universität
wien

# DIPLOMARBEIT

Titel der Diplomarbeit

# „A Systems-Theoretic Perspective on Functional Units in Molecular-Biological Networks“

Verfasser

Konrad Mönks

angestrebter akademischer Grad

Magister der Philosophie (Mag. phil.)

Wien, 2011

| | |
|---|---|
| Studienkennzahl lt. Studienblatt: | A 296 |
| Studienrichtung lt. Studienblatt: | Diplomstudium Philosophie UniStG |
| Betreuer: | Univ.-Doz. Dr. Manfred Füllsack |

# Contents

# List of Figures

# Chapter 1

# INTRODUCTION

In 2003, the first sequence draft of the human genome became publicly available. Since than, sequencing technology underwent truly revolutionary changes, presently culminating in what is known as 'next-generation sequencing' (cf. [1]). Through massive parallelization, decoding complete genomes became a matter of days, resulting in several thousand organisms being sequenced in less than a decade[1]. However, genome-centered technologies were not the only ones to experience a major shift from single-object analysis (e.g. one or few genes) towards high-throughput assessment of multiple molecular parameters (e.g. the whole genome). Based upon this novel experimental technologies also known as *omics methods*, and further enhanced by the progress made in information technology (cf. [2]), reconstructing 'complete' molecular interaction networks (*interactomes*) underlying cellular processes became feasible.

However, in order to *understand* a specific cellular function (e.g. DNA transcription, RNA translation, protein synthesis, meiosis, ...), the relevant components of the molecular network need to be known, i.e. the *functional unit* embedded in the network has to be identified. More precisely, to comprehend the difference between e.g. a diseased cellular state and the healthy counterpart, the molecular mechanism underlying this very difference needs to be deciphered. In general, it is assumed that only a rather reduced subnetwork of the complete interaction network is responsible for most phenotypic observables. In short, specific parts of the molecular networks, known as *functional units*, *functional modules* or *pathways*, are expected to encode specific cellular functions.

Presently, the most prevalent approach to identify such functional units in networks is based on the connectivity among the members of a unit, compared to the overall connectivity in the network. Thus, a unit is defined as a subnetwork which is densely interconnected (strong intra-connectivity) but only loosely related to the rest of the network (weak inter-connectivity). Apparently, this notion of a functional unit is driven by what can easily be measured, but not by a reasonable notion of *functionality* or *unit*. Given that the notion of a *functional unit* should not be determined by what can be measured, but rather should the objective of any measurement be determined by a principled understanding of functional unit, the question arises:

*What are functional, evolutionary units in molecular-biological networks?*

Even though we have outlined here merely the molecular-biological motivation of this question (cf. **subsection 1.1** for further details), there can hardly be any doubt

---

[1] As of today, the genomes of 11.743 organism are publicly available at the National Center for Biotechnology Information (NCBI-RefSeq: `http://www.ncbi.nlm.nih.gov/RefSeq/`)

that the quest for an understanding of *biological unity* is of general relevance and indeed dates back the ’primary substance’ of Aristotle. Furthermore, we explicitly do not refer to a specific type of biological network such as e.g. genetic networks, signal transduction networks or regulatory networks. Hence, given that networks as conceptual frameworks are suitable to represent and model an enormous variety of processes, it is reasonable to concentrate on the notion of *functional, evolutionary units* in a first step. Once this part of the question is properly addressed, it should be relatively easy to consider also the network specific aspects. Hence, the primary objective of this work is to elucidate the question:

> *What are functional, evolutionary units in biology?*

In order to address this question, a first guideline can be obtained from the theory of natural selection (cf. **subsections 1.2, 1.3, 1.4**). As a matter of fact, the Darwinian explanation of evolution as a process of selection and adaptation can arguably be seen as the single most important theory of biology. Hence, understanding biological units essentially entails understanding *evolutionary units*.

However, it has to be recognized that not all aspects of living units can be explained evolutionary. Indeed, natural selection itself has a series of prerequisites which have to be satisfied before any kind of evolution becomes feasible. Most importantly, the units upon which evolution is supposed to act have ’to be there’ in the first place. Thus, we stand in need of a theory explaining the emergence of units, or, stated more general, a theory explaining how the *precondition of natural selection* (cf. **subsection 1.5**) can be satisfied.

Finally, we will also examine an attempt to understand biological units beyond the context of evolution, but rather in the context of *cognition* (cf. **subsection 1.6**). Clearly, our notion of a functional unit in a molecular-biological network directly depends on our understanding of the *ontological status* of this very unit: e.g. in what sense does apoptosis, the functional module responsible for controlled cell death, ’exist’? If apoptosis is just an anthropogenic category, can we still expect to find *structural* criteria within a network that allow to identify such ’constructed’ units? But apart from the universal relevance of epistemological issues, the fact that organisms, including homo sapiens, are usually considered to be biological units as well leads to a rather peculiar situation: when trying to understand biological units, subject and object of cognition are identical. Apparently, trying to understand this very situation requires a specific notion of cognition.

In order to address this aspects, we have concentrated on three different research foci: First of all the concept of *Developmental Systems* is sketched as a possible notion of biological units (cf. **chapter 2**). In contrast to this approach originating from the philosophy of biology, several formal models proposed by the theoretical biologist *Stuart Kauffman* to simulate the emergence of units are outlined (cf. **chapter 3**). In **chapter 4**, several key concepts of the neurobiologist and philosopher *Humberto Maturana*, such as ’autopoiesis’ and ’structure determined systems’ are briefly introduced. Finally, in **chapter 5**, the major findings are summarized and discussed.

## 1.1 Functional units in molecular-biological networks

Throughout the last decade, the development of *high throughput techniques* capable of measuring simultaneously an elevated number of molecular parameters has revolutionized molecular biological research. These technologies can be seen as targeting different levels of cellular organization, namely the level of *DNA*, *RNA*, *proteins* and *metabolites*. As a characteristic, they attempt to account for the *totality* of ”objects” present at the level under investigation, i.e. the whole *genome*, *transcriptome*, *proteome* as well as the whole *metabolome*. Furthermore, not only methods to characterize the involved molecular entities, but also to identify *interactions* among them have been developed. This interactions may be of various different types[2], but due to experimental constraints, direct protein-protein interactions are currently the most frequently measured ones (cf. [4, p. 928]). Thus, driven by the constantly growing amount of measurable, molecular parameters and publicly available data, a novel research paradigm gained popularity: *systems biology* (cf. e.g. [5, 6, 7, 8]).

Through large scale data integration, a *holistic* (cf. [9]) understanding of cellular processes on the molecular level shall be achieved. To this end, the molecular states of the cellular machinery have to be mapped to phenotypic observables such as ”healthy” or ”diseased”. Hence, in accordance with ”traditional” biology, the ultimate goal of systems biology is to provide low-level (i.e. molecular level) models for high-level (i.e. cellular level) observables (cf. e.g. [10, 11]). Whereas there is typically little disagreement concerning the high-level observables, providing a suitable low-level model and linking it’s states to the observed phenomena can be seen as the key-challenge of biology. In the context of systems biology, *molecular-biological networks* represent the current gold standard model to address this very need (cf. e.g.[12]).

From a formal point of view, the nodes in such a network represent molecular entities (genes, RNA, proteins, metabolites, small molecules etc.), while the edges represent the interactions among these entities. Different *levels of abstraction* can be distinguished when modeling one and the same molecular system (cf. **figure 1.1**). More precisely, a level of abstraction is determined by the number of different *types* of nodes respectively edges which are distinguished in the network. Clearly, the impact of the chosen level of abstraction can hardly be overestimated. However, given the objective of this inquiry, we will not further discuss this issue but instead assume that *some* level of abstraction has been chosen. Furthermore, the problem we are about to address herein (*What are functional units in molecular-biological networks?*) is relatively independent of the specific network type and hence, our results are applicable to a wide range of different networks respectively levels of abstraction as well.

In order to gain a deeper understanding of network structure and dynamic, a wide range of analysis techniques has been developed (cf. e.g. [14, 15]). Apart from *statistical* measurements suitable to capture *global* topological network characteristics, as e.g. the amount of clustering present in the network (*average cluster coefficient*) or the *average shortest path length* between arbitrary pairs of nodes in the network,

---

[2] Indeed, ontologies (structured vocabularies) have been defined in order to capture the multitude of possibilities in which two or more molecular objects may ”interact” [3].

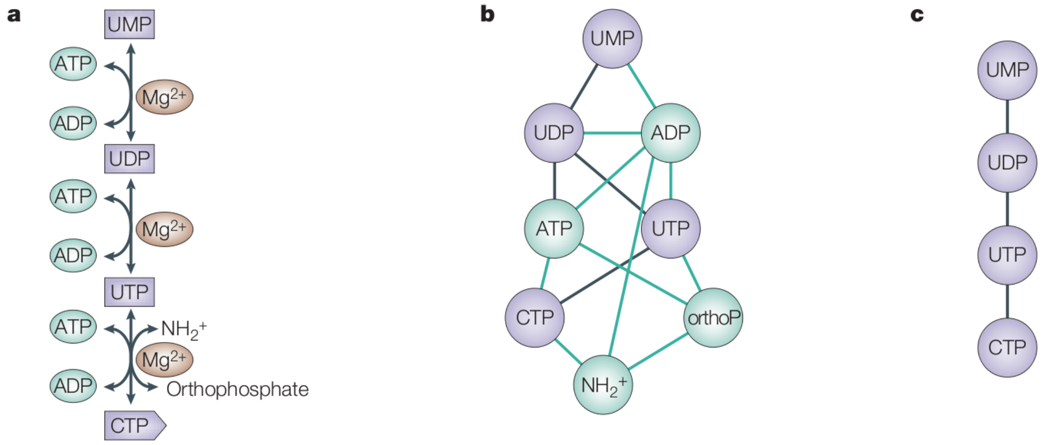


Figure 1.1: *Three network models corresponding to different levels of abstraction are shown for one and the same simple metabolic pathway. In a very fine-grained representation, all interacting metabolites and enzymes ($Mg^{2+}$) are included in the model* ***(a)****. However, depending on the research focus, a more abstract representation concentrating on the reaction's substrates and neglecting enzymes* **(b)** *and co-factors such as the high-energy donor ATP* **(c)** *may be more useful. Figure taken from [13]. For further details, refer to* section 1.1

measurements focusing on *local* properties have been devised as well. For example, the *node-degree* defined as the number of edges per node is suitable to identify highly interconnected nodes ("hubs")[3]. While the amount of clustering can be interpreted as a measurement for the *modularity* of a network, the edge degree is typically assumed to be related to the *relevance* of a node. As a consequence, the distribution of the edge degrees of all nodes of a network possibly reflects the network's *robustness* (redundancy), as a network with many hubs is more susceptible to perturbances (removal of a node) than a network in which all nodes are equally irrelevant. As an alternative to purely statistical measurements, an analysis based on frequent subgraphs (*network motifs*), typically comprising less than five nodes has been proposed as well (cf. e.g. [15]).

However, when confronted with a specific medical need as e.g. a disease phenotype, neither local (e.g. node degree or network motif) nor global (e.g. average clustering coefficient) analysis methods are suitable to provide a hypothesis on possible interventions. Instead, *meso-scale* approaches aiming to identify relevant *functional modules* are considered to be the method of choice (cf. e.g. [11, 17, 10]). The intuition behind this direction of research is rather obvious: in most cases, a phenotypic observable (e.g. a disease) will neither be related to just one molecular entity[4] (e.g. a single

---

[3] For a rather comprehensive list of existent topological network measurements refer to Platzer et al. [16]

[4] Clearly, there are cases in which a change in one single molecular entity is of phenotypic relevance, as can be seen for example in most heritable genetic diseases. However, the *mechanism* (pathway) which establishes the link between molecular cause and phenotypic effect *always* involves multiple interactors.

protein), nor to the totality of entities present in the network (e.g. the whole genome), but to a group of entities performing conjointly the disturbed cellular function. This is in accordance with the commonplace view that molecular networks are organized in *evolutionary modules* (cf. e.g. [18, 19]), i.e. that molecular networks can be decomposed in subgraphs which exhibit a certain amount of independence and which can be assigned a specific cellular *function* (cf. e.g. [20, 21, 22]). Identified functional modules, as for example the *apoptosis pathway* (cf. **figure 1.2**) can be found in *pathway databases*[5], such as *PANTHER* (cf. [24]) and *KEGG* (cf. [25, 26, 27]).

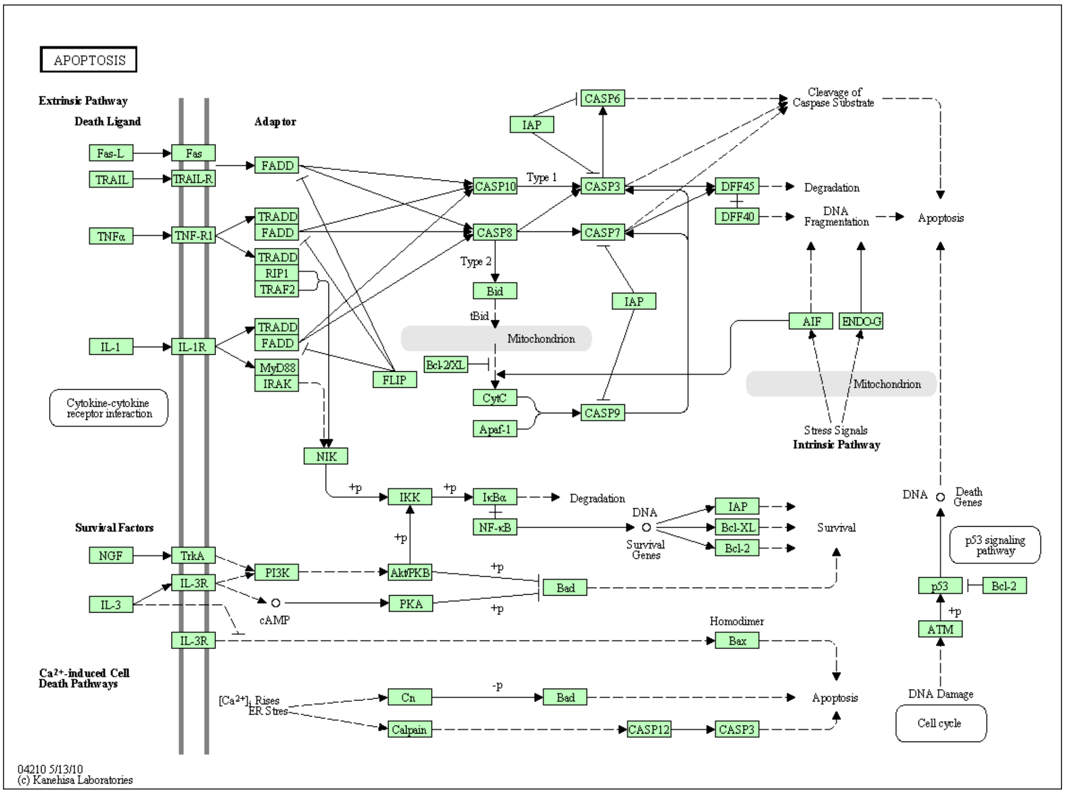


Figure 1.2: *The human apoptosis pathway as depicted in KEGG (cf. [25]). Apoptosis is a genetically controlled mechanism which induces cell death and hence, is of central relevance to tissue homeostasis. As can be seen, various causes such as e.g. DNA damage can trigger a series of molecular events leading to apoptosis. Given that his very mechanism is malfunctioning in tumor cells, it is one of the best studied functional modules in human. For further details, refer to* section 1.1

Apart from studying known pathways, there are also numerous attempts to identify novel functional modules in molecular-biological networks. Some of this attempts are based on the assumption that functional modules are also *evolutionary* modules: If a group of molecules conjointly perform a specific cellular function, this group is, *as a whole unit*, subject to evolution (cf. e.g. [19, 28]). Hence, when comparing

---

[5] A rather comprehensive list of currently available resources can be found at `http://www.pathguide.org/`. For an introduction into pathway databases refer to [23]

the molecular networks of different species, subnetworks exhibiting a similar structure in multiple species are taken to be functional units (cf. e.g. [29, 30]). However, in contrast to this approaches based on cross-species comparisons, the vast majority of methods proposed to identify functional modules is based on the assumption that such modules are densely interconnected subgraphs (cf. e.g. [18, 22, 12]). These methods are mainly heuristics for screening a given network for sets of nodes which exhibit a (statistically significant) high intra-module connectivity respectively a low inter-module connectivity, i.e. members of a module are more dependent on other members of the same module than on external entities.

Given that the relationship between functionality and conservation is beyond any reasonable doubt, the identification of functional modules based on a cross-species network comparison appears to be valid. In contrast to this, limiting the notion of functional modules to dense clusters seems to be problematic. The primary difficulty is to find a suitable *structural* criterion to identify *functional* units. Hence, while the concept of a specific function (e.g. apoptosis) is defined on the phenotypic level, the criterion we need to identify the corresponding subnetwork has to be defined on the molecular level. Thus, the challenge is to find a structural (low-level) description of functional units, departing from a generic (high-level) understanding of functional units. Thus, before concentrating on the identification of functional modules *in molecular-biological networks*, the notion of ”functional units” *in general* needs to be clarified.

To delimitate the scope of this inquiry, we will concentrate on *biological* and *evolutionary* units. Hence, in the next sections, we will outline the theory of natural selection, while putting some focus on it’s *universal validity* and it’s *necessary preconditions*.

## 1.2 Natural selection as optimization

Usually, it is claimed that three requirements need to be fulfilled for Natural Selection to take place.

> *”As seen by present day evolutionists, Darwin’s scheme embodies three principles [...]:*
>
> 1. *Different individuals in a population have different morphologies, physiologies, and behaviors (phenotypic variation).*
> 2. *Different phenotypes have different rates of survival and reproduction in different environments (differential fitness).*
> 3. *There is a correlation between parents and offspring in the contribution of each to future generations (fitness is heritable).*
>
> *These three principles embody the principle of evolution by natural selection. While they hold, a population will undergo evolutionary change.”* [31, p. 1]

First, certain variations have to be present in a population of individuals or, more precisely, a certain extend of variability is required for each individual. Some of these variations have to be linked to differences in the fitness of the individuals and, finally, some of these fitness-changing variations are heritable. Over the course of time the process of natural selection leads to populations of continuously increasing (average) fitness. If just one of these requirements is violated, natural selection can not take place. A generation of individuals which are exact copies of their ancestors is not able to change and adapt, variations which do not induce differences in fitness are evolutionary irrelevant and, finally, variations which can not be passed on to the next generation can not become traits of a species. An in-depth discussion of this requirements can be found e.g. in Godfrey-Smith's essay on the *'conditions for evolution by natural selection'* [32].

From a purely formal point of view, natural selection can indeed be understood as mechanism which continuously 'increases' a certain parameter, namely fitness.

> *"... we want to understand why selection has favoured particular phenotypes. The appropriate mathematical tool is optimization theory."*[33, p. 1]

As a matter of fact, the mechanism of natural selection represents a very generic approach to efficiently solve *optimization problems* (cf. e.g. [34]) and as such, can be formulated as a so-called *Genetic Algorithm* (GA). It is worth having a brief look at optimization problems and Genetic Algorithms as both concepts will provide us with a novel understanding of evolution itself.

From a formal point of view, an optimization problem consists in finding the extremes, i.e. maximum or minimum, of an *objective function* $f(x) = y$ (cf. **figure 1.3**). The objective function itself may represent any type of scalable property $y$ which depends on $x$ and shall be maximized respectively minimized. Note that each maximization problem can easily be converted into a minimization problem simply by using $f'(x) = (-1) \cdot f(x)$ instead of $f(x)$. Needless to mention that quite a number of procedures can be interpreted as optimization problems[6], but in order to *formally represent* them as optimization problems, the following two requirements need to be fulfilled:

1. Each possible solution to the problem can be represented by $x$, i.e. we are able to define a suitable *formal representation* for possible solutions.
2. We are able to evaluate the *quality* of each solution, i.e. we are able to define some $f(x) = y$ such that 'good' solutions obtain higher values of $y$ than 'bad' solutions

Now, lets have a look at how the principles of Natural Selection can be used to solve the problem depicted in **figure 1.3**. First of all note that the set of all possible solutions is given by [0..1] and their quality is measured by $y$. A (basic) Genetic Algorithm can be described as follows:

---

[6] Each time we enter a supermarket, we are in a certain sense solving an optimization problem: we choose a set of items which we believe to be *optimal* in some sense, e.g. in terms of price and / or nutritional value.

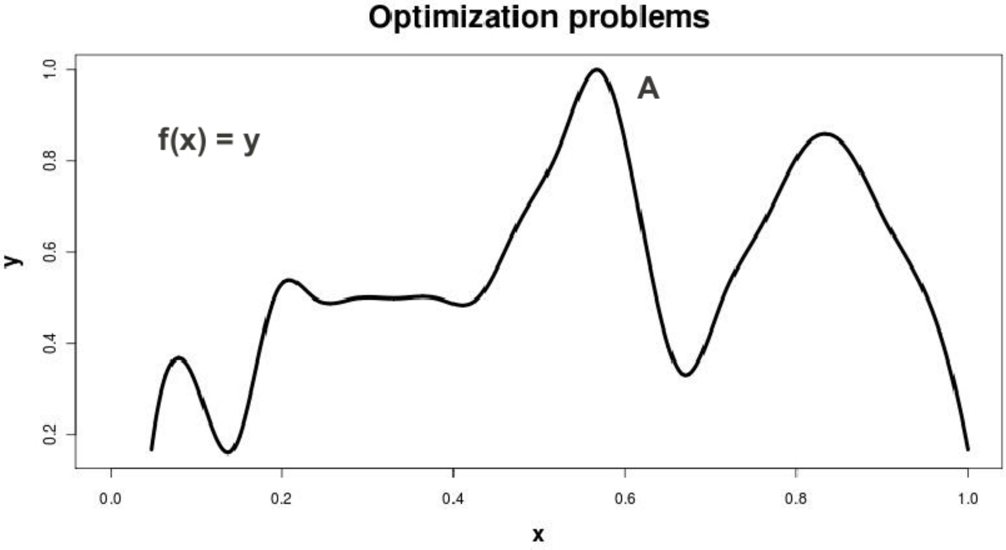


Figure 1.3: *An optimization problem consists in finding those values of x for which some* objective function $f(x)$ *is extreme. Assuming that in this case we seek the maximum (located at A:* $y = 1$*), the solution is found at approximately* $x = 0.58$. *In contrast to this illustration, the function to be maximized / minimized is usually high dimensional and / or discrete and thus analytical solutions are no longer applicable and dedicated methods, such as GAs, are needed. For further details, refer to* section 1.2

**Algorithm 1** Genetic Algorithm

```
1: $\bar{x} \leftarrow InitializePopulation()$
2: TerminationCriteria ← FALSE
3: while TerminationCriteria= FALSE do
4:     $\bar{y} \leftarrow f(\bar{x})$
5:     $\bar{z} \leftarrow selection(\bar{x}, \bar{y})$
6:     $\bar{x} \leftarrow mutation(\bar{z})$
7:     TerminationCriteria ← $terminate(x)$
8: end while
```

In a first step, a *population of individuals* is initialized. In principle, initializing the population can be done at random, i.e. any set of viable solutions can be chosen (e.g. $\bar{x} = \{0.1, 0.2, 0.8\}$. The rest of the algorithm consists of a loop, whose termination is controlled via some criteria defined in function *terminate()* (line 7), e.g. a maximum number of iterations ('generations') may be chosen. In line 4, the fitness of the population is evaluated (e.g. $\bar{y} = \{0.4, 0.45, 0.8\} = f(\bar{x})$) and in line 5, the 'fit' individuals are selected for reproduction. This novel generation ($\bar{z}$) still resembles exactly the parent generation ($\bar{x}$) and thus, in line 6, *small* variations are introduce. This variations are simply random modifications of the initial individuals. Note that, in order to maintain a population of constant size, we need to introduce novel individuals at some point of the process, given that in line 5 'unfit' individuals are removed from the population.

This can be done by replicating the chosen, fit individuals. The major consequence of this process is that over the course of time respectively with an increasing number of iterations (generations), the *overall fitness of the population increases*.

Thus, when switching back to biological terminology, we obtain the following picture: In the case of evolution, the objective function consists of the reproductive fitness of each individual, i.e. the *fitness function*, while the space of feasible solutions is the space of all possible biological forms, i.e. the *morphospace* or *phenospace*. In this picture, *the change of biological form over time is driven by the differential fitness associated with those forms*. Thus, evolutionary units can be thought of as 'traveling' on the *fitness landscape*, as illustrated in **figure 1.4**.

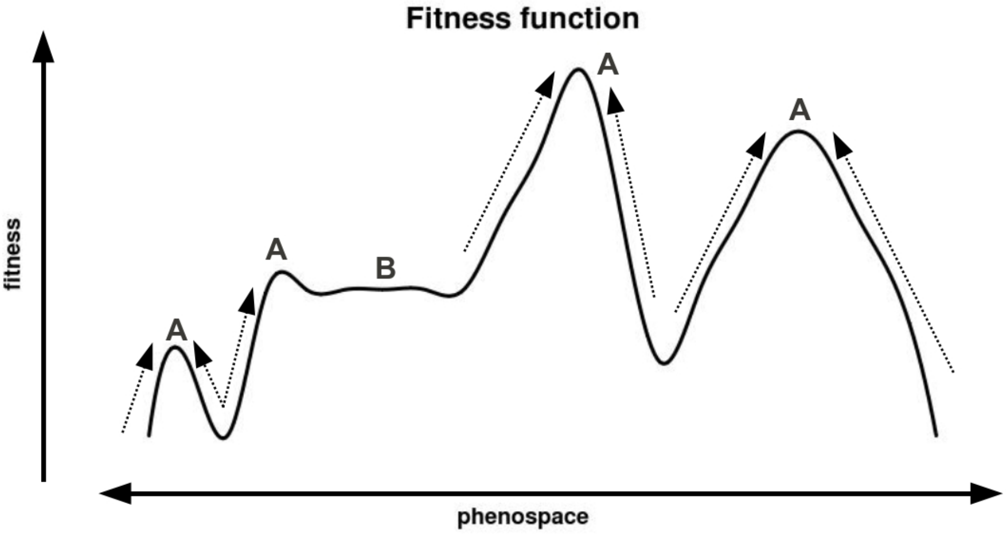


Figure 1.4: *Evolution as optimization. The process of Natural Selection 'pushes' species through the space of possible biological forms ('phenospace' or 'morphospace') and thus introduces a movements towards forms with increased fitness. Eventually, this mechanism will yield those species which exhibit optimal properties and thus, a maximal fitness (the peaks labeled with 'A'). Furthermore, the* inclination *of the fitness landscape yields interesting interpretations as well: The selective pressure always applies* vertically*, meaning that it does not apply at all to horizontal regions, like the* plateau *('B'). In this region, a high diversity of biological forms is possible, given that no differential fitness is present. For further details, refer to* section 1.2

As the notion of a fitness *landscape* suggests, the underlying morphospace can hardly be thought of as *unidimensional*. However, for the sake of visualization we are limited to uni-dimensional notions of morphospace (figure 1.4) or 2-dimensional ones (**figure 1.5**), bearing in mind that the space we are talking about is actually *high-dimensional*. Furthermore, note that even if we are not able to find a parametric representation of morphospace[7] (i.e. define morphospace in terms of measuarable variables),

---

[7] Personally I doubt that a reasonable parametric definition of morhpospace is possible, given the immense variety of biological forms.

the notion of a 'space' in which each 'dot' a.) corresponds to a possible biological form and b.) can be assigned a certain 'fitness' still remains valid. Consequently, the concept of a fitness landscape and all its derivatives, as illustrated in figures 1.4 and 1.5, still holds.

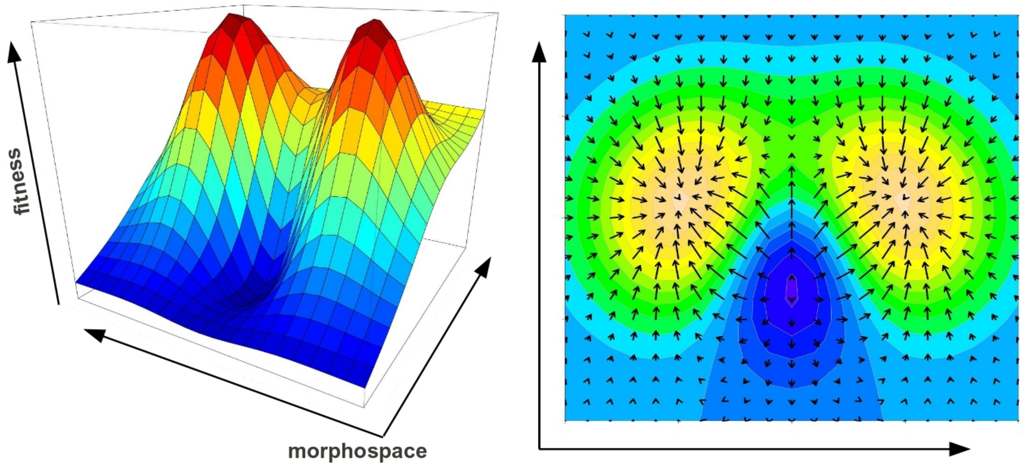


Figure 1.5: *On the left-hand side, a 2-dimensional illustration of a fitness landscape is shown, whereas on the right-hand side the evolutionary pressure is depicted as a force-field over morphospace. The force field is obtained from the gradients (direction and absolute value of steepest inclination) of the fitness landscape. Using this representation of the fitness landscape, the change of a species over time can be understood as it's trajectory towards an* atractor. *Furthermore, each atractor determines a* basin of attraction*, i.e. a set of points in morphospace which, when used as the point-of-departure of a specie's trajectory, will lead to the attractor itself. For further details, refer to* section 1.2

## 1.3 Speciation and convergence

If evolution is seen as an optimization procedure, several aspect of it can easily be explained, such as e.g.:

1. Speciation
2. Convergence ('Parallelism')

*Speciation* is the 'evolutionary process by which new biological species arise'[8] and thus represents one of the main phenomena that needs to be explained by any theory of evolution. In the context of the *phylogenetic tree*, speciation appears as branching and is the sole source for an increase in biological diversity. From a genetic point of

---

[8] http://en.wikipedia.org/wiki/Speciation

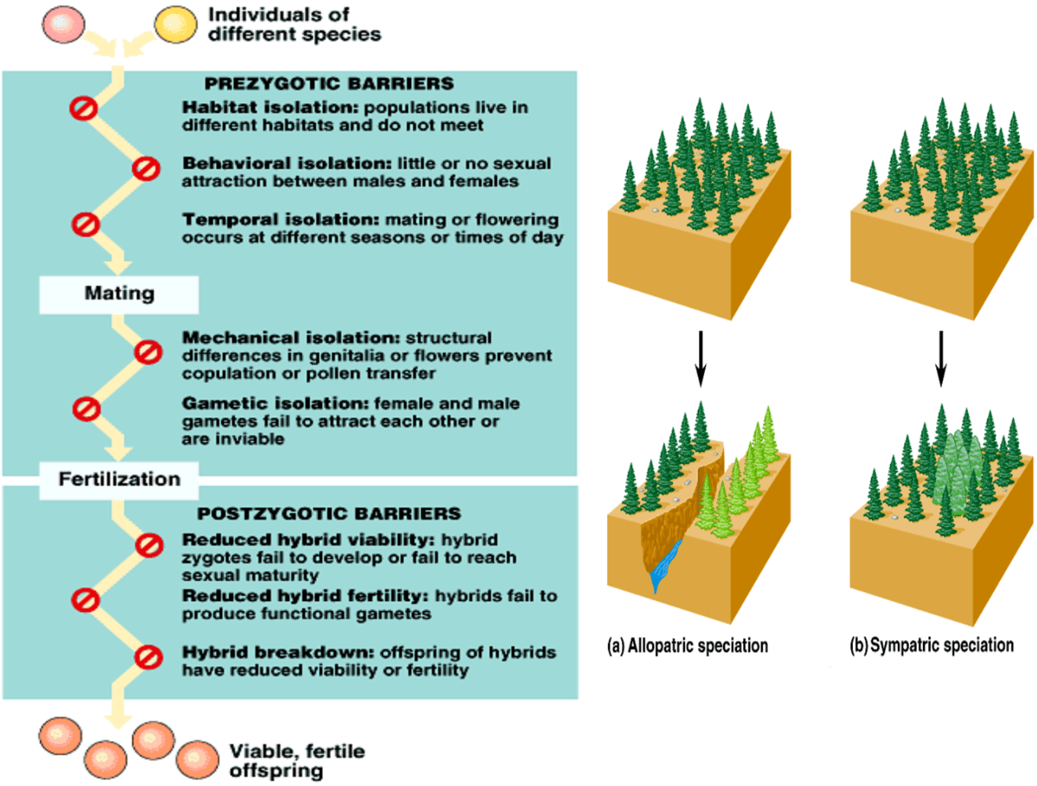


Figure 1.6: *On the left-hand side, the main barriers which may inhibit genetic interchange between species are shown, while on the right-hand side two forms of speciation, allopatric and sympatric, are illustrated. While allopatric speciation is driven by habitat isolation, sympatric speciation also occurs without geographical barriers. An example for sympatric speciation would be a population which develops* different *survival strategies in response to environmental changes. This strategies can also be thought of as alternative solutions to the same problem. As a consequence, the genetic interchange may dry out and speciation can occur. Both figures are taken from [35]. For further details, refer to* section 1.3

view, the main condition which needs to be fulfilled for speciation to occur is an *interruption of genetic exchange* (cf. [35]). As long as genetic material *can* be exchanged within all members of a species, the species will evolve as a whole, but no splitting of lineages will occur. If for some reason the flow of genetic material is blocked, speciation *can* take place. This blockade of genetic exchange may be cause by a variety of events, as illustrated in **figure 1.6**. However, an enumeration of all possible events leading to an interruption of the genetic flow is hardly possible. Furthermore, note that the blockade of genetic interchange represents just a necessary, but no sufficient condition of speciation, and ultimately could also be seen as a consequence, but not as a cause of divergence. In contrast to this, the notion of a fitness landscape provides one unified explanation for all cases of speciation: A splitting of lineages occurs whenever a population reaches a *border between two different basins of attraction* in the fitness landscape. In figure 1.5, this border is defined by all points which are located at the same distance from both maxima. Thus, exploiting the concept of evolution as optimization, we are able to formulate a sufficient condition for all cases of speciation, instead of a merely necessary one.

