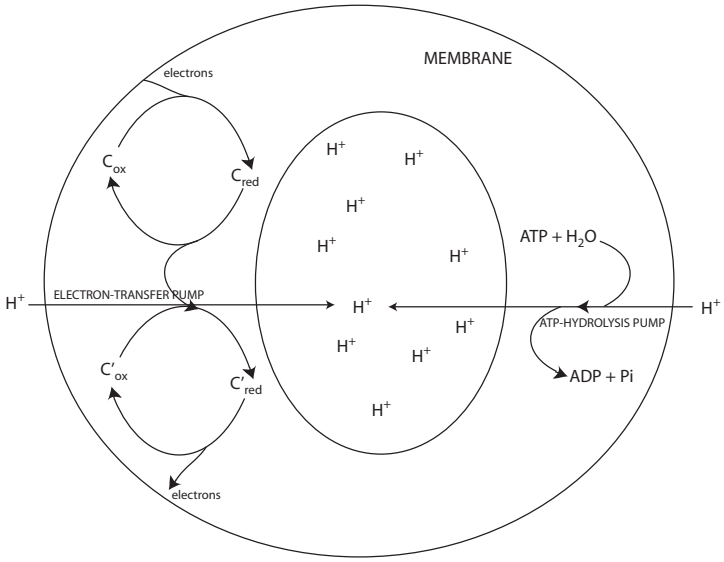

Fig. 7.10 Mechanism of the protonmotive force by making use of two coupled pumps. When the electron-driven pump builds a higher proton (H^+) potential, the ATP pump works in a reverse mode building ATP from ADP. Inspired by [*DE DUVE* 2005, p. 134].

given to a carrier that becomes reduced (C_{red}), i.e. electron rich, then from this (that becomes thereafter the oxidized C_{ox}) they are given to another carrier C', subject to an analogous procedure. The second pump transfers protons by hydrolysis of ATP. When, as is often the case, the electron-driven pump builds a higher proton potential than the ATP-driven pump, the latter functions in a reverse mode and synthesizes ATP.

Animals obtain ATP from the oxidation of biological molecules. This catabolic process occurs in three stages:

- The first one is the digestion, where the four organic molecules are derived.
- In the second stage, these subunits are broken down into acetyl CoA and small amounts of ATP and NADH (nicotinamide adenine dinucleotide) are produced (this can happen, for instance, in an aerobic way by glycolysis).
- In stage three, acetyl CoA is degraded into CO_2 and H_2O and a large amount of ATP and NADH are produced.

The *whole* metabolic process can be considered as a three-system or three-step process[47] (for the sake of simplicity, I only consider sugars here as fuel, and not fatty acids[48]):

- Energy is *stored* in long-term fuel molecules, i.e. glucose, through gluconeogenesis.
- Then, glucose is catabolicly *broken down* in a process called glycolysis releasing ATP.
- Finally, ATP is burned or anabolicly used[49] for *building* amino acids, polynucleotides, DNA, and RNA, selecting a specific destination among many possible ones. These molecules then regulate the building of new glucose.

Let us consider this cycle in green plants. Photosynthesis uses the sun's energy for deriving NADPH (nicotinamide adenine dinucleotide phosphate) from water and synthesizing ATP. Then, ATP and NADPH are used in forming sugar molecules by taking CO_2 from the air. The result for green plants can be synthesized as[50]

$$\text{photons} + 6CO_2 + 12H_2O \rightarrow C_6H_{12}O_6 + 6O_2 + 6H_2O\ . \tag{7.3}$$

In the course of glucose ($C_6H_{12}O_6$) breakdown through a series of oxidations, energy (ATP) and reducing power (in the form of NADH) are produced, in a sort of inverse transformation, whose net result can be synthetically written as

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O + \text{ATP}\ . \tag{7.4}$$

In this way, the organism, by acquiring free energy from the environment and by discharging entropy into it, in a circle of exergonic and endergonic reactions, is able to build itself as a structured and ordered system. It is a true feedback, self-increasing, circle [see Fig. 7.11]. It is this *systemic circularity* that must be maintained to preserve the unity and identity of the organism.[51] This circularity continuously brings the organism back to the same internal state, representing the tendency to preserve the same state against external fluctuations and changes (*homeostasis*).

Following Aristotle, F. Varela pointed out that a living organism is characterized by autopoiesis, i.e. it demolishes and rebuilds its own structures in a permanent process of self-production.[52] The

[47] [*BERG et al.* 2006, pp. 433–70].

[48] The fact that fatty acids are also fuels and do not only contribute to membrane-building is evidence that the membrane also originated from an autocatalytic process [*GÁNTI* 1987].

[49] [*DE DUVE* 1995, pp. 46–7].

[50] [*ALBERTS et al.* 1983, pp. 68–70].

[51] [*MATURANA* 1970, pp. 9–10] [*MATURANA/VARELA* 1980, pp. 78–9].

[52] [*MATURANA/VARELA* 1980, *MATURANA/VARELA* 1987] [MCMULLIN/VARELA 1997]. Varela and McMullin developed a computer model taking into account the following three elements: 1) substrate particles S, 2) catalysts K, 3) link particles L (made of two S) that can form chains and membranes, also making use of a rule called "chain-based bond inhibition." In this case, a free L particle cannot form a bond as long as it is within an existing chain but rather only at the end of the chain. This model has shown that cell-like closed structures spontaneously develop.

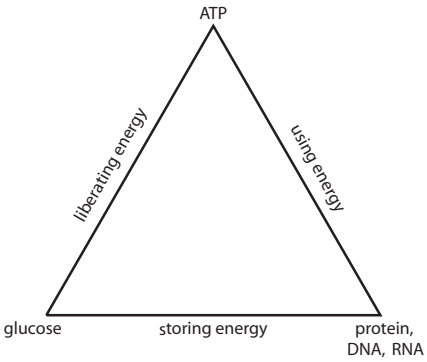

Fig. 7.11 Metabolic circle. This is an energetic circle. The direction is important here and cannot be inverted: It runs clockwise.

flux of external matter and free energy is transformed into an internal flux of self-production. An autopoietic organization is an autonomous and self-maintaining unity which contains component-producing processes. The components, through their interaction, generate recursively the same network of processes which produced them. An autopoietic system is operationally closed and structurally self-determined [Subsec. 3.2.2].

7.3.3 Organism Size and Metabolism

The size of the organism is relevant for its metabolism rate and structural properties.[53] In geometrical three-dimensional figures, the surface increases as the square, and the volume as the cube, of the linear dimensions. For this reason, at the very least, in living beings we see a problem of absolute magnitude, i.e. of scale, which depends on the organism itself and on its relation to the whole of its environment. In principle, there are forces acting on the surface and gravity acting on mass or volume. A large ratio of surface to mass in small animals would lead to an excessive transpiration if the skin were porous; for this reason, insects have hardened or thickened skins. The heat loss varies as the surface area does, whereas the heat produced by oxidation varies as the bulk of the animal does. Therefore, the ratio of loss to gain, like that of surface to volume, increases as the size of the specimen diminishes. For this reason, small animals, in order to produce more heat, need more food. A way of establishing a good balance between surface and volume is to alter the shape (for instance, by folding, a solution found in the human cortex). I should also mention that there are constraints on the skeleton of vertebrates, again due to the action of gravity on volume. Indeed, very big quadrupeds need a proportionally bigger skeleton than smaller vertebrates.

7.4 Genetics and Proteins

The genetic program of the organism consists in *codified* information, that is, (1) a linear combinatorics of (2) elementary and discrete unities (3) following syntactic rules [Subsec. 2.2.2]. Its basic elements are the nucleotide bases U (T), C, A, G [Subsec. 7.3.1], whose combinations build

[53] [*SPENCER* 1864–67, I, pp. 151–3] [*WEISMANN* 1889, I, p. 7] [*THOMPSON* 1942, pp. 35–53] [VON BERTALANFFY 1957].

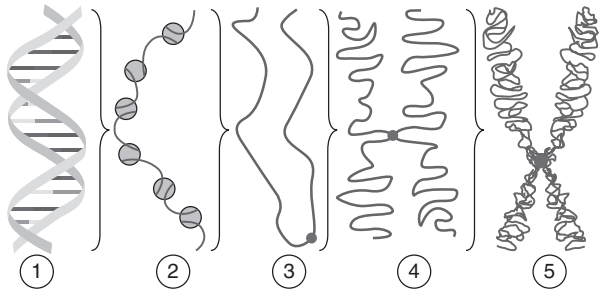

Fig. 7.12 Levels of DNA condensation. (1) DNA double-strand helix. (2) Chromatin strand (DNA with histones). (3) Condensed chromatin during interphase with centromere. (4) Condensed chromatin during prophase (two copies of the DNA molecule are now present). (5) Chromosome during metaphase (a stage of mitosis in the eukaryotic cell cycle in which condensed chromosomes, carrying genetic information, align in the middle of the cell before being separated into each of the two daughter cells). Adapted from http://www.all-science-fair-projects.com/science_fair_projects_encyclopedia/Chromatin.

triplets (codons), which, as words of language, give rise to proteins. While in prokaryotes DNA is packed into structures called nucleoids, in eukaryotes the DNA is located in the nucleus of the cell and densely packed together with histone proteins.[54] It incorporates the genetic information of the organism, while RNA is the means to use this information as a set of instructions for building proteins. Each region of the DNA helix[55] that produces a functional RNA molecule constitutes a gene. DNA is packed into nucleosomes, which consist of two full turns of DNA wound around an octameric histone core (histones act as spools around which DNA winds and they play a role in gene regulation) plus the adjacent linking DNA [Fig. 7.12].[56] Then, several nucleosomes are packed together to form a 30 nm chromatin fiber, and several chromatin fibers constitute a chromosome. Each DNA molecule that forms a linear chromosome must contain a centromere, two telomeres (the two ends), and a replication origin (a human chromosome is about 220 million base pairs long). Finally, many chromosome strings constitute the metaphase chromosome. This is a true information encapsulation [Subsec. 6.3.2], namely the way in which information and complexity are integrated in life: A variety of different labels are assigned to each package independently from the data contained (an example of hierarchical organization), allowing for a quick recovery of information.[57] In this way, as already understood by Waddington, such coherent structures allow for a more refined control of activity.[58] In each human somatic cell there are 22×2 chromosomes plus chromosomes X and Y for males or X and X for females (46 all together).

7.4.1 Genetic Information

The double helix of DNA [Fig. 7.13] is composed of the bases on the inside and sugars–phosphates on the outside.[59] The allowed connections from one strand to the other are A with T and G with C (always a purine with a pyrimidine [Fig. 7.5]). The fact that only T binds with A and only C with G [Fig. 7.14], allows that in each copying operation a negative of the original is provided, which,

[54][*ALBERTS et al.* 1983, pp. 211–19].
[55]This was discovered by Watson and Crick [WATSON/CRICK 1953a, WATSON/CRICK 1953b].
[56][*GILBERT* 2006, pp. 101–2]. [57][SHAPIRO 2002]. [58][VAN SPEYBROECK 2002].
[59]For an introduction to genetics see [*ALBERTS et al.* 1983, pp. 125–499] [*BERG et al.* 2006, pp. 107–29].

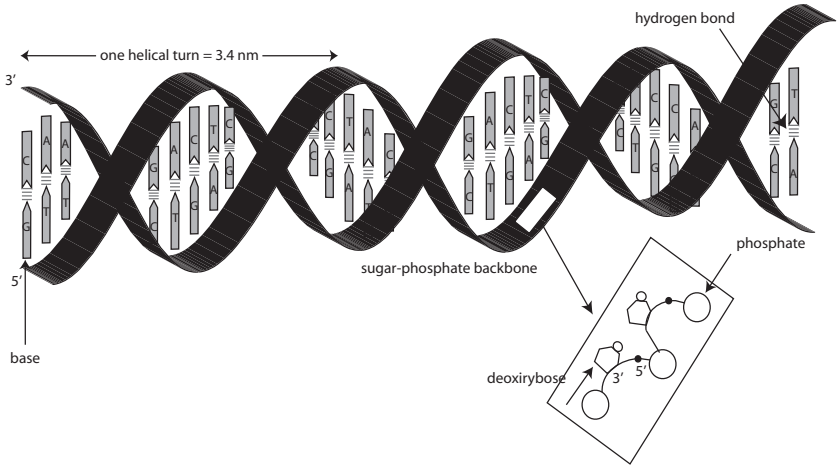

Fig. 7.13 DNA's double-helix structure. A helix is a common motif (actually, a type of wave) in biological structures.

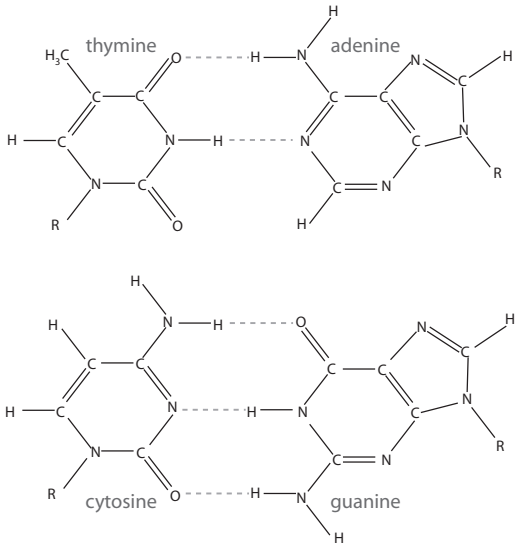

Fig. 7.14 The base T establishes hydrogen bonds (dashed line) with A (above) as well as C with G (below). Note that the upper hydrogen bond in the pair T–A as well as the upper and the lower bond in the pair C–G are of the type shown in the left corner of the bottom row in Fig. 7.3, while the lower hydrogen bond in the pair A–T as well as the middle one in the pair C–G are of the kind shown in the left corner of the top row in Fig. 7.3.

Table 7.1 The 20 amino acids and their bases constituting the genetic code. Since the combinatoric possibility of bases in the different codons (a triplet of nucleotides) is 4^3, that is 64, each amino acid may be coded using different bases (it is called degeneracy, even if it is a case of redundancy). Only 61 codons are employed to codify for amino acids, so that 3 combinations of bases (UAA, UGA, UAG) are free for providing the stop signal that terminates the translation process.

Amino acid	Abbr.	Symbol	Codons	Side chain polarity	Side chain acidity or basicity	Hydropathy index
Alanine	Ala	A	GCA, GCC, GCG, GCU	np	n	1.8
Cysteine	Cys	C	UGC, UGU	p	n	2.5
Aspartic acid	Asp	D	GAC, GAU	p	a	~ 3.5
Glutamic acid	Glu	E	GAA, GAG	p	a	~ 3.5
Phenylalanine	Phe	F	UUC, UUU	np	n	2.8
Glycine	Gly	G	GGA, GGC, GGG, GGU	np	n	~ 0.4
Histidine	His	H	CAC, CAU	p	wb	~ 3.2
Isoleucine	Ile	I	AUA, AUC, AUU	np	n	4.5
Lysine	Lys	K	AAA, AAG	p	b	~ 3.9
Leucine	Leu	L	UUA, UUG, CUA, CUC, CUG, CUU	np	n	3.8
Methionine	Met	M	AUG	np	n	1.9
Asparagine	Asn	N	AAC, AAU	p	n	~ 3.5
Proline	Pro	P	CCA, CCC, CCG, CCU	np	n	~ 1.6
Glutamine	Gln	Q	CAA, CAG	p	n	~ 3.5
Arginine	Arg	R	AGA, AGG, CGA, CGC, CGG, CGU	p	sb	~ 4.5
Serine	Ser	S	AGC, AGU, UCA, UCC, UCG, UCU	p	n	~ 0.8
Threonine	Thr	T	ACA, ACC, ACG, ACU	p	n	~ 0.7
Valine	Val	V	GUA, GUC, GUG, GUU	np	n	4.2
Tryptophan	Trp	W	UGG	np	n	~ 0.9
Tyrosine	Tyr	Y	UAC, UAU	p	n	~ 1.3

Legend: p = polar, np = non-polar, n = neutral, sb = strongly basic, wb = weakly basic, a = acidic.

through a new copying operation, is able to restore the original[60]: DNA replication (for producing other DNA) begins with a local separation of its two complementary strands. Each strand acts as a template for the formation of new DNA molecules by sequential addition of deoxyribonucleotide triphosphate. This also means that the genetic code has perhaps originated from two independent basic binary codes (for DNA, likely C–T and A–G) evolved into a quaternary code with positive–negative subcoding operations [Tab. 7.1]. Moreover, as mentioned, there is a third (RNA) basic code, represented by subcodes C–U and A–G. This suggests a former phase of life with a larger variety of codes, from which, through fine-tuning due to selective pressure, the actual codification has been born. This is also shown by the possibility of artificially producing new base pairs: The

[60] [*KÜPPERS* 1990, pp. 12–15].

so-called bases S and Y have been added to the four traditional ones by producing a sort of six-bases code allowing for the building of $6^3 = 216$ different amino acids.[61] These considerations are fundamental, since they show a basic property of any codified information to be true: The existence of *several* alternative codes as well as of translation rules between them [Subsec. 2.2.2].

Very appropriately, Shapiro and von Sternberg distinguish between three different information-storing processes that are performed by DNA[62]:

(1) Long-term ("genetic") storage involves DNA sequence information, stable for many generations.
(2) Intermediate-term ("epigenetic") storage occurs through the complexing of DNA with protein and RNA into chromatin structures that may propagate over several cell generations. Chemical modifications of DNA that do not change sequence data, such as methylation and demethylation[63], contribute to epigenetic storage.
(3) Short-term ("computational") information storage, which we may call a genetic-system information storage (genetic system for short), involving dynamic interactions of DNA with proteins and RNA molecules that can adapt rapidly within the cell cycle as the cellular environment changes (as it happens in many ontogenetic activities like chemotaxis). Information about recent conditions inside and outside the cell is maintained in the form of transient nucleoprotein complexes reflecting recent responses to internal and external signals.[64]

Programs (2)–(3) are not hardwired into the DNA sequence (there is in fact no linear relation genotype–phenotype, as was once assumed), and they sometimes permit the formation of very different organisms using a single genome (e.g. invertebrates have distinct larval and adult stages). As we shall see, even different proteins can be coded starting from the same sequence (a form of multifinality [Subsec. 6.3.3]). As already mentioned, the reason for all the processes described so far, of monitoring, computation, and decision-making, is to keep millions of chemical interactions from undergoing chaotic transitions and spinning out of control.[65] Moreover, genetic information is rewritten many times and then finally altered. Actually, cells are like natural genetic engineering machines. In other words, living cells could theoretically rearrange their genomes in any way that is compatible with the rules of DNA biochemistry.

It is possible to consider information coding in DNA as a metaphoric way of speaking.[66] Suppose that certain configurations of DNA molecules were privileged due to the fact that the bindings of their bases were much stronger than they would be for any other distribution of bases; then such a DNA would have no information content.[67] This is actually the case for atoms and ordinary chemical molecules [Subsecs. 2.4.1 and 6.2.1]: Since the orderly structure here is due to a maximum of stability, corresponding to a minimum of potential energy [Subsec. 6.2.4], the orderliness of such molecules lacks the capacity to function as a code. The pattern of atoms forming a crystal is another instance of order without appreciable information content. Therefore, whatever may be the origin of a DNA configuration, it can function as a code *only if* its order *is not* a deterministic consequence of the forces due to potential energy. The issue of a certain order or configuration of elements (bases) must be as *physically indeterminate* as the sequence of words is on a printed page: As the meaningful arrangement of words on a printed page is extraneous to the physics and chemistry of the printed page,[68] so too is the base sequence in a DNA molecule as extraneous to the chemical forces at work in the DNA molecule. As I have stressed [Sec. 2.1], information is a *formal*

[61] [HIRAO *et al.* 2002]. [62] [SHAPIRO/VON STERNBERG 2005]. [63] [HAJKOVA *et al.* 2008].
[64] [SHAPIRO 2006]. [65] [SHAPIRO 2002]. [66] [GRIFFITHS 2001]. [67] [POLANYI 1968].
[68] A point well understood long before the birth of information theory [*BOREL* 1920, p. 297].

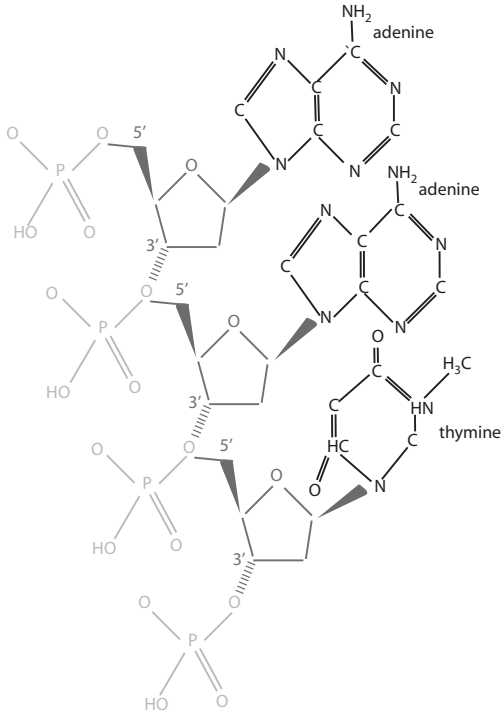

Fig. 7.15 The DNA sequence AAT. Note that the bases (on the right) remain essentially untouched here and chemical connections only concern the sugar–phosphate backbone. Therefore, base pairs essentially maintain the same structure in any sequence. It is precisely this property that allows DNA to be a good storage for information. Triangular bonds points toward us; hatched bonds away.

quantity. Then, the possibility to combine DNA "letters" in this way is guaranteed by the fact that base pairs enter into sequences without changing their shape [Fig. 7.15]: To obtain information codification, it is necessary that chemical bonding and information combinatorics be *separated*. As a matter of fact, the bases only pair across the two strands and *not along* the same strand. In other words, the *sequence* of bases cannot and indeed does not depend on chemistry and therefore can be arbitrary relatively to chemical reactions and bonds. This result can only be obtained with large (complex) molecules (which have significant mutual information among the components and which allow also for the possibility of *local* information codification [Subsec. 6.4.2]): A part of such molecules permits chemical bonds (the sugar–phosphate backbone) while the informational units will be kept separated (the DNA bases).

This is a true inventive generation, since with the genetic code the first classical codification of information of our world was born, determining a completely new situation which allows for further possibilities. Indeed, if it is a fact that quantum systems already codify information [Sec. 2.2], it is also true that, due to their nonlocal features (interference terms, entanglement), they show a too ordered configuration—actually a zero-entropy one in the case of pure systems [Subsec. 2.3.3]. This means that they are the simplest systems in nature, which in turn implies that they simply represent the information they are: Quantum systems cannot represent information that can somehow be

used as a set of instructions. Instead, the true novelty of the genetic code is that it is both *complex* (in its constituents and chemical constraints) and *linearly* (informationally) ordered; and it is precisely for this reason that it can give rise to *a set of instructions*, represented by RNA, which can even be considered the only true set of instructions existing in organisms.[69] Indeed, to have instructions, we need a further requirement relative to information codification: The arrangement of the words (the triplets) cannot be random, and, technically speaking, we need here a *permutation* (where the order of the codifying units is relevant) and not a simple combinatorics (where the order of the elements does not matter). This is necessary, since otherwise the sequence of DNA could not be read univocally. Indeed, by looking at Tab. 7.1, we can see that the letters A, C, U can give rise to histidine (in the arrangement CAU), leucine (CUA), isoleucine (AUC), serine (UCA), threonine (ACU), or tyrosine (UAC). I also remark that no *single* amino acid can be coded by the same basic elements differently permuted to avoid ambiguities.

Therefore, to accomplish this act of inventive generation, apart from the separation between chemical bonds and information combinatorics, some additional physical and chemical constraints are necessary so that only certain reactions and combinations occur starting from a higher initial variety. It is a selection and amplification process[70] [Sec. 6.3]. Those bases that allowed self-replication with the highest fidelity would be the ones that were finally preferred [Sec. 7.1]. These initial constraints can largely be found in the chemistry of the amino acids. Apart from redundancy (i.e. the fact that the same protein unit can be coded through different codons) [Fig. 7.16(a)], amino acids show important correlations concerning[71]:

(a) Hydrophobicity [Fig. 7.17(a)],
(b) Energy dependence with respect to the volume [Fig. 7.17(b)], and
(c) Correlation between these first two correlations.

These physico-chemical properties allow for the creation of the genetic code as a set of instructions: Only amino acids satisfying the requirements expressed by the first two correlations should enter into the genetic code table, in which their hydrophobicty and a volume parameter determine their respective positions in the code.

This examination of the constraints necessary for the emergence of a classical code together with what has been said about quantum information and complexity [Secs. 6.2 and 6.4–6.5] very nicely shows that,

> In order for classical information coding and processing to emerge in our chemical and macroscopic world, two conditions are necessary: (1) That quantum features must get lost (this happens through decoherence [Subsec. 1.2.8]), and (2) that both the operations of encoding and processing must be shielded against (nonlinear) fluctuations and noise coming from the environment.

Condition (2) allows for any classical code to be also a set of instructions. It is not easy to attain, and thus demands either complex chemical constraints like those we have mentioned here (as well as those active at the peripheral receptors for complex organisms), or those that are imposed in logical circuits made of silicon [Subsec. 2.4.1]. To fully appreciate the importance of both the connection and the difference between quantum and classical information, we must recall that codification is a common trait and that DNA is not completely shielded against quantum and complex[72] fluctuations, which remain a fundamental source of variation and innovation.

[69] Obviously, in an ancient RNA world this molecule was both a codifying chemical and a set of instructions. This, however, does not diminish the relevance of this conceptual distinction and on the contrary, as we shall see, helps us to understand how helpful (and even necessary) was the passage from such a world to the actual genetic mechanism.
[70] [*DE DUVE* 2002, pp. 65–6]. [71] [LEHMANN 2002]. [72] [PARISI 2006].

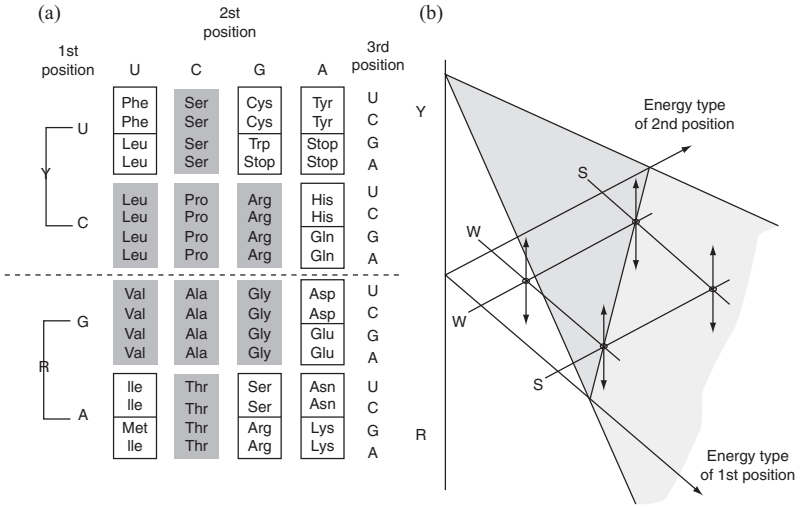

Fig. 7.16 Symmetrical representations of the genetic code. (a) The universal genetic code table written as the succession [U, C, G, A], which exhibits the redundancy symmetry shown by the dashed line: Fourfold redundant codon families are indicated by grey rectangles and twofold ones by squares. The mitochondrial code is more symmetrical, because UGR and AUR are also twofold codon families. (b) Schematic representation of the redundancy (mitochondrial code). Each small arrow indicates whether the considered codon belongs to a twofold (below the inclined plane) or a fourfold (above the inclined plane) redundancy family. The energy types of 1st and 2nd codonic positions are indicated by W (weak: A or U) and S (strong: G or C). The vertical axis indicates the R/Y type. Adapted from [LEHMANN 2002].

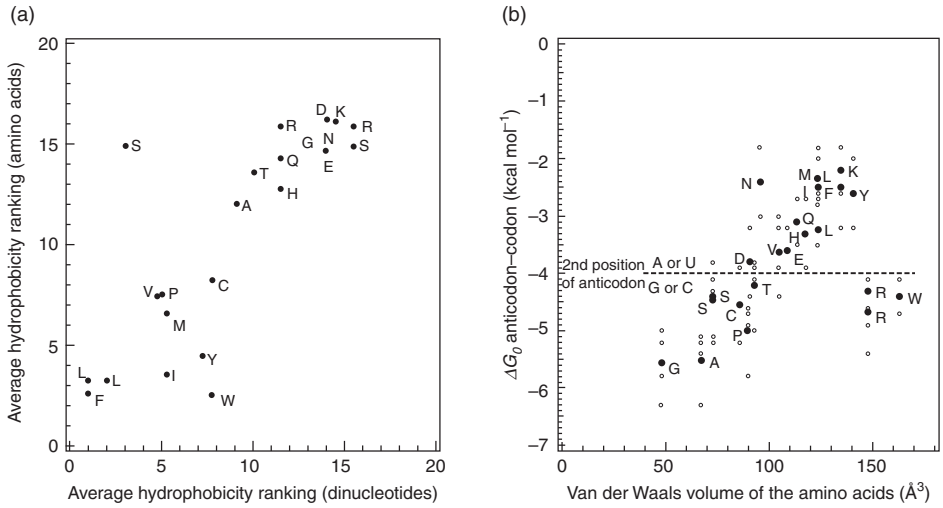

Fig. 7.17 (a) Average hydrophobicity of amino acids ranked as a function of the average hydrophobicity of their corresponding anticodonic dinucleoside monophosphates. (b) Correlation energy dependence–volume in amino acids. Letters follow the codification shown in the third column of Tab. 7.1. Adapted from [LEHMANN 2002].

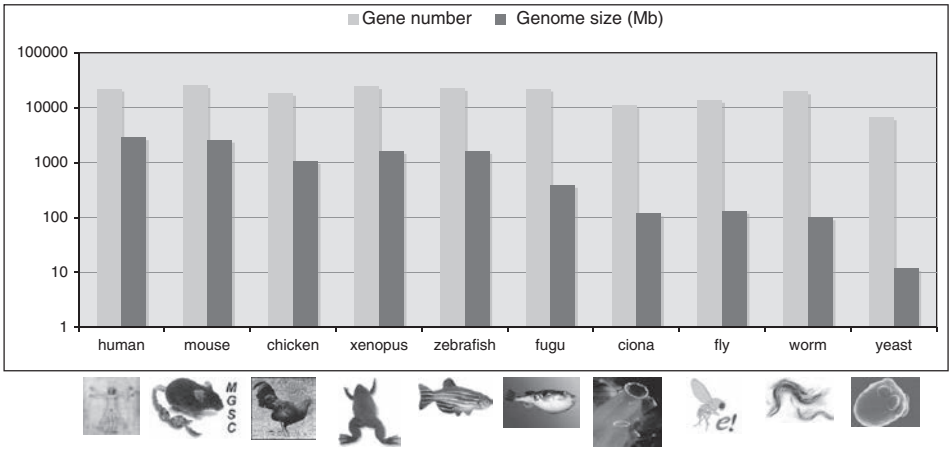

Fig. 7.18 Plot of gene number and genome size (courtesy of Graziano Pesole, University of Bari, Italy).

Since what is concerned here is a true information codification, there is no direct connection between the quantity of DNA and the complexity of the organisms[73] [Fig. 7.18]. Today it is also acknowledged that the same DNA sequence can code for different proteins, according to many factors, like where the stop signal is set, how the introns are recombined after splicing, how exons are dealt with, and so on. This shows that DNA only contains *potential* information [Subsec. 2.2.2] that must be activated under a set of conditions that determines which part will be transcribed into a set of instructions. Such an inquiry, however, is part of a wider investigation that we will turn to later.

In conclusion, it would be dangerous to reduce the function of the genetic system (comprehending DNA, RNA, and certain proteins) to codification. There are many additional functions in which it is involved, such as[74]: (1) Regulating timing and extent of coding sequence expression. (2) Organizing coordinated expression of protein and RNA molecules that work together. (3) Packaging DNA appropriately within the cell. (4) Replicating the genome in synchrony with the cell division cycle. (5) Accurately transmitting replicated DNA to progeny cells at cell division. (6) Detecting and repairing errors and damage to the genome. (7) Restructuring the genome when necessary (as part of the normal life cycle or in response to a critical selective challenge). Now, these additional capabilities involve specific kinds of interactions between DNA and other cellular molecules, in which semiotic aspects are deeply involved (about which much will be said below).

7.4.2 The Necessity of Random Mutations

G. Hardy and W. Weinberg[75] proved that the genetic pool of a given population remains constant if the population is large and there are no disturbing forces. This can be expressed by the formula [Fig. 7.19]

$$p^2(AA) + 2pq(Aa) + q^2(aa) = C \,, \tag{7.5}$$

[73][KNIGHT 2002]. [74][SHAPIRO/VON STERNBERG 2005].
[75][HARDY 1908] [WEINBERG 1908]. See also [*DOBZHANSKY* 1970, pp. 99–125]. [*DEPEW/WEBER* 1995, pp. 232–3].

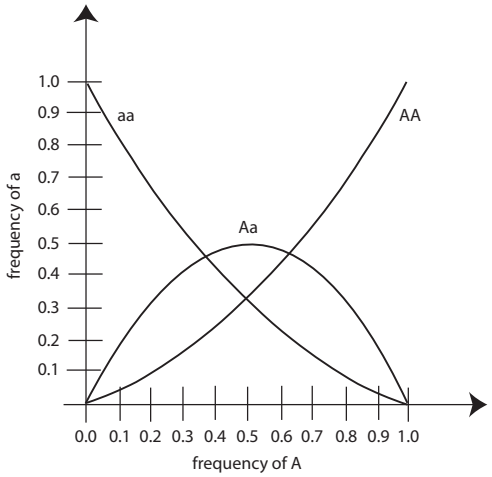

Fig. 7.19 Graphical representation of Hardy–Weinberg distribution (7.5).

where C is a constant, and $p(A)$ and $q(a)$ are the probabilities of obtaining allele A (pure dominant) and a (pure recessive), respectively (Aa being heterozygotes). Alleles are the different variants of a specific gene.

Nevertheless, random mutation is a necessity for life, a fact I shall account for in the following chapters. Let me give a preliminary example: Viruses.[76] Viruses are genetic elements (of DNA or RNA) enclosed by a protective coat (made of proteins) that enables them to move from one cell to another. Cells must have evolved before viruses, or at least viruses represent a first branching of the tree of life when organisms were formed for the first time. The precursors of the viruses were probably small nucleic acid fragments (plasmids) that developed the ability to multiply independently from the chromosomes in their host cells. Under the pressure of selection, such precursors could facilitate their own multiplication by acquiring nucleotide sequences from the host cells, including sequences that code for proteins. Viruses must *mutate* to survive the attacks of the host's immune system.[77] The viral mutation rate is optimized in an evolutionary trade-off between adaptability and genomic integrity. This is a consequence of a feature already observed in complex systems, according to which noise and errors may be amplified (through positive feedback) and play a positive role in becoming a source of new equilibria when such mutations result adaptive or can be integrated in the organism. We shall study the mechanisms of this process in the next chapters.

7.4.3 Activation–Transcription–Translation as a Feedback Circuit

Notwithstanding the possibility of errors, the genetic code is very reliable, since there is only one mistake (a mutation) in any 10^9 produced nucleotides.[78] DNA is error-free far longer than RNA because, representing only codified information and not a set of instructions, it allows the dissociation of transcription and replication,[79] and this prevents transcription from damaging the DNA template. Replication (the process through which DNA doubles itself, which is the basis for biological inheritance) is semiconservative (the original strand remains intact because in each round

[76][*LEVINE* 1992]. [77][BONHOEFFER/SNIEGOWSKI 2002]. See also [EIGEN 1993].
[78][*WAGNER* 2005, pp. 15–24]. [79][*DE DUVE* 2005, pp. 114–16].

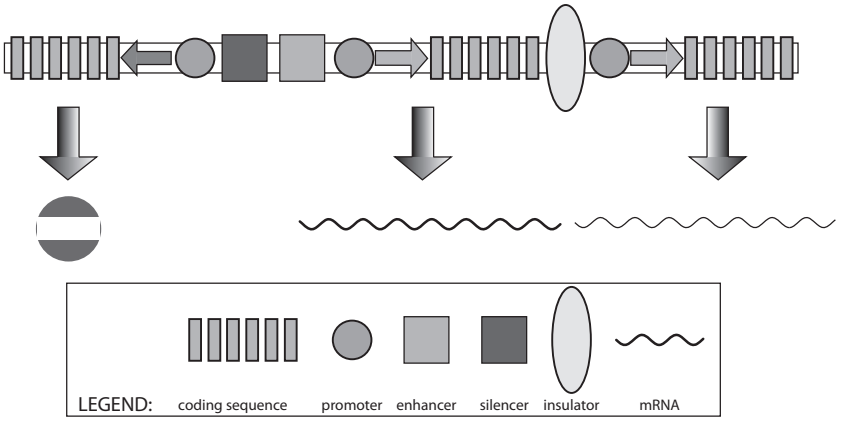

Fig. 7.20 Promoters, enhancers, silencers, and insulators along a strand of DNA: The four main elements involved in transcription (building an RNA sequence starting from DNA) together with gene regulatory proteins. The promoter contains specific DNA sequences and response elements that are recognized by proteins known as transcription factors. These factors bind to the promoter sequences, recruiting RNA polymerase, the enzyme, a nucleotidyl transferase that synthesizes the RNA from the coding region of the gene.

Some transcription factors (the so-called enhancer-binding protein) bind to regions of DNA that are thousands of base pairs away from the gene that they control. Binding increases the rate of transcription of the gene.

Silencers are control regions of DNA that, like enhancers, may be located thousands of base pairs away from the gene they control. However, when transcription factors bind to them, the expression of the gene that they control is repressed.

An insulator prevents an enhancer from inappropriately binding to and activating the promoter of some other gene in the same region of the chromosome.

of replication both strands are used): Since the direction in DNA replication is always 5′–*to*–3′, in the DNA fork there is a leading strand which is continuously reproduced, and a lagging strand which is discontinuously produced. In transcription-translation [Fig. 7.20], the nucleotides are arranged in an order corresponding to the order of amino acids in the protein that they specify.[80] Random errors also occur during DNA transcription.[81] Here, mechanisms of DNA repair are necessary, otherwise thermal fluctuations would cause major changes.[82] This is a typical negative-feedback mechanism.[83] In the simplest case, the RNA polymerase itself performs the correction during the transcription. The feedback signal is here represented by incorrect chemical affinities. Further correction mechanisms are also available once the transcription has been accomplished.

The whole process of transcription–translation can be considered in terms of feedback circuits[84] (here and in the following I shall mainly consider transcription and translation in eukaryotes). The transcription–translation–expression process can be summarized as:

(1) DNA activation for transcription is performed by RNA polymerase enzymes together with promoters, gene regulatory proteins, and eventually enhancers[85] [Fig. 7.21]. It binds very

[80][*GILBERT* 2006, p. 105]. [81][*WAGNER* 2005, pp. 25–38]. [82][*BERG et al.* 2006, pp. 804–12].
[83][*ALBERTS et al.* 1983, pp. 269–70]. [84][*MURRAY* 1989, pp. 143–8]. [85][*GILBERT* 2006, pp. 108–12].

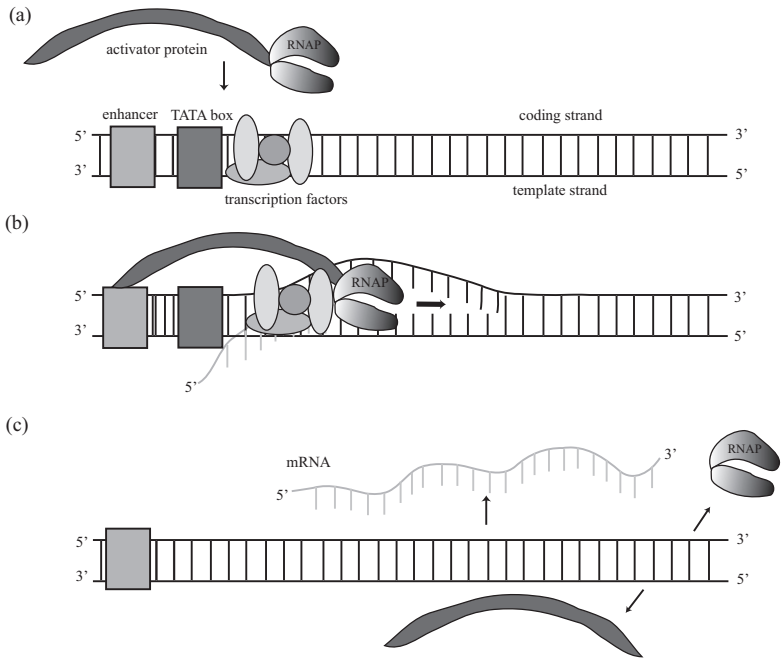

Fig. 7.21 (a) RNA polymerase (RNAP) with a protein tail, (b) activating a gene by acting on the region determined (in eukaryotes) by a complex of transcription factors or by a promoter (in bacteria), while attached protein acts as an enzyme on the enhancer region (other eukaryotic elements like mediators, chromatin remodeling complexes, and histone-modifying enzymes are not shown). (c) The final resulting mRNA. Only one of the two DNA strands is transcribed. This strand is called the template strand, because it provides the template for ordering the sequence of nucleotides in an RNA transcript. The other strand is called the coding strand, because its sequence is the same as the newly created RNA transcript (except for uracil being substituted for thymine).

The DNA template strand is read $3' \rightarrow 5'$ by RNA polymerase and the new RNA strand is synthesized in the $5' \rightarrow 3'$ direction until the stop signal. RNA polymerase binds to the $3'$ end of a gene (promoter) on the DNA template strand and travels toward the $5'$ end.

tightly when it contacts the promoter or the enhancer region, which contains the starting elements for RNA synthesis. Since three reading frames are possible (being codons triplets), there is always an initiation factor (AUG codon) and a stop codon (UAG).

(2) Then, we have *DNA transcription* into messenger RNA (mRNA). It is a process similar to DNA replication, since it starts with unwinding of the DNA helix in order to produce a single strand.

(3) Finally, the resulting mRNA, having left the nucleus through the nuclear pores toward the cytoplasm, can begin the *RNA translation* for producing the protein. In this way, mRNA is responsible for the transmission of the genetic code to the cytoplasm.

Since the polymerase and other proteins involved in activation are the result of a transcription–translation process themselves, we have the feedback circuit shown in Fig. 7.22. These feedback

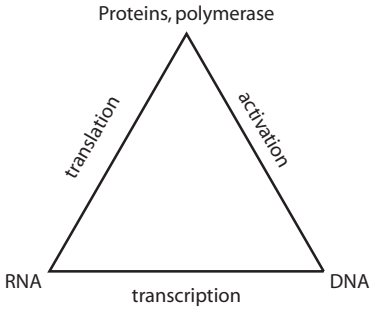

Fig. 7.22 The feedback circuit transcription–translation–activation (it runs clockwise).

circuits have special relevance during epigeny,[86] as we shall see below. I wish to recall here that DNA can function as an information processor only because it is activated through proteins and other chemicals. In itself, DNA is chemically inert. As mentioned, DNA only represents codified information that in its deep nature is pure potential [Subsec. 2.2.2].[87] Indeed, it has both coding and not coding sequences. RNA instead displays a set of active instructions for building a protein. The birth of DNA (in an RNA world[88] [Subsec. 7.3.1]) is due precisely to the fact that, being pure codified potential information, DNA can be activated or repressed, that is, DNA is much more controlled relative to RNA, which is always active and can therefore interfere with other operations.

Eukaryotic transcription could be seen as a threefold process: (1) First, we have the DNA. (2) Then, the information contained in the DNA is transcripted in a pre-mRNA, containing both exons and introns. (3) This step is then followed by RNA splicing that leads to the constitution of mature mRNA: The noncoding sequences, introns, are eliminated and only coding sequences, namely exons, remain.[89] This process can be done in very different ways for producing different mRNAs. Note that splicing has probably evolved from a self-splicing mechanism.

Also the whole contribution of RNA can be seen as a three-step process:

- The information contained in DNA is transcribed in mRNA.
- The translation of mRNA into a protein depends on an adaptor, the transfer RNA (tRNA), which has two extremes, an amino acid attachment and an anticodon that can be attached to the corresponding codon of the mRNA [Fig. 7.23]. It is a true mediator between information and function. Each type of tRNA can be attached to only one type of amino acid (but the same amino acid can be coded by different anticodons). The specific amino acid is attached to tRNA by an enzyme called aminoacyl-tRNA synthetase, which plays a simultaneous role in activation (an energetic issue) and translation (an informational issue).[90] The fidelity of protein synthesis is improved by two proofreading mechanisms.
- Protein synthesis happens on ribosomes. A ribosome is an organelle composed of ribosomal RNA (rRNA) and ribosomal proteins (known as a ribonucleoprotein or RNP), and with the help of tRNA, translates the information into a protein [Fig. 7.24]. The function of the rRNA is to

[86][*CHANGEUX* 2002, pp. 168–74].

[87]One of the first ideas about DNA as codifying potential information that needs to be activated (revealed!) can be found in [*MONOD* 1970, Ch. 5].

[88][HOLLAND/BLAKE 1987]. [89][*GILBERT* 2006, pp. 102–3]. [90][*DE DUVE* 2005, pp. 90–5].

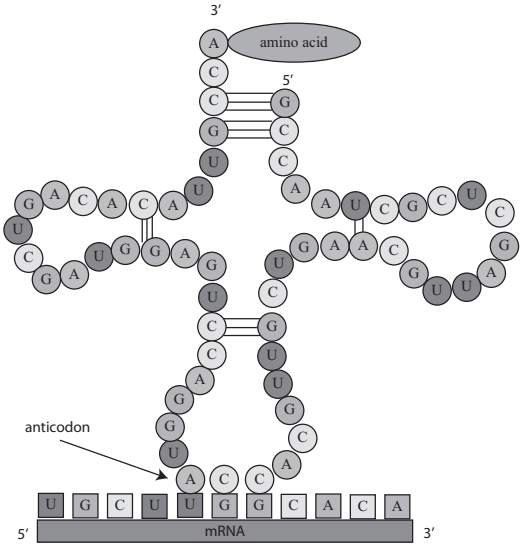

Fig. 7.23 A very schematic depiction of the tRNA structure. Some hydrogen bonds are shown.

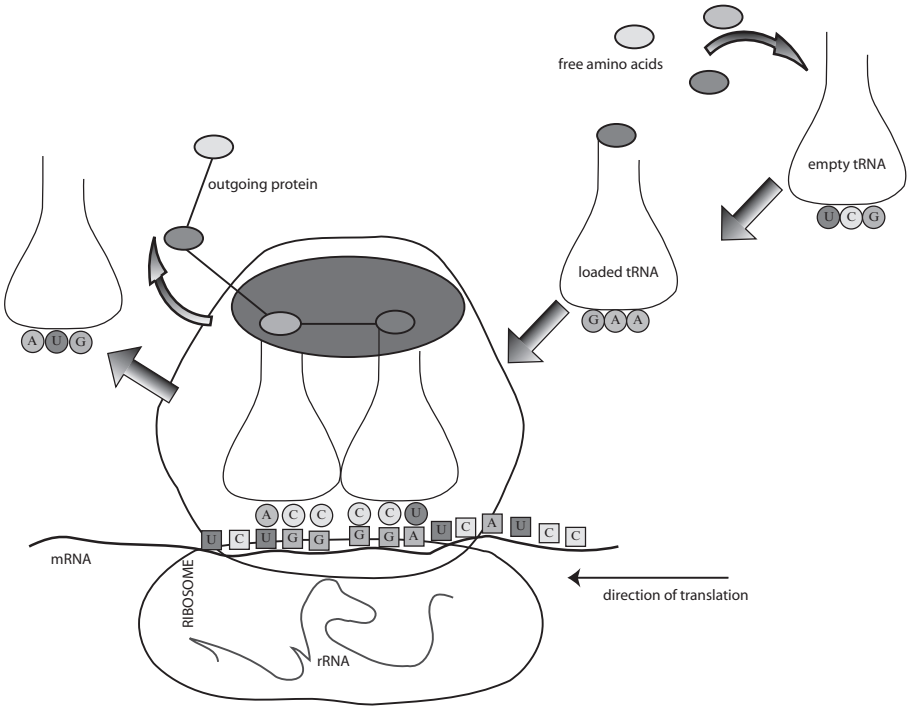

Fig. 7.24 How translation occurs. Incoming mRNA is read by tRNA loaded by an amino acid that joins the ribosome writing a triplet. Then it leaves the ribosomal complex. This is formed by a lower part made essentially of rRNA and an upper part made with a substantial contribution of proteins.

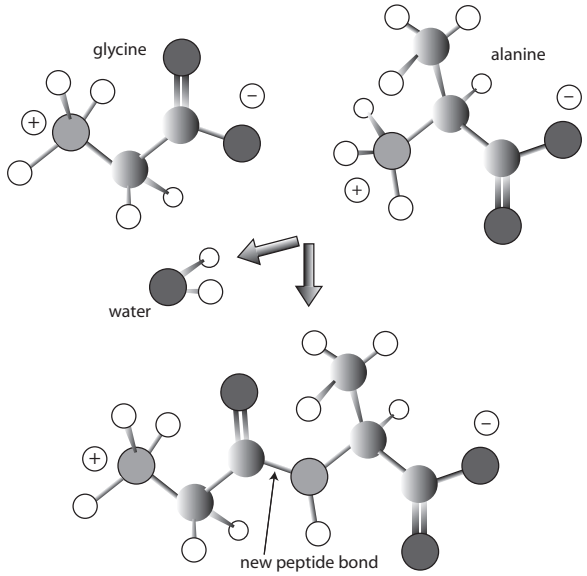

Fig. 7.25 A peptide bond is a covalent bond [see Fig. 6.6] formed between the carboxyl group of the previous amino acid and the amino group of the subsequent amino acid [see Fig. 7.4(c)]. It releases water. Note also the involved polarity. In the figure, the junction between glycine and alanine is shown. Hydrogen atoms are shown in white, carbon atoms in grayscale, nitrogen atoms in light gray, oxygen atoms in dark gray).

provide a mechanism for decoding mRNA into amino acids and to interact with the tRNAs during translation by providing peptidyl transferase activity (the enzymatic function that forms peptide links between adjacent amino acids using tRNAs during the translation process). Protein synthesis is the most free-energy consuming process in the cell [Subsec. 7.3.2]. The whole process synthesized here can be thought of as a kind of factory that builds a protein from a set of genetic instructions.

Ribosome's RNA (rRNA) originally may have served as the entire ribosome (therefore it could be a remnant of a very early stage in evolution). As a matter of fact, a ribosome is built from pieces coming from the inside of a specific part of the cellular nucleus (called the nucleolus), starting from a piece of rRNA to which several proteins are added in a complex process of cytoplasmic assembling and discarding.[91]

7.4.4 Proteins

Proteins are the most complex macromolecules that are known [see Tab. 7.1]. They show[92]

- A primary structure made of amino acids arranged in a linear chain and joined together by peptide bonds between the carboxyl and amino groups of adjacent amino acid residues [Fig. 7.25]. The linear sequence of amino acids is built according to the information translated by the mRNA.

[91] [*ALBERTS et al.* 1983, pp. 363–5]. [92] [*BERG et al.* 2006, pp. 25–59].

- A secondary structure: A bidimensional complex, consisting of regularly repeating local structures stabilized by hydrogen bonds. The most common examples are represented by the alpha helix and beta sheet: The alpha helix (α-helix) is a right-handed conformation, resembling a spring, in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues back. The β–sheet consists of beta strands laterally connected by three or more hydrogen bonds, forming a generally twisted, pleated sheet.
- The tertiary structure: The overall shape of a single protein molecule. The hydrophobic side is pushed into the interior of the protein and the polar side chains are arranged near the outside of the protein (meaning high reactivity). Folding happens in a specific environment (cytosol) where several parameters (temperature, salt concentration, and so on) must be appropriate. The process is often assisted by specialized proteins called chaperons (again a form of checking and control). Moreover, proteins may also be covalently modified through the attachment of functional groups other than amino acids (and therefore not resulting from the amino acid codification process).[93]
- The quaternary structure: When several proteins are joined for performing certain functions.

The information for starting many of the complex assemblies of macromolecules in cells must be contained in the subunits themselves, since under appropriate conditions the isolated subunits can spontaneously assemble in a test tube, to give rise to the final structure (for example the tobacco mosaic virus, where dissociated RNA and proteins subunits assemble spontaneously[94]). Purified proteins will generally refold properly, and this does not happen randomly: In fact, there are a vast number of possible conformations for any large protein that can be explored in the few seconds that are typically required for folding. Evidently proteins have also been selected for their ability to fold quickly. This also means that the protein represents an increase in complexity relative to the genetic information.[95] This is the problem of the cellular epigeny, as we shall see.

Proteins have very different specific properties and functions: For instance, collagen has enormous tensile strength, while elastin is very elastic. In principle, 10^{390} proteins could be made (a number much larger than the total number of particles of our universe, which is considerably less than 10^{100}), but only a small fraction would adopt a stable three-dimensional conformation, in accordance with the analysis developed in Subsec. 6.5.1. The wide majority would have many different conformations of roughly equal energy, each with different chemical properties [Subsec. 6.2.4]. Proteins with such variable properties would not be useful (in fact, as mentioned, proteins have very *specific* catalytic or structural functions that are crucial for information control, as we shall see). In other words, the interconnections (a problem of compatibility and stability) do not allow a continuum of possibilities, as is the case for classical mechanics. This amounts to saying that many constraints are at work in protein building.

Proteins have moving parts that are precisely engineered, whose mechanical actions are coupled to chemical events. The shape of a protein can be altered (allosteric transition) by binding to another molecule (the ligand). Two ligands which bind to the same protein affect each other's binding (enhancing or competing). Protein phosphorylation (a phosphate group becomes covalently coupled to another molecule) is a common way of driving allosteric transitions in eukaryotic cells (we shall see the importance of this for cell signaling). The structure of the cyclin-dependent protein kinase (Cdk)—a protein kinase is an enzyme that transfers the terminal phosphate group of ATP to a specific amino acid of a target protein—shows that a protein can function as a microchip (in

[93][*BERG et al.* 2006, pp. 57–8]. [94][NAMBA *et al.* 1985]. [95][*BARBIERI* 2003].

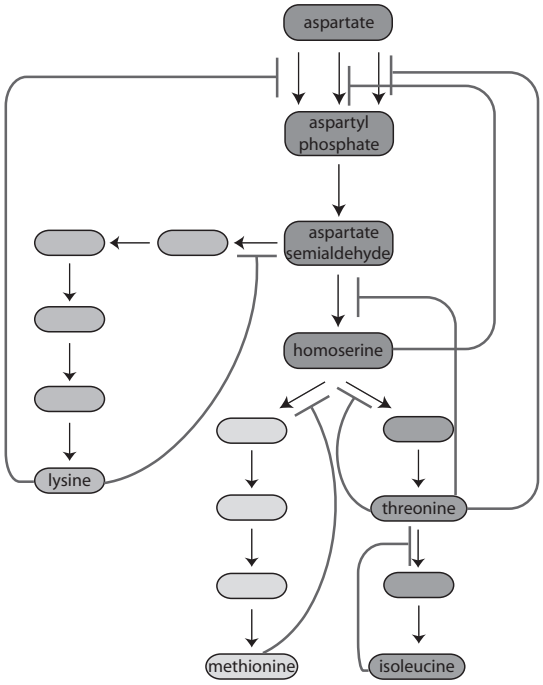

Fig. 7.26 Example of a protein feedback circuit: The dark gray interconnections represent inhibitory feedback. In this way some proteins are able to "control" their own production. Inspired by [*ALBERTS et al.* 1983, p. 170].

a vertebrate cell, individual Cdk enzymes turn on and off in succession as a cell proceeds through the different phases of its division cycle).

In general, proteins are the operators of the cell: They guarantee the lower level of performance relative to both the higher level of the control instances and the middle level of regulative instances (we shall consider this very complex mechanism later on). Proteins which do mechanical work use ATP for a work cycle, but only in one direction, otherwise a protein would walk randomly back and forth. Kinesin motors are specialized enzymes which generate force and movement along their cellular tracks, the microtubules.[96] Under this respect, proteins could be considered as true molecular machines. However, only a minor part of proteins is used for work: The rest of protein molecules is necessary for structural purposes, maintaining the polypeptide chain in the correct position, and regulation [Fig. 7.26].

7.4.5 Information and Function

When dealing with the whole process generating a protein, we must distinguish here between catalyzed assembly and codified assembly:

[96] [KIKKAWA *et al.* 2001].

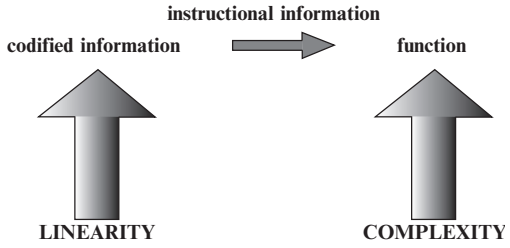

Fig. 7.27 From information to function: The two roots of biology: Linearity and complexity. Although complexity is also relevant for DNA and RNA, it is constitutive for any protein and its function.

- A *codified assembly* occurs during DNA transcription. In this case, a string containing linear information (mRNA) is produced that must transport (to the ribosome) and translate the instructions it carries.
- A *catalyzed assembly*, instead, occurs at the ribosome. Here the end-product of the process is no longer a linear string of information. It is true that the primary structure of the protein (as it comes out of the ribosome) is linear (as it still has the mark of the informational process started by the mRNA). However, the function of a protein is *not* to transmit instructions or to play an informational role [Fig. 7.27]. Notwithstanding the specific function of any protein, it does not pertain to the informational system of the organism but to metabolism: Its basic function is to perform or to help to perform work or to have a structural role, even when it is involved in typical informational activities like cellular transduction or genetic expression.

Therefore, in the transcription–translation process from DNA to the protein, we already have the unity of informational and metabolic processes that is the hallmark of organisms [Sec. 7.2]. This also raises a problem: There is no particular relation between the codified information in the DNA and the final protein displaying a function or *vice versa*. This point was well understood with Maynard Smith by reference to J. Monod[97]: He considered the example of inducers and repressors that show no necessary connection between their form (chemical composition) and their function (genes switched on or off). Moreover, to take another example, the triplet CAC codes for histidine, but there is no chemical reason why it should not code for glycine. However, the apparent arbitrariness, i.e. locality, of the codification is a prerequisite for speaking of information, since it is necessary that information could be codified otherwise elsewhere or in another context [Subsec. 7.4.1]. Nevertheless, the fact remains that we cannot easily account for a connection between codified and catalyzed assembly. To solve this difficult problem, let us first consider why a protein cannot be understood as an informational entity.

In DNA, apart from the backbone, we only have hydrogen bonds between informational molecules (bases). In tRNA, we have a chain of informational molecules plus some hydrogen bonds between them. It is important to understand that RNA can already show hydrogen bonds that do not fulfill the requirements of information codification.[98] RNA is indeed a macromolecule that can also have a function like a protein (and in the first stages of life, in the so-called RNA world [Subsec. 7.4.3], this aspect was much more important). The primary structure of a protein is

[97][MAYNARD SMITH 2000]. In [*MONOD* 1970, Ch. 1], the concept of teleonomic information is introduced as the quantity of information that must be transmitted in order for a certain structure or performance to be realized.
[98][*BERG et al.* 2006, p. 116].

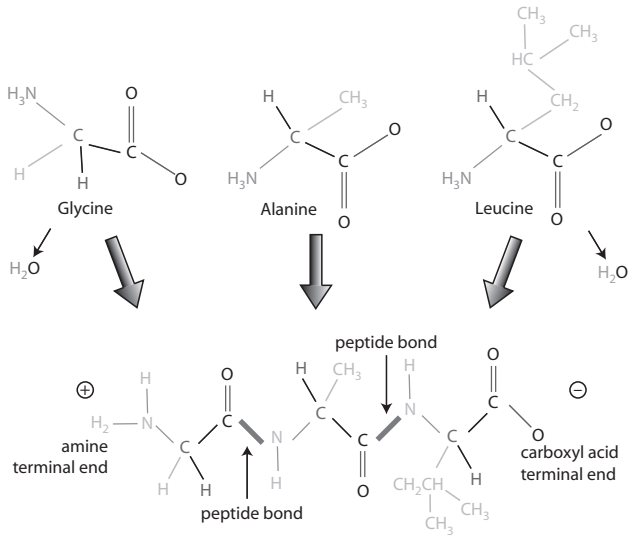

Fig. 7.28 The constitution of a three-amino acid sequence. Color conventions are those of Fig. 7.4(c): The part in blue is the amino group, the black carbon atom is the central atom with an added hydrogen atom (also in black), the carboxyl acid group in red, and the side chain in green. Note the repetition of the structure on the left and on the right of the peptide bonds until the end of the carboxyl group: A H–N molecule is connected to the central carbon atom which, apart from the side chain, is in turn connected to both or hydrogen atom and a O=C molecule. (This figure is reproduced in color in the color plate section.)

constituted by single peptides connected by peptide bonds, i.e. a carboxyl group of a molecule reacts with the amine group of another molecule releasing a water molecule [Fig. 7.25]. It is evident here that the only element that distinguishes the different segments is the side chain [Fig. 7.28]. Now, the point is that the structure appears as a whole whose segments can no longer be easily detached and recombined as we expect from true codified information. It is indeed the latter property that allows DNA and RNA to be good chemicals for dealing with information [Fig. 7.15]. The problem is even bigger in the final stage, where the protein folds thanks to non–covalent bonds (hydrogen bonds and van der Waals forces) [Fig. 7.29], a true self-organization process.

Given the previous point, the question is now whether or not it is possible to give rise to complex composed structures (the proteins) that do not themselves display informational character but a dealing-with-information (in their regulative or genetic-expression functions) whose *start* is with codified information (DNA). Fodor and Pylyshyn assumed that there is no dealing-with-information at all where there is no information codification because one could not compositionally employ structured elements having informational value.[99] However, one should distinguish between *concatenative* combinatorics, typical of the classical information approach, and a merely structural

[99][FODOR/PYLYSHYN 1988]. This criticism was actually addressed against connectionism and in particular against the idea that it is possible to give rise to representations in this way.

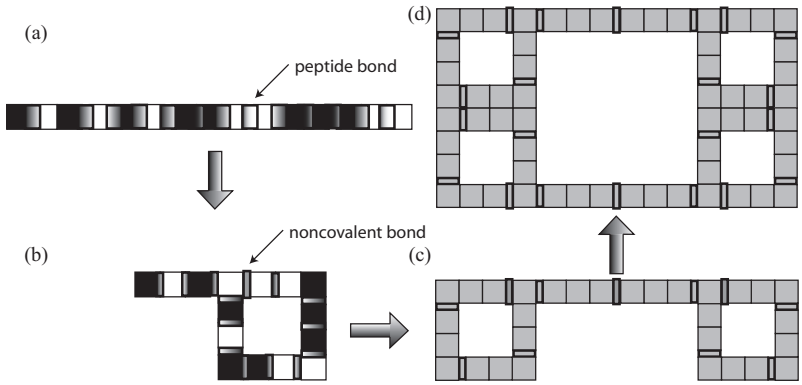

Fig. 7.29 The functional-combinatoric arising of hierarchically nested structures in life. I have shown here the building of an idealized protein.

(a) The process starts with a linear (primary) structure. For the sake of simplicity I have considered a binary code represented by black and with boxes representing triplets (peptides). The peptide bonds are shown explicitly.

(b) The protein folds thanks to noncovalent bonds (for the sake of simplicity, I have considered only a bidimensional (secondary) structure). Here the single boxes are no longer relevant, since these bonds can be established at very different places according to several constraints depending on the overall shape constitution.

(c) Several proteins are connected giving rise to quaternary structures. Here, the single triplets do not have meaning since it is only the points in which these bonds can be established that are relevant.

(d) When a higher complex structure arises (like the RNA polymerase) performing specific operations, it is only the overall shapes and functionality that matters.

or *functional* combinatorics.[100] Concatenative combinatorics preserves tokens of an expression's constituents (and the sequential relations amomg them) in the expression itself due to the linearity of information codification [Subsec. 2.2.2]. Instances of this variety are represented by natural and formal languages (mathematics, logic, and so on). The reason why they are so common is that they could *also* be functionally compositional (as RNA aptly shows). But most functionally compositional schemes are only functional and *not concatenative*. An example is given by a recursive autoassociative memory: When representing a tree, it is possible to store a part of it (a branch) in a stack separately, another in another stack, and so on. In this way, the whole tree can be represented recursively. The basic model is here the puzzle. This is a specific modality of representation. Another example is represented by Smolensky's vectorial representation. If each part of an object can be represented by a vector, the whole object can be represented by a product of vectors: Instead of using the outer product, in which the product of two vectors gives rise to a matrix [Eq. (1.17)] (a matrix always allows for the finding of its eigenvectors), we can use the vectorial product, giving rise to a new vector which can in turn be used recursively. For instance,

[100][VAN GELDER 1990].

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}. \tag{7.6}$$

Starting from the resulting vector we can no longer ascertain which were the single primary constituents (the $a_1, a_2, a_3, b_1, b_2, b_3$) of the two original vectors. The essence of functional compositionality is therefore the assemblage of constituent tokens. Obviously, while concatenative compositionality is infinite in potency, functional compositionality is not. Nevertheless, in this way it is possible to deal with functions that arise from some informational activity and can have further effects on the same informational activity contributing to information control, as we shall see. Moreover, in this way structures having informational value (DNA or RNA) and structures without informational value (the primary structure of a protein) may show important mapping. Indeed, these structures can share information [Subsecs. 2.3.3–2.3.4 and Sec. 6.4] without necessarily representing codified information themselves.

As we shall see, the connection between codified and catalyzed assembly is due to the fundamental semiotic nature of life, which implies that there is never a direct, Lamarckian instruction[101] of other biological systems or even subsystems *even inside* the same biological system. In other words, proteins or even a whole phenotype are a black box relative to the genetic system. The semiotic nature of the process is precisely due to the fact that the whole expression–transcription–translation process (as well as any other aspect involving information in the organism) is from the start both *metabolic* (because organisms are complex systems displaying functions) and *informational* [Secs. 6.3–6.4]. This, however, does not imply (as J. Maynard Smith believed) that genetic information is not codified. Indeed, the (semiotic) *use* of this information for giving rise to specific functions presupposes the role of the whole genetic system as information processor (and therefore also as codifier) [Subsec. 6.4.2].

7.4.6 An Information-Processing Mechanism

Genes (portions of DNA controlling the expression of a genetic character) can be activated or disactivated by a complex constituted by RNA polymerase, promoters, enhancers, and enzymes. This mechanics also has a central importance in epigeny, as we shall see. Let us consider the feedback circle introduced in Subsec. 7.4.3 from a pure informational point of view[102] in which [Subsec. 2.3.2]:

- The DNA codes the information (structural and potential information), and expresses it thanks to proteins and polymerases. It is a pure information-activation (communication) stage in a system, where the random mutation can be considered to be a message variation and splicing as an information-selection substep in the framework of this information communication.
- The RNA (mRNA, tRNA, and rRNA) ensures that the necessary bridge for this information is used further. Actually, it is a true interface that, through splicing and reshuffling of coding sequences (in eukaryotes), modulates information and establishes a bridge to the final product. The mRNA incorporates the set of instructions to be used (instructional information). The tRNA is the interface between instructional information and the building site (rRNA) of a new structure (the protein).

[101] [*LAMARCK* 1809, I, p. 235]. [102] [CRICK 1970].

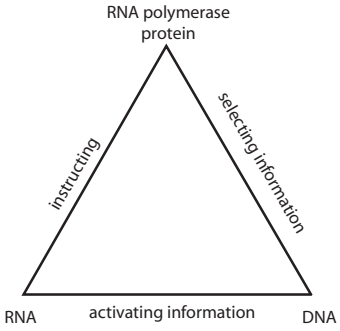

Fig. 7.30 Genetic system. It is a true information-transferring system (it runs clockwise).

- The outgoing proteins together with the RNA polymerase and promoters are the functional unities that select which part of the DNA will be further activated. This is *not* a transfer of information.

In this way, a single gene can act on a lot of genes, activating and inactivating them. The whole system constitutes a feedback circle [Fig. 7.30; see Fig. 7.22]. It is a global information-processing mechanism: As I have stressed, the genome alone is not in itself an information processor because it is inert, and the genetic *system* is made of DNA, RNA, and protein/polymerases.

Another point that I wish to stress is that it is an *irreversible* information-processing mechanism. Such irreversibility is due to information erasure [Subsecs. 2.2.1 and 3.2.2]. Also the normal biochemical mechanism by which RNA is destroyed when it is no longer needed is irreversible.[103] As indicated before, the synthesis of RNA by RNA polymerase is a logically reversible copying operation, and under appropriate (nonphysiological) conditions, it could be carried out at an energy cost of less than k_BT per nucleotide [Subsec. 2.3.3]. In principle, the thermodynamically most efficient way to get rid of an RNA molecule would be to reverse this process, i.e., to take the RNA back to the DNA from which it was made, and use an enzyme such as RNA polymerase with a slight excess of pyrophosphate to perform a sequence-specific degradation, comparing each RNA nucleotide with the corresponding DNA nucleotide before splitting it off. In this case, we would not have a feedback circuit but a reversible circuit not wasting energy. This process, however, does not occur in nature; instead RNA is degraded in a nonspecific and logically irreversible manner by other enzymes, such as polynucleotide phosphorylase. This enzyme catalyzes a reaction between an RNA strand and free phosphate (maintained at high concentration) to split off successive nucleotides of the RNA as nucleotide phosphate monomers. Because the enzyme functions in the absence of a complementary DNA strand, the removal of each nucleotide is logically irreversible. If it were to run backwards (in the direction of RNA synthesis), the enzyme would be as likely to insert any of the three incorrect nucleotides as if it were to reinsert the correct one. This informational irreversibility means that driving the reaction forward needs a phosphate concentration fourfold higher than it would be if it were required by a sequence-specific degradation. It is indeed this difference in the phosphate needed that keeps the enzyme from running backwards and synthesizing random RNA, but it also means that the cycle of *specific* synthesis followed by *nonspecific* degradation must

[103][BENNETT 1982].

waste about 1.4 $k_BT = k_BT \ln 4$ per nucleotide even in the limit of zero speed. For an organism that has already spent around $20k_BT$ per nucleotide to produce the RNA with near maximal speed and accuracy, the extra 1.4 k_BT is obviously a small price to pay for being able to dispose of the RNA very quickly, without taking it back to its birthplace. Moreover, keeping the operations of transcription and activation separated and only letting proteins and RNA polymerase perform the latter function, allows for an increase in information control relative to an RNA back-transcription. As we shall see, the latter operation is indeed possible, but less reliable than the mechanism described here [Sec. 7.2]. Summarizing, the information erasure—which determines informational irreversibility [Subsecs. 2.2.1, 2.3.2, and 2.4.1]—consists here in the fact that structures having informational value, like RNA, are destroyed, and therefore their informational value is erased.

7.5 Self-Reproduction

The relevance of self-reproduction and its very basic role for organisms justifies a specific treatment here. The basis of the modern approach to the problem of self-reproduction is due to W. Johannsen's distinction between genotype and phenotype [Sec. 7.1]. It is interesting to observe that in the original formulation, this distinction would acknowledge the importance of epigenetic interactions and processes in order to determine a phenotype.[104] Recall also that self-reproduction involves both the genetic and metabolic system of an organism.

7.5.1 Self-Replicating Automata and Artificial Life

The feature of life that has mostly impressed human beings throughout the ages is self-replication. Actually, at the beginning of the cognitive revolution, self-reproduction was thought of as the most salient aspect of a living system, and perhaps even as what defined life. Von Neumann tried to reproduce this feature of life by building self-replicating automata.[105]

This line of research started with the distinction between genotype and phenotype and proposed to generalize it into a distinction between a specification of machinery (the generalized genotype: GTYPE) and the behavior of that machine (the generalized phenotype: PTYPE). The PTYPE is a nonlinear function of the GTYPE and is a multilevel, hierarchical phenomenon. Therefore, there are two types of information contained in the description of a cellular automaton:

- Information which is interpreted, i.e. the set of instructions to be executed in the construction of offspring, and
- Uninterpreted information, i.e. passive data to be duplicated to form the description (program) transmitted to the offspring.

Otherwise, there would be an infinite regress, where the blueprint contains a miniature version of the blueprint, and so on [Sec. 7.1]. This conceptually justifies the functional distinction between DNA and protein as well as between DNA and RNA.

The greatest problem in this field is to distinguish a trivial proliferous machine from a significant machine, and nobody has yet devised a satisfactory answer, even if it is known that interesting reproducible patterns can emerge.[106] It is certain that a self-replicating machine must be complex,

[104] [*JABLONKA/LAMB* 1995, pp. 16–17].

[105] [*VON NEUMANN* 1966]. See also [LANGTON 1989b]. Mange and Sipper tried to show that Barbieri's theory is anticipated by von Neumann's work: The ribotype is the universal constructor [MANGE/SIPPER 1998].

[106] [SIPPER/REGGIA 2001]. See also [REGGIA *et al.* 1993, CHOU/REGGIA 1997].

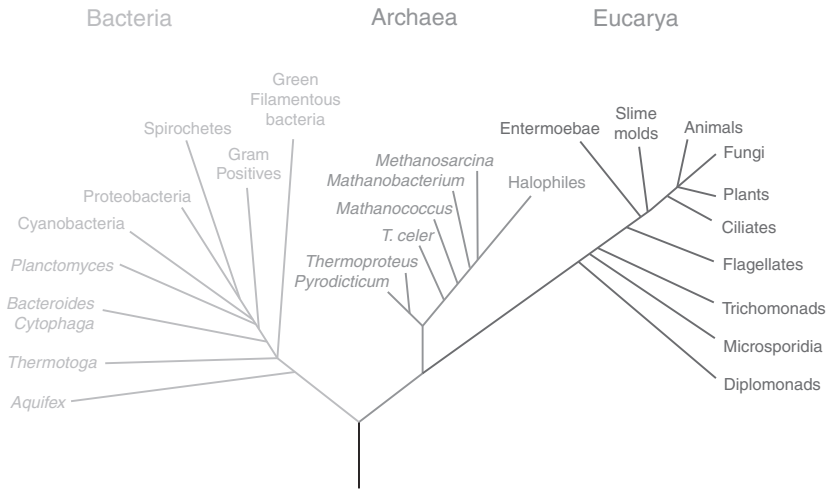

Fig. 7.31 Phylogenetic tree of life.

due to the hierarchical constitution of the phenotype. A major step was taken in 1984 when Langton observed that looplike storage devices could be programmed to replicate on their own. The device consists of two pieces: The loop itself and a construction arm.

This gave rise to an interesting bridge between the science of complexity and the sciences of life, namely so-called artificial life (AL), which is an attempt at artificially reconstructing some basic features of life, in particular self-organization and self-reproduction. This can throw light on some basic elements of life and perhaps help to distinguish it from artificial systems. AI has focused on the production of intelligent solutions without considering the relationship between the production and the way in which intelligence is generated in natural systems [Subsecs. 6.1.4–6.1.5], while AL is interested in spontaneous production of cognitive behavior[107]: Here, natural systems are essentially conceived of as parallel and distributed engines.

In 1997 Chou and Reggia investigated a primordial soup of randomly selected components in order to see if self-replicators could emerge spontaneously. They noticed that, above a certain threshold, small self-replicating loops appeared. Meanwhile, important steps have been taken for self-replication from assembling chaotic building blocks and a programmed error-correction procedure.[108]

7.5.2 Mechanism of Inheritance

It is time now to have a look at the basic mechanism of biological heredity at work in natural systems. First let us have a short look at the most general classification of living beings. This can turn out to be useful for dealing with the problems of this section. All living beings, according to Carl Richard Woese,[109] can be divided into prokaryotes and eukaryotes. Prokaryotes are further divided into bacteria and archaea, so that the whole gives rise to the so-called three-domain system [Fig. 7.31]. Prokaryotes lack a cell nucleus while eukaryotes have one. Eukaryotes also have a specialized structural development, constituted by the cytoskeleton (an array of proteins) for cells.

[107][LANGTON 1989b]. [108][GRIFFITH *et al.* 2005]. [109][WOESE *et al.* 1990].

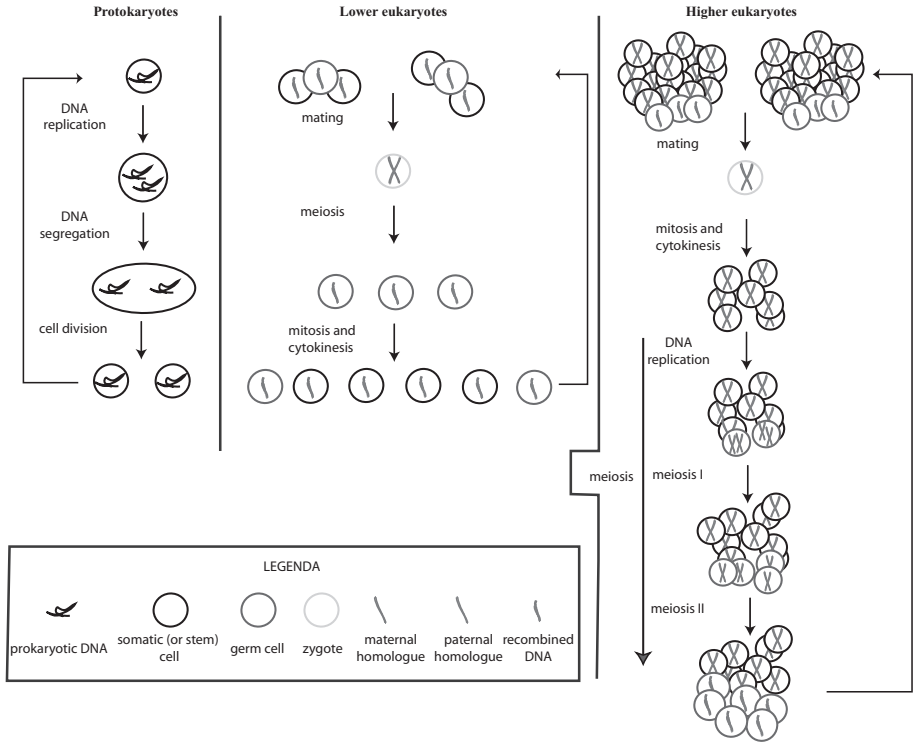

Fig. 7.32 The self-reproduction mechanisms. Multicellular organisms are represented as unorganized assemblies for the sake of simplicity. (The figure is reproduced in color in the color plate section.)

Prokaryotes follow binary fission.

Lower eukaryotes are essentially haploid organisms. Here, the case of multicellular sexual organisms is shown. However, low eukaryotes can also be unicellular and many eukaryotes are asexual. In this case, there is no longer a distinction between germ and somatic cells even if we still have the cycle shown in the central part of the figure.

Higher eukaryotes are diploid organisms.

Archaea are similar to bacteria in most aspects of cell structure and metabolism, but their genetic transcription and translation does not show the typical bacterial features, thus rendering them extremely similar to those of eukaryotes.[110] For this reason, today it is widely acknowledged[111] that eukaryotes and archaea have split successively to their separation from bacteria.

Biological hereditary transmission occurs either through a cell division (cytokinesis or binary fission) or through sexual mating and a recombination of the genetic pool [Fig. 7.32].[112] There are two types of genetic recombination: General recombination, when genetic exchange takes place between any pair of homologous DNA sequences (for instance, during meiosis), and site-specific recombination, which occurs in short, specific nucleotide sequences and DNA homology is not required (typical for viruses). The first form is a crossover. Although meiosis is typical for eukaryotic

[110][BARRY/BELL 2006]. [111][PACE 2006]. [112][*ALBERTS et al.* 1983, pp. 1053–113, 1269–304].

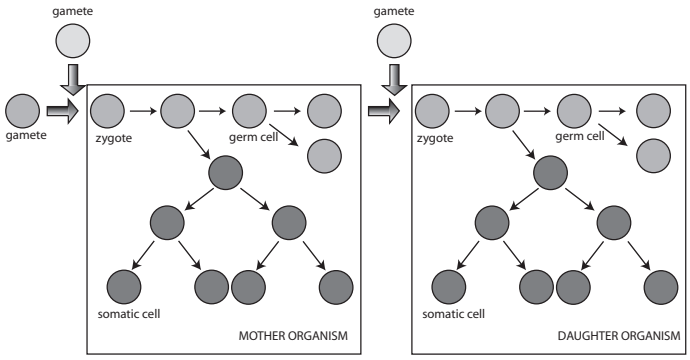

Fig. 7.33 Somatic and germ cells.

cells, bacteria also show a mix in which general recombination plays a role too. Therefore, life follows two main strategies for hereditary transmission:

- The first strategy, i.e. *binary fission*, characterizes prokaryotes. It is interesting to observe that, in order to duplicate itself a cell must first replicate its DNA and then segregate the two sets [Secs. 6.6 and 7.1]. In this way, it comes to a sort of quasi–multicellular organism, or at least something that is between a unicellular and a multicellular organism. It is a very common fact in evolution that new forms and solutions come out as stabilization of previous transitory processes. In this way, even if additional factors are necessary for explaining the transition from unicellular to multicellular organisms, this quasi-multicellular behavior in unicellular organisms can already show that there is a certain bridge. On the other hand, higher eukaryotes preserve particular unicellular stages, especially during meiosis and fertilization.[113]
- The second strategy is followed by sexual eukaryotes. Since haphazard genotypes produced by genetic recombination are as likely to represent a change for the worse as a change for the better (to a more survival-promoting form), then sexual reproduction helps a species to maintain a high level of species-specific *variability* (that is, more controlled than horizontal gene transfer, a typical bacterial strategy) and therefore to survive quite well in an unpredictably variable environment, without relying on pure chance. This is the evolutionary reason for sexual reproduction. It is appropriate to distinguish here between haploid and diploid organisms. A *diploid* organism has a complete set of homologue chromosomes coming from the mother and a complete set of chromosomes coming from the father, excluding the case of germ cells. This means that a diploid organism also has two copies of every gene, one from the mother and one from the father. A *haploid* organism, instead, only has a single set of chromosomes that is a combination of its father and mother's genetic pools. The different variants of a specific gene are known as *alleles*. If an organism inherits two alleles that are at odds with one another, and the phenotype of the organism is determined completely by one of the alleles, then that allele is said to be *dominant*. The other allele, which has no tangible effect on the organism's phenotype, is said to be *recessive*. This allows us to distinguish between two cases:

(1) *Lower eukaryotes* are haploid (some still unicellular) organisms. Sexual eukaryotes, when mating, give rise to a diploid zygote (the fertilized egg). Here, we must further distinguish

[113][*BONNER* 2000, p. 49].

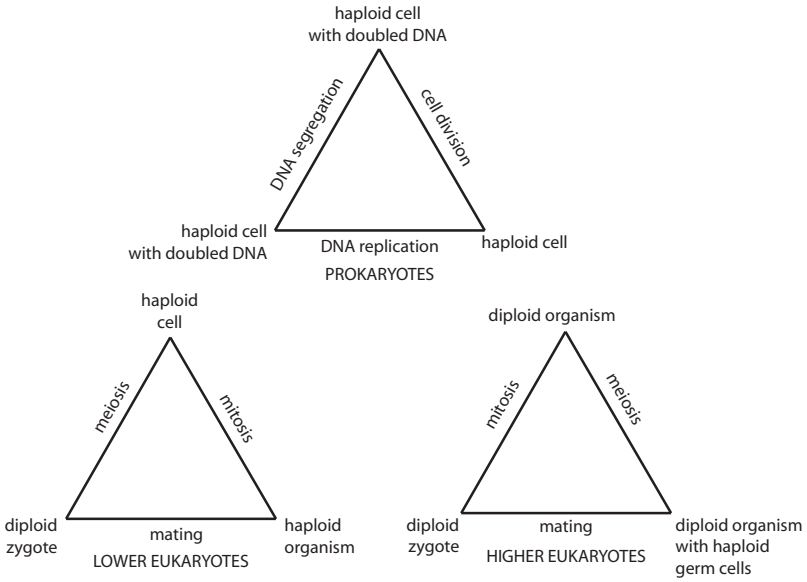

Fig. 7.34 The feedback circles of self-reproduction. All cycles run clockwise. It is interesting to observe that meiosis and mitosis are inverted in higher eukaryotes relative to lower sexual eukaryotes as far as sexual maturation in the former occurs at a relative later stage (during the developmental phase successive to epigeny); take into account that meiosis only concerns germ cells, as shown in Fig. 7.32.

between somatic and germ cells [Fig. 7.33]: Germ cells or gametes (oocytes and sperm) in all multicellular organisms are haploid, while somatic cells are diploid only in higher eukaryotes. Lower eukaryotes, through *meiosis*, recombination, and halfing of the genetic material,[114] produce both somatic and germ haploid cells from a diploid zygote.[115] Note that some very rudimentary eukaryotes like amoebae alternate sexual and asexual reproduction.

(2) *Higher eukaryotes*, instead, are *diploid* organisms: Having separated haploid germ cells and diploid somatic cells, they give rise, through mating, to the fusion of a haploid egg and a haploid sperm that builds a diploid zygote (as is already the case for lower eukaryotes). From the zygote, somatic cells are generated through mitosis (i.e. the nuclear division giving subsequently rise to daughter cells with the same parental genome), and then germ cells are generated through meiosis. In this way, a whole organism is constituted, having both somatic and germ cells. Sex determination occurs in several ways: By temperature acting on eggs (as for some turtles or alligators) or by chromosomal differences related to the chromosome X. For instance, in the *Drosophila* it is the ratio between the number of X chromosomes to the number of autosomal sets, a ratio that is determined during RNA splicing.[116]

[114]For a review of meiosis see [NEALE/KEENEY 2006].

[115]This separation between germ and somatic cells has led Weismann to state that the reproduction of multicellular organisms is essentially similar to the corresponding process in unicellular forms, consisting in the continual division of the reproductive cell, the only distinction being that in the first case we have two classes of cells [*WEISMANN* 1889, I, p. 75].

[116][*ALBERTS et al.* 1983, pp. 481–2].

Summing up, in all (lower and higher) eukaryotes the whole cycle can be considered as an alternance between haploid and diploid cells. In this way, we ultimately have three feedback circles, one for prokaryotes, and two for eukaryotes [see Fig. 7.34].

The transition from prokaryotes to eukaryotes allows a considerable increase in information control, especially through modularization of many crucial processes, as explained before [Sec. 7.2]. The process of mitosis in eukaryotes is indeed an interesting example of information control.[117] The problem is to ensure that each daughter cell receives one and only one homologue copy of each duplicated chromosome. This equal distribution is guaranteed by a checkpoint system delaying cytokinesis until the duplicated and paired homologues are aligned along the metaphase plate and attached by microtubules to opposite spindle poles. Proper alignment and attachment then leads to the distribution of one homologue to each daughter cell. When there is no proper alignment and attachment, chromosomes emit chemical signals that are interpreted by the cell-cycle control network and the homologue separation operator as a wait signal.

It is important to stress that prokaryotes and eukaryotes display another important difference that is the consequence of those previously shown. Bacteria tend to replicate indefinitely as a consequence of the accumulation of their growth [Sec. 7.1]. Some bacteria have a reproduction cycle of 20 minutes. This means that each 20 minutes 2 bacteria are born from a mother cell by binary fission. In 11 hours (33 intervals of 20 minutes) there would be 8,589,934,592 (2^{33}) individuals. In one week, starting from a single bacterium, the total mass represented by 2^{504} individuals would largely exceed the that of whole biomass of the Earth and even the mass of the Earth. Obviously there are external (environmental conditions) that damp such an exponential growth of the population: This is due to the fact that bacteria have a metabolism that needs certain environmental resources which are necessarily limited. Instead, eukaryotes have developed population growth regulation mechanisms that are also *internal* (again an increase in control). These are both more sophisticated forms of metabolism and in some cases constraints that are related to both mating and developmental issues. As a matter of fact, the progeny of eukaryotes are always subject to limits in number and to time windows of reproduction. The fact remains, however, that life shows a tendency not only to the growth of the individual but also of the population. This is the basis of the competition between individuals within the same population and therefore of natural selection.[118]

7.6 A Selection System

7.6.1 The Membrane

After having considered the metabolic and the genetic subsystems, we need to understand the third piece of the mosaic, the selecting subsystem. In the easiest case of unicellular organisms, this subsystem is constituted by a membrane and its annexes. The DNA does not code for the membrane, but only for the proteins building or regulating the membrane (and obviously for the gate proteins in the membrane). The membrane, playing the function of a decider (either allowing or preventing something to go in or out of the cell), has been evolutionarily built from lipidic material through pure physical mechanisms and, in the transmission to further generations, it is not subjected to genetic variability (it is a piece of physics recruited by life as an instance of generalized Darwinism [Subsec. 2.2.6]): *Omnis membrana e membrana*.[119] In other words, the lipids constituting the membrane are the only chemicals that do this work and therefore have

[117][SHAPIRO 2002]. [118][*DARWIN* 1859, p. 55]. [119][BLOBEL 1980]. See also [*DE DUVE* 2002, p. 36].

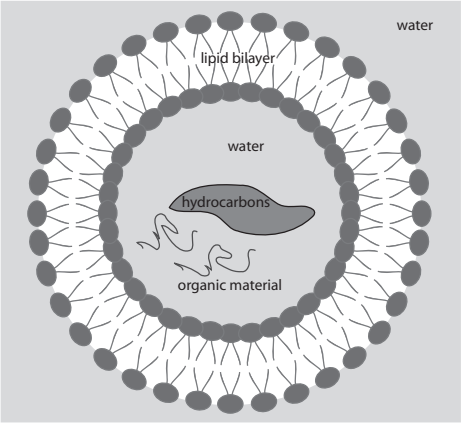

Fig. 7.35 Spontaneous formation of sack-like vesicles with organic chains inside.

essentially the same structure as they did in the ancestors, i.e. cannot be modified by the genome. The independence of the membrane from the genetic system is very important and even necessary, since any true selecting system must always incorporate a random element. In the case of the membrane this structure is not only independent of the genetic system, but it is also independent of the external environment, since it is rather a shield against it. To be more specific, the membrane is a lipid bilayer (constituted by phospholipids, glycolipids, and cholesterol) showing hydrophobic tails sticking together while the hydrophilic heads remain in contact with water.[120] Membranes spontaneoulsy generate sack-like vesicles, that is, relatively closed systems[121] [Fig. 7.35]. This is a self-assembly process [Subsec. 6.3.1].

The membrane proteins, allowing for both transduction and entropic fluxes from the interior to the exterior, are an example of allosteric proteins [Subsec. 7.4.4], that is, of proteins that can be in two different states (for instance, allowing or not allowing some ions to enter a cell), according to whether or not they are phosphorylized or dephosphorylized (phosophatase).[122] Very important insights into these mechanisms were brought to light by considering the family of the phosphoinositides, which are very relevant, among other things, in cell communication.[123] The lipid tail of these molecules renders the phosphoinositides obligately membrane-bound. This makes them well-suited for marking particular membrane partitions. Phosphoinositides achieve direct signaling effects through the binding of their head groups to cytosolic proteins or cytosolic domains of membrane proteins. Thus, they can regulate the function of integral membrane proteins, or recruit the membrane signaling components.

7.6.2 A Selecting System Serving the Organism

The membrane (understood here as the whole complex that regulates the fluxes between the exterior and the interior) is a true selection system in an organism and in this way allows also for information control on the external environment. Evidence for this is the fact that the most

[120][*BERG et al.* 2006, pp. 326–47].

[121]In [*AGENO* 1991] a segregation through hydrocarbon walls of a small population of complex molecules in water is hypothesized.

[122][MONOD *et al.* 1963, MONOD *et al.* 1965][*CHANGEUX* 2002, p. 20] [HUNTER 2000].

[123][MA *et al.* 1998] [MARTIN 1998] [DIPPOLD *et al.* 2009].

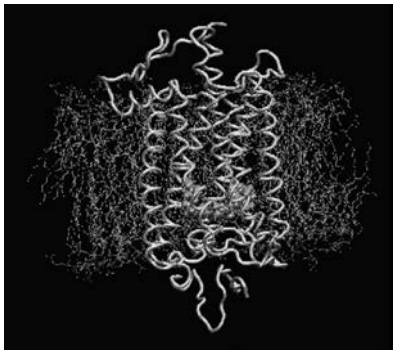

Fig. 7.36 A rhodopsin molecule (yellow) with a bound retinal (orange), embedded in a cell membrane (lipids shown as green, head groups as red/blue). This is an example of proteins embedded in the membrane. Adapted from http://en.wikipedia.org/wiki/Rhodopsin. (The figure is reproduced in color in the color plate section.)

important protein responsible for vision (rhodopsin) can also be found in bacteria embedded in the lipid membrane or in some mobile algae, with the exact same photodetection function [Fig. 7.36; Sec. 4.2]. Therefore, the root of information codification and processing of any more sophisticated sensory system is the information selection of the membrane.

To understand the way in which this system works, we need more general considerations. As will shall see in the following, the organism as a system and an individual unity represents the complexity that deals with physical perturbations in the external environment and is able to transform this in specific appropriate actions (or reactions) that are congruent with the goal of self-survival. The membrane selective system is somehow the bridge or the interface between this external perturbation and everything that may happen thereafter in the organism. It is what transforms this external perturbation (through transduction or, in higher eukaryotes, through particular sensory channels and organs [Chs. 3–4]), which in itself is a pure physical phenomenon, into codified information generating appropriate inputs that are then sent to the metabolic system where they are combined with further endogenous inputs. Obviously, in unicellular organisms the membrane also has another significance, that of being the device allowing for the coming in or out of elements that are necessary for the metabolism. But this is related to actions that organisms are able to perform *after* the codification has already happened and the information has gone through the whole organism. We should not mix these two very different issues, which in higher organisms become fully separated.

Schematically, cellular *transduction* consists of:

(1) A signal triggering the whole process (the so-called first messenger) that in most cases binds a receptor (gate protein) without itself entering into the cell;
(2) As a consequence, a second messenger (an intracellular molecule) is delivered inside the cell and may produce amplification of the initial signal (a sort of cascade); is probably the proper codification step, at least in terms of firing or not firing, as it happens for neurons [Subsec. 3.3.2];
(3) Finally, going into the inside of the cell there is activation of a response involving the metabolic system.[124]

[124][*BERG et al.* 2006, pp. 382–3].

Without considering now what further steps there are inside the organism, let us understand what the significance is of this codification-selecting system. Each codified input somehow represents a certain *surprise* (or surprisal) [Subsec. 2.3.3]. The surprise expresses the novelty represented by the codified input originating from the external physical perturbation. In other words, it is the mismatch or distance between the current state of the organism and this input information. This mismatch is coded in the selection system itself in the form of a mismatch between the state of the receptor (which stands here for the whole organism) and the new input information. Thereafter, the selection system informs the metabolic system about this kind of situation. The metabolic system in turn takes this information, together with another endogenous signal, as a sign to give rise to another series of procedures to preserve or restore its metabolically homeostatic state. In other words, an occurrence of information selection (coding and response) gives rise to a whole procedure, having a final thermodynamic and entropic character.

Mathematically speaking, we can express this procedure by employing a variable i which represents the new input as a function of some external (environmental) parameter k and as describing (determining) the state of the receptor of the sensory system [Subsec. 2.3.1].[125] The quantity A (which in a more sophisticated framework could be related to a matrix) represents the action that the organism undertakes in order to minimize the surprise given by the negative logarithm of the conditional probability

$$p(i, k|A). \tag{7.7}$$

It is important to understand that the action can be taken both on the environment or the organism itself. To be really efficacious, however, it always implies an action on the external environment as a way of avoiding further surprising inputs. Now, it can be shown that lowering the degree of surprise means to lower the following quantity that is an informational analogue of the free energy in thermodynamics[126] [Subsec. 6.2.4]:

$$g = -\left\langle \ln p(i, k|A) \right\rangle_{p'} + \left\langle \ln p'(k; s) \right\rangle_{p'}, \tag{7.8}$$

which implies that this lowering is in full accord with general statistical laws and can even be considered quite natural. The two expressions on the right-hand side are a (statistical) mean value of the logarithm of the relative probabilities: The quantity on the left is the surprisal [see Eq. (2.7)] while the quantity on the right represents the mean value of the logarithm of a probability distribution of both the environmental parameter k and of the internal parameter (state) s of the organism. Obviously, there can be many factors contributing to certain input information. However, this complication is unnecessary in such a context. Moreover, this distribution is always positive. The whole expression can also be reformulated as

$$g = -\left\langle \ln p(i|A) \right\rangle + D_{KL}\left(p'(k; s)||p(k|i, A)\right), \tag{7.9}$$

where the second term is the so-called Kullback–Leibler divergence (also called relative entropy) that here measures the distance of the probability distribution before the two vertical lines from the conditional probability (after the two vertical bars) of the external parameter k given that there is a certain input and a consequent action A. Given two probability distributions $p(j)$ and

[125]These interesting developments are due to Friston and coworkers [FRISTON 2005, FRISTON *et al.* 2006, FRISTON/STEPHAN 2007]. It is interesting to remark that this powerful formalism originally arose in the context of neurosciences. What I am trying to show here is that it can be generalized to apply to *any* organism.

[126][MACKAY 1995].

$p'(k)$, the classical Kullback–Leibler divergence expressed in binary logarithms (in the discrete case) is given by[127]

$$D_{KL}\left(p'(k)||p(k)\right) = \sum_k p'(k) \lg \frac{p'(k)}{p(k)}. \tag{7.10}$$

Note that the mutual information (2.13) can be expressed as the Kullback–Leibler divergence between the joint probability and the product distribution of two parameters

$$\begin{aligned} D_{KL}\left(p(j,k)||p(j)p(k)\right) &= \sum_j \sum_k p(j,k) \lg \frac{p(j,k)}{p(j)p(k)} \\ &= \sum_j \sum_k \left[p(j,k) \lg p(j,k) - p(j,k) \lg p(j) - p(j,k) \lg p(k)\right] \\ &= \sum_j \sum_k \left[p(j,k) \lg p(j,k) - p(j) \lg p(j) - p(k) \lg p(k)\right] \\ &= H(J) + H(K) - H(J,K). \end{aligned} \tag{7.11}$$

The surprise is implicitly conditioned upon the organism in question. It can be seen that by minimizing surprise one is effectively maximizing the probability of the selected inputs under a particular action (or state of the organism). In other words, lowering the amount of surprise means choosing a "model of the world" with the smallest g, while the latter has the highest marginal likelihood. This follows because g is an upper bound on surprise, given that the Kullback–Leibler divergence is nonnegative. This is easy to verify. Indeed, for any distributions $p(k)$ and $p'(k)$ we have

$$\begin{aligned} -D_{KL}\left(p'(k)||p(k)\right) &= -\sum_k p'(k) \lg \frac{p'(k)}{p(k)} = \sum_k p'(k) \lg \frac{p(k)}{p'(k)} \\ &\leq \lg\left(\sum_k p'(k) \frac{p(k)}{p'(k)}\right) = \lg\left(\sum_k p(k)\right) = \lg(1) \\ &\leq 0, \end{aligned} \tag{7.12}$$

which implies $D_{KL} \geq 0$. Therefore, minimizing the expression (7.8) amounts to minimizing the negative log-probability of the sensory input (reducing the mismatch between the expectation and the input).[128] This is precisely what we expect [Subsec. 6.1.1] any organism to do: It will expose itself selectively to those causes in the environment that it expects (or is programmed) to encounter. However, these expectations are limited to the repertoire of physical states that the system can occupy by preserving its homeostasis, and therefore the net result is that the inferred causes approximate the real causes.

The crucial point to understand here is that, in statistics, the minimization of the surprise is equivalent to a Bayesian probability computation, where the function g above can be used to approximate the likelihood function, i.e. the probability that, given a certain transduction, the parameters that may have caused it are those that the organism expects:

[127] The quantum-mechanical counterpart to this expression is $H(\hat{\rho}'||\hat{\rho}) = \mathrm{Tr}[\hat{\rho}'(\lg \hat{\rho}' - \hat{\rho})]$. See also fn. 48 to Ch. 2.
[128] [FRISTON *et al.* 2010].

$$p(k|i) = p(k)\frac{p(i|k)}{p(i)}, \tag{7.13}$$

where $p(k|i)$ is the likelihood function to have k given i, that is, to have the signal k as the possible cause of having in fact received the input i (for the sake of simplicity I do not consider the action A here), $p(i)$ is the *a priori* probability of the input (here, I do not consider internal parameters for the sake of simplicity), $p(i|k)$ is the *a posteriori* probability of having the input i given the parameter k, and $p(k)$ is the probability distribution of the environmental parameter k. Always in accordance with Subsec. 6.1.1, we may see that Eq. (7.13) can be derived from Eqs. (2.3)–(2.5) by making use of the result

$$p(k|j) = \frac{p(j,k)}{p(j)}, \tag{7.14}$$

so that we obtain

$$p(k|j) = p(k)\frac{\sum_{d\in D} p(j|d)p(d|k)}{p(j)} = p(k)\frac{p(j|k)}{p(j)}, \tag{7.15}$$

where we finally need to substitute i for j. It can be shown that by taking the mean value of the logarithm of both sides we obtain an equivalent expression of the mutual information (via the Kullback-Leibler divergence), so that minimizing the surprise is equivalent to reinforce the mutual information between the parameter k and the response j.[129] In our context here, the role of the data d can be played by internal parameters (due to contributions of the metabolic and genetic systems) together with the input in Eq. (7.13), so that we may write

$$p(k|A) = p(k)\frac{\sum_{j\in J} p(A|j)p(j|k)}{p(A)}, \tag{7.16}$$

where the j's denote these generalized inputs and I have inserted the action A. We shall consider the fundamental significance of the above conclusion. By now, we can see how relevant cognitive abilities (understood in a very broad sense) can be for biological systems allowing their plastic ability to cope with the environment helping in this way adaptation, as anticipated in the Introduction. To this extent, this investigation can be understood as a follow-up of AL [Subsec. 7.5.1].

A final problem is the following: Information selection should not be mixed with information capacity, otherwise we would fall back into a mechanist, classical view of the way in which biological systems process information. The former depends on the goals of the biological system, the latter is only the maximal amount of information that can be shared by two systems (in our case, a biological system and the environment). They correspond to two completely different aspects of dealing with information. Obviously, we must distinguish here between the selection of a *class of stimuli* and the specific problem of the selection of a *given stimulus* in certain conditions. As far as the first problem is concerned, the information capacity comes after such a selection and refers to the capacity of a system once a certain space of stimuli is given. As we shall see, this is also confirmed by studies of much more complex forms of dealing with information.

7.7 Concluding Remarks

In this chapter we have seen that organisms are characterized by a higher level of integration of entropic and informational aspects. In particular,

[129] [AULETTA 2011].

- Any organism integrates a metabolic system which plays the role of a regulator, a genetic system, which plays the role of an information processor, and a selection system like the membrane in unicellular organisms.
- The metabolic system is essentially an entropic–energetic feedback circle in which energy is acquired, stored as sugars (or lipids), and then used for building nucleic acids or amino acids, that is, functional entities or codified information.
- The genetic system is again a feedback circle in which there is first a transfer of information from DNA (the site of pure structural and potential information) to RNA (the vector of instructional information), and from this information a protein is built which can have feedback effects on the expression of DNA.
- The membrane is a selection system allowing for the appropriate entropic fluxes between the organism and the environment.
- A protein is a pure functional entity without any informational aspect, while DNA is a pure information-codification structure.
- DNA and RNA represent the birth of classical codified information. In order to obtain such a result several chemical constraints are necessary, especially the separation between chemical bonding and information combinatorics as well as some constraints on the codons, in particular concerning the hydrophobicity and the relation between energy dependence and volume.
- Self-reproduction happens essentially through three main feedback circles: DNA replication, DNA segregation, and cell division for prokaryotes; mating, meiosis, and mitosis in lower eukaryotes; mating, mitosis, and meiosis in higher eukaryotes.
- We have learned how to frame our study of biologically dealing with information in an appropriate way, both from information-processing and information-selecting points of view.

In this chapter we have established the basic elements that we shall meet in the next chapters. Without a deep understanding of these matters the cognitive capabilities of *biological* systems would not be rightly appreciated.

Appendix: The Stoichiometric Matrix

An important approach to chemical reactions in cells is represented by the stoichiometric matrix.[130] In any chemical transformation we have certain chemical compounds or elements and the reaction themselves. In any chemical reaction we have the fluxes

$$|\, v\rangle = (v_1, v_2, \ldots, v_n), \tag{7.17}$$

and the concentration of the chemicals

$$|\, x\rangle = (x_1, x_2, \ldots, x_m), \tag{7.18}$$

which are related by the stoichiometric matrix $\hat{S}$ as follows

$$\frac{d}{dt}\,|\, x_j\rangle = \sum_k s_{jk}\,|\, v_k\rangle\ , \tag{7.19}$$

and s_{jk} is the elements of the matrix $\hat{S}$. We therefore have m metabolites, the intermediate and products of metabolism, x_j and n fluxes v_k [Fig. 7.37]. Let us consider the example shown in Fig. 7.38.[131] In this case, the stoichiometric matrix can be written as

[130] [*PALSSON* 2006, pp. 89–99 and 136–41]. [131] [*WAGNER* 2005, pp. 121–5].

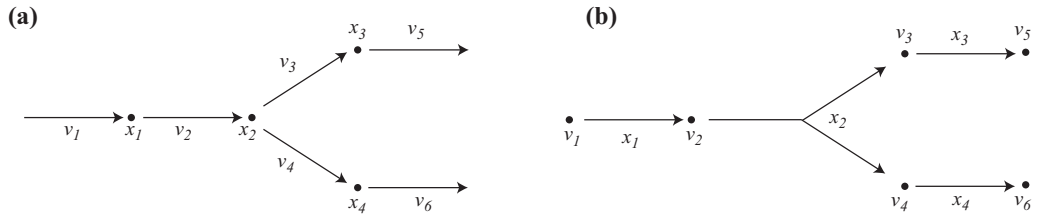

Fig. 7.37 Chemical processes can be understood in two ways: (a) The chemicals represent the points of the transformation and the fluxes represent the links. This process is ruled by stoichiometric matrix $\hat{S}$. (b) Alternatively, the chemicals represent the links and the fluxes represent the dots. This process is ruled by the negative of the transpose of the stoichiometric matrix $-\hat{S}^{\mathrm{T}}$.

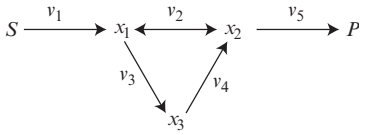

Fig. 7.38 Example of stoichiometric metabolic network $\hat{S}$: An initial substrate S gives rise to a reaction through a circuit constituted by metabolites x_1, x_2, x_3, whose final output is the product P.

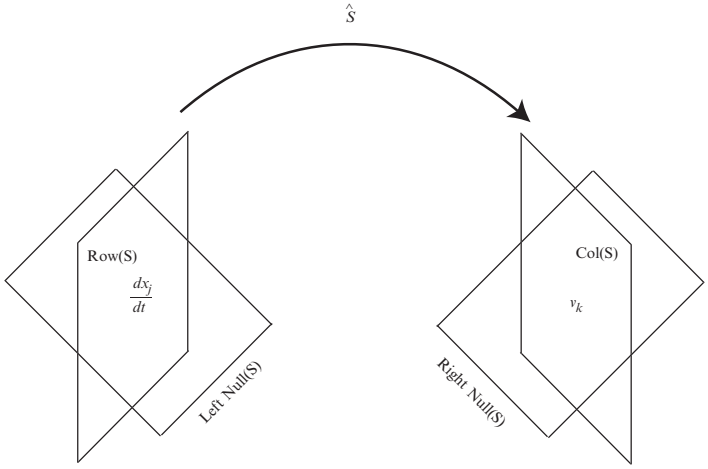

Fig. 7.39 The transformation induced by the stoichiometric matrix $\hat{S}$.

$$\begin{bmatrix} +1 & -1 & -1 & 0 & 0 \\ 0 & +1 & 0 & +1 & -1 \\ 0 & 0 & +1 & -1 & 0 \end{bmatrix}, \tag{7.20}$$

where the columns represent the reactions and the row vectors represent the three metabolites involved. For the sake of simplicity, any contribution is $+1$ or -1 (according to the direction of the reaction).

The stoichiometric matrix is an $m \times n$ matrix, whose rows are embedded in the vertical surface (Row(S)) on the left in Fig. 7.39, and whose columns are embedded in the vertical surface (Col(S)) on the right. The vectors $|l_j\rangle$ constituting the Left Null space (the horizontal surface on the left), are vectors orthogonal to the reactant vectors $|x_j\rangle$ and represent mass conservation. Moreover, any flux vector can be decomposed in a dynamic component $|v_j\rangle$ and in a steady-state component $|v_j\rangle_{ss}$. The row space of the matrix $\hat{S}$ is spanned by the dynamic vectors, while the Right Null space is spanned by vectors representing steady states.

8

The Organism as a Semiotic and Cybernetic System

In this chapter we shall deal with three main issues: (a) Organisms as semiotic systems, (b) teleonomy, information control, and teleology, (c) the notion of biological self. Organisms are essentially biological systems that are able not only to coadapt with other biological systems but also to control environmental information, that is, able to control the relevant parameters of the environment.[1] To do this, they have developed specialized systems for information selection and control [Subsec. 7.6.2]. Without information control no free energy would be acquired, and the organism could not survive. A change of some parameters (of external, environmental variables) will affect the system's stability in some way and could be dangerous, either directly by destabilizing the system, or indirectly by hiding some free-energy sources. In other words, the necessity to control environmental information arises because the environment always changes in a way that (1) cannot be predetermined by the organism and (2) often raises new problems and challenges. From a pure systemic point of view, it is this selective pressure that also produces new forms of adaptation and new solutions, like sexual reproduction at a phylogenetic level [Subsec. 7.5.2]. Therefore, adaptation, although going beyond the issue of information control, could not happen without the latter [Sec. 7.2]: An adaptive behavior is a special case of the behavior of a stable system [Subsec. 6.3.1], the region of stability being the region of the phase space in which all the essential variables lie within their normal limits[2] (this is called homeostasis). I finally stress that at the level of unicellular organisms, especially the bacterial level, a localized information-control instance does not exist. This function is executed by the whole organism as both a distributed and an integrated system.

8.1 The Concept of Sign

We have considered how problematic the classical concept of information-processing is [Sec. 6.1 and Subsec. 7.4.5]. The true mystery of life is how living beings are able to control environmental information by treating it as a sign of something that is fundamental for their own survival[3] and act accordingly. For example, a certain chemical gradient is a sign of a certain free-energy resource. *Signs* are any form of being or activity in which some pattern (icon) and an indexical relation to a referent (an object or event) are connected [Fig. 8.1]. According to Peirce's formulation,[4] a sign is any physical event or object that stands for something in a certain respect or capacity. Summing up, there are two different features to be distinguished in any sign:

[1] [AULETTA 2008a]. [2] [*ASHBY* 1952]. [3] [*VON HELMHOLTZ* 1867, p. 586].
[4] [*PEIRCE CP*, 2.228, 2.247–8, 2.304, and 1.540] [PEIRCE 1903c] [PEIRCE 1907]. Peirce was the father of semiotics.

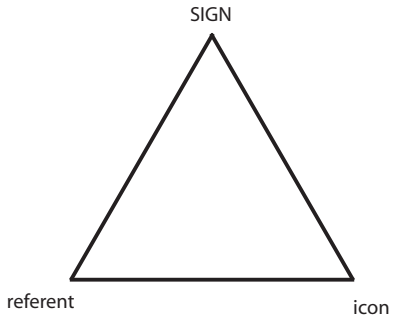

Fig. 8.1 General nature of the sign. The icon can be any pattern that can be associated with a certain referent. The referent can be any object, system, event, or outcome which can be physically interacted with. Ogden and Richards were the first scholars to make use of such triangles [*OGDEN/RICHARDS* 1923], but C. Cherry introduced this form [*CHERRY* 1957].

(1) The *referent*,[5] i.e. the thing that a sign stands for. A sign has an *indexical* relation with its referent: A mark expresses this relation[6] [Subsec. 5.1.4].
(2) A sign has such a relation with its referent in a certain respect. The *icon* is this respect, and it is the instructional or representational content of the sign, according to whether the direction is [Fig. 8.2]
   - From codified information (starting from the genome [Sec. 7.4]) to a referent (here represented by a specific goal to be attained [Sec. 5.3], for instance the production of a protein displaying a certain function) or
   - From some external stimuli that are codified in informational terms at the membrane (for instance, for a receptor to be activated or not) [Sec. 7.6], or also in some peripheral sensory system (sensory coding [Subsec. 3.3.1]), to an internal representational or quasirepresentational pattern or function. This is the way in which information acquisition serves the needs of the organism. The assumption here is that ANY external signal or variation in the environment can in principle be codified in this way.

   In both cases, codification always occurs at the source [Subsec. 2.3.2], which is internal in the first case and external (rather at the surface) in the second one. In any case, only codified information at the source allows for the kind of information control that is necessary for the organism and that will be described below. This is the evolutionary pressure that explains the emergence [Subsec. 2.4.2] of a new classical information codification.

The iconic content does not need to be an analogue of, or similar to, the referent. Anything to which the organism is sensitive can in principle stand for anything else if it is in a proper indexical relation with this something else and in a proper iconic respect relative to the organism and its needs. Speaking of animals' dealing with signs, D. Premack asserts that a piece of plastic becomes a sign whenever the properties ascribed to the item are not those of the piece of plastic but of the object that it denotes.[7] This is a very important intuition that can be generalized in this way: A

[5][*FREGE* 1892a, *FREGE* 1892b].
[6]Kripke thematized this referential relation in terms of the rigid indexicality of proper names [*KRIPKE* 1972] [AULETTA 2003a].

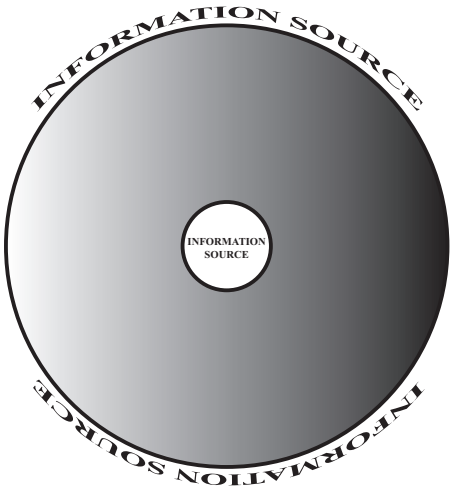

Fig. 8.2 There are two main information sources: One inside the organism and the other outside [see Sec. 7.2]. If we schematically depict the organism as a cell having a core, representing the genetic information (which is also true to a certain extent for unicellular eukaryotes), the first source is located in the most inner part of it (representing its program). The other source is the result of the interaction of the organism with the exterior environment and is depicted here as information codification taking place at the membrane (at the interface between interior and exterior).

sign does not represent physical or chemical properties but the *function* a certain item has, plays, or represents for the organism.[8] We can say that classical information-acquiring deals with the acquiring of information from a source, while representational semiotics, being a guess about the vital significance of an information source, can be conceived as kind of reversed Bayesian inference [Subsec. 6.1.1 and again Sec. 7.6]: This inference is expressed by the probability that, given a certain transduction, the parameters that may have caused it are those that the organism expects.

For this reason, physical laws and semiotic controls require disjointed, complementary modes of description.[9] Laws are global and inexorable [Sec. 2.4]. Controls are local and conditional. Life's semiotical activity is born for the purpose of allowing the organism's goal-directed activity,[10] which is the true core of the cybernetic research program [Subsecs. 3.2.2 and 6.1.5]. Semiotic controls require several aspects, none of which are functionally describable by physical laws that, unlike semiotic systems which rely on an information source, are based on parameters like energy, time, and rates of change. However, they are structurally describable in the language of physics in terms of constraints, degenerate states and processes, information-sharing and selecting, differential timing, and irreversible dissipative events.

The fact that semiotics cannot be dissociated from codified information (even if it is *not by itself* information but only dealing-with-information comprehending functions, outcomes, and processes that are not themselves codified) is of great relevance [Subsec. 7.4.5]. I mention here that some scholars have tried to build a theory of living beings in semiotic terms (so-called biosemiotics), but very often in sharp opposition to (at least classical) information theory.[11] It seems to me that in

[7][*PREMACK* 1976]. [8][AULETTA *et al.* 2008]. [9][PATTEE 1995]. [10][PATTEE 1997].
[11][*SEBEOK* 1991, *SEBEOK* 2001] [*HOFFMEYER* 1996, HOFFMEYER 1997].

this way no significant result can be obtained. For instance, we have already considered [Sec. 7.2] how the concept of a new complex system may arise through a new level of integration of previous aspects (information processing, regulating, and selecting) that are already present at the level of the pure information-acquisition processes [Subsec. 2.3.2]. Moreover, I have said that the most basic relation that grounds the semiotic processes of life can be found in the connection between codified information and function in the activation–transcription–translation process.

Among the first scholars to have understood life in semiotic terms was J. von Uexküll.[12] Indeed, only living beings grasp and produce signs. In the physical world there are patterns [Subsec. 6.3.1], but there is no agent other than an organism that is able to refer these patterns to objects, events, or operation outcomes through a proper indexical relation,[13] which is another way of saying that it is only organisms that can show adaptation processes.[14] This means that physical (or informational) patterns (i.e. signals) cannot be interpreted as signs *if not through* a biological system able to treat them *as* signs.

8.2 The Organism as a Teleonomic and Teleologic System

8.2.1 Teleonomy

We have remarked that prebiotic self-organizing systems like Bénard cells are dependent on entropic fluxes that are not controlled by themselves, and this is the main relation they have with the environment [Subsecs. 6.3.1 and 6.5.3]. These fluxes are also the core of the metabolic system in organisms [Secs. 6.6 and 7.3]. However, organisms entertain a relation with the environment that is also *informational*.[15] In particular, they are able to carve out environmental signals and stimuli that are in principle a source of disturbance and even noxious to their needs [Subsecs. 3.2.2 and 7.6.2]. Indeed, *any* variation in the environment is *a potential threat* for the organism's homeostasis.

Let me first summarize the results that have been found so far. As we have seen [especially Subsec. 7.4.4], biological systems make use of mechanical work to accomplish their task of information control and therefore, as classical systems, are causally ordered. However, there is an important proviso for organisms. There is here a fundamental distinction between the *current state* an organism is in and its *memory* (for instance, represented by the genetic memory) or its developmental *program*.[16] In fact, the information contained in a genome alone does not constitute the whole of the information necessary to build the corresponding phenotype [Subsec. 7.4.5], or all of the information contained in the embryo of a multicellular organism. Any organism (and especially a multicellular organism), in order to build its mature form, also depends on environmental cues and feedback circuits during epigeny.[17] At any stage of the development an organism shows a quantity of structural information [Subsecs. 2.3.4 and 6.4.1]—for instance, the body plan—bigger than that contained in its genome or its embryo. In other words, as for quantum systems [Subsec. 1.3.2], there is also some type of indeterminacy in biological systems, which is expressed here in terms of an insufficiency of the instructions contained in any embryonic state for fully determining the mature form of the phenotype or the three-dimensional structure of the protein even if those instructions *start* the whole process.

[12][*VON UEXKÜLL* 1926].

[13]Peirce seemed to believe, on the contrary, that semiotics could also be applied to the physical universe [PEIRCE 1866, pp. 471–5] [PEIRCE 1903b, pp. 193–4].

[14][AYALA 1998b]. [15][AULETTA 2010].

[16][AULETTA 2005b]. [17][*BARBIERI* 2003] [AULETTA 2006c].

Quantum and biological systems also show another important similarity: For both classes of systems an increase in determination may occur and this is allowed only through an interconnection with the environment. In the case of quantum systems, we have the emergence of the matter's structures [Sec. 6.2]. The reason for this commonality between quantum and biological systems is due to the fact that the latter ones are complex systems that are able to codify information [Subsecs. 6.5.1 and 7.4.1].

However, the way in which biological systems integrate complexity and information [Subsec. 7.4.5] makes them different from quantum systems. Indeed, quantum systems are able to generate structural information by downloading "interference information" (quantum features) into the environment and thereby losing their specific quantum character[18] [Subsecs. 1.2.8 and 2.2.3–2.2.5]. On the contrary, biological systems are able to *integrate* environmental cues into their own developmental path and therefore to put them—to a certain extent—at their service (similar considerations are also true for the internal environment during the different operations of cells that we have discussed above).[19] Similar processes are also common in the adult phenotype or in protein folding[20] [Subsec. 7.4.4]. The genome starts an informational process that leads to a stable final state (the mature form or the three-dimensional configuration of the protein) only by integrating environmental signals that are generated *independently* from the organism or by integrating cues that, for instance, come from the metabolic system and are also independent from the specific protein considered. The reason for this is that the phenotype or the proteins are a black box relative to the genome [Fig. 8.3]. We see again that there is no Lamarckian instruction even inside a *single* organism. However, the organism is also able to—partially—*canalize* the action of the external or internal environment,[21] producing a good fit (appropriate to the current operation, which is a phenomenon of *coadaptation*[22] (channelization).