*Convergent evolution*, also referred to as *parallelism*, describes the process by which phylogenetically dissimilar species develop resemblant forms, due to similar environmental constraints. Thus, convergence explains why evolutionary *non related* species exhibit morphological similarities (analogies), as opposed to inheritance, which explains why *evolutionary related* species exhibit morphological similarities (homologies). Frequently mentioned examples include the overall-shape of dolphins and fish, bat-wings and bird-wings or the (independently-developed) spikes of hedgehogs (Europe), porcupines (North America) and spiny anteaters (Australia). In contrast to speciation, this process decreases the diversity of biological forms and thus could also be termed *morphological confluence*. The explanation for this phenomenon is quite intuitive: some biological traits, such as spikes or wings, are useful in a variety of situations encountered in different environments. Stated in other words, if unrelated species are confronted with the same problem, they may indeed come up with (almost) identical solutions. When considering figure 1.5, it becomes clear that the morphological distance between species present in the *basin of attraction of one maximum* will eventually decrease, i.e. the trajectories of initially unrelated species are confluent if both are attracted by the same optimum.

Thus, apparently the notion of natural selection as optimization does have some explanatory power, as shown by the above examples of speciation and convergence. However, when having a closer look at these examples, we see that in the case of speciation, the fitness-landscape is located in *genospace*, while in the case of convergence, it is located in *pheno-* respectively *morphospace*. In the first case, we use the *genetic notion of a species* and thus can define speciation as an interruption of genetic exchange between populations. Hence, the units of selection in this scenario are (populations of) *genes*. In the second case, we concentrated on changes in the morphospace and explained phenotypic similarity despite genetic distance by similar requirements posed by the environment. Hence, in this scenario, the units of selection are (populations of) organisms respectively of morphological individuals. Apparently, when examining the evolution (e.g. speciation and convergence) of multiple species, the question needs to be addressed what a *species* actually is.

Probably the most well-known definition of a species is the one mentioned above, which is based on *genetic exchange*. Regrettably, this definition miserably fails to capture our preconception of 'species' when confronted with *agamospecies*, i.e. species which do not require genetic exchange for reproduction. Furthermore, genetic exchange indeed occurs in nature not only within members of one and the same species, but also among members of different species. This is known as *introgressive hybridization* and can be found not only among plants, but also among animals as e.g. in the family of *Poeciliidae*. In such cases, biologist refer to a different definition of species, based on *morphological similarity*. Indeed, there currently is *no general definition of 'species' accepted and applicable in all branches of biology* (cf. e.g. [35, p. 445]). Interestingly, a similar disagreement has to be diagnosed concerning the *gene-concept* (cf. [36] for a rejection of the notion of a gene). Given that 'gene' and 'species' are apparently hard to defined, returning to the *individual organism* as the ultimate unit may seem attractive at a first glance. However, from a cellular point of view, an organism is anything but *in-dividual*, not to mention those organisms which are capable of survival *only in the context* of a group, colony or swarm. Clearly, the consequence for mayor biological fields like taxonomy, phylogeny and the theory of evolution are anything but negligible.

Thus, given that there is no consensus concerning the notion of a *gene* or a *species* (or an organism), and given that natural selection is supposed to act upon *species* or *genes*, we can conclude that there is no consensus concerning the *units of evolution* either. In the following section, we will have a more detailed look at this very issue.

### 1.4 The units of natural selection

In the previous section we saw that understanding natural selection as an optimization mechanism which maximizes fitness provides convenient explanations for phenomenons like *speciation* and *convergence*. However, this notion also requires an answer to the question *'what exactly are the units of natural selection?'*. Several possible answers appear to be plausible, such as:

1. organism
2. species
3. gene

If the *individual organism* is seen as the unit of selection, several objections have to be made. First of all, *evolution*, in contrast to *development*, is defined as the intergenerational change of traits and thus, an individual organism does not 'evolve'. It merely develops (itself). Second, when considering *one single lineage* and the changes occurring in it over time, all we can see from an evolutionist point of view is that this lineage is either successful or not. Thus, if a single lineage is taken to be a series of individual organisms, all we can say over this series is that it either died out or survived. It either has been selected or not. But it has not been 'shaped' by natural selection.

Apparently, we have to consider *populations of individuals* instead, for selection requires that there is something to chose from. We need multiple individuals in order to obtain a notion of *differential fitness*. Indeed, the very feature which is maximized during natural selection, i.e. *fitness*, has to be understood as reproductive success.

However, given that natural selection as an optimization strategy is only applicable to populations of individuals, the question immediately arises how such a population can be identified. Stated in other words, what population of individuals is a unit of selection? As a first try, the totality of all living beings could be considered as the (one and only) unit of evolution. Opposed to this position it has been claimed that the notion of differential fitness requires some sort of *comparability*. It appears problematic to claim that e.g. nightingale exhibits a higher fitness than an elephant because it has a higher number of offspring. When recalling the notion of a fitness landscape (cf. section 1.2), we can see that the shape of this landscape is apparently *species-dependent*:

elephant and nightingale have different needs and different capabilities. Hence, they can apparently not be described by the same set of features and thus, can not be located on the same fitness landscape.

As a consequence, even though the notion of *fitness as reproductive success* holds for the entire biosphere, the units of selection have to be considered as populations of individuals which exhibit similar *needs*, i.e. they are single *species*. Individuals pertaining to one and the same species are located on the same fitness landscape and thus, their evolution exhibits (retrospective) a specific *direction*, whereas the evolution of the biosphere as a whole has no direction, but merely requires a minimal reproductive success, i.e survival. Nonetheless, as pointed out in section 1.3, defining a species is an open problem of biology, which in turn is related to a specific notion of *inheritance*, namely a gene-centered one.

As mentioned in section 1.2, *heritability* is a requirement for natural selection. Given that currently, a gene-centric notion of inheritance is prevailing in biological research, a gene-centric notion of selection has been proposed (cf. ”The selfish gene” by R. Dawkins [37]). Only traits which are heritable can become subject of evolution through natural selection, and *if* the mechanism of inheritance is assumed to be genetic, genes have to be seen as the ultimate units of natural selection. Nonetheless, departing from a principled notion of inheritance as *intergenerational continuity* it becomes clear that genetic inheritance is just one out of multiple inheritance systems. Hence, a genetic notion of the units of selection becomes questionable.

Indeed, the rejection of a purely genetic notion of inheritance is the point of departure for *Developmental Systems Theory*, which aims to provide an integrated notion of the units of evolution by means of the concept of ”Developmental Systems”. Not the individual organism, nor the species, nor the genes are proposed as the units of selection, but the *”...whole matrix of resources that interacts to reconstruct that life cycle”* [38, p. 2]. This approach will be outlined in greater detail in **chapter 2**.

## 1.5 The emergence of units

In order to obtain a principled definition of the units of selection, an investigation of the principles of natural selection has to be carried out. More precisely, given that we are trying to define the units of selection (among others) based on the role they have got to play in the context of natural selection, an inquiry concerning the *limits* and the *prerequisites* of natural selection is needed. To carry out this inquiry, we will follow a series of arguments provided by *S. Kauffman* (cf. **chapter 3**).

First and foremost it has to be noticed that natural selection just by itself does not generate units, but merely changes and modifies pre-existing ones. Hence Kauffman claims that natural selection, just by itself, can not be seen as the sole source of order. A second principle, *self-organization*, is needed to explain the *emergence of units* which are further shaped and molded by natural selection. The difficulties involved in providing a mechanistic explanation for the emergence of *living* units relies, according to Kauffman, primarily in their astonishing *complexity*. Even the most primitive free living organism exhibits an incredible level of complexity when compared to purely chemical systems. Hence, proposing a model of self-organization which is capable of explaining how such a living unit could have emerged in an exclusively chemical milieu is seen as the core-challenge. This challenge is addressed by the theory of *autocatalytic sets*, which basically relies on *circles of positive feedback* among the elements of the set. Even though Kauffman concentrates on the origin of life and hence, on the emergence of the first primordial *cell*, the basic structure of his model can be seen as a general answer to the question how higher order units can emerge from the autonomous interplay of lower-level components.

Second, as we have mentioned before (cf. section 1.2), *differential fitness* and *inheritance* are prerequisites of natural selection too. Given that differential fitness is understood as differences in the reproductive success of a unit, *reproduction* can be seen the most basic capability a unit has to exhibit in order to become a unit of selection. Units which do not reproduce themselves can not evolve. Furthermore, the offspring generated needs to *resemble* the parent, i.e. the process of reproduction needs to guarantee a certain amount of *stability*. On the other hand, were the offspring identical to the parent, evolution would not be possible either, i.e. reproduction also needs to allow for a certain amount of *variability*. Addressing this issues, Kauffman proposes a specific type of Boolean networks (namely NK-networks) to illustrate how a suitable balance between variability (chaos) and stability (order) could be achieved.

Third, when re-considering the familiar notion of a fitness landscape, Kauffman points out that this landscape has to exhibit a *certain shape* for natural selection to work. More precisely, a small change of the phenotype has to correlate with a small change in reproductive fitness, and vice versa. If a small change of the phenotype causes a major change in fitness, the corresponding fitness landscape is very mountainous, consisting of pinnacles, abysms and steep peaks. On such a landscape, continuously walking uphill will get stuck in poor local optima very soon. Gradual evolution through continuous accumulation of minor improvements is not feasible on such a landscape. Using modified boolean network models, Kauffman argues that a *modular organization* is require for evolutionary units in order to guarantee a fitness landscape which exhibits the desired correlation between phenotypic change and change in fitness.

## 1.6 Cognitive units

In contrast to Developmental Systems Theory (cf. chapter 2) and Stuart Kauffman (cf. chapter 3), *Humberto Maturana* (cf. **chapter 4**) introduces *epistemological aspects* into the discussion about the units of natural selection. He considers the units of natural selection as *cognitive* units and thus, his notion of life and his notion of cognition are intimately related. Most importantly, he takes into account that this very discussion is carried out by the units which are to be understood. Hence, he acknowledges the inherent circularity of the attempt to understand cognition.

As an immediate consequence, he rejects *representationism* for this model of cognition is apparently incapable of describing itself. Representationism itself can not be understood within a representationistic framework, as it considers cognition as a linear process from reality to knowledge: based on sensual perception and the reasoning about it, an image of the objective, external reality is generated. Given that the image corresponds reality, it is considered as a *true* image. As a second objection, the problem needs to be addressed how the *correspondence* between image and reality can be assessed. Apparently, some sort of 'direct', 'reliable', 'undistorted' or 'pure' access to reality is needed to be able to carry out the required comparison between reality and image. However, if we had such an access, there would be no epistemological problem. As a third and final objection, it has to mentioned that in there is no subject of cognition in the representationist's model. Knowledge as well as cognition are treated as if they had some subject-independent existence.

Maturana in contrast emphasizes that the subject of cognition is actively taking part in the process of cognition and hence, also contributes to the construction of knowledge itself. Whereas in representationism, knowledge is acquired from reality, reality itself is know seen as unknown and hence has to be (re-)constructed from 'knowledge'. Clearly, this situation entails a radical shift of the meaning of 'reality' and 'knowledge', but does not necessarily lead to *solipsism*. Indeed, Maturana claims that solipsism arises from demanding objectivistic properties from a non-objectivistic cognitive domain: if someone demands to know if the statement 'that there is no objective truth' is objectively true, he or she does not take the statement *seriously*. Against this background, *science* has to be understood as a specific way to answer questions. Both the question (or problem) itself as well as the criteria the answer (or explanation) has to satisfy are ultimately arbitrary. Luckily, the criteria an explanation needs to fulfill to be considered as scientific are already well-established by society. Hence, the existent socio-cultural context provides the criteria scientific *knowledge* has to satisfy to be considered as such.

Against this epistemological background, Maturana attempts to provide a scientific explanation of the mechanism underlying the emergence of cognitive units. As each unit that we (the observer) perceive as such is the result of an operation that we ourselves actively make, namely a *distinction*, the key question to be addressed is: 'what kind of distinction brings forth a *cognitive* unit?'. As a matter of fact this very distinction has just been made in the previous sentence. It is the operation of distinction which brings forth an observer and hence, a cognitive unit is defined by this very capability. According to Maturana, this very lines can be understood as the emergence of the observer, i.e. they establish a distinction (definition) which generates a cognitive unit.

Given the circularity of cognition, a cognitive unit has to be seen as autonomous, for it is self-defining. Taking the identity between living and cognitive units into account, this autonomy translates into the notion of *structure-determined* systems. Such systems, which include living, cognitive units, are determined by their very own, internal structure, and not by the environment or milieu they happen to be in. The interaction among multiple such autonomous units is characterized by the term of *structural coupling*. In contrast to the structure of a system, which is given by specific, concrete components and relations among them, the *organization* of a system determines the system's *class identity*. Multiple structures can realize one and the same organization. Now, the organization of living units is defined as an *autopoietic* one, i.e. living systems are not only self-organizing, but *self-creating*, just as we saw in the above definition of the observer as a cognitive unit.

## Chapter 2

# DEVELOPMENTAL SYSTEMS THEORY

The historical background against which *Developmental Systems Theory* (DST) argues is constituted by a *gene-centered vision of evolution*, which is summarized by P. Griffiths as follows (cf. also [39]):

> *"Natural selection occurs because individuals vary, some of these variations are linked to differences in fitness, and some of those variants are heritable [...]. Because variants that are not heritable cannot play a role in natural selection, and because the mechanism of inheritance is presumed to be genetic, evolution is defined as change in gene frequencies. In the 1960s and 1970s this gene-centered vision of inheritance was extended to yield a gene-centered view of selection [...]. According to gene-selectionism, the fact that individual genes are integrated into larger units, from genetic modules to entire phenotypes, is merely a special case of the fact that the fitness of any evolutionary unit is a function of the environment in which it happens to find itself. "* [38, p. 1].

The observation that offspring typically resemble their parents is hardly anything but obvious and probably has already been made some millennia ago. However, understanding the *hereditary mechanism* underlying this resemblance is anything but trivial and the first mayor breakthrough in this respect - the discovery of the double-helix structure of the DNA by Watson and Crick - was awarded the Nobel Prize in 1962. Even today, the exact details of the molecular machinery behind inheritance are not completely understood (cf. e.g. [40]), but nonetheless the central role of DNA is beyond doubt.

In the first decades after Watson and Crick, with an ever increasing understanding of the molecular mechanics governing inheritance, the understanding of evolution itself evolved: given that the mechanism of inheritance was genetic, the *units of evolution*, upon which natural selection acts, were no longer *species* (as proposed e.g. by K. Lorenz) nor individual *organisms*, but individual *genes*. The DNA molecule itself entered center stage as the ultimate convergence point of evolutionary change, while the 'enclosing organism' became part of the *environment* acting upon the DNA and selecting the 'fit' genes.

In this gene-centered view, we are nothing but vessels for our genes, carefully 'programmed' and 'designed' *by* them and *for* them. The DNA was identified as the 'blueprint' of an organism, whose only objective is to replicate this very blueprint, for DNA which fails to replicate will eventually leave the stage. In this picture, 'life' was understood as 'DNA-Software' which was executed by the 'Organism-Hardware'. Thus, as a second consequence of a purely genetic notion of inheritance, the organism itself and its *development* are as external to its genes as any other environmental factors such as *"rainfall of predator density"* [38, p. 1]. Consequently, the study of evolution and the study of development are completely separated.

In contrast to gene-selectionism, DST emphasizes the importance of *extra-genetic inheritance*, which in turn leads to a novel concept of the *units of evolution*:

> *"The result is an account of evolution in which the fundamental unit that undergoes natural selection is neither the individual gene nor the phenotype, but the life cycle generated through the interaction of a developing organism with its environment. In our usage, the developmental system is the whole matrix of resources that interacts to reconstruct that life cycle."* [38, p. 2].

Indeed, the claim that *extra-genetic inheritance* plays a crucial role evolution can be seen as the point of departure for DST and thus, will be discussed in **section 2.1**. Given that the relevance of extra-genetic inheritance is accepted - and we do accept it as we will see - the *evolutionary units* needs to be re-defined. **Section 2.2** is dedicated to the discussion of the evolutionary units proposed by DST, namely Developmental Systems.

### 2.1 Extra-genetic inheritance

Given that the notion of extra-genetic inheritance is a core element of Developmental Systems Theory, we have dedicated this section to arguments supporting this very view.

Departing from a principled notion of inheritance as the continuity between parent and offspring, *empirical* evidence can be provided supporting the existence of extra-genetic inheritance. Stated less formally, we claim that it can literately be seen that there is more than nuclear DNA being passed on from generation to generation (cf. **subsection 2.1.1**).

Furthermore, when concentrating on the notion of an evolutionary unit, we have to realize that the mechanisms of natural selection are ignorant with respect to the underlying mechanism of inheritance. Thus, from an evolutionary point of view, we can define 'inheritance' as a mechanism which establishes a specific relation between evolutionary units. The exact properties of this relation, namely a specific balance between stability and variability, are dictated by the principles of natural selection. Hence, assuming more than just one single, gene-based inheritance system appears reasonable (cf. **subsection 2.1.2**).

Finally, some authors are willing to accept that 'in a certain sense', extra-genetic inheritance exists, but argue that its *relevance* is a very limited one, when compared to genetic inheritance. In **subsection 2.1.3**, we will have a brief look at two common arguments, which try to reduced the relevance of extra-genetic inheritance. The first of this arguments is based on the notion of some 'developmental information' conveyed by DNA, in contrast to non-genetic inheritance mechanisms. More precisely, the concept of 'information' is employed to demonstrate that genes transmit information, while other developmental resources do not. We will counter this argument by showing that the concept of 'information' is *in general* inappropriate to talk about inheritance. Furthermore, it also has been argued that only genes have *unlimited evolutionary potential*, and thus, are the 'most relevant' heritable resource. Again, we claim that a gene which is never transcribed due to a lack of extra-genetic resources has not potential at all and thus, can hardly be seen as more relevant than any other, equally necessary resource.

### 2.1.1 More than nuclear DNA

To begin with, the trivial observation that a piece of DNA placed in a petri dish hardly suffices to 'produce' a living organism may be illustrative. Obviously, some more ingredients are needed to provide a viable package of resources which is able to give rise to a new individual. And these are exactly the extra-genetic resources this section is about, i.e.:

> *"What does an organism inherit? Certainly more than nuclear DNA."* [41, p. 195].

This claim for the existence of extra-genetic inheritance can be empirically verified on a variety of levels. First, we can observe that DNA is *always* embedded into the astonishingly complex machinery of a *cell*[1] and also passed on *only* within this environment. Outside of this protected environment, the fragile DNA molecule is neither stable nor can it replicate itself. Thus, apart from a variety of membranes and organelles which enable the cell to maintain precisely those chemical conditions which allow DNA to persist, the whole transcriptional machinery translating DNA to mRNA and mRNA to amino acid sequences needs to be present as well.

On a second level, we can observe that the presence of cellular resources is not yet enough:

> *"But unpacking the inherited resources in the cell is not the end of unpacking inheritance. In multi-cellular organisms the parental generation typically contributes extra-cellular resources."* [41, p. 195].

The range of examples illustrating the contribution of *extra-cellular resources* to a growing organism by its parents is almost unlimited: In viviparous organisms, the conditions provided by the womb are essential for the normal development of the offspring. But also for oviparous species, it holds that an egg positioned at a randomly chosen spot will hardly be able to survive, let alone all those oviparous species with extensive parental care. In a huge variety of (typically rather complex) organisms, parental care continues *after birth*. Furthermore, in most *social life-forms*, the resources needed for the development of an offspring have to be provided by the *population as a whole*, as a single parent is not able to do so. Stated in other words, some organisms (including

---

[1] At least in all free living life-forms, which excludes viruses. However, the limitations of 'naked DNA' are particularly obvious in viruses, as they need to rely on very specific 'foreign' resources in order to generate offspring. This intimate host-virus relationships and their well-known co-evolution nicely demonstrates the relevance of extra-genetic resources to virus evolution.

homo sapiens) are able to develop and survive *only* within the protective context of a group (herd, swarm, etc.).

But this is not yet the end of extra-genetic inheritance:

> *"Even after the resources created by the population as a whole are added in, a range of other factors must be present before the sum of the available resources adds up to a viable package. [...] While the evolving lineage cannot make these resources, it can still make them part of its developmental system."* [41, p. 196].

In the above quote, P. Griffiths refers to *environmental resources* which are 'independent'[2] of the species itself, like climatic conditions or soil quality. Given that such resources can not be produced by the species itself, even though they constitute a vital factor for it, the species needs to position itself on favorable spots where its requirements are met. Thus these resources become a (presumably constant) part of the developmental system.

Recapitulating, it can be said that extra-genetic resources frequently need to be *inherited* on multiple levels in order to generate viable offspring:

1. cellular resources
2. extra-cellular resources, including parental care
3. social resources, i.e. resources provided by the population as whole
4. environmental resources, i.e. resources which can not be generated by the species

Hardly anybody would question the relevance of these resources for the *development* and the life-cycle of an organism. Nonetheless, few people would say that these resources are inherited and thus, are subject to *evolution*:

> *"A traditional way to privilege genes over other causes in development is to argue that genes are the only things organisms inherit from their ancestors. Hence the biological nature of organisms must be in the genes. DST insists on a definition of inheritance that explicitly recognizes the wide range of resources that are "passed on" and are thus available to reconstruct the organism's life cycle."* [42, p. 3].

I suspect that one of the reasons for this controversy may be the following one: the *mechanism* of inheritance for genetic and extra-genetic resources is completely different. Taking into account that deciphering the molecular basis of genetic inheritance actually was a major scientific breakthrough, it becomes plausible that the notion of inheritance itself became related to the mechanism of DNA replication and variation. However, DST does not negate the relevance of genetic inheritance, but rather claims

---

[2] We are conscious of the difficulties involved in the notion of 'independent resources'. This very difficulties will be discussed in section 2.2.1

that the term 'inheritance' should be defined based on its *role* in evolution, and not based on some specific *mechanism* (type, form, etc.) of inheritance. Stated in other words, instead of interpreting natural selection based on an exclusively genetic notion of inheritance, the theory of natural selection should provide us with a generic understanding of inheritance:

> *"Many non-genetic resources are reliably passed on across the generations. Variations in these resources can be passed on, causing changes in the life cycle of the next generation [...]. DST applies the concept of inheritance to any resource that is reliably present in successive generations, ..."* [38, p. 3]

Thus, heritability is now defined as a *resource* which is *reliably passed on* or which is *reliably present* in successive generations. As such, it may also exhibit a certain *variability* and thus may lead to *differential fitness*.

However, this notion of inheritance poses the question what a *resource* actually is, or, more precisely, how can we know which resources are actually relevant for a specific Developmental System, and which are not? Apparently, this question directly relates to the quest for a definition of the *units of evolution*, which is discussed in section 2.2.

### 2.1.2 Natural selection

As we have seen, the mechanism of natural selection is not limited to any specific domain. Indeed, it represents a very generic optimization strategy which does not depend upon some specific mechanism of inheritance:

> *"It is important to note a certain generality in the principles [of natural selection]. No particular mechanism of inheritance is specified, but only a correlation in fitness between parent and offspring. The population would evolve whether the correlation between parent and offspring arose from Mendelian, cytoplasmic, or cultural inheritance."* [31, p. 1]

Consequently, as pointed out before, natural selection provides no clues what so ever with respect to the *mechanism* of inheritance, but rather requires that, over the course of the generations, the units of evolutions (traits, properties, organisms, ...) maintain a certain balance between *stability* and *variability*. Clearly, without a certain amount of stability, we hardly would speak of a unit which somehow remains the same over the course of time, i.e. an completely unstable unit would be no (evolutionary) unit at all. For what concerns variability, it represents the possibility respectively the ability to change, and thus, to evolve. As outlined in the previous section, these two properties are not only met by genetic inheritance. Note that, if inheritance is used to refer to the resemblance between parents and offspring, it is roughly equivalent to our notion of stability.

Thus, as an immediate consequence of this extended notion of inheritance as a *balance between stability and variability* it follows that ...

> *"... the power of selective explanations need not be limited to genetic changes. The range of phenomena that can be given selective explanation should be expanded to include differences dependent upon chromatin marking systems [..], prions [..], dietary cues in maternal milk, cultural traditions and ecological inheritance [..]."* [38, p. 4]

As we can see, we are clearly transcending the traditional notion of natural selection by claiming that also *cultural traditions* are - in principle - subject to the very same laws that govern the evolution of life. Even though we basically accept this claim as the logical consequence of natural selection as a generic optimization strategy, caution is needed: Can the emergence respectively extinction of cultural traditions be explained via their relevance for the 'fitness' of the 'carrier population'? If fitness is understood as reproduction rate, apparently not, given that cultural traditions, like religion, clothing habits or language, are not only passed on from parents to children ('inherited'), but are also directly interchanged among individuals or groups.

Thus, it might be useful considering the possibility of multiple interwoven systems of inheritance and evolution. This approach would allow to distinguish between e.g. 'cultural', 'biological' and also 'genetic' inheritance resp. evolution. However, this approach would also *require* to delimitate where one evolutionary systems ends and where another one begins. Stated in other words, if evolution is assumed to take place on multiple intertwined 'levels' or 'domains', we are confronted with the very objective of this inquiry: to develop a proper understanding of the units of evolution.

Now, when trying to identify the units of evolutions, it is instructive to have one more look at gene-centrism: assuming exclusively one specific mechanism of inheritance apparently leads to one specific notion of evolutionary units. If genes were the only thing which is reliably passed on from parents to offspring, they would be the only starting point for natural selection to do its job. Thus, gene-centrism provides us with a hint on how to tackle the problem of evolutionary units: they can possibly be defined via the identification of a *mechanism of inheritance*. In a certain sense, defining the mechanism of inheritance is equivalent to defining the units of evolution.

Summarizing, even though we recognize that the mechanism of natural selection is very generic and thus, can not be limited to genetic inheritance, two challenges need to be confronted if selective explanation shall be applied: a.) the *direction* of evolution needs to be identified, i.e. the suitability of 'fitness as reproductive success' has to be proven; b.) the *units* of evolution need to be identified. Clearly, these tasks are interrelated.

### 2.1.3 The relevance of extra-genetic inheritance

Even if the presence of extra-genetic inheritance is accepted, a variety of attempts to reduce its relevance exist. Basically, these attempts try to show that, in a certain sense, genes are the 'core', the 'center' or the 'most important part' of evolution. Within the following paragraphs, we will outline and reject the most frequent arguments used to diminish the role of extra-genetic inheritance.

### Only genes transmit information

Genes are frequently believed to contain a 'program' which is executed during development. They somehow contain the 'body plan' of an organism which is simply 'implemented' or 'unfolded' during an organism's life cycle. Traits of an organism are seen as 'encoded' by the DNA and thus, ultimately, the whole phenotype is just a mere 'reflection' or 'manifestation' of the *information* present in the genes (cf. e.g. [43, p. 35] for such a notion of information).

Now, the most natural way to investigate the idea that DNA represents some sort of information is to use *information theory*. Its mathematical formulation, as provided by Shannon and Weaver in 1949 (cf. [44]), is situated in the context of a sender-channel-receiver model. A sender transmits information via some channel to a receiver if the states of both the receiver and the sender are correlated. Just to give an example: if somebody is reading this lines without changing his or her (cognitive) state, no information has been transmitted.

When applying these concepts of information theory to biology, the *genes* (DNA) are typically taken to be the *sender*, while all the *developmental resources* needed to 'transmit' the DNA-information are seen as the *channel*. Finally, the life cycle of the organism represents the *signal* (respectively the *receiver*) itself. So far, so good. But an important aspect of information theory is the role of the *channel* and the *source*:

> *"... in information theory, the role of the source and channel condition can be reversed. A source is simply one channel condition whose current state the signal is being used to investigate."* [42, p. 5]

This basically means that the media - the channel, the context - which is used to transport or transmit a certain message is as important for the meaning of the message as the message itself. From the point of view of information theory, channel and sender can not be distinguished. Their roles can be interchanged. Applied to the above interpretation of DNA as the sender which transmits information via the channel of developmental resources, we have to recognize that the later are as much a source of information as the former. Stated in other words, the exclusive status of genes as a carrier of (some sort of) information can not be maintained in the light of information theory. Nonetheless, the apparent differences which exist between developmental resources as diverse as parental care and DNA shall not be neglected. But:

> *"The point is that these empirical differences between the role of DNA and that of other inherited developmental factors do not imply the metaphysical distinction between form and matter that is often inferred from them."* [38, p. 6]

Indeed, it is perfectly normal that in an experiment *either* the environmental conditions *or* the genotype, as e.g. in twin research, is held constant while varying the other factor. Thus, the actual contribution of each of this sources to the phenotype can be investigated. Summarizing it can be said that, ...

> *"...while many concepts of information can be applied to the role of genes in development, it appears unlikely that any of these captures the intuition that genes supply information and other developmental causes do not."* [38, p. 6]

**Only genes have unlimited evolutionary potential**

John Maynard introduces the distinction between 'limited' and 'unlimited' inheritance systems in [45]. It is claimed that the genetic inheritance system has an unlimited evolutionary potential, in contrast to other systems of inheritance:

> *"Most non-genetic inheritance systems [...] can only mutate between a limited number of states. In contrast, [...] genome and language both have recursive, hierarchical structures, and hence an indefinite number of possible heritable states."* [41, p. 200]

While DST indeed agrees with this observation, the relevance of this distinction should not be over-estimated for a number of reasons:

> *"Since the genome represents only a part of the entire developmental ensemble, it cannot by itself contain or cause the form that results. But then, neither can its surroundings."* [46, p. 19]

As already mentioned (cf. subsection 2.1.1), the DNA just by itself can not be seen as an inheritance system. Therefore, even if the genetic inheritance system bears an unlimited potential, it still depends upon the presence of limited, non-genetic inheritance to unfold this potential. As an immediate consequence of the interdependence of multiple inheritance systems, we find that:

> *"Adding one form of inheritance to another causes a multiplication of evolutionary possibilities, not just an addition to them."* [41, p. 201]

Thus, even though genetic inheritance exhibits a unique structure, it can hardly be seen as dominating other forms of inheritance. Stated in other words, non-genetic inheritance does not provide some sort of immutable basis, upon which the highly-potential genetic inheritance takes place, but rather a full-valued partner, i.e. both inheritance systems are equally subject to natural selection. Finally, it has to be noticed that the unlimited character of genetic inheritance is, in first place, visible on the *molecular level:*

> *"Second, from a selectionist viewpoint the combinatorial richness of an inheritance system must be measured in terms of the number of different phenotypic effects, not just the number of combinations of components"* [38, p. 7]

When it comes to the potential of an inheritance system, what ultimately matters is not some numerical or structural property of the mechanism of inheritance, but its relevance in the course of natural selection, i.e. its potential on the phenotypic level.

## 2.2 The units of evolution

In order to apply the mechanisms of natural selection, we apparently need to establish a boundary between some *(evolutionary) unit* and an *environment* which forces this unit to evolve. When tracing the historical development of the idea of natural selection, the following three different notions of an *evolutionary unit* can be identified:

1. Genotype or DNA (Gene-Centrism)
2. Phenotype or Organism (Traditional Darwinism)
3. Ecosystem, including Geno- and Phenotype (DST)

As outlined in the previous section, DST clearly favors the third option, which will be further discussed in the following subsections. A possible illustration of these three alternatives is depicted in **figure 2.1**. Please note that the term 'ecosystem' might be misleading, given that a Developmental System is defined as the whole matrix of resources needed to reliably reconstruct the life cycle of the system. However, in any case a Developmental Systems definitively *extends* the traditional notion of a phenotype and thus, intuitively resembles what is commonly considered an ecosystem *centered around* a species.