This coadaptation is also the reason why, from another point of view, the information contained in the genome is overabundant, in the sense that it is never fully expressed (recall here the accessibility principle [Subsec. 2.2.2]): For instance, only a subset of the possible proteins, potentially codified by the genome, is built inside a single cell or, in the case of multicellular organisms, any cell—during development—becomes specialized by expressing (selecting) a subensemble of the information it potentially contains. Again, this is a feature that we also find in the case of quantum systems, because a measurement outcome represents in general a selection out of the initial amount of potential information.

It is precisely the difference between memory or program and current state that gives to biological systems their distinctive temporality. By *temporality* I mean a rupture of time isotropy, such that

- Future and past states cannot be deduced from the system's current state.
- There is irreversibility, a feature that is strictly connected with the complexity of biological systems.

A. Lotka was one of the first scholars to understand that even the evolution of life is an irreversible process.[23] According to him, evolution is characterized by lags and leads that determine irreversibility, since in the case of lags, the actual and future state of the organism will depend on its previous history.

[18][AULETTA 2005b]. [19][*ARTHUR* 1984, pp. 33–5]. [20][BROOKS *et al.* 1998].

[21]A first idea of canalization at biomolecular level can be found in [*MONOD* 1970, Ch. 3].

[22][*SPENCER* 1855, pp. 385–8] [*LLOYD MORGAN* 1891, p. 119] [*RUSSELL* 1930, pp. 6–7]. See also [*ROBERT* 2004, p. 70].

[23][*LOTKA* 1925].

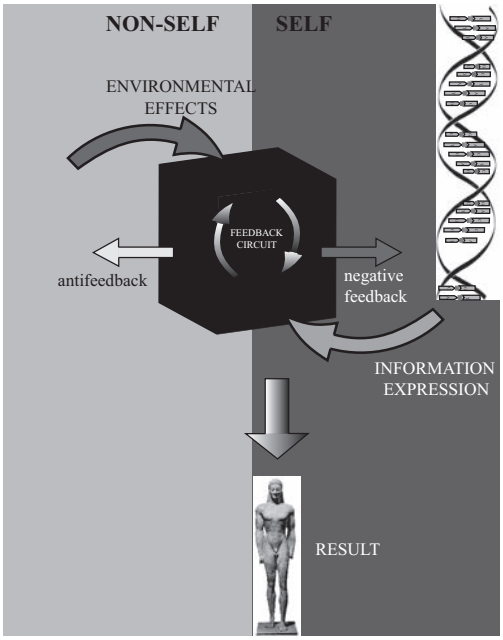

Fig. 8.3 A good example of teleonomic processes: Development (I only consider here the aspects relevant to this section and not others that also play a major role). The immature phenotypic system from which the mature form is developed is a black box to the genetic system (as well as to the environment). Nevertheless, certain informational processes coming from the DNA have been integrated in a feedback circle with independent environmental signals in order to give rise to a stable mature configuration. These two heterogeneous types of input have been selected together at a phylogenetic level (coadaptation). In other words, the genetic and environmental systems are both independent in their modality of action and tuned to one another due to an effect of natural selection. In this way, the final result is not simply the sum of these two independent modes of action. The antifeedback is the way in which the organism resists environmental pressures (negative feedback) and is able to tune the external environment according to its needs in order to restore its own homeostasis.

This nature of biological systems, and in particular potentiality and temporality, affects the nature of causality. In fact, the dynamical nature of biological temporality is essentially *teleonomic*[24] (though organisms also use mechanical causation for doing work, as mentioned), while classical systems are ruled only by a mechanical or efficient form of causality.[25] The difference between these two forms of causality is very important:

- In efficient causation, once the causes are given, the effects are linearly and certainly produced, so that any classically causal path is univocal.

[24][ROSENBLUETH *et al.* 1943] [*ASHBY* 1952] [*MAYR* 1988] [PITTENDRIGH 1958]. For a very good historical examination see [*CORDESCHI* 2002, Chs. 4–5].

[25][*SCHLICHTING/PIGLIUCCI* 1998, p. 27].

- On the other hand, where we have *teleonomic* causation, the current inputs of the system are insufficient to determine its future state. Moreover, a system ruled by teleonomic causation can follow different and alternative paths in order to arrive at the same final state.

While in efficient causation *only the path* (or the set of antecedent conditions) matters, in teleonomic causation *only the final state* does: Organisms are self-regulated systems that involve in their dynamics what is important for survival[26] [Subsec. 3.2.2]. In other words, efficient causality is sensible to past conditions but is robust in the ability to produce a certain effect given certain conditions (this is the determinism of classical systems). Teleonomic systems, instead, are robust relative to the final state or the task to be accomplished: Many variations of past conditions will not affect the realization of the task or the arrival at the final state (equifinality) [Subsec. 6.3.3].

I am here following the terminology of J. Monod,[27] according to whom teleonomy is the basic property of living beings as objects endowed with a project, represented by their structures and realized through their performances. Moreover, he stressed that such a property denotes the ability of organisms to build themselves through internal forces (in the expression of the genetic program) and not external ones, implying a freedom relative to exterior conditions (otherwise, they would only be sensible to past conditions, as happens for mechanical causality [Subsec. 5.3.3]). Therefore, teleonomy amounts to the quantity of information (and order) that must be realized, maintained and transmitted invariantly over the generations in order that this or that specific organism survives as a biological species. It is important to recall that Monod clarified that teleonomy does not imply having goals, i.e. a teleology, and firmly rejected this notion. I, on the contrary, shall use it later on. I believe that the reason for this rejection was that Monod underestimated the ontogenetic goal-directed activity of the individual organism (as in chemotaxis) and took only phylogeny and some very basic epigenetic processes into account (this was due to the situation of biology at that time).

Teleonomy can arise only when a circuit is established in which three types of feedback can work together in an appropriate way:

(1) A self-increasing positive feedback mechanism responsible for the building or growth of structures in the phenotype or of the three-dimensional configuration of a protein,
(2) A negative feedback having mainly the effect of blocking the expression of DNA, and
(3) An antifeedback directed towards the environment,[28] that is, a process activated by the organism with the aim of preventing environmental effects that are dangerous for the organism itself and aiming to restore its homeostasis.[29] In this way, antifeedback determines characteristic distortions of the way in which negative feedback contributes to new inputs, resulting in modulatory effects on both negative and positive feedback: Such a result is therefore the grounds for any further regulatory mechanism in life.

As stressed in the path-breaking paper of Ronseblueth *et al*,[30] when dealing with teleonomy (and, as we shall see, with teleology), negative feedback matters much more, because it represents a correction (coming from the environment or from a subsystem of the organism) and therefore a selection of the previous input. Instead, in nonbiological self-organizing systems [Sec. 6.3] the positive feedback is in general much more relevant (and here we often have amplification effects that are not kept under control). It is evident that several environmental cues can also induce

[26] [*ASHBY* 1956, p. 219]. [27] [*MONOD* 1970, Ch. 1]. But see also [AYALA 1970, p. 9].
[28] An expression in complex organisms is feedforward [*JEANNEROD* 2009, pp. 84–6].
[29] Also understood by Bichat [*BICHAT* 1800, pp. 1–3], who spoke of an internal principle of organisms able to counteract the spontaneous tendency towards destruction. See also [*RAMELLINI* 2006, pp. 32–4].
[30] [ROSENBLUETH *et al.* 1943].

DNA expression, and in this way they appear to represent positive and not negative feedback. However, this is the *effect* resulting from the ability of the organism, which thanks to antifeedback in particular and the feedback-control mechanisms shown in the previous chapter, is able to carve out a potential threat to its homeostasis in a *source of stability and growth*. In this way, we have the indirect emergence of new behaviors and structures [Subsec. 2.4.2]. Antifeedback exists only in organisms (being homeostatic systems). Not only significant aspects of development but also the whole mechanism of self-reproduction, especially in the cycle of higher eukaryotes, is a teleonomic mechanism.

Summing up, teleonomy, as I understand it, is a mechanism based on the attraction exercised by a "final" or next stable state on a biological system.[31] Actually, the final state of the system is not a reality existing by itself outside of the system and before the system implements it. Nevertheless, the crucial point is that it is able to regulate the dynamics of the system from the inside, and in this sense it is something more than a formal constraint acting only from the outside, as is still true for any other physical system [Secs. 1.3, 6.2, and 6.5]. The reason is that the dynamics has been selected here (through natural selection) in such a way that the organism already possesses, embedded in the cluster of relations characterizing it, the potential resources to deal with *whole classes* of external and future events in order to eventually reach the next stable state, namely the attractor of the system [Fig. 3.24], although, as we shall see, the latter cannot be understood in static terms: we must rather take into account a certain dynamic itinerancy of the organism. In other words, the reason is that a biological system can establish a channel with the environment even if in fully dynamical *independence* from it [Subsecs. 2.2.3, 2.2.6, and 6.5.1]. Then, when certain external signals occur (within a certain tolerance window), the organism is able to properly react and even to integrate them because it is prepared for specific environmental stresses and cues in an *anticipated* way,[32] a capability that plays a crucial role during developmental processes. This capability is a pure semiotic relation, that establishes an *equivalence class* (of external signals)[33] and grounds also an information-control mechanism—even if no information control is directly involved in teleonomic processes. Therefore, this semiotic relation with external cues and events makes them able to concur in determining the final output of the dynamics *as if* it were an outside force acting on this dynamics. Here, something that comes later on concurs to determine a process that starts because of this possible and subsequent event.

Therefore, it is teleonomy that accounts for the distinctive way in which organisms deal with the environment during the generations. Indeed, since the selected coadaptation between organism and environment is so complex and rather fragile, any genetic mutation and especially its expression in the phenotype (or in protein building) cannot be too abrupt. In other words, it cannot perturb this precarious balance. It must necessarily result in a further growth or different usage (or even use in a different context) of *already existing* structures [Subsec. 6.3.2]. This explains phenomena like exaptation (about which we shall say more in the next chapter), but also the general rule in biology that there is never a new beginning but always a further development of *already existing* structures. Often, it has been asked why nature does not work like an engineer, who is able to design a machine in an ordered way instead of using old existing structures. From an ideal point of view this is preferable, since it allows us to have a machine that works better at a lower cost. However, this is possible because the engineer probably has full control over the conditions for generating his machine (ability to plan, availability of materials and mechanical parts, or ability

[31] [AULETTA 2008d]. See also [*PIAGET* 1967, pp. 130–1].
[32] [*GILBERT/EPEL* 2009, pp. 129–34].
[33] [*MATURANA* 1970, p. 10].

to produce them if not, money for doing all that, a market interested in the product). However, organisms are not totally autarchic and depend on the integration of environmental cues, and the environment is in principle unpredictable (again a case of soft assembly). Indeed, during many of their processes organisms show teleonomic behavior without explicitly controlling the environment. This is the fundamental reason for the characteristic way in which evolution proceeds.

8.2.2 Information Control

We can speak of information control when there are two factors that cannot be reduced to any low-level element[34] [Subsec. 3.2.2]: 1) the formal structure or constraints determining a feedback control loop,[35] and 2) a goal to be reached. These two elements represent the way functional equivalence classes [Subsec. 5.3.2] are controlled in a top-down fashion [Fig. 6.16]. This means that only those constraints that satisfy the following requirements will work:

(a) An *operation* has to be executed in order to deploy the function needed by the organism. By *operation* I understand any spatially and temporally coordinated pattern of physical-chemical interactions (it is a complex structure) able to fulfill a function and therefore subjected to some top-down control.[36] Operations are therefore a bridge between chemistry and functionality [Subsec. 6.1.4] and display both efficacious work and variety. We could use the term "pathway" instead of "operation." However, in the prebiotic domain there are also pathways which are not necessarily operations since there is no information control. Moreover, the first step in information acquisition by the organism gives rise to a spontaneous pathway that is not controlled (it is a bottom-up process). The control comes only when such information is combined with endogenous signals.

(b) This is strictly linked to the fact that the organism has an inbuilt or ontogenetically acquired goal to reach. The most elementary goal of a biological system is its *self-maintenance*: A complex system that is forced to rule the entropic fluxes with the environment in a adaptive way, is also naturally pushed (through natural-selection mechanisms), for the sake of its self-maintenance, to act on the environment according to certain internal needs (i.e. which depend far more on its self-organization than on environmental parameters) [Secs. 7.1–7.2]. Therefore I stress that the category *goal* is not charged with any intentional meaning, but expresses the basic and irreducible fact that organisms are able to actively self-maintain themselves.[37] It is the result of a selective fork: Either control information or be destroyed. We come back here to the issue of emergence [Subsec. 2.4.2]. We have emergence when starting from some initial variety (diverse polymers), the process attains such a complexity [Sec. 6.3] that some particular constraints are activated that *canalize* the subsequent process that finally gives rise to new forms of interdependency (*channelization*) and new constraints.[38] This is the way in which goals have arisen as constraints on a complex system showing both informational and metabolic activity segregated by a membrane[39] [Subsecs. 7.3.1 and 7.6.2].

The same point has been understood by W. Wundt who, along with W. James, is the father of modern psychology, and said that, even if at the lowest levels of life the processes

[34][AULETTA *et al.* 2008]. [35][BATESON 1967]. [36][AULETTA *et al.* 2008].

[37][SHAPIRO 2007]. This was already very clearly understood by Oparin [*OPARIN* 1957, p. 350]. About the legitimacy of this terminology see also [AYALA 1970, AYALA 1998b].

[38][*MONOD* 1970, Chs. 5 and 7].

[39]In this way, we have the appearance of structures showing goals whose emergence is in itself not teleologic but teleonomic. In this way, we can deal with Kant's famous problem about the apparent teleology of finite beings [*KANT* 1790, Par. 77].

of consciousness are confined within extremely narrow limits, and the will is determined by the universal organic impulses only in the very simplest way, it is also true that, even among the lowest protozoa, the manifestations of life are explicable only upon the hypothesis that they possess a mind.[40] I would avoid using similar expressions here myself (unless we do not understand consciousness as the fundamental manifestation of the biological self, as we shall see below), but I support the idea of the organism as a goal-directed system.

(c) The organism needs to be able to verify somehow in a step-by-step fashion whether or not the function is actually deployed to the required degree. This will be done through regulative mechanisms able to measure the gap between the actual outcome of the operation and the expected outcome (determined by the goal). In this case, a successful outcome is taken to be a sign of the operation having been executed (as well as an unsuccessful one taken as a sign of something gone wrong). In this case, the operation will be repeated (with a possible correction).

So, we find here the triad controller-decider, comparator–regulator, performer–processor, in accordance with the previous schemes [Subsec. 2.3.2]. *Any* pathway of interactions that is framed in the above general structure of information control can do the job. Instead, when speaking of operation inside a cell, the chemical reaction here is the elementary unit of any operation and also of any biomolecular network. As I have mentioned in the introduction to the chapter, it is not necessary that a separated and somehow localized information-control system exist—this is often present in higher animals. On the contrary, any function must always be instantiated in some operation.

In order to have information control we also need some *codified* information. In Sec. 8.1, I have indeed explained that there is no semiotics without information. This is *a fortiori* true for information control (being based on semiotic processes). Now the fundamental question is: Where is this information nested in the scheme that I have presented here? Semiotic processes can obviously be very different according to the context, especially taking into account whether we are dealing with (genetic) instructional information or (ontogenetic) information acquisition [Fig. 8.4]:

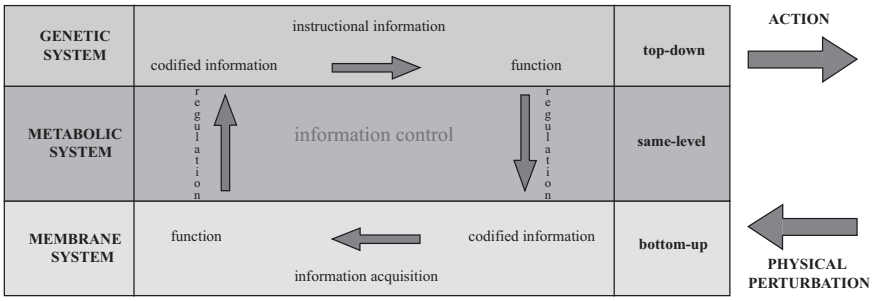

Fig. 8.4 The instructional information bridges between genetic information (the icon) and a given functionality (the referent) necessary for surviving, while information-acquiring allows the organism to be informed about the (induced) changes of the environment (the referent) thanks to appropriate functional steps (the icon). Through the regulatory system these functionalities back-react on the genetic system allowing for expression or repression. The whole distributed circuit displays information control. See also Fig. 7.27. (This figure is reproduced in color in the color plate section.)

[40] [*WUNDT* 1893, p. 29]. See also [*JAMES* 1890, v. I, p. 149] [*ROMANES* 1882]. In a similar language, it has been said that any organism is necessarily a cognitive system [*MATURANA* 1970, p. 10].

- In the first case organisms use (genetic) information for giving rise to and controlling functions (it is somehow information *controlling*): It is a top-down process aimed at the self-maintenance of the organism. Actually, any control needs to be performed in informational (non-mechanical) terms [Subsec. 5.3.3].
- In the second case the organism is monitoring environmental information (ontogenetic *control on* environmental information in the sense in which cybernetics understood it [Subsec. 3.2.2]): Here, information is acquired in a bottom-up way and carved out to give rise to functions necessary for expressing information and, in a feedback way, for controlling.

The whole scheme constitutes a feedback circuit through the regulatory system and it is *only this circuit as a whole that displays full information control*: The issue of information control is the vital significance of certain signal for an organism. It is true that I have spoken of genetic control of information or of control of external stimuli, that is, of mechanisms of control on the single steps along these circuits. It is very important to understand that we can speak of information control on the single steps and even admit some structural analogy between the way in which subsegments and the process as a whole works, *only because* the organism displays such information control *at the organismic level*. Finally, I stress that this double circuit presupposes both a separation and a connection between biological self and nonself, as will be explained in the last section of this chapter. I shall show now that, when there is information control, there are also equivalence classes.

8.2.3 Functional Equivalence Classes

Organisms as wholes do not act in mechanical ways, although many segments at the performer level can be executed mechanically [Subsec. 8.2.1]. However, teleonomic and teleologic processes never contradict mechanical causality. To deepen the distinction between mechanical causality and the latter processes, let us take the example of a thermostat. Although thermostats have feedback circuits, these circuits are designed and built by humans and have no goal in themselves. Specialized literature often speaks of the goal of a certain artificial device. However, this is only a figure of speech, since the goal is only in the humans who have built it. Instead, I assume that to have a true goal one needs a semiotic relation, i.e. an act of reference [Sec. 8.1]. This makes an important difference, since the "information control" exerted by a thermostat (maintaining a certain temperature) coincides here with a *single* path of physical interactions (the switch mechanism is turned on and off according to the temperature level attained): In other words, a mechanical device like that can only work in the way it has been arranged by humans; while in a system with true goals, there are in general several—physically or chemically *different*—operations that can lead to the same desired outcome (within a certain tolerance window).[41] We have already seen this aspect when dealing with teleonomy [Subsec. 8.2.1]. Let us now consider the concept of functional equivalence classes in all its generality.

The equivalence class is a rather formal concept that is commonly found in many fields. Its properties are

- Symmetry: If item a is equivalent to item b, then b is also equivalent to a,
- Reflexivity: a is equivalent to a, and
- Transitivity: If a is equivalent to b and b to c, a is also equivalent to c.

[41] [*WAGNER* 2005, pp. 195–216].

However, when speaking of equivalence at the most basic level of molecular interactions, such as in the circuit DNA–RNA–protein [Subsecs. 7.4.4–7.4.5] or at the level of the molecular mechanisms ruling the selection performed by the membrane [Sec. 7.6], we are actually dealing with *functional* equivalence classes.[42] A functional equivalence class is *context-sensitive*. It is a pure biological category, where different operations or signals are considered to be equivalent if they produce the same *outcome* for some functional purpose (a goal). Thus, I focus here on functional equivalence classes rather than purely formal equivalence classes, even if the formal properties previously defined must also hold for them. This, however, should not be understood in an absolute sense: Since operations that are equivalent with respect to a certain function are not automatically equivalent for other functions, it is important to identify the function concerned unequivocally and therefore also the context.[43]

Therefore, the criterion by which items are judged to be or not to be members of such classes is only a specific function, a problem that which can be reduced finally to the result of such a function. In this way, functional equivalence classes are characterized by a part–whole relation, which in turn links a particular kind of function directly with the issue of semiotics. As we have seen, with respect to equivalence classes, the outcome of the operation performed in a certain biological context may be considered a sign of the function being deployed successfully (or not), and hence that the operation really is (or not) a member of the equivalence class. In other words, of all the input information entrained in the physical-chemical properties of the molecules or of the interactions under consideration, a single feature is selected and taken as a *sign* of the function required by the inbuilt *goals* of the whole system. Properly speaking, this sign may be the outcome of the operation, or any of its features reliably associated to the outcome. Let us call this type of sign a *mark* of the operation. Systems able to perform a certain operation for a certain goal are acknowledged through a specific mark, and controlled in their mode of operation through such a mark. This is the reason why information control of an operation is a genuine semiotic process: Through the comparator, the control instance or the network as a whole needs to catch and select specific information as a sign of the fact that things are going in the right or wrong way. We have splendid examples of the way in which, at a molecular level, this marking system works. For instance, synaptic marking is a transient modification of a given synapse produced by short-term memory activation which marks a particular synapse in such a way that proteins are recognized and stabilized at that synapse, thus allowing for long-term memory[44] [Subsec. 3.8.2]. As we shall see, the brain uses some representations as marks or labels for clusters of other representations. The same is true for much more basic biological processes like those we are considering here.

8.2.4 The Concept of Function

There are no functions in physics. We may *use* certain physical objects (e.g. a stone) for a certain purpose (like an anvil). But this is dependent on the biological or even mental features of the agent and *not* of the stone. It is true that we can already speak of functional groups at a chemical level, comprehending several chemicals that share some fundamental properties[45]; for instance, some properties are common to the whole hydroxyl group (OH) and not just ethanol. However, we cannot speak here of functional *equivalence classes*. To be able to do that we need an

[42][AULETTA *et al.* 2008]. See [WEGSCHEID *et al.* 2006] for first experimental evidence of different operations displaying the same function. This kind of experiment has been further improved in [GOBERT *et al.* 2010].

[43]For the centrality of the concept of function in biology see [*VON UEXKÜLL* 1926, pp. 103–9].

[44][GOELET *et al.* 1986]. [45][*CLAYDEN et al.* 2001, pp. 31–7].

operation and its outcome (a far more complex chemical pathway than single molecules). Only the latter instantiates biological functions. Indeed, from the previous considerations we may define a biological function as some change or outcome to which an operation is ordered (from a control instance).[46] In the following it should be understood that when I speak of function, I am referring to a *biological function*. We may also infer that any biological function has the following general characteristics:

- It cannot be reduced to pure structure [Sec. 6.1], even if it depends on some specific structural elements that are crucial for the function and on the architecture of the whole[47]: Organisms are soft assembled. *Hubs* in networks are a manifestation,[48] something that contributes a great deal to the robustness of both networks and functions, since other elements can be changed or even fail without affecting the functionality[49]—functional equivalence classes are an immediate consequence. Something similar is already true for a tool's function [Subsec. 4.4.4]. Also, here there are parts that cannot fail or be changed without destroying its functionality. However, as mentioned, this functionality has been imposed and is used by humans.
- It is always *instantiated* in some operation: However, it does not depend on the operation, but rather on some formal constraints from above and the goal of this upper level [Subsec. 6.3.2]. So, it always presupposes a hierarchy.
- It is connected with the *needs of the whole organism* (or of the whole cell) and therefore relies on the cooperation of different subsystems, like the genetic, metabolic, and selecting systems.
- It depends on *information codification* but is not its immediate result [Subsec. 7.4.5].

A function represents a higher-order level that cannot be accessed or grasped by the pure local level of chemical interactions [Subsec. 6.5.1]. Obviously, the latter does not violate chemical laws but represents and helps a further level of organization than the level of operations and chemical interactions[50] [Subsec. 2.4.2]. In quantum mechanics, we need to *compare* different results obtained in two different locations in order to ascertain whether two particles are entangled or not [Subsec. 2.2.5]. Here, we need to compare different operations in the same organism or in different organisms in order to ascertain that they are functionally equivalent and thus to grasp the meaning and character of the related function.

This means that, in order to have a function we need both a selection and some constraints, which are again of two types: Environmental ones and systemic or intrinsic ones [Sec. 6.6]. This implies that we cannot explain the appearance of a function through selective pressure only, but we should consider the condition in which this function is relevant or even necessary.[51] A very common process for the emergence of new functions is the functional combination of subunits,[52] a process that is extremely relevant for the emergence of new structures and systems [Sec. 7.2].

[46] I owe this definition to my friend Andrew Pinsent at Oxford University.

[47] This shows that M. Behe is mistaken in his ideas according to which the (irreducible) complexity of organisms consists of the joint presence of all components of the system [*BEHE* 1996, pp. 42–3]. I remind the reader here of Spencer's law [Subsec. 2.4.4].

[48] [JEONG *et al.* 2000, JEONG *et al.* 2001].

[49] A point already understood by some fathers of biology [*SAINT-HILAIRE* 1830, p. 98]. See also [*MITCHELL* 2009, pp. 70–4].

[50] [*WUNDT* 1893, p. 31].

[51] For an examination of some of the philosophical difficulties related to this issue see [MILLIKAN 1989].

[52] [*GERHART/KIRSCHNER* 1997, pp. 214–18].

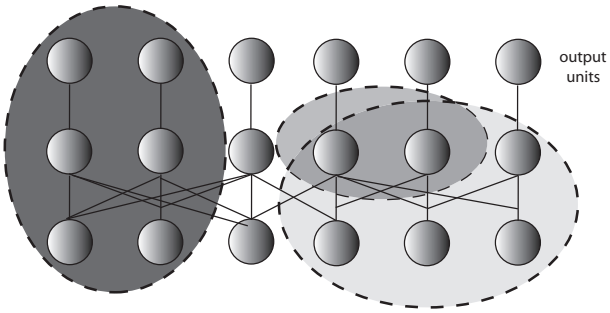

Fig. 8.5 Degeneracy can be understood as the fact that small subsets (dashed ellipses) of units (solid circles) are more tightly connected in such a way that they share a considerable amount of information with the output units (the first row of units above). The two rows at the bottom level represent the network.

8.2.5 Redundancy and Degeneracy

When speaking of functional equivalence classes, an interesting approach is to compare redundancy with degeneracy[53]: Unlike redundancy (the quantity that played a crucial role in classical information theory), which occurs when the same function is performed by identical elements, degeneracy, which involves structurally *different* elements [Subsec. 8.2.3], may yield to the same or different functions depending on the context. It is a prominent property of gene networks, neural networks, and evolution itself. It helps to clarify the relation between structure and function [Subsecs. 6.1.4 and 8.2.4].

Let us introduce some formal considerations.[54] Let us consider the network (it may be a genetic, metabolic, or even epigenetic one) shown in Fig. 8.5. The j-th set X_j^k composed of k units (here sets may also overlap) shares information with the output set O according to the usual formula for mutual information [Eq. (2.13)]

$$I(X_j^k : O) = H(X_j^k) + H(O) - H(X_j^k, O). \tag{8.1}$$

Let us choose to partition the network in sets of k elements. Then, relative to such a partition, the degeneracy of the whole network $X = \sum_j X_j^1$ (note that the sets X_j^1 are composed of only one unit) may be defined as

$$D(X, O) = \big\langle I(X^k : O) \big\rangle - \frac{k}{n} I(X : O), \tag{8.2}$$

where n is the total number of computing units and

$$\big\langle I(X^k : O) \big\rangle = \frac{1}{n} \sum_{j=1}^{n} I(X_j^k : O) \tag{8.3}$$

is the average mutual information of k-element sets. It is obvious that many elements will be found in different sets. Degeneracy is high when the mutual information $I(X : O)$ between the whole system and the output is not so high and at the same time the average mutual information between small subsets of the system (small values of k) and the output is higher than one would

[53][EDELMAN/GALLY 2001]. [54][TONONI *et al.* 1999] [TONONI/EDELMANN 2000].

expect from a linear increase over the increasing subset size. Now, let us consider redundancy. We may define it as

$$R(X, O) = \sum_{j=1}^{n} I(X_j^1 : O) - I(X : O), \tag{8.4}$$

which means that redundancy is high if the sum of the mutual information between each single unit and the output set is much larger than the mutual information between the entire system and the output. In other words, each of the elements of the system contributes similarly with respect to the output. Redundancy is zero if all elements of the system contribute to the output independently, which implies that the mutual information $I(X : O)$ between the entire system and O is equal to the sum of the mutual information between each element of the system and O:

$$\sum_{j=1}^{n} I(X_j^1 : O) = I(X : O). \tag{8.5}$$

As anticipated, degeneracy is the structural basis of functional equivalence classes, since several subnetworks can give rise to the same outcome and therefore in certain conditions display the same functionality, but in other conditions the same network can also display different functionalities, which shows that function is related to structure but in a loose way [Subsec. 6.1.4].

It would be useful here to consider a relevant example[55]: What happens when certain genes fail. In a number of cases (up to 30%), there is little or no evident phenotypic consequence despite the absence of the selected gene products [Subsec. 8.2.4]. Some examples include mice that are unable to make such seemingly important proteins as myoglobin, tenascin C, vimentin, gelsolin, and a neurofilament subunit. Similarly, in a systematic screen of single gene deletions at more than 500 loci in yeast, fewer than half showed any quantitative growth defects in either a rich or minimal medium. The plausible hypothesis is that the gene networks of the affected animals are degenerate, allowing widespread compensatory adjustments. Another relevant example is represented by the fact that the yeast $\alpha 2$ protein and *Drosophila*'s engrailed protein, with almost the same structure and functionality, only share 17 amino acid residues over 60.[56] This means that degeneracy allows for a fundamental character of life: Robustness against environmental fluctuations.[57]

8.2.6 Equivalence of Stimuli

Until now, I have mainly considered instructional semiotics. Let me briefly consider representational semiotics, where by this term I understand a mapping world–organism, without necessarily assuming the existence of representations in the true sense of the word [Sec. 8.1]. Representations are indeed very complex icons that are the result of the integration between different sensory modalities, and therefore can be found only in animals with a brain. However, for the reasons explained above, any single cell must be able to treat certain signals as equivalent, that is, as pertaining to the same class. This is especially relevant for stimuli coming from the external environment, although the following considerations also apply, to a certain extent, to cells' internal signals [Subsec. 7.6.2]. The equivalence of stimuli[58] is the very fact that grounds the semiotics of life[59]: Different stimuli may arouse the same response by an organism. Indeed, as I have said, the cornerstone of semiosis is that several signs can point at the same target or referent. I also recall

[55][EDELMAN/GALLY 2001]. [56][WOLBERG *et al.* 1991]. [57][*WAGNER* 2005, pp. 62–89 and 228–46].
[58][LASHLEY 1942]. [59][*HEBB* 1949, pp. 38–59].

that von Hayek[60] followed Klüver's theory according to which there is a class of stimuli and a class of possible responses with a one-to-many function.[61] We can express this requirement in a very simple way. If S, S' are two different stimuli, C is a certain complex of environmental conditions, and R is a response, we have

$$\{[(S \wedge C) \rightarrow R] \wedge [(S' \wedge C) \rightarrow R] \wedge \neg [C \rightarrow R]\} \rightarrow (S \leftrightarrow S')\,, \tag{8.6}$$

where $\wedge$ is the logical symbol for conjunction (AND), $\rightarrow$ the symbol for implication, $\neg$ for negation, and $\leftrightarrow$ for equivalence. The meaning of this formula is that if two different stimuli S and S' are followed by the same response R in the same context C, and this context alone does not suffice for evoking the response, then this implies that S and S' are treated as equivalent by the organism. Note that this is a sufficient but not necessary condition of the equivalence between the stimuli. According to traditional AI,[62] the single law of a cognitive system is that it takes the actions necessary for attaining its goals, using for this purpose all of the knowledge that it has. I think, instead, that a biological system (and not only a cognitive one) acts according to the equivalence classes *determined by* its goals, and the role of knowledge is eventually to update these classes (to update their features and their subdivisions).

8.2.7 Teleology and Teleonomy

By *teleological causality* I mean the mechanism through which a system exercises an informational control on another system in order to establish an equivalence class and select some specific information for *its metabolic needs*. In other words, it is a semiotic process through which a system refers to another external one in relation to *its own goals*, which are informationally determined[63] [Subsec. 3.2.2] and thermodynamically supported. A goal denotes therefore an active being-ordered-to-something, and so it is very different relative to constraints and contexts which are not active but play a crucial role in teleonomy. It was the historical merit of F. Ayala to have stressed that teleological explanations are appropriate and indispensable in biology, and cannot be reduced to nonteleological explanations without loss of explanatory power although at the same time they must not be in contradiction with mechanistic explanations.[64] Moreover, Ayala was the first scholar to use both the concepts of teleonomy and teleology as explanatory tools for biology.[65]

Processes like phylogenetic and transcription–translation processes considered in themselves are only teleonomic [Subsec. 8.2.1], while many processes in which metabolic, genetic, and information-selection aspects are involved all together are teleological processes. The former are still semiotic and functional but *without goals*.[66] Given the above definition, it is evident that *only individual organisms* as wholes display a teleologic behavior at an ontogenetic level [Subsec. 7.2], while those biological systems that are not organisms are involved in teleonomic processes of coadaptation. The root of teleonomy is mutual information (channelization), while the root of teleology is information selection (canalization) [Subsec. 2.2.3]. Semiotics is the connection between teleology and teleonomy, because in both cases (with and without information control, respectively) external

[60][*VON HAYEK* 1952, p. 17]. [61][*KLÜVER* 1933].
[62][*NEWELL* 1990, p. 50].
[63]I therefore support an internalist point of view on goals, like that expressed in [*WOODFIELD* 1975].
[64][AYALA 1970]. [65][AYALA 1998b].
[66]Apart from some terminological differences, I support Woodfield's distinction between function and teleology [*WOODFIELD* 1975].

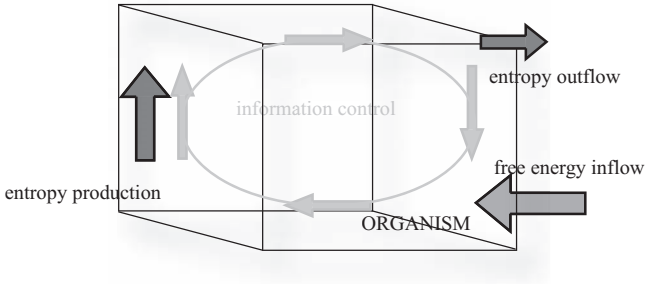

Fig. 8.6 A schematic and qualitative representation of the energetic and entropic fluxes through an organism. Thanks to the circuit of information control [Fig. 8.4], the organism will acquire some free energy and produce some entropy, part of which is downloaded into the environment.

signals are driven and embedded into the relational web of the organism. Moreover, to exercise information control both teleonomic coadaptation and teleology are necessary.[67]

Upon these grounds, organisms could be defined as structurally programmed complex systems thanks to the genetic system, acting as an information-processing device, presenting both information-control and thermodynamic aspects.[68] We can represent this fundamental connection between information and thermodynamic entropy as in Fig. 8.6. Summing up several previous results [Subsecs. 6.1.1 and 7.6.2, Sec. 8.1], we can say: An external, physical disturbance of the organism's homeostasis gives rise to a whole process though which the organism tries (in accordance with statistical laws) to lower the surprise caused by an appropriate codification of that stimulus, and, in doing so, it is able to give rise to an action that is thermodynamically favorable (order-building). The reason for the last step is that the organism will expose itself selectively to those causes in the environment that it expects (is programmed) to encounter, but on the other hand, these expectations are limited to the repertoire of the physical states that the system can occupy, which are precisely those states that are *compatible with its order-building metabolic requirements*. In other words,

- An external disturbance that is potentially disrupting, that is, entropy-increasing, is followed by
- A representational or quasirepresentational information process, which is in accordance with the laws of statistics (and thermodynamics), which gives rise to
- An entropy-lowering, order-preserving action.

The second step is teleologic. The coupling of the second and third steps is the result of natural selection (combined with the constraints of the organism) and is ultimately the effect of a teleonomic adaptation process. We are now in the position to understand the fundamental difference between formal information processes and dynamic entropic processes [Secs. 2.1 and 6.2].

In conclusion, let me stress that this dynamical view of the organism is probably rooted in the general way that nature behaves. I have already stressed that nature operates selections [Subsec. 2.2.3]. An instantiation of this principle is represented by the fact that any detected stimulus will be coded *in a certain way* and will determine the selection of some action or reaction

[67] Therefore, what I here call teleonomy corresponds to what, following F. Ayala, has been called consequence-etiological teleology [DEPEW 2011].

[68] [GILBERT 1982].

by the organism. However, I have also stressed that any stimulus is a perturbation of the organism's homeostasis and every organism would actually like to remain in its own state. The reason is that the organism is a complex system displaying correlations. I have even assumed that there are correlations at several levels of complexity, as an instantiation of a universal correlation principle [Subsec. 6.5.1]. We may then say that any perturbation of a system is precisely a perturbation of the web of the interdependencies that constitute it as a system. I have also stressed that any organism tries to dynamically integrate this perturbation as much as possible into the web of its regularities by lowering the degree of surprise. Due to the generality of both the selection and the correlation principles, my guess is that the latter point also speaks to a general behavior of nature, that could be summarized as[69]

In appropriate conditions, any system is pushed to minimize the distance between the selected option and the less expensive one or to lower the level of perturbation of its constitutive web of interdependencies provoked by the selection.

This principle means, in particular, that the organism maintains its independence from the environment precisely through changing. Indeed, the perturbation, being out of the system's control (otherwise, it would not happen at all), cannot ever be completely washed out. This implies that the system will never fully recover the previous equilibrium state. Therefore, it is a general *principle of dynamicity* and it is precisely the kind of generalization of dynamics we are looking for [Subsec. 2.4.1]. The concept of itinerancy, which is very widely used in the theory of complex and chaotic systems, can be very useful here in helping us to overcome the traditional fixed-point attractors in optimal control theory.

I would like to stress that this principle of dynamicity helps us to refine the notion of system [Sec. 1.1, and Subsecs. 2.4.4 and 3.2.2]: A system is what, when perturbed, shows a tendency to come back to the previous equilibrium state. To this extent, life is deeply rooted in thermodynamics (even if the bridge between entropy increasing and entropy lowering processes is constituted in organisms through an informational step, as explained). Indeed, the Le Chatelier–Braun principle[70] states that any perturbation of a factor contributing to equilibrium induces a compensating change in an opposing factor. The specificity of organisms is, however, in the fact that, thanks to feedback and control mechanisms, such a dynamic integration is or can be *adaptive*, that is, a change for the better [Subsec. 3.2.2 and 6.1.5].

The principle of dynamicity can also be understood as a general reformulation of the complementarity principle that is also true for other classes of systems. Indeed, I have reformulated the latter as a principle dynamically connecting global features and local events [Subsec. 1.3.3]. We can recognize in that case the aspects considered in the dynamicity principle, although at a lower level of generality. This principle also bears a certain analogy to the general principle of persistence, already known in classical physics and philosophy, especially when one takes the generalized form of the Le Chatelier–Braun principle: Any change in the status quo prompts a counterreaction in the responding system.[71] Although in my formulation it is rather a principle of stabilization that takes into account the impossibility of restoring the initial state; a least there is always a noticeable

[69][AULETTA 2010]. For a formulation that contains many elements developed here see also [*ULANOWICZ* 2009a, p. 29].

[70][*ULANOWICZ* 1986, p. 24].

[71]Spinoza's formulation is: *Unaquæque res, quantum in se est, in suo esse perseverare conatur* [*SPINOZA* 1677, Pars III, Prop. vi]. On the contrary, Descartes, Newton, and Leibniz have preferred to stress inertia (the resistance of a system to change its state of motion) while the French school (Maupertuis) has stressed the principle of least action. All these formulations, however, are not so general and are indeed not satisfied in quantum mechanics, where motion trajectories cannot be defined in most cases. Moreover, they lack any active component, which in fact is ubiquitous in nature.

difference when the system is sufficiently complex, and in this respect, the above formulation seems to be more general than other formulations.

8.3 Types of Information Control

In order to control environmental information in the way I have explained, any organism must be able to correct errors.[72] According to Sec. 6.1, true error correction does not consist in simply changing the relationships between elements of a given (eventually representational) pattern, but rather concerns the way in which an organism *associates* this response with some events in the external environment and is able to act accordingly.

8.3.1 Reactive Behavior

A form of information control is provided by the reaction of the organism as a whole to external stimuli. Metabolism would otherwise be impossible because the organism could not regulate the energetic and entropic fluxes it exchanges with the environment, as explained in the previous section. As a general law, once the organism has somehow individuated interesting information about its environment (revealing something about possible resources or dangers), it must act or react accordingly. At this level, the action or reaction might be very elementary, e.g. it might be a very basic form of motion: Bacteria can navigate in certain media following temperature, light, or even magnetic gradients.[73]

At this level, we have response patterns that are directly connected with some operation important for survival. A very good example is provided by the chemotaxis of the bacterium *Escherichia coli*.[74] Such an organism is unable to choose the direction it swims in by itself. Moreover, it is unable to preserve a straight movement for more than a few seconds due to the fluctuations (Brownian motion) of the external fluid. These bacteria alternate tumble and swim phases. In the presence of a chemical gradient (sugar concentration), they will base their motion on this parameter. If this organism senses to swim in a direction, it will preserve a straight line as long as possible before tumbling. If, on the contrary, it senses that it is swimming in the wrong direction, it will tumble sooner. It is a sort of induced "choice" based on the information control that the organism is able to exercise on the environment and therefore on the way it is able to treat certain signals. Concerning the motor output, the helical nature of the single flagellar filaments allows these two types of movement [Fig. 8.7]. Let us now consider the response mechanism.

Chemical gradients are sensed through multiple transmembrane receptors constituted by the methyl-accepting chemotaxis proteins (MCPs), which vary in the type of molecules that they detect[75] [Sec. 7.6]. These receptors may bind attractants or repellents directly or indirectly through interaction with proteins of the periplasmatic space between the exterior and the interior membranes. The signals from these receptors are transmitted across the plasma membrane into the cytosol, where Che proteins are activated. The Che proteins are able to alter the tumbling frequency [Fig. 8.8]. Signals are codified and passed from the transmitter module of one protein to the receiver module of a second protein via phosphotransfer. In the involved pathway, a family of related transmembrane receptors act as the input module by binding either small chemotactic molecules or their periplasmic binding proteins. Once these effectors are bound, the activity of a

[72] [SHAPIRO 2007]. [73] [BLAKEMORE/FRANKEL 1981].
[74] [ALON 2007a]. For another example see [*DRETSKE* 1988, pp. 63–4]. For eukaryotic chemotaxis see [WEINER 2002].

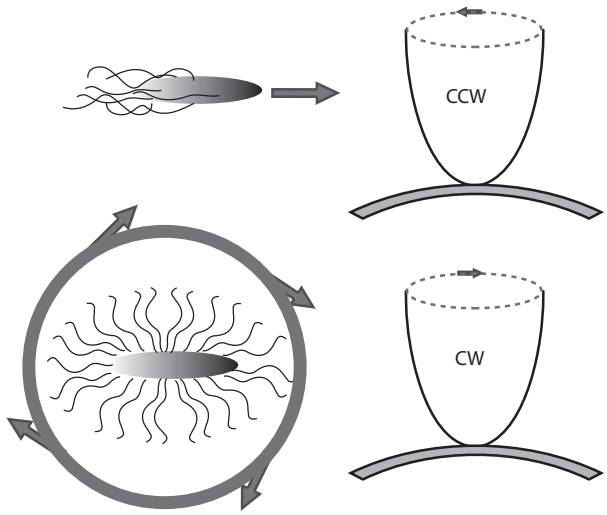

Fig. 8.7 *E. coli*'s movement. Above: Straight swim. In this case, the flagella turn counterclockwise. Below: Tumbling. In this case, the flagella turn clockwise.

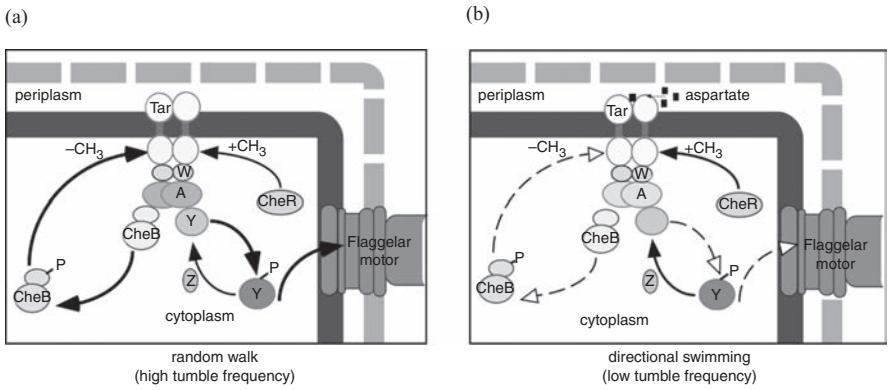

Fig. 8.8 Activated forms of the proteins are shown in the darker color and solid arrows are used for indicating activation. (a) The high level of phosphorylated CheY, due to the activity of CheA, increases the frequency of switching to clockwise flagellar rotation and thus determines tumbling. (b) When a receptor binds ligand and/or is unmethylated CheA is inactive. The levels of phosphorylated CheY are reduced leading to more counterclockwise flagellar rotation and more running. With CheB inactive, the methyltransferase activity of CheR serves to decrease receptor sensitivity. Adapted from [JURICA/STODDARD 1998].

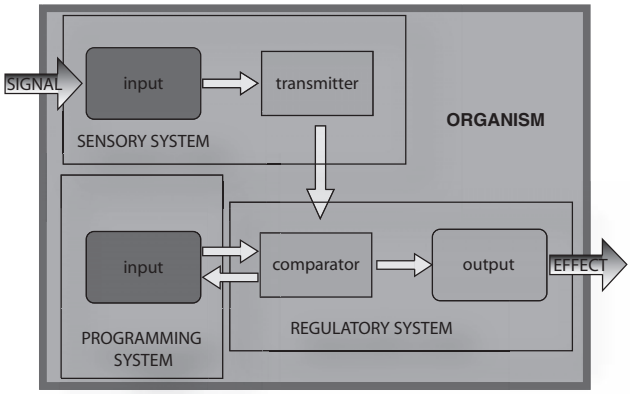

Fig. 8.9 The sensory component processes an environmental signal through its input module to activate the transmitter module. Phosphoryl transfer from the transmitter to the receiver–comparator module of the response regulator component, jointly with other inputs coming from processes inside the programming (genetic) system, activates the output module, and therefore triggers a final motor output [see also Fig. 5.10]. Inspired by [JURICA/STODDARD 1998].

transmitter histidine kinase (CheA) that is associated with the cytosolic domain of the receptor(s) is rapidly modulated. Changes in the activity of this kinase lead to transient increases or decreases in intracellular levels of phosphorylated CheY (the *response regulator*) which directly affects flagellar rotation and the frequency of their reversal. Slower habituation of this response, effected at the level of receptor signaling, is induced by the reversible methylation and demethylation of a specific group of glutamate residues within predicted coiled-coil regions of the receptor cytosolic domains. These covalent modifications are catalyzed by an S-adenosylmethionine-dependent methyl-transferase (CheR) and a partner methylesterase (CheB), having opposite effects in damping or increasing the signal respectively. CheB, which is another response regulator, is also a substrate for CheA kinase activity. The protein is most active as a methyl-esterase in the phosphorylated form and further serves as a feedback regulator of signaling.

Apparently, this behavior is very mechanical and does not seem to display the elements that we expect from true information control. However, its simplicity should not induce us to misanalyze the situation.[76] The first point to stress is that here a regulatory component is also involved [Fig. 8.9], and is determined by the opposite effects of the proteins CheB and CheR. This prevents us from considering the whole as a pure input-output mechanical engine. However, this does not suffice for establishing information control. Here the two crucial questions are: Is there a true comparator nested in this regulatory part? Is there some form of information codification here? When we look more closely, we shall remark that [Sec. 8.1; see also Eq. (7.16)]:

- To make a temporal comparison of chemoeffector levels (a pure informational operation), the bacterial cell requires a sensory adaptation mechanism that cancels chemoreceptor signal outputs in a static environment, no matter what chemoeffectors may be present[77] (whether attractive or

[75][JURICA/STODDARD 1998]. [76][JURICA/STODDARD 1998]. [77][PARKINSON 1993].

repulsive). This enables the bacterium to reset the threshold sensitivity of the signaling system in order to detect any *new* change (any surprise) in the chemical environment. It is a true information-erasing mechanism [Subsec. 2.2.1] necessary for information-acquiring, a kind of very elementary representational function.

- Another very important element is reaction timing: Because Brownian motion of the fluid medium can randomly reorient the bacterium, this requires very short response latencies. It is here that genetic (instructional) factors play a role by enhancing and dampening protein production. It has indeed been observed that the protein CheZ plays a very important role in enhancing the rate of CheY (which is directly involved in the change of motion) dephosphorylation.[78] Morevoer, in the overall regulation the transcription factor NarL also plays a role.
- Finally, the output activity of the protein CheA is dependent not only on the inputs but also on a feedback circuit that allows to maintain such an activity as constant *against* environmental inputs (it is an anti-feedback mechanism).[79]

This means that the regulation and selection activity is indeed very sophisticated and that informational aspects are involved in signal transduction and information erasure, in the comparison of chemoeffector levels, in DNA activation or repression, and keeping the level of a certain activity no matter which the environmental inputs are. This also means that several comparisons are executed along the whole path going from transduction to reaction: Between concentration levels of nitrate or phosphates and the required ones, between the proteins that are at available and the required ones, and so on. The single segments and operations can obviously be explained in pure chemical terms. But the *whole* pathway or network involved here, in which both cybernetic and informational aspects play a major role, is an instance of information control [Subsec. 6.5.1]. I stress again that, to accomplish all of this, it is not necessary for there to be a "brain" as such.[80]

Forms of reactive behavior are still present in any higher forms of life and behavior [Subsec. 5.3.1], since, as I have already stressed [Subsec. 3.3.1], in any change occurring in the nervous system after a stimulus, we must distinguish between a reaction and a plastic change (the plastic change is perception, and perception occurs in any nonpure reactive representational activity). Examples in animals are the reaction of the male stickleback to red spots, which means that rivals are present, or the reaction of frogs to flying black dots, which might be insects.[81] Obviously, especially in the case of the frog, it is very likely that we already have some form of true representation. In this sense, reaction is integrated here into a more complex instinctive ability. When there is a reactive behavior, we can indeed speak of semiotic marks (stable aspects or properties connected with a certain object having survival value) able to trigger the appropriate reaction.

We can generalize these results by saying that any elementary reaction behavior presupposes three fundamental aspects: (1) Sensation, (2) response selection, and (3) response production.[82] Response selection is actually the outcome of regulation and sensation, a pattern-production process (a sort of search). Therefore, in response production comparative activity is also involved. This is the reason why humans can perceive two stimuli simultaneously and even generate two actions simultaneously, but *not two selections*. The minimal time interval needed to select a second response is called a psychological refractory period.

Although reactive behavior is very limited in its control on the stimuli coming from the environment, this does not mean that the behavior is instructed by the environment, but only that bacteria have a very narrow range of possible reactions evolutionarily tuned into very specific

[78][RAO *et al.* 2004]. [79][AULETTA 2011]. [80][BEER 1995a].
[81][*MILNER/GOODALE* 1995, pp. 6–11]. [82][*PASHLER* 1998] [*WILLINGHAM* 2001, pp. 138–9].

forms of stimulus,[83] even if they already show an astonishing capacity to extract covariant signals.[84] Regularities of the physical environment are progressively encoded into the structure of networks.

8.3.2 Active Behavior

Things stand in a different way with eukaryotes. If we consider unicellular organisms like amoebae, we see that there are important structural differences with prokaryotes. I do not need to consider here the modularization of the three main systems (metabolic, genetic, and selective) [Sec. 7.2], but I shall confine my examination to aspects relevant for eukaryotes' dealing with the environment. Eukaryotes have flagella, but these are made of a protein (flagellin) that is different from the bacterial analogue. This means that eukaryotes' flagella are capable of *autonomous* sinusoidal movements. Moreover, eukaryotes also use other protuberances for their movement that are called pseudopods, which allow them to crawl (and even enable them to ingest material particles of food, a process called phagocytosis). Their locomotion presupposes a cytoskeleton (very often either absent in or present in bacteria in a very rudimentary way). It is this internal structure that allows for the appropriate and controllable deformations of the cell's shape that are necessary for autonomous locomotion, especially contraction-movement, which is ultimately self-generated.

Therefore, eukaryotes are not only able to react to external signals but also to *act* on their own. It is very important not to confuse the issue of action in general with that of spatial movement (which is a specific expression of the former). Indeed, plants very often do not move but still act. By *action* I understand the capacity of an organism to undertake certain procedures on the external environment that are not started by an environmental signal, nor driven (where by this expression I am not excluding instructed behavior only) by those signals (as in the case of reactive behavior only) even if they may make use of different forms of environmental cues. In other words, they are endogenously generated.[85] Also W. Wundt stressed a similar point when he observed that contractile movements arise in eukaryotes sometimes at the instigation of external stimuli but sometimes also in the absence of any apparent external influence, so that they seem to be spontaneous actions of the lowest forms of life, and are therefore the results of forces that are resident in the contractile substance itself.[86] This is what makes eukaryotes *autonomous agents*. We have here a higher manifestation of what I have already pointed out for any organism: Even if the system is partly relying on external inputs, its circular reactions as displayed by information control will also affect the single components, and in doing so subsequent behavior's trajectories will be affected accordingly, so that it is not strictly true that any state of the system depends solely on those inputs[87] [Sec. 8.2]. We shall see how this general ability, in subsequent evolutionary history, shall take two different paths in plants and animals.

I finally wish to stress that autonomy has different dimensions and gradations.[88] A creature is more autonomous if:

(i) Its behavior is not directly determined by the environment even when there is an external stimulus but mediated by inner mechanisms partly dependent on its previous history,
(ii) The control mechanism is self-generated rather than (phylogenetically) prefigured, and
(iii) The inner mechanism is also dependent somehow on a particular information control and selection.

[83][*ROMANES* 1882, pp. 3–4]. [84][TAGKOPOULOS *et al.* 2008].
[85]A possibility that Lamarck acknowledged for animals [*LAMARCK* 1809, I, p. 82].
[86][*WUNDT* 1893, pp. 29–30]. [87][*ULANOWICZ* 1986, pp. 55–6]. [88][BODEN 1995].

The second condition is the manifestation of teleological top-down processes in the organism, that is, of the ability to modify available information-control mechanisms if the environmental pressure is sufficiently strong.

8.4 The Self

One of the most important features of life is that a living organism is a structured and somehow self-contained system. Any organism is separated to a certain extent from its environment upon which it has some effects in a selective manner. This is what allows the organism to exercise an informational control upon the entropic fluxes between itself and the environment, which constitutes an important difference relative to any inorganic self-organizing system, that cannot control these fluxes and is therefore directly dependent on environmental parameters [Sec. 8.2].[89] This means that organisms may be understood as instances of a specific class of physical system,[90] besides quantum-mechanical and classical-mechanical ones: Biological systems, showing complexity [Secs. 6.3–6.6], but also instructional and representational information processes [Sec. 8.1]. I wish to stress here that the future of the biological sciences depends very much on their ability to avoid the opposite dangers of mechanist reductionism and anthropomorphism and to find their own independent conceptual and methodological foundations.

Therefore, an organism is a biological system characterized by a sharp separation between self and nonself. By *self* I mean the systemic totality of the living organism,[91] excluding everything that lies outside of it or that does not fully depend on this systemic organization. I wish to stress that the self is a *functional unity*, and for this reason it is only possible at a biological rather than physical or chemical level. The distinction between self and nonself is characteristic not only of the basic structures of life, but reaches all levels of organization and complexity, up to consciousness or the human language. So, an organism, by integrating in a new way systems that can be found separately in the physical world [Sec. 7.2], somehow *duplicates* the world and is constituted as a universe apart. The separation between a self and a nonself induces a new situation in our world. Prebiotic physical systems must somehow "find their place" in a given context, as they are put into some medium. For instance, any system will reach (through entropic exchanges) the temperature level of the environment and find a final equilibrium [Subsec. 2.4.1]. The same is true for biological systems. The only difference consists in the fact that, since there is now a boundary between self and nonself, there is a systemic self-referring circle of reactions and control mechanisms[92] such that the operation of an organism finding its own place in the environment happens in two different directions: *Accommodation* of the organism to the external environment (whose root is teleonomy) and *assimilation* of the environment by the organism (whose root is teleology)[93]: These two processes together constitute adaptation. When I say assimilation, I am not only speaking of feeding but of the ability of any organism to monitor and carve out the environment according to its metabolic needs. Only organisms show assimilation processes. As we shall see, very often we have a combination, an optimal trade-off, between these two opposite processes, as stated above in the principle of dynamicity [Subsec. 8.2.7]. In this way, any organism maintains its independence from the environment not through absence of change but *through change*, as beautifully understood

[89]This was well understood by Thomas H. Huxley [*RAMELLINI* 2006, pp. 37–8].
[90][AULETTA 2006c]. For one of the first ideas in this direction see [*AGENO* 1986].
[91][*LEDOUX* 2002, pp. 26–1].
[92][*MATURANA* 1970, p. 10]. For some early intuitions of this mechanism see [*BERNARD* 1865, pp. 109–10 and 127–32].

by J. Woodger.[94] This dynamicity is rooted in the fact that, apart from some crucial hubs, many components of the organism can change or even fail without disrupting the relative functionality [Subsecs. 6.3.2 and 8.2.4].

I would also like to remark that metabolism in the proper sense of the word is not possible without a separation between self and nonself. Indeed, viruses, which do not have a metabolism, only show a very rough separation between self and nonself and a rudimentary reproduction program. The point is that viruses do not constitute a self, since (1) they are ever-growing systems without self-regulation (they are not able to constitute a cybernetic circle, a notion that I shall consider in this section), and (2) succeed in reproducing themselves not by their own mechanisms, but rather by enslaving an external system (the host cell, which is, instead, able to autonomously self-replicate), with a characteristic reversal of the mechanism characterizing nonbiological self-regulating systems (which are dependent on environmental context) [Subsec. 6.5.3]. A reversal, however, is not a truly new solution. This point shows how dangerous was the tendency of the last decades to put the emphasis solely upon the self-replicative aspects of biology, so as to conceive the organism essentially as a bag containing DNA[95] [Sec. 7.1].

It is important to consider that none of the main systems constituting the organism can work in isolation. In this case, they would lose even their meaning and function. The genetic system works in order to (contribute to) determine the selection system and the regulatory system (for deploying information control). In other words, genes do not reproduce genes, otherwise the organism would reduce itself to genetic information or to self-replication of genetic information [Sec. 7.5]. The metabolic system cannot regulate anything if there is not a genetically established background already in place and a selection system which is appropriately tuned. The selection system cannot work if there are no reference items relative to which it is programmed to control and select information. Moreover, without the regulatory activity provided by the metabolic system the organism would finally be controlled by the environment and not *vice versa*. The main point here is that this global system consisting of three subsystems (which are sometimes, as we have seen, still further articulated) is a true self-organizing complex system: It provides for its own self-production and even self-reproduction and production [Fig. 8.10]. Therefore, we have three main forms of change[96]:

- *Self-production* is the emergence of a self-referential perspective through which the system is able to care for its self-maintenance.[97]
- *Self-reproduction*[98] is the synthesis between the metabolic and the genetic aspect of an organism.
- *Production* is the effect of the biological self on the external environment and the result of a connection between the control and the genetic systems. This is the birth of agency [Sec. 6.6], i.e. of an entity provided with *goals* [Subsec. 8.2.2] that is able to produce effects in the world that leave a trace of its actions.

In mechanical terms, this triad could also be understood as (a) maintaining and repairing, (b) building, (c) working.[99] It is a true feedback system in which any pair of subsystems mutually interact, and for this reason a *whole circuit* is established: The succession no longer has relevance

[93][*PIAGET* 1967, pp. 169–71]. [94][*WOODGER* 1929, p. 483]. See also [*RAMELLINI* 2006, p. 59].
[95][*DAWKINS* 1976].
[96][*LAMARCK* 1809, I, pp. 91–2; II, pp. 125–6]. Roederer has proposed that life should be considered a triad: Encapsulation (containment), self-adaptation (metabolism), and reproduction (genetics) [*ROEDERER* 2005, p. 126].
[97][*OPARIN* 1957, pp. 354–7].
[98][*OPARIN* 1957, pp. 359–63]. [99][*VON UEXKÜLL* 1926, p. 121].

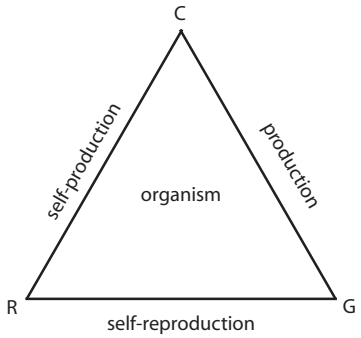

Fig. 8.10 The three modalities of production of the organism establish a true cybernetic circle that can be cast in this way:

We have essentially three centers: The center of metabolic activity is the regulatory (metabolic) system (R), represented by ribosomes in the cell and by the hormonal system in higher organisms; the center of the heredity system is the genome (G) and in higher organisms the whole genetic and epigenetic system; finally, the center of the information-selection system is protein mechanism controlling signal transduction or, in higher organisms, the sensory CNS (C).

The connection between the heredity and metabolic systems guarantees the *self–reproduction* of the organism (this is especially evident for sexual organisms that need specific hormones for reproduction, but it is also true for pure binary fission).

The connection between the regulatory system and the information-selection system guarantees the *self-production* of the organism (its maintenance, the control of food and free energy sources, structural information, and so on).

The connection between the genetic system and the information-selection system guarantees the *production* by the organism—that is, all its activities that are directed toward the modification—of its immediate environment and to assure an appropriate niche in it.

In conclusion we can speak of a triad replicator–interactor–structurator as already anticipated in Sec. 7.1.

here since we have a single triadic relation instead of three dyadic relations [Subsec. 2.4.4]. This is what I call a *cybernetic circle*, which is a far more complex form of organization than a simple feedback, self-increasing circle, like some of those considered in the previous chapter. The former is a *functional* circuit, e.g. that shown in Fig. 8.4. The latter form of circuit has no self-regulative mechanisms (as it happens for viruses). These are grounded on both negative feedback and antifeedback circuits. Due to such regulation, only certain parts of biological systems are unlimited and self-potentiating (and any regulative aspect, like genetic expression regulation in those feedback circuits, is a consequence of the regulative activity of the *whole* organism). For instance, germ cells are totipotent. On the other hand, an interesting but dramatic effect of positive but uncontrolled wave-like feedback is represented by cancer. Therefore, negative feedback always acts as a stop signal that is discontinuous and point-like in principle, while antifeedback, combined with the other forms of feedback, as a regulator.

We have previously seen that it is only organisms that are stable, while subordinate or superordinate forms of biological systems are not [Sec. 7.2]. Where is the difference? Why are these forms more subject to anarchic tendencies? The reason is that organisms alone have both a metabolism and an informational system as well as a strict separation between energetic (through metabolism) and informational aspects, whereas all subordinate forms of biological system lack

a metabolism, while the superordinate ones have "subsystems" in which *both* metabolic and informational aspects are present (for instance, single organisms in a society of organisms). It is this specificity of the organisms which allows for the constitution of their self, and gives rise to a cybernetic circle, in which all functions are separated but all subsystems interconnected.

I recall that this triadic structure of the organism is a higher manifestation of the basic structure of dealing-with-information that we have learned by studying quantum information and the correction it provides to the traditional, classical theory [Subsec. 2.3.2 and Sec. 6.1]. As we shall see in the next three chapters, self-reproduction is the basis of *phylogeny*. Production is the basis of *ontogeny*. Here and in the following, I understand by ontogeny the whole pertaining to the individual's life. Finally, self-production is the basis of *epigeny*, the first phase of development (and of ontogeny), through which the organism produces its own existence.

8.5 Concluding Remarks

The main results of this chapter are:

- Life is essentially a semiotic activity. Any sign is an indexical relation with a referent that is associated with an iconic aspect.
- Biological systems display a teleonomic form of causation, in which different paths can lead to the same result (the task or the final state), by integrating information processes from the inside of a biological system and external cues. It rests on circuits integrating positive, negative, and antifeedback.
- Organisms also display teleological causation when top-down processes and information control for some metabolic need are involved.
- Any information-control mechanism combines instructional information with information acquisition.
- Any organism's information control establishes functional equivalence classes between operations.
- Functions are both related to and partly independent from structures. Structures instantiating functions show degeneracy.
- The organism is a biological self which is both autarchic and dependent on environmental cues. The cycle of self-reproduction, self-production, and production is a true cybernetic circle in which any subsystem of the organism is connected with the other ones.
- We have learned a fundamental principle ruling nature: There is a dynamic trade-off between local perturbation and the correlations constituting any true system.
- This previous principle has a corollary for biological systems: Change is intrinsically rooted in the concept of life and manifests itself in the ability of the organism to deploy an integration between accommodation and assimilation.

This chapter, and in particular the dynamicity principle with its consequences, sets the foundations of the following three chapters and also establishes a framework for Chs. 12–17, which deal with representational semiotics.

9
Phylogeny

Before entering into details, I wish to recall what has already been mentioned in the introduction to Ch. 7, namely that the issues concerned in what follows now and in the next two chapters should be understood as a critical evaluation of phylogeny (this chapter), ontogeny [Ch. 10], and epigeny [Ch. 11] from the point of view of dealing-with-information. In other words, it is not my aim to give a systematic account of these difficult matters but only to take advantage of some important results in this field for the goal of further exposition. It is well known that T. Dobzhansky considered evolution as crucial for understanding any aspect of biology.[1]

After having introduced the general concept of evolution, I shall consider the action of natural selection. The issue of populations of genotypes and phenotypes as well as the problem of variation are examined. Then, I shall provide a short summary of the concept of species. I shall also consider the relations between evolution and complexity theory and a very helpful approach to evolution: Game theory. Another important question that shall be dealt with is: How many systems of inheritance are there? Then, an investigation into the problem of entropy and order at different time scales follows. Finally, the Baldwin effect is discussed and some general principles of evolution are considered.

9.1 Evolution as Cumulation of Phylogenetic Changes

We have already established two key general notions: Any system is both informationally shielded and somehow correlated with the environment [Chs. 2 and 6]. These two features have a particular relevance for biological systems and are integrated in a higher dynamical process [Ch. 8]. We have also seen that self-reproduction is a natural consequence of a complex system that displays these general features [Secs. 7.1–7.2 and 8.4]. *Self-reproduction* guarantees both for hereditary transmission, i.e. self-replication [Sec. 7.5], and variation [Subsec. 7.4.2] which makes the genetic system a true information processor:

- If *self-replication* were the sole mechanism of connection between different generations, the result would be, at the most, the degeneration of the initial "prototype" (as in any physical process of copying), which is the case for any propagation of signals from a source [Subsec. 2.3.1].
- It is *variation* (through mutations and sexual and asexual DNA recombinations) together with the capability of the organisms to assimilate such variations in their self-production processes that guarantees that each generation is *its own* prototype.

However, self-reproduction is only one side of the coin. Evolution would not occur without corresponding action of the environment on the organisms, since variation would very quickly diversify

[1] [*DOBZHANSKY* 1973].

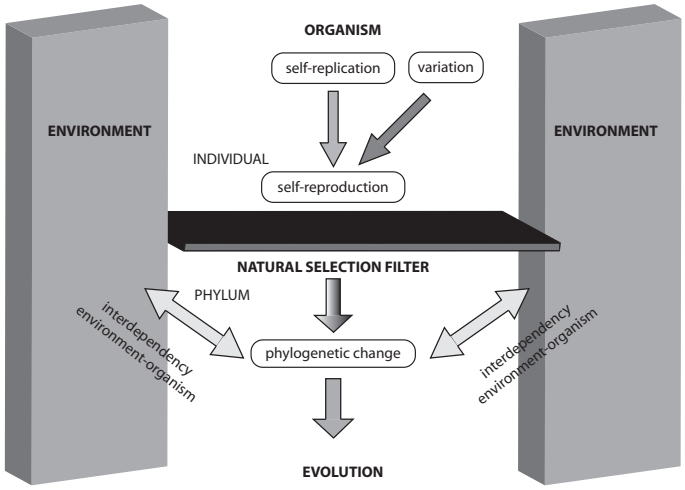

Fig. 9.1 The scheme for evolutionary change. Note that both several phylogenetic changes and several interdependencies between organism and environment are necessary in order to have evolution.

species too much, which would endanger their survival. This action is represented by natural selection. It is natural selection, combined with self-reproduction, that fixes (1) phylogenetic variation and (2) the basic interdependency between the environment and the organism (an aspect of which is represented by the adaptation of the organism to the environment). The *phylogenetic accumulation* of variations and interdependencies with the environment is what we call *evolution* [Fig. 9.1]. Therefore, in nature, the evolutionary process occurs when the following conditions are satisfied[2]:

(1) There is a population of self-replicating entities.
(2) There is some variety (random mutations and DNA recombinations) among the self-reproducing entities.
(3) There is some phenotypical difference in ability to survive (in a given environment) associated with this variety, so that the organisms survive and transmit those parts of their heritage that are more apt to survive (phylogenetic change).

I wish to stress that when we speak of evolution we should not understand it merely in terms of straightforward progress.[3] In fact, the course of evolution is much more complicated than a pure rectilinear process, presenting ramifications[4] and possible partial backward fluxes.[5] This does not exclude the fact that certain parameters could have shown a significant increase in their value during evolutionary history; a problem we shall turn to now.

9.2 Natural Selection

According to Mayr,[6] Darwin's theory consists of 5 hypotheses: (1) methodological naturalism, (2) transmutation, (3) monophyletic descent, (4) natural selection, (5) causal pluralism. Here,

[2][LEWONTIN 1970]. [3][*GOULD* 1996]. [4][FORD DOOLITTLE 1999].
[5][TEOTÓNIO/ROSE 2000] [BULL 2000]. [6][*MAYR* 1988].

I understand naturalism as a general requirement for scientific enquiry and causal pluralism as an expression only of the ensemble of possible effects of the environment on the organism, which ultimately all have selective natures. Therefore, it is only points (2)–(4) that I strictly adhere to, according to the previous examination.

9.2.1 Darwin's Contribution

Let us first present a brief historical sketch of the development of the evolution theory.[7] The main point for Darwin[8] was to prove that organisms are ordinarily subject to environmental stress and competition for resources, given the tendency of any population to grow [Subsec. 7.5.2]. He found in classical economics[9] the idea of a selection of competing individuals in a population.

Generally speaking, natural selection consists of (1) The action of the physical environment on species, (2) elimination by enemies of other species (predators, parasites, and so on), (3) competition among individuals of the same species.[10] Between-species selection often turns out to be very weak compared with selection between individuals of the same species.[11] Note also that it was clear to Darwin that one should add sexual selection to natural (environmental) selection.[12] Many shapes in nature are only understandable by taking into account the game of mating and reproduction.[13] We shall consider sexual selection in various parts of this book.

It is also important to stress that, according to Darwin, when a species becomes dominant in a given environment (and so is subject to less selective pressure), it has the tendency to widen and to differentiate much more than less dominant species.[14] As a consequence, there is a higher level of variability, inducing a breakup of the large species into smaller species. This confirms what was said in the previous section: Variation to a certain extent is opposite to the selection pressure in the generation of new species. Moreover, according to Darwin, natural selection is a differential and necessarily local change, since it punishes one individual or species by giving advantage to at least another individual or species.[15] In this way, at a global level, the biosphere is not immediately touched by natural selection.

Notwithstanding these provisos, in Darwin's contribution to biology there is an important novelty relative to the classical framework of science. The classical framework, dominated by classical mechanics, was deterministic, whereas Darwin introduced chance into science in order to explain the emergence of a pool of differences upon which natural selection could act. In this way, chance together with natural selection was the mechanism that could explain the diversity of species in space and time. It is important to stress that chance enters into Darwin's theory not only as source of variations[16] but also as the contingent match (interdependency) between variations themselves and environmental constraints (Mayr's point (5) above). In fact, the cornerstone of Darwinism is the idea that environmental changes and mutation in organisms are two independent processes that become involved with each other only through the selective action of the environment on living beings. In this way, there is no positive (instructive) feedback but only a negative (selecting) one. This is its main difference with previous evolutionary theories, such as Lamarckism,

[7]For this matter see [*DEPEW/WEBER* 1995, pp. 1–160], whose results I synthesize here.
[8][*DARWIN* 1859, p. 55]. See also [WALLACE 1858] [*LAMARCK* 1809, I, pp. 99–101].
[9][*MALTHUS* 1798]. [10][*LLOYD MORGAN* 1891, pp. 79–80]. [11][MAYNARD SMITH 1996].
[12][*CRONIN* 1992]. See also [*LLOYD MORGAN* 1891, pp. 197–209]. [13][*GOULD* 1985, pp. 40–55].
[14][*DARWIN* 1859, pp. 48–51]. [15][*DARWIN* 1859, pp. 73–4].
[16][*DARWIN* 1859, pp. 12–13 and 35]. On this especially De Vries' contribution is important, giving rise to a first saltationist theory [*DE VRIES* 1901–3, *DE VRIES* 1904].

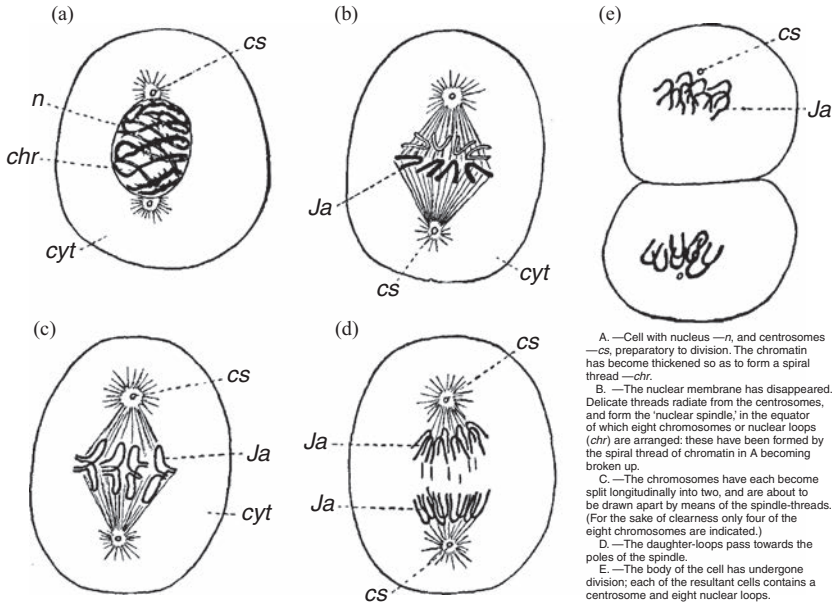

Fig. 9.2 Weismann's explanation of nuclear division [*WEISMANN* 1893, p. 27].

which were instructive, and therefore assumed that living systems are informationally permeable to external instructions.[17]

There is a distinction made currently[18] between different forms of selection, in particular between stabilizing, directional, and disruptive selection. Stabilizing selection is strictly related to functions.[19] Another similar classification is stabilizing, directional, random natural selection.[20] No matter which classification we choose, selection is a general mechanism of nature and the natural consequence of the informational closeness of a physical system together with its interaction with an environment. In other words, it is not confined to biology. Bénard cells are examples of physical selection [Subsec. 6.5.3], whereas Belousov–Zhabotinsky reaction is an example of chemical selection [Appendix to Ch. 6]. This is due to the central role played by chance in any situation in which several systems or behaviors interact and compete. This is what I have called a generalized Darwinism [Subsec. 2.2.6].

9.2.2 Neo-Darwinian Synthesis

I shall not discuss here the modern evolutionary synthesis and the related basic connections between Darwin's theory of natural selection and Mendel's genetics.[21] On the contrary, I would like to stress the centrality of Weismann's work underlying this wider theoretical conception.[22] Weismann, on

[17][*LAMARCK* 1809, I, p. 235]. [18][*SCHMALHAUSEN* 1949] [*FUTUYMA* 1998][*RIDLEY* 1993].
[19][*CONWAY MORRIS* 2003, p. 115].
[20]See [HENDRY 2008] for new methods for taking into account directional selection.
[21][MENDEL 1866]. The obscure work of Mendel was rediscovered only toward the end of the 19th century, when De Vries had already developed some of the main ideas of genetic heritage by himself [*DE VRIES* 1901–3, *DE VRIES* 1904].
[22][*WEISMANN* 1889, I, pp. 76–105] [*WEISMANN* 1893, pp. 2–3, 11–12, 22–30].

the basis of the data available at that time, realized that the cells that produce the so-called germ plasm (which is itself segregated, namely in the nucleus), or gametes (such as sperm and eggs in animals), separate from the somatic cells that go on to make other body tissues at an early stage in development [Subsec. 7.5.2]. Since he could see no obvious means of communication between the two types of cells, he asserted that the inheritance of acquired characteristics was therefore impossible; a conclusion now known as *Weismann's barrier*.

In so doing, Weismann rejected two of Darwin's ideas: (1) the possibility of transmission of acquired characteristics (which is still present in Darwin's original formulation[23]), and (2) the so-called pangenesis[24] (which is the idea that inheritance of characteristics due to the migration of particles carrying them, going from all cells of the body to the reproductive cells). These two issues are related, but, as we shall see, Weismann's rejection of pangenesis is also connected with his opposition to the concept of epigeny. This picture is far from obvious. In particular, more recent works have shown that the separation between somatic and germ cells is not universal (for instance, it is not true for plants) and that there are three main modalities of germ cell formation: Early and rigid germ-line determination (the variety considered by Weismann), late germ-line determination, and somatically derived germ cells.[25] In the following, I shall consider some of these developments and also have occasion to come back to the main tenet of Weismann.

9.2.3 How Is Natural Selection to be Understood?

There are several problems with the original formulation of evolution theory and its neo-Darwinian refinement. We will consider some of them in the next pages. However, here I shall be concerned only with those connected with natural selection.[26] The first one was the relation with the environment. Lewontin showed that this relationship cannot consist merely of a passive registration by the organism of certain environmental changes but is rather an interactive process. Moreover, organisms construct their environment.[27] To a certain extent, this can be understood as a form of inverted Lamarckism (this will be the subject of the next chapter).

Furthermore, there was an ambiguity in Darwin's original formulation. Are the survivors of the selective action of the environment actually the fittest? In a strict competitive model such as the Malthus's model of economics (which inspired Darwin's own work[28]) this is necessary.[29] However, in this way, again, we risk to surreptitiously introduce some form of instructionalism, while today the constraints of survival and reproduction are acknowledged to be too weak to be able to provide an account of evolution by themselves. Here, we must pass from a prescriptive to a proscriptive theory: Selection discards what is *not compatible* with survival and reproduction but there are many possible solutions that are compatible with those constraints. This is why there is *biological diversification* in spite of natural selection, a circumstance which cannot be explained by mere economical considerations. The consequence is that there are only *satisfying* solutions and *not optimal* ones[30] [Subsec. 8.2.1].

9.3 Populations of Genotypes and Phenotypes

In the words of M. A. Nowak, neither genes, cells, organisms, nor ideas evolve. Only populations can evolve.[31] This tenet is also expressed in Mayr's definition of species, as we shall see.

[23][*DARWIN* 1859, Ch. 1]. [24][*DARWIN* 1868, Ch. 27]. [25][*JABLONKA/LAMB* 1995, pp. 37–48].
[26][*DEPEW/WEBER* 1995, pp. 359–91]. [27][*LEWONTIN* 2000, p. 48]. [28][*DARWIN* 1859, Ch. 4].
[29][*MALTHUS* 1798]. [30][JACOB 1977]. See also [*VARELA et al.* 1991, pp. 185–214]. [31][*NOWAK* 2006, p. 14].

9.3.1 A Combinatorics of Discrete Units

Mendel discovered the mechanism of inheritance on the basis of experimental data (with peas) and a pure abductive (hypothetical) reasoning.[32] He discovered in fact that traits that were silent in several generations of peas could reappear much later. This could be explained only by assuming that there must be a genetic mechanism that is both separated and relatively independent from the mature organism (otherwise certain traits would progressively disappear). In this way, the distinction phenotype/genotype was established for the first time [Subsec. 7.5.1]. I understand the term *phenotype* as denoting a system composed, at a very basic level, of a structural part, which also guarantees the sensory transduction, a regulatory metabolic subsystem, and a decisional subsystem [Secs. 7.2 and 8.4]. Mendel's discovery of the mechanism of genetic inheritance had some surprising consequences.[33] Indeed, the possible combinations of traits in successive generations follow the binominal distribution given by the Hardy–Weinberg formula (7.5) (the physical model used here was the theory of gases), that states that both allele and genotype frequencies in a population remain constant or are in equilibrium from generation to generation unless specific disturbing influences are introduced, like assortative mating, particular selective pressures, limited population size, genetic drift, and gene flow [Subsec. 8.2.7].

J. Haldane and R. Fisher drew the conclusion that mutation is not the principal agent of evolution and that combination of traits is sufficient even if the selection pressure is very weak. Natural selection rather plays the role of fixing traits. It is important to distinguish here between a succession of favorable deviations from the laws of chance, and the continuous and cumulative action of these laws. It is on the latter that the principle of natural selection relies. If an organism is really adapted to the place it fills in the environment to a high degree, the adaptation will be constantly menaced by undirected agencies (representing chance).

Charles Darwin still accepted a blending theory of inheritance. However, Fisher quickly understood that, with blending inheritance, bisexual reproduction will tend rapidly to produce uniformity without variation [Sec. 9.1]. Therefore, causes of variability must continually be at work, and a principle of variability [Subsec. 2.3.2] must be added to a principle of uniformity or of persistence, according to which any system remains in the same state if not perturbed[34] [Subsec. 8.2.7]. As mentioned, Fisher[35] understood that random mutations cannot determine the direction of evolutionary change; their function is rather to maintain the stock of genetic variance at a certain level, which in turn can determine the *speed* of evolution. Fisher[36] was the first to introduce statistical models on the population of genes and the concept of fitness maximization on the model of entropy maximization[37] [Subsec. 2.3.4]. *Fisher's law* states that the rate of increase in fitness of any organism at any time is equal to its additive genetic variance in fitness at that time. However, this result still needs to be fully assessed, since it cannot imply that the average fitness of a population always increases. Still, he confirmed Darwin's idea that more numerous species are more variable. The crucial point is that mutant genes can be immediately dominant, but in general they are recessive. Being potential at an initial state, mutations eventually have time to become dominant later on (a point also stressed by Baldwin, as we shall see). Indeed, visible effects of mutations depend both on the gene substitution itself and on the whole genetic endowment of

[34][PEIRCE 1891, p. 296] [PEIRCE 1892a, pp. 308 and 310] [*PEIRCE CP*, 1.159, 1.174, 1.302, 1.405, 5.119, 6.91, 6.97]. See also [*DE VRIES* 1901–3, *DE VRIES* 1904] [*BOREL* 1920]. Lloyd Morgan was among the first scholars to study individual and interspecific variability, in particular by analyzing phylogenetic modifications of the bat's wing [*LLOYD MORGAN* 1891, pp. 63–75].

[35][*FISHER* 1930].

[32][MENDEL 1866]. [33][*DEPEW/WEBER* 1995, pp. 169–328]. [36][*FISHER* 1930]. [37][*FRIEDEN* 1998].

the organism. Consequently, the organism can evolve in such a way as to modify its reaction to any particular gene substitution. This means that many variations can be stored in a potential or latent form (the inactivated potential information codified by the DNA), as already postulated by Mendel, and nicely confirmed by more recent results. It is important to consider that the mutation rate needed to maintain a given amount of variability is, if one assumes a combinatorics of discrete units, many thousands of times smaller than the one required in assuming the blending theory. Therefore, all the main characteristics of the Mendelian system flow from the assumption of *particulate inheritance* of the simplest character [Subsec. 2.4.1]. Fisher succeeded in calculating the reproductive value of an organism at a given age as a function of a parameter that measures the fitness needed in order to survive. Summing up, long before the genetic code was discovered, Fisher understood, on the basis of pure statistical calculations and some reasonable guesses, that the genome must represent a combinatorics of discrete unities, which is one of the fundamental requirements of codified information [Sec. 7.4].

Schrödinger followed a line of thought that is to a certain extent parallel to (and independent from) Fisher's research.[38] He discovered that the material carrier of life is a crystal (which later turned out to be DNA), but also assumed, as Fisher did, that this structure is subjected to random variations.[39] While a periodic crystal repeats the same configuration, an aperiodic crystal does not, and for this reason Schrödinger guessed that the molecule responsible for heredity was an aperiodic crystal. In order to account for the variations, he introduced for the first time quantum fluctuations as a possible explanation.[40] As we have seen, Darwin's error was to think that there are small continuous variations and that these could be inherited (a blending theory of inheritance). Schrödinger understood that only significant and discontinuous mutations are inherited and that these are due to quantum jumps in the gene molecule: Mutation is a *single* and *local* event (in a ray of about ten atoms) that spreads to a change of basis. At a basic level, when there is a jump, a molecule is driven into another configuration [Subsec. 6.2.4] of the same basic atoms (the two molecules are called isomeric). A discrete amount, i.e. a quantum of energy, is needed in order to accomplish this transition, although such a discrete change will affect a whole basis and therefore a genetic sequence. Schrödinger also ascertained that mutations must be extremely rare, otherwise, since mutations are generally disadvantageous, they could rapidly drive organisms to extinction [Sec. 9.1]. This is connected with the issue of evolution timing, a problem I shall deal with below.

9.3.2 Between Populations of Genes and of Phenotypes

Another phenomenon was discovered by Sewall Wright[41]: In small and isolated populations a chance departure from the statistical distribution of traits becomes possible (it is a statistical fluctuation that is independent from natural selection). This is called *genetic drift*, which might be the basis of exaptation, a concept I shall introduce below. In this way, instead of striving toward a fitness maximum, nature tries out various solutions. It is a pluralist point of view [Subsec. 9.2.3]. The relevant parameter here is not the rate of mutation, or the rate at which selection eats up genetic variation (as is the case for Fisher), but it is rather the phenotypic population size and its migration rate. Therefore, we naturally pass here from the consideration of populations of genes to that of populations of phenotypes, modifying the meaning of population genetics.[42] Wright thought that pure statistical fluctuations could do a good part of the work, and in this sense is reminiscent of Peirce, who assumed that statistical arrays have ordering properties of their own.[43]

[38] [*SCHRÖDINGER* 1944]. [39] [LINDAHL 1993]. [40] See also [*MONOD* 1970, Ch. 6].
[41] [*WRIGHT* 1969, *WRIGHT* 1978]. [42] [LEWONTIN 1989]. [43] [PEIRCE 1891, pp. 288–93].

The origin of population thinking is to be found in the Russian school. According to Chetverikov, the master of Dobzhansky, a species soaks up heterozygous mutations like a sponge, while remaining externally (phenotypically) homogeneous or presenting at least some subvariants. For this reason, nature may harbor a good deal more potentially adaptive variations than Malthusian Darwinians were disposed to admit. Dobzhansky[44] extended and Darwinized this point of view, assuming that drift produces variations that are not important and are confined to small subpopulations until they spread to the whole population by natural selection; as we shall see, this is an interesting connection with Baldwin's approach and even with Waddington's epigenetics.[45] In other words, phenotypic population dynamics preserves and eventually scatters variations. For Dobzhansky, then, species became the real distributed entities on which classification is anchored. Dobzhansky treated the adaptive landscapes of Wright (a point on which I shall say more below) as actual biogeographic distributions, so that nonadaptive valleys become actual geographic barriers. Each species occupies a peak of these distributions and the valleys are empty. Taxonomy is therefore subordinated here to a populational and evolutionary conception of species.[46]

The modern synthesis (whose most authoritative exponent is E. Mayr) had inherited these ideas, and (to be schematic here) is based on the assumption that species are real entities, spatially and temporally bound populations held together by genetic links in a well-defined ecological niche (we have already considered some of the problems connected with this idea). Starting with Dobzhansky's contribution, Mayr[47] shifted the attention much more to phenotypes. For him, the most important point is interbreeding (reproductive isolation). The emphasis is therefore placed even more on effective spatial distribution. He pointed out that populations living peripherally differ most radically from other populations in the group, and called this phenomenon peripatric speciation. Such genetically unbalanced populations may shift to new niches. Then, isolation causes a whole reorganization of the genetic structure of a population. In other words, in contrast to Dobzhansky, the speciation may be complete before the new species reconnects with its parent population.

9.4 The Concept of Species

To define a concept of species is not an easy job (it is not by chance that Darwin explicitly avoided this[48]). The successors of Darwin have tried to provide such a definition. In particular, the classical definition of species, the so-called *biological species concept*, is essentially due to the work of T. Dobzhansky[49] and was made popular by E. Mayr.[50] The main challenge faced by Mayr is to give a definition that avoids two opposite problems:

- That of mixing species that are morphologically very similar but unrelated, and
- That of exaggerating the demarcations between species (if one takes into account the high variability that is intrinsic to populations).

[44][*DOBZHANSKY* 1970]. [45][*PIAGET* 1967, p. 122]. [46]See also [*LAMARCK* 1809, I, pp. 102–4].
[47][*MAYR* 1963]. [48][*DARWIN* 1859, p. 39].
[49][DOBZHANSKY 1937a, *DOBZHANSKY* 1937b]. I owe this remark to F. Ayala, whom I thank.
[50][*DOBZHANSKY* 1970, p. 357] [MAYR 2000]. I shall make use of this synthetic paper as one of the last statements of Mayr's approach. The book in which it has been published also contains other interesting papers that I shall quote in this section.

As mentioned, the concept of population is the true center of Mayr's definition. He defines the biological species as a group of interbreeding natural populations that are reproductively isolated from other such groups (it is a reproductively cohesive assemblage of populations).[51] This definition attempts to connect two very different issues: The genetic aspect (interbreeding is indeed based on genetic relationships) and the populations dynamics. The main concern of Mayr is to avoid transforming a species into a class that is separated from the individuals that constitute the population. However, there is a problem here (of which Mayr was perfectly aware): This concept only applies to organisms that reproduce sexually, that is, to most eukaryotes [Subsec. 7.5.2]. This cuts out the majority of living beings, represented by bacteria. Mayr's justification is that the issue of species is the protection of harmonious gene pools. Now according to him, the bacterial genotype does not require any protection because it is not threatened by destruction through outcrossing. I will not discuss here the implications that this statement has for microbiology, but I shall stress that it does make the species something more than a population able to interbreed.

These problems are the reason why other options have also been explored. One of the most interesting is represented by the definition given by Willi Hennig: All individuals connected through tokogenetic relationships (i.e. individuals bound through some descendance relation) constitute a (potential) reproductive community, and such a community should be called a species.[52] This definition potentially covers both uniparental and two-parental species. Indeed, there have been interesting further developments with regards to the concept of species as the smallest aggregation of (sexual) populations or (asexual) lineages diagnosable by a unique combination of character states.[53] By stressing tokogenetic or phylogenetic relations, a stronger continuity among different species is also established, and the only discontinuous events, from this point of view, are represented by species extinctions.

The above problem can also be seen from another more speculative perspective. Mayr added that the term interbreeding indicates a *propensity.*[54] The reason is simple: Spatial and temporal barriers between species (that determine their isolation) can in principle be overcome, and sometimes it happens so. Now, if previous isolated species are able to interbreed, then the definition above is weakened, since it only consists of a mere spatial or temporal separation, without involving any genetic aspect. On the contrary, if we would like to preserve some connection with the genetic aspect of the problem, then we will need something *more* than a mere spatial or temporal separation. It is here that the concept of propensity comes into play to fill this empty place. However, this concept does not seem especially clear in this context. Does it apply to populations or to individuals? If to individuals, it is at most an operative criterion for recognizing members of the same species and not a definition of species. If to populations, then these are no longer mere collections of individuals and to a certain extent become individuals themselves.[55] Instead, if we admit that the members of a species *share* information (DNA) [Subsec. 2.2.3] and therefore are subjected to common constraints, then we can understand that it can be a population of individuals also submitted to *common* selective pressure without assuming it to be a sort of individual or quasi-individual.[56]

[51]There are already several interesting considerations about species segregation developed in [*LLOYD MORGAN* 1891, pp. 99–110].

[52][*HENNIG* 1950]. See also [*LAMARCK* 1809, I, p. 54], and, for recent examination, [*MINELLI* 2009, pp. 10–12].

[53][WHEELER/PLATNICK 2000]. See also [MEIER/WILLMANN 2000].

[54][POPPER 1959, *POPPER* 1990].

[55]This seems the move made in [*ELDREDGE* 1985].

[56]For an examination of the problem see [HULL 1976].

9.5 Constraints on Evolution

In this section, I shall show that there are constraints during evolution that are not a consequence of natural selection. On the contrary, they often restrict the scope of the latter. These constraints are mainly of structural and functional types, the latter ones giving rise to evolutionary convergences. One of the first scholars to understand these problems was G. Mivart.[57] As a consequence of convergences, Mivart was also led to the view that evolution is not only capable of jumps and leaps, but also long periods of stability.[58] Another scientist to have pointed out the relevance of convergences for evolution was Lloyd Morgan.[59]

9.5.1 Punctuated Equilibria

Breeding experiments with domestic animals showed that there is a distinct limit beyond which further changes become impossible or at least increasingly difficult. A species, as stressed by Richard Owen,[60] comes abruptly into being and its latest forms are hardly distinguishable from its earliest forms. In other words, a species has a great *stability* over time. If everything is subject to random changes, then why are there characteristics that, once fixed, remain subsequently immune to any variation? A big problem for the gradualism of the original formulation of evolutionary theory is that organisms are wholes. Cuvier's principle of *teleonomic correlation* among the different parts of the organism [Subsec. 8.2.1], which helped to reconstructs fossils, took the idea that organisms are wholes as a sheer fact[61]: The assumption was indeed that structures such as those developed by living beings could not be the result of mere chance.

As a matter of fact, it is very difficult to understand how organisms could evolve by single specific mutations, given the interdependency between the different parts.[62] Also in transitional forms (like lungfishes or monotremes) or hybrids, there are characteristics, say for the lungfish, which are basically typical for fish and characteristics which are typically amphibian, but it is difficult to find characteristics which are between the two.[63] It is not always possible to form a sequence from fish to mammals: In fact, mammals' eggs are closer in their initial pattern to those of a frog than to any reptile's.

For these reasons, S. Gould abandoned gradualism[64]: According to him, emergences of species are geologically instantaneous events followed by long stasis periods, because speciation is related to a sudden reorganization of developmental programs. These events occur mostly in isolated populations—as pointed out by Dobzhansky and Mayr[65] [Sec. 9.4]. Later, the distinction between replicators and interactors [Sec. 7.1] was introduced by Hull and Wimsatt[66] in order to account for the feedback effects necessary for explaining this sudden reorganization. This post-neo–Darwinian expanded synthesis has a genealogical and an ecological part, as we shall see in details.

Fossils have confirmed the basic pattern of discontinuity: Species appear suddenly and are relatively stable phenomena. Random extinctions could not have eliminated almost all the intermediate forms. Eldredge showed that the origin of the new species (genetic incompatibility) is short

[57] See [*MIVART* 1871, pp. 35–75] for the issue of structural constraints and [*MIVART* 1871, pp. 76–110] for convergences, called *concordant variations* by Mivart.

[58] [*MIVART* 1871, pp. 111–41]. [59] [*LLOYD MORGAN* 1891, pp. 117–19]. [60] [*OWEN* 1866, v. 1, Preface].

[61] [*CUVIER* 1817, v. 1, pp. 5–6 and 9]. Actually, Cuvier stressed that finality is the *causal* criterion for dealing with the conditions of existence of an organism, i.e. those conditions that, put together in a coordinated way, explain not only the organism as such but also its correlation with the environment. In this context, I prefer to use the concept of teleonomy instead of finality.

[62] [*DENTON* 1985, pp. 93–141]. [63] An argument often used for rejecting evolution [*JOHNSON* 1991].

[64] [ELDREDGE/GOULD 1972] [GOULD/ELDREDGE 1993][*ELDREDGE* 1985]. [65] [*MAYR* 1988, pp. 319–20].

[66] [WIMSATT 1986] [HULL 1981b].

term and does not involve a great deal of morphological changes (which happen more often within a species until a certain stability is acquired). Speciation events are *propulsive* of evolutionary changes and they should not be considered as mere results of evolution; Eldredge pointed out that the latter conclusion depends on erroneously taking into account successful speciation only. Indeed, not all morphological innovations represent adaptations.[67] If a part of a larger population becomes isolated in a suboptimal habitat, natural selection will drive it into the direction of an increasing fit, where with phenotypic *fitness* I roughly understand the reproduction rate. Therefore, if a certain random mutation does help this adaptation, adaptive speciation will be the result. So, speciation as such is not necessarily related to adaptive change, whereas *successful* speciation is.[68] Summarizing, while local populations always become distinctive in some way, the event of speciation fixes their innovations and eventually gives them historical validity, so that, once it has happened, each species is subsequently free to accumulate more variations, at least until a stable state is reached [Subsec. 9.3.1].[69]

Generally, clades show a greater diversification at the beginning; then they reach an equilibrium. The reason is that an initial lack of competition with other individuals or proximate species (due to a low density of populations in that region of the fitness landscape) permits unusual opportunity for diversification, so that clades diversify rapidly. Successively, selection establishes a situation of equilibrium.[70] The *fitness landscape* is a multidimensional space assigning a fitness value to any genome, function, or organism (giving rise to genetic, functional, and phenotypic landscapes, respectively). The fitness of an organism is indeed the measurable capacity of the organism to carry out some definite function in a given environment, and the distribution of such a fitness measure over the space of possible organisms (or of proteins for a single cell) is, for instance, the fitness landscape with respect to that specific function.

The above results are also confirmed by analyzing mutation rates. Diversification rates in evolution are markedly less variable than extinction rates at cycles shorter than 25 million years. Then, there are intrinsic speed limits for diversification rates.[71] The very low nucleotide mutation rate limits natural selection to those alleles that currently differ from the current fixed allele by a single nucleotide. As a consequence, evolution, if only natural selection is considered, should proceed in a series of bursts.[72] A typical burst is shown to involve on average about 1.5 to 2.5 allelic substitutions. This elevates the variance to a level that is commonly observed.

Summing up, it is not true that anything at an evolutionary level can be explained by adaptation.[73] The constraints of the inherited form and developmental pathways may channel any change so that the channel can represent in itself the primary determinant of subsequent evolution, a fundamental teleonomic process [Subsecs. 8.2.1–8.2.2]. On the other hand, these arising structures may represent no adaptations to a previous environmental context, and may even appear as simple byproducts of evolutionary changes. In fact, the pool of nonadaptive mutations that are somehow preserved must be far greater than that of direct adaptations. In general, it can be assumed that every trait preserved across evolution is partially adaptive and partially an epiphenomenon created by a stronger selection occurring on other traits.[74] This very important explanation is known as *exaptation*[75]: Structures now indispensable for survival may have arisen for other reasons and been shifted to a different functional role, for example, the hominid's brain, which was probably already capable in itself of high abstraction, and so on, at least 1.5 million years ago, but was used for these

[67][*TATTERSALL* 1998, pp. 82–103]. [68][*ELDREDGE* 1995, pp. 93–123].
[69]In [AHLBERG/CLACK 2006] an impressive account of the evolutionary transformations from fish to tetrapods is reported. See also [AYALA/AVISE 2009].
[70][GOULD *et al.* 1987]. [71][KIRCHNER 2002]. [72][GILLESPIE 1984].
[73][GOULD 1982]. See also [*JOHNSON* 1991]. [74][*OSTER/WILSON* 1978]. [75][GOULD/VRBA 1982].

purposes much later.[76] This explanation would still be reinforced if we assumed that a significant part of mutations are neither adaptive nor noxious but *neutral*.[77]

9.5.2 Major Transitions

Another consequence of the above considerations is that there are major transitions in evolution (so-called macroevolution). In particular, according to Maynard Smith and Szathmáry, evolution is characterized by the following major transitions[78]: From replicating molecules to chromosomes, from prokaryotes to eukaryotes, from binary fission to sexual reproduction, from unicellular to multicellular organisms. Obviously, the most important one is the first, which I would rephrase as the zeroth transition, from macromolecules to organisms. I cannot claim here to give an answer to this huge problem. As I have often stressed, I am rather interested in the necessary conditions of life than in explaining its origin, and I have also emphasized some difficulties in accounting for this transition [Subsec. 7.3.1]. Nevertheless, some educated guesses about certain aspects of this process can be made. Life probably began with an autocatalytic molecule[79] [Sec. 6.6]. It is very likely that it was a chemoton, i.e. an autocatalytic cycle that already had an informational part.[80] Such a cyclic activity may have conferred a certain autonomy to this molecule. The initial autocatalytic molecules could have been of the family of amino acids (peptides, in order to distinguish them from actual proteins[81]), able to bridge from a protometabolism to a true metabolism. However, since proteins do not have an informational part [Subsec. 7.4.5], it is more likely that this molecule was an RNA molecule, which can also display programming activity.

It is likely that many "experiments" were spontaneously performed before some more or less optimal solutions in terms of biochemicals and their functionalities were found. I have already pointed out, for instance, that we can assume the existence of a larger variety of codifying molecules in the first steps of life than we can currently observe [Subsec. 7.4.1]. The progressive reduction of alternatives and the final selection of some of them is a stabilization process of a teleonomic kind [Subsec. 8.2.1] that is also fully in accordance with the requirements of information transmission [Subsec. 2.3.2]. It is also likely that there was a common ancestral community of primitive cells, so that there is probably no common root of the tree of evolution.[82] Moreover, as far we know, eukaryotic cells were developed by symbiogenesis[83] between protokaryotic cells and new aerobic cells (which became mitochondria in animals and chloroplasts in plants), even if many details are not at all clear and caution is still demanded.[84] Symbiosis is actually a variant of a larger and very important phenomenon called horizontal gene transfer, which comprehends all forms (through bacteria and viruses) of nongenealogical transfer of genetic material from one organism to the other.[85] This is an economical way of inducing high genetic variability, as I mentioned in Subsec. 7.5.2—note also that horizontal gene transfer makes Mayr's previously quoted statement, about bacteria not being threatened by genetic exchange, obsolete today, and therefore his definition of biological species becomes rather inadequate [Sec. 9.4]. In the eukaryotic cells compartments were established to protect DNA and RNA from the cytoskeleton as the cell moves. In so doing, the number of internal surfaces augmented relative to the external cell membrane [Subsec. 7.3.3], which is a way of allowing increasing information control and capability to act

[76] See also [JACOB 1977]. [77] [KIMURA 1968, *KIMURA* 1983].
[78] [*MAYNARD SMITH/SZATHMÁRY* 1995, *MAYNARD SMITH/SZATHMÁRY* 1999].
[79] [EIGEN 1971, EIGEN/SCHUSTER 1977]. [80] [*GÁNTI* 1987]. [81] [*DE DUVE* 2002, pp. 63–4].
[82] [FORD DOOLITTLE 2000].
[83] [*MARGULIS* 1970] [*MARGULIS/SAGAN* 1986]. See also [*MARGULIS/SAGAN* 2002] [MARTIN/RUSSELL 2002].
[84] [POOLE/PENNY 2007]. [85] [DOBRINDT *et al.* 2004] [GOLDENFELD/WOESE 2007].

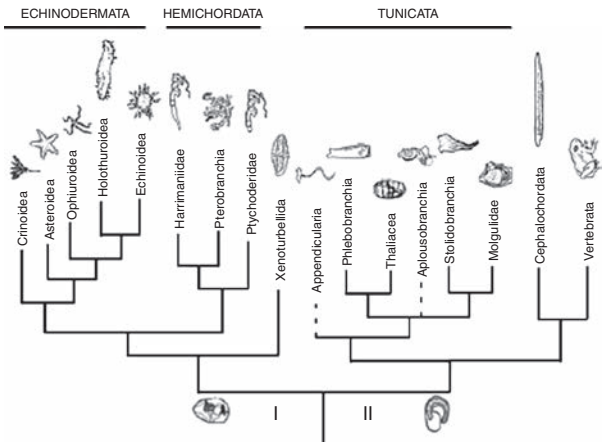

Fig. 9.3 The emergence of the different body plans. Ciliated *Ambulacraria* larvae (I) and *Tunicata* tadpole larvae (II) are likely to have separate origins. Adapted from http://www.nature.com/hdy/journal/v97/n3/figtab/6800872f1.html.

[Subsec. 8.3.2]. Therefore, this modularization process [Sec. 7.2] goes together with the capacity to establish new and higher levels of integration, according to Spencer's law [Subsec. 2.4.4]. In particular, I emphasize that the nucleus can be considered a self inside the self [Sec. 8.4].

As we shall see, the passage from unicellular to multicellular organisms often proceeded through association in colonies (green algae exist in unicellular, colonial, or multicellular forms), even if the mechanisms in place were different.[86] In multicellular organisms, cooperation and cohesion are indeed very important. Therefore, here the traditional cell membrane is replaced by more plastic plasma membranes. With the advent of the multicellular organism, we also have the spectacular emergence of different basic body plans and relatively self-organizing epigenetic processes (the so-called Cambrian explosion, about 530 million years ago) of almost all living forms (and much more, since many forms have since disappeared) [Fig. 9.3].[87] This means that, after the constitution of organisms, two crucial major steps in evolution (from prokaryotes to eukaryotes and from unicellular to multicellular organisms) were accomplished without genetic heredity or traditional natural-selection processes playing a role, but mainly through physical mechanisms of integration and requirements stemming from systemic organization. It is also possible that the passage from binary fission to sexual reproduction could be explained by a physical mechanism: Cells may fuse and give rise to a diploid organism [Subsec. 7.5.2]. If this occured in eukaryotes, which are characterized by the segregation of the DNA in the nucleus, the steps of meiosis and mitosis would be a quite natural consequence.

9.5.3 Evolutionary Convergences

In evolution there are both divergences (called, at an evolutionary level, homologies) and convergences (analogies). Darwin's theory was concerned with homologies. We have already seen

[86] [*BONNER* 2000].

[87] [SHERMAN 2007]. G. Saint-Hilaire was among the first scholars to understand that there are general body plans in animals [*SAINT-HILAIRE* 1830].

in the introduction to this section some precursors of the idea of evolutionary convergence. In this context, note also the more recent (post-neo-Darwinian) contribution of G. De Beer,[88] who pointed out that homologies cannot be the sole explanation of evolutionary processes. Still more recently, this has become the core of S. Conway Morris's research, who has also provided many examples of convergence at an evolutionary level. One of the most remarkable ones is represented by the fact that 15 million years ago plants, notably grass, passed everywhere on the Earth from so-called C_3 to the C_4 photosynthesis (the figures refer to the numbers of carbon atoms in the first compound to be formed) as a reaction to the decline of carbon dioxide's presence in the atmosphere.[89] Another good example is provided by the naked mole rat (*Heterocephalus glaber*), a mammal the existence of which had been predicted by R. D. Alexander by making an extrapolation from the eusociality of some insects like ants and bees.[90] Very light animals, like insects, need a stabilization system for flight to avoid being too exposed to fluctuations in the air. Both dipteran and strepsipteran insects have developed halteres, an equilibrium system acting as a vibrating gyroscope for detecting and compensating for angular acceleration. However, in dipteran insects it evolved from the hindwings, and in strepsipteran insects from the forewings. Convergence in eye structure is also remarkable. For instance, despite the absence of a nervous system, the visual apparatus of some dinoflagellates is convergent on the animal eye. Subsequently we see a similar occurrence in the olfactory and auditory systems.[91] Furthermore, mammal's viviparity can also be shown to be a convergent process.[92] However, the most striking example of convergence is represented by the social organization of animals.[93] Indeed, even activities somewhat similar to agriculture and military organization can be found in animals like ants.

Evolutionary convergence is due to the fact that, even if potentially infinite forms of life (of organisms, proteins, and so on) exist, the number of biological systems that can truly exist in a stable form is relatively small [Subsec. 7.4.4]. A good example is represented by the finite variation of all forms of the skeleton.[94] The reason is that all biological systems obey two important constraints[95] [Subsec. 8.2.4]:

(1) There are particular sites (hubs) or properties of biological systems fundamental for certain functions and therefore crucial for survival, and
(2) A given function depends on an architecture of the whole that is highly recurrent.

We see here again a special combination of local and global features [Subsecs. 2.2.5, 6.5.1, and 8.2.7]. This means that the huge range of theoretical possibilities has many places that remain forever empty. Evolutionary convergences are therefore deeply rooted in the concept of function. My guess is that, when there is some true evolutionary convergence, this is the hallmark of the emergence of a biological function.

As such, the general explanation of convergences could be the following: The molecular or organismic degeneracy of several structures (even in distant species) relative to any function [Subsec. 8.2.5] allows for a pool of preexisting structures to be plus or minus apt, by successive exaptation [Subsec. 9.5.1], for deploying a certain function. This assures the necessary variability at the start (random structural modifications are the source of variety here). Consequently, the fact that any function has a crucial structural hub for displaying itself [Subsec. 6.3.2], ensures that a subset of

[88] [*DE BEER* 1971] [BRIGANDT 2006]. [89] [*CONWAY MORRIS* 2003, pp. 109 and 293–4].
[90] For this and most of the following examples see [*CONWAY MORRIS* 2003, pp. 141–214].
[91] [*CONWAY MORRIS* 2003, pp. 190–4]. [92] [HEAP 1994]. [93] [*TINBERGEN* 1953].
[94] [THOMAS *et al.* 2000]. For another example see [ABZHANOV *et al.* 2006]. [95] [*CONWAY MORRIS* 2003, p. 10].

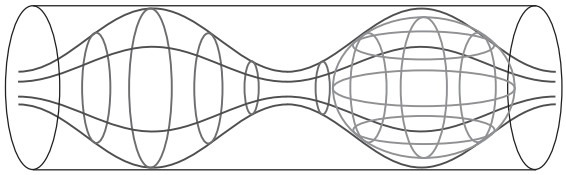

Fig. 9.4 A cylinder (in black) ideally contains an unduloid which ideally contains a spheroid (shown in the right bulb of the unduloid): The symmetry rotation axis is horizontal. The unduloid represents here the wave-like perturbation to which we submit the cylinder. When the wave amplitude (and therefore the distance between peaks and valleys) increases, so that it exceeds the ray of the cylinder, this becomes an unduloid. If the wave amplitude still increases, the unduloid breaks down into spheroids.

this pool will be selected. In other words, variability plus crucial structural elements will determine teleonomic convergence. When I introduced the concept of *emergence* [Subsecs. 2.4.2 and 8.2.2], I indeed emphasized that new functionalities arise when, from an initial variety, both a canalization process and some channelization come out.[96] This also raises the very difficult problem of whether there is any directionality in evolution, a philosophical issue that will occupy us much later.

It also remains true that many convergences may be explained by latent homologies due to unexpressed or partially expressed potentialities of regulatory networks[97] [Subsec. 9.3.1]. This is acknowledged in part by Conway Morris, who supports the principle of inherency, according to which the building blocks of complex structures are available long before they are recruited for more sophisticated tasks.[98] This is again a special case of exaptation (and teleonomy) and can be very useful for explaining the emergence of new functions.

9.5.4 Structural Constraints

Having shown that there are functional convergences, it is now interesting to observe that there are also physical and structural constraints [Subsec. 2.3.4] on the organism, an aspect that reduces the window of the effective initial random mutations and can partly explain those convergences or at least their initial appearance.[99]. D'Arcy Thompson rejected the explanation of the emergence of many forms being due to heredity only, because many structures present no selective advantage. For instance, cells—like drops—tend to a definite size in order to balance the surface tension,[100] and this is the reason why nerve cells are almost all the same size in all mammals. The surface tension shrinks the cell until it is counterbalanced by some other forces. *A priori*, there is no limitation of the possible forms that represent surfaces of a minimal area, but, if we limit ourselves to the surfaces of revolution, i.e. to shapes that are symmetrical about a given rotation axis, they are six in all: The plane, the spheroid, the cylinder, the catenoid, the unduloid, and the nodoid. If, by perturbating a cylinder, the wave amplitude exceeds the diameter, the cylinder turns to an unduloid and then breaks in spheroids [Fig. 9.4]. In a system of equal spheres in contact with each other on a common plane, any sphere is in touch with six others around it. If the system is subject to uniform pressure, then the contact points become lines and we obtain a system of hexagons

[96]Lack of consideration of this kind of dynamical process has led to the assumption that there is a high improbability of the emergence of life and new functions from which theoreticians of Intelligent Design have subsequently inferred the necessity of a Designer [*DEMBSKI* 1998].

[97][*WILKINS* 2002]. [98][*CONWAY MORRIS* 2003, pp. 5 and 166].

[99]Haeckel seems to be the first scholar to have taken seriously morphology as a relevant factor of organisms' adaptation [*HAECKEL* 1866]. See also [MATSUNO 1984].

[100][*THOMPSON* 1942, pp. 88–131].

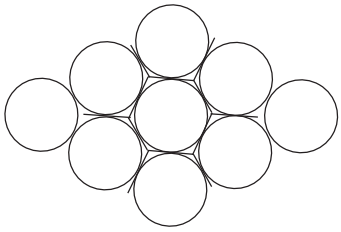

Fig. 9.5 How hexagonal cells arise on a surface. Inspired by [*THOMPSON* 1942, p. 103].

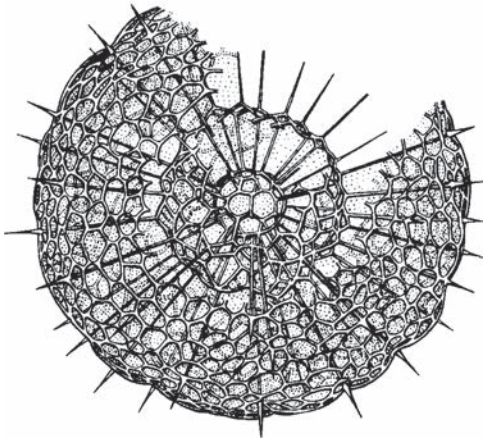

Fig. 9.6 Example of hexagonal repartition of a surface with sparse pentagons: The skeleton of the *Actinomma arcadophorum*. Adapted from [*THOMPSON* 1942, p. 157].

(as for honeycombs) [Fig. 9.5]. It is a geometrical fact that no system of hexagons can completely cover a 3D space, so that hexagonal skeletons also show some pentagons [Fig. 9.6].

Another recurrent form in living systems is the spiral.[101] All spirals may be classified as equable spirals (the spiral of Archimedes), where the point traveling from the center has a uniform velocity (here the ray r is equal to $a\theta$, where a is a constant and θ is the whole angle through which it has revolved); or equiangular or logarithmic spirals (typical of *Nautilus* [Fig. 9.7(a)]), where the speed of the point travelling from the center is increasing (and here we have $r = a^{\theta}$). The nautilus is also an example of a simple algorithm called Fibonacci's series, which is followed by many natural forms, like many flowers[102]: Here, each number n_i of the series is the sum of the two previous occurring ones ($n_i = n_{i-2} + n_{i-1}$). An interesting development of Thompson's ideas is represented by the fact that many biological patterns (somehow connected with metabolism, like heart size, life span, and so on) vary as quarter–powers of body mass.[103] This is connected with efficient blood transportation through the whole body.

Thompson finally showed how it is very easy to change the shape of an animal by a simple deformation (torsion, elongation, and so on) of its diagrammatic representation[104]: Due to the interdependence between the different parts, it suffices to exert a simple (geometrical) transformation [Fig. 9.8]. This means that a local mutation affecting some structural change, if not resulting

[101][*THOMPSON* 1942, pp. 172–201]. [102][KLAR 2002]. [103][WEST *et al.* 2001].
[104][*THOMPSON* 1942, pp. 266–325] [*ARTHUR* 1984, pp. 21–6].

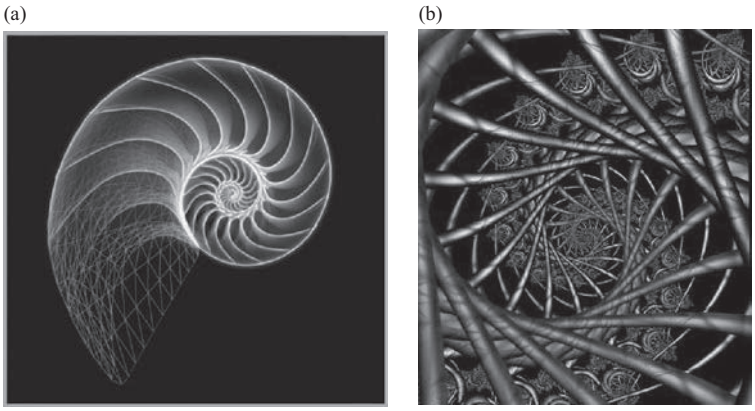

Fig. 9.7 (a) A splendid artistic reproduction of a *Nautilus* taken from the gallery in http://www.todman.dircon.co.uk/. (b) A fractal with the shape of a spiral, resembling the double DNA helix. The figure is taken from http://www.fractalus.com/paul/, a webpage of the artist Paul DeCelle. There are several pages showing an incredible number of beautiful fractal shapes. See also http://bugman123.com/Fractals/Fractals.html. (This figure is reproduced in color in the color plate section.)

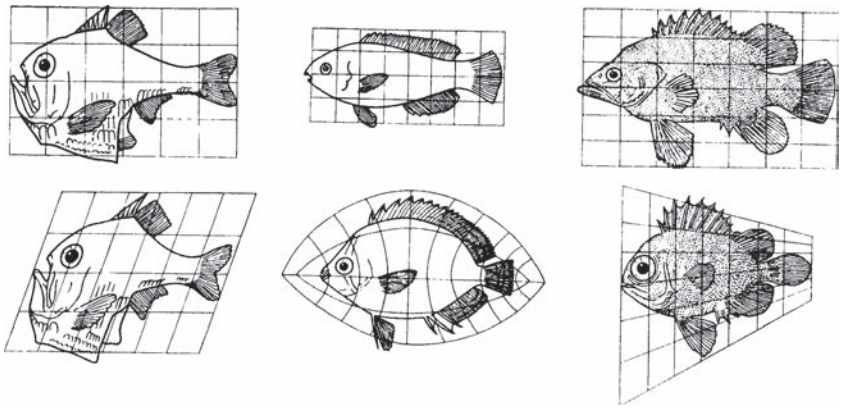

Fig. 9.8 Changes of fish morphology by simple deformation. All coordinates on the left are the usual Cartesian ones.

Left: A change to oblique coordinates (by 70°) allows the transformation from *Argyropelecus olfersi* (top) to *Sternoptyx diaphana* (bottom).

Middle: A change to coaxial circles allows the transformation from *Scarus* (top) to *Pomacanthus* (bottom).

Right: A change to triangular coordinates allows the transformation from *Polyprion* (top) to *Pseudopriacanthus altus* (bottom).

Adapted from [*THOMPSON* 1942, p. 299].

in a noxious or at least a neutral change, will induce a rearrangement of the epigenetic pathway, thus shifting to a new morphology.

Another interesting form of structural constraint in living beings that was not known at Thompson's time was the form interdependency at different scales [Subsec. 6.3.2], that is, the fact that living beings show structures or patterns [Subsec. 6.5.2] that are recurrent at different scales.[105] Today, this feature is very much studied in mathematical terms under the name of "fractals" (i.e. geometrical shapes whose dimensions are fractional), a type of structure that is very common among complex and chaotic systems.[106] It is sufficient to know that bacterial colonies[107] or blood-circulation systems can be modeled in this way. For instance, consider the spiral-forming fractal in Fig. 9.7(b), Fig. 6.15, or Fig. 6.22. Fractals (of which Fibonacci's series is a particular instance) are a consequence of the rhythmic, wave-like, nature of complex systems [Subsec. 6.3.1] and of small asymmetries in the application of initial forces or the initial conditions [Subsec. 8.2.7]. The fractal dimension of living beings represents a sort of fourth dimension of life able to explain the finding that scaling of most of the biological parameters follows a quarter power (all values found in organisms are multiples of 1/4).[108] For instance, diameter of tree trunks and aortas scale at $M^{3/8}$, where M is the body mass while blood circulation and heartbeat scale at $M^{-1/4}$.