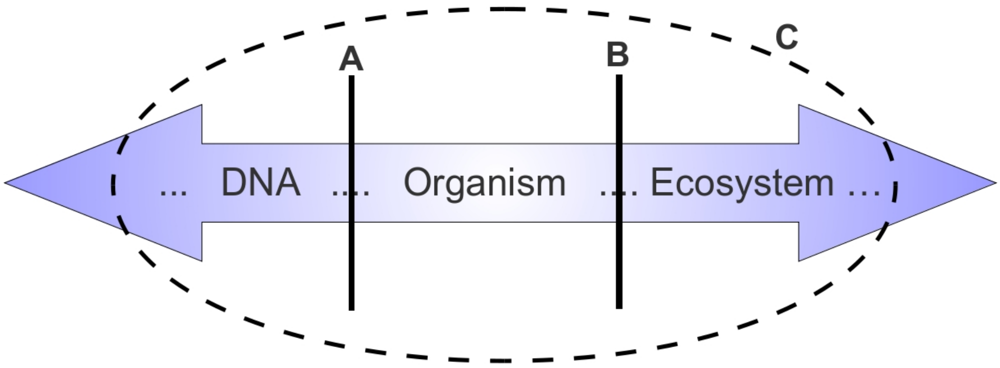


Figure 2.1: *Three alternatives to establish unit-environment boundaries: Darwin identified the individual organism as the unit upon which the environment acts (boundary B), while according to gene-centrism, both the organism (including its development) and the ecosystem exert selective pressure upon the genes (boundary A). Based on an extended notion of inheritance, Developmental Systems Theory proposes the totality of developmental resources (DNA, Organism, Ecosystem) as the fundamental unit of evolution (boundary C). For further details, refer to section 2.2*

In the following **subsection 2.2.1**, we will have a brief look at the change of the concept of an evolutionary unit on a rather formal level, i.e. neglecting any empirical aspects. We will see that DST, in a certain sense, dissolves the unit of evolution by merging it into the environment. Even thought there are indeed strong arguments favoring this notion of a somewhat 'dissolved' or 'open' unit, a series of difficulties arise as well, which will be examined in **subsection 2.2.2**. In some respect, tackling this difficulties means *reconstructing* the units de novo by re-introducing a system-environment boundary. Finally, in **subsection 2.2.3**, we will outline the concepts proposed by DST to model the *internal structure* of evolutionary units.

### 2.2.1 Dissolving units

Informally stated, the evolution of the notion of a Developmental System (DS) is driven by the insight that in nature, hardly anything is completely independent and hardly any interaction is purely uni-directional. This intuitive view on DS is roughly sketched in **figure 2.2**. A slightly more abstract and thus, possibly less intuitive view on DST is shown in **figure 2.3**. Please note that the basic concepts and their relations are identical in both figures, i.e. they represent alternative yet equivalent ways to illustrate three different notions of a functional *unit*.

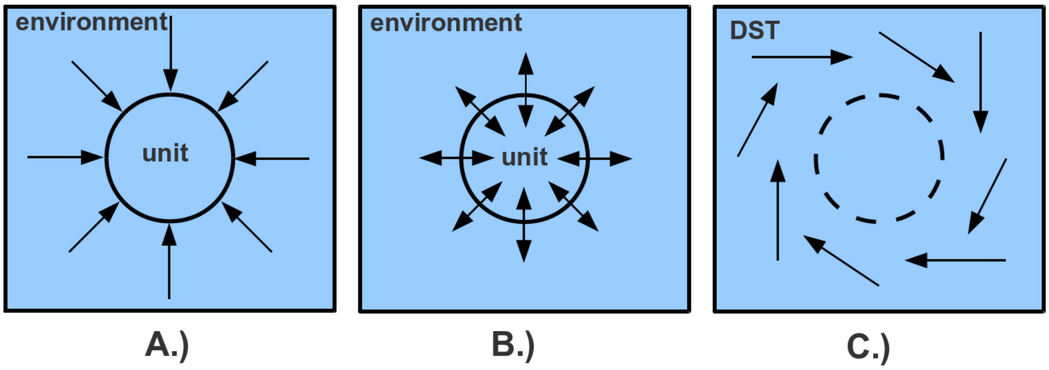


Figure 2.2: *Three different views on evolution: A.) an evolutionary unit is forced to adapt to an independent environment, B.) an evolutionary unit and its environment are co-evolving and C.) the strict boundary between unit and environment has been dissolved and the resulting Developmental System (DST) evolves according to its own internal dynamics. For details, refer to section 2.2.1*

Metaphorically speaking, if the amount of *bi-directional dependence* present in evolution is increased, we obtain a series of different scenarios corresponding to the subfigures A.) - C.) in figure 2.2 and 2.3. In the following paragraphs, we will have a more detailed look at each of theses scenarios.

#### Units are shaped by the environment

This notion of evolution has be characterized by R. Lewontin as the 'lock and key' model of adaptation (cf. also [47, 48]):

> *"... the model of the world in which there is an outside force, the pre-existen environment, that dictates the "problems" organisms must solve, and inside forces of variation that generate the organisms' "solutions" to the "problems". Organisms map the autonomous external changes in the world."* [49, p.100]

It corresponds to the concept of an organism-independent environment which defines a fitness landscape, on which the organisms are traveling. P. Griffiths [38, p. 10] called this concept *"traditional neo-Darwinism"* and pointed out that even within this framework, change in the environment can take place, but not as a consequence of the state of the organism in it. Likewise, the fitness landscape can change, but, once more, this change is not drive by the evolution of the organism. From a formal point of view, the change of the environment over time is exclusively a function of the (previous state of) the environment itself. A thorough discussion of this model, which translates into the traditional notion of *adaptation* has been provided by Godfrey-Smith, cf. e.g. [50, 51].

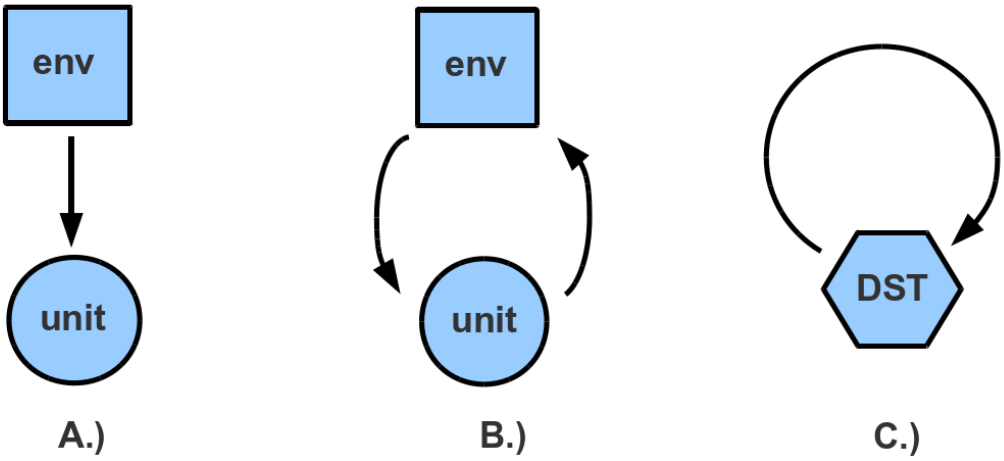


Figure 2.3: *Illustration of three possible environment ("env") - unit relationships: Traditionally, evolutionary units are seen as determined and shaped by their environment (A). If feedback mechanisms between unit and environment are taken into account, a situation of mutual dependence is obtained (B). If a clear distinction between unit and environment is rejected as in DST, a self-referential 'unit' is the consequence (C). For details, refer to section 2.2.1.*

### Co-evolution of units and environment

The model of *niche-construction* has been introduce by R. Lewontin to emphasize the interplay between environment and organism:

> *"... through niche construction organisms not only shape the nature of their world, but also in part determine the selection pressures to which they and their descendants are exposed."* [52, p. 117]

This notion of evolution incorporates *feedback loops* between unit and environment, as present e.g. in predator-prey scenarios. As a consequence, the *shape of the fitness landscape* is no longer independent of the position of the organism on it. The landscape changes while the organism travels on it, even though it remains unclear what type of change occurs. Metaphorically speaking, valleys may appear in place of steep mountains, and summits may rise in the middle of low-lands. 'Walking uphill', i.e. changing oneself, is no longer enough; it is rather the environment that needs to be manipulated such that the own current position on the fitness landscape becomes an optimum. Prominent examples of such feedback mechanisms are all those species which were *to successful to survive*, i.e. which affected their environment in such a way that the resources needed for their own survival disappeared[3]. But also the current situation of homo sapiens has to be characterized as rather changing the environment than changing itself.

Indeed, we have to notice that the *roles of environment and unit are interchangeable* in this picture (cf. subfigure B in 2.3): from the 'outside', an observer is able to identify two interacting components, but he or she can not tell which is the unit, and which is the environment.

#### Developmental Systems

Based on the arguments in favor of a.) extra-genetic inheritance and b.) the presence of a strong interdependency (feedback) between unit and environment, a central claim of Developmental Systems Theory is obtained:

> *"Since we claim that there is no distinction between organism and environment, where do evolutionary pressures on the developmental system come from? [...] everything that impinges on the process is an element of the system itself. [...] all change in the system must be endogenously driven ... "* [38, p. 12]

As expressed in the above quote, the change over time of a Developmental System can no longer be described as some sort of relation or interaction between unit and environment. As a consequence, the central evolutionary function of the environment, i.e. the selection of *fit* individuals, needs to be integrated into the Developmental System itself. It is no longer the environment, that decides which units are fit and which are not; it is rather the DS itself. From a formal point of view, this may appear a bit odd, given that DS are now supposed to carry out some sort of 'self-evaluation'. They should, by themselves, exert evolutionary pressure on themselves. Thus, in a certain sense, evolutions defines and solves exclusively its very own problems.

### 2.2.2 Reconstructing units

Claiming that there is no distinction between organism and environments is definitely inspiring but, on the other hand, also posses a series of challenges. In **figure 2.4**, three problems, which are clearly not independent from each other, are exemplarily depicted.

First, a naive person might want to know how to identify a DS as such when standing right in front of one. Or, may be even more illustrative, one might want to know

---

[3] According to J. Diamond, the population of the Easter Island apparently managed to reach this very situation 3 centuries ago [53].

how to tell that something is not (part of) a DS. In other words, defining something without telling where it ends appears to be no definition at all. Thus, we would like to know what is *outside* of a DS.

Second, we are apparently surrounded by *more than just one single* DS [4]. Thus, it would be nice to know where one DS begins and another one ends.

Third, can we tell something about the *internal structure* of DS? Stated in other words, there are quite a few organisms resp. DSs, which exhibit an astonishing amount of structure and organization within themselves. We feel that this fact should be addressed somehow by a general framework of evolution and development, as presumably presented in Developmental Systems Theory.

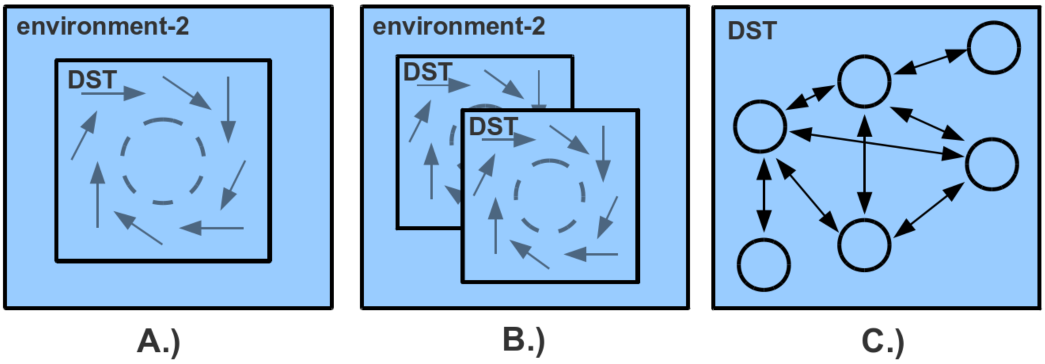


Figure 2.4: *Based on the claim of DST that there is no distinction between environment and organism, a series of questions arise which are visualized in the above figures. (A) What is* outside *of a DS? Are not we standing in need of some notion of the* environment *of a DS? (B) There are, apparently, multiple DS. How can they be distinguished and how are they related to each other? (C) Finally, what about the internal structure of DS? Please note that this figure rather serves to illustrate a possible view on DS, than to precisely define what problems may arise from DS theory. For details, refer to section 2.2.2*

The following paragraphs are dedicated to a discussion of these three difficulties: namely the problem of the *outside*, of *multiple DS* and of the *internal structure* of Developmental Systems. In a certain sense, all these difficulties seem to force us to *reconstruct* Developmental Systems as units which are clearly separated from *their* environment.

### The environment of Developmental Systems

From a *pragmatic-epistemologic* point of view we have to ask ourselves: What have we *gained* by replacing the traditional notion of an organism with the novel concept of a Developmental System *if we can not tell where it ends*? A possible reply could

---

[4] J. E. Lovelock has pointed out that the whole biosphere could indeed be considered as one single organism, taking its high level of interdependence and feedback into account (cf. [54]). However, in the given context we will stick to a somewhat more common notion of biological units respectively DS, in order to maintain the ability to view e.g. a dog or a cat as individual units.

be that our previous notion of an organism was 'simply wrong', as it was based on an *arbitrary* distinction between unit (organism) and environment. In this sense, we agree that the notion of a Developmental System is driven by valid arguments and thus, it can indeed be considered as an improvement or as a gain. On the other hand, the living world has lost its contours completely and became a rather homogeneous soup of unlimited Developmental System(s). Apparently, we have *lost* an understanding of the (living) world as a structured whole in which units and entities can be distinguished. Furthermore, it appears difficult to *empirically specify* a DS, for not a single example could be found in the literature reviewed by the author.

Furthermore, when considering the whole matrix of resources relevant to a DS, we have to notice that some of these resources are more important than others. However, the relevance of a resource is apparently no property of the resource itself, but rather determined by the function it fulfills for respectively within the DS. Informally stated, some resources are of vital importance to a DS, while other are completely irrelevant. Odling-Smee (cf. e.g.[52]) incorporated this observation into his framework, and, consequently, re-introduce the notion of organism-independent environment. More precisely, he assigned ...

> *"... separate roles to the environment of a particular lineage of organisms and what he calls the 'universal physical environment'. The former, organism-referent description of the environment is the source of evolutionary pressures on that organism, and the organism is the source of niche-constructing forces on that environment. The later, the universal physical environment, is a source of exogenous change in the organism's environment."* [41, p. 205,206]

Thus, *organism* and *organism-referent environment* are interdependent, just as they are in the concept of Developmental Systems. The *universal physical environment*, in contrast, is independent of the organism while driving *exogenous* change. In this scenario, the border between 'unit' and 'environment' is defined via the units ability to modify the environment, i.e. the universal physical environment can not be changed by the organism per definitionem. However, just as the organism-referent environment depends upon the organism, one might claim that the *relevance* of an aspect of the universal physical environment is organism-dependent too. Therefore, even though an organism may not be able to (actively) change the universal physical environment, it still can be able to 'decide' upon the 'meaning' of it: An increase in rainfall can (presumably) not be changed by a tree, but if this increase is positive, negative or simply completely irrelevant - 'invisible' - for the tree is not a property of the rainfall itself. Thus, assuming that the rainfall-increase is of no relevance for the trees well-being, it can hardly be seen as the 'source of exogenous change'. More importantly, the *universal* physical environment can apparently be a different one for each organism resp. each DS. Consequently, within the universal physical environment, a visible (or 'relevant') and a invisible (or 'irrelevant') part can be distinguished. *Both* parts are organism respectively DS dependent.

Finally, if no distinction between unit and environment is made, *the fitness landscape* behaves identical as in the case of co-evolution: it is reacting to the movements of the DS on it. In other words, the fitness of a DS is now *completely* dependent on the position of the DS in morphospace. This may appear good news, but indeed, it is not: without pressure, there is no change. At first sight, it remains unclear, why such a system should change at all. And as we all know, biological systems do indeed change. If we still want to claim that, during the course of evolution something is 'optimized', we need to provide a radical new meaning for the term 'fitness'. At a first glance, it may be tempting to replace 'fitness' with 'overall energy', as we know that according to the second law of thermodynamics chemical systems will, by themselves, eventually assume a low-energy equilibrium state. However, this sort of evolutionary pressure would be an *external* one, as it precisely assigns, to each state of the DS, a certain amount of 'energy'. In contrast to this, internal pressure can not be defined a priori which means that the very assignment of energy (fitness,...) to specific states of the system is impossible.

Stated in others words, the problem here is not that free living systems are far away from a low-energy equilibrium state, and thus, their dynamic can not be understood as driven by energy-minimization. The problem is rather that the definition of *any* objective function, no matter if we label it fitness, energy or something else, appears to be impossible without an environment providing an outside evaluation-criterion.

Thus, *a boundary between unit and environment needs to be established*, which allows for external pressure. Furthermore, this boundary should not be based on the criterion of 'complete independence' between unit and environment, in order to avoid the problems we saw above.

An interesting alternative approach to define the environment-unit boundary based on the immune-response of an organism can be found in [55], but given that this criterion is based on the presence of an immune system, it seems to transfer the problem rather than solving it: instead of defining the unit, we now stand in need of defining what an immune system is. In the case of primitive organism, this might be a rather tedious task.

### Distinguishing multiple Developmental Systems

First of all notice that, once we are talking about multiple DS, we implicitly accept the notion of a DS-dependent environment. Trivially, for each DS, all the other DSs belong to the environment. Consequently, if multiple Developmental Systems are considered, the notion of *differential* fitness re-gains plausibility:

> *"According to Darwin's theory, in a world in which there are interacting entities with the properties of multiplication, heredity, and heritable variation that affects the chances of multiplication, natural selection will necessarily occur, ..."* [56, p. 9]

This notion of fitness as the differing capacity of multiplication respectively reproduction appears to be problematic, given that the distinction between unit and non-unit (i.e. environment) is, by itself, difficult. More specific, as a prerequisite of reproduction, there has to be some unit or entity that can be reproduced:

> *"According to DST an evolutionary individual is one cycle of a complete developmental process - a life cycle. [...] Developmental systems include much that is outside the traditional phenotype. This raises the question of where one developmental system and one life cycle ends and the next begins."* [38, p. 14]

Note that, when discussing multiple DS, we may also be talking about parent and offspring: as long as both rely on the same resource, their DSs are apparently 'overlapping'. However, in order to talk about overlapping DS, we need a notion of the internal structure of Developmental Systems.

### The internal components of Developmental Systems

As we already know, Developmental Systems embrace the whole matrix of *resources* need for their reconstruction. Thus, there is no single specific component which needs to be present in any DS. Furthermore, if something is a resource for a given DS or not is determined by the DS. In other words, 'being a resource' is not a property, but a relation between a component and the whole system. The whole system, in turn, is exclusively constituted by its components and their interdependencies.

> *"Developmental systems theorists...define inheritance as the reliable reconstruction of interactive causal networks"* [57, p.961]

Thus, the internal structure of DS can be understood as reproducing causal networks among resources. Once more we can see the self-referential, emergent character of DS:

They are *emergent* insofar as they constitute the point-of-reference to decide upon the relevance of resources. More precisely, they decide - metaphorically speaking - if something constitutes a resource or not, if something is of any relevance to them or not. Consequently, in this sense the distinction between system-relevant 'inside' and system-irrelevant ('invisible') outside is drawn by the system itself (cf. the notion of system-referent environment in subsection 2.2.2). In this sense, DS are defining themselves and emergence can be seen as the very act of self-defining.

On the other hand, DS are also *self-referential* insofar as the whole is constituted by its parts, while its parts are constituted *as such* by the whole. Stated in other words, the parts can not be identified as units (components, resources, ...) without the whole: they do not exist as parts. I would like to emphasize this last point once more: in general, it is easy to see that the whole is constituted by its parts: if some parts are missing, the whole does not exist. In the context of DS, one would say that, if a vital resource is missing, the affected DS will fail to reproduce and thus 'die out'. For what concerns the relevance of the whole for its parts, it is less intuitive to say that they 'do not exist' if the whole is not present. However, any living organism provides a nice example of a systems whose parts - organs, tissues, cells - do not exist without the whole that they constitute.

The difficulty in seeing an equivalence between [part → whole: the parts imply the whole] and [part ← whole: parts are implied by the whole] may be the direction of determination. The relation between part and whole is typically seen as a *hierarchical* one. Furthermore, we are used to *bottom-up causation* and thus, to *top-down reductions*: explaining complex phenomenons by dissection into ever smaller and simpler parts [5]. However, in the current situation, both directions bottom-up and top-down, are completely equivalent as they can not be distinguished: *the parts are as much an emergent phenomenon as the whole*.

### 2.2.3 Structuring units

As we have seen in section 2.2.2 we stand in need of a boundary, a *distinction between unit and environment*. We have noticed that the concept of Developmental Systems, as long as it entails a 'union of environment and unit', is rather incompatible with evolution as optimization. We need an outside, thus we have to establish a conceptual framework including this very difference between 'outside' and 'inside'. Concerning this distinction, we consider the notions of *unit and environment* and *DS and environment* to be equivalent.

Furthermore, we also saw that an *internal structure* of DS's is needed to explain the differences between multiple DSs. Within the following paragraphs, some approaches focusing on the emergence of DS from causal interaction networks among resources are presented.

#### Repeated assemblies and trait groups

According to DST an evolutionary unit is 'one cycle of a complete developmental process - a life cycle' [38, p. 14]. Following this definition, the main criteria of a DS is its cyclic structure. Given that there is a wealth of cyclic phenomena and 'repeated assemblies' in nature, we apparently need to further specify what exact type of cycles are DSs and which are not. A number of similar cells integrated into an organ or a group of individuals forming a swarm are apparently repeated assemblies. However, we claim that a notion of DSs which is incapable to distinguish these types of assemblies from a traditional physiological individual would reflect a (too) wide notion of "unit". Therefore, the question is

> *"... what makes a repeated assembly a developmental system in its own right, as opposed to a part of such a system or an aggregate of several different systems."* [38, p. 14]

In the above quote, the question how to distinguish a repeated assembly and a DS is re-cast into the question of identifying the right level of granularity: defining a DS (intermediate level) means distinguishing it from its environment (upper level) as well as from its parts (lower level). This problem is well-known as the 'level of selection problem' (cf. [61, 62]), and in order to address it, Wilson and Sober (cf. [63, 61]) introduce the concept of a *trait group*. A trait group is ...

---

[5] For a discussion of possible concepts of top-down causation, as opposed to bottom-up causation, refer to [58, 59], whereas the implications of a hierarchical model of the world are discussed in [60].

> *"... a set of organisms relative to which some adaptation is, in economic terms, a public good."* [38, p. 14]

The idea behind the concept of a trait group is to ask for whom a certain trait is beneficial, and for whom it is not. Most importantly, some trait of an organism may increase the fitness not only of the organism itself, but may also be beneficial for other organisms too (cf. **Figure 2.5**).

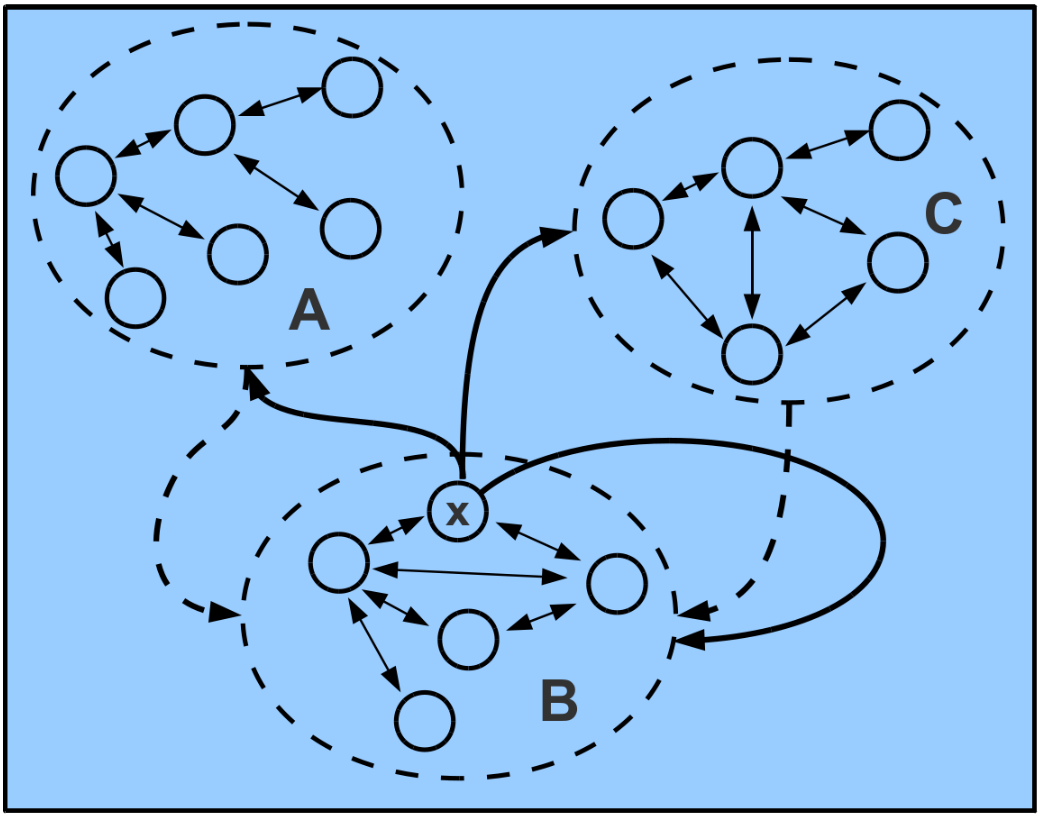


Figure 2.5: *Each organism - A, B and C - consists of several components resp. traits, which in turn are interdependent. Some specific trait* X *of an organism B may be relevant for the fitness of this organism itself, but also for the fitness of other organisms A and C. All organisms, for which a specific trait is beneficial, are forming the trait group. If the fitness of these other organisms A and C is of relevance for B's fitness, the trait X in B may be selected not because of its* direct *contribution to B's fitness, but because of its* indirect *one. For details, refer to section 2.2.3*

However, in order to characterize a trait group as an evolutionary unit, we need to add some positive feedback on the fitness-level. Stated in other words, for a trait group to become a unit of evolution, the relationship between the 'elements' (e.g. organisms) of a trait group needs to be *mutually advantageous*: For of each member of the group, increasing the fitness of its group members is the most efficient way to (indirectly) increase its own fitness. On a causal level, this sort of economic cooperation is realized via the interaction among specific traits of the involved group members. Eventually, this sort of feedback leads to functional differentiation and the dependence of individual elements on the whole group:

> *"Not every trait group is a superorganism. There are a number of features that seem to mark the difference between mere trait groups and superorganisms, such as the functional differentiation of parts and the dependence of parts on the whole for their viability."* [38, p. 14]

As a consequence of *cooperation* and mutually advantageous interactions, *competition* between the components disappears. Thus, the individual components respectively base-level evolutionary units are no longer competing for access to one and the same limited resource, but rather identified each others as resources. However, as the amount of economic cooperation present within a group of components can range from 'virtually non-existent' to 'completely interdependent', the distinction between a repeated assembly, and a DS which originated from trait group selection is not a crisp one:

> *"We suggest, then, that a repeated assembly is a developmental system in its own right, as opposed to a part of such a system or an aggregate of several different systems when specific adaptations exist, presumably due to trait group selection, which suppress competition between the separate components of the assembly. This account of the evolution of individuality can actually explain why the distinction between a colony of organisms or a symbiotic association and an individual organism is not a sharp one."* [38, p. 14]

Hence, according to this approach, evolutionary units are basically the result of economical cooperation. It is easy to see that, once such an economically positive feedback exists between a set of components, trait group selection is likely to increase the level of (economical) interdependence. Thus, as a consequence, we have to assume that Developmental Systems are highly integrated, i.e. they are formed by strongly interdependent components. Also intuitively appealing, this notion of highly integrated systems poses a serious problem to natural selection, as we will see in the following section.

### Modularity and gradualism

When thinking about a unit as opposed to the environment, the interdependence among the unit's components gains relevance: We can clearly see that an organism or parts of it are interacting with elements of the environment. Nonetheless, the organism-environment boundary is apparently established due to a somewhat stronger interaction encountered among the parts of the organism. In the previous section, we saw an approach centered around 'mutually advantageous' traits respectively components, which attempts to explain how such a highly integrated system can emerge. The system to be explained has been characterized by suppressed competition, functional differentiation and dependence of parts on the whole for their viability.

In contrast to this, R. Lewontin observed that a tight integration of the organism's components may hinder evolution. He claimed that *quasi-independence* among the parts is required for evolution through natural selection:

> *"Quasi-independence means that there is a great variety of alternative paths by which a given characteristic may change, so that some of them will allow selection to act on the characteristic without altering other characteristics of the organism in a countervailing fashion;"* [31, p. 230]

If all traits of an organism are strongly interdependent, a small change in just one of these traits may lead to a big change of the organisms overall fitness. Hence, a certain *modular* composition of the evolutionary unit, which allows for quasi-independent modifications of the individual components, is required. Otherwise a *gradual* improvement become impossible, because the situation - 'small change, big effect' - actually renders any systematic optimization procedure, including Natural Selection, infeasible. The fitness-landscape becomes a rugged mountain landscape full of pinnacles and abysms. Traveling uphill is as impossible as finding a 'reasonably good' local optima, because there are too many of them. On such a landscape, *random search* is as efficient as any other heuristic [6]. Stated in other words, as a prerequisite for any heuristic, including Natural Selection, the *locality criteria* (cf. e.g.[64]), sometimes also referred to as the *requirement for continuity* (cf. [31, 65]) needs to be satisfied. This criteria requires a certain correlation between the distances in morphospace and the differences of the associated fitness-values. Thus, if very similar body-plans exhibit very different fitnesses, the locality criteria is violated and natural selection becomes infeasible.

The notion of locality, as a formal requirement for any heuristic, needs to be distinguished from *gradualism*, as the assumption that changes in nature occur gradually, continuously in small steps which accumulate over the course of time[7]. Without locality, natural selection itself becomes infeasible and can no longer explain the course of evolution, no matter if the later is seen as a gradual process or not. With locality, natural selection becomes feasible, but the feasibility of natural selection does not imply a gradual, stepwise evolution.

Thus, the discussion about the gradual nature of morphological change, as opposed to a punctuated equilibrium scenario (cf. **Figure 2.6**), has no implications what so ever concerning the validity of natural selection as a cause of evolutionary change. However, this discussion does have an impact on the shape of the *fitness-landscapes*, *if natural selection is seen as the sole source of morphological change*:

- *Phyletic Gradualism* implies continuous evolutionary pressure which translates into an intermediate inclination of those parts of the fitness-landscape, which are relevant for the evolving species.

- In contrast to this, a *Punctuated Equilibrium* scenario, as proposed by Edredge and Gould (cf. [67, 68]), reflects a mixture of evolutionary stable phases and phases of rapid morphological change. Whereas the former (no morphological

---

[6] In computer science, a *heuristic* is any type of search strategy. Such search strategies, as Genetic Algorithms e.g., are typically used to find good solutions for formally defined problems

[7] Even though the concept of gradualism may appear almost trivial today, it has played an important role during the initial formulations of Natural Selection in the 19th century, as it implies that the differences among species - including homo sapiens - are merely gradual ones [66].

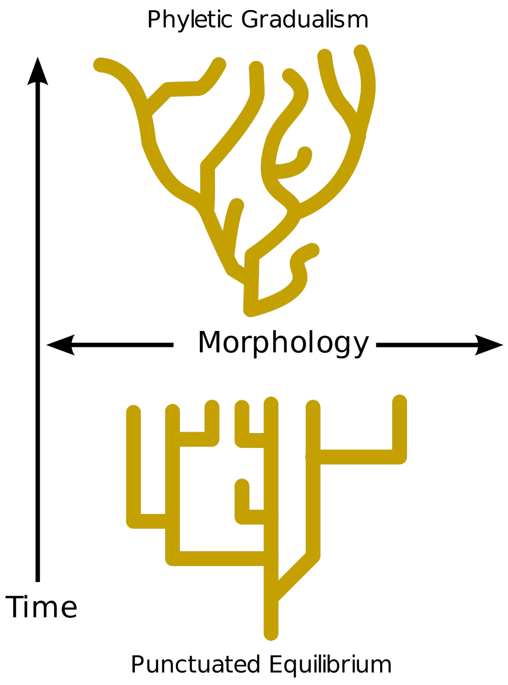


Figure 2.6: *Phyletic gradualism assumes that species are continuously changing over time, while the punctuated equilibrium scenario proposes periods without any alteration of biological form, interrupted by sequences of rapid morphological change. Both scenario are consistent with the concept of natural selection, even though the fitness landscape driving evolutionary change will exhibit a different shape depending on the chosen scenario. Figure obtained from* `http://en.wikipedia.org/wiki/Punctuated_equilibrium`*. For details, refer to section 2.2.3*

change) translates into (plateaus of) local optima, the later (rapid morphological change) corresponds to very steep inclinations.

As the final consequence, the notion of highly integrated units needs to be questioned: *The principles of natural selection require a certain level of modularity*, of independence among the traits of an evolutionary unit.

### Developmental constraints

One of the central dogmas of evolutionary biology is the correlation between functionality and conservation. Nature selects for outcome, hence what we see in todays nature exhibits a higher functionality than those species, features and genes which went extinct. In particular, those aspects of biological form which are able to withstand evolutionary pressure over prolonged periods of times have to be competitive. Otherwise, they would have been replaced. Therefore, functionality implies conservation. Conversely, if some feature of an organism is highly conserved throughout the course of evolution, it is assumed to be highly functional.