9.5.5 Complexity and Evolution

Let us now consider the previous issues at a more general level. Alberch[109] was among the first scholars to understand that, during phylogeny, the possible stable configurations toward which an organism can evolve are limited and strictly connected with transformations during development. The main conclusions of Alberch were[110]: (1) In the phenotypic space, phenotypes cluster around major themes, (2) while there is dispersion around a theme, the variability in any trait is limited, (3) when new themes arise, the transitions between them are not random, (4) these properties are largely a result of epigenetic interactions during development. Summing up, the pool of stable systems at an evolutionary scale is relatively limited and the set of those that can be accessed by starting from a given organism are very few. It is precisely this fact that allows us to speak of teleonomic processes at an evolutionary scale [Subsecs. 9.5.1–9.5.4], and to introduce the concept of the next stable state which has the function of attractor [Fig. 3.24] without having recourse to finality [Subsec. 8.2.1]: The fact that *some* steady states (even if they are few) are accessible to certain initial evolutionary conditions [Subsec. 6.3.3] and *not* only one, although the subsequent dynamic process will lead to a single state, interdicts to speak of any information control on evolution or of teleological causation. This is also clear from the fact that at the evolutionary scale there are no goals [Subsec. 8.2.2] to be attained, otherwise species would become kinds of superorganisms [Sec. 9.4].[111]

S. Kauffmann has pointed out that natural selection works only by restricting and selecting among possibilities and entities that are already self-organized [Sec. 6.3].[112] Generally, order is conserved *despite* selection [Subsec. 9.3.1], because either selection awards random mutations which are counterbalanced by self-organization or the possible movements are trapped in a little phase

[105][*GERHART/KIRSCHNER* 1997, p. 147]. [106][MANDELBROT 1967, *MANDELBROT* 1977] [*BARNSLEY* 2006].
[107][BEN-JACOB *et al.* 2000]. [108][WEST *et al.* 1999]. [109][ALBERCH 1980].
[110][OSTER/ALBERCH 1982].
[111]This seems to me to be the main problem with the so-called Intelligent Design approach [*DEMBSKI* 1998, *DEMBSKI* 1999]. See also [AYALA 1998a, *AYALA* 2006] [*FUTUYMA* 1998, p. 342] [*ALEXANDER* 2008, Chs. 14–15]. On the concept of design see [*RUSE* 2003].
[112][*KAUFFMAN* 1993, pp. 9–25, 34–6].

space where a structure is preserved. This point was also stressed by D. Layzer: Random mutations either diminish the fitness or leave it unchanged.[113] As we have seen, structural analysis shows that a continuum of forms does not exist, and true neighboring morphologies in evolution reflect transformations into neighboring forms in the related families, generated by the underlying and relatively autonomous developmental mechanisms.

Moreover, the idea of the selection of a genetically predetermined program is unable to explain why the organism (differently from a computer program) can suffer substantial variation and yet still function[114] [Subsec. 8.2.4 and Sec. 8.4]. We may distinguish here between frozen dynamics, oscillatory dynamics, and chaotic dynamics. Structurally stable systems[115] can adapt on correlated (ordered) fitness landscapes while chaotic systems, which are not structurally stable, adapt on uncorrelated (disordered) landscapes. At the edge of chaos we have emergent global self-organizing, i.e. complex, structures like organisms [Subsec. 6.5.1]. We are interested in systems that spontaneously box them into small volumes of their state spaces (attractors) in absence of outside work. However, in contrast to prebiotic self-organizing systems [Subsecs. 6.5.3 and 8.2.1], this does not imply that organisms are insensitive to the environment and therefore fixed once and for all. Indeed, living systems show a trade-off between fixity (order) and freedom (disorder) [Subsec. 8.2.7], implying a character that is typical of them and only them: Evolvability and developmentability as a consequence of an itinerant dynamics. Mutations can then be understood as ways of walking through a fitness landscape.

Kauffman's model can be mathematically represented as a lattice determined by the positions of genes and their relationships (fitness contributions of the alleles). A *lattice* is a partially ordered set as shown in Fig. 9.9. As the organism's complexity grows, the number of conflicting constraints grows too, with the consequence that the adaptive walks in the fitness landscape become short and poor, and only local optima can be reached—a nonoptimal adaptation[116] [Subsec. 9.2.3]. The extreme case is a completely random fitness landscape with many peaks ($K = N - 1$, where N is the number of elements or parts and K represents the number of local interactions between parts). When we have such a rugged landscape that the different parts are completely uncorrelated, there will be too many local optima (in reality they are very flat) and we have the complexity catastrophe (for this reason, there cannot be a true atomistic evolution). At the opposite extreme, we have a completely global correlated fitness landscape with a single universal optimum (in this case we have $K = 0$). Here, there is no elasticity at all and the organism is subjected to a single fate: Either reach this peak or die. In intermediate cases, it results in the highest optima being nearest to one another, which means that the optima are not randomly distributed. They also have the biggest drainage basins, which explains Alberch's intuition. Convergent evolution (a common trait shared by different and independent species [Subsec. 9.5.3]) can be explained as an adaptation climbing from different initial points to either the same peak or nearby peaks in the fitness landscape.[117] Then, in order to evolve a system should keep the number K of local interactions lower than the number N of elements.

Summing up, as different mutations simultaneously occur, an organism can jump from one position in the landscape to another, which accounts, at a general level, for jumps in evolution [Subsec. 9.5.1]. If it jumps beyond the correlation lengths of the space, then, even if the landscape is smooth, the result would still be a fully uncorrelated random landscape (the time for finding a new peak doubles after each run). On the other hand, if the landscape is smooth, the rate

[113][LAYZER 1988, p. 35].

[114]See also [LAYZER 1978]. This raises the question of whether or not the genetic code is some kind of frozen accident [GOULD *et al.* 1987].

[115][*THOM* 1972, THOM 1980]. [116][*KAUFFMAN* 1993, pp. 36–120]. [117][POELWIJK *et al.* 2007].

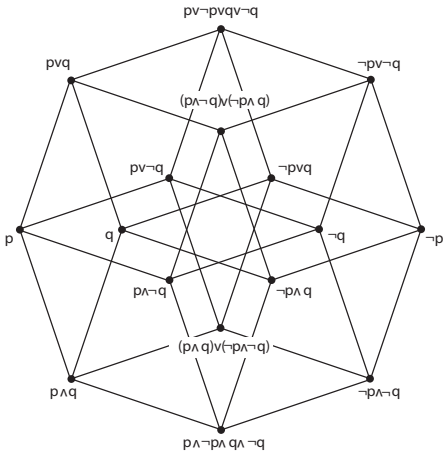

Fig. 9.9 A lattice that could represent two events, both occurring (p and q) and nonoccurring ($\neg p$ and $\neg q$) with all possible combinations (the couples p and $\neg p$ as well as q and $\neg q$ could also represent different alleles). We may think of these events as genetic mutations, the AND ($\wedge$) relations as meaning their joint occurrence while the OR ($\vee$) relation refers to the possibility that at least one of the disjoint events occurs. Note that any AND point is the starting of a divergence while any OR point is a convergence. The lowest element is the impossible event (meaning a contradiction since the two events both occur and do not occur). The highest element is the necessary event (at least the occurrence or nonoccurrence of one of the two events must happen). We may also interpret the lattice as a representation of a fitness landscape, in which case the highest element would be the peak in fitness, while the lowest one would be the point in which conflictual constraints make the survival of an organism impossible.

of finding successively fitter nearby variants decreases (smaller variations). At the beginning of evolution long jumps have a better chance of finding fit peaks. In the midterm it is better to find nearby hills. And on the longer time scale, the process, before it can proceed, must wait for either a successful jump to a better hillside some distance away or a landscape deformation. *Von Baer's law* states that development progresses from the general to the special, so that divergences between species manifest themselves progressively and therefore occur more in late embryos than in early embryos.[118] This can be explained by the fact that today early mutants are adapting on a highly uncorrelated landscape while late mutants are adapting on a well-correlated landscape, i.e. the rate of finding better mutants decreases very rapidly for fundamental alterations (there are branching pathways). Therefore, at the beginning of the era of multicellular organisms (the so-called Cambrian explosion), poor organisms could rapidly explore the whole space (resulting in the creation of 100 phyla; today there are only 30 surviving) [Subsec. 9.5.2].[119] Later on, there was only a place for lower mutations inside a given, established phylum (families and so on). In my opinion, this should be explained at the most general level: namely, early developments are more fundamental but less fine-tuned than later developments. This is likely to be a general law of evolving and developing systems that can be applied to biology, culture, science, art, and so on.

[118][*GOULD* 1977]. See also [*SPENCER* 1860–2, p. 301]. This does not mean that early embryos are necessarily very similar or even identical, as caustically shown in [*WELLS* 2000, pp. 81–109]. Nevertheless, as we shall see, genes are and are called homologue genes.

[119]See also [LEVINTON 1992] [*VALENTINE* 2004].

I finally stress that any landscape is also dependent on the fitness of other species. In the coevolutionary process the adaptive landscape of one actor deforms that of another actor, even if any actor acts locally on its fitness structure[120]: It is a teleonomic process of coadaptation [Subsec. 8.2.1]. It is again an application of the rule, typical for complex systems, that local actions can affect distant parts of the system [Secs. 6.3 and 6.5]. The mechanism here is the building of environmental niches that can host several species, as we shall see in the next chapter.

9.6 Game Theory

Another very important issue is that of "egoism versus altruism" in evolution. The traditional neo-Darwinian paradigm heavily stressed egoism in the struggle for survival.[121] However, already in the 1960s new explanations began to substantially modify this framework. I recall here Hamilton's work,[122] centered on a genetic mathematical model which allows for interactions between relatives promoting one another's fitness—I remind the reader here that I understand *fitness* roughly as the rate of reproduction: It is therefore a concept related to phenotypes. Making use of Wright's coefficient of relationship as the measure of the proportion of replicant genes in a relative, he found a quantity which incorporates the maximizing property of Darwinian fitness. This quantity was called *inclusive fitness*, i.e. the sum of direct (individual's) and indirect (on somehow related partners') fitness.[123] Species following this model should tend to evolve a particular behavior such that each organism appears to be attempting to maximize its inclusive fitness. This implies a limited restraint on selfish competitive behavior and possibility of limited self-sacrifices. Among the most known examples of inclusive fitness is that of social insects. This behavior, however, could again be explained in pure egoistic terms: For preserving *my own* genetic pool, in the same situations it would be better to let a strictly relation survive, like a brother or a sister.[124]

In order to explain this behavior, it is useful to refer to the mathematical game theory. Game theory was applied to evolutionary theory by Maynard Smith and Price.[125] In game theory, when two players can choose between a strategy A and a strategy B, we have the payoff matrix shown in Tab. 9.1. In other words, strategy A has payoff a when playing against A and payoff b when playing against B, while B gets payoff c when playing against A and payoff d when playing against B.

An interesting situation is when we have a Nash equilibrium[126]: We say that A has a Nash equilibrium when $a \geq c$, and that B has a Nash equilibrium when $d \geq b$ (we have a strict Nash equilibrium when we have $>$ instead of $\geq$). It is easy to see that the left part of Tab. 9.2 represents a Nash equilibrium for A, while the right part does so for B.

Another interesting game, proposed by Maynard Smith,[127] is represented by the Hawk–Dove (H-D) game: The rule is that H escalates fights, while D retreats. The benefit is b and the cost of injury is c. When two hawks meet, the payoff for each of them is $(b - c)/2$: One wins and the other is injured. If both hawks are equally strong the two probabilities are equal to $1/2$. If a hawk meets a dove, the payoff for the hawk is b, for the dove (which retreats) is 0. If two doves meet, the expected payoff is $b/2$. These results are summarized in Tab. 9.3. Here there is no Nash equilibrium

[120][*KAUFFMAN* 1993, pp. 237–79].

[121]Already Thomas H. Huxley employed the metaphor of the gladiatorial fight to depict the struggle for existence: Quoted in [*KROPOTKIN* 1902, p. 4].

[122][HAMILTON 1964a, HAMILTON 1964b]. [123][*DUGATKIN* 1997]. [124][*WILSON* 1975].

[125][MAYNARD SMITH/PRICE 1973]. See also [*MAYNARD SMITH* 1982]. Throughout this whole section I have taken into account the excellent Chs. 4–9 of [*NOWAK* 2006].

[126][NASH 1951]. [127][*MAYNARD SMITH* 1982, pp. 11–20].

Table 9.1 General payoff matrix in game theory.

	A	B
A	a	b
B	c	d

Table 9.2 Examples of payoff matrices when A is a Nash equilibrium (left) and when B is a Nash equilibrium (right).

	A	B		A	B
A	6	1	A	4	0
B	3	2	B	5	2

Table 9.3 Payoff matrix for Hawk–Dove game.

	H	D
H	$2^{-1}(b-c)$	b
D	0	$2^{-1}b$

for the Dove: If anybody else plays the Hawk, it is better to play the Dove, but if somebody else plays the Dove, it is better to play the Hawk.

One of the most famous games is the so-called Prisoner's Dilemma. The game is played by two persons suspected of a crime. If one confesses to the crime and the other remains silent, the latter will get 10 years in jail and the former will go free. If both confess the crime, both are condemned to 7 years. If both remain silent, both receive 1 year. Let us summarize the situation as in Tab. 9.4, where the two strategies in general terms are called cooperative (C) when one prisoner supports the other one, and defective (D) when one does not support the other. I might also mention that an interesting variant was proposed by A. Rapoport, and is called the Tit–for–Tat strategy. The

Table 9.4 Payoff matrix for the Prisoner's Dilemma. The numbers represent the possible losses in terms of jail years.

	C	D
C	-1	-10
D	0	-7

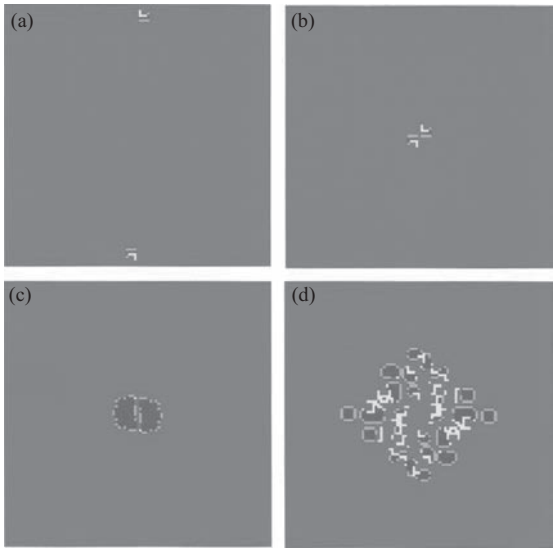

Fig. 9.10 Simulation of a collision of two walkers, giving rise, from the early stage (a), to an explosion of cooperative behavior (d) in a space occupied by defectors (red color). A walker is a cluster of 10 cooperators (blue color with yellow contour) with a form of a leg or pipe. Adapted from [*NOWAK* 2006, p. 160]. (The figure is reproduced in color in the color plate section.)

rule is the following: One starts cooperatively and always does what his opponent does. In many situations this has been shown to be the most profitable strategy.

The interesting point in the Prisoner's Dilemma is the following: No matter what one player does, it is better for the other one to defect. Indeed, if the first player cooperates, the other one can obtain 1 year if he also cooperates but no prison at all if he defects. When the first player wishes to defect, the other one obtains 10 years if he cooperates and 7 years if he defects. For this reason, it does not matter what the partner chooses, it is rational for the other one to choose the defective strategy. Obviously, the fact remains that, if *both* had chosen a cooperative behavior, both would have got only 1 year. But how can I know that my opponent will also cooperate if I do? The problem could be solved if we have been selected or educated to do that. Therefore, the game shows that, when we speak of a rational strategy, we are rather speaking of a rational strategy *from an egoistic point of view*, or from the point of view of the *individual* interest. This is not always the rational strategy *from a group point of view* (which can also turn out in many cases to be the best strategy also from an individual point of view).

The problem here is that there are *different levels* of selection and the traditional neo-Darwinian approach can account at the very most for selection and fitness of individuals (within-group selection). However, altruistic behavior can reinforce a group much more than purely selfish behavior,[128] and in this way the "gene for altruism" can evolve if the benefit for the group is much bigger than the individual cost (I stress that the relevant concept here is *group* and not species as such [Sec. 9.4]). In other words, cooperative behavior is intrinsically relational and global [Subsec. 2.2.5 and Sec. 6.3] and cannot strictly be reduced to the measure of individual

[128] [WILSON 1983] [*SOBER/WILSON* 1998].

interest. This does not mean that it is not efficacious. This can be seen very nicely through cellular automata models (so-called spatial games), an example of which is shown in Fig. 9.10: A small group of cooperative people in a large society of defectors will expand; we can think about the fact that cooperative people can easily deal with the Prisoner's Dilemma, cooperating in the confidence that the other partner will also cooperate (this is precisely what makes the struggle against small destabilizing groups in our society so difficult). A bridge to such winning altruistic strategies could be represented by the so-called generous Tit–for–Tat[129]: It is just like Tit–for–Tat but occasionally one responds to defection with cooperation.[130] If the effect is positive, then a cooperative behavior could become *evolutionarily stable*. Another bridge to altruism could be represented by mechanisms of indirect reciprocity.[131] Obviously, these remarks are much more valid when certain activities (like hunting) may require a mutualist behavior for being successful.[132] We shall see in the next chapters how these considerations for understanding social behavior are relevant.

9.7 Variation and Heredity

The connections between heredity, ontogeny, and epigeny are still not completely understood. While we have provided a general framework for this, it is not completely clear what the specific mechanisms at play are. The genome of higher organisms contains a large excess of DNA: No more than a small percentage of the mammalian genome can be involved in regulating or encoding RNA molecules or essential proteins. Moreover, the amount of DNA in the haploid genome of an organism has no systematic relationship to the complexity of the organism [Fig. 7.18]. In fact, some amphibian and plant cells contain 30 times more DNA than human cells. Furthermore, the genomes of different species of amphibians can vary 100-fold in their DNA content.[133] Finally, many related sequences of DNA are structurally and functionally equivalent, a fact that is still not completely explained.[134]

9.7.1 Mobile Information

Every normal protein-making gene has a complementary DNA sequence that sits on the other side of the ladder and is not usually transcribed into RNA. This can be taken to be a backup copy, because the cell can use it to repair damage to the gene. In some cases, however, while the DNA strand is producing an RNA message, its alter ego can produce an "antisense" RNA that has a complementary sequence. Whenever matching sense and antisense RNAs meet, they mesh to form their own double-stranded ladders, effectively interfering with the gene's ability to express a protein. These competing RNAs may suppress a gene just by tying up the gene's messenger RNA[135]: When double-stranded RNA appears in a cell, enzymes dice it up, peel the two strands apart, and use one RNA fragment to seek out and destroy any other RNA messages that stick to its sequence. The system protects cells against viruses, which often deliver their payloads in the form of double-stranded RNA. However, G. Rotman *et al*[136] showed that an RNA-interference machinery also has general regulative effects on genome expression.

[129][*PLOTKIN* 1997, pp. 116–18].
[130][*DUGATKIN* 1997, pp. 29–30]. [131][NOWAK/SIGMUND 2005].
[132][*DUGATKIN* 1997, pp. 31–34]. See also [TRIVERS 1971]. As we shall see, this is a point also raised in [*KROPOTKIN* 1902].
[133][KNIGHT 2002]. [134][SZOSTAK 2003]. [135][GIBBS 2003].
[136][YELIN *et al.* 2003] showed that it is RNA interference which also provides a handy way for scientists to shut off any gene at will (the so-called RNA-interference technique).

Other sequences appearing to be selfish are transposable DNA segments, some originally coming from bacteria (transposons).[137] They seem to make up at least 10% of higher eukaryotic genomes. Although most of these elements move only very rarely, there are many elements whose movements have a major effect on the variability of a species: More than half of the spontaneous mutations examined in *Drosophila* are due to the insertion or removal of a transposable element in or near the mutant gene—I stress that such a behavior would be impossible for proteins (since it would lead to sudden destruction or ineffectiveness), confirming that they do not represent codified information [Subsec. 7.4.5]. Thus, they cause a variety of short additions and deletions of nucleotide sequences.[138] A DNA sequence rearrangement caused by a transposable element is often observed altering the timing, level, or spatial pattern of expression of a nearby gene without necessarily affecting the sequence of the protein or RNA molecule that the gene encodes.

Transposable or mobile elements show a tendency to undergo long quiescent periods, during which they remain fixed in their chromosomal position, followed by a period of intense movement. Such cataclysmic changes in genomes, called *transposition bursts*, can involve near simultaneous transpositions of several types of transposable elements. When they occur, they induce *multiple changes* in the genome of an individual progeny. This increases the probability that two new traits that are useful together but of no selective value separately will appear in a single individual (a type of exaptation [Subsec. 9.5.1]). In several types of plants there is evidence that transposition bursts can be activated by a severe environmental stress, generating a variety of randomly modified progeny organisms, some of which may be better suited than the parent for surviving in the new conditions. Generally speaking, genes react to external shocks (x-rays, heat) in an unpredictable way.[139] In particular, the cells are sensitive to broken ends of DNA and try to unite these ends to one another, generating new mutations. We shall see the importance of these considerations for the Waddington effect and for epigeny in general. It seems that here a mechanism has evolved in order to activate transposable elements that serve as mutagens that produce an enhanced range of variant organisms when this variation is most needed. Recently, it was supposed that these mobile elements can also account for individual diversity.[140]

McClintock[141] was the first scholar to discover that genes may jump or move from one chromosome to another, implying that the genome can be understood as a turbulent superstructure showing the typical aspects of complexity [Secs. 6.3–6.5 and Subsec. 7.4.1]. This may cause rapid changes in developmental patterns and therefore account for rapid evolutionary changes in regulatory programs, speciation in small populations, and punctuated (discontinuous) patterns [Subsec. 9.5.1].

9.7.2 Genetic Recombination and Other Effects

Only about one nucleotide pair in a thousand is randomly changed every 200,000 years [Subsecs. 7.4.2–7.4.3 and 9.5.1]. Even so, in a population of 10,000 individuals, every possible nucleotide substitution will have been tried out in the course of a million years.[142] However, while point-like mutation is an efficient mechanism for fine-tuning the genome, evolutionary progress in the long term must depend on more radical types of genetic change.

A particular case of transposition is called genetic recombination and causes the genome to undergo major rearrangements with surprising frequency: The genome can expand or contract

[137][*ALBERTS et al.* 1983, pp. 305–26]. [138][*JABLONKA/LAMB* 2005, pp. 68–70].
[139][MCCLINTOCK 1984]. [140][MUOTRI *et al.* 2005]. [141][MCCLINTOCK 1956].
[142][*ALBERTS et al.* 1983, pp. 305–26].

by duplication or deletion, and its parts can be transposed from one region to another in order to create new combinations.[143] Component parts of genes—their individual exons and regulatory elements—can be shuffled as separate modules to create proteins that have entirely new roles. By these means, the genome can evolve *as a whole* so that it becomes increasingly complex and sophisticated even if the expression of each gene is controlled according to its own specific rules and information codification is not impaired. Higher eukaryotes contain an efficient enzymatic system that joins the two ends of broken DNA molecules together, so that duplications (as well as inversions, deletions, and translocations of DNA segments) can also arise as a consequence of the erratic rejoining of fragments of chromosomes that have somehow become broken in more than one place. DNA duplication followed by sequential unequal crossing-over underlies DNA amplification. Exons separated by introns greatly facilitate the evolution of new proteins.[144] The duplications necessary to form a single gene coding for a protein with repeating domains, for example, can occur by breaking and rejoining the DNA anywhere in the long introns on either side of an exon encoding a useful protein domain. We shall see below the fundamental role played by introns.

The eukaryotic genome contains not only introns but also other copies of seemingly nonessential DNA sequences that do not code for proteins.[145] Repetitive DNA (in humans it represents more than 50% of the genome) also plays an important role in computing and controlling tasks.[146] It influences chromatin structure. Sometimes this DNA is arranged in a dispersed form. However, it can also be organized in tandem. In this case, it can affect expression of genetic loci also at distances of many kilobase pairs. About one-third are the tandemly repeated satellite DNAs. Satellite DNA sequences generally are not transcribed and are located most often in the heterochromatin associated with the centromeric regions of chromosomes [Sec. 7.4]. In some mammals a single type of satellite DNA sequence constitutes 10% or more of the DNA and may even occupy a whole chromosome arm, so that the cell contains millions of copies of the basic repeated sequence. Satellite DNA sequences seem to have changed unusually rapidly and even to have shifted from their positions on chromosomes in the course of evolution. No function has yet been found for satellite DNA. It has been proposed that it can be an extreme form of selfish DNA sequences, whose properties ensure their own retention in the genome.

9.7.3 Information-Processing and -Controlling Requires Noncoding Sequences

We have seen that, during the eukaryotic transfer of information from DNA to RNA, noncodifying introns are removed from pre-mRNA through splicing [Subsec. 7.4.3]. Introns have probably invaded eukaryotes relatively late in evolution.[147] and can function as transposable elements (as we have seen): They are both the frozen remnants of history and the sites of future evolution.[148] Nuclear introns derived from self-splicing group II introns, which then evolved in conjunction with the spliceosome. This was possible once there was the separation between transcription and translation [Sec. 7.4]. However, if introns did colonize eukaryotes after their divergence from prokaryotes, then the question at hand is what their general significance for eukaryotic biology would be.

Let us make some general considerations at a pure computational level. Multitasking is employed in every computer in which control codes of n bits set the central processing circuit

[143] [*GERHART/KIRSCHNER* 1997, pp. 218–27]. Chromosomal rearrangement can turn out to be very helpful for explaining speciation events [AYALA/COLUZZI 2005].

[144] [GILBERT 1978] [HOLLAND/BLAKE 1987]. [145] [MATTICK 1994]. [146] [SHAPIRO/VON STERNBERG 2005].

[147] [MATTICK 1994]. But it has also been supposed that primitive ancestors of prokaryotes already had introns [HOLLAND/BLAKE 1987].

[148] [GILBERT 1978]

for processing one of $2n$ different operations. Sequences of specific and even modular control codes (programs) can be internally stored in memory, creating a self-contained programmed response. With the distinction between central processor and program (a problem of hierarchy in the system), reprogramming required the loading of new control codes into the memory and appropriate communication between nodes for synchronizing and integrating network activity.[149] In theory, gene networks could exploit a similar strategy, using internal controls to multitask components and subnetworks for the generation of a wide range of programmed responses, such as in differentiation and development. Existing genetic circuit models, although sophisticated, often ignore endogenously controlled multitasking and consider each molecular subnetwork (involving a few genes, for instance) to be sparsely interconnected and either be on or off, so that only one dynamical output is expressed. In contrast, multitasking allows for a wide range of programmed responses to be obtained from limited numbers of subnetworks (and genetic coding information). The imbalance between the exponential benefit of controlled multitasking and the small linear cost of control molecules makes it likely that evolution will have explored this option. Indeed, this may be the only feasible way to lift the constraints on the complexity and sophistication of genetic programming.

Prokaryotes have limited genome sizes (upper limit ca. 10 Mb) and low phenotypic complexity, suggesting that advanced integrated control strategies are not widely employed in these organisms. Potential cellular control molecules enabling multitasking and system integration must be capable of specifically targeted interactions with other molecules, must be plentiful (as limited numbers impair connectivity and adaptation in real and evolutionary time), and must carry information about the dynamical state of cellular gene expression. These goals are most simply achieved by spatially and temporally synchronizing control molecule production with gene expression. Most protein-coding genes of higher eukaryotes are mosaics containing one or more intervening sequences (introns) of generally high sequence complexity. Introns, therefore, fulfill the above conditions for system connectivity and multitasking:

- The potential for specifically targeted interactions as a function of their sequence complexity. Sequences of just 20–30 nt should generally have sufficient specificity for homology-dependent or structure-specific interactions;
- Multiple output in parallel with gene expression;
- Large numbers, especially if, as is likely, they are further processed into smaller molecules after excision from the primary transcript.

Introns are therefore excellent candidates for, and perhaps the only source of, possible control molecules for multitasking eukaryotic molecular networks, as genetic output can be multiplexed and target specificity can be efficiently encoded, which also allows for interesting original combinations that give rise to new proteins. This is especially shown by non-coding RNA, which is crucially involved in regulating functions.[150]

Summing up, the three critical steps in the evolution of this system could have been: (1) The entry of introns into protein-coding genes in the eukaryotic lineage, (2) the subsequent relaxation of internal sequence constraints through the evolution of the spliceosome and the exploration of new sequence space, and (3) the co-evolution of processing and receiver mechanisms for transacting RNAs, which are not yet well characterized but are likely to involve the dynamic modeling and remodeling of a chromatin and DNA coding sequences, as well as RNA–RNA and RNA–protein interactions in other parts of the cell.

[149] [MATTICK/GAGEN 2001]. [150] [AMARAL/MATTICK 2008].

9.8 Systems of Inheritance?

The central dogma of neo-Darwinian molecular biology[151] is a very important specification of Weismann's barrier[152] [Subsec. 9.2.2]. According to this dogma, information only flows from genes to proteins (and phenotypes). Such an assumption was never proved in its general form, and, for this reason, it is legitimate to ask whether or not there could eventually be some exceptions. As we have seen [Subsecs. 7.4.3–7.4.5], there are two different pieces that are held together in the whole transcription–translation process: A pure informational process (mainly the transcription) and a process (the translation) culminating in the construction of a three-dimensional protein having a specific function [Secs. 8.1–8.2 and 9.1]. The process as a whole can be summarized in the following form:

- From codified structural information to a messenger of instructions, that is, of instructional information (transcription);
- From this instructional information to a linear structure (the primary structure of the protein) mapped to the codified information through a set of instructions, but no longer being codified information itself;
- From this linear structure to a three-dimensional shape, the tertiary structure, which neither codifies information, nor is mapped to it; it is the structure performing a given function.

Therefore, I would like to ask three questions:

(1) Can the Weismann barrier be violated? In the 1960s retroviruses were discovered: They possess an RNA genome [Subsec. 9.7.1] and replicate themselves via a DNA intermediate so that RNA reversely transcribes DNA through the enzyme reverse transcriptase. It is well known that Crick[153] congratulated himself in not claiming that reverse transcription was impossible: Indeed, such an operation is still allowed by the central dogma. The fact is that the genome of an organism is not totally modularized but is part of a genetic *system*—comprehending mRNA, tRNA, and rRNA, with all related operations [Sec. 7.4]—which is in turn part of the organism's general developmental-physiological adaptation to the environment[154] [Ch. 8]. This also means that there is gene–gene interaction [Subsec. 9.7.2]. Moreover, a gene can give rise to a characteristics as a result of induction by environmental stimulus (which in itself does not carry any instruction about the character that will arise) and this product stimulates further activity of the gene even when the external stimulus has disappeared.[155] This means that Weissmann's barrier can indeed be violated.

There are many examples of possible exceptions to the Weismann barrier. Changes in gene activation might be transmitted in sexual reproduction[156]: For instance, the members of a clone of *Daphnia* develop spines when there are predators in their environment and these changes are then transmitted. Another example is represented by flax (*Linum*): Its morphology changes when it is treated with high levels of fertilizer, and the new morphology is then inherited when the treatment is finished. Recently, there has been a successful transformation of human skin cells into cells that are virtually indistinguishable from human embryonic stem cells.[157] More recently, experiments have been done that do not use viral vectors but plasmids.[158] Plasmids, which are extragenomic DNA material, are very common among bacteria, allowing genetic exchange without sexual combination horizontally, i.e. in the same generation [Subsec. 9.5.2],

[151][CRICK 1970]. [152][*WEISMANN* 1889, *WEISMANN* 1893]. [153][CRICK 1982]. [154][GOTTLIEB 2001].
[155][JABLONKA/LAMB 1998]. [156][*MAYNARD SMITH* 1998, pp. 8–12].
[157][DO/SCHÖLER 2004] [BAKER 2007]. [158][OKITA *et al.* 2008].

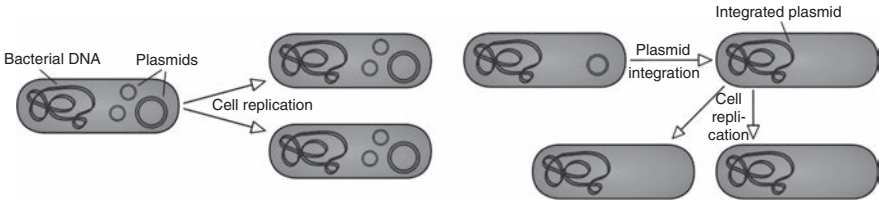

Fig. 9.11 On the left is ordinary binary fission. On the right, integration of a plasmid in the bacterial heritage. Adapted from www.commons.wikipdia.

as a way to adapt to a particularly difficult environment (for instance, where some bactericide is used). A very rough scheme of plasmid manipulation is shown in Fig. 9.11. It is an interesting case concerning a general law, about which I shall say more below: More primitive organisms anticipate more sophisticated solutions when dealing with those problems relative to which those solutions are appropriate answers.

These processes, however, can easily be dealt with without any violation of dogma if the genetic activity is inserted in an epigenetic and organismic way, for coping with the environment. This means that the central cornerstone of Darwinism—that environment is not instructive—*does not imply* that there is no interaction, at least indirectly, between phenotype and genotype. As we have seen, R. Dawkins[159] introduced the idea that organisms are only vehicles of genes [Sec. 7.1]. The terminology is not correct,[160] because it does not consider the role of interaction in selective processes.

(2) Can the central dogma as such be violated? Remember that the codified structural information in DNA is only potential and DNA is an inert chemical that in eukaryotes is activated only when the separation between codifying and not codifying sequences occurs (splicing) [Sec. 7.4]. Up to this point the DNA works at most as a pure information processor in the way I have explained in Subsec. 9.7.2. The final segment of the process is the building of a unit having functional value (the protein) with the RNA bridging between information and function. The whole process above is irreversible. If it weren't, then we should have at least one case in which it was also invertible (at least for statistical reasons). Proteins can allow the expression of a gene or block its expression; they can have a role in translating (other) proteins, but cannot back-translate *themselves*, recovering the information that started the process of their own building and starting a subsequent reverse-transcription into DNA [Subsec. 8.2.1]: This would be an instance of information control on their own building process, and the reason for its nonoccurrence is the impossibility of back-transforming a functional unit (the protein) into the codified information that gave rise to this unit, since this information has been lost in the process through which the functionality has arisen [Subsecs. 7.4.5–7.4.6 and 8.2.4]. It is known that the sequence of amino acids in a protein can be reconstructed with a method known as automated Edman degradation.[161] It is important to understand that this is not easy work and that in many cases we can succeed in such an extrapolation only by using a recombinant DNA technology, in which long stretches of DNA are cloned and sequenced in order to reveal the amino acid sequence of the protein encoded by the gene (it is a complementary experimental approach). In other words, we can recover the amino acid sequence by using *additional* codified information, which shows here that information control needs some information from the outside. This is also the reason

[159][*DAWKINS* 1976]. [160][*HULL* 1988b, p. 413]. [161][*BERG et al.* 2006, pp. 78–84].

why, for information control to be performed, some information codification is necessary *at the start* and this control cannot rely solely on functions [Sec. 8.1 and Subsec. 8.2.2]. Information control requires the possibility of different encodings and of instructions to be used, while any function is preassigned to the protein. This also means that information transfers from protein to protein, from protein to RNA, and from protein to DNA are all impossible processes. The conclusion is that the central dogma cannot be violated.

(3) Given the violation of Weismann's barrier but not of the central dogma, how many systems of inheritance are there? At a general level, there are several ways to transmit information: Through the genetic system, epigenetic processes, resonant behavior (as in the case of bird song or English tits opening milk bottles[162]), social transmission, and the symbolic system.[163] Obviously, we must sharply distinguish between (a) different systems of inheritance and (b) the issue of whether or not there are instructions from the environment. If the Lamarckian dimension proposed by Jablonka and Lamb is understood in the first sense, it is fully compatible with a generalized Darwinism in which there several different selection mechanisms [Subsecs. 2.2.6 and 9.2.1]. If in the second sense, it is no longer compatible. I endorse the first option. Given this enlarged understanding, the range of phenomena that can be given selective explanation should also be expanded to include cultural tradition[164]:

- Genetic activation state, chromatin marking system (the chromatin marks are clonally inherited; recall that chromatin is the portion of the cell nucleus that contains all of the DNA of the nucleus in animal or plant cells [Sec. 7.4], and for normal gene transcription to occur, DNA must be accompanied by the chromatin marking system[165]), DNA methylation, paramutation[166] (a genome of one parental source that will not be inherited, influencing a genome of the other parental source that will be inherited), RNA interference for silencing specific genes [Subsec. 9.7.1], epigenetic variations, for instance in morphology, like the plant *Linaria vulgaris*;
- Structural inheritance, of which we have already seen some examples: Existing cell structures (in general, protein-made) are used to guide or template the formation of new similar structures, for instance, different patterns of cilia in ciliates or prions (proteinaceous infectious particles);
- The information about the whole organism is inherited, for instance, the developmental legacy of the mother in mammalians is transferred to her daughters, as we shall consider below;
- There is also enviroment-induced inheritance like dietary cues in maternal milk, and even ecological inheritance.

Maynard Smith and his collaborators acknowledged that epigenetic inheritance played a crucial role in the major transitions in evolution, for instance, in the symbiotic origin of the eukaryotic cell[167] [Subsec. 9.5.2]. Instead of speaking of several parallel systems of inheritance, it is perhaps suitable to see the life cycle as a system of resources where different channels are strongly interwoven.[168] What matters is the combination of several forms of inheritance that give rise to a number of different phenotypic effects and not the combination of single elements as such. We shall consider these effects further in the chapter on epigeny.

[162][*JABLONKA/LAMB* 2005, pp. 166–72]. [163][JABLONKA 2001].
[164][*JABLONKA/LAMB* 2005, pp. 119–6 and 162–6].
[165][GRIFFITHS/GRAY 2001] [*ALBERTS et al.* 1983, pp. 471–7].
[166][SOLOWAY 2006]. [167]See also [RIVERS/LAKE 2004]. [168][OYAMA 2001, OYAMA 2003].

9.9 Entropy and Order in Evolution

Evolution involves not only transmission of information but also irreversible changes in entropy and order,[169] for instance, in structural information (the protein network and the DNA considered as incorporating information) or in the complexity of the organism. Darwinism only requires that genetic variations are not causally linked to selection, but it does not imply that all of them are necessarily random. We have seen in the previous sections that evolutionary theory must be emended in order to explain the relationship between form and function in evolution[170] [Secs. 8.2 and 9.5]. Apart from structural information [Sec. 6.4], organisms also have instructional information, which can exist only when there is a whole semiotic system that allows codified information to become a set of instructions for doing something else [Sec. 7.4]. Information takes precedence over energy here, since it is instructional information, together with additional information coming from the metabolic system or the selection system, that determines what chemical reactions are allowed in the organism [Secs. 8.1–8.3] (it is a typical top-down process [Subsec. 6.3.2]). Moreover, codified information interacts with the regulatory system to determine how free energy and entropy will flow through an organism. As I have stressed, organisms, in terms of exchange of information, are closed systems, since the environment cannot directly cause changes in codified information, whereas they are open systems in terms of energy/matter (entropic) exchange [Subsec. 3.2.2 and Sec. 8.4]. In this process, the organism's organization (its structural information) is made possible because of the increase in the maximal level of entropy attainable when several systems are combined [Subsec. 2.3.4 and Sec. 6.4]. This is what has happened in the transition to eukaryotes and multicellularity and this is also what currently happens in any sexual reproduction. No system can violate entropy superadditivity: When two systems merge the entropy of the resulting compound system cannot be smaller than the sum of the entropies of the two initially separated systems [Eq. (2.12)]:

$$H(J,K) \leq H(J) + H(K), \tag{9.1}$$

which does not necessarily imply that the entropy of the resulting system is the same as that of the two subsystems taken separately.[171] This requirement is called homogeneity. Indeed, if entropy decreases as a result of a certain work performed in the merging of the two systems, there can be a violation of homogeneity, in which case the compound system can exhibit equilibria which are not entropy maxima, and this allows for further order. Indeed, the so-called Prygogine's principle tells us that, in the neighborhood of thermal equilibrium, the steady-state configuration is such that the entropy production is minimized.[172] Though the principle does not possess an absolute generality, it can suggest that evolution will promote organisms that minimize their entropy production (i.e. a state much lower than the maximum level attainable). These organisms should turn out to be precisely those that are able to exert more information control on the environment. Reciprocally, information control is made possible though entropic fluxes maintaining order.[173] This contributes to the clarification of the tight and circular connection between information and entropy in organisms and the dynamicity principle that we have established above [Subsec. 8.2.7].

The evolutionary mechanism through which those conditions are established was suggested by A. Lotka. At the evolutionary scale, natural selection should operate in such a way as to increase

[169][COLLIER 1988]. [170][*BROOKS/WILEY* 1986, pp. 2–59]. See also [BROOKS *et al.* 1984].
[171][LANDSBERG 1984a, LANDSBERG 1984b] [LANDSBERG/TRANAH 1980]. See also [LAYZER 1976, LAYZER 1977, LAYZER 1978].
[172][*PRIGOGINE* 1947, *PRIGOGINE* 1955] [*ULANOWICZ* 1986, p. 24]. [173][*ULANOWICZ* 1986, p. 87].

the total mass of the organic system, to increase the rate of circulation of matter through the system, and to increase the total energy flux through the system, so long as there is an unused residue of matter and available energy.[174] This may be expressed by saying that natural selection tends to make the energy flux through the system a maximum, as far as this is compatible with the constraints to which the system is subject. It is not lawful to infer immediately that evolution tends to make this energy flux a maximum. For in evolution two kinds of influences are at work: Selecting influences and generating influences. The former selects, the latter furnishes the material for selection. It appears probable, however, that among the very large (if not infinite) variety of abstract biological types subject to selection, the ones that will survive are those which give the opportunity for selection to operate in the direction indicated, namely to increase the total mass of the system, the rate of circulation of mass through the system, and the total energy flux through the system. If this condition is satisfied, the law of selection also becomes the law of evolution, and the latter goes on to make the total energy flux through the system a maximum compatible with the constraints, even if there is also an increase in dissipation.[175] The increase in the total energy flux will help organisms to maintain lower levels of entropy. In a circular feedback loop, increase in information control will in turn help in maximizing the total energy flux through the organism.

It is convenient to distinguish among three types of thermodynamic entropy and entropy-growing processes [Subsec. 6.2.4] involved in life[176]:

(1) *Configurational entropy*, which deals with the increment of variation and aperiodicity through dispersive and permutative processes (increment number and variety by different permutations of a certain number of molecules of an organism). It is particularly relevant for interaction among bases and for structural properties of proteins. It is a structural arrangement and therefore a matter randomization, as it is displayed in Fig. 2.7. Here, we can have an increase of heterogeneous complexity.
(2) *Thermal entropy*, which is concerned with the distribution of energy in a system and is therefore connected with chemical interconnections (structuring reactions lowering thermal entropy are in fact accompanied by the movement of thermal energy from translational modes to less densely spaced vibrational modes).
(3) *Energy randomization* due to the ontogenetic ability to constitute environmental niches, especially aiming at constituting environmental stores (fuel supply). Metabolic self-sufficiency is especially important in plants: In fact, as we have seen, animals have lost the ability to synthesize a significant fraction of the 20 essential amino acids [Subsec. 7.4.1]. Behavioral range (important for animals) and metabolic self-sufficiency (as in plants) cannot be jointly optimized.

These three forms of entropy are especially connected with the genetic system, the metabolic system, and the selecting system, respectively.

Another way to express these concepts is the following. As I have said, biological systems are dissipative.[177] Zotin[178] proposed a decomposition of the total dissipation function ψ of biological systems in external dissipation function (ψ_e) and bound dissipation function (ψ_μ). Brooks and Wiley in turn proposed[179] a decomposition of the latter into a part pertaining to the accumulation of biomass (the living matter, i.e. the material aspect of life) and another pertaining to the accumulation of genetic information (ψ_μ^b and ψ_μ^i, respectively), so that we finally have

[174][LOTKA 1922a]. [175][*ULANOWICZ* 1986, pp. 115–16].
[176][WICKEN 1980, WICKEN 1984] [PULSELLI *et al.* 2009]. [177][*PRIGOGINE* 1955].
[178][*ZOTIN* 1972, *ZOTIN* 1990] [*LAMPRECHT/ZOTIN* 1985]. [179][*BROOKS/WILEY* 1986, pp. 60–107].

$$\psi = \psi_e + \psi^b_\mu + \psi^i_\mu. \tag{9.2}$$

During short time intervals (physiological cyclic time), ψ_e predominates (and this corresponds to energy randomization due to behavior). Viewed in the intermediate time intervals (epigenetic time), ψ^b_μ predominates (and this corresponds to thermal entropy). But in the longest time intervals (evolutionary time) ψ^i_μ predominates (and this corresponds to matter randomization). This means that we have (1) phylogenetic, (2) epigenetic, and (3) ontogenetic (phenotypic) time scales. All of these three processes are irreversible and, as we saw in the previous section, and each one is bound to a specific form of transmission or heritage: Genetic, epigenetic, or behavioral (which, as we shall see in the next chapter, also comprehends to the transmission of an environmental niche).

According to Wicken,[180] any process in life must be coupled to relative reductions in energy randomization, thermal entropy, or matter randomization. Energy randomization and matter randomization in particular, provide the two entropic generative forces in evolution, provided that they can somehow be controlled (through teleologic causal processes, especially for energy randomization) or canalized (through teleonomic processes, for matter randomization). Note that the second law of thermodynamics does not mandate either increases or decreases in complexity with time, neither does it mandate increases in randomness. The complexity of a chemical structure depends not only on the number of elements it contains, but also on their variety and the aperiodicity of their interconnections [Sec. 6.4]. Matter randomization always promotes molecular heterogeneity; once a quasistable genome becomes connected with the metabolism of a primitive cell or protocell, this principle will force random alterations in the genome as a way of generating molecular heterogeneity [Subsec. 7.4.2 and Sec. 9.7]. Matter randomization promotes two kinds of reactions which are essential to molecular evolution:

- The first might be called *dispersive* reactions: they involve the formation of large varieties of biomonomers such as amino acids and nucleotides from small sets of reactants [Sec. 7.3].
- The second might be termed *permutative* reactions [Subsec. 7.4.1]: they involve the generation of ensembles of *alternative* molecular sequences from a given basis-set of chemical elements.

The thermodynamic force behind each kind of reaction is provided by the opening up of new configurational microstates whenever a new chemical species is formed, regardless of whether this new species happens to be more complex than its precursors but opening spaces objectively to new complexity. Summing up, all biological systems are nonequilibrium systems that operate—and autocatalytically produce themselves—by degrading energy resources. Natural selection is based on competitive success in autocatalytically converting resources into organization and order.[181]

9.10 The Baldwin Effect

Let us now discuss a problem that somehow represents the bridge between phylogeny and ontogeny: The Baldwin effect. As I have mentioned, one should distinguish between two flows: One from the environment to genes (accommodation) and the other the other way around (assimilation) [Sec. 8.4]. Since environment and genetic systems cannot act directly on one other (they represent the unknown and independent parameters relative to each other) [Subsecs. 8.2.1 and 9.2.1], they

[180] [WICKEN 1979, WICKEN 1988]. See also [WICKEN 1980]. [181] [WICKEN 1985].

always do this through phenotypes (which does not mean that phenotypes are instructed by the environment either):

- Natural selection acts on the phenotypes.
- On the other hand, the gene contributes to the determination of the phenotype,
- Which acts on the environment by changing it according to its needs (through what is called niche construction).

The final result is that the plastic ontogenetic action of phenotypes smooths or favors a given action of natural selection and teleonomically carves it in a determinate direction by bringing it to expression or by accelerating certain genetic predispositions, and in this way indirectly influencing their subsequent generations. This is the essence of the Baldwin effect.

Such an effect was simultaneously discovered by Baldwin, Lloyd Morgan, and Osborn.[182] According to Baldwin,[183] novelties coexist alongside established specialization. The question here is why organisms should adopt a new and eventually clumsy alternative. The reason is that some individuals would do even worse without this novelty, so that they adopt a lower fitness option when it is likely to be more profitable than other options. The subsequent evolution will then eventually refine this adaptation through exaptation [Subsec. 9.5.1]. The modularity of some traits allows for the independent selection of these traits. However, a corollary is the principle of coevolution of traits that are coexpressed: Traits that are expressed together are also selected together and evolve together as a set.[184]

According to Lloyd Morgan there are three levels of selection[185]:

(1) Germinal (epigenetic) selection (the differential survival of germ-like variants as they struggle *in utero* for scarce maternal resources), which goes in the opposite direction relative to ontogenetic adaptations,
(2) Organic selection (now known as the Baldwin effect), and
(3) Natural selection.

In the long run only the germinal elements that reinforce the direction of organic selection are retained. There is then a loop between ontogenetically adapted organisms and the arrow of evolution.

The motivation of many supporters of the Baldwin effects is that they are looking for a mechanism that is able to get a population into a "hard-to-find" part of the fitness landscape [Subsec. 9.5.5] in which—apparently very unlikely—evolutionary products like the human brain, language, and mind can rapidly evolve. The social and intelligent component considered by the Baldwin effect opens a breathing space for genetic variations and selection. On the contrary, Lamarckism would mean loss of flexibility since it would imply an immediate and effective fixation of certain traits because they are adaptive.

This raises the issue of *what* evolves during evolution.[186] Some authors have proposed that it is complexity,[187] others intelligence.[188] Both answers are probably correct but do not completely catch the root of the problem. Moreover, intelligence is something that does not fit well with elementary organisms like bacteria and probably not with plants and elementary animal species either. As we have seen, Baldwin also proposed that evolution goes in the direction of increasing

[182][DEPEW 2003]. [183][*BALDWIN* 1894, *BALDWIN* 1902]. [184][*WEST-EBERHARD* 2003, pp. 165–8].
[185][*LLOYD MORGAN* 1896, pp. 262–79]. [186][*PIAGET* 1967, pp. 122–3]. [187][SAUNDERS/HO 1976].
[188][*BALDWIN* 1902].

plasticity, whose utility seems so great as to outweigh all other biological characteristics.[189] The reason is due to the fact that the main issue in evolution is how populations deal with unknown situations. Then, evolution goes into the direction of increasing the organisms' information control on their environment [Subsecs. 8.2.2 and 8.2.7]. Indeed, to control more environmental parameters means that unexpected events are better dealt with, which makes it easier to survive. Moreover, it determines a circulation of more energy fluxes in accordance with the analysis of the previous section. This demands more complex solutions, increasing plasticity, and, at a higher level, intelligence (which becomes a specific and later *consequence* of this general trend of evolution[190]), since in order to control more information (that is, to have both more informational sources and more information about the same source, which in principle is reducible to the first aspect), the organism must:

(1) Become more receptive with regard to external information, and
(2) Be better able to carve out original sources of disturbances in cues that are useful for surviving and growing [Subsecs. 8.2.1 and 8.2.7].

Summing up, *plasticity* can be defined as the trade-off between information control and responsiveness to environmental stimuli, so that what grows during evolution is information control and not plasticity as such.

It is only by an application of natural selection from a set of overproduced functions that focal directed behavior can effect selective adjustments without violating the closeness of physical laws and the noninstructivity of the organism by the environment. Baldwin called this *functional selection*, i.e. it finally results in the selection of specific functions [Subsecs. 8.2.4 and 9.5.3]: Only those reactions of the organism will persist that could either be used for higher functional needs, or which at least would not stand in the way of the exercise of higher functions. This statement can only be completely understood when taking into account the results of the next two chapters.

Summing up, the Baldwin effect requires natural selection operating upon variations in the direction of more information control without loss of plasticity, which allows selective adjustments through the further operation of natural selection upon the organism's functions [Subsec. 9.3.1]. In other words, there are three general aspects involved here:

- Variation together with natural selection,
- Information control with plasticity (increase in assimilation), and,
- In between, accommodation by functional selection.

As anticipated, information control, which is a typical ontogenetic and individual activity, becomes the *result* of a phylogenetic process. Instead, during functional selection, *teleonomic* processes are mainly involved through which the organism is able to give rise to new functionalities and even covergences of functionalities [Subsec. 9.5.3]. Therefore, the range of possible accommodations of the organism as a whole becomes wider, and its congenital impulses less fixed as evolution goes on. The hereditary impulse is only a starting platform and the evolutionary succession of phenotypes becomes, in its hereditary character, more and more indeterminate with respect to what will be produced. The individual effort that reinforces a certain adaptation to the environment can help to establish the direction of evolution. Therefore, the Baldwin effect is strictly related to and, in the middle stage, even coincident with what is called *genetic accommodation*: The fine-tuning of a certain variation by eliminating its negative side-effects and integrating it phenotypically.[191]

[189] [*BALDWIN* 1902] [*SPENCER* 1855, pp. 388–406]. See also [*GODFREY-SMITH* 1996].
[190] In this way there is a certain progress without finality. For the opposite view see [*GOULD* 1996].
[191] [*WEST-EBERHARD* 2003, pp. 140 and 147–57] [*GILBERT/EPEL* 2009, pp. 384–5].

Baldwin's proposal is a middle way between Lamarckism and neo-Darwinism: It acknowledges that evolution reflects individual progress without adopting Lamarckian inheritance of acquired characteristics, but it does not follow the preformationist idea either, according to which the individual organism simply shows the unfolding of what its genetic endowment has made possible (this would be right only if the issue were a continued reproduction without evolution [Secs. 9.1 and 9.8]: Ultimately, a type of mechanics).

It was G. Simpson, under the influence of Schmalhausen,[192] who rediscovered the Baldwin effect and showed that it was fully compatible with a Darwinian theory. He wrote down three conditions under which the Baldwin effect could be effective:

(i) The ability to acquire a new character has in itself a *genetic basis*,
(ii) Selection for the ability to acquire an adaptive character *narrows* the developmental range in which the character would usually or invariably appear,
(iii) There is a certain *balance* between fragility and stability of developmental ranges and norms in evolution.[193]

Simpson's version of the Baldwin effect describes how natural selection in a changing environment may produce something looking like a Lamarckian transition in populations. On the surface, it seems that initially a few individuals acquire a new and beneficial trait. Over time, the frequency of individuals that possess the trait eventually increases. Almost simultaneously, alternative states become less diverse. This continues until almost all of the individuals have the beneficial trait [Subsec. 9.3.2]. What seems to be non–hereditary eventually reveals a genetic component. The quantitative model demonstrates that this phenomenon can be explained by natural selection on the phenotypic plasticity of a population in a three-step process:

- Individual organisms interact with their environment to produce behavioral or structural modifications.
- Genetic mutations in the involved populations produce hereditary characters similar to or supporting those modifications.
- Finally, the fitness benefit of the new relative-optimal trait outweighs the cost of phenotypic flexibility, causing a spread of that character in the population and a selection for narrow phenotypic ranges that contain the relative-optimal genetic trait.

Together the two processes—the (random) search for the new relative-optimal phenotype and the subsequent convergence on the phenotype—determine the transition time between equilibria in different successive environments. We shall see the relevance of these considerations when dealing with epigeny, since these random searches are most likely to occur in conjunction with developmental processes. Immediately following the initial emergence of a new evolutionary situation (induced by a particular environmental change) and until a single individual encounters what in certain circumstances and to a certain extent can be considered the new phenotypic optimum, that is, a good solution, the population will drift, maintaining a narrow distribution of phenotypic ranges. The first individuals to have reached this solution inside their norm of reactions typically lie on the very high end of the distribution of range sizes. *Norms of reaction*[194] are the various patterns of phenotypic expression of a single genotype across a range of environments. Strong selection for these relatively optimal types temporarily skews the distribution towards more extended norms of

[192] [SIMPSON 1953b] [*SCHMALHAUSEN* 1949].
[193] Simpson's interpretation of Baldwin has been put in quantitative form by L. Ancel [ANCEL 1999].
[194] [GOULD/LEWONTIN 1979] [GUPTA/LEWONTIN 1982].

reaction. After the population swings over to the new relative optimum, most individuals realize the fitness benefit of the new equilibrium, and selection for decreased norm length narrows the distribution to the new equilibrium.

It is not necessary to assume that evolutionary changes must occur at the genomic level, as Simpson still does. As mentioned, it suffices to say that an adaptive change happens somewhere in the developmental system, which has sufficient robustness and reliability to ensure recurrence of the adaptive phenotype in succeeding generations.[195] The division between inherited traits and ontogenetically acquired traits can then even appear arbitrary. The challenge is to understand how traits are expressed during each organism's life. Nonfunctional parts of the genome may undergo much more rapid changes than functional ones do [Sec. 9.7]: If the environmental situation changes, parts of the genome which formerly were expressed become nonfunctional and *vice versa*.[196] Thus, evolution, by way of forgetting unused items, shows that such an adaptation is more general than evolution solely via natural selection: A high percentage of newborns survive even if some genomic solutions have never been checked by natural selection. Summing up, the organism generates some of its own environmental inputs and incorporates them into a new developmental achievement. For instance, upright walking may determine some of the neurobehavioral asymmetries that are typical of our species.

9.11 General Principles of Evolution

According to Lewontin,[197] Darwin's theory comprises three principles [Sec. 9.1; see also Sec. 7.1]:

(1) Phenotypic variation (the present state),
(2) Differential fitness (on the basis of mutations that have already occurred),
(3) Heritage of fitness (transmission towards future generations).

He suggested that these principles are rather general insofar as the mechanism of inheritance is not specified [Sec. 9.8]. Both individuals and populations can be of very different kinds [Sec. 9.3]. Also, the reasons for differential fitness are not stated at this general level, and many forms involve many features. The three principles of Lewontin can perhaps be mapped to Campbell's ones of blind variation and selective retention,[198] since they in turn consist of[199]: A mechanism for introducing variation, a selection process, and a mechanism for preserving and propagating selected variants. H. Plotkin proposed a G–T–R heuristic: (1) *Generation* of variants, (2) a *test* (selection) phase, (3) *regeneration* of variants. The Lewontin–Campbell–Plotkin principles are the bridge to the generalized Darwinism we are looking for[200] [Subsec. 2.2.6]. In my opinion, the point (3) should be reformulated as a dynamical bound between the first two principles. In this case, we would have a G–S–T mechanism[201]: Generation of variants, selection, *transmission* of the survived variants [Fig. 9.12]. In this way, the Darwinian mechanism becomes a sort of Bayesian model in which the generation of variants is the hypothesis tested and corrected by natural selection and eventually further applied (transmitted) [Subsecs. 6.1.1 and 7.6.2]. Since selection implies here the establish-

[195] [MOORE 2003]. [196] [HOFFMEYER/KULL 2003]. [197] [LEWONTIN 1970].
[198] [CAMPBELL 1960]. [199] [*PLOTKIN* 1993, pp. 82–6].
[200] Hull proposed considering selection as a mechanism acting at different biological levels, and extended it as an explanatory tool, even to the dynamics of scientific knowledge [*HULL* 1988b].
[201] [TURING 1950, p. 456].

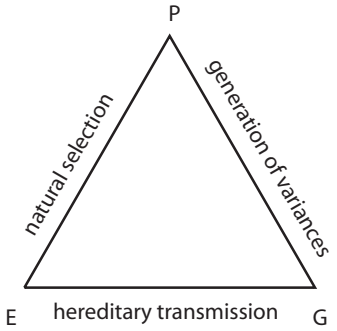

Fig. 9.12 Phylogenetic relationships between environment (E), phenotype (P), and genotype (G). The genotype is the source of variations, which are then exposed to the environment. The environment selects phenotypes and determines an indirect selection of the genetic pool through hereditary transmission. The informational "flow" is from the environment to the genetic system, even if these two do not interact directly but only through the phenotype. There is no direct relation between genotype and environment, apart from the indirect flow above. Recall that any action (even purely mechanical) of the environment on the phenotype, becomes information once it is integrated to the self as a cybernetic system [Secs. 8.1–8.2].

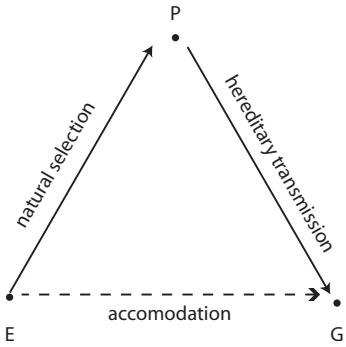

Fig. 9.13 Phylogenetic accommodation: Given the selective action of the environment on the phenotype and the fact that the survival of phenotypes is fundamental for genetic transmission (of mutations), as a final result the environment acts indirectly on the genotype and in this way it is accommodated to the external environment [Subsec. 8.2.1].

ment of a coadaptation, this is also in accordance with Peirce's investigation,[202] so that we could also speak in general terms of a Darwin–Peirce–Lewontin mechanism. Moreover, this mechanisms shows a good agreement with the three general (heuristic) principles I have formulated: Selection as meaning the occurring of random choices [Subsec. 2.2.3], establishment of correlations [Subsec. 6.5.1], and itinerant dynamics [Subsec. 8.2.7]. Obviously, the *indirect* action of the environment on the genotype (through selection of phenotypes) has as a result the final accommodation (one side of adaptation) of the organism to the former [Fig. 9.13]. Note that to say that there is indirect action does not eliminate the fact that here the organism is treated as an open system.

[202] [*PEIRCE W*, v. VI, p. 202]. Another interesting triad was introduced by Huxley in 1942: Origin–polarity (temporal order of the manifestation of alternative states)–spread [*WEST-EBERHARD* 2003, p. 198].

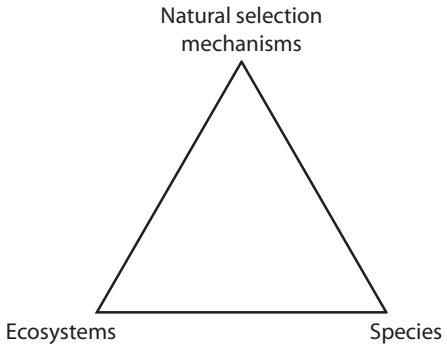

Fig. 9.14 The environment as a biological system.

Waddington proposed considering a feedback circle between 4 systems[203]: The exploitive (phenotypic) system modifies the environmental niche; this has an influence on the epigeny by revealing certain potentialities; the natural selective system acts as a filter; and finally the genetic system accounts for inheritance. This in turn acts on the exploitive system. This is conceived as a feedback circuit. I maintain that there are three systems in play here[204]—genetic, environmental, and phenotypic—since we cannot speak of an epigenetic system properly, being that epigeny is rather (at least a part of) the building process of the ripe organism, and also due to the fact that there is no existing evolutionary system or ontogenetic system. They are rather *processes* that involve different systems as well as other processes.

Let us now consider the reason for speaking of environment, genotype, and phenotype as three biological systems. The biological *environment* is a true biological system,[205] since it comprehends [Sec. 7.2]

- Species, which act as the information processor (the source of variety),
- The ecosystem, which functions here as the regulator with which the species have to fit: It is indeed the interface among the ontogenetic action of individuals of various species (assimilation) on the one hand, and between these species and the selective action of the external environment (accommodation), on the other,
- The external natural-selection mechanism (actually many different mechanisms), which is the decider [Fig. 9.14].

The genotype is also a biological system, since it consists in codifying information in the DNA, in the transmission and variation of information through a set of instructions given to the RNA, and finally in protein building through translation [Sec. 7.4]. Furthermore, the phenotype is a true biological system, since it is constituted by a mechanism of signal transduction, a metabolic subsystem, and a selective system, which in the most elementary case is the membrane system [Sec. 8.4].

Given the variants of generation–selection–transmission mechanism, we could point out three general ways to deal with environmental changes and uncertainties. Organisms can act

(1) *On the generation of variance*, for instance by multiplying the descendants. Prokaryotes and low eukaryotes have favored this strategy. Even when bacteria strongly transform their

[203][WADDINTON 1961b, WADDINGTON 1968b]. [204][BATESON 1963].
[205][*VERNADSKY* 1998] [LINDEMAN 1942] [*MARGALEF* 1968].

environment, this is still a consequence of a multiplication strategy and not a direct transformation of it [Subsec. 7.5.2]. With such a strategy, genetic variability is privileged (horizontal gene transfer).[206]

(2) *On the selection mechanism* (only indirectly), by transforming the environment (construction of a niche). Multicellular organisms have adopted this second strategy. There are many ingenious ways in which social insects, amphibians, reptiles, birds, and mammals do this. Here, epigenetic variability is privileged, since it is the phenotype that is involved in niche construction.

(3) *On the type of transmission*, by finding alternative forms of transmission, for instance social and cultural transmission [Sec. 9.8]. Humans have adopted this third strategy, centered on cultural variability.

9.12 Concluding Remarks

In this chapter we have discussed the fundamental contribution of Darwin and some subsequent developments towards focusing on the issue of the relation between evolution, information, and entropy:

- I have understood natural selection as that which discards what is not compatible with survival and reproduction: However, we should remember that there are many possible solutions that are compatible with those constraints.
- The major contribution of Fisher is that he pointed out the necessity to explain self-reproduction in terms of a combinatorics of discrete elements.
- We have dealt with the concept of species and arrived at the conclusion that tokogenetic relations are to be taken seriously into account.
- We have also seen that not everything happens by chance at an evolutionary scale, since there are many structural constraints at a physical level as well as those stemming from the complexity of organisms, especially related to its functionalities. This has several consequences, like many convergences (analogies) as well as punctuated equilibria (bursts of phylogenetic change followed by long periods of relative stasis). The concept of exaptation is also very important, i.e. many useful adaptations consist in a functionality developed by some structures that did not arise for that purpose.
- Major transitions (macroevolution) are difficult to explain within the traditional neo-Darwinian framework.
- According to the neo-Darwinian tradition, the competition for life is essentially egoistic. This is true for individuals. However, results due to game theory show that small communities of altruistic organisms can have more reproductive success than the surrounding individual competitors.
- Many parts of DNA, even those not coding, are fundamental in the role of information processor played by the genome and of the organism as controlling instance.
- There are different forms of inheritance (of which epigeny and cultural transmission are especially relevant). However, this does not imply the necessity of going back to Lamarckism. As a matter

[206] [DOBRINDT *et al.* 2004].

of fact, the Weismann barrier can be overcome, although the fundamental dogma of molecular biology is not violated.

- There are several modes and time scales for dealing with the necessity of avoiding the effects of growing entropy.
- The Baldwin effect tells us that serve reactions of surviving organisms would either serve for higher functional needs, or at least would not stand in the way of the exercise of higher functions.
- The general principles of evolution are: Generation of variants, selection, and transmission of the survived variants.
- With the action of to natural selection on the phenotype and also on its hereditary transmission, the environment acts indirectly on the genome by causing the organism to accommodate to it.

We have spoken of three different time scales of inheritance: Phylogenetic, ontogenetic, and epigenetic [Sec. 9.9]. In this chapter we have focused on the first form of inheritance. In the next two chapters we shall deal with the other two and show that there are fundamental interconnections between these three processes.

Appendix: The Eigen–Schuster Equation

Many evolutionary processes (of genomes or phenotypes) in a fitness landscape are regulated by the Eigen–Schuster equation.[207] Consider all genomes of binary sequence of length L. They constitute a set of $n = 2^L$ variants. If x_j represents the relative abundance or frequency of those organisms containing the j-th genome, the whole population at a certain time can be represented by the population's state vector

$$|\, x\rangle \,= \sum_{j=1}^{n} |\, x_j\rangle \,. \tag{9.3}$$

If f_j represents the fitness of the genome j, we may represent the landscape by a fitness-landscape state vector (it is like a second physical system relative to the population)

$$|\, f\rangle \,= \sum_{j=1}^{n} |\, f_j\rangle \,, \tag{9.4}$$

while the average fitness of the population is given by the vector

$$|\, \phi\rangle \,= \sum_{j} |\, f_j\rangle \otimes |\, x_j\rangle \,. \tag{9.5}$$

The probability that the genome j is (randomly) changed into the genome k is given by the matricial element m_{jk} of the mutation matrix

$$\hat{M} = \begin{bmatrix} m_{11} \ldots m_{1n} \\ \ldots \ \ldots \ \ldots \\ m_{n1} \ldots m_{nn} \end{bmatrix}, \tag{9.6}$$

[207] [*EIGEN/SCHUSTER* 1979]. See also [*NOWAK* 2006, 31–42].

which is a stochastic matrix. A stochastic matrix is characterized by the properties: (1) equal number of rows and columns, (2) each number in the matrix represents a probability (a real number between 0 and 1), and (3) each row sums to 1. For instance,

$$\hat{M}\,|\,x(0)\rangle = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_4 \\ x_4 \\ \dots \\ x_n \end{pmatrix} = |\,x(t)\rangle\,, \tag{9.7}$$

where $|\,x(0)\rangle$ represent the initial state of the genetic pool and $|\,x(t)\rangle$ represents the transformed state at an arbitrary time t. Similar considerations hold for the vector for the fitness of the genome. Obviously, the elements m_{jj} do not change any vectors (they are elements of the identity matrix).

The equation then takes the form

$$|\,x_j(t)\rangle = \langle f_j|\sum_k m_{kj}|f_j\rangle \otimes |x_j\rangle - \sum_{k\neq j}\langle f_k(t)|f_k(t)\rangle \otimes |x_k(t)\rangle. \tag{9.8}$$

10
Ontogeny

We have already taken ontogenetic processes like metabolism and information control into consideration [Chs. 7–8]. Let us now rather focus on some specific aspects that are related both to phylogenetic and epigenetic issues. I am using the word *ontogeny* in a slightly different sense from current usage. As it is often understood, ontogeny is synonymous with development. On the contrary, I am using the term as indicating the total biography of an individual from conception to death [Sec. 8.4]. Indeed, I have pointed out that the organisms' production is the basis of ontogeny. Then, problems like niche construction demand a specific ontogenetic treatment.

After having considered the flow of information as an irreversible process during the life of an individual, I shall more specifically deal with the three main stages of ontogeny: Development, maturity, and aging. Then, I shall consider the intrinsic randomness of the organism as one of the main sources of variations among individuals. Finally, I shall consider the most important aspect in which the ontogenetic information control of the organism is displayed: The construction of environmental niches.

10.1 Preliminary Entropic Considerations

Unicellular organisms like bacteria can seem truly immortal since, when they grow and the task of self-maintenance becomes more complex [Secs. 7.1 and 9.11], they split, restarting the whole process of information expression and transmission again. Things stand in a different way with multicellular eukaryotes. It is here that phylogeny is no longer the dominant life process and ontogeny and epigeny emerge in their autonomous significance. Let us therefore study ontogeny in multicellular eukaryotes.

Zotin and Alekseeva[1] distinguished between inducible processes (reversible deviations from the current steady state) and constitutive processes in ontogeny. The latter consists in an irreversible approach to a final steady state, the mature form, and, in a further stage, to the death of the organism [Subsec. 8.2.1]. The flow of information from DNA through RNA to protein synthesis involves an increase in configurational and thermal entropy[2] [Sec. 9.9]—leading to the expression of a subset of the original potential codified information (that arises from the reading of the actually transcribed triplets).[3] In fact, as we know, not all DNA is transcribed. This selection is the information cost of the transmission of information during the life of an organism [Secs. 2.2–2.3]. The process is irreversible. Therefore, the information flow cannot be inverted *once the transmission has occurred* [Sec. 9.8].

[1][ZOTIN/ALEKSEEVA 1985]. [2][JOHNSON 1987a]. [3][*BROOKS/WILEY* 1986, pp. 108–75].

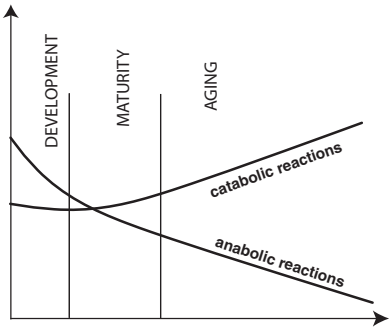

Fig. 10.1 The three stages of the ontogenetic path can be understood as being characterized by the relations between anabolic and catabolic reactions: During development the former prevails, while during senescence the opposite occurs with a growing divergence between the two curves. Maturity is the relatively stable state in between.

However, during epigeny we also have a significant ability to increase complex order even if the rate of gene activity (number of working genes per cell) falls. The genome is an organized unity, where successful changes in any part of the developmental program are those which are integrated with the unchanged parts (compensatory changes). All genes active in the early stages are part of the entropy-producing ensemble at all subsequent stages. Consequently, the sequential activation of portions of the genome during epigeny can be understood as an entropy producing access to *more and more* microstates in a growing ensemble (enlargement of phase space due to growing expression of potential genetic information). However, access to new microstates affects a progressively smaller proportion of the system due to its *growing complexity and growing size*. In this way the difference between the actual entropy and the maximal attainable entropy also grows [Subsec. 2.3.4 and Sec. 6.4].

Let us now consider the whole free-energy/entropy flux through the organism during its life. As we have seen [Secs. 7.3 and 9.9, Subsec. 8.2.7], organisms acquire free energy, produce a certain entropy during their metabolic processes, and download a part of this entropy into the environment. This means that the variation dS/dt of the whole thermodynamic entropy S (that is, the difference in entropy across the time t) [Subsec. 6.2.4] is given by the sum of two factors,[4] namely

$$\frac{dS}{dt} = \frac{dS_i}{dt} - \frac{dS_e}{dt}, \tag{10.1}$$

where dS_i/dt is the variation of the entropy internally generated by the organism and dS_e/dt is the variation of the entropy downloaded into the environment (it therefore has a negative sign since it diminishes the whole entropy present in the organism). Now, from the developmental phase across maturity through aging, there is a decrease of anabolic (order-building) reactions and an increase of catabolic (entropy producing) reactions due to the increasing needs of metabolism[5] [Fig. 10.1; Subsecs. 2.4.1 and 7.3.2], with the result that the entropy-energy balance becomes less favorable after development.

[4][NICOLIS 1979]. [5][*BENGTSON/SCHAIE* 2008].

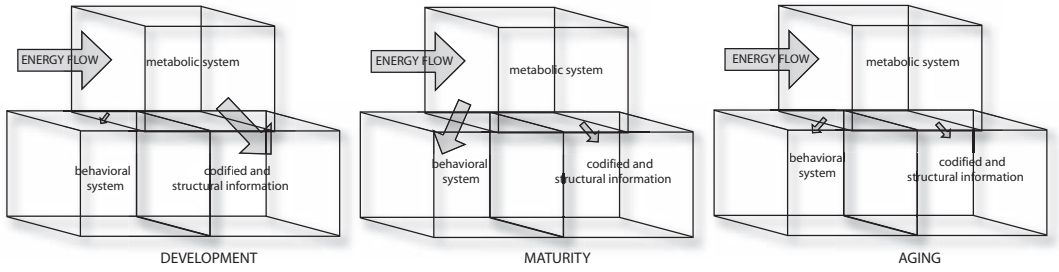

Fig. 10.2 During development we can assume that most of the free energy is used for information codification and its expression (building of structures). During maturity most of the free energy is used selectively and is goal-directed for controlling environmental information. Consider that there is a proportional decrease of the absorbed energy during life (not shown in the figure). The term *behavioral system* is provisional and only has a practical value here. An animal, especially a vertebrate, is able to display a very complex range of activities that go much further than the simple membrane-based selective system of bacteria. This complexity will be explored in the next chapters. We shall see that the brain plays a central role. Finally, during aging, most of the free energy should be used for metabolism.

10.2 Ontogenetic Path and Information

10.2.1 Three Stages of Life

Let us consider the whole of the energy flux in particular during the whole ontogenetic path. I assume here that at the beginning of life almost all free energy is used for constructing structures, while during maturity it is used for ontogenetic behavior (which shows that life uses the following methodology: First structures then functions), and finally for the metabolic waste of energy during aging [Fig. 10.2]. Here, we find the three main types of thermodynamic entropy and the use of free energy that was considered in Sec. 9.9. From an ontogenetic point of view, life could then be divided into three major phases[6]:

(1) Development:
- At the beginning of an organism's life, a maximum of energy flow goes into creating organization while a small part is used in metabolism. After an initial increase, the specific metabolic energy—the energy flow divided by the biomass—it is likely to decrease.
- In the early stages of development, the organism shows a high degree of plasticity: The ability of an organism with a given genotype to appropriately change its phenotype in response to changes in the environment is called phenotypic plasticity. In general terms, *plasticity* [Sec. 9.10] is an optimal trade-off between information control (preservation of the ontogenetic path) and information sharing with the environment (responsiveness to environmental stimuli). There is a convergent increase in complexity, that is, in the size and number of the organism's components and their interactions. Here the organism shows an increasing degree of control on its ontogenetic path by driving it teleologically [Sec. 8.2]. The integration of a high level of plasticity and of an increasing level of control is a definition of development.

[6][*SALTHE* 1993, pp. 95–137]. See also [*WICKEN* 1987] [SCHNEIDER 1988].

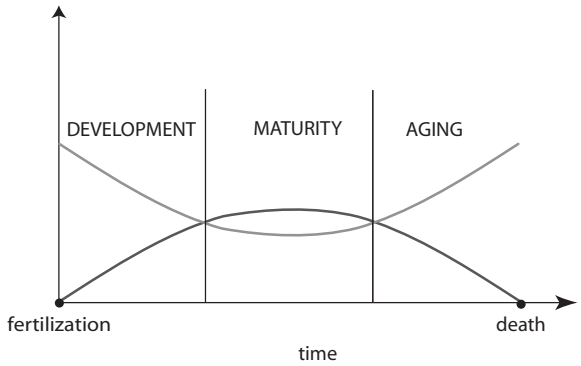

Fig. 10.3 Complementarity between channelization (light gray) and canalization (dark gray) along the ontogenetic path. The graph shows that there are probably specific values of both channelization and canalization below which and above which the organism can no longer survive. If this guess is correct, it implies that the death of the organism at the end of its life can occur due to a too-high channelization, a too-low canalization, or both. My understanding is that both maturity and aging are started when the two curves cross. Experimental data can be useful for confirming or disconfirming this hypothesis.

(2) Maturity: The rate of development slows down and the organism finds a relatively stable equilibrium [Sec. 10.1] where small environmental fluctuations, which can be dealt with, tend to accumulate slightly and are progressively more difficult to integrate. Moreover, in such a stable state, a sufficiently strong environmental shock may still cause the breakdown of the organism.
(3) Senescence or aging: is determined by the accumulation of the disruptive effects produced by environmental fluctuations. The accumulation of errors and shocks together with the fixation of certain structures (increase in rigidity and loss of plasticity) during the lifetimegrow more and more.

These three stages are the ontogenetic instantiation of a general succession of stages in ANY dynamic process out of equilibrium: Growth, steady state, decline.[7] This is the general, system-theory result about the kind of dynamic processes involved in biology and cognition and can be considered a consequence of the dynamicity principle [Subsec. 8.2.7]: It is the specific form that itinerancy takes in evolvable and developmentable systems [Subsec. 9.5.5]. As we have seen, energy and entropic fluxes are determinant here, as they also proved to be in the previous section. However, when the ontogenetic path of an organism is at play, information acquiring and control is also a major factor. As such, the whole ontogenetic trajectory of an organism from an informational point of view could be understood as a compromise between two complementary features [Subsecs. 2.2.3 and 2.2.5]: Channelization with the environment and local canalization [Fig. 10.3]:[8]

(i) At the beginning of development, the organism is practically uncanalized and very open to external cues (it can easily be disrupted). During development, canalization increases while channelization decreases. In particular, we may distinguish three phases: (1) Just at the beginning, when canalization is very low, any environmental fluctuation is very dangerous, since the newly formed organism displays a very low level of information control. This corresponds

[7][*ODUM/ODUM* 1976, pp. 62–3]. [8][AULETTA 2005a].

to the stage of development called fertilization. (2) Later, environmental fluctuations of the same strength as those causing the disruption of the organism during fertilization can induce the choice of a new path. This stage is what I strictly call *epigeny*. We have here a high level of plasticity with a high level of information control. It is very important to stress that during the first phases of development (cleavage and gastrulation) environmental stress can be lethal, while later on (in a period between weeks 3 and 8 of human gestation) there is maximal susceptibility to teratogens (disruption factors inducing malformations).[9] (3) In the third phase of development, it is increasingly difficult to change paths. At this stage, where canalization and channelization curves tend to cross, it corresponds to the postepigenetic phase of development called maturation, the bridge to sexual maturity (my guess is that this passage is determined precisely by that intersection). Here the metabolic energy slows down and plasticity also begins to decrease as behavior is fixed.

(ii) During maturity we have a sort of equilibrium between the two complementary aspects of canalization and channelization.

(iii) Aging is the reverse process relative to development, and can be seen as an increase of channelization at the expense of the capability to canalize external inputs. The beginning of aging is likely to be characterized by a new intersection of the canalization and channelization curves. When death occurs, we have zero information control and maximal channelization, since the organism goes back to the environment.

In other words, development is a process in which canalization and channelization converge, maturity is a stage in which they are in relative equilibrium, and aging is a stage in which they diverge. It is important to stress that here canalization and channelization are the ontogenetic categories corresponding to assimilation and accommodation at a phylogenetic level [Sec. 8.4].

10.2.2 The Homeorhetic Plateau

Life can then be understood by means of the homeorhetic plateau [Fig. 10.4]. Homeostasis is the tendency of an organism to stabilize some parameters against environmental fluctuations. Waddigton[10] preferred to speak of homeorhesis, that is, of an ever-changing, dynamic tendency to *transient* equilibrium [Sec. 8.4]. As I have said, it is a case of itinerancy, according to the dynamicity principle [Subsec. 8.2.7]. In the first phase of life (development), positive feedback prevails, giving rise to the epigenetic process, and negative feedback from the environment is controlled (thanks to antifeedback,), promoting the self-organization and growing complexity of the organism [Subsec. 8.2.1]. In the mature phase, negative and positive feedbacks are in equilibrium, and there is a relatively high antifeedback (which has a regulatory function). This ensures self-conservation and enables the organism to be on the homeorhetic plateau. Finally, we can assume that in aging uncontrolled negative feedback (environmental fluctuations) and autonomous (no longer able to deal with external stimuli and therefore unable to provide for canalization) antifeedback (with the consequent crystallization of structures and behavior) disrupt the organism. That is, during senescence the organism increasingly lacks both control and plasticity.

Let us consider the problem from this point of view: Channelization decreases during development and increases during aging, while canalization is the opposite. Considering the negative environmental feedback as constant, this amounts to saying that, in the first stage, positive feedback prevails, that is, environmental stimuli can be integrated as cues in the developmental process of the

[9] [*GILBERT/EPEL* 2009, pp. 171–2]. [10] [WADDINGTON 1974].

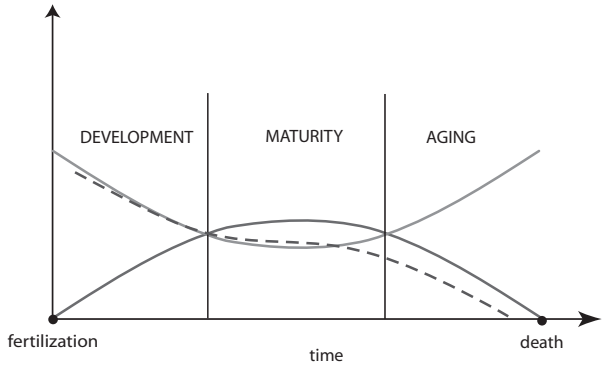

Fig. 10.4 The homeorhetic plateau. The dark gray and light gray curves again represent canalization and channelization. My guess is that the amount of plasticity of the organism reaches a maximum in the developmental stage, finds a relative stable state during maturity (the homeorhetic plateau, where a relative equilibrium between anabolic and catabolic reactions allows for maximal investment in behavioral energy randomization [Sec. 9.9]), and decreases during aging. This is represented in the ontogenetic path of the organisms, shown with the dotted line which represents a growing canalization (in the sense of a growing fixation) of the organism.

organism, while, during aging, antifeedback blindly prevails, with the consequence that negative feedback has disruptive consequences.

10.2.3 Chreods

During development, in the first phase of the ontogenetic path, the organism gradually approaches a species-specific steady state, which ends the first period of life. Toward the end of development, it also reaches its specific individual equilibrium state among the many it potentially had access to. This phenomenon was called canalization by Waddington with the meaning of a growing determination and fixation of phenotypic characters[11]. This is a meaning slightly different from the one that I use which is more general, implying also the organismic ability to carve out environmental stimuli.

The fact that the organism can use different environmental cues for its development shows that it is able to select specific environmental features [Sec. 7.6 and Subsec. 8.2.6]. But which cues are selected in turn depends on the developmental path and on related feedback circuits, that is, on the interplay between organism and environment.[12] Accordingly, one cannot take the entire environment as a supply of information that is able, together with the initial state of the organism and the genetic memory, to determine a univocal dynamics [Subsecs. 8.2.1 and 8.2.7]. Here, the final result is not simply given by adding the environment to the organism since the dynamic interplay between organism and environment (included its own environmental niche and other species) is crucial. This means, as already stressed, that ultimately the genome is *blind* relative to the final result of the epigenetic process—the species-specific phenotype. This is a true emergent system—emergent through this interactive process—and the only thing that does matter is the fact that it is an apt solution, even if not necessarily the optimal one [Subsecs. 9.2.3 and 9.5.5], given its environmental conditions. This is what *adaptation* really means.

[11] [WADDINGTON 1961a]. [12] [*THELEN/SMITH* 1994].

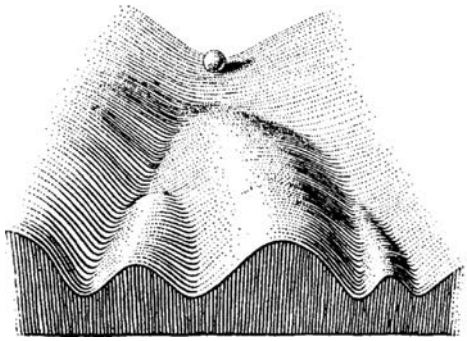

Fig. 10.5 Waddington's epigenetic landscape. The ball (figuratively representing an individual organism) can roll down along different paths. The more they progress away from the top, more divergent the different valleys (chreods), representing different epigenetic paths (the epigenetic segment of an ontogenetic path), become. This means that a bigger environmental shock becomes increasingly necessary in order to change epigenetic path.

Let us consider this interactive process. For Waddington,[13] whose contribution to the understanding of epigeny has recently become very highly appreciated,[14] the organism (especially in its developmental phase), following a disequilibrium depending on an environmental change, is able to produce an original response among the many possible ones (already potentially present in the genes). However, it is not the environmental stimulus alone that produces such an effect. A certain specific sensibility to the stimulus and a certain reservoir of possible responses are also necessary. This is the norm of reaction [Sec. 9.10]. Instead, in the Lamarckian and behaviorist framework, the response is simply a copy of the stimulus.[15] During development, it becomes more and more difficult to alter the development if it diverges from the norm: The canalization of an organism may be seen as a ball rolling on a rugged landscape, where there are different alternative pathways (chreods) [Fig. 10.5].

This is the reason why Waddington preferred to speak of a dynamic homeorhesis instead of a homeostasis. As time flows, the walls of the chreods become more and more steep, but, by catastrophic or highly shocking events, a system may still be displaced from one chreod to another.[16] When maturity is attained this becomes very difficult, since the shocking event that becomes necessary here could destroy the organism itself. Fig. 10.4 shows that, while control on the environment is still weak during development (if the organism jumps from one developmental path to another, this is indeed due to the environment's uncontrolled action), the inversion of aging is due to the fact that the curves diverge so that it becomes impossible to bring an environmentally uncontrolled effect to some alternative stability represented by another chreod.

10.2.4 Individuality Emerging

As a complex multicellular organism develops, it becomes more clearly *an individual* because in the course of its transactions with the environment it acquires more and more particular phenotypic traits; this is especially true for mammals, which show, as we shall see in the next chapter, strong regulative epigenetic processes. At the beginning of life, there is vagueness and

[13][WADDINGTON 1953, *WADDINGTON* 1957]. [14][REIK/DEAN 2002]. [15][*EDELMAN* 1989, p. 12].
[16][WADDINGTON 1961a, WADDINGTON 1974].

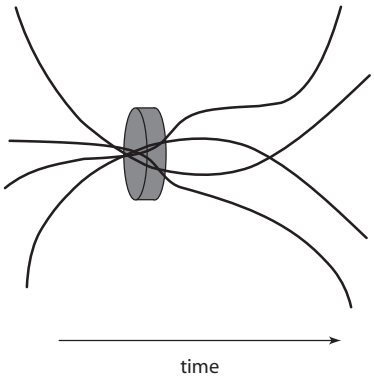

Fig. 10.6 Ontogenetic paths of different but closely related species: Starting from different points, the ontogenetic paths converge in a broad area (the ring) representing the phylotypic stage that in some species can be the end of development, and then diverge again with aging (growing specification).

indetermination (a form of generality). Therefore, development always presents both predictable aspects and new ones, that is, the growth of the particular differentiation of individuals in a species. Developmental plasticity leads to structural and behavioral divergence between individuals and thus to individualization. Such a process can be understood by considering the fact that genes only specify a norm of reaction, which is based on the variety of possible individual phenotypes thanks to the interaction with different environments and given a certain genetic endowment[17] [Sec. 9.10]. This is widely acknowledged today.[18]

This also means that there cannot be a complete individuation in the course of life: This would be represented by death, which is the definitive fixation of the organism[19] by cessation of any metabolic activity.[20] Thus, organisms are constitutionally open, they are never complete, and are therefore able to interact with the nonself [Sec. 8.4].

Among species there is a similar mechanism. As a matter of fact, closely related species differ widely in the ways that their early developmental stages proceed. In a later developmental stage they converge towards a shared step (the so-called phylotypic stage), only to diverge again, culminating in different adult forms[21] [Fig. 10.6; Subsec. 9.5.5].

10.2.5 Intrinsic Randomness

Differences between individual organisms are traditionally attributed either to *genetic variations* or to *environmental fluctuations*. For instance, it has been pointed out that genetic information is not transmitted with perfect accuracy nor expressed with 100% reliability[22] [Sec. 7.4]: There are indeed errors in gene expression (in transcription and in protein synthesis). Recently it has been shown[23] that there is noise in the eukaryotic gene expression. Moreover, according to Lewontin,[24] ontogeny is due to the interaction between genetic regulation, environment, and random events within the cell at the scale of molecular interaction. Again, this should be the effect of quantum and complex fluctuations [Subsecs. 6.5.1 and 9.3.1]. Finch and Kirkwood[25] proposed *intrinsic chance* at the onto-

[17][GOULD/LEWONTIN 1979] [GUPTA/LEWONTIN 1982]. [18]For plants see [SULTAN 2000].
[19][*SALTHE* 1993, pp. 139–244]. [20]Weismann spoke of the arrest of life [*WEISMANN* 1889, I, p. 114].
[21][*MINELLI* 2003, p. 124]. [22][SYMER/BENDER 2001]. [23][BLAKE *et al.* 2003]. [24][*LEWONTIN* 2000, pp. 17–18].

genetic (and also epigenetic) level as a third irreducible factor, in addition to genetic variations and environmental fluctuations. Such interactions between molecules that influence key outcomes in cell differentiation and development are inescapably governed by chance. Furthermore, during aging, chance acts through random damage to DNA and other molecules. A remarkable source of variance is due to differences in molecular control, for instance in error correction *during the lifetime*.

Let me give some examples:

- A detailed study on humans shows that the first variation across individuals is the life span, that can differ by up to 35% in identical twins and cannot be attributed to genetic variance or differences in environment and food resources.
- The age of reproduction among individuals varies a lot.
- The degree of constancy in the number of cells in adult tissues varies strongly. In the case of oocytes, the death cells will be not replaced, but rather the decay process is random, as is the case with radioactive decay [Subsec. 1.2.8].
- There are errors in DNA methylation that can cause aging syndromes, for instance cancer.[26]
- In the brain there are similar random processes in cell death.
- The brain structure in monozygotic twins already shows unpredictable differences in fetal development. There are even strong differences between twins in the functional circuits of the brain. As we shall see, this is consistent with Edelman's hypothesis that variations in neuron numbers are the source of individual variations in neural circuitry, including cortical maps.
- Among different individuals, there is a high variability of heterochromatin in the human chromosome.[27]
- There are even variations within a single individual. For instance, the lateral line system of cool-blooded vertebrates shows a high variance in the individual number of cells in each node. Also, an asymmetry in cellular division can occur (for instance, in the case of multicellular organisms, during early development). Daughter cells show different development potential and also some structural differences. There is also asymmetry in the position of body organs, an asymmetry that can be altered depending on several factors occurring in development. The same is true for brain asymmetries.

10.3 Niche Construction

10.3.1 What is the Relation between Organism and Environment?

I recall here what we have observed when the concept of teleology was introduced [Subsec. 8.2.7]: Organisms need specific causal modalities to control their environment. Schrödinger[28] understood that the behavior of individuals is relevant for evolution because it represents a feedback relative to random mutations: Changes in behavior strongly reinforce the usefulness of certain mutations. We have already considered this aspect from the point of view of phylogeny: The so-called Baldwin effect [Sec. 9.10]. Baldwin has indeed attributed much importance to the reaction of the organism to environmental stimuli. There is therefore a strict relationship between the use of a mutation and its being further implemented. In these cases, the information control of the organism manifests itself in its highest form, especially (in the case of multicellular organisms) during maturity.

[25][*FINCH/KIRKWOOD* 2000]. [26][*GILBERT/EPEL* 2009, pp. 267–83]. [27][CRAIG-HOLMES/SHAW 1971].
[28][*SCHRÖDINGER* 1958, pp. 107–14].

Recall also that the vast majority of morphological, behavioral, and physiological differences among individuals do not "Mendelize." The reaction of many biologists has traditionally been to recur to polygenic control. However, there is no evidence that this can stand as a general explanation.[29] All that genes do is specify a norm of reaction over environments.[30] What we know shows that the developmental responses of different genotypes to varying environments are nonlinear. Phenotypes are not determined even when the genotype and the environment are completely specified [Subsecs. 8.2.1 and 10.2.4–10.2.5]. Moreover, when considering species, there are no means for determining beforehand exactly which environment will be occupied by phenotypes: We only recognize a specific environment once we know the organism living in it. In fact, the environmental niches are made by the organisms themselves: Organisms (1) determine what is relevant, (2) alter the external world, (3) transduce the physical signals of the external world and metabolically assimilate free energy.

The mathematical law describing the interaction between organisms and environment has traditionally been formulated as[31]

$$\frac{dO}{dt} = f(O, E)\ , \quad \frac{dE}{dt} = g(O, E), \tag{10.2}$$

where E represents the environment, O a species, and f and g are two functions. Eqs. (10.2) are very general and expressed in neutral form, so that they tell us little about the specific mechanisms at play here. I recall that the traditional neo-Darwinian point of view, according to which the organism has no effect on the environment, can be formulated in terms of Eqs. (10.2), that is, more explicitly as

$$\frac{dO}{dt} = f(O, E)\ , \quad \frac{dE}{dt} = g(E). \tag{10.3}$$

What Eqs. (10.3) tell us is that the organism depends on the environment but not *vice versa*. However, active niche construction is a fact of life: Merely by existing (as in the case of bacteria), organisms must change their local environment to some degree (for instance, plants produce oxygen and some bacteria decompose animal and vegetal tissues). This is a consequence of the entropic openness of organisms, and therefore of their metabolic activity [Sec. 9.9]. Moreover, the niche represents the amount of parameters that are under informational control by the organism (can be predicted) and therefore it represents the class of interactions that can guarantee the homeorethic path.[32]

There are mainly three ways by which an organism changes its environment: Through perturbation of the environment, reallocation of resources (which is typical of animals), and reshaping and reinventing the environment, as happens in human culture. The latter behavior will be considered in the next part of the book. Many animals that dig burrows or build nests exhibit characteristics that are anatomical or behavioral adaptations to their ancestors' niche construction.[33] An interesting special case of the action of the organisms on the environment is represented by stigmergy

[29][LEWONTIN 2001].

[30][*LAMARCK* 1809, I, p. 8]. The contribution of Lamarck on this point has been stressed by Piaget [*PIAGET* 1967, pp. 104–14].

[31]See also [*LEWONTIN* 2000, p. 101][LALAND *et al.* 2001].

[32][*MATURANA* 1970, pp. 10–11].

[33][*ODLING-SMEE et al.* 2003, 69–101].

[Subsec. 2.4.2]: The indirect interaction of some individuals with other individuals through a particular environmental modification that can affect the behavior of the latter.[34]

Recall that Lloyd Morgan[35] correctly distinguished here between natural *selection* and natural *elimination*. The latter amounts to what is ordinarily called natural selection today and consists of negative feedback [Subsec. 9.2.3]. The former, instead, consists of the individual choice of organisms which becomes very relevant when their action on the environment is considered.

Taking into consideration these aspects, Odling-Smee, instead of Eqs. (10.2)–(10.3), proposed another set of equations:

$$\frac{dO_{\text{pop}}}{dt} = f(O_{\text{pop}}, E_{\text{pop}}), \tag{10.4a}$$

$$\frac{dE_{\text{pop}}}{dt} = g(O_{\text{pop}}, E_{\text{pop}}), \tag{10.4b}$$

$$\frac{dO_{\text{pop}}E_{\text{pop}}}{dt} = h(O_{\text{pop}}, E). \tag{10.4c}$$

The main point here is to introduce the distinction between an environment of a particular population or lineage (E_{pop}), i.e. its environmental niche, and the universal environment (E), which also comprehends the ecological niche—we have made a similar distinction when speaking of fitness landscapes [Subsec. 9.5.5]. The universal environment acts eventually catastrophically and unpredictably on populations, while the ecological niche acts in a canalized way on individuals, that is, canalized through their own niche-building [Fig. 10.7]. In this sense, an ecological niche is a functional entity advantageous for the survival of a species.[36]

Even with the independent renewal and depletion of key resources, the effects of niche construction can override external sources of selection to create new evolutionary trajectories and equilibria as well as to produce time lags for plastic adaptation,[37] as mentioned. Niche construction indeed introduces feedback into the dynamics of evolution and ecosystems through the accumulating action of individuals (and therefore, as a result, through the action of a species). When it creates time lags in the response of a trait to modified selection pressures,[38] it contributes to the Baldwin effect [Sec. 9.10]. In other words, an organism acts on its environment by smoothing some features and by sharpening other ones, in order to modulate the environmental effects.[39] For this reason, niche construction (at the ontogenetic level) and natural selection (at the phylogenetic level) show interesting interferences.

I wish to stress that this phylogenetic canalization is always an *indirect* effect of the organism, and in this way is a pure teleonomic process through which antifeedback finally translates into regulatory effects [Subsec. 8.2.1]. Both the broad physical environment and the action of individuals represent true dynamic causes acting directly on populations and the environment, respectively, while the resultant niche is a complex of interconnections in the sense of structural and formal causes [Subsec. 6.3.2] canalizing certain phylogenetic courses.[40] This allows us to make use of the concept of canalization both for the ontogenetic path and for evolution.

[34][BONABEAU *et al.* 1999, pp. 14–17][BONABEAU/THÉRAULAZ 2000].
[35][*LLOYD MORGAN* 1891, p. 79][*LLOYD MORGAN* 1896, pp. 152–6 and 270]. [36][*ELTON* 1927].
[37][LALAND *et al.* 1999]. [38][*ODLING-SMEE et al.* 2003, 114–15]. [39][*LEWONTIN* 2000, p. 60].
[40][LALAND *et al.* 2008].

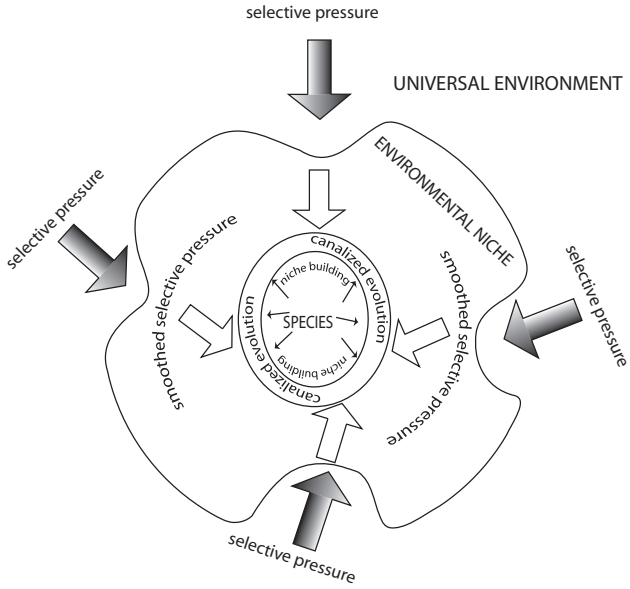

Fig. 10.7 Canalization of the evolutive process as an imbalance of a smoothed selective pressure (through the environmental niche) and the antifeedback effect of the individuals of a given species when they build the niche.

This concept of niche is strictly related with the German term *Umwelt*, which could be understood as the sum of problems which the organism faces.[41] Therefore, the concept of niche is also strictly related to that of *affordance* [Subsec. 4.4.3], which we shall consider in detail later on.

Summing up,[42] niche construction can (1) cause evolutionary inertia and momentum, (2) lead to the fixation of otherwise deleterious alleles, (3) support stable polymorphisms where none are expected, (4) eliminate what would otherwise be stable polymorphisms, and (5) influence disequilibrium. As mentioned, these results suggest that the changes that organisms bring about in their niche can themselves be an important source of natural selection pressures, and imply that evolution may sometimes proceed in alternate cycles of selective pressure and niche construction.

10.3.2 Ecological Networks

Therefore, an environmental niche can be considered both as

- The sum of all ordinary natural selection pressures to which a certain population is exposed and
- The accumulated result of individual organisms' work.[43] In this way an ecological inheritance may also be established [Sec. 9.8].

Different lineages experience the same environmental change quite differently. In other words, given the existence of ecological niches, environmental variations do not act as mechanical inputs, but

[41] [*VON UEXKÜLL* 1909, pp. 89–90]. [42] [LALAND *et al.* 1996]. [43] [*ODLING-SMEE et al.* 2003, 37–50].

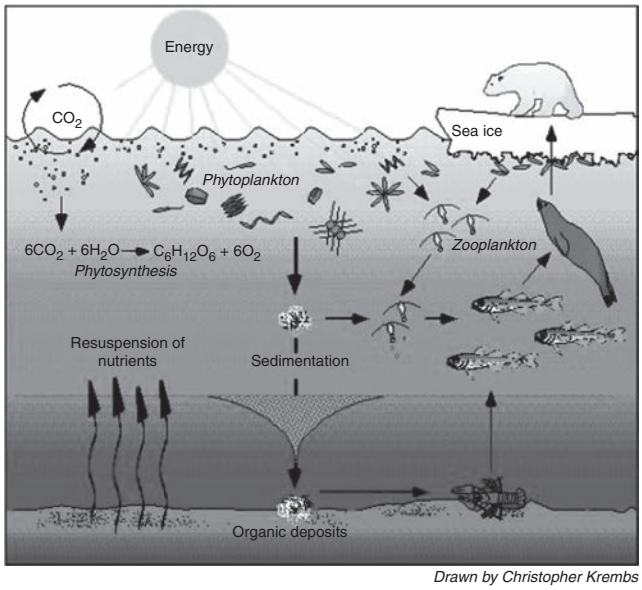

Fig. 10.8 Example of the arctic foodweb. Artist: Christopher Krembs. Adapted from http://oceanworld.tamu.edu/students/forams/images/arctic_marine_food_web_90.jpg.

rather their effect depends on the specific way a species is able to cope with them.[44] A fascinating subject is the symbiotic relations between different species, which constitutes whole entrenched niches or subniches in this way. Here, we not only have cases of parasitism but also of mutualism, in which both partners benefit from the interaction.[45]

It is interesting that gene pools can also interact or influence other gene pools through some niche factor, for instance genes influence prey choice in flamingos, and the food (crustaceans) influence genes for pigment extraction, determining their characteristically pink color. As a consequence, we have a model in which a certain genetic network G_1 influences another genetic network G_2 through environmental resources R.[46] In this way, there is a certain genetic modification (at least in expression) without the intervention of natural selection and an ecological-induced genetic inheritance.

Therefore, different species and niches interact in a larger environment constituting an ecological network (or ecosystem) as a functional system including an ecological community of organisms together with the physical environment. Ecosystems are characterized by the flow of energy through food webs, production and degradation of organic matter, and transformation and cycling of nutrient elements [Fig. 10.8; Sec. 9.9]. This production of organic molecules provides the energy base for all biological activity within ecosystems. In this respect, we distinguish between [Subsec. 7.3.2] autotrophic plants that are the *producers*, heterotrophic animals that are the *consumers*, and heterotrophic bacteria and fungi that are the *decomposers* or detrivores.[47]

Ecological networks, being centered on both exchange of information and energetic and entropic fluxes, are ruled by principles that are slightly different from other networks, especially relative to

[44]See also [GRIFFITHS 2003]. [45][*GILBERT/EPEL* 2009, pp. 79–114].
[46][*ODLING-SMEE et al.* 2003, pp.133–58]. [47][LINDEMAN 1942].

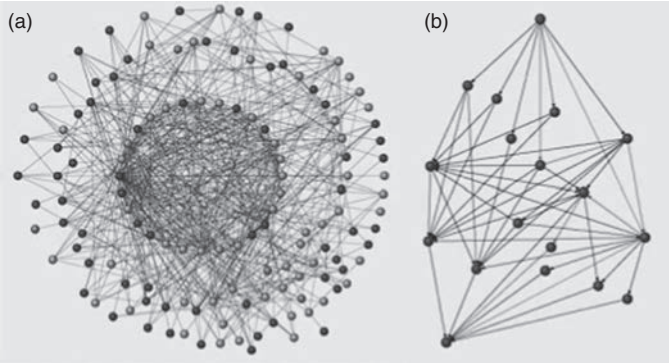

Fig. 10.9 Example taken from the Ythan estuary food web.

(a) The most connected species (the fluonder *Platichthys Flesus*) is shown in red (the trophic direction of the links—what eats what—is ignored). Dark green dots represent species that are one link apart; light green two links; and blue three links. The central circle represents the densest subnet consisting of 28 species with at least 7 links with the other species inside the subweb.

(b) Food chain between basal species of *Enteromorpha* (red node bottom) and the top predator, the cormorant *Phalacrocorax carbo* (red node top). Shortest path in blue (2 links), longest path in red (6 links). Adapted from [MONTOYA *et al.* 2006]. (The figure is reproduced in color in the color plate section.)

those dealing with information exchange only, like neural networks [Subsec. 3.8.3], genetic [Sec. 7.4], and epigenetic ones (to be considered in the next chapter), or even the worldwide web.[48] The two main principles of ecological networks are[49]:

- No species is too distant from the most connected species [Fig. 10.9(a)],
- No top predator is too distant from a species at the base of the web [Fig. 10.9(b)].

By "distance" I mean the minimum number of links connecting two species.[50] In contrast to webs like the internet, where the rich get richer, in ecological networks feedback effects between predators and prey assure an equilibrium. If the vector $|\, n(t_0)\rangle$ describes the population (the number n of individual organisms) of a species at an initial (arbitrary) time t_0, the rate of change is given by[51]

$$\frac{d}{dt}\,|\, n(t_0)\rangle \;=\; \text{births} \;-\; \text{deaths} \;\pm\; \text{migrations}. \tag{10.5}$$

By taking into account the limitation processes (in food resources, for instance), we obtain [see also Eqs. (10.4a) and (6.16)]:

$$\frac{d}{dt}\,|\, n(t_0)\rangle \;= r\left(1 - \frac{n}{K}\right)|\, n(t_0)\rangle\;, \tag{10.6}$$

where r and K are positive constants. The variable K represents the carrying capacity of the environment determined by the available sustaining resources while $r\,(1 - n/K)$ is the *pro capite* birth rate.

Note that ecosystems and niches show a certain spontaneous *directionality* in evolution.[52] Indeed, the metabolic processes of an ecosystem are an example of a thermodynamic system

[48] [MILO *et al.* 2002]. [49] [MONTOYA *et al.* 2006]. [50] [*BARRAT et al.* 2008, pp. 7–10].
[51] [*MURRAY* 1989, pp. 1–4].

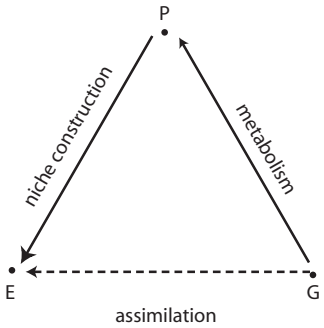

Fig. 10.10 Ontogenetic circle. The phenotype, as a complex result of the genetic process, builds its environment in an informationally and energetically suitable way. Here, the informational flow is from the genetic system to the environmental system, even if the environment and gene do not interact directly.

far from equilibrium [Subsecs. 6.3.1 and 6.5.1]. As such, the long-term succession of an ecosystem exhibiting a metabolic exchange of its constitutive species with the external environment proceeds in the direction along which the irreversible decay rate (given by the ratio between the irreversible outflow of degradation of biological matter on the one hand and the biomass on the other) decreases. The material flow through the ecosystem is due to the photosynthetic biomass production and the outflow due to its degradation through herbivores, carnivores, and detrivores. If the ecosystem has a sufficiently macroscopic dimension so that the system may be homogeneous in biomass distribution, one can theoretically deduce Margalef's principle, which states that the ecosystem evolves in the direction along which the ratio of the *photosynthetic* biomass production rate per unit time in a unit area to the *total* biomass present in the same area *decreases*.[53] This is strictly connected with Lotka's intuition [Sec. 9.9] and raises a very important problem in the distribution of energy. Considering again Fig. 10.9(b), we may distinguish different levels of the ecosystem, starting from a primary level, represented by the heterotrophic plants and going up to the highest predators. Considering any food-cycle level n, we have a flow of energy both entering that level and leaving it.[54] The rate of change of the energy content of Λ_n at this level is represented by

$$\frac{d\Lambda_n}{dt} = \lambda_n^i + \lambda_n^e, \tag{10.7}$$

where λ_n^i is the positive part and represents the energy contribution of the previous levels $n-1$ to n as well as n's own production (if any), while λ_n^e is the negative part and represents energy that is either dissipated at that level n or handed on to the following level $n+1$. The efficiency in the productivity of any level n relative to the previous level $n-1$ is given by $\lambda_n^i/\lambda_{n-1}^i$. Considering $n=0$ as the level of photosynthesis, we necessarily have $\lambda_0^i > \lambda_1^i > \ldots > \lambda_n^i$. This is the so-called Eltonian pyramid.

10.3.3 Summing Up

In ontogeny, the relations between the environment, phenotype, and genotype can be considered as reversed relative to the phylogenetic circle [Fig. 10.10]. In phylogeny, the source of the variations is

[52][MATSUNO 1978]. [53][*MARGALEF* 1968]. [54][LINDEMAN 1942].

the genotype with its random mutation, upon which the environment acts with an indirect selective action, and here the environment, at least in a certain time window, can be thought to be stable [Fig. 9.13]. On the contrary, in ontogeny it is the environment that changes, while the genotype can be taken to be stable (in the short time window of the behavioral action of the organism). Obviously in this second case only a relatively small portion of the environment is considered. In both cases, the phenotype represents the unit of accommodation and assimilation [Sec. 8.4].

I would like to add here a few words on the relation between teleological and teleonomic causality. Teleonomy [Subsec 8.2.1] is the ability of organisms to reach stable states by integrating information and different types of feedbacks from different sources (from the self and nonself). Teleology makes use of causal processes to exercise information control on the nonself [Subsec. 8.2.7].

Both processes play an important role in phylogeny and ontogeny. However, the essence of phylogeny is not in teleology but in teleonomy (where different stable states are possible as the result of an evolutionary process: multifinality). The issue here is the accommodation of the organism to the environment. In other words, even if organisms are able to influence their own evolution through niche construction and in this way contribute to the Baldwin effect, they *cannot control* their own evolution. They would only do this if they could exercise perfect, full control of their environment, other species included. Such perfect control is possible only when the environment is completely assimilated to a given species. In this sense, however, it becomes a dead environment, and the organism will lose its plasticity by becoming a form of machine.

Instead, the essence of ontogeny is in teleology, especially if we consider maturity (which is the core of ontogeny, at least in the case of multicellular organisms). Indeed, development comprehends epigenetic processes which we shall deal with in the next chapter. The most specific trait of mature ontogeny is information control on the environment, and therefore the ontogenetic canalization of environmental cues, the building of an environmental niche, and the assimilation of the environment to the organism. This assimilation is basically ruled by the metabolism, and in its most basic modality of action consists in feeding. However, when organisms become more and more sophisticated during evolutionary time, this assimilation takes other forms and consists in general in the transformation of the environment according to the needs of the organism. In the case of humans, it amounts to true environmental shaping and reshaping.

10.4 Concluding Remarks

In this chapter I have introduced the idea of the ontogenetic path of the individual organism. This trajectory can be considered as determined by an irreversible information transformation (selection). Moreover,

- During the ontogenetic path we have increasing catabolic reactions and decreasing anabolic reactions.
- It is very important to distinguish between species-specific and individual trajectories.
- The whole ontogenetic path is articulated in development, maturity, and aging.
- During development we have high channelization with the environment but low and growing canalization in a certain ontogenetic path.
- During maturity, we have a relatively stable equilibrium between channelization and canalization. During this stage we have the maximal intervention of the organism on its environment (through niche construction).

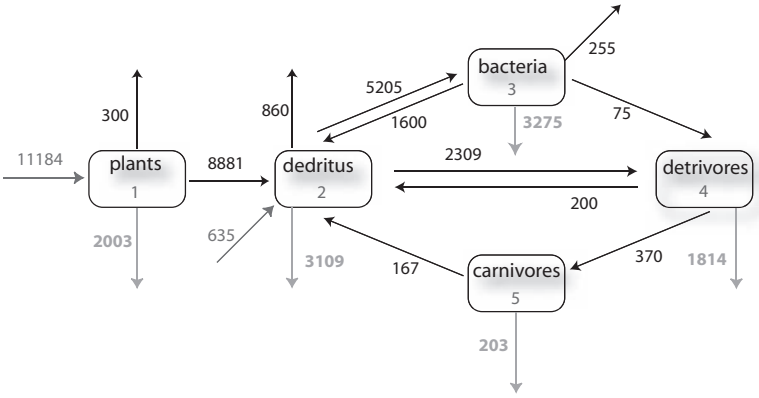

Fig. 10.11 The so-called Cone Spring ecosystem. Flows are measured as a function of kcal. 1–2–3–4–5 the easiest sequence. In dark gray we have the energy inflows and in light gray the energy spent for respiration. Lengths of arrows are not related to the quantities exchanged.

- During aging, the organism is easily disrupted by small environmental fluctuations: It is a stage characterized by low plasticity as well as by high channelization and low canalization.
- The organism is not passive relative to the environment but is able to give rise to a species-specific environmental niche with the effect of partly and indirectly canalizing its own evolution or at least smoothing the effects of natural selection.
- Several individuals and species interact in building whole ecosystems that show intrinsic irreversibility.
- In this way, the genetic system, giving rise to a metabolic activity which has the environmental niche as one of its effects, is able to act indirectly on the environment by assimilating it to the organism.

Appendix: Networks

Let us show, in a few words and by using an example, the way in which energetic fluxes are calculated in an ecological network. In the network shown in Fig. 10.11 there are five nodes and eight internal transfers. We may represent the transfer of energy from compartment i to compartment j with the element T^E_{ij} of a 5×5 matrix[55] [Subsecs. 1.2.3–1.2.4]

$$\hat{T}^E = \begin{bmatrix} 0 & 8,881 & 0 & 0 & 0 \\ 0 & 0 & 5,205 & 2,309 & 0 \\ 0 & 1,600 & 0 & 75 & 0 \\ 0 & 200 & 0 & 0 & 370 \\ 0 & 167 & 0 & 0 & 0 \end{bmatrix}, \tag{10.8}$$

or also the transfer of energy to component j from component i with the element $\hat{T}^I_{ij}$ of the matrix

[55] [*ULANOWICZ* 1986, pp. 30–52].

$$\hat{T}^I = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 8,881 & 0 & 1,600 & 200 & 167 \\ 0 & 5,205 & 0 & 0 & 0 \\ 0 & 2,309 & 75 & 0 & 0 \\ 0 & 0 & 0 & 370 & 0 \end{bmatrix}. \tag{10.9}$$

Both $\hat{T}^E$ and $\hat{T}^I$ are matrices representing so-called directed graphs, i.e. networks with directional arrows.[56] Moreover, we represent with column vectors the environmental inputs to the compartment i (each given by the line I_i), its environmental exports (each given by the line E_i), and its respiration for metabolic needs (each given by the line R_i) as

$$|I\rangle = \begin{pmatrix} 11,184 \\ 635 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |E\rangle = \begin{pmatrix} 300 \\ 860 \\ 255 \\ 0 \\ 0 \end{pmatrix}, \quad |R\rangle = \begin{pmatrix} 2,003 \\ 3,109 \\ 3,275 \\ 1,814 \\ 203 \end{pmatrix}. \tag{10.10}$$

If the system is in a steady state, for each node j we have

$$T_j^I = T_j^E, \tag{10.11}$$

where

$$T_j^I = I_j + \sum_{i=1}^{n} T_{ij} \quad \text{and} \quad T_j^E = \sum_{k=1}^{n} T_{jk} + E_j + R_j \tag{10.12}$$

are the input flow to and the output flow from the unit j, respectively. This is the case for the example chosen. For instance, for the unit 2 we have the balanced input and output fluxes

$$\left.\begin{array}{rcl} T_2^I &=& I_2 + T_{12} + T_{22} + T_{32} + T_{42} + T_{52} \\ &=& 635 + 8,881 + 0 + 1,600 + 200 + 167 \\ T_2^E &=& T_{21} + T_{22} + T_{23} + T_{24} + T_{25} + E_2 + R_2 \\ &=& 0 + 0 + 5,205 + 2,309 + 0 + 860 + 3,109 \end{array}\right\} = 11,483. \tag{10.13}$$

[56] [*BARRAT et al.* 2008, p. 3].

11
Epigeny

The discipline studying development, especially in its first stages, is known today as epigenetics or epigenesis. I refer to the process itself that is the object of such disciplines as *epigeny*, which seems to be on the same footing as *phylogeny* and *ontogeny*. I recall that the basis of epigeny is the organism's self-production [Sec. 8.4]. After an examination of the *status quaestionis*, I shall deal with the issue of the general nature and significance of epigeny. Then, we shall consider some fundamental mechanisms operating during the epigenetic process. The core of this chapter is represented by analysis of the steps of epigeny (especially embryogenesis and organogenesis). A specific but very relevant aspect concerns the way in which the brain develops and how a functional organization of the brain is superposed on a developmentally earlier one having a simple structural character. Another important issue is the significance of the Waddington effect, one of the first phenomena to have attracted the attention of scholars to this field. Finally, the relations between phylogeny, ontogeny, and epigeny are framed in a wider synthesis.

11.1 A Big Change

11.1.1 The Problem

Epigeny has its roots in the borderline between ontogeny and phylogeny, in a way representing the confluence of them, the point at which we have interactive relationships between the three systems—environment, phenotype and genotype—so that a true cybernetic circle is constituted [Fig. 8.10]. The great discovery of Waddington was that during epigeny one can force an organism to inherit certain acquired characteristics. Waddington took another path with respect to the traditional genetic explanations, as he focused on branching events in development, that is, on binary choices that determine the following developmental course, instead of a linear determination of the phenotype through the genome.[1]

In this way epigeny ensures that the general principles for building an organism (which have their source in the coding genes but are actually deployed through the network of regulatory genes) are enacted in a specific and individual organism (able to survive, i.e. to be operative in a certain context) *through interaction and cooperation with environmental stimuli* that are unrepeatable by definition [Subsecs. 8.2.1, 8.2.7, and 10.2.4]. This individuality of stimuli is also the basis of the higher cognitive functions. Even if all traits are the result of the correlation between genes and environment (i.e. indirectly and mediated by the phenotype), a trait could still be said to be genetic if a genetic difference is responsible for its variability in a given population.[2]

[1] [*WILKINS* 2002, pp. 105–8]. [2] [GIFFORD 1990].

11.1.2 A Little Bit of History

In his book *On the Generation of Animals* Aristotle understood that there is a certain progression in the building of the organism and that "all the parts are first marked out in their outlines and acquire later on their colour and softness or hardness, exactly as if Nature were a painter producing a work of art, for painters, too, first sketch in the animal with lines and only after that put in the colors" (Book II, Ch. 2). A lesson further worked on by K. von Baer across the first half of the 19th century and easily forgotten in the 20th century, in which a genetic-centric biology has hindered a correct evaluation of epigeny for many decades.[3] Weismann's dictum that epigenetic development is impossible[4] is well known although he understood that development implies a certain differentiation due to the effects of the various somatic cells on genetic material.[5] The reason is that he thought that epigeny would imply a material subdivision of the so-called germ plasm[6] or a material retroaction of somatic cells on germ cells, the so-called pangenesis [Sec. 7.1 and Subsec. 9.2.2]. In other words, Weismann was not able to conceive this process as a transmission of *information* that can determine effects even if both somatic and germ cells remain blind to each other [Subsec. 2.2.3]. This exchange of information is necessary for the fundamental self-regulatory (cybernetic) mechanisms of the organism[7] [Ch. 8]. This is very excusable for that time (in which no information theory existed). The underestimation of epigeny was, and is, much less excusable in some of his modern followers. The irony here is that a genetic determinism such as that molding the neo–Darwinian synthesis has been proclamed the sole possible scientific approach in a time when such a determinism no longer existed in physics [Ch. 1]. The main point here is that, due to the double blindness between genotype and phenotype [Subsec. 8.2.1], the role of genes in building an organism is rather an *indirect* one: Although the DNA (i.e., the starting information) is indeed codified, life as a transmission phenomenon is not a coded program in itself [Sec. 9.8]. This means that the central dogma of molecular biology is still valid,[8] even if Weismann's barrier, implying a strict inability of the phenotype to act on the genome, is not.