This very conclusion - conservation implies functionality - lies at the heart of numerous present day research projects in molecular biology: given that 'functionality' is a completely generic property, it can hardly be measured if the system under investigation is not yet thoroughly understood. Indeed, *understanding* the function of a component of some system usually represent the final objective of an inquiry. In turn, identifying conserved features merely requires a comparison of multiple evolutionarily related systems and thus, is a convenient point to start from.

However, this procedure has been questioned, because...

> *"...there are many highly conserved features of biological lineages that are not plausibly explained by stabilizing selection."* [69, p. 8]

In more general terms, we need to question the *extend of adaptive explanations*. Stated in other words, random variation and natural selection might be not enough to explain every aspect of the living. For what concerns the 'conservation without functionality' situation described above, the concept of *developmental constraints* has been proposed:

> *"A constraint can be defined fairly uncontentiously as a bias in the production of variation in a population"* [69, p. 8]

First of all it has to be noticed that *a constraint is, by its very definition, opposed to adaptation*. Hence, as intended, it may serve as an explanation for conserved and apparently non-functional features of an organism. Stated in other words, using the concept of developmental constraints we are also able to provide an explanation for cases of maladaptation.

However, in contrast to this notion of developmental constraints it has been argued that ...

> *"... constraints can only ever be temporary, since evolution can reconstruct the developmental system of the organism so as to achieve whatever outcome is selectively optimal."* [69, p. 9]

From this *adaptationist* point of view, a developmental constraint is not opposed to adaptation, but rather a mid-term result, which may turn out disadvantageous on the long run.

Summarizing the above discussion on developmental constraints, the following aspects need to be emphasized: First of all we saw that functionality implies conservation, but conservation does not necessarily imply functionality.

Second, the adequacy of the concept of 'developmental constraints' to analyze evolution can be questioned. Empirically, it is impossible to decide if there is a bias in the production of variations in a population, i.e if a developmental constraint 'exists'. On a purely conceptual level, however, we have to recognize that from an adaptationist point of view, there can be no non-functional features and thus, there can be no developmental constraints (as something that constrains functionality) either. Conversely, claiming that there actually are non-functional features, already requires some non-adaptationist perspective. Thus, recurring to the existence of non-functional features can not be used as an argument to support a non-adaptationist position. Likewise, 'functionality' is not an empirical category, but the assignment of a role in a readily available model. Given that the concept of a 'developmental constraint' is intimately related to some notion of functionality, the difficulties originating from the former are supposedly rooted in the later. More precisely, if 'functionality' and 'adaptation' are seen as a deeply subject-dependent construction, the problems associated with 'developmental constraints' might be alleviated.

Finally, if we consider an organism as a system made up of different modules each of which exhibits its own dynamics, evolution can be understood as a *multi-objective* *optimization* procedure. Once we acknowledge that there is more than one notion of *fitness*, more than one notion of a *unit* upon which selection acts, the concept of developmental constraints is no longer needed.

## 2.3 Summary

The point of departure for Developmental Systems Theory is the rejection of a gene-centered vision of inheritance. To support this view, a series of arguments have been outlined (cf. **section 2.1**).

First of all, *empirical evidence* indicates that nude DNA can hardly be seen as an autonomous inheritance system. Obviously, a set of additional resources is needed to allow for self-sustained reproduction. A single DNA strand is not capable of reproduction, nor is there any organism which, at any point of its life cycle, exists as such. The smallest unit during an organisms life cycle, which is traversed by *all* organisms, is a single *cell*. But to allow for normal development, frequently a series of *extra-cellular resources* needs to be present as well, not to mention *parental care* and further resources generated by the enclosing population as a whole (cf. **subsection 2.1.1**). All this resources, no matter if they are genetic, cellular, extra-cellular or provided by the whole population, have to be seen as equally essential to the unfolding of an organism's life-cycle, i.e. to its *development*. Hence, evolution and development have to be studied in an integrated fashion, for an extended notion of inheritance also questions the crisp dichotomy of ontogeny and phylogeny.

Furthermore, *natural selection* itself provides a second mayor argument in favor of multiple, including non-genetic, inheritance systems (cf. **subsection 2.1.2**). Already the observation that Darwin formulated his theory without any notion of genetic inheritance indicates that, for natural selection to work, the details of inheritance are irrelevant. Indeed, from an evolutionary point of view, the only requirements for 'inheritance' are a.) some resemblance between the units of a lineage, and b.) that this resemblance concerns features (properties, characteristics, ...) which are relevant for the unit's fitness, i.e. for it's reproductive success. Any mechanism which guarantees that this requirements are satisfied has to be considered as an inheritance system.

However, some author agree on the existence of extra-genetic inheritance but try to minimize their *relevance* (cf. **subsection 2.1.3**) by either claiming that a.) only genes transmit information or b.) that only the genetic inheritance system has an unlimited evolutionary potential. We rejected a.) by showing that using the concept of 'information' in the context of inheritance is problematic in general and b.) by pointing out that also the genetic inheritance systems depends upon further, non-genetic inheritance systems.

The first mayor conclusion to be drawn from the above discussion of multiple inheritance systems is the equivalence between *defining a unit of evolution* and *defining a mechanism of inheritance*. Apparently, if we can identify a mechanism of inheritance, we are also able to identify the units of evolution, for this very units are also the subjects of inheritance. Characteristics or properties, which are either not inherited (not passed on from one unit to the other), or which remain unchanged, can never become subject to evolution. Consequently, if the mechanism of inheritance is presumed to be exclusively genetic, the only possibly unit of evolution has to be genetic as well. As can be seen at R. Dawkins, a gene-centered vision of inheritance necessarily entails a gene-centered vision of selection. However, the insight that the question *"what are the units of evolution?"* can be faithfully translated into the question *"what is the mechanism of inheritance?"* is of absolutely no use to answer either of the questions. In both cases, a prior knowledge concerning the answer of one of the questions is required to be able to address the other one. In order to define a mechanism of inheritance, we need a prior notion of 'unit' (e.g. organism, species, ...), while in order to define the units of evolution, a prior understanding of the mechanism (e.g. genetic, cellular, cultural, ...) is required.

Consequently, accepting the presence of extra-genetic inheritance entails a rejection of gene-selectionism, but provides no further hints concerning a fruitful understanding of 'evolutionary unit' (cf. **section 2.2**). Indeed, defining a developmental system as the whole matrix of resources needed to reliably reconstruct a casual network, i.e. to complete one life cycle, may be consistent with natural selection, but apparently lacks an applicable criterion to distinguish between system and environment. Given the part-whole relationship between unit and 'required resources', identifying the required resources is impossible without a prior definition of the unit itself. Conversely, the unit itself can not be identified as long as it's parts (required resources) are not yet identified. Thus, Developmental Systems theory seems to lead to a step-wise dissolution of the boundary between unit and environment (cf. **subsection 2.2.1**). A transition can be identified from the strict dependence of the unit from the environment to a scenario of co-evolution of unit and environment. If the notion of interdependence underlying co-evolution is taken seriously, the final dissolution of the system-environment boundary, as present in a Developmental System, is the consequence.

However, there is a series of problems which can not be solved without a clear distinction between system and environment (cf. **subsection 2.2.2**). First of all, natural selection can not be understood as somewhat 'internal selection', but clearly requires an environment excerting evolutionary pressure on the unit. Furthermore, as selection is related to (the unit's) fitness, we also have to notice that fitness in this context can only be defined as differential reproductive success. Apparently, the notion of reproduction requires the notion of a unit which can be reproduced. Hence, the question remains were to draw the distinction between unit and environment. This difficulty, also known as the problem of the 'level of selection', has been addressed various times, either proposing genes (e.g. R. Dawkins), organism (e.g. C. Darwin) or species (e.g. K. Lorenz).

In order to address this problem, a series of arguments focusing on the internal structure of evolutionary units have been sketched (cf. **subsection 2.2.3**). In order to distinguish between actual evolutionary units and mere 'repeated assemblies', the concept of *trait groups* has been proposed. A trait group is a group of traits (or components, elements,...), for which a certain trait is, in economic terms, a public good. Hence, the presence of this trait is in evolutionary terms advantageous for the corresponding trait group. Now, assuming a multitude of trait groups, it is easy to imagine a scenario of *positive feedback* among a set of traits. As a consequence of this feedback, the evolutionary interdependence within this traits increases until they form a unit in it’s own right. Needless to mention that this model requires the pre-existence of some evolutionary units in the first place, namely the traits upon which selection can act. Hence it can not truly be seen as a model explaining the *emergence* of evolutionary units, but merely their *differentiation*.

Furthermore, a relative independence (called ’quasi-independence’ by R. Lewontin) among the traits of an evolutionary unit, i.e. a *modular structure*, was identified as a prerequisite for natural selection. Without such a structure, a continuous improvement is impossible as the fitness-landscape lacks the smoothness needed for stepwise optimization. Finally, the notion of a *developmental constraint* as a cause of mal-adaptation has been outlined and rejected, for extinction is seen as the only reasonable criterion to identify cases of mal-adaptation. Hence, the suitability of this concept to define evolutionary units seems to be limited.

# Chapter 3

# STUART KAUFFMAN

Stuart Kauffman is probably best known for claiming that natural selection just by itself is insufficient to explain life. *Self-organization* has to be assumed as an additional principle, which is not only involved into the emergence of life as an historical event, but which is also constantly present in all of todays life forms. In order to support this view, Kauffman heavily relies on formal models and computer simulations. Thus, he tries to provide a *mechanistic* explanation for life, i.e. he sketches a mechanism (model, simulation) capable of generating the phenomenon to be explained (life).

However, in order to truly appreciate the proposed explanation, an explicit statement of the question to be addressed is needed. Thus, in a first step, we will try to outline Kauffman's notion of the phenomenon to be explained (cf. **section 3.1**). When doing so, we have to keep in mind that the phenomenon to be explained is not, roughly stated, 'life itself', but only those aspects and characteristics of living units which can presumably not be explained by natural selection. Thus, we are not only trying to outline those aspects of life which seem to require an explanation, but also to delineate the explanatory limits of natural selection. The first thing to notice during this endeavor is the apparent *improbability* of life (cf. **subsection 3.1.1**), if life is seen as a closed thermodynamic system. Hence we seek a model which makes the emergence of living units not only possible, but also probable. For what concerns this very emergence of living units, Kauffman proposes a holistic view (cf. **subsection 3.1.2**), i.e. he claims that we seek a model which explains the origin of life as a one-step process. Finally, in **subsection 3.1.3**, we recapitulate Kauffman's observation that there seems to be a certain *minimal complexity* for living units below which life is impossible.

After summarizing the obtain problem description (cf. **section 3.2**), the models employed by Kauffman to address this problem(s) are outlined in the following subsections. In **subsection 3.2.1**, the theory of *autocatalytic sets* and a more recent extension of it (*Graded Autocatalytic Replication Domains - GARD*) is presented. This models primarily address the emergence of living units. In **subsection 3.2.2** *Boolean NK-Networks* are outlined, and their suitability to formalize the notions of 'stability' and 'variability' of living units is discussed. Finally, in **subsection 3.2.3**, the modularity of units in relation to the fitness landscape is examined using an *adjusted Boolean Network* model.

We will conclude this chapter with a summary of the central results in **section 3.3**.

## 3.1 Natural selection and self-organization

One of the central claims of Stuart Kauffmann is his rejection of natural selection as the *sole* source of order. Instead, he proposes *self-organization* as an *additional* driving force in evolution:

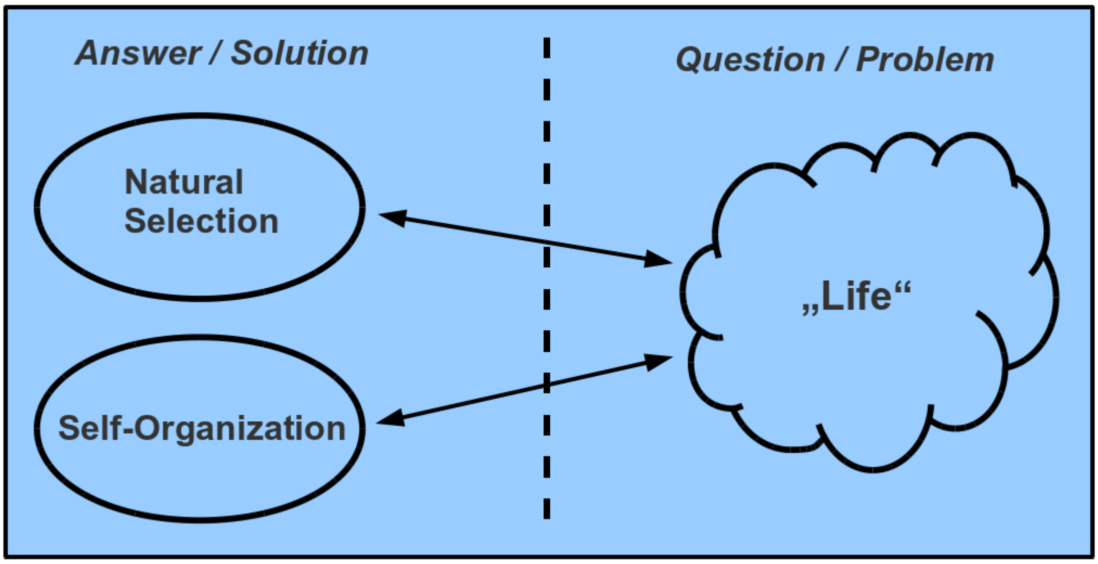


Figure 3.1: *Illustration of the 'Question-Answer Model': in order to evaluate the suitability of an answer (solution), the question (problem) needs to be formulated explicitly. Hence, in a first step we will outline Kauffman's notion of 'life' before discussing, in a second step, his understanding of self-organization. Note that the later includes clarifying the relation between natural selection and self-organization. Clearly, as can be seen in the above figure, question and answer are always interdependent: proposing a solution necessarily includes defining the problem. For details, refer to section 3.1*

> *"Since Darwin, we turn to a single, singular force, Natural Selection, [...]. Without it, we reason, there would be nothing but incoherent disorder. [...] I shall argue in this book that this idea is wrong. [...] Laws of complexity spontaneously generate much of the order of the natural world. It is only then that selection comes into play, further molding and refining. [...] this emergent order underlies not only the origin of life itself, but much of the order seen in organisms today. "* [70, p. 8]

As expressed in the above lines, the laws of complexity and self-organization are in a certain sense *prior* to natural selection, which is designated as a merely 'refining' force. However, self-organization is not seen as a single historical event, but rather as a specific process which is continuously present in all of yesterdays, todays and tomorrows life forms.

In order to fully appreciate role and relevance of self-organization, a more specific description of Kauffman's explanandum - life - is needed (cf. **Figure 3.1**)

Stuart Kauffman sees *order* as the central aspect of life, which needs to be explained. In this context, the key characteristic of order is its inherent *improbability*. Would order be something which is very likely to appear, something that we expect to arise 'naturally', then there were no reason for us to seek an explanation for it. The expected needs not be explained. Thus, in a first step we will outline in which sense life is *ordered*, and in which sense this order appears to be *improbable* (**subsection 3.1.1**).

Another aspect of life which is typically seen as non-explicable by means of natural selection is its *beginning*. Apparently, natural selection explains how life *evolves*, but fails to explain how it *emerged* from the non-living. The problem seems to be rooted in the absence of evolvable units upon which natural selection can act. Consequently, the main question around which current research is centered concerns the formation of conditions which allow natural selection to take place. In **subsection 3.1.2**, a very brief look at two theories is provided.

Finally, Kauffman points out that even the simplest free living organism today exhibits an astonishing level of molecular organization. Based on this observation, he claims that there is a certain *minimal complexity*, below which life as such can not exist. In **subsection 3.1.3**, this position and its main implications are sketched.

### 3.1.1 The improbability of order: dissipative systems

The second law of thermodynamics states that in closed physical systems differences, e.g. in temperature or chemical concentration, will eventually disappear (cf. e.g. E. Schrödinger's explanation in [71]). This decrease of differences is equivalent to an increase in *entropy*, which can be considered as a measure of the amount of disorder present in a closed system. A familiar - even though physically not truly correct - example of such a closed system would be a cup of hot coffee with some cold milk in it. As we know, the drop of milk does not remain in a precisely defined area of the cup, clearly separated from the coffee. They mix up to form what is known as 'cafe latte' and, eventually, will exhibit the same temperature. Thus, the second law does *not* state that *closed* physical systems will eventually assume a low-energy state. This is impossible due to the first law of thermodynamics (energy conservation). It rather states that for purely statistical reasons, in a closed equilibrium system, *ordered states* are rather improbable:

> *"The consequence of the second law [of thermodynamics] is that in equilibrium systems, order - the most unlikely of the arrangements - tends to disappear. If order is defined as those coarse-grained states that correspond to only a few fine-grained states [...], then at thermodynamic equilibrium, those delicate arrangements disappear because of the ergodic wandering of the system through all its microstates."* [70, p. 9,10]

Given that life is considered to be a highly ordered phenomenon typically exhibiting a variety of differences with respect to temperature, chemical potential and pressure, it seems to be contradicted by the second law of thermodynamics. Stated in other words, if life is considered as a closed equilibrium system, the second law of thermodynamics is applicable which in turn renders the ordered character of life highly improbable.

One of the first approach addressing the improbability of life while rejecting religious explanations was focused on *time*:

> *"The important point is that since the origin of life belongs in the category of at least once phenomena, time is on its side. However improbable we regard this event, or any of the steps which it involves, given enough time*

> *it will almost certainly happen at-least-once. And for life as we know it, with its capacity for growth and reproduction, once may be enough. [...] Given so much time, the "impossible" becomes possible, the possible probable, and the probable virtually certain. One has only to wait: time itself performs the miracles."* [72, p. 48]

However, even though the chances of success clearly increase if we are given more than one chance, the chances of success for one single attempt still need to be known before *"the impossible becomes possible"*. Stated in other words, knowing that we can throw the dice a hundred times is not yet enough to claim that we will 'certainly' obtain a six at least once. We also need to know the chances to obtain a six in one single throw. Transfered to our question, we need an estimate of the odds to obtain, by chance, a primitive organism like e.g. *E. coli*. Such an estimate was proposed by Hoyle and Wickramasinghe, based on the chances of a random assembly of a number of functioning enzymes. The resulting probabilities are so incredibly low, that *"even in the lifespan of the universe"*, the spontaneous assembly of life is still virtually impossible. Stated in Hoyles own words:

> *"The chance that higher life forms might have emerged in this way is comparable to the chance that a tornado sweeping through a junkyard might assemble a Boeing 747 from the materials therein."*[73]

Confronted with this conclusion, some people claim that intelligent design is the only answer. Others, like Kauffman, conclude that there must be something wrong with the conclusion: We are alive, hence, life can not be impossible. Consequently, the second law of thermodynamics may not be applied in the case of living systems. Hence, living systems do not represent closed equilibrium systems.

Instead, Kauffman characterizes living systems as *open, nonequilibrium* systems, which continuously dissipate matter and energy. However, even though this notion of life as a *dissipative system* alleviates the problems associated with the second law of thermodynamics, it still does not explain how dissipative systems emerge. In particular, our notion of life as highly *ordered* system remains untouched.

### 3.1.2 The origin of life

As illustrated in **Figure 3.2**, theories concerning the origin of life are trying to close the gap between a 'primordial soup' and the first protocell. Before going into the details, it is worth mentioning that the process we are trying to understand exhibits the following two characteristics:

a.) *It can no longer be observed:* apparently, the transition from non-living matter to living organisms is no longer taking place in nature, at least not on our planet.

b.) *It may have occurred just once:* for all we know, there is just one single common ancestor to all of todays (and yesterdays) life forms. Even though we may speculate that there were multiple pre-biotic (semi-biotic) 'species', just one of them actually became the founder of todays life.

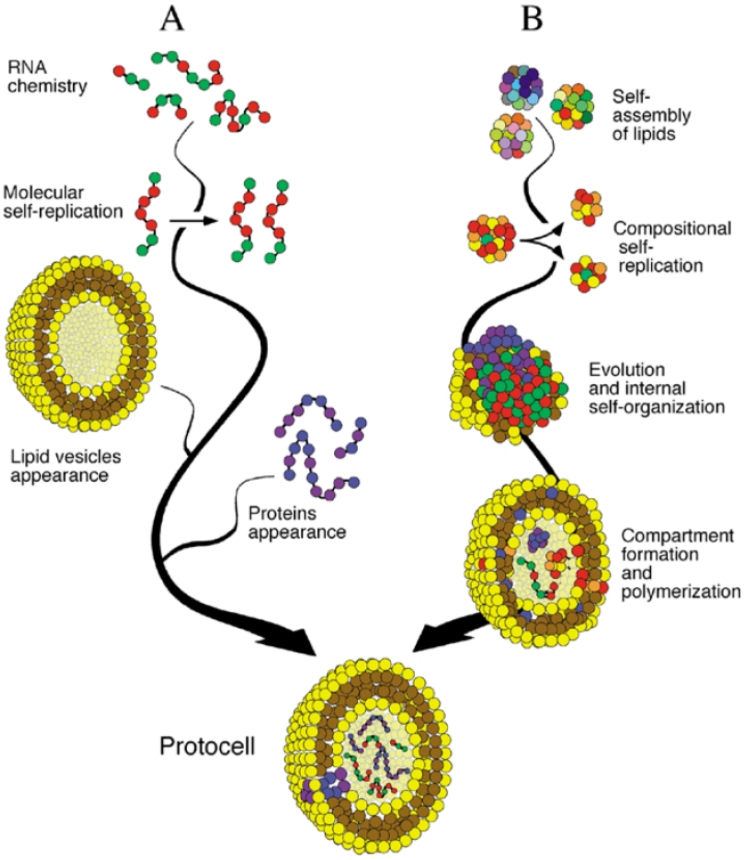


Figure 3.2: *(A) The "biopolymer first" scenario, according to which the emergence of self-replicating informational strings such as RNA and proteins are assumed to have had an independent origin from that of lipid encapsulation. (B) The "lipid world" scenario, which maintains that the roots of life could have been aggregates of spontaneously assembling lipid-like molecules capable of compositional inheritance. Taken from [74]. For details, refer to section 3.1.2*

From a gene-centric point of view (cf. section 2.1) DNA is seen as the ultimate unit of evolution, and thus, the main problem to be addressed is the origin of DNA-replication. Without replication, natural selection can not occur. Regrettably, DNA-replication is a quite complicated mechanism, and thus, simpler scenarios have been sought. Obviously, the most simple DNA-replication scenario would be a scenario involving exclusively DNA, and no other molecules. Ideally, DNA would catalyze the formation of DNA. Regrettably, this is not the case, but RNA - an almost identical molecule - indeed is capable of enzymatic activity. This type of RNA is known as *ribozymes* respectively *catalytic RNA*. Thus, a scenario of self-replicating RNA-molecules seems to be a feasible candidate. However, macro-molecules like DNA or RNA are very unstable and can only persist under a narrow range of chemical conditions. A protective vesicle is needed, whose evolution - what ever it may be - appears to be completely separated from that of RNA-replication. This 'mainstream' view on the origin of life can also be called the *"biopolymer first"* scenario, as it concentrates on mechanisms for the replication of complex biopolymers like RNA (cf. Figure 3.2).

Faced with this difficulties, the *"lipid world"* scenario has been proposed:

> *"One is compelled to consider an alternative: that self-replication has never been a property of individual molecules, but rather one of molecular ensembles. [...] The crucial enigma seems to be related to self-organization and the self-reproduction of supra-molecular structures ..."* [74, p. 217,218]

Thus, instead of concentrating on the replication of *complex, individual molecules*, the afore mentioned protective *vesicle* entered center stage as a *primitive, supramolecular unit*. Such a vesicle could be obtained from a self-organizing process of lipid-like amphiphiles, i.e. molecules with lipophylic tails and hydrophilic head groups. In an aqueous solution, the hydrophilic heads are likely to align on the 'outside', while the hydrophobic tails tend to point to the 'inside'.

Apart from vesicle formation, the second key-concept of this 'lipid-world' scenario is that of *compositional self-replication*: if a mutually catalytic set of molecules undergoes a random split, we can expect that the resulting (two or more) sets of molecules will exhibit similar chemical concentrations as the parent assembly (cf. subsection 3.2.1 on GARD). For example, imagine a cocktail of different chemicals. If a sufficiently large sample is taken from that cocktail, it will reflect more or less faithfully the composition of the cocktail itself. Compared to the mechanism needed to duplicate a molecule comprising approximately 4-billion base-pairs (human DNA), replication and inheritance seems to become almost trivial in the case of a 'compositional genome'.

Now, assuming a sufficient diversity of such assemblies, natural selection can take place allowing for a *gradual* evolution.

> *"The crucial origin of life question then becomes how natural selection was initiated by some molecular assortments, irrespective of their exact chemistry. [...] protocellular assemblies [...] could have constituted the first systems capable of information storage, inheritance and selection."* [74, p. 217,219]

Summarizing, we can see that the key problem that needs to be addressed, is the formation of a *supra-molecular structure*, i.e. a *unit*, capable of *inheritance* and thus, of evolution through natural selection. The mechanism which leads to such a unit is required to be independent of the presence or absences of specific chemicals.

### 3.1.3 Minimal complexity

In the previous subsection 3.1.2, the idea was introduced that self-replication never was a property of individual molecules, but rather of molecular assemblies. Supporting this *holistic* approach is Kauffman's notion of a *minimal complexity*:

> *"... all living things seem to have a minimal complexity below which it is impossible to go."*[70, p. 42]
>
> *"Life emerged, I suggest, not simple, but complex and whole."*[70, p. 47 et seq.]

As a possible candidate for the simplest free-living organism, the bacterium *Mycoplasma pneumoniae* has been proposed[1]. It has been characterized as *"a blueprint*

---

[1] Kauffman himself proposed *"Pleuromona, a highly simplified bacterium"*[70, p.42] as the simplest free-living cell. As this organism is unknown to the NCBI taxonomy, probably because the name is outdated, I have chosen to use M. Pneumoniae to illustrate his point.

*of the minimal cellular machinery required for life"*[75, p. 1235] and is illustrated in **Figure 3.3**. Its genome comprises 816,394 nucleotides, containing 689 protein-coding genes, not to mention all types of no-coding mRNA and the whole machinery needed for transcription and protein synthesis. It is further endowed with a fully functional metabolism and a complex cell membrane allowing for a controlled interaction with it's environment. Naturally, this bacterium is capable of DNA-replication and cell division - otherwise, it would not be *alive*. Clearly, discussion are possible about the exact meaning of 'simple' in this context, but it seems to be obvious that *"there is no such a thing as a simple bacterium"*[76, p. 1201].

Indeed, there are 'simpler' organism than M. pneumoniae, like the insect symbiont Hodgkinia cicadicola known for its small genome ([76, p. 1201]), or host-dependent bacteria. Furthermore, viruses in general are magnitudes 'simpler' than any bacterium, but all these life forms can hardly be seen as *free-living*, independent organism. Hence, they provide poor models for an ancestor organism that emerged from the primordial soup. Furthermore, note that 'small' or 'simple' e.g. in term of genome size does not correspond to old in evolutionary terms. In contrast, the low redundancy of very reduced bacteria and viruses probably represent the result of an evolutionary optimization process, and not the 'point of departure' for evolution.

Clearly, discussions are possible concerning a.) the exact meaning of *free-living* and b.) what measurement to use to identify the *simplest* free-living organism. But independent of the outcome of such a discussion, the main point here seems to be beyond doubt: what ever organism we choose to be the *most simple, free-living* one, it will be a supra-molecular structure (unit) of *astonishing complexity*.

As a first immediate consequence of this holistic notion of a minimal complexity follows a non-gradual, non-continuous account of the origin of life. Second, as a concluding observation on this issue, we would like to briefly mention the idea that only a rather elevated complexity allows for adaptability. Stated in other words, complexity can be seen as a requirement for adaptability.

## 3.2 Formal models and simulations

After the discussion of the previous section, we are now able to provide a more precise description of the problem we seek to solve. Furthermore, we can also formulate a series of requirements that possible answers have to fulfill:

1. How is the emergence of *ordered, dissipative systems* not only possible, but *probable* (cf. subsection 3.1.1)?

2. We assume that this emergence was a *single-step event*, and not a gradual evolution (cf. subsection 3.1.3).

3. The resulting dissipative system needs to be *capable of evolution* through natural selection, i.e. it needs to exhibit the following properties (cf. **Figure 3.4**):

    (a) a *supra-molecular unit* (cf. subsection 3.1.2)

    (b) capable of *reproduction* (cf. subsection 3.1.2)

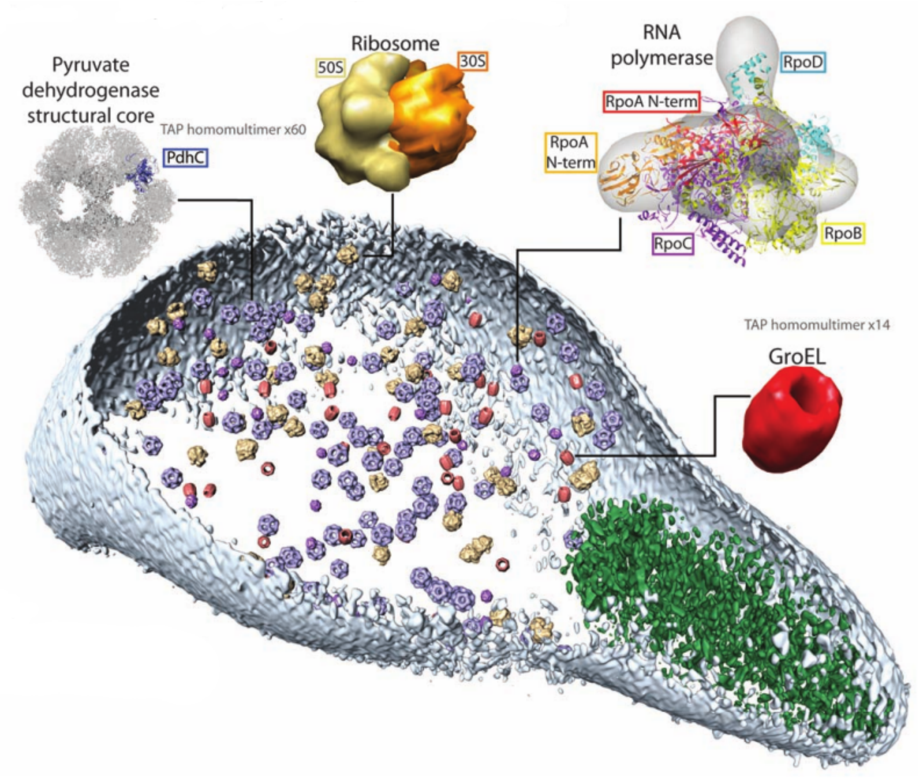


Figure 3.3: *Illustration of M. pneumoniae. Even without going into the details of it's main components and their interplay, it becomes clear that this example of one of the most 'simple' free-living cells is anything but simple. Please not that also it's basic components as e.g. RNA polymerase represent extremely complex biopolymers, which are never found outside a living organism. Taken from [75, p. 1238]. For details, refer to section 3.1.3*

(c) with a certain amount of *stability and variability* (cf. subsection 3.1.2)

(d) which is *modular* such that a gradual evolution becomes feasible (cf. subsubsection 2.2.3).

Furthermore, we have seen that a compositional genome might be a suitable model for reproducing units (cf. subsection 3.1.2). Likewise, we suspect that variability requires a certain amount of complexity, thus supporting our notion of emergence as a single-step event (cf. subsection 3.1.3).

In the following subsections, we will have a look at a series of models trying to address one or more of the above issues. The theory of *autocatalytic sets*, as outlined in **subsection 3.2.1** aims at providing an answer to question number 1. and 2., while the GARD model (ibid.) also provides hints on how a primitive form of reproduction and inheritance (question 3.b) might be feasible. Using *boolean networks*, the notion of stability and variability can be given a rather straightforward interpretation (cf. **subsection 3.2.2**). Finally, question 3.d concerning the relation between gradualism and modularity can be approached using a model which integrates this two properties (cf. **subsection 3.2.3**).