One of the first scholars to have pointed out the relevance of epigenetic aspects for evolution was Richard Benedict Goldschmidt.[9] An important step was Jacob and Monod's discovery of the first regulatory genes.[10] This strand of research started when Monod became aware that the bacterium *Escherichia coli* is able to discriminate between glucose and lactose: Given a mixture of the two sugars, the bacterium first consumes all glucose and then digests lactose.[11] When glucose is available, a membrane-associated protein involved in transporting glucose into the cell also phosphorylates this sugar molecule. The transport protein itself is most of the time in an unphosphorylated form, but during this process it becomes phosphorylated. When the sugar is consumed, glucose is no longer available as an acceptor of the phosphate group; therefore, the protein remains phosphorylated and becomes able to convert ATP [Subsec. 7.3.2] into cAMP (cyclic adenosine monophosphate), thus raising the cellular concentration of cAMP. The cell uses the phosphorylated transport protein and a high cAMP concentration as signs (a typical semiotic activity [Sec. 8.1]) indicating that glucose is no longer available. Namely, the cAMP concentration is read by the cAMP receptor protein (CRP), which binds to the CRP site in *lac* only in the presence of abundant cAMP [Fig. 11.1]. This then stabilizes the contact between *lacP* (the –10 and –35 regions of the canonical 70 promoters shown in Fig. 11.1) and RNA polymerase, and so signals that the *lac* operon (*lacO*) is ready for transcription. In the absence of lactose, however, transcription

[3][*MINELLI* 2003, pp. 21–42] [GOODWIN 2000]. [4][*WEISMANN* 1893, pp. xiii–xiv].
[5][*WEISMANN* 1893, pp. 32 and 68]. [6][*WEISMANN* 1893, pp. 2–5]. [7][*PIAGET* 1967, pp. 114–18].
[8][CRICK 1970]. [9][*GOLDSCHMIDT* 1940]. [10][JACOB/MONOD 1961]. [11]See also [SHAPIRO 2002].

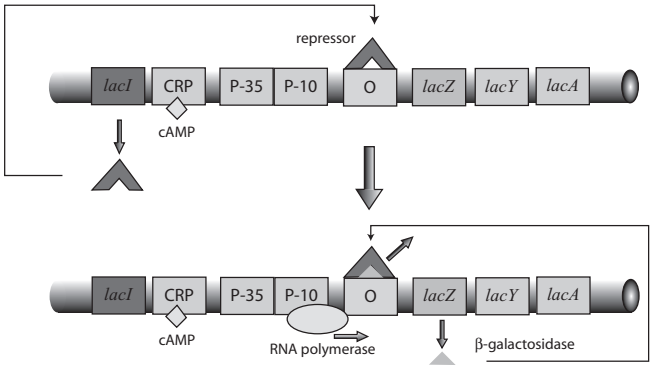

Fig. 11.1 The genetic part of Jacob and Monod's model of *lac* operon. O (O1 and O2) represents the operator sequences (*lacO*), binding sites for dimers of the *lacI* repressor. P-10 and P-35 are the –10 and –35 regions of the promoter *lacP*. The protein *β*-galactosidase coded by *lacZ* allows for blocking the repressor and therefore starting expression. (This figure is reproduced in color in the color plate section.)

events are rare, since the *lacI* repressor molecules bind to two of the operator sites (O1 and O2) and create a loop in the DNA, blocking access to the promoter *lacP*. The cell is able, however, to sense the presence of lactose (a manifestation of information control on its environment). Then, the protein *β*-galactosidase, coded by *lacZ*, converts some of these sugar molecules to a related sugar called allolactose, which can bind to the *lacI* repressor, inducing a change in the shape of the repressor that makes it unable to bind to *lacO* and so freeing *lacP* for transcription [Fig. 11.2]. Here, the repressor *lacI* is the prototype of allosteric proteins [Subsecs. 7.4.4 and 7.6.1], which exist under different conformations.[12]

Summing up, Jacob and Monod discovered that in the absence of lactose, the repressor gene codes for a protein that binds to the promoter of the gene coding for the enzyme that is able to digest lactose, thus preventing its transcription. When the *E. coli* is in a solution of lactose, this substance is allowed to enter the cell and to bind the repressor proteins inhibiting the transcription of the genes ecoding the lactose metabolizing enzyme. This shows that a molecule activates an expression that is necessary for its own metabolization, a beautiful example of feedback circuits. Note that the whole mechanism, apart from the information control on some environmental parameters, relies on teleonomic causality.

Another interesting example is when virus proliferation is kept under control by bacteria. The case of the bacteriophage λ has been studied.[13] After a certain time, a kind of alternative emerges: Either the bacteriophage pursues its reproduction indefinitely (this response is called lytic), and in this case the host cell finally dies, or virus reproduction is kept under control trough production of the λ repressor which is expressed by promoters p_{RE} and p_{RM} (a response that is called lysogenic).

This research, though fundamental for our present understanding of epigeny, unfortunately was not connected with evolutionary biology at the time, and its relevance for epigeny was not properly understood since bacteria themselves do not show interesting developmental processes.[14] After these studies, two of the first scientists to have stressed the centrality of epigeny for evolution

[12][MONOD *et al.* 1963] [MORANGE 2002]. [13][ARBER 1983].
[14]Actually Jacob understood the relevance of the model for development [GILBERT 1996].

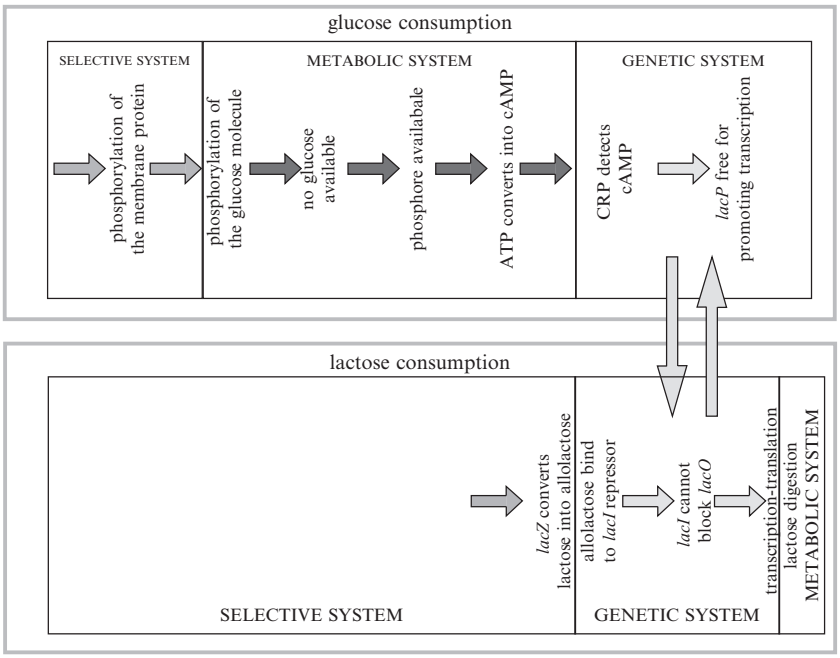

Fig. 11.2 Jacob and Monod's cybernetic model of *lac* operon.

are Ho and Saunders[15] (leaving aside Waddington, whose work we shall deal with extensively below). Contrary to the neo-Darwinian view, they pointed out that the variations of the phenotype, upon which natural selection could act, do not arise totally at random; they are produced mostly by interactions between the organism and the environment *during development*. They proposed, therefore, that the intrinsic dynamic structure of epigeny itself, due to interaction with the environment, is the source of non–random phenotypic variations which direct evolutionary change, and that a proper study of evolution would consist in working out the dynamics of the epigenetic processes and their response to environmental stimuli as well as the mechanisms whereby novel developmental responses are canalized.

Today, there are many studies that support this point of view. Let me give a specific and very useful example. A recent study[16] found that, although twins are epigenetically indistinguishable during the early years of life, older monozygous twins exhibited remarkable differences in their overall content and genomic distribution of 5-methylcytosine DNA (formed by methylation of cytosine) and histone acetylation (addiction of an acetyl functional group), affecting their gene-expression portrait: The methylation marks inhibit the genetic activity by making DNA coding sequences inaccessible.[17] Such a process entails the addition of a CH_3 group to a substrate, represented here by a nucleotide basis (indeed, thymine may be derived by methylation of uracil at the 5′ carbon [Fig. 7.5]). Instead, histone acetylation of the lysine residues at the N terminus of histone proteins removes positive charges, thereby reducing the affinity between histones and DNA,

[15] [HO/SAUNDERS 1979]. See also [GOODWIN 1982].
[16] [FRAGA *et al.* 2005].
[17] [*GILBERT/EPEL* 2009, pp. 38–46].

thus making it easier for the RNA polymerase and transcription factors to access the promoter region and enhancing transcription (while histone deacetylation represses transcription). These findings confirm how an appreciation of epigenetics is missing from our understanding of how different phenotypes can be originated from the same genotype. It is also important to stress that in some instances these epigenetic changes can be inherited[18] [Sec. 9.8].

11.1.3 Some Problems Today

The Genome Project[19] has produced very relevant results but has also shown the necessity of taking a further step in particular towards a sort of proteinomics, which would help to lead us to a modern understanding of epigeny. It must, however, be stressed that this domain of investigation is still obscure and too little is known even today about epigenetic mechanisms. What is certain is that a large part in this process is represented by the noncoding sequences of DNA (junk DNA), which have complex back-actions on the coding DNA[20] [Sec. 9.7]. Another feature, as we have seen, is represented by the methylation of histones (in chromatin), when lack of methylation on the promoter and enhancer regions of the gene is mostly connected with transcription.[21] Recently it has been shown that impaired functions in memory and learning can be recovered by increased environmental stimulation that in turn determines chromatin modifications, i.e. increased histone-tail acetylation.[22] Finally, there is the action of RNA polymerase on the promoter regions. In all these cases, we have no codification. Regulative activities cannot be codified since they consist in the connection between previously independent systems or processes, especially between protein and RNA displaying specific functions and the codifying genetic system [Subsec. 7.4.5]. It is interesting to remark that modifications to histone proteins may also not happen locally,[23] while other aspects of the epigenetic process are global.

11.2 The Nature and Significance of Epigeny

11.2.1 The Transition from Unicellular to Multicellular Organisms

As in many fields of life, for epigeny there are a lot of exceptions and specific trends for any individuated regularity.[24] However, any true advancement in science necessarily requires the finding of these regularities [see the Introduction to the book]. I think that we are justified in this operation if most of the phenomena in the field are captured by regularities, so that we can consider the remnant in terms of intermediary forms between regular ones or as the necessary fluctuations that accompany any regularity in nature[25] [Subsecs. 6.5.1 and 8.2.7]. However, we must nevertheless avoid the opposite danger of unconditional generalization.

Morphological traits cannot be transmitted in all generality through non-genetic means. This is evident in the case of animals.[26] Animals go through a unicellular stage that cannot preserve the memory of a full adult. Animals are gametogamic and not gamontogamic (as ciliates are). This means that they are usually[27] the result of the fusion and morphological modification of two gametes, rather than the exchange of gametic nuclei between two conjugants without disruption

[18] [*JABLONKA/LAMB* 1995, pp. 133–57]. [19] [IHGSC 2001]. [20] [BIÉMONT/VIEIRA 2006].
[21] [*GILBERT* 2006, pp. 116–19] [BECKER 2006]. [22] [FISCHER *et al.* 2007].
[23] [VOGELAUER *et al.* 2000] [BERGER 2000]. [24] A supporter of this view is especially Minelli [*MINELLI* 2003].
[25] [*PEIRCE CP*, 1.158–62]. [26] [*MINELLI* 2003, p. 29].
[27] There are also parthenogenetic species, such as aphids and *Daphnia*, which do not have males. The egg has all the genetic information.

of the cellular integrity—this does not mean that we should consider life as originating in a self-assembly process, which is not true even in the case of a simple cell: Indeed, organisms are not only self-organizing systems but complex systems [Sec. 6.3] based on whole cybernetic circuits [Sec. 8.4]. As a matter of fact, cells only form through growth and division of preexisting cells.

It is likely that there was a primitive phase of multicellular organisms during which they were much more susceptible to environmental fluctuations (they had less information control on the environment [Sec. 9.10]) and in which the early morphogenetic features were the result of immediate physical and chemical properties of the cells.[28] In an epoch preceding the biochemical canalization of developmental pathways and the stabilization of phenotypes, the interaction of multicellular clusters with their physical-chemical environments dictated a many-to-many mapping between genome and morphology. At a very primitive stage, these forms would have been generated by rudimentary epigenetic mechanisms: Initially, physical processes among chemically active materials would predominate; later on, conditional, inductive interactions among the organism's constituent tissues would become more and more dominant. Let us consider this transition a little bit closer [Subsec. 9.5.2]:

- The most ancient multicellular forms must have been simple cell aggregates that arose by adhesion of originally free-living cells, or by the failure of the same to separate after mitosis [Subsec. 7.5.2]. Today, we have the example of the amoeba *Dictyostelium* that alternates a unicellular with a multicellular stage. The specific chemical or physical nature of the adhesive interaction would have been unimportant, as long as it served to keep the organism's cells from dispersing. Indeed, the advent of a cell–cell adhesion mechanism early in the history of multicellular life, although certainly dependent on the preexistence of particular gene products, did not require additional gene sequence change as such. This is a general rule, as we shall see: What can be done downstream and with more economic tools is generally preferred.
- If the first organisms showing a true epigenetic mechanism were Mendelian, in the sense that genotype and species-specific phenotype are inherited in some close correlation, and their morphological changes were therefore correlated with genetic change [Sec. 9.8], the polymorphic metazoan ancestors postulated here would have constituted a pre-Mendelian world of organisms, whose genotypes and morphological phenotypes were connected only in a very loose fashion and therefore showed greater variety than today. In this exploratory period of organismal evolution, the mapping of a given genotype to a morphological species-specific phenotype would have been one-to-many, rather than many-to-many, as in the previous pre-epigenetic stage, or one-to-one (always from a species-specific point of view), as is the case today. The transition between this second stage and the current one would have been a true convergence process [Subsec. 9.5.3], helped by the action of natural selection, toward the phenotypic forms that would result in being much more stable, starting from a certain range of initial conditions.

Let us now consider what happened in this second stage. Once one or several adhesive mechanisms were at play, other more complex morphological consequences could have followed, simply by virtue of variations in cell adhesivity brought about by random processes like metabolic noise, and by means of the action of the relevant physical laws on such heterogeneous cell aggregates.[29] Cells with different amounts of adhesion molecules on their surfaces, for example, tend to sort out into islands of more cohesive cells within lakes composed of their less cohesive neighbors. Eventually, by random cell movement, the islands coalesce and an interface is established, across which cells will not intermix and multilayered structures can form [Fig. 11.3.A–B]. Two of the five

[28][NEWMAN/MÜLLER 2000] [*BONNER* 2000]. [29][*SPENCER* 1860–62, pp. 228 and 377–8].

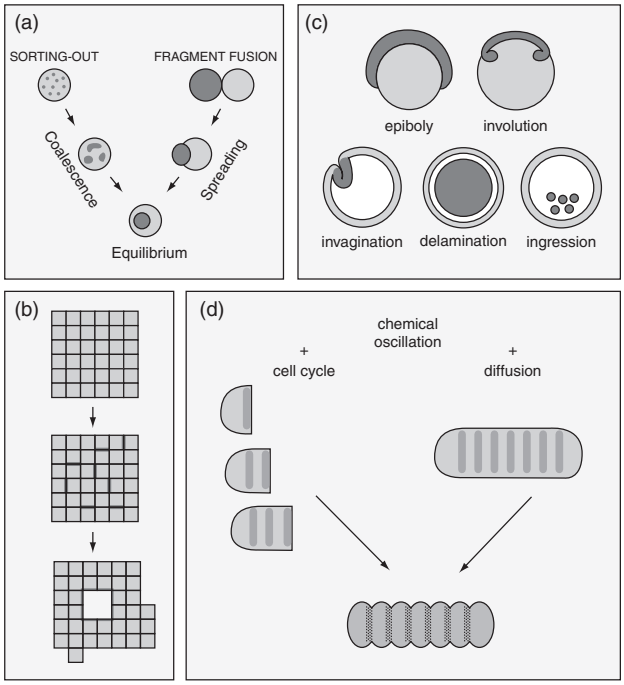

Fig. 11.3 Newmann and Müller's model to account for the beginning of epigenetic processes. Adapted from [NEWMAN/MÜLLER 2000].

major types of gastrulation seen in modern metazoans, i.e. epiboly and involution (and possibly a third, delamination), could have originated as simple consequences of differential adhesion [Fig. 11.3.C]. Furthermore, if variations in metabolic or biosynthetic activity, rather than being purely random across the tissue mass, affected cell–cell adhesion in a temporally or spatially periodic fashion, then compartmentalization takes the form of segmentation [Fig. 11.3.D]. This periodicity is quite normal for complex systems and, given appropriate physical and chemical conditions, would have arisen very naturally [Subsecs. 6.5.1–6.5.2], provided that this bottom-up explanation would be integrated by a parallel top-down one, which would be necessary in order to integrate processes and avoid an excess of differentiation [Sec. 6.3], which, in turn, could have been the natural consequence of those spontaneous physical and chemical processes.

- Once major body plans were established, selection for biochemical integration, which promoted physiological homeostasis and developmental reliability, stabilized the relationship between the genotype and the ecological setting referred to as fitness or adaptedness. Homology, the principle of morphological organization, is a consequence of the interplay between genetic, morphogenetic templates and evolving, stabilizing biochemical circuitry under the control of the growing organism. Fixed at the body-plan level, with their molecular and developmental bases free to drift, homologues persevere and become attractors of morphological design.

It is important to stress that, in modern organisms, there is a one-to-one correspondence between genome and *species-specific* phenotype, but modern organisms also show certain one-to-many mappings, especially considering their *individual* development, so relevant in mammals [Subsec. 10.2.4].

I recall that Gupta and Lewontin[30] have indeed developed the idea of the norm of reaction, that is, of the array of individual phenotypes that will be developed by the genotype over an array of environments [Subsec. 10.3.1]. Therefore, my guess is that we may distinguish three evolutionary stages of epigeny (apart from the initially pre-epigenetic step):

- A pre-Mendelian stage, in which phenotypes were weakly connected with genotype (a one-to-many species-specific mapping).
- A Mendelian stage, in which there was a one-to-one correspondence between genotype and grossly individual phenotype. This is still the main mechanism for lower forms of life in which driving-force processes predominate (which I shall introduce in a few pages). These are mechanical and almost pure feedforward (bottom-up) processes, and therefore essentially teleonomic. The fact that epigenetic mechanisms are the generative agents of morphological character origination helps to explain findings that are difficult to frame in the standard neo-Darwinian model, e.g., the burst of body plans in the early Cambrian [Subsec. 9.5.2], the origins of morphological innovation, homology, and rapid change of form during development.
- A post-Mendelian stage, in which there is a one-to-many mapping from the genome to individual phenotypes and *regulatory* processes dominate, as it happens at least for mammals. These are mechanisms in which feedback and regulation are much stronger, and teleological causal processes exercising information control by the organism are much more relevant, even if it is already present at a bacterial level.

11.2.2 The Significance of Epigeny

The evolutionary significance of development is related to the fact that young organisms exhibiting structural and behavioral development survive and reproduce better than other organisms whose behavior changes less considerably with age and with size[31] [Secs. 10.1–10.2]. For this reason, epigeny is essentially a compromise between the continuity of the species and the discontinuity of evolutionary or epigenetic novelties. This can bring us to a deeper understanding of epigeny:

- For the reasons indicated previously, epigeny can indeed be understood as a compromise between cell proliferation (the continuity of self-reproduction and self–production) and the necessity to control cell division (discontinuity, consisting in negative feedback against growth), especially of some specific cells. For instance, early metazoans blocked proliferation by differentiating a primitive blastula in a ball of ciliated cells.[32] The only cells that could still differentiate were those already in or migrating into the inner cavity of the ball, a sort of gastrulation. In this way, cilia also acquire a developmental role: They help in maintaining the animal shape. This compromise obviously has an evolutionary significance[33]—compare this with the behavior of viruses [Sec. 8.4] or even bacteria [Subsec. 7.5.2].
- Epigeny can also be understood as a compromise between environmental inputs (the discontinuous aspect here) and control genes providing anti-feedback (the continuous aspect, from the point of view of the organism). Environmental factors like temperature, nutrition, pressure and gravity, light, presence of predators, and the presence or absence of conspecifics are very important for the development of an organism.[34] The action of the environment on the organism can affect transcriptional regulation, the neuroendocrine system, or involve direct cellular induction[35]: They correspond to an action on the genetic, metabolic, and selection system, respectively. The presence of these factors explains why epigeny is not a full teleologic causal process but also

[30][GUPTA/LEWONTIN 1982]. [31][*FAGEN* 1981]. [32][*BUSS* 1987]. [33][*MINELLI* 2003, pp. 12–13].
[34][*GILBERT/EPEL* 2009, pp. 13–32]. [35][*GILBERT/EPEL* 2009, p. 38].

consists of considerable teleonomic processes. However, this does not imply that the environment somehow guides or instructs the organism: The environment only provides negative feedback, and it is only thanks to those teleonomic processes that the organism is able to canalize and even make those environmental stimuli positive for its own growth [Subsec. 8.2.1].
- Continuity and discontinuity are also important from a further perspective. It is perhaps convenient to distinguish between initial information (the continuity across the generation) and information conditioned by the specific metabolism of an individual (the discontinuity). Evidence for this is shown by the fact that differences in position and surroundings in the presence of a faithful genetic duplication that results in differences in phenotypic patterns are regulatory differences that are responsible for divergences in developmental processes [Subsec. 10.2.5].[36]

It is crucial to understand that epigeny is based on:

1. Cellular memory: Initially, memory consists almost completely of genetic information; as epigeny goes on, epigenetic memory progressively grows. It is, therefore, a parallel process and, to a certain extent, also a cyclic (wave-like) one.
2. Cellular machinery, which in turn also depends on cellular memory.[37] It is also true that every cell starts its own version of life anew, since its configuration depends on (both temporally and spatially) local context much more than on the genetic information it brings, and this represents the point-like, discontinuous aspect: Each generation event and each environmental input are discontinuous from the point of view of the epigenetic control,[38] while the developmental program is the continuous aspect.

The ability of a cell to detect and react to a specific location thanks both to its memory and its current state is called topobiological *potency* (another form of potential information). For this reason, development cannot be understood as a mere sum of cellular behaviors but is a very complex feedback network in which sophisticated regulatory processes are at play.

11.2.3 Convergence and Divergence

Phylogenetic transmission is a divergent process in its own essence (new species always arise), even if we have seen that convergent aspects also play a role [Sec. 9.5]. Ontogeny, understood as the whole life trajectory, is basically a convergent process [Sec. 10.2]. Epigeny, in its own nature, is a convergent–divergent increase of complexity[39] [Fig. 11.4]. The different genetic and epigenetic switches can be understood as true logical operations ruled, for instance, by *if . . . then* or AND operators.[40]

During epigeny there are many possible paths that can lead (converge) to the same species-specific result, a behavior called *epigenetic degeneracy*[41] [Subsec. 8.2.5]. Degeneracy is also a basic property not only of the genetic code but also of the brain and the mind, especially in their higher functions.[42] Recall that this feature, at the level of system theory, is called by von Bertalanffy the "principle of equifinality"[43] [Subsec. 6.3.3]. Epigenetic degeneracy means that the initial conditions given by the genetic memory of the system, its initial state and a given environment, do not suffice for singling out the developmental path of an organism [Subsec. 11.1.1]. However, I recall that, as

[36][*WEST-EBERHARD* 2003, pp. 209–10]. [37][*MINELLI* 2003, pp. 3–4].
[38]The otherwise interesting study [*GORDON* 1999] seems to overlook the issue of control.
[39][*ARTHUR* 1997, pp. 123–25]. [40][*DAVIDSON* 2001, pp. 56–61]. [41][*MINELLI* 2003, pp. 231–2].
[42][*EDELMAN/TONONI* 2000, pp. 86–87]. See also [LAUGHLIN *et al.* 2000].
[43][*VON BERTALANFFY* 1969b]. See also [PEIRCE 1902].

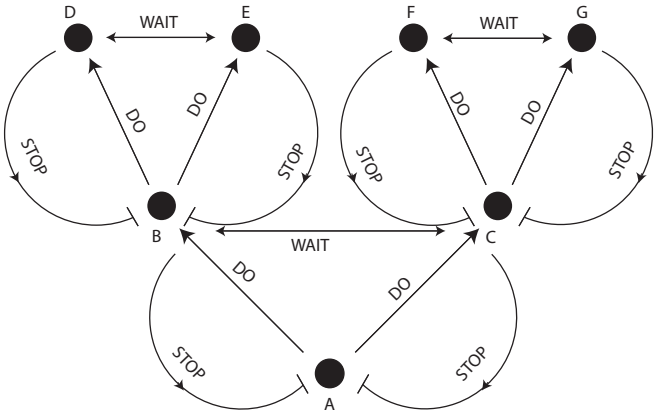

Fig. 11.4 A very schematic drawing of the main typologies of epigenetic interactions: The gene A is able to activate genes B and C, which in turn respectively activate genes D and E, on the one hand, and F and G, on the other. Any downwards gene negatively feedbacks on upstream genes. Genes located at the same level of the activation cascade influence reciprocally with waiting signals for regulation and fine-tuning of the process. Interestingly, genes E and F may be activated synchronically (and in this way share information); even their parent genes B and C are not causally sequential [Subsec. 2.2.3]. Inspired by [*ARTHUR* 1997, p. 124].

the organism approaches its mature form, it will become increasingly difficult to change course, and the danger that any significant change may turn out to be disruptive to development becomes greater [Subsecs. 10.2.1–10.2.3].

Obviously, the opposite is also true, that is, a single gene, through epigenetic mechanisms, can give rise to a set of *different ontogenies* when it is exposed to internal or external environment, i.e. the developmental reaction norm [Sec. 9.10, Subsecs. 10.3.1 and 11.2.1]. However, this does not apply solely to single genes.[44] Wright was the first scholar to understand that there are effects arising thanks to circuits that are actually constituted by a network of interacting genes (a phenomenon called epistasis) [Fig. 11.5].[45] Therefore, gene complexes must have coadapted in the course of evolution. In this way, as I have already pointed out, evolution does not present a single peak in the fitness landscape but multiple peaks of various heights [Subsec. 9.5.5].

Summing up,[46] the *interactive* (organism–environment, cell–cell, chemical aspects–genetic regulation) dynamic process is crucial here and the final (mature) species-specific steady state attained at the end of development is an attractor [Fig. 3.24]: We have a dynamical basin of attraction in which all information necessary for joining the attractor is not present from the start as a set of sufficient instructions[47] [Subsec. 8.2.1], but is rather a process of self-organization ruled by the dynamicity principle [Subsec. 8.2.7], through which a complex system emerges out starting from some initial instructions and the attractor itself may be changed during this dynamic process [Sec. 6.3]:

(1) The set of initial instructions allows for the building of the first elements that give rise to a cascade process (positive feedback) through which further genes and signals are activated and propagated.

[44] See also [*SCHLICHTING/PIGLIUCCI* 1998].
[45] [WRIGHT 1931, WRIGHT 1932].
[46] [AULETTA 2010].
[47] [*BARBIERI* 2003].

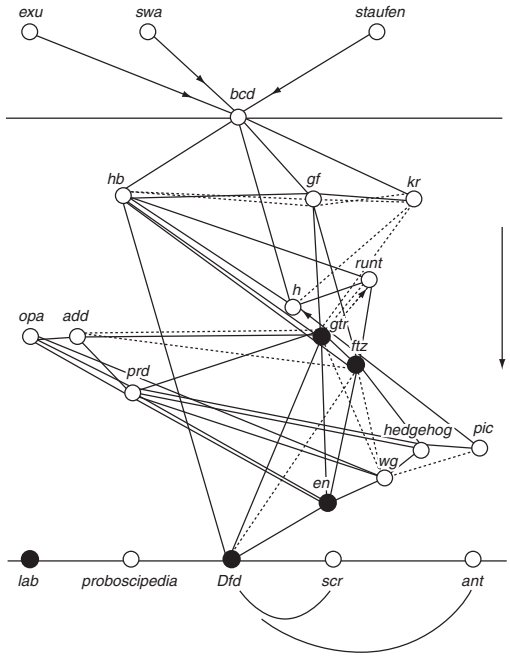

Fig. 11.5 Network of interacting genes that control the expression of the *deformed* (*Dfd*) gene (bottom center) in the development of the *Drosophila* body plan. Regulation proceeds from top to bottom. Dashed lines represent negative regulation, solid ones represent positive regulation. Black circles are possible autoregulatory genes. Many other genes (not shown here) are also activated or repressed by this network. Adapted from [*SCHLICHTING/PIGLIUCCI* 1998, p. 6].

(2) The cellular multiplication process proceeds by a successive and parallel but *hierarchical* building of different levels of commands and containment (body plan, organs, tissues, single cells), where negative-feedback effects are at work. The process is governed by the principle of information accessibility [Subsec. 2.2.2], allowing for different levels of information encapsulation.

(3) During this process, many events happen that have multiple effects establishing new interconnections and therefore a huge network of shared information, both horizontally and vertically [Secs. 6.4–6.5].

(4) The organism is characterized by the fact that it actively searches for the environmental cues (temperature, light, food, and so on) that allow its own development, and it is here that (through anti-feedback) information control and teleologic causation come into play, assisting epigeny in the process (cognition, in the wide sense of the word that I am using in this book, assists epigenetic processes).

(5) The whole can be seen as a process tending to a final stable state through a trajectory where the distance from the final species-specific steady state is minimized through the active concourse of the organism [Secs. 10.1–10.2]. However, since each stable state is provisional, we have an itinerant dynamics [Subsec. 8.2.7].

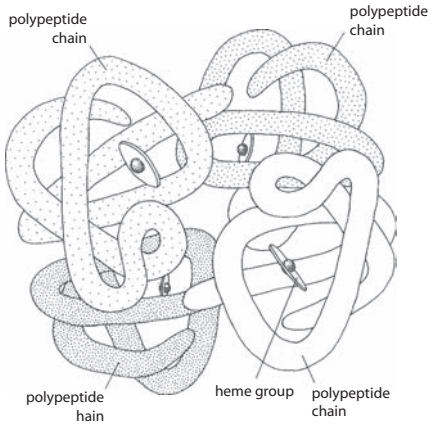

Fig. 11.6 Example of entrenchment: the hemoglobin molecule exemplifies the intimate connection between genetically specified factors (polypeptide chain) and elements of environmental origin (the iron in heme groups). Adapted from [*WEST-EBERHARD* 2003, p. 500].

This is the language of physics that we need here [Sec. 8.1]; a language based on concepts like constraints, degenerate states and processes, information sharing and selecting, anti-feedback, differential timing, and irreversible dissipative events. The problem of the increasing complexity in epigeny can then be reduced to a very specific problem of modularization and integration [Subsec. 2.4.4]: It is this difference in information (between the memory and the current state as well as between the current state and the final state), in an opportune teleonomic network transforming the mechanical inputs in a controlled set of instructions, that needs to be further implemented in an interactive cascade process.

11.2.4 Environments and Developments

I have stressed that epigeny is an interactive construction in which environmental cues play a central role. The process by which the environment supplies materials that become essential for development alongside genetic factors is called *entrenchment*[48] [Fig. 11.6], and is one of the highest manifestations of teleonomy.

Moreover, we must distinguish between the external and internal environments. The *external environment* is the nonself. Here, as mentioned, the organism is sensible to many external cues, like temperature, humidity, sound, gravity, and many others. Sometimes, the role of genes has been overestimated, while experiments in gene control can succeed only in the right ecological environment (channelization), as stressed by Lewontin.[49] Therefore, it is necessary to distinguish between a universal environment and an ecological niche, which plays a direct role in development[50] [Sec. 10.3]. For this reason, when we speak of a genetically determined character, the most we can say is that this character determination is highly influenced by the genotype *in the conditions observed*.[51] The second aspect is the *internal environment*. Here the problem consists of differences in the cells' environments within the embryo.

[48][*WEST-EBERHARD* 2003, pp. 500–3]. [49][*LEWONTIN* 2000, p. 31] [*ROBERT* 2004, p. 7].
[50][*WEST-EBERHARD* 2003, p. 98]. [51][*WEST-EBERHARD* 2003, pp. 101–4 and 135–8].

The relations between epigenetic, genetic, and environmental processes are governed in particular by *epigenetic buffers*, that is, by proteins that are able to smooth the effects of environmental changes and genetic variation on the phenotype, as well as to accumulate different phenotypic variants in a neutral way and syncronize their conversion to a nonneutral state. A typical buffer is represented by the heat-shock protein 90 (Hsp90).[52] Facing a variation in selective pressure, this protein may provide an avenue by which populations can evolve different genotypic states—from those that produce a particular trait to those in which the trait dynamically responds to the environment, and from here to those in which a different developmental endpoint has become a fixed characteristic.

11.2.5 Types of Change

Summing up, we may have very different types of epigenetic interaction and process. For the sake of clarity, we can distinguish between different types of change during development[53]:

1. Change in time (heterochrony): For instance, a development of the forelimbs that takes place very early allows dolphins to develop their flippers.[54]
2. Change in place (heterotopy): It consists in changing the place of a gene's expression. For instance, ducks have an inhibitor gremlin in the webbing of their hindlimbs allowing them the webbed configuration of their feet.
3. Change in amount (heterometry): There is an established correlation between the dimension of beaks in Darwin's finches and the amount of *Bmp4* expression.
4. Change in type (heterotypy): There are changes in regulating proteins. An example is provided by the difference in number of legs between insects and spiders. This is due to the insertion of a polyalanine sequence in the Ubx protein of the former family.
5. Change in control (heterocyberny)[55]: Environmentally induced traits that are integrated into the developing organism. Examples are represented by the Waddington effect and phenocopies, which will be discussed below. I stress that none of these phenomena imply an environmental instruction to the organism. This is evident by the fact that all of these changes effect results from environmental negative feedback (stress) on the organism [Subsec. 8.2.1].

It is evident from this list that epigeny is a compromise between teleonomic and teleologic causal processes, as I have already mentioned [Subsec. 11.2.3].

11.3 Mechanisms of Epigeny

A very appropriate distinction has been introduced by Edelman[56]: In epigeny, cellular division, cellular motion,[57] and cellular death constitute the *driving* force processes, while cellular adhesion, differentiation, and induction are *regulatory* processes. I also add cellular induction and transduction as *informational processes*.

Moreover, there are at least three possible models for the development of a morphology: Reaction diffusion (due to A. Turing), positional information (due to L. Wolpert), mechanical

[52][QUEITSCH *et al.* 2002]. [53][*ARTHUR* 2004, pp. 82–3] [*GILBERT/EPEL* 2009, pp. 342–54, 372–91].
[54][DE BEER 1938] [BRIGANDT 2006]. [55][LALAND *et al.* 2008]. [56][*EDELMAN* 1988].
[57][MURRAY *et al.* 1983, MURRAY/OSTER 1984].

propagation of the configuration (due to Oster and Murray).[58] Turing[59] proposed a reaction diffusion (a pure wave-like) mechanism with at least two chemical species, say A, B. In the absence of diffusion, the two chemicals tend to a linearly stable uniform state; then, under certain conditions, spatially inhomogeneous patterns can evolve by diffusion-driven instability if $D_A \neq D_B$ (D_A and D_B being their respective diffusion rates). Diffusion is usually conceived as a stabilizing process. The explanation of reaction diffusion is not completely adequate because it does not clarify how cells move and adhere, and there is also no connection to genetics. It is a similar case for the model of Oster.[60] We shall consider the reaction diffusion model and in the appendix to this chapter see how it works for explaining coat formation in many mammals. It is a model that explains most of the pattern formation in living organisms.

It is likely that during epigeny all of the three processes mentioned above are used. As a matter of fact, apart from positional information which we shall discuss below, some experimental evidence has also been found for the Turing model.[61]

11.3.1 Cellular Induction and Transduction

Cells influence each other. This process is called *cellular induction* and occurs through cellular signaling. Here we find the general features of signal transduction that I have already pointed out[62] [Secs. 3.3 and 7.6]. Cellular signaling is therefore a universal feature of life that goes beyond neural aspects and even touches a wider domain than that represented by multicellular organisms. There are three main ways cell signaling occurs[63]:

- Direct contact: Here, the cells stick together through some molecules on their surface [Subsec. 11.2.1].
- Gap junction: The signal may pass from cell to cell through relatively small gaps, as in neuron–neuron transmission.
- Diffusion: Here, signals spread and are transported through diffusion mechanisms, as it is usual for endocrine interactions.

We distinguish therefore between direct cell contact (*juxtacrine interactions*) and cell signaling through diffusion of proteins over short or long distances. The latter two forms of signaling are examples of *paracrine interactions*. Note that paracrine factors may produce a new set of paracrine factors in other cells that cause the first group of cells to change. This is called *reciprocal induction*, and it is the foundation of organ formation. Here I shall consider the paracrine mode of interaction.

As we know, no *information* enters the cell from the exterior. Even hormones, in order to be active, must in general interact with intracellular receptors, though there are cases, especially when hormones having lipid structures, in which they can overcome the membrane barrier. That is, first (external) messengers must be "translated" into second (internal) messengers (this is the proper signal transduction) [Figs. 11.7 and 11.8; Subsec. 7.6.2]. The only elements that are normally allowed to enter the cell are units recognized as having negentropic value (necessary for metabolic reasons). Depending on the context, the same signal may give rise to different effects as well as several

[58] See also [EDELMAN 1976, *EDELMAN* 1988]. [59] [TURING 1952].
[60] [ODELL *et al.* 1981, OSTER/ALBERCH 1982, OSTER/MURRAY 1989] [61] [LI *et al.* 2001].
[62] [SIEBENLIST 2001].
[63] [*WOLPERT et al.* 2002, pp. 141–2] [*GILBERT* 2006, pp. 145–69] [*ARTHUR* 1997, pp. 102–20].

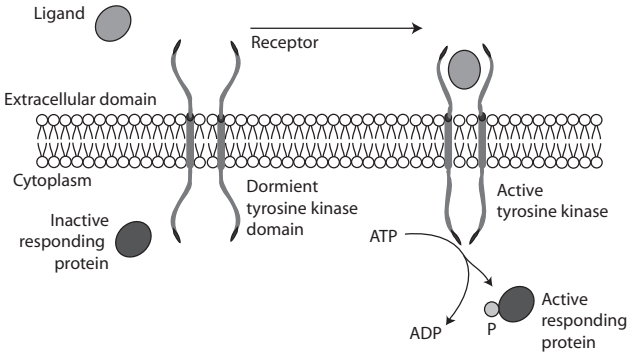

Fig. 11.7 The basic and general mechanism of paracrine induction. The external inducer gives rise to an enzymatic activity. Usually, this is a kinase activity using ATP to phosphorylate (here P represents a phosphorus atom) specific kynase residues of certain proteins. Inspired by [*GILBERT* 2006, p. 147].

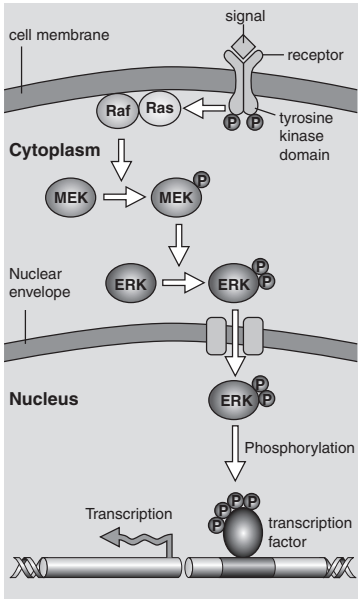

Fig. 11.8 A specific mechanism of signal transduction during epigeny. The signal (a chemical) binds to an appropriate receptor that sets a cascade of protein phosphorylations in motion. First, a Ras protein is activated with the result that a Raf protein binds to it, which in turn results in the phosphorylation (P represents a phosphorus atom) and activation of the protein kinase MEK, which phosphorylates another kinase, ERK, which finally enters the nucleus and activates gene expression. Adapted from [*WOLPERT et al.* 2002, p. 299].

signals being codified in the same manner: The transmission is a function many-to-one (for this reason, we have both an inducer and a responder). Further, the same signal can act on cells at different times or ages causing different responses (because of an intracellular clock). Further, when acting on some cell types, the paracrine factor *Bmp4* causes bone formation, on other cell types it causes sensory nervous differentiation, and on others it causes cell death.[64] Often, the activation of a single gene can produce proteins that have activation and inhibition effects on different portions of the genome, and even on other cells. There are also processes of reciprocal and sequential induction.[65] In this way, the cascade and self-regulation processes—typical of epigeny—arise[66] [Subsec. 11.2.3].

Therefore, the effect of signals is not automatic, it does not depend on the energy or solely on the information that they carry; rather, the response they produce is very selective and is strongly dependent on the current state of the receptive cell, which works, as a whole, as a selector and decider [Subsec. 2.3.2] on the significance of the input signal[67] [Sec. 8.1]. This also has the consequence that the sources of the impinging stimuli—being either the environment or the genome—has little significance, and many sources may be interchangeable to a certain extent.[68] This is also true when one considers general information-theory aspects: Since the output entropy of the signal's sender is in general much less than the input entropy of the receiver (due to entropy growing during any process of selection [Subsecs. 2.3.1, 2.3.3, and 7.4.6]), signals are often not sufficiently diverse to control the communicative behavior of receivers.[69] The set of all final output signals must then be a subset of all possible outputs. From the point of view of the organism's activity, it is impossible to understand the mechanism of cell communication without considering that different stimuli or operations are treated as equivalent [Subsec. 8.2.6].

It is worth stressing here a very important mechanism through which cells and even parts of cells are able to acknowledge each other and to select the proper reaction in a specific interaction context, especially considering that the origin of the signal can have little significance. The general way in which this happens is by individuating or releasing specific signals that can be acknowledged by other cells or parts of cells as the marks that individuate a specific operation or stimulus [Subsec. 8.2.3]. We shall consider the general importance of this fact. Let me add that the mark is an active connection with some context: It is the way in which an organism can dynamically find its path to a suitable situation (or avoid and escape from an unsuitable one).

11.3.2 Cellular Motion and Position

Cells differentiate according to where they are in the spatial organization of the embryo.[70] Although the mechanism should be genetically controlled, the genes themselves cannot create the pattern [Subsec. 6.5.2]. They only provide a blueprint or recipe for the pattern generation. Moreover, as we know, molecular genetics does not explain the origin of functional properties which epigenetically build a phenotype as a living being.

Following the positional information theory of Wolpert,[71] cells are preprogrammed to react to a chemical concentration and differentiate accordingly. For instance, in plant roots different cell types are organized in a well-defined pattern: Each cell knows exactly where it is and what it should do.[72] In such a process we distinguish between three aspects[73]:

[64] I owe this remark to S. Gilbert. [65] [*GILBERT* 2006, pp. 53–67]. [66] [*WOLPERT et al.* 2002, pp. 293–327].
[67] [*OYAMA* 1985, pp. 15–16]. [68] [*WEST-EBERHARD* 2003, pp. 100 and 117–28].
[69] [*HAILMAN* 1977, pp. 21–155]. [70] [*MURRAY* 1989, pp. 372–414]. See also [SAUNDERS 1984].
[71] [WOLPERT 1969, WOLPERT 1971, WOLPERT 1977] [LEWIS *et al.* 1977] [SMITH/WOLPERT 1981].
[72] [HAKE 2001]. [73] [*WOLPERT et al.* 2002, pp. 20–2].

Fig. 11.9 The orientation and vectorial direction of reading does matter in epigeny.

- A chemical, called *morphogen*, whose concentration (gradient) is involved in pattern formation.[74] This is the necessary variability at the source.
- The fact that each cell is regulated for responding to a certain threshold concentration.
- The final selected response. This means that, relative to the same morphogen, different patterns can be developed, so that, as explained before, the final step is not an immediate consequence of the second one. For instance, the same morphogen can be interpreted as a French-flag or a Dutch-flag structure (depending on the orientation and direction from which the pattern is considered) [Fig. 11.9].

The system can even regenerate the same pattern even if the original pattern is cut in half [Subsec. 6.3.3]. In order to produce patterns, cells must inhibit the birth of similar structures in the cell immediately adjacent to them.

11.3.3 Cellular Differentiation

Cellular differentiation is the result of the different cell signaling combined with the mechanical driving-force processes previously considered (i.e. cellular division, motion, and death). The problem here is to know how the elements in the DNA sequence are used. Cell differentiation depends on changes in *gene expression*[75] [Subsec. 7.4.3], regulated, as we have seen, by RNA polymerase and proteins rather than gene loss, even if the genetic material is sometimes changed from one cell to another [Subsec. 9.7.2], as is the case with B and T cells of the immune system, whose DNA is irreversibly altered[76]: Most cells become different because they synthesize and accumulate different sets of RNA and protein molecules without altering the sequences of their DNA.[77] As we know, DNA is packed in highly compacted chromatin, including heterochromatin, which contain special proteins that make the DNA usually inaccessible to gene activator proteins. There is a cell memory [Sec. 9.8] because the choice of a particular cell typology will generally be maintained by many subsequent cell generations, which means that the changes in gene expression are somehow recalled [Subsec. 11.2.3].

The gross structural diversity of the organism is encoded in genes called *selectors* (they select distinct developmental pathways) that give rise to structures such as eyes, wings, and so on [Fig. 11.10]. The cells need to know not only which structure they are making but also where they are located within the structure. The positional information is controlled by a small set of intercellular signaling pathways.[78]

Humans show more than 200 different cell types (bones, blood, skin, muscles, hepatic cells, neurons, and so on). Cells are often determined for a future specialization long before they differentiate overtly. When cells are differentiated, they usually become regionally specified and

[74]For evidence of the existence of morphogens see [CHEN/SCHIER 2001].
[75][*ALBERTS et al.* 1983, pp. 411–99].
[76][*JABLONKA/LAMB* 2005, p. 68].
[77]See also [NOVINA/SHARP 2004].
[78][AFFOLTER/MANN 2001, GUSS *et al.* 2001].

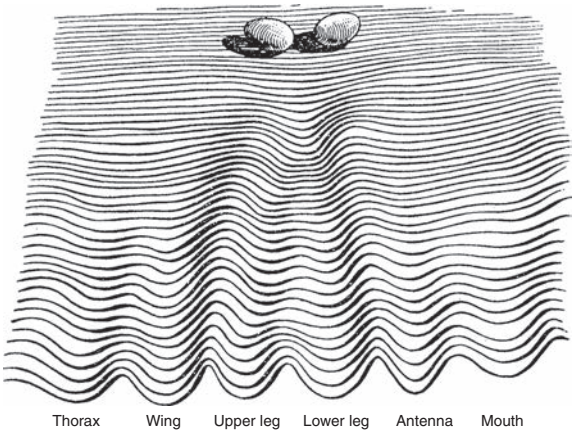

Fig. 11.10 Epigenetic landscape showing cell differentiation. Adapted from [*THELEN/SMITH* 1994, p. 123].

acquire a positional value that reflects their location in the body (this is again a form of memory). Another important feature is intercalation: Discontinuities of positional values provoke local cell proliferation, and the newly formed cells take on intermediate positional values so as to restore continuity in the pattern.

I have already spoken about the way cells relate to the internal environment [Subsec. 11.2.4]. Let us now consider the problem from the point of view of the different modes of differentiation. There are three ways in which this is done[79]:

- *Autonomous specification.* If a determinate blastomere is removed from an embryo in early development, this blastomere will produce the same type of cells that it would have made if it were still part of the embryo. This is called mosaic development.
- *Syncytial specification.* It is typical of insects, and consists in a division of the egg cytoplasm, creating many nuclei in a single large cell. There are also morphogen gradients in order to establish some positional information. For instance, high concentration of proteins bicoid in the anteriormost portion and nanos in the posteriormost portion of the *Drosophila* embryo leads to establishing the anterior–posterior axis.
- *Conditional specification.* It is the context-dependent specification. Here, cells are able to take over the role of other missing cells. This ability to change the cells' fate is called regulation or regulative development. It is also obvious from what I have said before that we expect the regulative development to predominate when ascending the ladder of complexity in evolution, and to determine what I have called post-Mendelian organisms [Subsec. 11.2.1].

In most animals the first and third forms of specification are combined. Once that a specific expression has been obtained, several mechanisms for maintaining cell differentiation have also evolved.[80] Motion (with subsequent adhesion) is an especially important driving force for the differentiation of cells.[81] Cell adhesion is provided by three classes of molecules: Cadherin molecules,[82] proteins from the immunoglobulin superfamily of proteins (N-CAM), and integrins. Since cells adhere more

[79][*GILBERT* 2006, pp. 53–67]. [80][*GILBERT* 2006, pp. 169–71].
[81][*WOLPERT et al.* 2002, pp. 253–8] [EDELMAN 1984a, *EDELMAN* 1992]. [82][*GILBERT* 2006, 71–4].

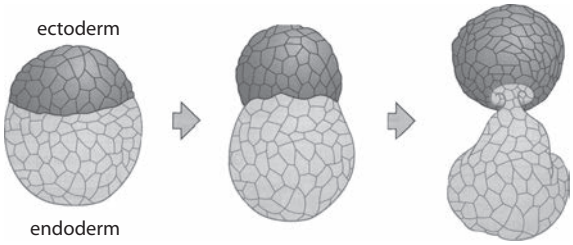

Fig. 11.11 When cells from early ectoderm (dark) and early endoderm (light) are placed together, they initially fuse but then separate until only a narrow strip connects them. Adapted from [*WOLPERT et al.* 2002, p. 256].

tightly with some groups rather than others, tissue differences emerge [Fig. 11.11]. In early stages of epigeny, the mechanical driving-force processes, such as cellular division, motion, and death, predominate.

11.4 The Stages of Epigeny

11.4.1 Preliminary Considerations

Though we already have many developmental aspects in microbial biology, it is in multicellular organisms that epigeny shows spectacular manifestation.[83] It is a process in time during which the organism changes its morphology by passing from a single fertilized cell to the adult form. As we have seen, it is ruled by several switch points (epigenetic buffers) allowing for the change from a default to an alternative developmental pathway[84] [Subsecs. 11.2.3–11.2.4 and Sec. 11.3]. Individual development always begins with an inherited bridging phenotype, that is, a responsive and organized cell.[85] An animal egg or a plant seed often has specialized physiological capacities and an adaptive external morphology. Moreover, cytoplasmic components can include organelles, ribosomes, proteins, and messenger RNAs. In this way, many parental features are transmitted in a way that is different from the genetic one [Sec. 9.8]. It is worth mentioning that the whole genetic complex represents less than 1% in the insects' mature egg volume; all the rest comes via the hemolymph of the maternal soma and in general reflects environmental variables like food, temperature, and so on, so that one could also speak of inherited environmental effects. The morphological, biochemical, and behavioral phenotype of the spermatozoan is a product of the paternal phenotype. It is also interesting that maternal gene transcripts continue to be used for some functions after embryonic gene expression begins, which points out the important continuity of the phenotype. This shows that there is also some crossgenerational continuity of information through phenotypes [Subsec. 11.2.2]. It is important to consider that the advantage of the genetic transmission line is indeed not in continuity but in the relative immutability and faithfulness of its replication. Indeed, while the genetic (allelic) variation is discrete [Subsec. 9.3.1], polygeny (and phenotypic inheritance) is continuous [Fig. 11.12].[86] An important point, as we know, is that phenotypic structures are the units of reproduction. Indeed, genes replicate but cannot reproduce themselves across the generations.

[83] [*MINELLI* 2003, p. 1]. [84] [*WEST-EBERHARD* 2003, pp. 67–8 and 129–35].
[85] [*WEST-EBERHARD* 2003, pp. 90–8 and 112–13]. [86] [*ARTHUR* 1984, p. 32].

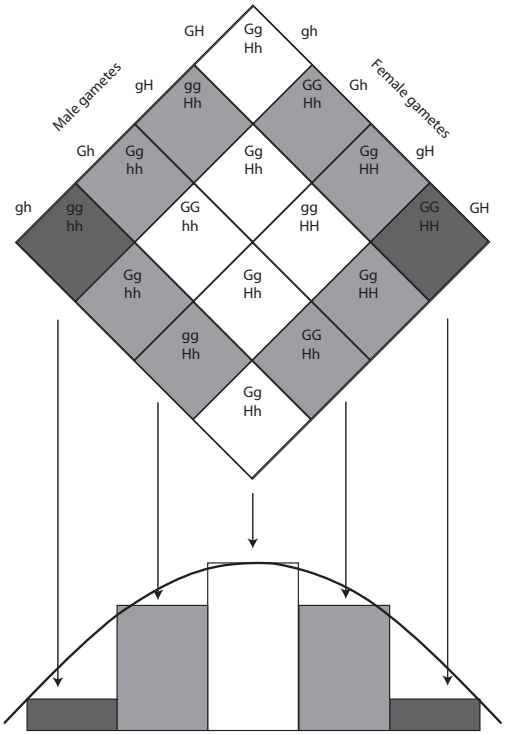

Fig. 11.12 Genetic variation and polygeny, representing discontinuity and continuity in heritage, respectively. By increasing the number of gametes to be combined (and therefore the number of squares on the top part of the figure), we approach more and more the continuous bell-shaped curve below.

Table 11.1 Development in animals with a CNS.

Development					
Fertilization	**Epigeny**				**Maturation**
	Embryogenesis			**Organogenesis**	**Larvation**
	Cleavage	**Gastrulation**	**Neurulation**		

11.4.2 Development

Development is the first stage of ontogeny [Secs. 10.1–10.2], and for vertebrates, can be divided into three general stages[87] [Tab. 11.1]:

[87]See [*WOLPERT et al.* 2002][*GILBERT* 2006] [*ALBERTS et al.* 1983, pp. 1305–415] for some good introductions to these issues.

- *Fertilization*, which represents the start of the whole epigenetic process, where environmental fluctuations are out of control. Here, contact and recognition between sperm and egg (or oocyte for earlier mammalians) must happen, followed by regulation of sperm's entry in the egg in order to give rise to the fusion of genetic material.[88] Finally, the egg metabolism is activated which gives rise to development.
- *Epigeny*, the stage at which at high level of plasticity is accompanied by increasing information control. For the sake of concision, I consider the whole process that comprehends embryogenesis, organogenesis, and larvation as epigenetic.
  (1) *Embryogenesis*, in which a whole form emerges from the fertilized egg.
  (2) *Organogenesis* in animals (and germination in plants), in which the organs (or body parts) are differentiated and formed.
  (3) *Larvation*. In animals we have *metamorphosis* for insects and some vertebrates, and *postnatal growth* in some reptiles, birds, and mammals. It is likely that the maximal plasticity of an organism is displayed between the end of organogenesis and the first steps of larvation.
- Sexual *maturation*, the postnatal stage of many organisms in which plasticity begins to decrease. It is the bridge to the full maturity of the organism.

For this reason, development can be considered as the set of modifications occurring before a multicellular organism reaches its final sexual maturity.[89] Here, besides fertilization, I only consider the first two big steps of epigeny. Larvation and maturation will be discussed—for humans—in the next part of the book. Moreover, I stress that I am essentially considering animal development. I wish also to point out two issues (which will be technically discussed in the Appendix to this chapter):

- The conformation of the embryo and its maturation can be understood as a wave-like propagation phenomenon. This does not at all mean that the point-like, discontinuous aspect is absent, as I have already mentioned [Subsec. 11.2.2]. Actually, during epigeny there are many critical moments. In general, the interaction with the environment will produce more or less violent shocks, as is evident for the Waddington effect that we shall discuss below.
- The morphogenetic patterns (especially after the conclusion of cleavage) are highly structured entities, inserted in further processes—being an organism far more than a mere collection of patterns, as is often the case for self-organizing systems [Sec. 6.3]. These patterns can be understood to a certain extent as the physical basis of any representational process in multicellular organisms, as we will consider later.

11.4.3 Fertilization

In almost all animals there are certain asymmetries from the start that allow for the establishment of important differences during development. These asymmetries are either due to internal anysotropies built into the egg during oogenesis or those resulting from external cues like the advent of the sperm during fertilization.[90] During the first stage of the development of some animal families, three axes are established: Anteroposterior (from the front to the rear), dorsoventral (vertically from back to belly), and mediolateral axes (from the medial plane outward to the left or to the right) [Fig. 11.13]. Animals showing bilateral symmetry are called bilaterian, which today can be subdivided into (eucoelomate) protostomes (ecdysozoans, platyzoans, lophotrocozoans), and

[88] [*GILBERT* 2006, pp. 175–206]. [89] [GRIESEMER 2000]. [90] [*DAVIDSON* 2001, pp. 90–4] [GURDON 1992].

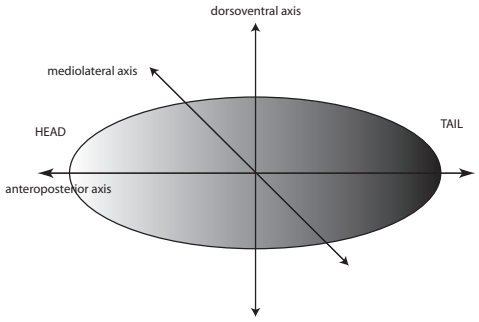

Fig. 11.13 The three animal axes.

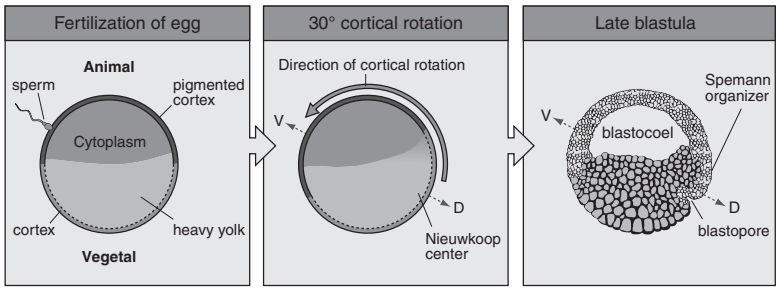

Fig. 11.14 The sequence from the fertilized egg to the blastula in amphibians. Adapted from [*WOLPERT et al.* 2002, p. 69].

deuterostomes (like echinoderms and chordates, whose first opening, which in protostomes then becomes the mouth, here becomes the anus).[91] In plants only the axis going from roots to the growth direction is established. It is interesting to observe that the three animal axes represent a reference frame (the axes are orthogonal one to another) and should be strictly connected with the fact that the animal is a mobile system in a three-dimensional space (in this way, those axes, to a certain extent, could be considered to be a rudimentary representation of space). The axis from head to tail is determined by what will be its forward direction of movement.

An interesting study case is presented by amphibians. In the frog *Xenopus*, an unfertilized egg shows a vegetal pole (destined to form internal tissues), which is the lower end of the egg, and an animal pole (destined to form external tissues such as the skin), constituting the upper end [first panel of Fig. 11.14]. Therefore, the animal–vegetal asymmetry of the egg accounts only for the anteroposterior axis of the embryo. In nonmammals, the yolk (food) is concentrated in the lower region. Fertilization in amphibians triggers a distortion of the egg contents which creates the dorsoventral asymmetry (it is determined by the point of sperm entry): The outer, actin-rich cortex is rotated relatively to the central pole so that the animal pole is shifted toward the future ventral side [second panel of Fig. 11.14]. Cells firstly are multiplied in very small cells that together finally form the blastula [last panel of Fig. 11.14]. In *Drosophila*, before starting to grow, the asymmetry

[91]For a better and deeper understanding of this tangled subject matter see [*MINELLI* 2009, pp. 53–70]. See also [*VALENTINE* 2004, pp. 138–41].

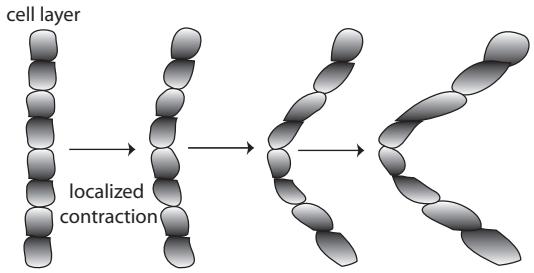

Fig. 11.15 Cell contraction starts the gastrulation. It is a self-increasing process ruled by the elastic and adhesion properties of cells.

head–tail is determined by the higher concentration in the head pole of *bicoid* mRNA able to produce the *bicoid* protein.[92] Though mammals like the mouse follow different principles, the site of the second polar body and the point of the sperm entry may define axes in the fertilized egg.[93]

11.4.4 Embryogenesis

In animals, *embryogenesis* is a complex process, whose main steps have been distinguished as cleavage, gastrulation, and neurulation. Let us now consider them. A fertilized cell *cleaves* (splits in a series of successive subdivisions) producing a complex of small cells. During this process the original totipotent zygote (possessing the capacity to build the whole organism) gives rise to pluripotent primitive stem cells (having the potential to differentiate into any of the three germ layers: Endoderm, mesoderm, ectoderm).[94] Until a relatively recent time it was assumed that embryonic stem cells were homogenous self-renewing cells. More recent studies[95] show instead that they appear to be in a metastable state and shift between inner-cell-mass-like and epiblast-like states while retaining pluripotency.

In many animals, the final result of this first step is the blastula (an epithelium which surrounds a cavity) [right panel of Fig. 11.14]. It is interesting to note that, during this stage of rapid cellular division, we have a relaxation of the cellular defenses[96] [Sec. 10.2] and that differentiation only begins when a multiplication process has already occurred. One can even say that, to a certain extent, proliferation and differentiation are mutually exclusive.[97] This is probably due to both a potential conflict between DNA replication and transcription as well as to the total reorganization of the cytoskeleton during mitosis, preventing a mitotitc cell from contributing to morphogenetic mechanisms [Subsec. 7.5.2].