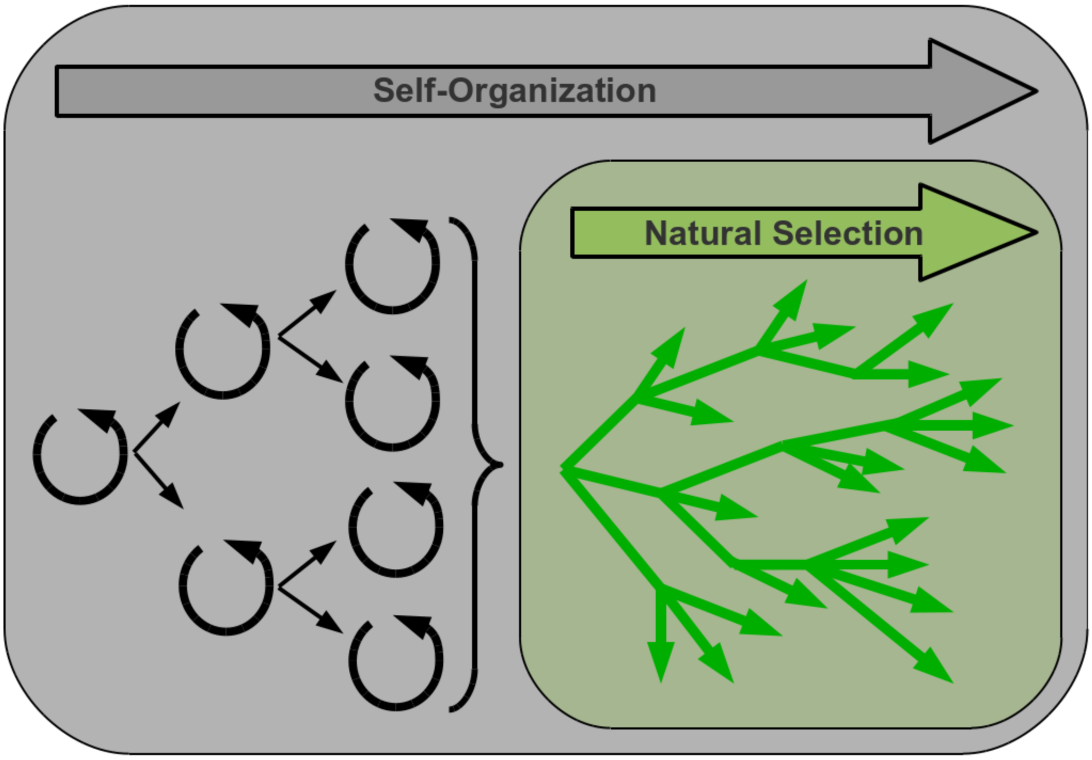


Figure 3.4: *According to Stuart Kauffman, self-organization creates and maintains the conditions which are required for natural selection to take place. Thus, as illustrated above, evolution through natural selection is embedded in self-organization. As shown on the left hand side, self-organization aims to provide an answer to the question how* multiple, relatively stable, self-replicating units *emerge and thus enable natural selection. In this context,* positive feedback *or* self-reference *is seen as one of the core mechanisms of self-organization. For details, refer to section 3.2*

### 3.2.1 Emerging units: autocatalytic sets

Kauffman proposes the model of collectively autocatalytic sets as an answer to the question, how stable supra-molecular structures could have emerged in a single step. To begin with, imagine a set of chemicals, that are reacting with each others, thus forming novel chemicals, but also changing the concentration of those chemicals which were already present. We can represent such a set of inter-reacting chemicals as a network: each *node* of the network represents a specific chemical, while the *weight* (thickness) of the *directed edges* between chemicals represents the corresponding rate constant, i.e. the speed at which one chemical is transformed into the other one. Some reactions involve more than just two chemicals, i.e. they are not of the form $A \to B$ but rather $A+B \to C$ or $A+B+C \to D$. As can be seen, the more reactions are taking place, the higher the expected chemical diversity, and the higher the chemical diversity, the more reactions are becoming feasible. At this point, the notion of *catalysis* is introduced: if a chemical is a catalyst to some reaction, it speeds up this very reaction. Thus, catalysis increases the rate constant of the catalyzed reaction and thus accelerated the formation

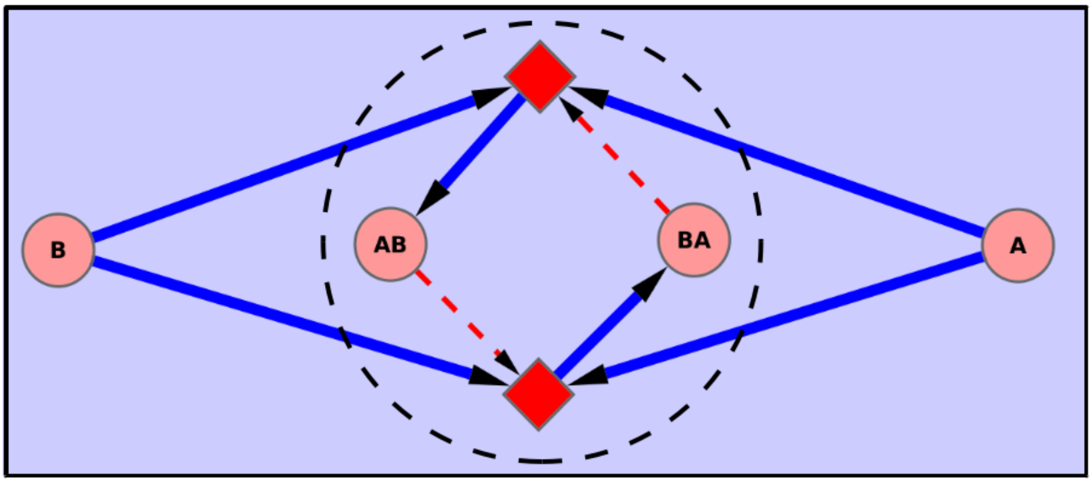


Figure 3.5: *Two types of food molecules ('A' and 'B') are reacting to form the molecules 'AB' and 'BA'. The reaction itself is shown as a red diamond. Both 'AB' and 'BA' have a catalytic effect on the formation of each other (shown as dotted red arrows). Hence, 'AB' and 'BA' form a (minimal) autocatalytic set. The achieved catalytic closure, i.e. the proposed supra-molecular unit, is symbolized by a dotted circle. For details, refer to section 3.2.1*

of the reaction product. Now, if two or more chemicals are mutually catalytic, they speed up each others formation. Hence, *mutual catalysis* represents a case of *positive feedback* (cf. **Figure 3.5**). One of the most important aspects to notice is that, from a chemical point of view, such a system of mutually catalytic chemicals can deviate significantly from the initially expected equilibrium state, following its own, internal dynamics.

Not all members (chemicals) of such a reaction network need to be part of the autocatalytic set. Indeed, the key aspect of this model is that, if the reaction network is 'big enough', a subset of chemicals, which achieves catalytic closure, is very likely to emerge. Stated in other words, the *probability of catalytic closure* depends on two parameters:

> *"... [A] an increase in either the diversity of molecules or [B] the probability that any molecule catalyzes any reaction."*[70, p. 64] *"Simpler systems simply do not achieve catalytic closure."*[70, p. 69]

*Ad [A]:* Clearly, the higher the diversity of molecules, of involved chemicals, the more reactions are feasible. Thus, the chances that some of these reactions are catalyzed by some molecules increases and consequently, the emergence of positive feedback loops gains probability too.

*Ad [B]:* If the chances that any molecule catalyzes any reaction increase, so do the chances that a set of molecules is mutually catalytic. Note that the probability to achieve catalytic closure increases exponential in both parameters, which implies that a *phase transition* can be expected, if these parameters are changed: While a simple chemical soup will hardly allow for autocatalytic sets, catalytic closure suddenly becomes highly probable if the primordial soup reaches a certain minimal diversity and / or a minimal level of catalytic interdependence. Indeed, Kauffman claims that in a sufficient complex chemical soup, *"The emergence of autocatalytic sets is almost inevitable"*[70, p. 61]. This clearly matches the previously introduced requirement for a minimal complexity:

> *"Simpler systems simply do not achieve catalytic closure. Life emerged whole, not piecemeal, and has remained so. [...] we have a hope of explaining why living creatures seem to have a minimal complexity ..."*[70, p. 69]

As another advantageous aspect of this model, he points out that the mechanism underlying autocatalytic sets is independent of any specific chemistry, but rather depends upon *mathematical* properties. On the one hand, this independence can be seen as an advantage, given that no problematic assumption concerning the ingredients of the primordial soups need to be made. On the other hand, when using a mathematical model for a chemical process, it has to be guaranteed that the model actually 'represents' the process it is supposed to 'represent'. Given that, for what concerns the origin of life, the process itself is unknown to us, establishing such a correspondence between model and process-to-be-modeled is difficult:

> *"... the particular details of the chemistry may not matter. We will be showing that the spontaneous emergence of self-sustaining webs is so natural and robust that it is even deeper than the specific chemistry that happens to exist on earth; it is rooted in mathematics itself."*[70, p. 60]

A more recent application of these ideas can be found in the **Graded Autocatalytic Replication Domain - GARD** (cf. e.g. [78, 79, 80]). In this model, the basic ideas of autocatalytic sets has remained unchanged, while introducing some slight modifications which allow for evolution through *natural selection*. Thus, the primary objective of the GARD model is not to show that catalytic closure is very likely to emerge under certain chemical conditions, but to show how mutual catalysis can be achieved by *different* sets of chemicals. Such a diversity of units is one of the primary requirements for natural selection. More precisely, for each GARD (cf. **Figure 3.6**), a certain amount of randomly assigned catalytic relations among the involved chemcials is assumed. As a consequence of this random assignment, catalysis will be of different efficacy in different domains (GAR*Domains*), i.e. the different GARD systems are graded (*Graded*ARD) according to their catalytic success and according to their chemical composition. Assuming that a GARD will split in two GARDS once a certain maximal number of molecules is reached, catalytic efficacy can be directly linked to reproductive success. Ultimately, this success depends upon the combination of chemicals (dimers) present in the replication domain in specific concentrations, i.e. on the GARD's compositional genome. An illustration of this different concentration levels, i.e. of different compositional genomes, can be seen in **Figure 3.7**.

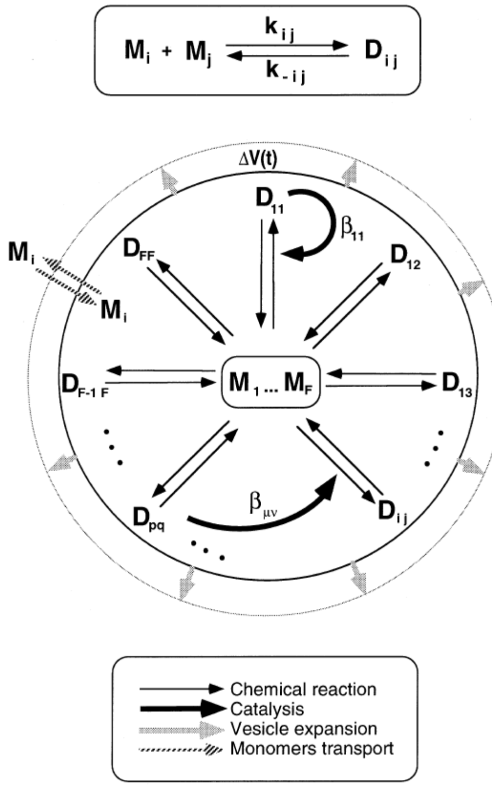


Figure 3.6: *The GARD model of an autocatalytic set. As can be seen on the top, the basic chemical reaction is the formation of dimers* ($D_{ij}$) *from monomers* ($M_i, M_j$)*, together with their corresponding rate constants (k). This rate constants may be changed through catalysis* ($\beta_{u,v}$)*. All chemicals are enclosed in a flexible membrane, which is to a certain extent permeable for monomers ('food molecules'), but not for dimers. Thus, if the number of dimers increases, the vesicle will expand and eventually split into two novel vesicles. The novel vesicles will exhibit similar concentrations of dimers as the parent vesicle. Thus, the abundance levels of dimers represents a* compositional genome*. From [77]. For details, refer to section 3.2.1*

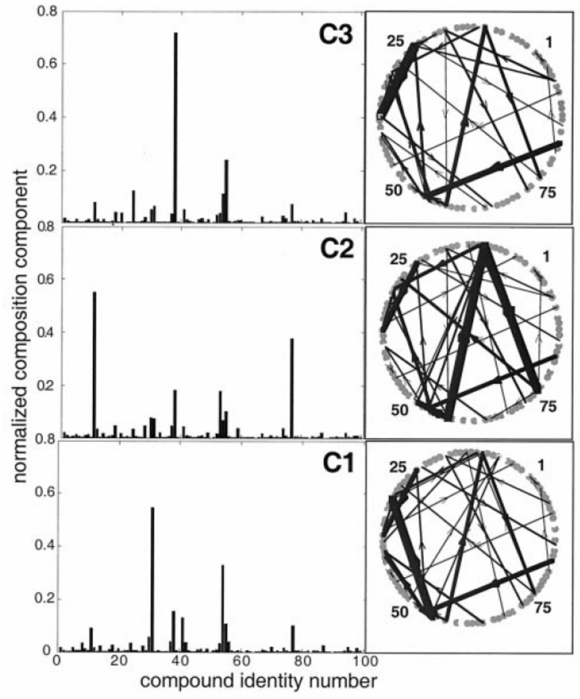


Figure 3.7: *Compositional genomes in three different GARDs. Each GARD consists of* 100 *different chemical compounds, which may exert a certain catalytic effect on the formation of other compounds. In the right column, the strength of this force is visualized as the thickness of an edge linking the compounds, which are arranged in a circle. In the left column, the resulting compositional genomes are shown as the profile of concentrations: on the abscissa, the compounds (*$1-100$*) are listed, while the ordinate measures the obtained concentration. From [81]. For details, refer to section 3.2.1*

Finally, note that the concept of a *compositional genome* illustrates that the distinction between (information coding) genotype and (information executing) phenotype can be omitted. As observed by Kauffman, this is aspect is not only typical for the GARD system, but for autocatalytic sets in general:

> *"With autocatalytic sets, there is no separation between genotype and phenotype."*[70, p. 73]

The structure itself, or, stated in more general terms, the organization of the system itself represents its 'blueprint'. Thus, the challenge of inheritance is to reproduce this very structure respectively organization, and no matter how this challenge is addressed, the use of 'information talk' in this context seems to be not only problematic (cf. section 2.1.3), but also unnecessary.

### 3.2.2 Stable units: attractors and state cycles

As mentioned before (cf. subsection 3.1.2), a certain level of stability is a requirement for evolution by natural selection:

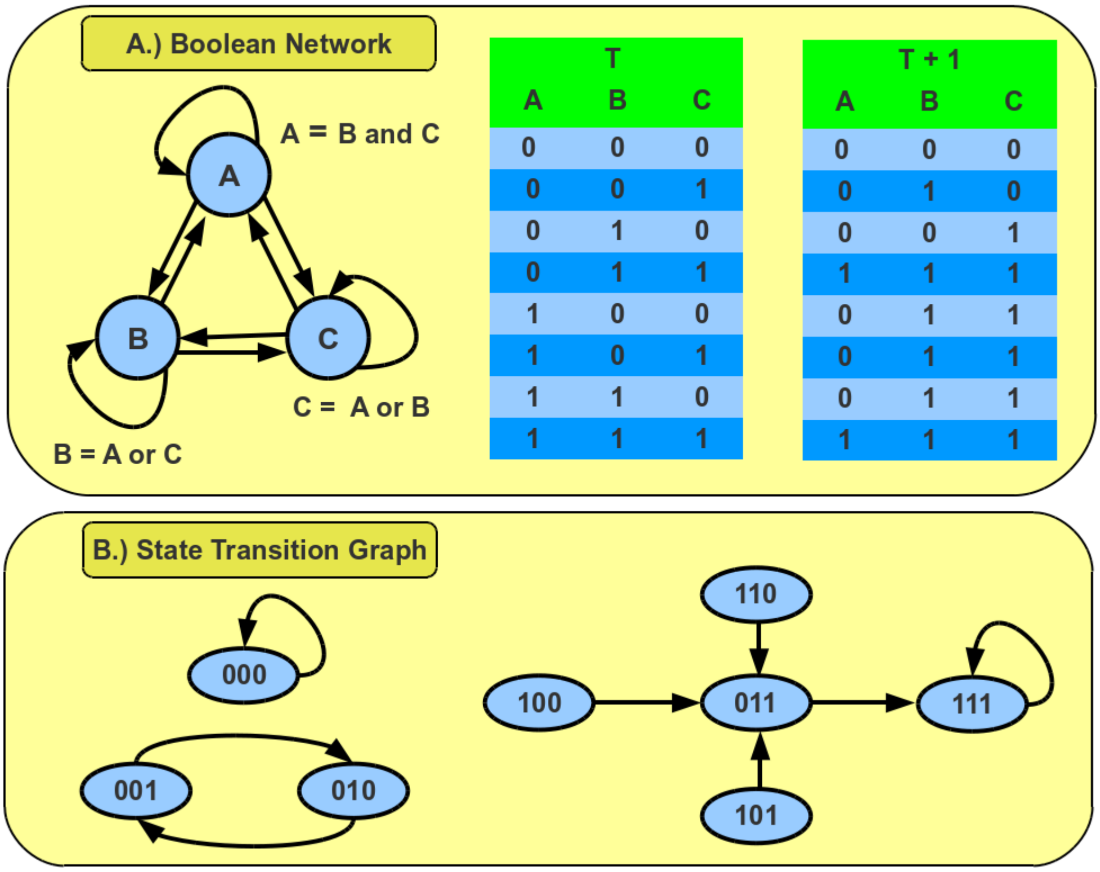
> *"To engage in the Darwinian saga, a living system must be able to strike an* internal *compromise between maleability and stability."*[70, p. 73] *"However life started, with nude replicating RNA molecules or with collectively autocatalytic set, this stability cannot be imposed from outside by natural selection. It must arise from within as a condition of evolution itself."*[70, p. 80]

In order to find hints on how living systems achieve stability, Kauffman analyzes the behavior of *Random Boolean Networks* (cf. e.g. [82] for an introduction). As pointed out by Kauffman ([83]), such networks can be used to model *gene regulation* and are still in use for this very objective today (cf. e.g [84, 85]). Given that in most multi-cellular organisms most cells contain the same set of genetic instructions, i.e. of DNA, the question arises how the formation of different cell types during development is achieved. Stated in other words, on the way from the fertilized zygote to the adult organism, a variety of specialized tissues, as diverse as nerve cells, bones and skin, is developed[2]. Given that each of these cell types holds the same set of genes, it is assumed that the phenotypic differences are accounted for by *regulatory* differences, i.e. differences in gene activity. As a matter of fact, in a nerve cell other genes are 'on' than in the intestinal epithelium. Furthermore, gene products (protein or mRNA) of one gene may have a regulatory impact on the activity of another gene, i.e. the activity of gene A may change (increase or inhibit) the activity of gene B. Apart from the ongoing popularity of Boolean Networks as a means to model such regulatory networks, they are also suitable to examine the behavior, e.g. 'stability' and 'variability', of dynamical systems in general.

In general, Boolean Networks consist of a set of $N$ binary variables, $\{x_1, x_2, ..., x_n\}$, and thus, the number of possible states (e.g. $x_1 = 0, x_2 = 1, ..., x_n = 0$) for such a

---

[2] In homo sapiens, approximately 260 different tissues can be identified [35, p. 392 et seq.].



| T |  |  | T + 1 |  |  |
|---|---|---|---|---|---|
| A | B | C | A | B | C |
| 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 0 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 |

Figure 3.8: *A.) Example for a Boolean Network comprising three variables (A,B,C), each of which is determined by the states of the* other *two variables. Combining the three individual (local) transition rules yields a global rule, describing the transition of the whole system from one state ($T$) to the next ($T + 1$). This rule can be written as a table (top right). B.) Based on the global transition rule, the system's trajectories through states space can be illustrated as a state transition graph. Three attractors of differing size can be identified, with* 111 *being the only* stable *attractor with a basin of attraction. For details, refer to section 3.2.2*

network is $2^n$. For each variable, a *transition rule* from state $t$ to state $t + 1$ is defined. This rule can be seen as a boolean function defined on (possibly) all variables of the network (e.g. $x_1 = x_1 \wedge x_2 \vee ... \wedge x_n$ or $f(x_1, x_2, ..., x_n) \mapsto x_1$) and thus, is a mapping from $2^n$ to $\{0, 1\}$: $2^n \mapsto \{0, 1\}$. Naturally, a transition may be based on a subset of variables too, e.g $x_1 = x_1 \wedge x_2$ or simply $x_1 = x_2$. Taking all the individual transition rules together, the *global transition rule* is obtained, i.e for each of the $2^n$ states of the network, the consecutive state at $t + 1$ is defined: $f(x_1, x_2, ..., x_n) \mapsto \{x_1, x_2, ..., x_n\}$, $2^n \mapsto 2^n$. In the upper part of **Figure 3.8**, a network of 3 variables (i.e. $N = 3$), it's individual transition rules and the global transition rule have been sketched.

Applying the global transition rule, for each possible state of the network the consecutive state can be determined. In order to illustrate the behavior of the system, a *state transition graph*, as shown in the lower part of figure 3.8, can be used. For each

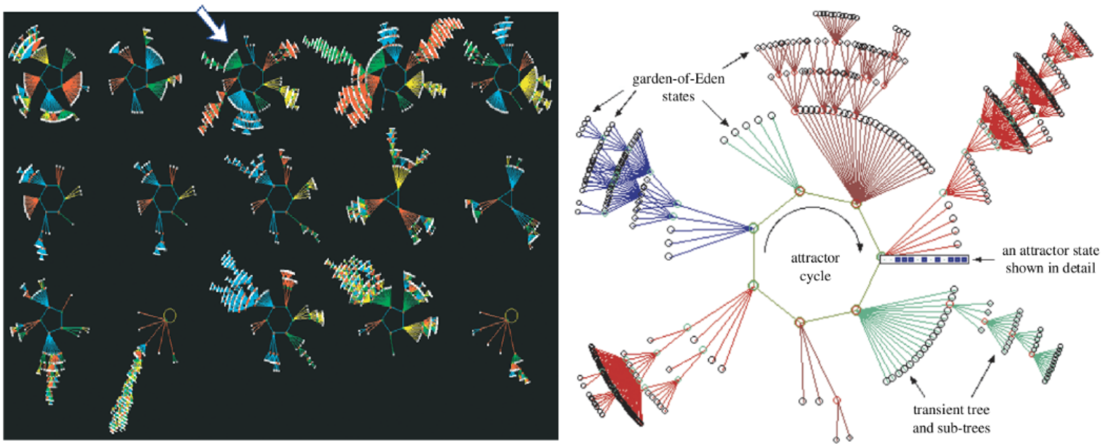


Figure 3.9: *On the left hand side, the complete state space of a boolean network with* $N = 13$ *(i.e. 4192 states) is shown, each of which flows into one of 15 attractors. The attractor marked with an arrow is shown in more detail on the right hand side. As can be seen, some state have no precursor states (”garden-of-eden states”) and the vast majority of states is only transient, forming tree-like structure which make up the basin of attraction for the attractor cycle. From [84]. For details, refer to section 3.2.2*

possible state of the network, a node is drawn. Given that each state leads to exactly one following state (possibly itself), a directed edge can be used to designate state transition.

Examining the properties of such a state transition graph, we find that, not matter in which state the system is started, it will necessarily end up circling around in a number of states. This circle may exhibit very differing lengths, either consisting of just one single state, repeating itself over and over again, or comprising the entire state space of the system. Furthermore, multiple such attractors may exist in a single Boolean Network (cf. **Figure 3.9**). Thus, an examination of the state transition graphs yields information on the system’s behavior in terms of i.) how many attractors are there, ii.) what cycle lengths they have got and iii.) what sizes their corresponding basins of attraction have.

Now, equipped with the knowledge of Boolean Networks and their behavior, we can return to our initial question concerning the *”compromise between maleability and stability”* each living system has to strike. Using our notion of attractors in Boolean Networks, we can clearly see that such an attractor suits our notion of stability: perturbing a system only slightly, such that it does not leave its current basin of attraction, will not have a catastrophic impact on the systems behavior. In general, a slight modification will not send the system into a completely different region of its state space, but rather will cause an only temporary deviation from its current attractor. Without further perturbation from the outside, the system will *stabilize* itself again. However, if the amount of external perturbation surpasses a certain threshold, the system may be pushed into a different basin of attraction.

Thus, we can indeed provide a meaningful interpretation of *stability* and *variability* exploiting the concepts of *attractors*, *state cycles* and, in particular, *basins of attrac-*

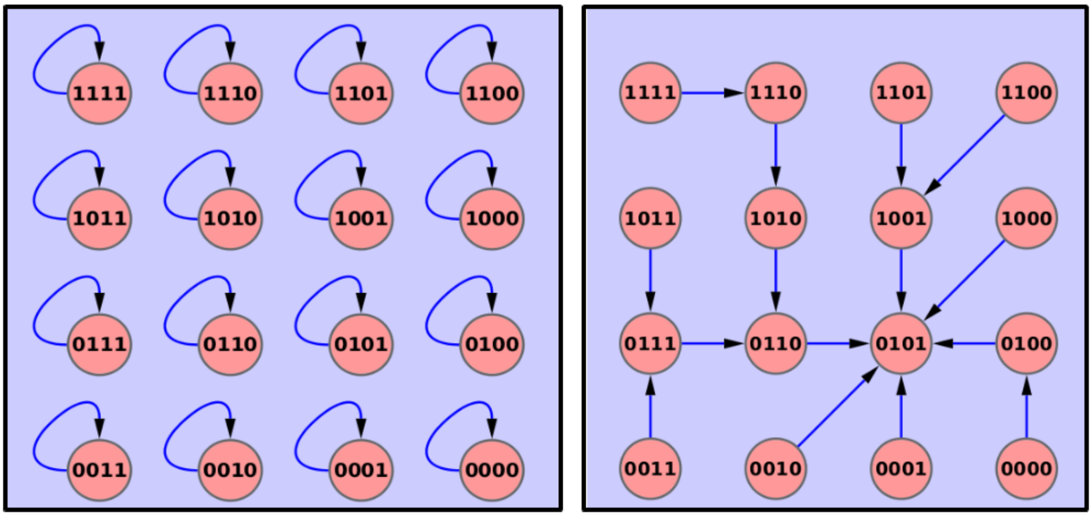


Figure 3.10: *Two extreme state transition graphs, for a Boolean Network with* $N = 4$ *nodes and* 16 *different states. On the left hand side, a maximum number of 'attractors', each comprising one single state, is shown. On the right hand side, a system with one single attractor (*0101*), who drains the whole state space, is depicted. While in the first case, no stability to perturbations is given, in the second case, no flexibility (adaptability) can be observed. For details, refer to section 3.2.2*

*tion*[3].

However, it is crucial to notice that such a reasonable interpretation of stability requires a specific structure of the state transition graph. Apparently, the following two extreme examples of state transition graphs do not match our intuition of a stable yet flexible system (cf. **Figure 3.10**): If there are either too much attractors or too few ones, the system is either very sensitive to small changes (i.e. unstable or *chaotic*), or does not respond to perturbations at all (i.e. inflexible or *ordered*). Thus, the key question is: what makes the difference between a 'good' network, exhibiting the desired mixture of stable yet flexible behavior, and a 'bad' network? What parameters can be used to characterized those Boolean Networks which are flexible yet stable? Stated in Kauffmans own words:

> *"From a biological perspective, the most interesting networks lie at or near a critical point in parameter space that divides ordered from chaotic attractor dynamics."*[86, p.68] *"The transition from the ordered regime to the chaotic regime constitutes a phase transition, which occurs as a variety of parameters are changed. The transition region, on the edge between order and chaos, is the* complex *regime. "*[87, p. 174]

---

[3] Clearly, boolean networks are just *one* possible approach to simulate dynamical systems and thus, exemplify terms like 'attractor', 'basin of attraction' and the like. However, given that they are among the most simple ones, they are well-suited to satisfy our needs. Furthermore, the central results obtain on this model can easily be transfered to more sophisticated ones as e.g. differential equations.

It turns out that one of the most important parameters of boolean networks is the number $K$ of inputs per node[4]. $K$ specifies on how many components of the system, i.e. nodes of the network, the consecutive state of a single node depends. For example, in a network with $K = 3$, the state of each node depends on three other nodes, possibly including itself. Interestingly, the network properties which are of relevance to us - flexible yet stable - are largely determined by $K$, independent of the exact type of local transition rules. Thus, the local transition rules may be chosen at *random* (*Random* Boolean Networks) without changing the global network behavior. Specifying $K$ is sufficient to determine if the network will be locate in the *chaotic*, in the *ordered*, or in the *complex* regime. The class of Random Boolean Networks determined by $N$ (number of nodes) and $K$ (number of inputs per node) is called *NK-Networks*.

Apparently, the maximum possible number of inputs for a single node is $N$, i.e. if $K = N$, the state of each node of the network depends on the states of all the nodes of network, including itself. Such a system is deep in the chaotic regime, characterized by elevated number of attractors, which have typically small basins of attraction and rather long cycles. Thus, these attractors are unstable. Reducing the level of interdependence by lowering $K$ introduces more stability. Around $K = 2$, the *complex* regime is reached, before at lower values of $K$ the network 'freezes in' and assumes an *ordered* state (cf. [88]):

> *"Thus systems with millions of elements can crystallize order if each element is affected by only a few others. The emergence of order does not require that all details of structure and 'logic' be controlled precisely. Hence a rich vein of order lies available for selection's further sifting."*[87, p. 174]

Thus, under certain conditions order may emerge from purely stochastic processes. Systematically randomized behavior on the lower level of a system may be a solid basis for lawful behavior on higher levels. Schrödinger points out that a similar relationship can be found in thermodynamics, where stochastic models on a molecular level are the basis for deterministic *laws* governing the relations between supra-molecular properties as e.g. heat and density (cf. [71]). Summarizing the results obtain from simulation with Random Boolean Network, two things have to be noticed:

a.) Not all possible states of a system are equally probable, which stands in stark contrast to the claim upon which the 'improbability of order' is based (cf. section 3.1.1). As a consequence, the system ends up squeezed into tiny portions of the state space. Thus, the true probability of a state (as determined by the systems dynamics) may deviate significantly from the 'naively expected' probabilities.

b.) Rather loosely interconnected systems ($K = 2$) are poised at the edge between 'chaos and order' and thus, exhibit the mixture of variability and stability needed

---

[4] In the given context, we will abstain from a rigorous formal presentation of Random Boolean Networks, as this would go beyond the scope of this discussion. Such a presentation would also include a more precise definition of the defining parameters and the resulting system behavior, which can be found e.g. in [82]

for evolution through natural selection. The details of the systems architecture can be determined in a randomized fashion and thus, can be seen as relatively irrelevant.

### 3.2.3 Modular units

In the previous section, certain properties of the units themselves were examined. More precisely, the optimum balance between stability and variability as a requirement for evolution through natural selection has been discussed in the context of randomized NK-networks. However, for evolution to work, not only the units of evolution, but the fitness landscape itself needs to exhibit a certain property: it needs to be *correlated* (cf. 2.2.3):

> *"Things capable of evolving [...] all live and evolve on landscapes that themselves have a special property: they allow evolution to 'work': These real fitness landscapes, the types that underly Darwin's gradualism, are 'correlated'. Nearby points [in morphospace] tend to have similar heights [fitness values]."*[70, p.169]

The basic idea of this argument is simple: if minor changes of an organism's traits, i.e. a small movement in morphospace, cause major changes in the organism's fitness, a continuous, gradual improvement of the overall design (bodyplan) via the accumulation of multiple minor improvements is not possible. Thus, as a *requirement for gradualism*, a correlated fitness landscape is needed[5].

Given that our goal is to develop a more precise notion of a functional, evolutionary unit, this knowledge concerning the shape of the unit's fitness landscape can be used as a point of departure to examine the unit itself. Stated in other words, the question is *what types of units exhibit correlated fitness landscapes?* What internal organization of a unit is needed to achieve the required correlation between morphospace and overall fitness? In order to tackle this question, it is worth noticing that the characterization of a *morphospace* via certain features or traits can be seen as a first level description of the unit. On this level, a characterization of the organization, the structure and the mechanisms determining the units second level behavior is given. *Fitness*, on the other hand, appears as a certain aspect of this second level description and thus, could be termed a second level feature or trait. As such, fitness is a property of the whole unit and not only of separate parts and the notion of a correlated fitness landscape translates into a specific relation between first level structure and second level behavior. Kauffman's approach to analyze this relation is to *decompose* the notion of fitness. He distinguishes between the *overall fitness* of an organism and a *trait-specific fitness*: even though the overall fitness is a property of the organism as a whole, for selection to act upon individual features of the organism, these features must be related somehow to the overall fitness. Hence, it is conceptually sound to assign individual fitness values to individual traits. Thus, the overall fitness can be seen as a *non-linear combination* of the trait-specific fitness values:

---

[5] See [87, p. 45 et seq.] for a detailed discussion of the properties of uncorrelated fitness landscapes.

> *"... the contribution to overall fitness of the organism of one state of one trait may depend in very complex ways on the states of many other traits."*[70, p.170]

Thus, the trait specific fitness is not independent of the other traits of an organism: if some feature is advantageous or not depends on multiple other features. As an example, the interdependence between brain, hands, vocalization and upright gait has been discussed (cf. e.g. [89]): a highly developed brain is rather useless without the necessary tools (hands and voice) to manipulate the world in a sophisticated way. Likewise, hands capable of accurate movements are are not truly helpful without the upright gait, which is not only freeing the anterior limbs from the tedious job of carrying the body, but which places them in the field of vision too.

Thus, when trying to draw conclusions concerning the structure of evolutionary units from the requirements for gradualism, we are apparently trapped in the following conflict:

a.) For evolution through natural selection to act upon single features of a unit, those features must contribute to the overall fitness of the unit in a relatively autonomous way. Otherwise, small changes in one single feature would have a strong impact on the overall fitness of the organism and thus, gradual evolution via the accumulation of minor changes would not be possible. In this case, evolution would only be possible as the simultaneous change of all features of an organism.

b.) A brief look at nature tells us that the utility of individual traits indeed depends upon multiple other traits of an organism: Wings are as useless to a dolphin as gills to a horse. I.e. evolutionary units are *functionally integrated*.

For the sake of gradualism and thus, for evolvability, functionally independent traits are required (a). In contrast to this, we have to concede that our very notion of functionality implies the *integration* of parts to a functional (fit) whole (b).

Kauffman considers this situation as a task consisting of *conflicting constraints*, and proposes the following model to simulate the correlation between i.) a units *internal structure*, namely the level of interdependence present among its parts, and ii.) the corresponding *fitness landscape*: The morphospace of our model organism is defined by set of binary features, variables which may assume just two different states. The individual fitness contribution of each feature is not only determined by the state of the variable itself, but also by the state of $K$ other randomly chosen variables. To each possible configuration of this $K + 1$ binary variables (i.e. $2^{K+1}$), a fitness value between 0 and 1 is assigned at random. Thus, if one feature is changed, on average $K$ individual fitness values are affected. The overall fitness is simply defined as the sum over the individual fitness values. Using this model, the relation between i.) the amount of interdependence present measured by $K$, and ii.) the shape of the fitness landscape can be examined. The results are illustrated in **Figure 3.11**.

For low values of $K$, highly correlated fitness landscapes are obtained, posing optimization problems which are easy to solve. With $K$ increasing, the corresponding

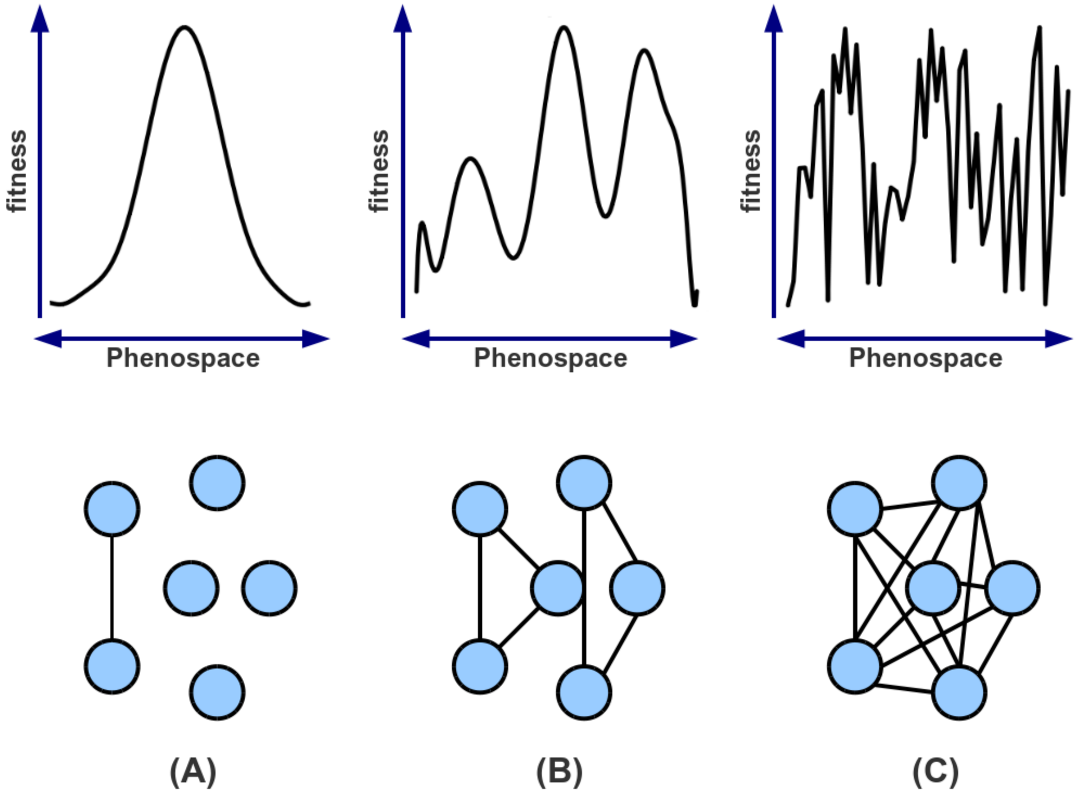


Figure 3.11: *Correspondence between the level of functional interdependence present in a system and the resulting fitness landscape. If the level of functional interdependence is very low, as shown on the left hand side (A), gradual improvements are possible and the maximum on the fitness landscape can easily be found. This can be seen as an optimization problem without any conflicting requirements, i.e. each feature can be adapted independent of the others so as to attain a maximal overall fitness. If the level of interdependence increases (B), the fitness landscape becomes more rugged, and finally (C) violates the requirements for gradualism: small difference in morphospace can cause catastrophic changes in fitness. For details, refer to section 3.2.3*

fitness landscapes become more and more jagged, finally leading to completely randomized surfaces for $K = N$: if the change of one single feature affects the fitness values of all other features, the overall fitness - as the sum of all individual fitness values - is decorrelated from the underlying morphospace. Almost identical combinations of features, e.g. two organism differing in just one single trait, may exhibit a completely different overall fitness.

As we know that the traits of a unit are indeed functionally integrated, we have to exclude those models with very low values of $K$. Even though gradualism would work for such units, we know that the secret of success lies in the functional integration of parts into a whole. An evolutionary unit is not a collection of more or less useful traits which are independent from each other. On the other hand, we can see that highly integrated units, represented by those models with elevated values of $K$, are not capable of gradual evolution.

Thus we may tentatively conclude that each evolutionary unit has to exhibit a certain level of modularity, which optimizes the trade-off between evolvability (indepdence of features) and functional integration.

### 3.3 Summary

Before examining the answers proposed by Kauffman in the form of formal models and computer simulations, we have tried to formulate an explicit problem statement in a first step (cf. **section 3.1**). Among the problems associated with living units, the origin of life is a particularly well-suited example for the *circularity* inherent to question and answer, to phenomenon and model. In this case, we are confronted with the task to provide an explanation for an event that occurred billions of years ago and which can no longer be observed. Hence, defining the problem in first place can arguably be seen as more difficult than providing suitable models for an already well-defined phenomenon. Indeed, we may even ask if there is a problem at all. After all, it is not 'nature itself' asking questions and demanding answers, but our very own understanding of nature which exhibits 'gaps' that call for a theory to fill them. Thus, we have to diagnose in the first place that the origin of life constitutes such a gap in our knowledge.

We are confronted with something that appears to be highly improbable in the light of our current knowledge (cf. **subsection 3.1.1**), *yet we are here*, no matter how improbable it may appear to us. Hence, we are looking for a theory which explains the emergence of living units as something sufficiently probable. Given that life is highly improbable if seen as a closed thermodynamic system, we concluded that closed thermodynamic systems are no suitable model for life, and adopted the notion of open *dissipative systems* instead. Kauffman also asserts that the origin of living units can not be understood as a gradual, continuous process, but rather has to be modeled as a holistic, emergent phenomenon. He argues that the organization of todays life forms is a circular one, in which hardly any component is superfluous. More precisely, he claims that an accumulative evolution of a system formed by mutually interdependent components has to be seen as impossible. If all components of a system are equally important for the system's functioning, the system can not be assembled in step-by-step, one-by-one fashion. Ergo, we seek a model which explains the emergence of life as *whole, living units* (cf. **subsection 3.1.2**). What we can also learn from recent species, according to Kauffman, is the existence of a certain *minimal complexity* required below which no living unit can exist (cf. **subsection 3.1.3**). Independent from the exact definition of 'simple' and 'free living' that is employed, the claim that there is no such thing as a simple, free living organism seems to be beyond all reasonable doubt.

Thus, as summarized in **section 3.2**, the main problem is to provide a model that

a.) makes the emergence of open, *ordered, dissipative systems probable*,

b.) in which this emergence is a single-step event (*holistic*) and

c.) which allows for *natural selection* to take place.

The last claim addresses the requirements which need to be satisfied such that evolution through natural selection becomes feasible. This includes primary a certain mechanism of *reproduction*, a suitable balance between *stability and variability*, and, above all, a certain *fitness-landscape* which allows for *gradual improvement*.

The theory of *autocatalytic sets* outlined in **subsection 3.2.1** represents a model addressing b.) in the first place. Mutual catalysis among a set of chemicals can lead to positive feedback loops, which in turn stabilize the formation of the catalytic network itself. This network is *holistic* insofar as mutual catalysis requires the simultaneous presence of all interacting members of the network. Thus, the spontaneous formation of such a network, which is according to Kauffman highly probable under certain conditions (a.)), can be seen as the single-step emergence of living units. A more recent extension of autocatalytic sets can be found in the Graded Autocatalytic Replication Domain (GARD) model, which also addresses question c.), i.e. the requirements for natural selection. The concept of a *compositional genome* is proposed as a model for a primitive form of inheritance. As the name suggests, the 'information' contained in such a genome is encoded in the composition of various chemicals, which are the elements of a catalytic network. From a chemical point of view, such a network is nothing but a cocktail of different ingredients. If a random split of this cocktail is performed, the composition of the contained ingredients remains on average identical (stability), but also allows for a certain variability. Another account for the subject matter of stability versus variability are Random Boolean Networks (cf. **subsection 3.2.2**), exhibiting a suitable number of attractors. Most importantly, Kauffman demonstrates that the details of the network architecture need not be specified in order to obtain reasonable network behavior in terms of 'stability' and 'variability'. Furthermore, we have seen that natural selection (cf. question c.)) requires a certain type of *fitness landscape*. More precisely, any type of optimization technique based on local information is doomed to fail if the function to be optimized lacks a certain 'smoothness'. Figuratively speaking, on a rugged landscape made up of pinnacles, canyons and abysms, simply walking uphill will lead to a very poor, local optimum. On such a landscape, a random search may be just as good as natural selection. Now, good landscapes on which natural selection can work are *correlated landscapes*, i.e. a small change in phenospace leads to a small change in fitness. Based on the model proposed by Kauffman (cf. **subsection 3.2.3**), such correlation landscapes require a modular structure of the unit to be optimized. Stated in other words, for natural selection to work, a certain independence of the features contributing to the unit's fitness is required.

# Chapter 4

# HUMBERTO MATURANA

In contrast to the previous chapters, Maturana introduces *epistemological* aspects into the concept of an evolutionary unit. Metaphorically speaking, Maturana takes a step back and before asking the question *"what is an evolutionary unit?"*, he tries to answer the question *"what does it mean to ask a question?"*. By doing so, he assumes a *constructivistic* point of view[1], which in turn fundamentally alters his notion of an evolutionary unit.

More precisely, Maturana rejects the notion of cognition as a representation of some external reality and claims that each *unit* is brought forth by an observer through an act of distinction. Given that each observer is seen as a living unit, cognition as the operation of distinction becomes the fundamental capacity of life. Consequently, *understanding life* and *understanding understanding* become challenges which can not be addressed separately (cf. **section 4.1**). Maturana's approach to this challenges can be found in the concept of *autopoiesis*, i.e. the concept of *self-creation*. Autopoiesis, as the defining property of living units, can also be understood as the structure governing the 'understanding of understanding'. In this spirit, Maturana's theory of living units (cf. **section 4.2**) as autopoietic systems *constructs* the emergence of the observer itself, i.e. in a constructivistic sense, it *describes* the phenomenon of cognition.

## 4.1 Knowledge and cognition

The objective of this section is to outline *relativism* respectively *constructivism* as opposed to *objectivism* respectively *representationism*. Note that we are treating this terms, namely relativism - constructivism and objectivism - representationism, as equivalent, assuming that each of them highlights a different aspect of the same position. More precisely, we will briefly outline Maturana's arguments against representationism (cf. **subsection 4.1.1**), followed by a discussion of the main difficulties originating from constructivism in **subsection 4.1.2**. As a conclusion of this epistemological inquiries, Maturana's notion of *science* as a defeat of *solipsism* is presented in **subsection 4.1.3**.

### 4.1.1 Representationism

As the primary requirement of *representationism*, an objective reality, i.e. a subject independent reality has to be assumed. Apart from claiming the existence of an objective reality, as a second requirement of representationism, it has to be assumed that we do

---

[1] We follow the notion of 'constructivism' as coined by H. von Foerster: *"Die Umwelt, so wie wir sie wahrnehmen, ist unsere Erfindung." (The world, as we perceive it, is our own invention - K.M.)* [90].

have some sort of *cognitive access* to this reality. Stated in other words, claiming that an objective reality exists is not yet enough for representationism to work. We also have to assume that we are able to know something about this objective reality, namely that we can *"make cognitive statements"* about this reality. Thus, the core concept of representationism can be summarized in the following claim:

> *"We exist in an objective world that can be known and about which we can make cognitive statements that reveal it as an independent reality whose validity is, therefore, independent from us as observers."*[91, p.1]

Note that representationism is *not self-satisfying* or circular, i.e. the underlying notion of reality as an objective, subject-independent one is a prerequisite for representationism, but none of its consequences. The claim that there is an external, subject-independent reality is none of the cognitive statements that we can make based on representationism, neither does it follow from representationism that we have some sort of cognitive access to this reality. Thus, both the objectivity of reality as well as our cognitive access to it have to be postulated, and can not be derived within the epistemological framework provided by representationism.

Furthermore, representationism also incorporates the antique, Aristotelian concept of *truth as correspondence* between statement (claim) and reality. Thus, cognition as the process of gaining knowledge about the outside reality can be seen as the task of constructing the 'right' mapping between statements and reality. Naturally, the question arises how correct and erroneous mappings can be distinguished and one of the criteria proposed is that of 'successful predictions'. If we are able to predict what the reality will be like in the future based on our representation of it, this capability to predict is a proof for the correctness of our representation. However, Maturana claims that ...

> *"... a successful prediction does not prove that the operation through which we make it reflects an objective reality, or constitutes an expression of our cognitive access to this reality."*[91, p.1]

Indeed, the problem to distinguish between correct and incorrect *representations* is not alleviated, but merely restated, if we are using the distinction between correct and incorrect *predictions* to solve it. The basic problem underlying the antique truth criteria of representationism is that of *comparing* a statement, i.e. the representation, and reality, i.e. that which is represented. A representationist needs to explain what it is exactly that shall be compared here, and how this comparison shall be carried out. The correctness of a statement which is supposed to describe a current state of the world (representation) does not become any more comprehensible when the statement is supposed to describe a future state of the world (prediction). Stated in other words, if we want to proof that *'X is the case'*, we do not gain anything by examining if *'X will be the case'*, as the problem is basically left unchanged by referring to different points in time. Indeed, a 'post-diction' (*'X was the case'*) would be just as helpful in this respect as a pre-diction.

Maturana also points out that this notion of cognition as a mapping hardly can be 'mapped' onto our *biological understanding* of human perception. From a neurological point of view, the question naturally arises how this mapping-process is reflected on the 'biological basis'. For Maturana it is obvious that the notion of a mapping or a correspondence between reality and knowledge is inconsistent with a biological perspective on human cognition (cf. [92, p.21 et seq.]).

If our senses are seen as detectors of filters which inform us about the external reality, we also have to concede that the information we obtain about this outside-world is not objective, because ...

> *"... it is a general feature of all filters that their organization determines what passes through them [...]. If our sense organs function as filters we cannot use them to make any objective statement about the external world ..."*[91, p.2]

Apart from the fact that, according to Maturana, the human senses can not be understood as sensors or filters, one central observation concerning representationism can be extracted from the quote above: As a requirement for representationism, we need some sort of *direct, objective access* to the external world. If we want to proof the correctness of a mapping by a comparison between statement and reality, we need some direct, cognitive access to this reality in the first place. This means that the reality we want to 'discover' during the cognitive process needs to be a *known reality*, i.e. it already needs to be part of what we know in order to serve as a criteria for faithful representations. Thus, the harder we try to build upon objective reality as the basis of cognition, the softer it gets.

Summarizing, the root of this comparison problem seems to lie in the very notion of objective reality as something which is per defintionem not part of our cognitive domain, but external to it. Thus, as we know only what we know, and as we are not able to compare something that we suspect to be true (i.e. a statement concerning reality) with the unknown reality itself, the projected grounding of knowledge in reality itself is doomed to fail. Apparently, we are trying to compare something we do not know (i.e. a non verified statement) with something else that we do not know either (i.e. reality itself) in order to gain knowledge (i.e. a verified statement).

In order to circumvent the problems involved in outlining how a direct comparison between statement and reality is possible, an indirect correspondence derived from the *functionality* ('efficiency') of our knowledge has been proposed. The basic idea behind this notion is that, if we are able to efficiently handle something, i.e. if we know how something 'functions', than we can justifiably assume that we posses a correct representation of its internal structure. From an evolutionary point of view, the correctness of a representation is crucial to our success as a biological species, for it determines our behavior. Metaphorically speaking, surviving reality appears to be difficult with an erroneous notion of reality, just as 'solving a problem' appears to be difficult if the problem itself is unknown.

> *"We obtain knowledge through our sense organs by a process of mapping the objective external reality onto our nervous system, accommodating our behavior to the structure of the world revealed through this mapping."*[91, p.1]

As we will see in the following section, Maturana indeed assumes a close relationship between cognition and evolutionary success, but rejects that *efficiency* implies *faithful representation*. Just as a *successful* prediction does not represent some direct feedback from reality itself, a statement does not become a true statement - in the representationist's meaning - just because it is *useful* for those who believe it to be true.

### 4.1.2 Constructivism

The core concept of Maturanas epistemological position can be seen in his claim that every statement is *subject-dependent*. If something is known, it is always known by someone, and any cognitive action is always carried out by someone. Hence, there is no subject-independent epistemological process.

> *"Alles Gesagte ist von jemandem gesagt." (Everything said is said by somebody - K.M.)*[92, p.31]

Given that knowledge and cognition are always subject-dependent, the possibility of objective, subject-independent knowledge is negated. As a consequence, the claim that 'there is an objective reality' has to be seen as a subject-dependent claim, that can neither be verified, nor be explained in a representationistic framework. As the only notion of reality that we have is our subject-dependent knowledge about reality, the claim that reality itself is objective can impossibly be verified. Departing from the very definition of objective reality, it follows that this reality can never be known in a subject-independent manner. Hence, any statement about the 'objective reality' is, by its very definition, unprovable. Similar to Kant's *"Ding an sich"*, objective reality disappears into the depth of some unknowable metaphysical realm. Summarizing, we have to notice that negating objective knowledge renders the notion of an objective reality obsolete or, as the early Wittgenstein would have termed it, 'meaningless'. Figuratively speaking, objective reality becomes invisible and thus, can not be used to explain what we can or can not see, simply because ...

> *"Wir sehen nicht, daß wir nicht sehen." (We do not see that we do not see - K.M.)*[92, p.23]

However, apart from deconstructing objectivity, the concept of cognition or *observation* as a process which does not alter its object is rejected, as Maturana claims that ...

> *"Jedes Tun ist Erkennen und jedes Erkennen ist Tun." (To act is to recognize and to recognize is to act - K.M.)*[92, p.30] *"... jeder Akt des Erkennens eine Welt hervorbringt." (...each act of cognition brings forth a world - K.M.)*[92, p.30]

Cognition is indeed seen as a process which actively *generates* or *constructs* its object, instead of passively recognizing or perceiving it. Thus, the notion of reality undergoes a major change here, as it is re-defined as a construction, a product of the subject's activity. The basic idea behind this claim of a 'constructed reality' is the following one: once the possibility of objective knowledge is rejected, the notion of an objective reality becomes a 'metaphysical phantasm', i.e. something that we can not know anything about. Believing in an objective reality may be *useful* for a variety of reasons, but this believe can impossibly be verified within representationism. Thus, the only reality we have got left, is the one we know about, and everything we know depends on us, i.e. is subject dependent. Hence, if the term 'reality' shall be used in a meaningful way, 'reality' has to be considered as a subject-dependent construction.

In a certain sense, Maturana has inverted the traditional notion of cognition here: whereas normally some external or objective reality serves as the basis for knowledge, Maturana points out that 'all we know is what we know', i.e. our knowledge is the only basis we have got to comprehend what 'reality' is. Thus, instead of considering cognition as a process in which, departing from reality, knowledge about this reality is accumulated, cognition is seen as a process in which, departing from knowledge, reality itself is actively evolved by the subject.

> *"Living systems are cognitive systems, and living as a process is a process of cognition."*[93, p.7]

Cognition, understood as activity which constructs reality, is thus not one single, specific ability of homo sapiens but rather a very basic capacity of all living beings. As such, it is not constrained to the nervous system, but rather can be identified with the activity of being alive itself. Consequently, life itself is a cognitive process which in turn implies that cognition itself is subject to evolution.

> *"Erkennen ist effektive Handlung, das heißt, operationale Effektivität im Existenzbereich des Lebewesen." (Cognition is effective action, that is, operational effectiveness in the domain of existence of living beings - K.M.)* [92, p.35]
>
> *"Erkennen als wirksame Handlung, das heißt, als eine Handlung, die es einem Lebewesen in einem bestimmten Milieu erlaubt, seine Existenz darin fortzusetzen, indem es dort eine Welt hervorbringt." (Cognition as effective action, that is, as an action that allows an animal in a certain environment to continue its existence in it, by bringing forth a world in it - K.M.)* [92, p.36]

When defining cognition as an *effective* action, the traditional notion of *truth* as correspondence is replaced by a functional concept of truth. In contrast to the representationist claim, that a correct representation of the outside world is necessary for survival, survival itself is seen as a requirement for cognition, but this does not entail a correct representation of some external reality. Whereas in the case of representationism, the objective reality is seen as a given problem that needs to be understood in order to be solved, in the case of constructivism constructing and solving the problem is seen as one and the same action. Thus, the correctness of cognition is no longer a requirement for survival, but survival has become the only requirement for 'correctness', i.e. for 'effective action'.

Hence, due to the crucial role the subject plays in Maturana's epistemology, and due to the fact that this subject is understood as a biological one, the questions "*what is cognition?*" and "*what is a living unit?*" are intimately interrelated.

> *"Cognition is a biological phenomenon and can only be understood as such; any epistemological insight into the domain of knowledge requires this understanding."*[93, p.6]

However, before proceeding to the concept of a cognitive, biological unit (cf. section 4.2), the implications of a strictly subject-dependent notion of knowledge have to be examined.

### 4.1.3 Solipsism and science

As emphasized in the previous section, Maturana has *relativized* the distinction between 'cognition' as passive perception of reality and 'taking action' as an active change of reality. Consequently, when using this distinction, we now have to take into account that the relationship between cognition and action is a circular one, as, in a certain sense, we can not distinguish between perceiving and changing reality. This circular, constructivistic nature of cognition apparently affects this very inquiry too, because ...

> *"... die Zirkularität zwischen Handlung und Erfahrung auch für das, was wir hier und jetzt tun, gilt." (...the circularity between action and cognition also applies to what we are doing right now - K.M.)*[92, p.31]

Stated in other words, we can no longer act as if we were standing 'outside', as if we could describe some external reality. We have to take into account that, *along this very lines, the object of this investigation is constructed.* This relativity could lead to the uncomfortable situation of complete arbitrariness: Without some fixed point as e.g. an external, objective reality, our knowledge seems to disappear into mere believes, all of which are equally 'wrong' or 'right'. The world around us, as our own private construct, seems to be completely at our mercy, with ourselves, i.e. with the *subject* at it's center. However, Maturana points out that such a *solipsism* only arises if constructivistic ideas are intermingled with objectivistic remains.

> *"... solipsism arises as a problem only if we insist on demanding from a subject dependent cognitive domain the properties of a subject independent cognitive domain."*[91, p.4]

One such property would be e.g. the notion of an absolute truth grounded in a faithful representation of the objective reality. For the sake of illustration, let us assume that we have just abolished the difference between 'true' and 'false' statements, namely by means of the following statement *A: "There are no such things as 'true' or 'false' statements"*. From an objectivistic point of view, this statement has to be seen as *contradictory* as it can not be neither true nor false, and thus, needs to be rejected.

However, this objection is only possible if the statement itself is taken as an objectivistic one: if it is 'taken seriously', the distinction between true and false can no longer be used when judging the statement and thus, claiming that this statement is contradictory has to be considered as a *misunderstanding* of the statement itself. Indeed, if proposition $A$ is read as a *pre*-scription rather than as a *de*-scription, the difference becomes more obvious: objectivism aims at providing descriptions, but this very claim is abandoned by constructivism.

Likewise, if the distinction between an objective and a subjective reality is dissolved as outlined above, solipsism disappears. Thus, the *ontological status* of constructivistic statements can no tbe assessed using an objectivistic ontology. Given that the notion of an external, objective reality is revealed as an unprovable claim, the distinction between 'objective' and 'subjective' reality becomes *meaningless:*

> *"The question, What is the object of knowledge? becomes meaningless. There is no object of knowledge. To know is to be able to operate adequately in an individual or cooperative situation."*[93, p.29]

Nonetheless, even though solipsism can be avoided using Wittgensteinian arguments, the question remains how a positive notion of truth and, consequently, of *science* is possible. Stated in other words, even though relativism and complete arbitrariness can be avoided pointing out that the underlying terms and concepts can no longer be applied, the question remains if and how science is still possible. First of all it has to be noticed that science, just like any other type of knowledge, is subject dependent too:

> *"... since I want to give a scientific description of the observer as a system capable of descriptions (language), I must take the subject dependent nature of science as my starting point."*[94, p.2]

Apart from this notion of science as a subject dependent endeavor, the above formulation of Maturanas goal is worth noticing as well: we seek a subject-dependent description of the *"observer as a system capable of descriptions"*, thus, we are apparently trying to provide a self-describing description. Stated in other words, understanding living units and understanding 'understanding itself' can be seen as corresponding tasks from this point of view. Metaphorically speaking, looking at one of these two phenomena, living units *or* cognition, will facilitate the understanding of the other one.

Returning to the problem of science, Maturana exploits the concept of *question* and *answer* to outline his understanding of science as a non-arbitrary, but nonetheless subject-dependent process. Science is now defined via a special type of explanations, namely *scientific explanations*, which, in first place, provide answers to questions, just like any other type of explanation.

> *"... an explanation is always an answer to a question about the origin of a given phenomenon, and is accepted or rejected by a listener who accepts or rejects it according to whether or not it satisfies a particular implicit or explicit criterion of acceptability that he or she specifies. Therefore, there are as many different kinds of explanations as there are different criteria of acceptability ..."*[95, p.4]

In contrast to the objectivistic notion of science, this formulation emphasizes the *pragmatic* and *dynamic* aspects of science as a *social practice*, which is taking place within a specific cultural, sociological context. Science is no longer characterized by some privileged access to objective reality, but is seen merely as a special form of communication, following the spirit of Wittengsteins' *Sprachspiele*, which can be analyzed and described using the ordinary concepts of 'question' and 'answer'.

Thus, we now have defined science via its use of scientific explanations, which are in turn defined via a specific criteria of acceptability. Maturana formulates four conditions which, taken together, represent the criterion of acceptability for scientific explanations. For an explanation to be considered as scientific, it needs to satisfy the following four requirements[2]:

a. *"The specification of the phenomenon to be explained through the stipulation of the operations that a standard observer must perform [...] in order to also be a witness of it ..."*[95, p.5]

b. *"The proposition [...] of [...] a generative mechanism, which when allowed to operate gives rise as a consequence of its operation to the phenomenon to be explained, to be witnessed by the observer .... This generative mechanism [...] is usually called the explanatory hypothesis ..."* [ibid.]

c. *"The deduction [...] of other phenomena that the standard observer should be able to witness ..."* [ibid.]

d. *"The actual witnessing [...] of the phenomena deduced in (c) ..."* [ibid.]

First of all it is worth noticing that the explanation sought in the second step, i.e. the *explanatory hypothesis* represents a *mechanistic* one: it describes the mechanism underlying and generating a certain phenomenon. Consequently, the compatibility between hypothesis and phenomenon can still be seen as the criteria for scientific correctness. However, none of these two components is a fixed point of reference any longer, nor is there any substantial difference between the two aspects 'phenomenon' and 'explanation'. The specification of the phenomenon as well as the description of the generative mechanism are operations carried out by the observer. They are both subject-dependent manipulations of the observers cognitive domain, i.e. self-manipulations, which - according to Maturana - happen to comply with *our* socio-cultural determined notions of scientific practice.

Hence, scientific practice reflects the intrinsic circularity of cognition, but implements this circularity in a specific, socio-cultural determined manner which can be characterized by the four criteria sketched above.

Now, as a result of this preliminary epistemological considerations, not only our notion of science, but also the very objective of this inquiry has been transformed. According to Maturana, *living units* have to be understood as cognitive systems at any time, which implies that, during the course of our investigation, we have to take into account that we as the observers providing a description are cognitive, living systems

---

[2] An alternative formulation of these requirements can be found in [92, p. 34].

as well. Hence, we now seek a *scientific explanation* for the phenomenon of cognition, which, as we will see in the next section, corresponds to explaining the *emergence of the observer*.

## 4.2 Living units

As we have seen in the previous section, the explanation we seek is a *circular* one. In stark contrast to this we have to recognize that the language we have to use to provide this very explanation exhibits a *linear* structure. No matter if we consider the spoken or the written language, we must proceed step by step, one by one. Furthermore, we are not used to circular arguments, but rather to linear ones, just as we are accustomed to explicit definitions. However, without a fixed point of reference, we have to accept that there actually is no difference between explicit and implicit definitions and that indeed the only possible definition is based on *usage*. More precisely, the meaning of a term is determined by the way it is used in a specific context:

> *"In what follows I shall not offer any formal definitions for the various terms used, such as cognition, life, or interaction, but I shall let their meaning appear through their usage. This I shall do because I am confident that the internal consistency of the theory will show that these terms indeed adequately refer to the phenomena I am trying to account for, and because I speak as an observer, and the validity of what I say at any moment has its foundation in the validity of the whole theory, which, I assert, explains why I can say it. Accordingly, I expect the complete work to give foundation to each of its parts, which thus appear justified only in the perspective of the whole."*[93, p.6]

The structuralistic flavor of this notion of *meaning as usage* is apparent and gets further reinforced by the corresponding notion of validity as *internal consistency*. In a strict sense, the structure of the argument is not merely a circular one, but consists in a part-whole relationship, which resembles the popular notion of a self-fulfilling prophecy.

Consequently, as we will see in the following **subsection 4.2.1**, the *distinction* between two units respectively between part and whole is located at the very heart of Maturanas explanation. However, *understanding* units can hardly be achieved by simply claiming that every unit is the result of an operation of distinction. We still seek a *mechanistic* explanation of the internal dynamics of a unit and the way it interacts with its environment (cf. **subsection 4.2.2**). Once the central concepts of *"structural coupling"* and *"structure determination"* are introduced, we can finally proceed to Maturanas concept of *autopoiesis* as the organizational principle governing living units (cf. **subsection 4.2.3**).

### 4.2.1 Observer and unit

When approaching Maturanas notion of living systems, it is worth recalling in first place his notion of cognition as a subject dependent activity, i.e. as the activity of a living unit. Thus, in order to understand living units, we need to understand cognition as the fundamental biological capacity distinguishing the living from the non-living. On the other hand, in order to understand cognition, we need to understand living units. Thus, neither any question concerning life itself nor any epistemological problem can be tackled in isolation:

> *"Anything said is said by an observer. In his discourse the observer speaks to another observer, who could be himself; [...] The observer is a human being, that is, a living system, and whatever applies to living systems applies also to him."*[93, p.7]

Note that this circular relation between living systems and cognition implies that any attempt to understand cognition necessarily includes the attempt to gain self-awareness. Metaphorically speaking, we are trying to formulate a theory which is capable to 'understand itself', i.e. which does not only exhibit internal consistency, but in which an observer is constructed. The fundamental epistemological operation carried out by this observer is the operation of distinction and hence, the theory we seek needs to begin with the distinction which is necessary to give rise to an observer. The distinction which brings forth the observer is nothing but the distinction between the observer 'himself' and the rest. It is the definition of the observer by its defining property, and this very property is the operation of distinction. Hence, note that in the following quote the observer himself is constructed by a distinction, i.e. the unit which is brought forth in the following lines is the observer:

> *"The basic operation that an observer performs in the praxis of living is the operation of distinction. In the operation of distinction an observer brings forth a unity (an entity, a whole) as well as the medium in which it is distinguished, ..."*[95, p.9]

Clearly, this notion of a *unit* is anti-essentialistic and anti-realistic. There are no unchangeable ideas manifesting themselves in the things we perceive, neither do we capture these ideas in a cognitive act. There is no independent, a priori reality which we do access in cognitive acts and comprehension does not mean to get hold of the 'essence' of something. Consequently, a 'correct' definition of some phenomenon can not be identified by means of some correspondence between definition and phenomenon, as the phenomenon as a such does, strictly speaking, not exist. This is rather easy to understand when applied to social or cultural phenomena, such as 'religion' for example. Apparently, all we can do when defining the phenomenon of religion is to reformulate our implicit assumption on religion in a more precise way. We can gain a higher level of self-awareness with respect to our own believes by stating explicitly what we previously knew only implicitly. But claiming that we can distinguish between 'right' and 'wrong' definitions of religion in a correspondence-theoretic meaning apparently makes no sense.