The next phase is *gastrulation*, in which a differentiation of the organism begins that is initially driven by mechanical forces, in particular by cell contraction [Fig. 11.15]. In vertebrates endoderm, mesoderm, and ectoderm are constituted. They will give rise in a further step to (a) gut and respiratory tubes, (b) the organism's structure (the skeleton, heart, kidneys, muscles, and gonads), and (c) the outer skin and the sensory-CNS system, respectively. In particular, in both invertebrates and vertebrates both the endoderm and the mesoderm move from the outer surface of the embryo to the interior. Thus, it is interesting to note that a mechanism starting from the genetic system gives rise to the organism's metabolism (endoderm), to the ectoderm, and to the structural part of an

[92] [IRION/ST JOHNSTON 2007] [RUSTEN/STENMARK 2007].
[93] [*WOLPERT et al.* 2002, pp. 269–79]. The point is controversial, however.
[94] [SURANI *et al.* 2007].
[95] [HAYASHI *et al.* 2008].
[96] [*GILBERT/EPEL* 2009, pp. 125–9].
[97] [*MINELLI* 2003, p. 108].

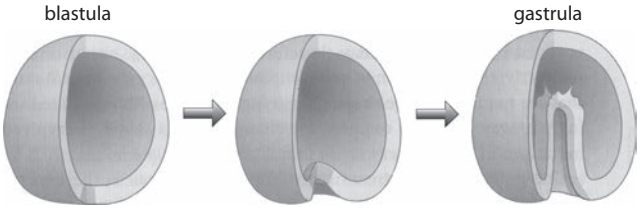

Fig. 11.16 Sea urchin's gastrulation. Adapted from [*WOLPERT et al.* 2002, p. 269].

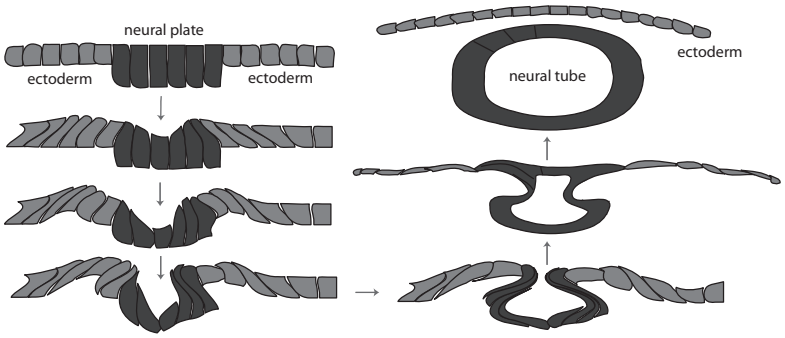

Fig. 11.17 Neurulation.

organism. In invertebrates like the sea urchin, gastrulation generates a multilayered structure with a central gut tube (going from the mouth to the anus), constituted by the endoderm [Fig. 11.16]. The gut is the first component of the organismic metabolism. Gastrulation also gives rise to bilateral symmetry in bilaterians.

As shown by Kauffman[98] [Subsec. 9.5.5], an intrinsic limitation on a partial differentiation in the organism is due to the fact that, by increasing the number of interacting parts, the number of conflicting constraints also increases, with the consequence that only poor structural equilibria become possible. This explains why there is a strong systemic segmentation or compartimentation [Sec. 3.6] in living beings, so that the number of kinds remains relatively low even if the number of individual exemplars can be relatively high.[99] This implies that compartimentation or modularity is always hierarchical[100] [Sec. 6.3].

During *neurulation*, vertebrates develop the ectoderm in the neural tube and several different neural cells differentiate.[101] Again, at the start mechanical cell-contraction forces come into play, in particular on that part of the ectoderm known as the neural plate [Fig. 11.17]. As I have said, the ectoderm gives rise to the entire nervous system. In neurulation a broad central region of the ectoderm thickens, rolls up in a tube (the neural tube), and pinches off from the rest of the cell sheet. Along the line where the neural tube pinches off from the future epidermis, a number of ectodermal cells break loose from the epithelium and migrate as individuals out through the mesoderm. These are the cells of the neural crest. The passage from mitosis to a true neural population with axon growth is provided by the annihilation of a gene regulator called Id2 that inhibits the protein complex E12-E47, which initiates the expression of neuron-specific genes.[102]

[98][*KAUFFMAN* 1993]. [99][*MINELLI* 2003, pp. 86–91]. [100][*WEST-EBERHARD* 2003, pp. 60–5].
[101][LIVESEY/CEPKO 2001]. [102][LASORELLA *et al.* 2006] [JACKSON 2006].

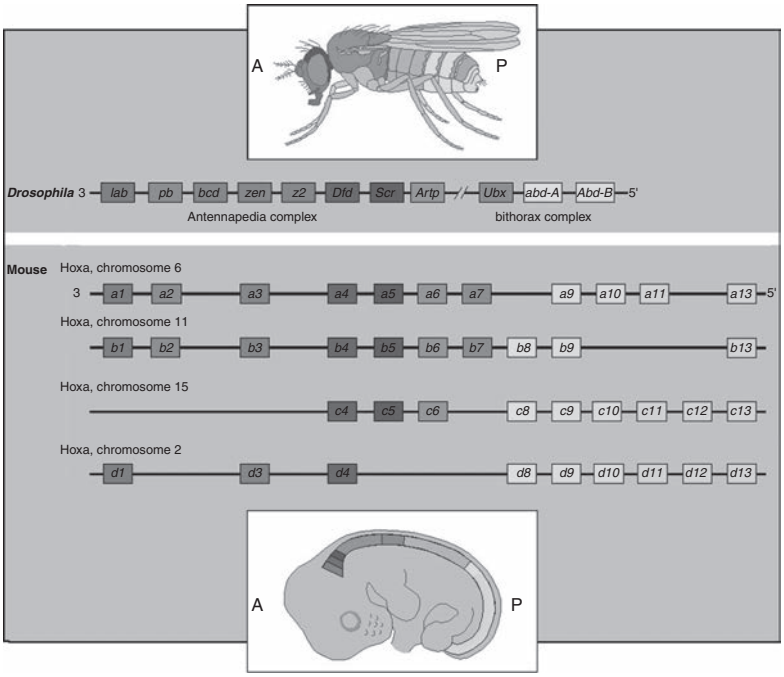

Fig. 11.18 The action of homeobox genes in the segmentation of the *drosophila* and the mouse. The homeobox genes for the *drosophila* are: *labial* (lab), *proboscipedia* (pb), *deformed* (Dfd), *sex combs reduced* (Scr), *antennapedia* (Antp), *ultrabithorax* (Ubx), *abdominal-A* (abd-A), and *abdominal-B* (abd-B). (This figure is reproduced in color in the color plate section.)

It has also been speculated[103] that vertebrates are truly dual animals with an ectodermal, somatic (neural) part and a endodermal, visceral part. It is interesting to compare the autonomous and local organization of the visceral part with the central organization of the nervous system.

11.4.5 Organogenesis

In the early stages of organogenesis, the cellular organization of some species is controlled by Hox genes [Fig. 11.18][104] and Pax genes. However, at a certain point, the control of connectivity, especially of neural connectivity, becomes even more interactive.[105] Hox genes are clustered together and are mapped to the mature organism structure in both spatial and temporal (sequencing) dimensions[106] [Figs. 3.15 and 3.16], which are also both collinear in the several copies. This ensures the robustness of developmental information against disruption. It also shows how patterns born in an epigenetic or metabolic context can acquire, in a different situation, a representational value.

I cannot enter here into the details of organogenesis (the construction of limbs and internal organs).[107] It is interesting to know that the Hox gene expression pattern in appendages resembles

[103]See [*MINELLI* 2003, p. 147]. [104]See also [SZATHMÁRY 2001] [*ARTHUR* 1997, pp. 154–6].
[105][*EDELMAN* 2004, p. 29]. [106][LEWIS 1978] [PATEL 2004].
[107]See [*GILBERT* 2006, pp. 443–527] [*WOLPERT et al.* 2002, pp. 331–70]. These authors consider organogenesis as part of embryogenesis [*WOLPERT et al.* 2002, p. 467]. However, in the case of plants, embryogenesis ends before germination [*WOLPERT et al.* 2002, p. 55], which to a certain extent corresponds to a part of the animal's organogenesis.

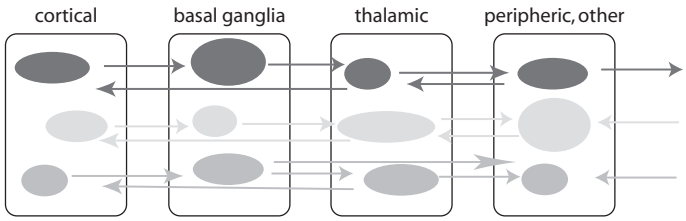

Fig. 11.19 How functional modules are superimposed on structural modules. The dark circuits (top) represent here the motor circuit, the light-gray (middle) the auditory circuit, and the middle-gray (bottom) the visual circuit. Note the cross relations established among the different areas. Inspired by [REDIES/PUELLES 2004].

that of the main body axis.[108] I wish to stress that the genes that control the building of an organism can be considered to be the same in all species and the unique differences across species are in the promoter and enhancer genes regulating expression.[109]

An interesting example of the hierarchical organization [Subsecs. 6.3.2 and 11.2.3] of gene regulation during organogenesis is given by the genes controlling the development of wings in the *Drosophila*[110]:

- A single selector gene (*Ubx*) establishes the segmental identity, at the first layer, and therefore whether the cell will develop a wing (the regulator T2, in the normal development) or haltere (the regulator T3, in the normal development) [see also Subsec. 9.5.3].
- The second layer is constituted by the selector complex VG–SD (the protein products of *vestigial* and *scalloped* genes, respectively). It defines the cell as part of the flight organ and elicits within it the developmental program necessary for building the organ.
- The third layer is constituted by the compartmental selector genes *engrailed* and *apterous* and helps to pattern the structure by defining the anteroposterior and dorsoventral compartments of the wings.
- Finally, genes, like those of the *achaete scute* complex, define the destiny of the individual cells in the organ.

This hierarchical structure helps us to understand how a complex regulation network may have phylogenetically arisen by *successive recruitment* of other regulators or protein complexes. It is also very important to realize that such structures arise and are operative both under the action of one or more selectors [Subsec. 11.3.3] and through the help of intercellular signal transduction [Subsec. 11.3.1], resulting in a beautiful example of combinatorial gene regulation.

Summing up, the final morphology of the organism cannot be reduced to mere adaptability (natural selection), nor to genetics alone. Structural genes only specify local rules (for building proteins), and not the global structure of the organism; yet, it is this global configuration that is the evolutionary basis of morphogenesis and the final aim of epigeny, not for growth and cell differentiation as such.[111] This mature form of the organism, as the result of development, is crucial for ontogenetic processes like niche-building [Sec. 10.3]. In this way, the whole epigenetic building is again an application of the law that rules complex systems, according to which local actions and interactions, inserted in feedback circuits and kept under control, can give rise to global structures [Secs. 6.3–6.5 and 8.2].

[108][*MINELLI* 2003, p. 164]. [109][BROCCOLI *et al.* 2000, p. 59]. [110][NELSON 2004a]. [111][*EDELMAN* 1988].

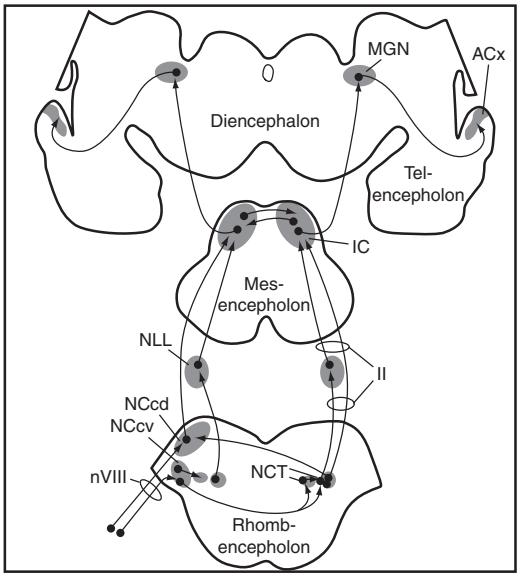

Fig. 11.20 Module from hearing. Auditory information reaches the rhombencephalon via the auditory nerve (nVIII), which terminates in the cochlear nuclei (NCcd and NCcv). From here, the auditory information is conveyed to both sides of the brain along specific fiber connections to other hindbrain nuclei (NCT and NLL) and along the projection in the lateral lemniscus (ll) to the midbrain auditory center, the inferior colliculus (IV). The information is further transmitted to a relay nucleus of the diencephalon, medial geniculate nucleus (MGN), and finally to the auditory cortex (ACx) of the telencephalon. Adapted from [REDIES/PUELLES 2004, p. 160].

11.5 Epigeny and the Brain

11.5.1 Further Investigation of Modularity

Some specific words on the epigeny of the brain seem appropriate here. We have already discussed the problem of the modularity of the brain [Sec. 3.6]. Recall that large and structural modules are not very important in the mature brain. Indeed, they reflect an early organization of the brain.[112] In the embryonic stage a modular structure of the brain (for instance, in vertebrates) is clearly visible and dominant. With regard to growth into the mature form, it is clear that a growing functional organization becomes superposed to the purely structural one [Subsecs. 3.4.1–3.4.2 and Fig. 11.19]. Here, plastic and transversal subnetworks are built which are recruited for different operations and can therefore cooperate. This does not mean that the structural compartmentalzation is not relevant. It corresponds indeed to many primary or basic features that will be the object of investigation later on. Moreover, some specific structural elements can be crucial for certain functions [Subsec. 8.2.4]. Concrete evidence of a functional circuit is provided in Fig. 11.20.

11.5.2 Genes and Learning

Development of the brain goes through the following steps[113]: (1) Proliferation, cell death, and migration, (2) aggregation, specification, and transient connections, (3) cell death and establish-

[112][REDIES/PUELLES 2004]. [113][RAKIC 2000]. See also [BOURGEOIS *et al.* 2000].

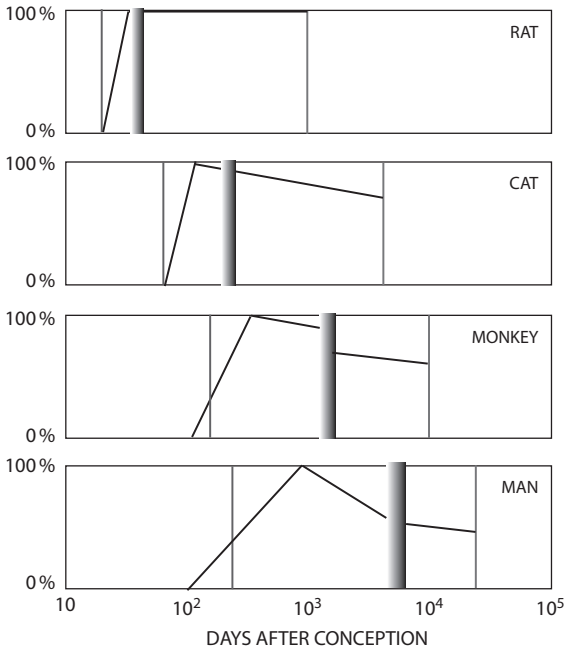

Fig. 11.21 Pictorial representation of the growth of synaptic density in different mammals. The thin vertical line on the left is birth, the one on right the death. The large gray bar in between represents puberty. Phase (3) in primates is more marked in the post-natal period (between the two thin bars). Adapted from [*CHANGEUX* 2002, p. 190].

ment of topography and synaptic strengthening. While phases (1) and (2) proceed in an orderly way in each individual according to a species-specific timetable, phase (3) is influenced by activity-dependent mechanisms which, especially after birth, involve individual experiences [Fig. 11.21]. During the second part of phase (2) and phase (3) the mechanism of synaptogenesis becomes more and more experience-expectant and a suitable motor activity becomes necessary for the proper final adjustment of the cortical circuitry. The transition from intrinsic to extrinsic regulation most likely involves the cellular mechanisms underlying learning and memory.

Hubel and Wiesel have provided evidence of the fact that, after a short period of plasticity extending from birth to six months of age, the connection from thalamus to cortex becomes fixed.[114] It is tempting to extend this from thalamocortical projections to all connections and functional properties of the primary sensory cortex. However, current evidence shows that these connections are highly dynamic [Subsec. 3.4.3]. It is puzzling that in the case of retinal lesions, there is a quick recovery (surrounding cells expand their activity in order to cover the silent region) because here cells seem to integrate information over a larger part of visual space than that covered by their specific receptive fields[115] [Sec. 4.3 and Subsec. 4.4.1]. The solution is that the definition of the receptive field is stimulus-dependent and a cell's response can be modulated by stimuli lying outside the classic receptive field, so that a cell's response to a complex visual stimulus cannot be

[114][WIESEL/HUBEL 1963]. [115][GILBERT/DARIAN-SMITH 1995].

predicted from its response to a simple stimulus, like that of a single short line segment. In fact, cells have overlapping receptive fields; for instance, removing inputs into the receptive field center may "unmask" and enable the expression of portions of the receptive field that are peripheral to the original receptive field center.

Neurocognitive development, therefore, relies on several complex interplays between genetic and environmental events, and the degree of interplay is highly variable across different neurocognitive systems.[116] This requires an appropriate distinction to be made: The systems mediating the representation of peripheral visual space (dorsal pathway) may be more modifiable than those representing central visual space (ventral pathway) [Sec. 4.3]. This could be so due to the fact that the representational part of vision can develop as a partly independent system (where fixed structures are established), whereas the referential (dorsal) part is more dependent on experience. The processing of magno (dorsal) stimuli could be selectively enhanced in congenitally deaf subjects, especially when attention is required. As a matter of fact, the parvocellular layers (leading to the ventral pathway) mature in humans earlier than those in the magnocellular laminae (leading to the dorsal pathways).

11.5.3 Theory of Neural Group Selection

According to the previous results, the foetus has many more axons than an adult: Development consists in a reinforcement of some connection and in the exclusion of others through neuronal death, elimination of collateral branches of neurons, and elimination of synapses by surviving neurons [Subsec. 3.4.4]. A case of information selection.

A very important deepening of these ideas is represented by Edelman's neural theory (the theory of neural group selection, TNGS), whose fundamental principles are[117]:

(1) Selection during epigeny allowing the constitution of a *primary repertory*;
(2) Postnatal, experiential reinforcement of some connections as a result of the interaction with the environment, which is a *secondary repertory*. Experiential selection does not occur, like natural selection, in evolution as a result of differential reproduction, but rather as a result of differential amplification of certain synaptic populations in individual organisms.
(3) Connections between several maps through *reentrant mechanisms*. Reentry is very common between different brain areas and neuron groups. It is especially relevant when information already processed in higher areas gives feedbacks into primary receiving areas [Secs. 4.3–4.4]. Reentry allows for establishing neural circuits. It does not simply consist of feedback (that occurs along fixed loops using previous instructionally derived information for control and correction) but occurs in selectional systems across multiple parallel paths where information is not prespecified.[118] Reentrant mapping is a nonalgorithmic relation between different maps that allows perceptual categorization: The sensorimotor activity acts in a selective manner and the resulting coordination between maps is the basis of the categorization.

TNGS meets the general requirements of evolution theory [Sec. 9.11].[119] It is a form of neural Darwinism and therefore an instance of generalized Darwinism [Subsec. 9.2.1]. Edelman, with his theory of neural group selection, represents an important advancement in understanding these processes, having shown that: (1) behavior is selected from a wider array of possibilities, (2)

[116][NEVILLE/BAVELIER 2000].
[117][CHANGEUX/DANCHIN 1976] [*EDELMAN* 1987, pp. 43–69] [*EDELMAN* 1992].
[118][*EDELMAN/TONONI* 2000, pp. 79–110]. [119][*EDELMAN* 1987, pp. 17–19].

dynamic perception–action mappings are primary in early life (joint firing of neurons, according to the Hebbian rule [Subsec. 3.8.2], reinforces certain connections), (3) multimodal exploration is a key process for acquiring new forms, (4) creation and exploitation of variability are crucial elements. In this way we can obtain a bridge between traditional dichotomies: innate vs. acquired, learning vs. maturation, evolution vs. development, genes vs. environment.

11.6 The Waddington Effect

An evolutionary process that goes on requires an initially heterogeneous environment, whose heterogeneity is continually increased since [Subsec. 10.3.2]: Different populations evolve, the epigenetic organization of the phenotype takes place, and in doing so it produces effects on the environment, and gene mutations, which are impossible at earlier stages and occur at the later stages.[120] Organisms placed in a different environment may first show changes in the phenotype as a response to the new conditions (this is indeed more economic than to change something in the genome, since it only requires some modification of the regulatory network). Subsequently, these changes may be genetically inherited (transferred to the genotype) through selection [Subsec. 9.7.1]. Waddington called this process *genetic assimilation*,[121] which is why it is known as the *Waddington effect*. For Waddington genetic assimilation does not run against the accepted Darwinian idea of evolution, since the environment rather induces a change in gene *expression* and not in the gene codification itself. The Waddington effect is of the highest importance as far as it establishes a connection between the environment and genes that cannot be found either in phylogeny, or in postnatal ontogeny. The departure point of Waddington was that sometimes it is impossible to account for evolutionary changes only in terms of simple point-like changes.[122] One of the preferred examples of Waddington was the comparison between the arm of a gibbon and that of a pangolin, which shows that a mere lengthening of the pangolin's arm would not turn it into that of the gibbon; instead, we are dealing here with a *precise set* of carefully coordinated changes involving different bones of the limb and of the shoulder girdle. No natural selection alone can produce such a coordinate change without assuming a canalization process occurring during development [Secs. 9.5 and 10.2]. For instance, ostriches show a characteristic callosity under their feet. Such a phenomenon could be the initial response of skin cells to friction and pressure during walking (representing the external stimulus here). Later on, this stimulus has been superseded by an internal genetic or epigenetic factor able to induce the same thickness of the skin at the right place.[123]

As Waddington showed, environmental pressures may lead to the expression of different genomic elements through alternative epigenetic operational pathways. At least in the laboratory, it is possible to observe these processes within single organisms, thus involved in the typical discontinuous generational events. When an environmental stress impinges on the developing organism during epigeny, by virtue of both information control and teleonomic mechanisms, the organism is able to recur to genetic resources otherwise not exploited. This will turn into a different-characterized phenotype showing a relative increase in fitness with respect to the considered environment. Note that, whereas organisms are systems featuring information control (they indeed adaptively change

[120] [WADDINGTON 1967].

[121] [WADDINGTON 1961a]. See [RUTHERFORD/LINDQUIST 1998]. The word "assimilation" here has an opposite meaning to my use in this book. I mean assimilation of the environment to the organism, while Waddington is speaking of an assimilation of the genome-expression to the environment. The fact is that assimilation can always be understood in both ways. In any case, since he speaks of *genetic* assimilation, I hope that no misunderstanding arises.

[122] [**?**].

[123] [WADDINGTON 1942].

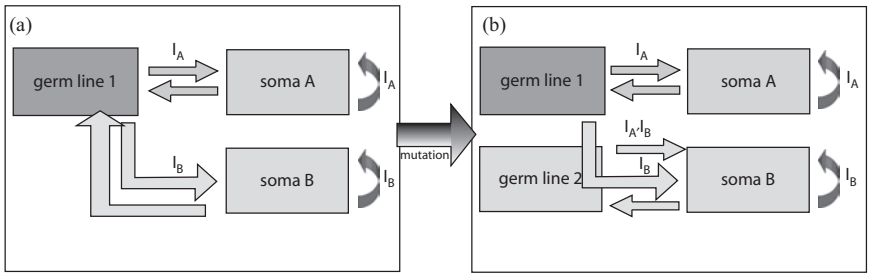

Fig. 11.22 Inheritance of epigenetic characteristics. System (a) represents a higher plant, with two types of differentiated somatic cell, each able to give rise to germ cells, which in turn can be converted into somatic cells. A single mutation can convert this into (b). I_A and I_B represent different environments, and soma A and soma B different adult phenotypes. Soma A breeds properly in environment I_A, but is converted to soma B if development takes place in environment I_B; soma B develops properly in any environment. Inspired by [*MAYNARD SMITH/SZATHMÁRY* 1995, p. 249].

thanks to this), the overall system represented by genotype, phenotype, and environment does not present a further higher-level control instance and is therefore a teleonomic network [Subsec. 10.3.3]. The natural balance we often observe is nothing but the result of many mutual and gradually fine-tuned interactions of the only informationally controlled systems (the organisms), and the nonliving physical matter they experience (without being directly instructed by that).

Waddington pointed out for the first time the relevance of the homonym effect when proved that eggs of *Drosophila* subjected to environmental temperature stress create mutant flies.[124] Another clear confirmation of Waddington's point of view may be found in the phenomenon of genetic imprinting[125]: In some loci, only the gene inherited from the father is active, while in other ones only the gene inherited from the mother. Since assortment at meiosis is random, the maternal chromosomes in the germ line of a male must be relabeled, just as for paternal chromosomes in the germ line of a female. For this reason, the germ line is accessible to reprogramming and the new labels may be transmitted through meiosis (again, a true form of semiotic marking). Maynard Smith developed these ideas into a dual inheritance-system model. Supposing that an inducer (representing the environment) can start a process giving rise to a given soma from a given germ line and also from another, then changed phenotypes are produced that are also subsequently inherited [Fig. 11.22].

It is interesting to observe that, while the Baldwin effect leads to an increase in plasticity [Sec. 9.10], Waddington's genetic assimilation is the opposite,[126] since it leads to fixation of characters. Moreover, the Waddington effect is concerned with a more active role of the environment for evolution (impacting on phenotypes), while the Baldwin effect is concerned with the reaction of organisms to complex changes in the environmental conditions. However, both refer to change in regulation, often affecting more than one trait (*pleiotropy*).

Budd[127] proposed an alternative model having affinities to Waddington's genetic assimilation but invoking discrete rather than continuous shifts that are in control of a particular morphology. The object of the study is the evolution of the Hox gene expression, which stresses the need

[124] [WADDINGTON 1952]. See also the excellent summary in [*JABLONKA/LAMB* 1995, pp. 32–37].
[125] [*MAYNARD SMITH/SZATHMÁRY* 1995, pp. 247–50]. [126] [*WEST-EBERHARD* 2003, pp. 151–63].
[127] [BUDD 1999].

for incremental functional integration. A surprising implication of the model would be that mutations in Hox genes and their regulators have virtually no primary role in driving morphological evolution. Rather, morphological change through microevolutionary adaptation comes first, with Hox expression shifting only afterwards, presumably to make the building of the new body pattern more efficient or more stable.

In conclusion: Is the Waddington effect a proof of some Lamarckian instructive action by the environment? Some scholars seem to think so. Jablonka and Lamb speak of a causal connection between the environment and genetic expression.[128] The point here is what we understand by "causal connection." If we understand the Waddington effect as evidence that the environment can have such an impact on an organism that the acquired changes (in the expression of the genome) are inherited, then it is so far Lamarckian in a sense that could be shared today [Sec. 9.8]. If the question is: Is information imported from the environment? My answer is no (and it appears that neither Jablonka nor Lamb is of this opinion). In fact, the environment has acted as a *stimulus* apt at *provoking* a certain epigenetic *response* in the organism.

11.7 A Wider Framework

From the previous analysis the necessity of integrating phylogeny, ontogeny, and epigeny into a wider framework arises. There is a growing body of evidence that morphological novelties in evolution originate as regulatory ones, and that evolutionary change is based on intraspecific developmental change, that is, innovative phenotypes arising from preexisting phenotypes within a developmentally variable population.[129]

11.7.1 Genetic and Developmental Networks

A developmental pathway is the sequence of causal events that propels a particular developmental process.[130] A developmental genetic pathway (*genetic pathway*, for short) is a sequence of key (mostly regulatory) gene activities that underlies a developmental pathway. Most of the highly conserved patterning genes act as intermediate steps in the genetic pathways (between genes coding for proteins and the production of proteins themselves). Therefore, the evolution of developmental pathways is to a large extent a matter of evolved differences in pathway components surrounding these key regulators. Changes upstream of the conserved regulatory genes can alter either their timing or the spatial domains of expression (or both). Changes downstream of the conserved regulators affect the sets of target genes (often regulatory themselves), and turn them on and off [Fig. 11.23].

Due to these structures as well as the multiple regulations of different control genes and feedback controls [Subsecs. 11.2.2–11.2.3], we have true *genetic* and *developmental networks* in which different pathways cross.[131] Developmental networks are the highest manifestation of plastic self-organization inside the organism [Secs. 6.3, 9.5, and 9.10]. This indeed ensures that a single-point mutation may determine true functional (cascade-like) alterations at very different taxonomic levels,[132] like phylum, order, family, and species. Obviously, the majority of these mutations will be deleterious, and the developmental network will show a certain robustness against such

[128] [*JABLONKA/LAMB* 1995, p. 31]. [129] [*WEST-EBERHARD* 2003, pp. 23 and 51]. [130] [*WILKINS* 2002, pp. 9–11].
[131] [*DAVIDSON* 2001, pp. 125–85]. Waddington had understood the relevance of this point [VAN SPEYBROECK 2002]. See also [KERZBERG/CHANGEUX 1994] [*CHANGEUX* 2002, pp. 168–74].
[132] [*DAVIDSON* 2001, pp. 188–240] [*WILKINS* 2002, pp. 118–21].

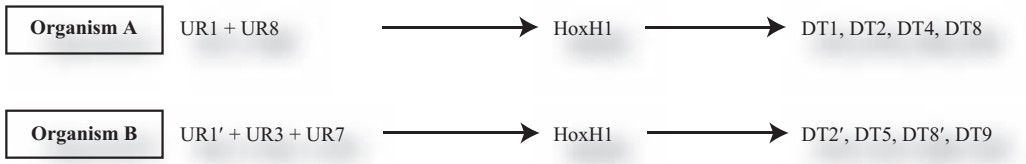

Fig. 11.23 Developmental genetic pathways in two different organisms. The prime symbol indicates that the "same" gene is operant in the two organisms (they are known as orthologue genes). DT stands for downstream target gene, UR for upstream regulators. Adapted from [*WILKINS* 2002, p. 10].

Table 11.1 Comparison of homologue genes in *Drosophila* and vertebrates controlling eye-building.

Drosophila	Vertebrates
PaxA: pox neuro	
PaxB: sparkling	*Pax 2,5,8*
PaxC: twin of eyeless, eyeless	*Pax 4,6*
PaxD: pox meso, gooseberry , gooseberry neuro, paired	*Pax 3,7*
sine oculis	*Six 1,2*
optix	*Six 3, Optx 2*
six 4	*Six 4,5*
eyes absent	*Eya 1,2,3,4*
dachshund	*Dach 1, 2*

alterations[133] [Subsecs. 8.2.4–8.2.5]. However, some can be neutral when they are not functionally relevant, which is also the case for those potentially deleterious mutations to which the network is robust. Since neutral modifications are more common than useful ones, a change of a positive activator in another positive activator is a much rarer change than mutational inactivation of an inhibitor molecule. Often, heterochronic changes have this character [Subsec. 11.2.5]. An interesting case is when we have a *master control gene*, i.e. a gene whose expression is sufficient to direct the development of a complete and properly formed organ or subsystem.[134] An important mechanism is when old genes are recruited for new functions[135] [Subsec. 11.4.5]. In this case, it is probable that first a potential target gene will enter into a previous chain and then successively come under the control of a regulatory gene [Fig. 11.24].[136]

An interesting example is provided by the network of *eyeless* (*ey*), *sine oculis* (*so*), *eyes absent* (*eya*), and *dachshund* (*dac*) in the *Drosophila* and their respective homologue *Pax*, *Six*, *Eya*, and *Dach* genes in vertebrates[137] [Tab. 11.1]. The network in the *Drosophila* is such that *ey* controls both *so* and *eya* as well as *dach*, and *eya* also controls *dach*. However, it is a nonlinear network in which all the four genes are required for eye development. Indeed, both *eya* and *dach* are able to induce expression of genes initially upstream in the network (*eya* on *ey* while *dach* on *eya* and *ey*).

[133] [*WAGNER* 2005, pp. 143–91]. [134] [*WILKINS* 2002, p. 151]. [135] See also [*VALENTINE* 2004, pp. 109–12].
[136] [*WILKINS* 2002, pp. 155–68]. [137] [KARDON *et al.* 2004].

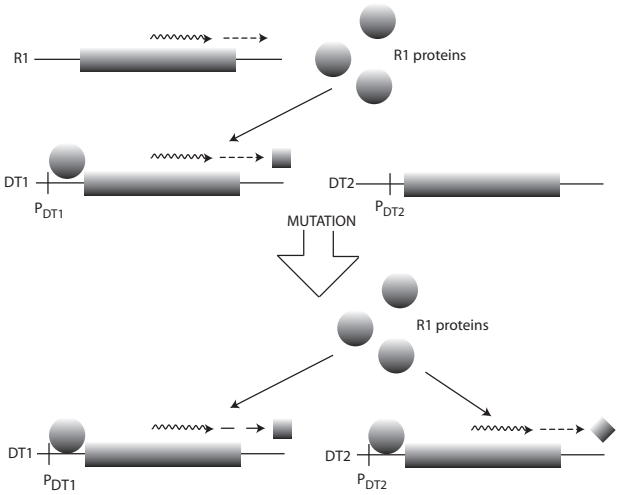

Fig. 11.24 The mechanism according to which old regulatory genes controlling a target gene giving rise to a key developmental function come to control new functions. A rare mutation in the promoter P_{DT2} of a downstream target gene DT2 allows the latter to come under the control of the regulatory gene R1, which already controlled DT1, giving rise to a new function. Considering that in general we are concerned with networks, a single regulatory gene can come to recruit several target genes as well as drop others along the history of evolution. Inspired by [*WILKINS* 2002, p. 156].

Now, how much of the inherited information must be continuous between lineages? Evolution can dissociate pieces of regulatory circuitry from their previous developmental functions (modularity). As such, we are obliged to admit that the most relevant aspect of genes is their *potential to give rise to functions* (their codification) rather than their *actual expression*; this again shows the relevance of potential information [Subsecs. 2.2.2, 7.4.1, and 7.4.5]. This also means that some presumptive convergences can be explained as the result of the independent expression of homologous genes that were previously unexpressed [Subsec. 9.5.3]. Nevertheless, even in those cases, homologous genes are not sufficient for explaining such a wide phenomenon[138] and we could not have phenotypic convergence without some environmental and phenotypical *constraints* allowing expression and threfore that specific phenotypic form [Subsec. 2.4.2]. This is evident especially when considering functions having a specific adaptive value in certain environmental contexts. For instance, even if the rhodopsin, the molecule in our retina that is responsible for fundamental aspects of vision in animals, is present also in bacteria and plants [Fig. 7.36; Sec. 4.2], it is only in animals that it gives rise to vision since this is necessary for controlling motion. In bacteria it is rather used to drive the proton pump to transfer hydrogen atoms as the basis for synthesizing ATP [Fig. 7.10]. Moreover, as it is evident from Tab. 11.1, most horthologous genes are not identical and are associated with other genes necessary for constituting specific developmental networks. In some cases, it is also possible that morphological and genetic homology dissociate.[139] Homology or horthology as such may ultimately consist in a combination of shared key genes plus shared developmental functions (through teleonomic processes) for which these genes became crucial. Actually, there is also evidence that the duplication of genetic material

[138] [*DE BEER* 1971]. [139] [*VALENTINE* 2004, pp. 131–2]. See also [CHÁVEZ *et al.* 2006].

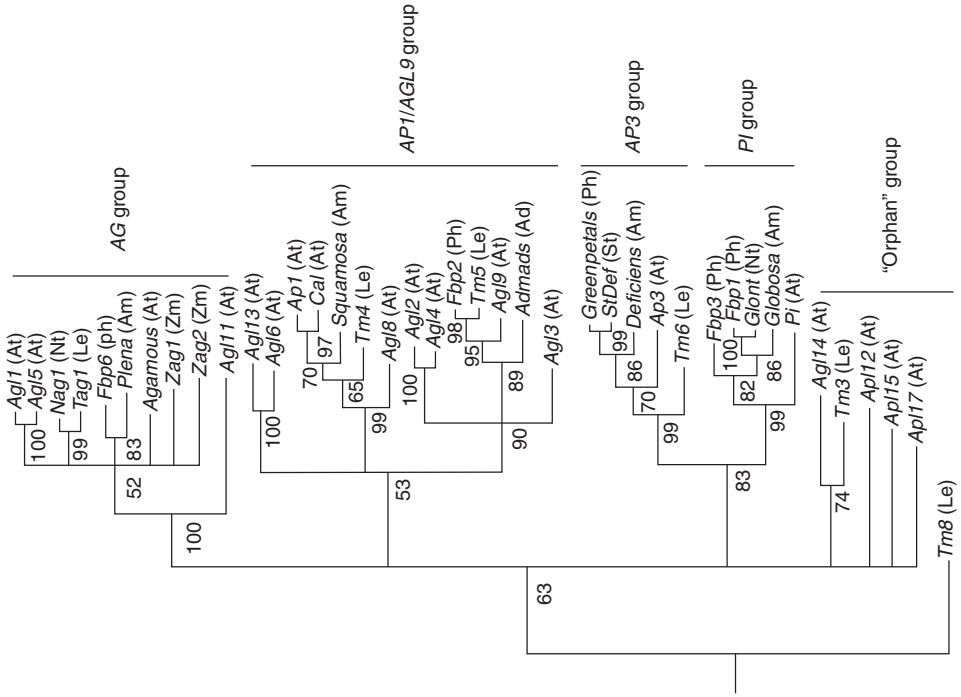

Fig. 11.25 Gene duplication in the evolution of the MADS-box plant gene family. Groups: *agamous* (*AG*), *apetala* (*AP*), *pistillata* (*PI*). Abbreviations in parentheses refer to species of flowering plants. Adapted from [*SCHLICHTING/PIGLIUCCI* 1998, p. 26].

allows for divergence during evolution.[140] Indeed, tens of thousands of genes of vertebrates can be cast into 1,000 gene families. Many genes of plants can also be grouped, showing their evolutionary dependencies [Fig. 11.25].

11.7.2 Piaget's Contribution

A very important contribution to the understanding of epigeny was provided by Jean Piaget, who was influenced by both Baldwin and Waddington. According to Piaget,[141] genes, in their regulatory activity, give rise to the process that determines the phenotype. During epigeny, interaction with the external environment occurs [Subsec. 9.5.5]. If the interaction leads to phenotypical forms showing some instability (there is some mismatch between the genes and environment so that phenotypic canalization fails in part), a stressful negative feedback is sent to regulatory genes, and a random search for other solutions begins [Sec. 9.7].[142] Under the constraints of this negative feedback, either recent random genetic variations are integrated into the epigenetic mechanisms (Baldwin effect) or new genetic expressions are determined and inherited (Waddington effect) [Fig. 11.26].

[140] [*SCHLICHTING/PIGLIUCCI* 1998, pp. 24–7].
[141] [*PIAGET* 1967, pp. 172–4] [*PIAGET* 1974]. See also [*PLOTKIN* 1993, pp.113–15].
[142] [*JABLONKA/LAMB* 2005, pp. 79–102]. However, some of the effects considered here seem to be overtargeted.

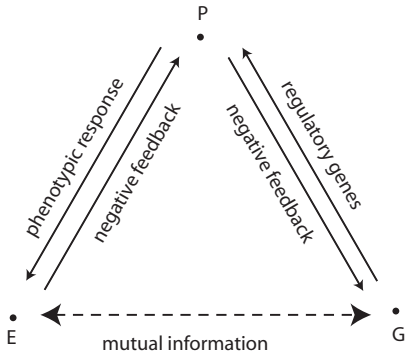

Fig. 11.26 In epigeny all the three systems constitute a true cybernetic circle and any can interact with any other. Here the mechanical action of the environment on the phenotype and the teleological action (behavior) of the phenotype on the environment are finally unified in a true teleonomic cybernetic system.

Piaget spoke of phenocopies (a *phenocopy* is an individual whose phenotype shows an induced and not genetic-based trait which is identical to another trait of another phenotype which is, instead, determined by its genotype), a concept that was originally due to Goldschmidt.[143] Here, it is also relevant to introduce the concept of *genocopies*, which is when mutant genes mimic environmental-induced effects.[144] In this way, the genetic search mechanism is activated but *not instructed* by the feedback. The genetic variations, though independent of the phenotype, can eventually produce a *coherent* response to the environmental solicitation, that is, may give rise, through epigenetic pathways, to a stable phenotypic form that is able to integrate the different levels that are necessary from the genome to the ripe phenotype. It is the combination of a teleonomic causal process with a teleologic one [Subsec. 8.2.7]. Let us call this *Piaget's law* for short. Obviously, any response that is not sufficiently coherent (unstable) will be eliminated through negative feedback (selection) of the environment on the phenotype (and therefore indirectly on the genotype). In this way, the environment and genotype come to an almost perfect correspondence, that is, to share information, even if they have *never directly interacted* with one another [Subsec. 2.2.3]. This also accounts for the fact that, when the phenotypic mutation is stable, no genetic change is necessary, as I have already anticipated, but a pure epigenetic and phenotypic accommodation (jointly with a suitable assimilation) suffices: Nature always prefers the most economical ways. This is what happens perhaps in the case of the human culture and, at least in part, for language. Such a mechanism explains two important aspects of evolution:

- Why there are often not intermediate forms [Subsec. 9.5.1].
- How random variations can be integrated into a complex organism.

In this way, we could understand epigeny as the confluence of phylogeny and ontogeny as well as of teleonomic and teleologic causality:

(1) With regard to the relation with phylogeny, environmental effects having evolutionary value are much more likely to influence the organism during its developmental period [Subsec. 11.1.2].

[143] [*GOLDSCHMIDT* 1940]. [144] [*WEST-EBERHARD* 2003, pp. 116–17].

Moreover, as I have said, changes in the regulatory pathways are more economic than those occurring at a genomic level. Finally, even changes in the genome always imply important (compensatory) changes in the regulatory network.

(2) Concerning the relation with ontogeny, development represents the first stage of the ontogenetic path as a matter of fact, whose successive steps are maturity and senescence [Secs. 10.1–10.2]. Epigeny provides the basic structure that will be used during maturity to deploy information control on the environment.

(3) With regard to the issue of teleonomy–teleology, the passage from development to maturity is mainly due to an increase in information control on the environment through teleological causation. On the contrary, the first stages of epigeny are dominated by teleonomic aspects. Finally, at an evolutionary scale, while less evolved species rely on mechanical, feedforward, driving-force processes, a suitable combination of teleonomic and teleologic aspects dominate for higher organisms (especially mammals), where regulatory feedback mechanism are crucial [Subsec. 11.2.4].

The cybernetic circle shown in Fig. 11.26 also allows for a certain reconsideration of the evolutionary mechanism. It is clear that natural selection perfectly corresponds to the informational closeness of organisms. In other words, the reason why organisms adapt is the random interaction with the environment once their variation in informational content has occurred with complete autarchy. In other terms, channelization—the eventual agreement with the external environment—occurs only after the organism has "chosen" a determined epigenetic and phenotypic canalization [Subsec. 2.4.2].

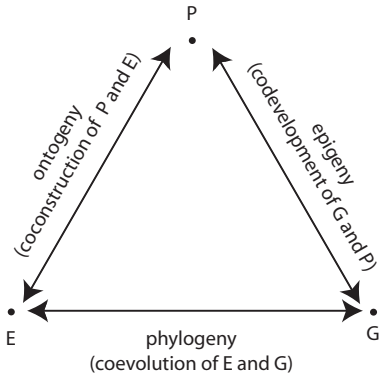

Fig. 11.27 Phylogeny–epigeny-ontogeny in their mutual relations.

Phylogeny is a co-evolution of the environment and genome: The hidden joint is represented here by the phenotype. Indeed, all evolutionary niches built at the ontogenetic level, as well as all genome selections, are due to the environmental selective pressure on changes in the phenotypes during evolutionary time, in a way that finally becomes coordinated.

The phenotype and the environment are built together at the ontogenetic level: Here the hidden bridge is the genome. Indeed, the environment is built and rebuilt by phenotypes that are ultimately the result of genetic information.

The genome and the phenotype codevelop in epigenetic time: The genome is expressed, silenced, influenced by environmental cues and stresses on the phenotype, and the phenotype is built thanks to the same effects.

There is also some evidence for this cybernetic circle.[145] As is well known, different species can show similar phenotypic characteristics because since they descend from a common ancestor (homology) or because they converge to common traits [Subsec. 9.5.3]. Now, the cybernetic mechanism is very similar in both cases since in each context a similar gene is activated through regulatory mechanisms.

Evolution can therefore be better understood as a passage from a multidimensional genotype space to a multidimensional epigenetic space that is in turn mapped to a multidimensional phenotypic space that is finally mapped in a one-dimensional fitness space (where only the parameter of fitness is of relevance).[146] The final result of this integration of phylogeny, ontogeny, and epigeny can be cast as in Fig. 11.27.

11.8 Concluding Remarks

Epigeny flourished as a consequence of the emergence of multicellular organisms, a form of stabilization of previous evolutionary stages in which several similar organisms find themselves together in certain particular conditions:

- We distinguish between a pre-Mendelian stage where the phenotype is weakly connected to the genotype; a Mendelian stage in which there is one-to-one correspondence between genotype and phenotype; and a post-Mendelian stage, in which regulatory processes dominate and there is a one-to-many correspondence between genome and individual phenotypes.
- Epigeny can be considered a compromise between environmental inputs and control mechanisms of the organism. Therefore, it is also a combination of teleologic and teleonomic causal processes.
- Epigeny shows both continuous and discontinuous aspects.
- Among the main mechanisms of epigeny are cellular induction and transduction, cellular motion and position, cellular differentiation.
- Epigeny is the developmental stage between fertilization and sexual maturation. It consists in embryogenesis, organogenesis, and larvation. Embryogenesis can also be considered as articulated in cleavage, gastrulation, and neurulation.
- The brain development shows a functional organization superposed on a structural one. The postnatal development is particularly relevant for its functional organization. TNGS provides an interesting insight into postnatal brain development by stressing the role of reentrant maps for higher cognitive functions.
- The Waddington effect (genetic assimilation) is to a certain extent the opposite relative to the Baldwin effect (consisting in part of genetic accommodation) and consists of some environmentally induced stressful stimuli in the embryo that are successively inherited. It is a form of epigenetic inheritance.
- Such a mechanism can be understood when considering two crucial elements. The first is the fact that coding genes are inserted into networks of regulatory genes that determine which genes are expressed and with which timing. Very important feedback effects are determined in this way.
- The second aspect is Piaget's discovery that environmental stresses (negative feedback when there is some mismatch between organism and environment) act on the organism by inducing a search in both the genetic and epigenetic spaces. Stable epigenetic solutions are more economic, and therefore, when found, easily stabilized and eventually, at an evolutionary scale, genetically grounded.

[145] [PRUD'HOMME *et al.* 2006] [WRAY 2006]. [146] [WADDINGTON 1968b].

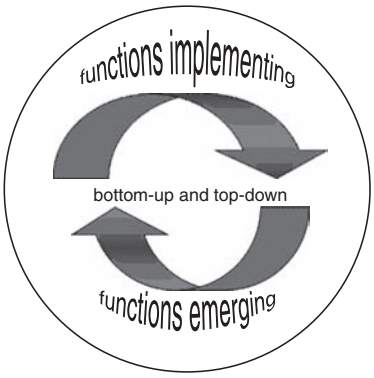

Fig. 11.28 The global circle of biological causality allowing new functions (and species) to emerge as well as providing the minimal biomolecular variety required for implementing such functionalities. Both processes show bottom-up and top-down aspects.

- In this way mutual information is created between environment and genome in a cybernetic circle such that phylogeny is the coevolution of genome and environment, ontogeny the coconstruction of phenotype and environment, and epigeny the codevelopment of genome and phenotype.

Through the very complex forms of interplay we have considered in this chapter and in the last two, functionalities continuously emerge and variety is displayed allowing their implementation (for instance, through exaptation) in some working operation [Sec. 8.2]. It is the circle between top-down and bottom-up processes at the level of environmental networks and even of the whole biosphere [Fig. 11.28].

With this chapter we have concluded what could be called the first block of the second part of the book [Chs. 6–11], which has dealt with the basic biological aspects grounding higher cognitive processes. With the next chapter, we shall deal with representational and cognitive aspects mainly in mammals.

Appendix: Morphogenesis

Let us make use of a form of the Fisher equation for population diffusion (genes, phenotypes, morphogenes), which is a development of Eq. (10.6) and, in nondimensional terms , may be written as[147]

$$\frac{\partial}{\partial t}n = r\left(1 - \frac{n}{K}\right)n + D\frac{\partial^2}{\partial x^2}n\,, \tag{11.1}$$

where D is the diffusion rate of some stuff. The steady states are $n = 0, 1$, which are, respectively unstable and stable. Let us write, in analogy, the equation system in the form[148]

$$\frac{\partial A}{\partial t} = F(A, B) + D_A\nabla^2 A, \tag{11.2a}$$

$$\frac{\partial B}{\partial t} = G(A, B) + D_B\nabla^2 B. \tag{11.2b}$$

[147][*MURRAY* 1989, pp. 236–41]. [148][*MURRAY* 1989, pp. 372–414].

We have at least three different approaches. The simplest kinetics is given by

$$F(A,B) = k_1 - k_2 A + k_3 A^2 B, \qquad G(A,B) = k_4 - k_3 A^2 B, \tag{11.3}$$

where the k's are the positive rate constants, and A is created autocatalytically by the $k_3 A^2 B$ term in $F(A,B)$. Alternatively, we can make use of an activator–inhibitor mechanism, namely

$$F(A,B) = k_1 - k_2 A + k_3 \frac{A^2}{B}, \quad G(A,B) = k_4 A^2 - k_5 B, \tag{11.4}$$

where A is the activator and B is the inhibitor, and the term $k_3 A^2/B$ is again autocatalytic. The third approach is the so-called substrate-inhibition mechanism, given by

$$F(A,B) = k_1 - k_2 A - H(A,B), \quad G(A,B) = k_3 - k_4 B - H(A,B), \tag{11.5a}$$

$$H(A,B) = \frac{k_5 AB}{k_6 + k_7 + k_8 A^2}. \tag{11.5b}$$

Let us now introduce L as a typical length scale and choose:

$$a = \frac{k_1}{k_2}\left(\frac{k_3}{k_2}\right)^{\frac{1}{2}}, \quad b = \frac{k_4}{k_2}\left(\frac{k_3}{k_2}\right)^{\frac{1}{2}}, \quad d = \frac{D_B}{D_A}, \quad \gamma = \frac{L^2 k_2}{D_A}, \tag{11.6a}$$

$$u = A\left(\frac{k_3}{k_2}\right)^{\frac{1}{2}}, \quad v = B\left(\frac{k_3}{k_2}\right)^{\frac{1}{2}}, \quad t^* = \frac{D_A t}{L^2}, \quad \mathbf{x}^* = \frac{\mathbf{x}}{L}. \tag{11.6b}$$

Note that d is the diffusion coefficient ratio.

Dropping the asterisks for convenience, the dimensionless reaction diffusion system becomes [Figs. 11.29–11.31]

$$\frac{\partial u}{\partial t} = \gamma(a - u + u^2 v) + \nabla^2 u = \gamma f(u,v) + \nabla^2 u, \tag{11.7a}$$

$$\frac{\partial v}{\partial t} = \gamma(b - u^2 v) + d\nabla^2 v = \gamma g(u,v) + d\nabla^2 v, \tag{11.7b}$$

where f, g are defined by these equations. An appropriate nondimensionalization of Eqs. (11.4) and (11.5a) is given by

$$f(u,v) = a - u - h(u,v), \;\; g(u,v) = \alpha(b - v) - h(u,v), \tag{11.8a}$$

$$h(u,v) = \frac{\rho u v}{1 + u + K u^2}, \tag{11.8b}$$

where a, b, α, ρ, and K are positive parameters.

The parameter γ can have any of the following interpretations:

- $\gamma^{\frac{1}{2}}$ is proportional to the linear size of the spatial domain in one dimension. In two dimensions γ is proportional to the area.
- γ represents the relative strength of the reaction terms.
- An increase in γ can also be thought of as being equivalent to a decrease in the diffusion coefficient ratio d.

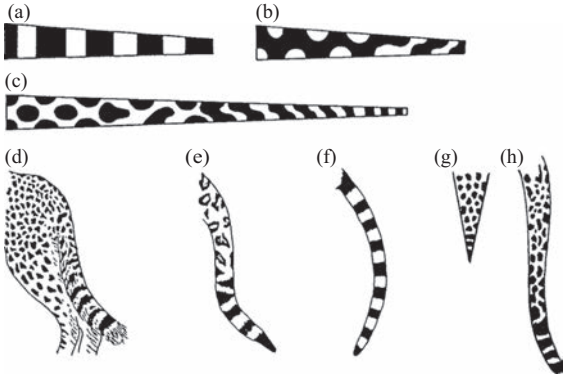

Fig. 11.29 Felines' tails [see also Subsec. 6.5.2]. The dark regions represent concentrations of the morphogen u [see Eqs. (11.7a) and (11.9)] above the steady state u_s. Fixing the parameters of the problem as $\alpha = 1.5$, $K = 0.1$, $\rho = 18.5$, $a = 92$, $b = 64$ (which implies a steady state $u_s = 10, v_s = 9$), $d = 10$, we have: (a) Scale factor $\gamma = 9$, (b) $\gamma = 15$. Notice that the pattern bifurcates into more complex patterns as γ increases. (c) $\gamma = 25$ (here we have the spot-to-stripe transition). Here the dark regions have $u < u_s$. (d) Typical tail marks of an adult cheetah (*Acinonyx jubatis*). (e) Typical adult jaguar (*Panthera onca*) tail pattern. (f) Prenatal tail markings in a male genet (*Genetta genetta*). (g) Typical tail markings of an adult leopard. Adapted from [*MURRAY* 1989, p. 441].

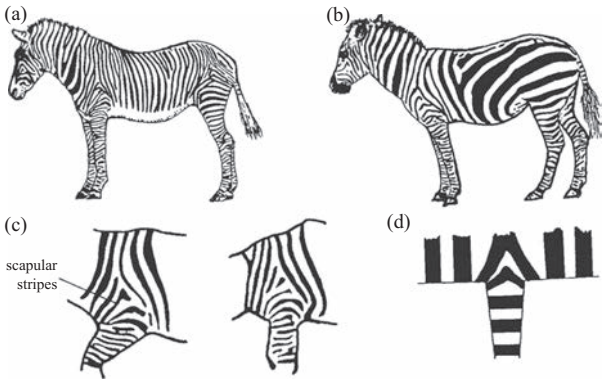

Fig. 11.30 Zebra patterns. (a) *Equus grevyi*. (b) *Equus burchelli*. (c) Typical examples of scapular stripes on the foreleg of a zebra (*Equus zebra zebra*). (d) Predicted spatial pattern from the reaction diffusion mechanism. Adapted from [*MURRAY* 1989, p. 442].

It can be shown that spotted animals have tails with stripes (this depends on the size of the animal).[149] First, let us define the nondimensional system by making use of Eqs. (11.7a) and (11.8a):

$$\frac{\partial u}{\partial t} = \gamma f(u, v) + \nabla^2 u, \quad \frac{\partial v}{\partial t} = \gamma g(u, v) + d\nabla^2 v, \tag{11.9a}$$

[149] [*MURRAY* 1989, pp. 435–68].

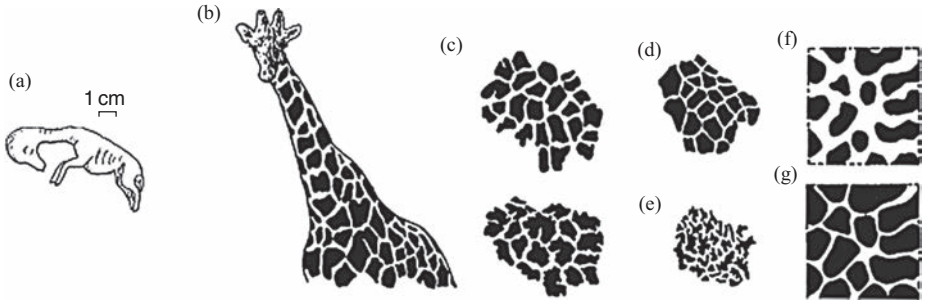

Fig. 11.31 (a) Giraffe (*Giraffa camelopardalis*): 35–45 day embryo. (b) Typical neck spots on the reticulated giraffe (*Giraffa camelopardalis reticulata*). (c)–(e) Tracings of trunk spots of *Giraffa camelopardalis* (c) *rotschildi*, (d) *reticulata*, (e) *tippelskirchi*. (f) Spatial patterns obtained from the model mechanism (11.9) with the same parameters as in Fig. 11.29. (g) Spatial pattern obtained when there is a lower threshold than in (f) is considered to initiate melanogenesis in the same simulations which gave (f). Adapted from [*MURRAY* 1989, p. 444].

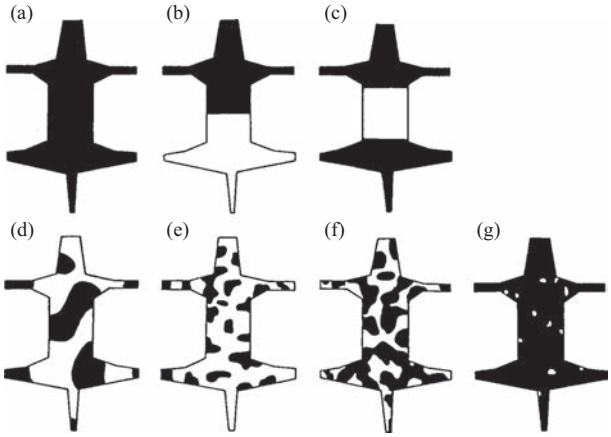

Fig. 11.32 Following D'Arcy Thompsons intuitions [Subsec. 9.5.4], we show here the effect of body surface scale on the spatial patterns formed by a reaction diffusion mechanism (11.9), with the parameter values $\alpha = 1.5$, $K = 0.125$, $\rho = 13$, $a = 103$, $b = 77$ (steady state $u_s = 23$, $v_s = 24$), $d = 7$. Domain dimension is related directly to γ. (a) $\gamma < 0.1$, (b) $\gamma = 0.5$, (c) $\gamma = 25$, (d) $\gamma = 250$, (e) $\gamma = 1250$, (f) $\gamma = 3000$, (g) $\gamma = 5000$. Adapted from [*MURRAY* 1989, p. 445].

$$f(u,v) = a - u - h(u,v)\ , \quad g(u,v) = \alpha(b-v) - h(u,v), \quad h(u,v) = \frac{\rho uv}{1+u+Ku^2}. \tag{11.9b}$$

Consider the surface of a tapering cylinder of length s with $0 \leq z \leq s$ and with the circumferential variable θ. The linear eigenvalue problem requires the solutions $\mathbf{W}(\theta, z; r)$ of equation

$$\nabla^2 \mathbf{W} + k^2 \mathbf{W} = 0\ . \tag{11.10}$$

Since we are only concerned here with the surface of the tapering cylinder as the domain, the radius r of the cone at any point is a parameter which reflects the thickness of the cylinder at given z. Then, we have

$$\sum_{n,m} \mathbf{C}_{n,m} e^{\lambda(k^2)t} \cos(n\theta) \cos \frac{m\pi z}{s}, \tag{11.11}$$

where

$$k^2 = \frac{n^2}{r^2} + \frac{m^2\pi^2}{s^2}, \tag{11.12}$$

$$\gamma L(a, b, d) = k_1^2 < k^2 < k_2^2 = \gamma M(a, b, d), \tag{11.13}$$

and L and M are functions only of the kinetics parameters of the reaction diffusion mechanism. If the tapering cylinder is very thin everywhere, then r is very small, which in turn implies that the first circumferential mode with $n = 1$ and all others with $n > 1$ in (11.11) lie outside the unstable range defined by the (11.13). In this case the unstable modes involve only z-variations. This is equivalent to the one-dimensional situation [Fig. 11.32]. If however r is large enough near one end so that $n \neq 0$ is in the unstable range, θ-variations appear. We thus have a situation in which there is a gradation from a two–dimensional pattern in z and θ at the thick end to the one-dimensional pattern at the thin end [Fig. 11.29]. In other words, by observing the coat of spotted animals, one sees that, by nearing to the thin end of the tail, spots are substituted by strips. Also the wing coloration of butterflies can be deduced by these methods.

Very small and very big animals present no differentiation pattern in their coat. By reaction diffusion, messages can be transmitted more quickly than by sole diffusion. Finally, I would like to stress that reaction diffusion is a pure wave-like propagation process.