> *"For the observer an entity is an entity when he can describe it. [...] the observer can describe an entity only if there is at least one other entity from*

> *which he can distinguish it and with which he can observe it to interact or relate. This second entity that serves as a reference for the description can be any entity, but the ultimate reference for any description is the observer himself. [...] The observer can define himself as an entity ..."*[93, p.7]

When referring to something in our daily life, we usually do not think about 'making a distinction' but rather about 'making a description'. According to Maturana, these terms are interchangeable, as at the very hear of each description lies the distinction between the described entity and the rest. Without this distinction, we would not be able to identify the entity as such. Thus, the non-entity or, as termed in the quote above, the *"second entity"* serves as a point of reference which is necessary to identify the 'first entity'. Apparently, first and second entity are interchangeable and thus, we can not tell which one is the entity to be defined and which is the point of reference. Consequently, describing an entity necessarily brings forth a point of reference as well. Ultimately, this point may be either the observer itself or the rest, i.e. the second entity from which the observer distinguishes itself. In this context, the 'I' (ego) also has to be considered as a construction brought forth by an act of distinction.

> *"Living systems are units of interactions; they exist in an ambience. From a purely biological point of view they cannot be understood independently of that part of the ambience with which they interact: the niche; nor can the niche be defined independently of the living system that specifies it."*[93, p.7]

In accordance with the concept of co-evolution of niche and species introduced in subsection 2.2.1, Maturana once more affirms the interdependence between these two entities. In a strict sense, the moment an observer distinguishes between niche and species, both of them are brought into existence. As pointed out previously, the distinction as the observer's creative action generates both the unit and the rest, i.e. first and second entity. Clearly, this also applies to the concept of species and niche.

### 4.2.2 Structure and organization

There are innumerable possibilities to carry out a distinction, just as there are innumerable possibilities to describe something. In order to distinguish, and hence describe, life itself, Maturana introduces a distinction between *organization* and *structure*. To begin with, organization ...

> *"...refers to the relations between components that define and specify a system as a composite unity of a particular class, and determine its properties as such a unity. Hence, the organization of a composite unity specifies the class of entities to which it belongs."*[94, p.3]

While the organization of a unit determines what that unit *is*, i.e. to what class of objects or systems it belongs, a unit's *structure* is merely the realization of a specific organization. More precisely, structure ...

> *"...refers to the actual components and to the actual relations that these must satisfy in their participation in the constitution of a given composite unity."*[94, p.3]

Within the context of Maturana's theory of life, both terms play complementary roles: while the organization of a systems determines the class of the system, e.g. 'living' or 'non-living', the structure of a system represents the purely individual, concrete realization of its organization. Thus, *class-identity* is provided by the organization, while individual identity is found in the system's structure. Hence, the twin terms of structure and organization can be seen as tools used to account for the general problem of differences in uniformity respectively, uniformity despite differences: 'All humans are humans, even though there are no two identical humans'. In the given context, the terms are used to describe the class of all living beings.

Apart from being cognitive subjects, all living beings exhibit, according to Maturana, the same organization, but may differ in their individual structure. In other words, a certain type of organization is the criteria used to distinguish between the living and the non-living, while the differences among organisms are accounted for by structural differences. Note that both structure as well as organization have to be understood in a *mechanistic* way: describing organization and structure of a systems implies describing the mechanism governing that very unit. As we are here discussing the organization of living systems, we are confronted with a mechanistic account of live, and not a e.g. vitalistic one.

As an immediate consequence of such a mechanistic understanding, the concept of *structure-determined systems*, which entails a certain notion of *autonomy*, emerges [3]:

> *"... an external agent that interacts with a composite unity only triggers in it a structural change that it does not determine. Since this is a constitutive condition for composite unities, nothing external to them can specify what happens in them; there are no instructive interactions for composite unities. [...] the structure of a composite unity also determines with which structural configurations of the medium it may interact. [...] It follows from all this that composite unities are structure determined systems in the sense that everything that happens in them is determined by their structure."*[95, p.11]

> *"For an observer, the organization and structure of a structure determined system determine both its domain of states and its domain of perturbations as collections of realizable possibilities."*[94, p.5]

The structure of a system determines the possible states the system can enter, as well as the perturbations from the outside which are possible. Thus, the environment may *"trigger"* changes in the systems behavior through perturbations, but it can not

---

[3] A more recent account of living systems as autonomous systems can be found in [96]. In accordance with Maturana, a mechanistic explanation centered around a specific notion of organization is used to define life in a systems biological context.

*"determine"* or *"instruct"* the system's trajectories. Metaphorically speaking, the environment can only choose which path a system will follow, but it does not determine the domain of possible paths. However, Maturana concedes that some perturbations may *disintegrate* the system, which implies that the system ceases to exist as a unit. However, in this case the environment does not determine the system, for at this point the system is no longer existent.

As a further consequence, the system's possibilities to perceive changes in the environment are determined by the system's structure as well: a modification in the environment which does not trigger any response within the system is - from the point of view of the system - no change at all. Hence, Maturana's notion of structure-determined systems also entails the notion of a *system-referent environment*, in contrast to a *universal physical environment*. Note that this exact distinction is also made by Developmental Systems Theory (cf. subsection 2.2.2).

Indeed, the concept of a system-referent environment can be seen as equivalent to the claim outlined above that every observation is subject dependent. Recall that 'observation' is not merely passive perception, but rather an active interaction with the environment. Hence, structure-determination does not only imply a form of passive autonomy in the flavor of 'we can only see what we can see', i.e. we can only perceive the world in the way our structure permits us to perceive it, but it does also imply an active form of autonomy: the spectrum of possible actions is determined by the system itself.

Now, we have emphasized at various occasions that the role of system and environment, of unit and niche respectively ambience (cf. subsection 4.2.1), are interchangeable. Thus, if confronted with multiple (i.e. at least two) autonomous, structure-determined systems, the question arises how the interaction among these systems can be understood. In such a case, apparently ...

> *"... changes of state of one system become perturbations for the other, and vice versa, in a manner that establishes an interlocked, mutually selecting, mutually triggering domain of state trajectories."*[94, p.5]

Thus, both systems maintain their autonomy, even though their state trajectories are *structurally coupled*. They do not determine each other, but each system selects which of the possible states of the other system will be chosen. Note that not only the set of possible states to chose from is determined by the system itself, but also the set of triggering changes performed by the selecting system: if a change in system A triggers or selects a change in system B depends upon the structure of system B, because it is up to system B to decide if there was a change in system A at all. Stated in other words, the state and the structure of the system itself determines structure and possible states (perturbations) of it's environment. A change in the environment which can not be perceived by the system is no change at all. Thus, *mutual structural coupling* leads to a situation of *mutual invisibility*, which ultimately implies that the precise sequences of states of both involved systems can only be reconstructed as a subject-dependent sequence. Each system, including the observer of two structurally coupled systems, perceives only what it's structure permits to perceive and hence, identifying the 'true' sequence of mutually triggering states is impossible.

This concept of structural coupling also leads to a novel understanding of *adaptation*:

> *"I call structural coupling or adaptation the relation of dynamic structural correspondence with the medium in which a unity conserves its class identity (organization in the case of a composite unity, and operation of its properties in the case of a simple one), and which is entailed in its distinction as it is brought forth by the observer in his or her praxis of living. Therefore, conservation of class identity and conservation of adaptation are constitutive conditions of existence for any unity (entity, system, whole, etc.) in the domain of existence in which it is brought forth by the observer in his or her praxis of living."*[95, p.13]

First of all notice that, in accordance with the Darwinian understanding of the term, *adaptation* does not mean that some organism or systems is 'optimized' or 'well-suited' for some purpose or environment. In fact, adaptation has to be understood as a *constructed* property, which is *derived* from an organisms capacity to survive. As an organism's capacity, we can not see 'adaptation', but rather have to deduce it from a comparison of multiple organism-environment pairs. But even though there seems to be some sort of 'fit' between organism and environment, the only criteria we actually have to define 'adaptation' is the existence of a concrete organism-environment pair. Stated in other word, the only thing common to all organism-environment pairs, which may be used to define adaptation, is their presence. Thus, as long as an organism or species survives, it has to be considered as adapted. Given that we are not able to predict or define in some way what determines the survival of a system *in general*, we are also not able to provide a more precise meaning for 'adaptation'. Hence, adaptation has to be seen as a binary property common to all living systems which somehow managed to *exist as such*, i.e. which managed to conserve their class identity. A system respectively a unit which fails to adapt will loose its class identity and thus, will cease to exist as such, i.e. it will disintegrate.

### 4.2.3 Autopoiesis of living systems

In contrast to the rather well-known concept of *self-organization*, Maturana and Varela introduced the concept of *self-creation*, i.e. of *autopoiesis* as the defining criteria of living systems. Given that the class identity of any system is defined by the system's organization, autopoietic systems are defined by their autopoietic organization. This sort of organization is characterized by Maturana in the following way:

> *"There is a class of dynamic systems that are realized, as unities, as networks of productions (and disintegrations) of components that: (a) recursively participate through their interactions in the realization of the network of productions (and disintegrations) of components that produce them; and (b) by realizing its boundaries, constitute this network of productions (and disintegrations) of components as a unity in the space they specify and in which they exist. Francisco Varela and I called such systems*

> *autopoietic systems, and autopoietic organization their organization (Maturana Varela, 1973). An autopoietic system that exists in physical space is a living system ...*"[94, p.5 et seq.]

Thus, the two key characteristics of autopoietic organization are a.) a recursive network of productions of the system's components and b.) the system's boundaries. A certain resemblance with rather familiar biological concepts, such as metabolic networks and cellular membranes can hardly be neglected and is anything but a coincidence. Indeed, the concept of autopoieses does not reject any findings of todays biological research, but rather provides a novel framework to re-interpret well-established results. In what follows, we will have a brief look at some of this re-interpretations.

First of all, by pointing out that both a *network* of productions (metabolism) as well as a clearly defined *system-boundary* (protective cellular membrane) is needed for live to work, the holistic view of Kauffmann (cf. subsection 3.2.1 on emerging units) is affirmed. Thus, the emergence of a living unit can not be understood as a sequential process, in which e.g. in a first step some system-boundary appeared, which was, in a second step, filled with some network of productions. Both components are equally essential to live and thus, do not allow for a separate origin (cf. [92, p. 47 et seq.]). Hence, based on this view, a continuous, gradual transition from the non-living to living systems appears improbable.

Furthermore, the *circularity*, which provides the basis for the autonomy (cf. subsection 4.2.2) of living systems is omni-present in modern biology: there is hardly any biological process without feedback loops, which are frequently essential for a system's stability. Note that a certain type of circularity is also central to Kauffmann's findings on stable networks (cf. subsection 3.2.2) and hence, a certain type of self-reference (self-regulation) could be seen as necessary to autonomy, stability and homeostasis:

> *"This circular organization constitutes a homeostatic system whose function is to produce and maintain this very same circular organization ...*"[93, p.8]

When considering the above mentioned aspects of autopoietic organization, namely the notion of a *circular, self-referring network* of productions enclosed by some *membrane*, the correspondence to *cellular* organization is obvious. Indeed, Maturana explicitly states that cells are *first order* autopoietic systems (cf. [92]), in contrast to *second* and *third* order autopoietic systems. The term *"metacellular"* is introduced to describe second order autopoietic systems as for example individual organisms. By using the term metacellular Maturana emphasizes that multiple cells, as hidden in the common term 'multicellular', are not yet enough to constitute a second order *unit*. Finally, third order units are systems composed of second order units, as e.g. ant colonies. Thus, the individual cell is a first order unit, a bee made up of cells is a second order unit, and the beehive is a third order unit. Now, Maturana derives the primacy of the individual cell from the following observation: during its life cycle, *every* organism undergoes a single cellular stage (cf. [92]). Hence, the fundamental *unit* in life is not the organism or the population, but the cell. Note that it has been claimed that the major achievement of evolution was the formation of the cell, compared to which everything else, including the formation of higher order organisms, has to be considered as almost trivial.

However, apart from this accordance concerning the cell as the central unit of life, Maturana does not assume an evolutionary position. More precisely, neither *evolution* nor *reproduction* are seen as essential characteristics of life itself:

> *" Evolutionary change in living systems is the result of that aspect of their circular organization which secures the maintenance of their basic circularity, allowing in each reproductive step for changes in the way this circularity is maintained. Reproduction and evolution are not essential for the living organization ..."*[93, p.9]

Thus, when trying to understand life, both reproduction and evolution have to be seen as mere consequences, but not as the underlying principles. Indeed, reproduction can be understood as a side effect of autopoiesis, which in turn allows for evolution as a purely mechanistic process.

Furthermore, Maturanas notion of *inheritance* is based on his distinction between organization and structure. The basic *organization* of life remains unchanged throughout the generations, but as this organization can be realized by a variety of *structures*, change and variability are possible. Indeed, the circular architecture of autopoietic organization is not changed during evolution:

> *"...what changes is the way in which the basic circularity is maintained, and not this basic circularity in itself."*[93, p.8]

Given that autopoiesis is the defining characteristic of life, each living unit must maintain this very type of circular organization in order to maintain its class identity, i.e. in simple terms, each living units needs to survive as such. As long as this requirement is satisfied, all the other aspects may change. Hence, in a certain sense, Maturana provides a minimalistic, almost formal notion of inheritance as the continuing capacity to survive.

## 4.3 Summary

Departing from a thorough discussion of representationism, a constructivstic notion of cognition and science has been developed. Based on this novel understanding, the notion of autopoiesis was introduced in order to describe living systems. This notion, in turn, allowed for the re-interpretation of a series of central biological terms, such as adaptation, evolution and inheritance.

According to Maturana, *representationism* is characterized by the assumption that there is some subject-independent, external reality which can be accessed by the cognitive subject. Based on this access, some internal representation of this reality can be obtained which is regarded as (true) knowledge, if it faithfully reflects (i.e. 'correctly represents') that reality (cf. **section 4.1**).

First of all it has to be noticed that this notion of cognition fails to prove its own validity, i.e. in a representationist's framework, the truth of the above statement can not be shown. In this sense, representationism has to be seen as inconsistent. More precisely, for representationism to work, some sort of direct, subject-independent access to reality itself would be required. However, given that in the context of representationism the possibility of such an access is a precondition for any insight, the existence of this access itself can never proven. In other words, a direct access to reality is a requirement for representationism, but no conclusion of it. The root of this problem lies in the comparison that needs to be carried out in order to distinguish correct and incorrect representations: how to assess if some 'access to reality' is a *direct* one? We would have to *compare* the access in question with ..., well, with *the* direct one. Apparently, the comparison required by representationism is impossible if we do not know yet which access is the direct one, respectively which reality is the objective one. However, without a direct access, correct (true) and incorrect (false) representations can no longer be distinguished. Furthermore, the notion of an objective, external reality becomes metaphysical, as it refers to something which can not be known (cf. **subsection 4.1.1**).

The first conclusion that has to be accepted if the above argument is taken seriously is the 'primacy of knowledge over reality'. We can no longer act as if we were deriving knowledge from reality, but rather have to admit that we are deriving reality from our own, subject dependent knowledge. Whatever reality we thing of, it will always be a *known reality*, for 'we do not see what we do not see'. Hence, any knowledge as well as any reality is always *subject dependent.* Furthermore, the notion of cognition as a somewhat passive perception of reality has to be abandoned in favor of an understanding of cognition as an active, self-referring act of construction. 'Reality' becomes a product of the cognitive subject. Given that this subject is seen as a living unit, epistemological and biological questions become intimately related (cf. **subsection 4.1.2**).

The circular, self-referring aspect of cognition is reflected in Maturana's notion of *science* as the interplay between a specific type of 'question' and 'answer'. The reproach of *solipsism* is considered as meaningless, as it involves applying subject-independent concepts to a subject dependent cognitive domain. Consider for example the statement *A.) "There is no objective truth"*. Claiming that this statement itself can not be proven to be 'objectively true' is rather a misunderstanding of the statement (as a descriptive one), than a valid argument against it (cf. **subsection 4.1.3**).

In a certain sense, the point of departure for constructivism can be seen in the deconstruction of the representationist's epistemological framework and the entailed distinctions. As we have seen above, the notion of an objective reality has been rejected as a metaphysical one. This in turn implies that the strict dichotomy between 'subject' and 'object' is dissolved, as well as the distinction between 'passive perception' and 'active creation' or between 'knowledge' and 'reality'. However, this de-construction of differences leads to some sort of *"all is one"* philosophy, which fails to provide a *scientific* explanation of the phenomenon of cognition. Hence, in a second step Maturana proposes a solution to the problem of cognition, which entails his notion of a living unit as a cognitive subject (cf. **section 4.2**).

During the following (re-)construction process, the first distinction introduced is the definition of an observer, as a unit which *makes distinctions* and thus, creates units. Hence, the very definition of the observer can be seen as a constructive act which brings forth the observer itself (cf. **section 4.2**). This very distinction represents the core element of Maturana's theory of cognition. Apart from this self-description, a description of the observer as a cognitive, biological unit *as if it were external to the observer providing this description* can be provided. I.e. two levels have to be distinguished: a.) strictly speaking, the circularity of knowledge (cognition) affects this very lines as well and hence, each object of cognition, including cognition itself, has to be seen as a construction of the observer (meta-level). Thus, self-description and self-creation (*auto-poiesis*) are equivalent b.) However, making a distinction entails to act *as if* the object of cognition were external to the observer (object level). Thus, the following description of cognitive, living units is located on the object level.

On the object level, Maturana provides a series of novel interpretations of living units. First of all, *autonomy* and *adaptation* of living units is explained via the concepts of structure determined systems and structural coupling (cf. **section 4.2.1**). Given that the structure of a unit determines the space of possible states this unit can assume, the role of the environment is reduced to merely selecting possible states. This structure determination includes cognition as well, i.e. what is and can be known is determined by the observer (cognitive subject) itself, and not by some external reality. However, given that the role of unit and environment are interchangeable, adaptation is seen as the co-evolution of two independent systems, which are structurally coupled (cf. **section 4.2.2**). Furthermore, a living system is characterized as a *network of productions* which is enclosed by a clearly defined *boundary*. In this understanding of an *autopoietic* system, the biological *cell* is identified as the prototype of a living system (first order system), which can be distinguished from organisms (second order systems) or colonies, swarms and the like (third order system). As a consequence, neither evolution nor reproduction (and inheritance) are seen as defining properties of life, but rather as secondary phenomenons generated by the autopoietic structure of living units (cf. **section 4.2.3**).

# Chapter 5

# DISCUSSION

In accordance with his notion of the autopoietic organization of cognition, Maturana also outlined a circular notion of science. In simplified terms, science is solving its very own problems and providing answers to self-formulated questions. In this spirit, we have to diagnose that the sources we employed in this work represent answers to related, yet different questions. While Developmental Systems Theory (cf. chapter 2) concentrates on the relation between development, evolution and inheritance, Kauffman (cf. chapter 3) heavily relies on computer models to clarify the mechanism underlying cellular organization and its origin. Maturana (cf. chapter 4) on the other hand, introduces epistemological aspects into the discussion and attempts to understand life as a cognitive process.

The question that we are pursuing in this work, i.e. *"what are functional, evolutionary units in biology?"*, both in general as well as in molecular-biological networks in particular, is merely partially aligned with the research foci of the above directions respective authors. However, given the central relevance of this question, it has been addressed in some way, direct or indirect, by all of them. Even though different aspect of our question are discussed and highlighted, two fundamentally different notions of a biological unit can be identified: On the one hand, a unit can be understood as an object of evolution and therefore, has to be described in the context of natural selection (cf. **section 5.1**). On the other hand, given that the mechanism of natural selection is not restricted to living units, it appears reasonable to seek alternative ways to capture the specific 'nature' of living units (cf. **section 5.2**). Finally, in **section 5.3**, an attempt is made to apply the obtained results to functional units in molecular-biological networks.

## 5.1 Evolutionary units

The results discussed in this section can be understood as a reply to the following question:

> *What is an evolutionary unit in biology?*

For what concerns this notion of units, two types of results have to be distinguished: First of all, it became clear that the *mechanism of inheritance* as well as the specific *level of selection*, may it be e.g. gene-, group- or behavioral-selection, is completely independent from the mechanism of natural selection. On the other other hand, both the mechanism of inheritance and the level of selection directly entail a specific notion of evolutionary unit (cf. **subsection 5.1.1**), i.e. defining a mechanism of inheritance corresponds to defining an evolutionary unit and vice versa.

Second, the mechanism of selection does indeed impose a series of requirements upon evolutionary units. The most apparent ones are probably best summarized by E. Jablonka (cf. **subsection 2.2.2**):

> *"According to Darwin's theory, in a world in which there are interacting entities with the properties of multiplication, heredity, and heritable variation that affects the chances of multiplication, natural selection will necessarily occur, ..."* [56, p. 9]

Given a population of units which exhibit differing rates of reproduction, and which reproduce in a manner that transfers the reproduction-rate, natural selection will take place. However, apart from this obvious requirements, several other, maybe less apparent implications of the mechanism of natural selection can be identified (cf. **subsection 5.1.2**).

### 5.1.1 Universality of natural selection

As sketched above, natural selection has to be seen as a mechanism which will necessarily occur given that certain very general requirements are fulfilled, namely the presence of a population of units with heritable, variable reproductive capacity. However, several other aspects, including the mechanism of inheritance as well as the specific level of selection are left untouched by natural selection. Moreover, a circular relation between these questions (inheritance, level of selection) and a certain notion of evolutionary unit can be identified.

#### Inheritance

Deciphering the molecular structure of the DNA by Watson and Crick in 1953 was arguably a major step towards answering why offspring resemble their parents. However, hidden behind this answer is a specific notion of evolutionary unit: both 'parent' as well as 'offspring' are understood as *organisms* in the traditional sense. Confronted with the progress in genetics, this very understanding of evolutionary unit changed. Given that only genes are assumed to be be reliably passed on from generation to generation, any non-genetic variation can impossibly become subject to evolution through natural selection. Hence, a gene-centric view of selection resulted from a gene-centric notion of inheritance.

However, given that the mechanism of natural selection is completely indifferent with respect to the mechanism of inheritance, the strict dichotomy between genotype and phenotype, between phylogenesis (evolution) and ontogenesis (development) has to be seen as the result of an unjustified pre-conception of 'evolutionary unit' (cf. **section 2.1**).

#### Level of selection

The problem of the 'level of selection' can be stated as follows (cf. subsection 2.2.3):

> *"... the key question concerns the level of the biological hierarchy at which natural selection occurs. Does selection act on organisms, genes, groups, colonies, demes, species, or some combination of these?"* [62, p. 349]

Given that natural selection can indeed act upon multiple different types of units, the question apparently arises which is 'the ultimate' unit of selection. The only reasonable answers is *multi-level selection* (cf. [62, p. 350]), as natural selection will, under certain circumstances (cf. 5.1), *necessarily* occur. Clearly, the *direction* of selection varies from level to level: e.g. a trait which is beneficial for the individual may be disadvantageous for the group. However, whatever the biological hierarchy pre-supposed in this discussion looks like[1], it has to be noticed that *the hierarchy itself is a result of evolution*. Hence, the question is not to decide if genes *or* organisms *or* groups are the units of selection, but rather to explain how the *"major transitions in evolution"* (cf. [98]), e.g. from single-cellular life to multicellular organisms or from solitary individuals to groups, occurred.

Thus, irrespective of the level of selection that we consider, the *emergence of novel units* of selection needs to addressed (cf. paragraph "The origin of units" in **subsection 5.1.2**). Clearly, this issue can not be explained in general by the theory of natural selection. However, when thinking about the emergence of a unit on a 'higher' level, it is illustrative to notice that, *from the point of view of the unit there is no 'up' or 'down', but just one single environment.* Organs and cells of an animal are as 'external' to this individual as rainfall and predator density. Likewise, the traditional 'inside-outside' distinction has to be abandoned in favor of one single environment in which the unit itself is located.

Hence, a hierarchical notion of the biosphere can not be inferred from a purely evolutionary notion of a unit, but rather has to be rejected as conceptually deceptive in an evolutionary context. Stated in other words, the theory of natural selection distinguishes between unit and environment, but neither implies nor requires a hierarchical structure of the biosphere. Nonetheless, if emergence is understood as a cross-level phenomenon, a theory of emergence would also yield a theory of a hierarchical structured domain.

### 5.1.2 Prerequisites of natural selection

Apart from the apparent prerequisites of natural selection outlined at the beginning of this section (i.e. **section 5.1**), several less obvious implications can be identified.

First and almost trivially, a distinction between unit and environment has to be made to allow for natural selection to work. Second, the units which are shaped by natural selection according to their differing reproductive success have to come into being in the first place. The question of their origin respectively their emergence needs to be addressed. This can indeed be considered as *the* central question of this inquiry. Third, it has been claimed that evolutionary units exhibit a certain modular structure to allow for adaptation through the accumulation of minor improvements.

---

[1] Compare e.g. [97] for a different hierarchical structure of the biosphere.

Within the following paragraphs, we will have a more detailed look at this prerequisites.

### Unit and environment

First of all a distinction between unit and environment, respectively between one unit and the other, has to be made to allow for a reasonable notion of an evolutionary unit. Stated in other words, natural selection is a mechanism acting upon *discrete entities*, and can not be understood as some sort of 'change over time in a homogeneous, continuous media'. Clearly, 'live' could be seen as a continuous phenomenon, but in that case a re-formulation of the theory of natural selection would be needed. Thus, probably the most fundamental distinction which needs to be made to allow for natural selection to work is the distinction between environment and unit, which entails the distinction between multiple different units.

Even though the interdependence between unit and environment, as well as the interchangeability of their roles, have provided important insights such as the notion of *niche-construction* or the rejection of the lock-key model (cf. **subsection 2.2.1**), dissolving the boundary between unit and environment is no option (cf. **subsection 2.2.2**). Differential fitness as differing reproduction rates, together with the complementary notion of selective pressure from the environment - which corresponds to an external-driven change in reproduction rates - are essential components of evolution through natural selection, which do require a well-defined notion of unit.

Thus, instead of adhering to one specific notion of unit, may it be reductionist (gene-centric) or holistic (Developmental Systems - cf. **section 2.1**), it appears reasonable to seek a problem-specific understanding of unit.

### The origin of units

In continuation of the paragraph on the "Level of selection" (cf. **subsection 5.1.1**), the origin of units can be seen in a hierarchical context. More precisely, when seeking a model to explain the emergence of units, we can distinguish two alternative views (cf. **Figure 5.1**), which can be seen as hierarchical insofar as they represent *bottom-up* respectively *top-down* notions of the formation of units.

On the one hand, the formation of a unit can be seen as a process which, departing from multiple interacting low-level units, leads to the formation of a higher-level unit. This corresponds to the traditional notion of emergence and has been exemplified by Kauffman's autocatalytic set theory (cf. **subsection 3.2.1**). Also the matrix of interacting developmental resources reconstructing a life cycle can be seen as such a *bottom-up* notion of emergence (cf. **subsection 2.2.1**). In particular, the concept of *trait groups* (cf. **subsection 2.2.3**), which is based on the selection of mutually beneficial traits, i.e. on *positive feedback*, can be seen as the prototype of this type of emergence.

On the other hand, the formation of a unit can also be seen as based on the introduction of a *distinction* between unit and environment. This concept is primarily driven by epistemological considerations, but can equally be understood as a mechanistic explanation. Both aspect are condensed in Maturana's concept of autopoiesis (cf. **subsection 4.2.3**), which is meant to be understood both as a cognitive as well

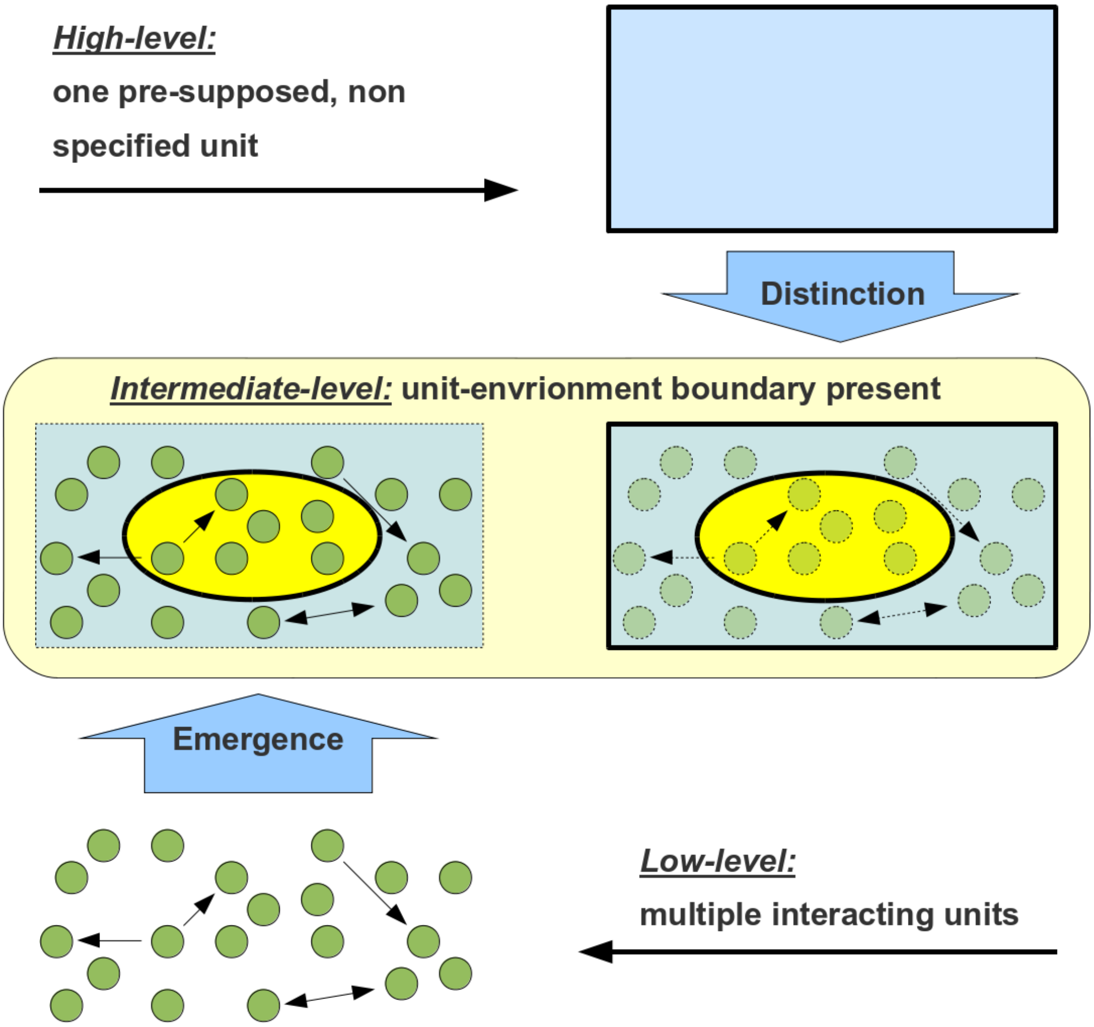


Figure 5.1: *On the left hand side, a bottom-up notion of emergence is illustrated, while on the right hand side, the process of unit-formation understood as internal differentiation (top-down) is depicted. Ultimately both scenarios lead to similar results, namely to the establishment of a unit-environment boundary. For details, refer to subsection 5.1.2, "The origin of units".*

as a biological process. Indeed, both evolution (phylogenesis) as well as development (ontogenesis) can easily be understood as the *internal differentiation* of 'life itself' respectively 'one single organism'.

However, no matter if a bottom-up notion of emergence or a top-down notion of internal differentiation is favored, the pre-existence of some sort of unit(s) is required. Bottom-up emergence relies on multiple low-level units, while internal differentiation requires the presence of a high-level unit. Among the compared concepts, Maturana's notion of autopoiesis as self-creation appears to yield the most suitable approach to this difficulty, in particular if combined with N. Luhman's claim that 'the horizon [high-level unit] always remains invisible'.

### Modular units

Both S. Kauffman (cf. **subsection 3.2.3**) as well as P. Griffiths et al. (cf. **subsection 2.2.3**) have claimed that evolutionary units need to exhibit a certain modular structure respectively need to be composed of *"quasi-independent"* components to allow for natural selection. The argument behind this claim can be recapitulated in the following three steps:

1.) Natural selection acts upon certain *discrete traits*, which also define the dimensions of phenospace.

2.) Given that each position in phenospace is associated with a certain fitness, a change (variation) in phenospace corresponds to a change in fitness. If a small change in phenospace corresponds to a big change in fitness, the resulting fitness landscape does not allow for the *accumulation of minor improvements*. 'Walking uphill' on such a landscape will soon get stuck in a poor local optimum. Likewise, if a big change in phenospace corresponds to a small change in fitness, there is no selective pressure at all, for the fitness landscape is indifferent to any change in phenospace. Thus, phenospace and fitness need to be correlated such that a small change of one trait causes only a small change in fitness, while a big change of the phenotype yields a big difference in fitness.

3.) If an evolutionary unit is *highly integrated*, a change in one single trait greatly alters the fitness of the unit as a whole, for the combination of traits matters. Hence, evolutionary units have to be *modular*, i.e. their traits have to be *quasi-independent* from each other to allow for separate adaptation of the individual traits and accumulative, gradual evolution.

Just by itself, this argument would pose no problem. However, in general, evolutionary units as e.g. organisms are considered to be *highly integrated*. For example, during the evolution of homo sapiens, a strong positive feedback among hands, cognitive capacities and upright walking is assumed. The fine motor skills of the hands are useless as long as the hands are needed for walking, just as the most capable brain is of no benefit without proper tools (speech, hands) to interact with the environment. Indeed, these traits are mutually amplifying their overall utility and thus, can hardly be understood as (quasi-)independent. Consequently, re-examining the above argument appears to be necessary.

The critical step of the above argument seems to be that, departing from the required correlation between phenospace and *fitness* (which we accept), a certain conclusion concerning phenotypic *organization* is drawn, namely quasi-independence. Given that phenospace is defined via discrete traits, this conclusion relies

a.) on an identification of *traits* with *components*, i.e. with *parts* of the evolutionary unit and

b.) on a non-linear notion of the overall fitness. This means that the traits respectively parts of the units can not be adapted independently in order to maximize the overall fitness, i.e. the optimization problem can not be decomposed into partial problems.

Given that *b.)* corresponds to our holistic notion of an evolutionary unit, the root of the apparent contradiction might be hidden in *a.)*. Even though we are currently not able to provide a truly satisfying answer to this problem, the following might serve as a first trace: *Both the dimensions of phenospace as well as the parts (components) that a unit is made of are subject to evolution as well.* During evolution, novel parts emerge within an evolutionary unit, i.e. the internal organization of a unit undergoes a process of differentiation. Thus, the problem of modularity outlined in this paragraph appears to be related to the problem of the emergence of novel units discussed in the previous one.

However, we might tentatively concluded that nature does not only *"select for outcome"* (reproductive success), but also for *how this outcome is achieved* (modular organization). It is tempting to suspect that phenotypic integration itself is subject to natural selection. Even though this solution to the modularity-problem has also been proposed by Kauffman himself as a 'self-tuning of evolution' (cf. **subsection 3.2.3**), it appears to us misleading, given that the above argument precisely targets a *prerequisite* of natural selection. Hence, it appears to us that a different mechanism, complementary to natural selection, has to be proposed to explain how modular yet integrated units necessarily emerge. Note that also a specific *balance between stability and variability* (cf. **subsection 3.2.2**) is a prerequisite of natural selection, and hence can not be understood as its consequence. Thus, concerning the network-model proposed by S. Kauffman to elucidate how this balance is achieved, the question needs to be asked how this type of networks emerged, for natural selection is only capable to *fine tune* an already pre-stabilized balance.

## 5.2 Living units

In contrast to the previous section, the results discussed herein can be understood as a reply to the question:

> *What is a living unit?*

Hence, the evolutionary aspect of biological units is completely neglected and thus, the answer given by Maturana (cf. **section 4.2**) does not involve concepts like *reproduction*, *adaptation*, *inheritance* or *DNA*[2] as the defining criteria of life.

Instead, Maturana proposes *autopoietic organization* as the defining criteria of living units. A living unit is hence defined as a unit which is organized such that it creates itself. This *self-creation* is not understood as a single historical event, but rather as a continuous activity. *Reproduction*, *differential reproduction rates* and hence, *evolution* are seen as secondary phenomena, which are not 'essential' to life, but merely a consequence of continuous autopoiesis.

Furthermore, autopoiesis is understood as a fundamentally cognitive process and thus, living units and cognitive units are seen as identical. This notion of living units

---

[2] Which is in particular remarkably when compared to R. Dawkins gene-centric concept of evolutionary units. Indeed, most biologists would be flabbergasted by this notion of life.

leads to the problem of understanding cognition, which is an intrinsically *circular* problem: how to understand understanding? Maturana concludes that a representationistic approach to this question must fail (cf. **subsection 4.1.1**), because the circularity of cognition (autopoiesis) requires to unite *description* and *construction*: only if cognition is seen as a constructive as well as a descriptive process, a consistent description of cognition (autopoiesis) is possible: as self-description (cf. **subsection 4.1.2**). This was formulated by Maturana as the *"emergence of the observer"* (cf. **subsection 4.2.1**). The central operation which brings forth simultaneously the observer (unit) and its world (environment) is the *distinction*, and by writing (respectively reading) this very lines, according to Maturana, the observer as such emerges.

Besides emphasizing the subject dependent nature of cognition, Maturana rejects *solipsism* as a misunderstanding of constructivism (cf. **subsection 4.1.3**). The misunderstanding of constructivism which leads to solipsism relies in demanding objectivistic properties from a subject dependent cognitive domain: if e.g. someone demands to know how the 'objective truth' of the statement *"There are no objectively true statements"* shall be proven, he (or she) apparently misunderstood the statement by reading it as an objective description. In this context, *science* is no longer the guardian of the truth itself, but merely one specific, socio-culturally determined way to answer self-formulated questions.

Hence, apart from autopoiesis, the scientific answer to the above formulated question is that units are, in general, constructed by the observer. Consequently, the emergence of a unit has to be understood as a cognitive phenomenon, i.e. the question is not how a unit comes into being, but how or why the observer makes the underlying distinction.

Even though this observation highlights an important aspect of a variety of biological units, namely the arbitrariness contained in their definition, the *explanatory potential* of this approach seems to be rather limited. Do we *really* gain a deeper understanding of living units if 'units' in general are seen as necessarily subject dependent construction? Metaphorically speaking, this approach appears to solve the problem by proving that the problem does not exist. Likewise, the notions of *structure determined* (autonomous) systems and *structural coupling* appear to us rather problematic (cf. **subsection 4.2.2**). The distinction between 'defining the possible state transitions' (unit) and 'choosing one of the possible state transitions' (environment respectively another unit) upon which Maturana's notion of autonomy and self-determination is based hardly captures the common notion of *causation*. Even if a system determines the space of possible future states it may enter, as long as the environment *determines* (selects) which of this states the system actually enters, the system can hardly be seen as autonomous.

## 5.3 Functional units in molecular-biological networks

We are now returning to our initial question, namely:

*What are functional, evolutionary units in molecular-biological networks?*

As we will see in the following lines, most of the results presented in the previous sections (i.e. **sections 5.1** and **5.2**) can directly be applied to molecular networks. But before doing so, it is worth recapitulating some *biological* aspects of molecular-biological networks (cf. **section 1.1**).

First of all there is a broad agreement among biologists that molecular networks exhibit a modular (cf. [10, 17, 12, 19]), possibly even hierarchical structure (cf. [99, 4, 8]). Hence, it is not the network as a whole which is responsible for every aspect of cellular behavior, but rather do subnetworks exist which performer specific, dedicated functions. These subnetworks, also referred to as *"functional modules"* or *"units"*, are assumed to be subject to *evolution* and hence, they can be considered as a trait passed on from parents to offspring (cf. [20, 28, 18]). As functional traits are conserved by natural selection, a cross-species comparison of molecular interaction networks can be used to identify similar, i.e. conserved and thus, hopefully also functional subnetworks. Stated in other words, as functionality implies conservation, conservation provides strong evidence that the corresponding trait (body plan, DNA sequence, protein structure, network architecture, etc.) is functional (cf. [29, 30]). An alternative, less convincing approach to identify functional modules in molecular networks is based on the assumption that functional modules are dense clusters which are only loosely related to the rest of the network (cf. e.g. [22]).

This biological understanding of functional modules as evolutionary units is in accordance with our notion of the *universality of natural selection* (cf. **subsection 5.1.1**). As outlined in the paragraph on multiple levels of selection, functional units will be subject to natural selection, given that several requirement (cf. ibid.) are met. Hence, not only the nodes of the network (proteins, genes or the like) or the network as a whole have to be understood as shaped by evolution, but also the modules within the network. However, as already pointed out previously (cf. **subsection 5.1.2**), the most basic requirement of natural selection is the presence of a unit as such.

Interestingly, the presence or 'existence' of functional modules within a network also represents a requirement for the evolution of the network as a whole. As outlined in the paragraph on *modular units* (cf. ibid.), there is strong evidence supporting the notion of modular evolutionary units. According to this arguments, only modular units are capable of evolution through natural selection and hence, if a complete molecular-biological network (e.g. the metabolic network of a cell) is considered as an evolutionary unit, the existence of modules within this network is an immediate consequence.

Thus, both for the evolution of the unit itself, as well as for the evolution of the network it is embedded in, the *origin of the unit* (cf. **subsection 5.1.2**) has to be clarified. This origin respectively emergence can, once more, be understood in two different ways: Either as an *assembly of multiple low-level entities*, such as genes, proteins and the like, or as a differentiation of high-level functionality, i.e. as a *division of cellular labor* (cf. **figure 5.1**). While the first approach leads to a *structural* notion of units, the second one yields a *functional* understanding[3]. However, no matter if the *origin of a*

---

[3] Just as e.g. a chair can be either be defined via certain *structural* characteristics or via the *function* it fulfills.

*unit* is understood as an assembly of low-level *units* or as a division of high-level *units*, in both cases the *pre-existence* of some unit(s) is required. A first step towards solving this apparently paradox situation might be the following *rejection of hierarchies*, i.e. the claim that in a certain sense there is no difference between 'low-level assembly' and 'high-level division'.

As pointed out before (cf. **subsection 5.1.2**), the notion of a hierarchy in which high-levels and low-levels can be distinguished does not follow from the theory of natural selection. Indeed, when focusing on the evolutionary aspects of functional network modules, the only applicable distinction is the one between unit and environment. As a consequence, the environment of a unit has to be understood as the union over both high- *and* low-level entities. Hence, the environment of a functional unit can not be understood as 'all the nodes which *are not part* of the unit', just as the unit itself can not be understood as 'all the nodes which *are part* of the unit'. A protein such as *p53*, which plays a central role in apoptosis (cf. figure 1.2), is as external to *apoptosis as a functional unit* as the general condition of the cell. Stated in other words, a mutation in p53 is just as much an environmental factor to apoptosis as a change in the general state of the cell which might trigger the process, e.g. toxic stress. Likewise, the individual cells making up a certain organ are as external to this organ as the organism around it.

As the ultimate consequence, we have to agree with Maturana (cf. **subsection 5.2**) who proposes a *subject dependent*, but nonetheless *non-arbitrary* understanding of functional units. As follows from the above rejection of hierarchical structures, the only criteria left to define a unit is the *distinction* between unit and environment. Thus, the most basic requirement every unit has to fulfill to be rightfully considered as such is constructed through the observer respectively *subject dependent* operation of distinction. What makes apoptosis a unit is not the existence of a 'natural boundary' between apoptosis and the remaining molecular processes of a cell, but rather the *utility* of the underlying distinction for the observer. As the term 'utility' suggests, the operation of distinction is *non-arbitrary*, for the *utility* (functionality) of a distinction depends on the cognitive context of the observer (cf. Maturana's notion of *science* in subsection 4.1.3).

However, in the cognitive context of applied sciences, the utility of Maturana's understanding of living units as *cognitive units* (cf. ibid.) appears to have only limited potential. In particular, the status of functional units in molecular-biological networks remains unclear, given that his focus is clearly on *cells* as the primary units of live and autopoiesis.

# Appendix A

# ABSTRACT IN ENGLISH

The rise of high-throughput techniques in biological research has allowed the simultaneous assessment of multiple molecular parameters and eventually led to the reconstruction of molecular-biological interaction networks. These networks are assumed to be organized in a modular manner, where 'modules' are seen as dedicated functional units responsible for specific cellular behaviors. Hence, the identification of functional units is a crucial step towards understanding the molecular mechanisms behind phenotypic observables.

Various screening methods exists to detect functional units in networks, but in the context of applied sciences, none of these methods involves a thorough discussion of the employed notion of *functional unit*. In order to address this very need, i.e. to answer the question *'what are functional units in biology?'*, the following authors respectively approaches were investigated: (A) *Developmental Systems Theory* is a branch of the philosophy of biology which rejects a gene-centric notion of biological unity and replaces it with the concept of a 'developmental system'. As this concept typically extends the traditional phenotype (i.e. the organism), the need for a principled definition of biological unity is inherent to this approach. (B) In contrast to this, *Stuart Kauffman* concentrates on the limits of natural selection as a means to explain the order present in biological units. Given that in biology, 'functionality' ultimately refers to reproductive success, his investigations are centered around an evolutionary understanding of units. (C) Finally, departing from epistemological considerations, *Humberto Maturana* proposes the concept of autopoiesis (self-creation) as the (self-)defining criteria of living units.

The major results obtained from a comparison of this sources are the following ones: Given that natural selection is a generic mechanism which is not limited to any specific *level of selection* (genetic, cellular, ...), seeking the 'ultimate unit of evolution' is meaningless. Hence, if functional units satisfy certain preconditions, most notably a *modular structure*, they are subject to evolution as well. However, their *origin* is not clarified by natural selection. In this respect, a bottom-up notion of emergence as the *'assembly of low-level entities'* (e.g. genes or proteins) can be distinguished from a top-down notion of emergence as the *'division of cellular labor'*. From a systems-theoretic perspective, both notions can be understood as complementary, which is in accordance with Maturana's claim that ultimately, each 'unit' is constructed by the observer dependent operation of distinction. From this point of view, a structural criteria to identify functional units in general, as seeked by the applied sciences, can not exist.

# Appendix B

# ABSTRACT IN GERMAN

Durch die Entwicklung von Hochdurchsatz-Verfahren in der Biologie wurde die gleichzeitige Messung einer Vielzahl molekularer Parameter möglich, was zur Rekonstruktion molekular-biologischer Interaktions-Netzwerke führte. Man geht davon aus, daß diese Netzwerke in *funktionale Einheiten* bzw. *Module* zerlegt werden können, welche jeweils für ein spezifisches zelluläres Verhalten verantwortlich sind. Die Identifikation dieser Einheiten ist daher von zentraler Relevanz.

Im naturwissenschaftlichen Kontext existieren verschiedene Methoden zur Erkennung derartiger Einheiten, aber keine dieser Methoden beinhaltet eine explizite Definition des zugrunde liegenden Verständnisses von *'funktionaler Einheit'*. Um genau dieses Defizit zu adressieren, d.h. um die Frage zu beantworten, *was im biologischen Bereich unter funktionaler Einheit zu verstehen ist*, wurden die folgenden Disziplinen bzw. Autoren herangezogen: (A) *Developmental Systems Theory* ist ein Zweig der Philosophie der Biologie, der ein gen-zentrisches Verständnis von biologischer Einheit ablehnt und durch das Konzept des 'Developmental Systems' ersetzt. (B) *Stuart Kauffman* hingegen konzentriert sich auf jene Aspekte einer biologischen Einheit, die nicht mittels natürlicher Selektion zu erklären sind und diskutiert Selbst-Organisation als das zugrunde liegende Prinzip. (C) *Humberto Maturana* positioniert die Frage nach der biologischen Einheit im epistemologischen Kontext und führt somit erkenntnistheoretische Aspekte in die Diskussion ein. Eine lebendige Einheit wird von ihm immer auch als kognitive Einheit gesehen.

Die zentralen Ergebnisse aus einem Vergleich dieser Quellen sind folgende: Natürliche Selektion ist ein *allgemeingültiger Mechanismus*, der in unterschiedlichsten Bereichen gleichermaßen wirksam ist, sofern gewisse Bedingungen erfüllt sind. Insbesondere müssen die Einheiten der Selektion *modular* aufgebaut sein. Der *Ursprung der Einheiten* wird durch natürliche Selektion nicht geklärt. Bezüglich dieser Frage lassen sich zweierlei Arten von Emergenz unterscheiden: als Kombination vieler 'kleiner' Einheiten (Gene, Proteine, ...) oder als Ausdifferenzierung einer 'großer' Einheit ('zelluläre Arbeitsteilung'). Aus systemtheoretischer Perspektive, i.e. sofern der Begriff der Einheit auf die fundamentale Unterscheidung zwischen 'innen' und 'außen' reduziert wird, werden beider Varianten als komplementär erkennbar. Dies wird gestützt durch Maturan's Feststellung daß jegliche Einheit durch die Beobachter abhängige Operation der Unterscheidung erschaffen wird. Insofern muss die Suche nach allgemeinen strukturellen Merkmalen funktionaler Einheiten in Netzwerken als problematisch angesehen werden.

# Appendix C

# CURRICULUM VITAE

## Personal

Born on may 24, 1980 in Munich, Germany.

## Education

Studying Philosophy and Linguistic (French, Spanish, Portuguese) in Munich and Dresden 2000 - 2004

Studying Computational Intelligence at the TU - Vienna, 2004 - 2009

Studying Philosophy at the University of Vienna, 2007 - 2011

Academic Excellence Scholarship of the TU - Vienna, Oct. 2007

Academic Excellence Scholarship of the TU - Vienna, Oct. 2009

Master Thesis: *Identifying functional molecular-biological networks using classification algorithms, P-Systems and topology-based approaches*

M.Sc., grade: 1.0, Dec. 2009

## Employment

Internship at emergentec biodevelopment, Feb. 2008 – Feb. 2009

Research assistant at the TU of Vienna, Feb. 2009 – Nov. 2009

Ph.D. position at emergentec biodevelopment, Dec. 2009 – Mar. 2010

Ph.D. position at the Medical University of Innsbruck, Apr. 2010 - Dec. 2010

Ph.D. position at emergentec biodevelopment, Jan. 2011 - today

## Contributions

Mühlberger I. et al., *Omics profile integration for characterizing kidney diseases*, exhibited at the 17th International Conference on Intelligent Systems for Molecular Biology (ISMB ECCB 2009)

Bernthaler A. et al., *Analysis of omics profiles using protein interaction networks*, exhibited at the 11th International Conference on molecular systems biology (ICMSB 2009)

Bernthaler A. et al., *A dependency graph analysis in the context of disease classification*, exhibited at the 11th International Conference on Systems Biology (ICSB 2009)

Mönks K. et al., *Inference of Drug-Disease-Gene Networks from Literature Data*, exhibited at the International Metabolomics Austria (IMA 2010)

Kainz A. et al., *Regulierung des m-TOR Wegs bei Nierentransplantation*, exhibited at Austrotransplant 2010

Mühlberger I. et al., *Integrative bioinformatics analysis of proteins associated with the cardiorenal syndrome*, International Journal of Nephrology, 2011:809378, Oct. 2010

Mönks K. et al., *Computational reconstruction of protein interaction networks*, in Dehmer M., Emmert-Streib F., Graber A. & Salvador A. (Eds.), *Applied Statistics for Network Biology: Methods in Systems Biology*, April 2011

Mönks K. et al., *Mycophenolate moeftil associated molecular profiles and diseases*, Journal of Computer Science and Systems Biology, accepted 2011

Bernthaler A. et al., *Linking molecular feature space and disease terms for the immunosuppressive drug rapamycin*, Molecular BioSystems, submitted 2011

## International Experience

Studying at the Pontificia Universidad Catolica del Peru, Feb. 2003 - Sept. 2003

Studying at the Pontificia Universidade Catolica de Sao Paulo, Oct. 2003 - Feb. 2004

# Bibliography

[1] Neil Hall. Advanced sequencing technologies and their wider impact in microbiology. *The Journal of Experimental Biology*, 210(Pt 9):1518–1525, May 2007.

[2] Gert Mayer, Georg Heinze, Harald Mischak, Merel E. Hellemons, Hiddo J. Lambers Heerspink, Stephan J. L. Bakker, Dick de Zeeuw, Martin Haiduk, Peter Rossing, and Rainer Oberbauer. Omics-bioinformatics in the context of clinical data. *Methods in Molecular Biology (Clifton, N.J.)*, 719:479–497, 2011.

[3] Henning Hermjakob, Luisa Montecchi-Palazzi, Gary Bader, Jerome Wojcik, Lukasz Salwinski, Arnaud Ceol, Susan Moore, Sandra Orchard, Ugis Sarkans, Christian von Mering, Bernd Roechert, Sylvain Poux, Eva Jung, Henning Mersch, Paul Kersey, Michael Lappe, Yixue Li, Rong Zeng, Debashis Rana, Macha Nikolski, Holger Husi, Christine Brun, K. Shanker, Seth G. N. Grant, Chris Sander, Peer Bork, Weimin Zhu, Akhilesh Pandey, Alvis Brazma, Bernard Jacq, Marc Vidal, David Sherman, Pierre Legrain, Gianni Cesareni, Ioannis Xenarios, David Eisenberg, Boris Steipe, Chris Hogue, and Rolf Apweiler. The HUPO PSI's molecular interaction format–a community standard for the representation of protein interaction data. *Nature Biotechnology*, 22(2):177–183, February 2004.

[4] Soon-Hyung Yook, Zoltan N. Oltvai, and Albert-Laszlo Barabasi. Functional and topological characterization of protein interaction networks. *Proteomics*, 4(4):928–942, 2004.

[5] Hiroaki Kitano. Systems biology: A brief overview. *Science*, 295(5560):1662–1664, March 2002.

[6] Trey Ideker, Timothy Galitski, and Leroy Hood. A new approach to decoding life: Systems biology. *Annual Review of Genomics and Human Genetics*, 2(1):343–372, 2001.

[7] Douglas B. Kell. Metabolomics and systems biology: making sense of the soup. *Current Opinion in Microbiology*, 7(3):296–307, June 2004.

[8] Zoltan N. Oltvai and Albert-Laszlo Barabasi. Systems biology: Life's complexity pyramid. *Science*, 298(5594):763–764, October 2002.

[9] Derek Gatherer. So what do we really mean when we say that systems biology is holistic? *BMC Systems Biology*, 4:22, 2010.

[10] Shi-Hua Zhang, Chao Wu, Xia Li, Xi Chen, Wei Jiang, Bin-Sheng Gong, Jiang Li, and Yu-Qing Yan. From phenotype to gene: detecting disease-specific gene functional modules via a text-based human disease phenotype network construction. *FEBS Letters*, 584(16):3635–3643, August 2010.

[11] Hiroaki Kitano. Computational systems biology. *Nature*, 420(6912):206–210, November 2002.

[12] Wenyuan Li, Ying Liu, Hung-Chung Huang, Yanxiong Peng, Yongjing Lin, Wee-Keong Ng, and Kok-Leong Ong. Dynamical systems for discovering protein complexes and functional modules from biological networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics / IEEE, ACM*, 4(2):233–250, June 2007.

[13] Albert-Laszlo Barabasi and Zoltan N. Oltvai. Network biology: understanding the cell's functional organization. *Nature Review Genetics*, 5(2):101–113, February 2004.

[14] Mark Newman, Albert-Laszlo Barabasi, and Duncan J. Watts. *The Structure and Dynamics of Networks*. Princeton University Press, 1 edition, April 2006.

[15] Shihua Zhang, Guangxu Jin, Xiang-Sun Zhang, and Luonan Chen. Discovering functions and revealing mechanisms at molecular level from biological networks. *Proteomics*, 7(16):2856–2869, August 2007.

[16] Alexander Platzer, Paul Perco, Arno Lukas, and Bernd Mayer. Characterization of protein-interaction networks in tumors. *BMC Bioinformatics*, 8:224–224, 2007.

[17] Carol V. Robinson, Andrej Sali, and Wolfgang Baumeister. The molecular sociology of the cell. *Nature*, 450(7172):973–982, December 2007.

[18] Monica Campillos, Christian von Mering, Lars Juhl Jensen, and Peer Bork. Identification and analysis of evolutionarily cohesive functional modules in protein networks. *Genome Research*, 16(3):374–382, March 2006.

[19] Like Fokkens and Berend Snel. Cohesive versus flexible evolution of functional modules in eukaryotes. *PLoS Computational Biology*, 5(1):e1000276, January 2009.

[20] Jose B. Pereira-Leal, Emmanuel D. Levy, and Sarah A. Teichmann. The origins and evolution of functional modules: lessons from protein complexes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 361(1467):507–517, March 2006.

[21] Jose M. Peregrin-Alvarez, Xuejian Xiong, Chong Su, and John Parkinson. The modular organization of protein interactions in escherichia coli. *PLoS Computational Biology*, 5(10):e1000523, October 2009.

[22] Victor Spirin and Leonid A. Mirny. Protein complexes and functional modules in molecular networks. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21):12123–12128, October 2003.

[23] Carl F. Schaefer. Pathway databases. *Annals of the New York Academy of Sciences*, 1020:77–91, May 2004.

[24] Paul D. Thomas, Michael J. Campbell, Anish Kejariwal, Huaiyu Mi, Brian Karlak, Robin Daverman, Karen Diemer, Anushya Muruganujan, and Apurva Narechania. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Research*, 13(9):2129–2141, September 2003.

[25] M. Kanehisa and S. Goto. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1):27–30, January 2000.

[26] Minoru Kanehisa, Susumu Goto, Masahiro Hattori, Kiyoko F. Aoki-Kinoshita, Masumi Itoh, Shuichi Kawashima, Toshiaki Katayama, Michihiro Araki, and Mika Hirakawa. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Research*, 34(Database issue):D354–357, January 2006.

[27] Minoru Kanehisa, Susumu Goto, Miho Furumichi, Mao Tanabe, and Mika Hirakawa. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Research*, 38(Database issue):D355–360, January 2010.

[28] S. R. Paladugu, V. Chickarmane, A. Deckard, J. P. Frumkin, M. McCormack, and H. M. Sauro. In silico evolution of functional modules in biochemical networks. *Systems Biology*, 153(4):223–235, July 2006.

[29] Johannes Berg and Michael Lässig. Cross-species analysis of biological networks by bayesian alignment. *Proceedings of the National Academy of Sciences*, 103(29):10967 –10972, July 2006.

[30] Wenhong Tian and Nagiza F. Samatova. Pairwise alignment of interaction networks by fast identification of maximal conserved patterns. *Pacific Symposium on Biocomputing*, pages 99–110, 2009.

[31] R. C. Lewontin. The units of selection. *Annual Review of Ecology and Systematics*, 1:1–18, January 1970.

[32] Peter Godfrey-Smith. Conditions for evolution by natural selection. *Journal of Philosophy*, 104:489 – 516, 2007.

[33] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1st edition, December 1982.

[34] J. Maynard Smith. Optimization theory in evolution. *Annual Review of Ecology and Systematics*, 9:31–56, January 1978.

[35] Neil A. Campbell and Jane B. Reece. *Biology*. Benjamin Cummings, 7th edition, December 2004.

[36] Lenny Moss. Deconstructing the gene and reconstructing molecular developmental systems. In P.E. Griffiths, R.D. Gray, and S. Oyama, editors, *Cycles of Contingency*, pages 85 — 97. MIT Press, 2003.

[37] Richard Dawkins. *The Selfish Gene*. Oxford University Press, USA, 2nd edition, October 1990.

[38] Paul Griffiths and Russell Gray. The developmental systems perspective: Organism-environment systems as units of development and evolution. In M. Pigliucci and K. Preston, editors, *The Evolutionary Biology of Complex Phenotypes*. Oxford University Press, USA, 2004.

[39] Paul E. Griffiths and Russell D. Gray. Discussion: Three ways to misunderstand developmental systems theory. *Biology and Philosophy*, 20:417 – 425, 2005.

[40] Christoph Bock and Thomas Lengauer. Computational epigenetics. *Bioinformatics (Oxford, England)*, 24(1):1–10, January 2008.

[41] P.E. Griffiths and R.D. Gray. Darwinism and developmental systems. In P.E. Griffiths, R.D. Gray, and S. Oyama, editors, *Cycles of Contingency*, pages 195 – 218. MIT Press, 2003.

[42] Susan Oyama, Paul E. Griffiths, and Russell D. Gray. *Cycles of Contingency: Developmental Systems and Evolution*. The MIT Press, March 2003.

[43] Daniel R. Brooks and E. O. Wiley. *Evolution As Entropy*. University Of Chicago Press, 1st edition, October 1988.

[44] C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, July 1948.

[45] J. Maynard Smith and E. Szathmary. *The Major Transitions in Evolution*. Oxford University Press, 1995.

[46] Susan Oyama. *The Ontogeny of Information: Developmental Systems and Evolution*. Duke University Press Books, 2nd edition, February 2000.

[47] R.C. Lewontin. Adaptation. *Scientific American*, 239:212 – 230, 1978.

[48] R.C. Lewontin. The organism as the subject and object of evolution. *Scientia*, 118:65 – 82, 1983.

[49] R.C. Lewontin. *The Triple Helix: Gene, Organism, and Environment*. Harvard University Press, 1st edition, 1998.

[50] Peter Godfrey-Smith. Adaptationism and the power of selection. *Biology and Philosophy*, 14(2):181–194, April 1999.

[51] Peter Godfrey-Smith. Three kinds of adaptionism. In *Adaptionism and Optimality*, pages 335–357. Cambridge University Press, 2001.

[52] K.N. Laland, F.J. Odling-Smee, and M.W. Feldman. Niche construction, ecological inheritance, and cycles of contingency in evolution. In P.E. Griffiths, R.D. Gray, and S. Oyama, editors, *Cycles of Contingency*, pages 117 – 126. MIT Press, 2003.

[53] Jared Diamond. *Collapse: How Societies Choose to Fail or Succeed.* Viking Adult, 1st edition, December 2004.

[54] J. E. Lovelock. *Gaia: A New Look at Life on Earth.* Oxford University Press, USA, reprint edition, December 1987.

[55] Thomas Pradeu. What is an organism? An immunological answer. *History and Philosophy of the Life Sciences*, 32(2-3):247–267, 2010.

[56] Eva Jablonka. The systems of inheritance. In Susan Oyama, P.E. Griffiths, and R.D. Gray, editors, *Cycles of Contingency*, pages 99 — 116. MIT Press, 2003.

[57] J.S. Robert, B.K. Hall, and W.M. Olson. Bridging the gap between develomental systems theory and evolutionary developmental biology. *BioEssays*, 23:954 – 962, 2001.

[58] F. Craver Carl and William Bechtel. Top-down causation without top-down causes. *Biology and Philosophy*, 22:547–563, 2007.

[59] Max Kistler. Mechanisms and downward causation. *Philosophical Psychology*, 22(5):595, 2009.

[60] Stuart Glennan. Mechanisms, causes, and the layered model of the world. *Philosophy and Phenomenological Research*, 81(2):362–381, 2010.

[61] Elliott Sober and David Sloan Wilson. A critical review of philosophical work on the units of selection problem. *Philosophy of Science*, 61(4):534–555, December 1994.

[62] S. Okasha. Recent work on the levels of selection problem. *Human Nature Review*, 3:349 – 356, 2003.

[63] David Sloan Wilson. Biological communities as functionally organized units. *Ecology*, 78(7):2018–2024, 1997.

[64] J. Gottlieb and G. R. Raidl. Characterizing locality in Decoder-Based EAs for the multidimensional knapsack problem. *In Proceedings of artificial evolution: Fourth European Conference*, pages 38—52, 1999.

[65] Robert N. Brandon. The units of selection revisited: The modules of selection. *Biology and Philosophy*, 14(2):167 – 180, 1999.

[66] Franz M. Wuketits. *Charles Darwin und der Darwinismus.* Beck'sche CH Verlagsbuchhandlung Oscar Beck, September 2005.

[67] S.J. Gould and N. Edredge. Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology*, 3(2):115 — 151, 1977.

[68] N. Edredge and S.J. Gould. Punctuated equilibria: An alternative to phyletic gradualism. In *Models in Paleobiology*, pages 82–115. Freeman, Cooper and Co., 1972.

[69] Paul E. Griffiths. The philosophy of molecular and developmental biology. In P. Machamer and M. Silberstein, editors, *The Blackwell Guide to the Philosophy of Science*, pages 252–271. Blackwell Publishers Ltd, 2008.

[70] Stuart Kauffman. *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. Oxford University Press, USA, August 1995.

[71] Erwin Schrödinger. *Was ist ein Naturgesetz? Beiträge zum naturwissenschaftlichen Weltbild*. Oldenbourg, 6th edition, July 2008.

[72] George Wald. The origin of life. *Scientific American*, pages 45 – 54, August 1954.

[73] Sir Fred Hoyle and Chandra Wickramasinghe. *Evolution from Space*. Flamingo, August 1983.

[74] D. Segre and D. Lancet. Composing life. *EMBO Reports*, 1(3):217–222, September 2000.

[75] Sebastian Kühner, Vera van Noort, Matthew J. Betts, Alejandra Leo-Macias, Claire Batisse, Michaela Rode, Takuji Yamada, Tobias Maier, Samuel Bader, Pedro Beltran-Alvarez, Daniel Castano-Diez, Wei-Hua Chen, Damien Devos, Marc Güell, Tomas Norambuena, Ines Racke, Vladimir Rybin, Alexander Schmidt, Eva Yus, Ruedi Aebersold, Richard Herrmann, Bettina Böttcher, Achilleas S. Frangakis, Robert B. Russell, Luis Serrano, Peer Bork, and Anne-Claude Gavin. Proteome organization in a genome-reduced bacterium. *Science (New York, N.Y.)*, 326(5957):1235–1240, November 2009.

[76] Howard Ochman and Rahul Raghavan. Systems biology. Excavating the functional landscape of bacterial cells. *Science (New York, N.Y.)*, 326(5957):1200–1201, November 2009.

[77] D. Segre, D. Lancet, O. Kedem, and Y. Pilpel. Graded autocatalysis replication domain (GARD): kinetic analysis of self-replication in mutually catalytic sets. *Origins of Life and Evolution of the Biosphere: The Journal of the International Society for the Study of the Origin of Life*, 28(4-6):501–514, October 1998.

[78] Daniel Segre, Dafna Ben-Eli, Yitzhak Pilpel, Ora Kedem, and Doron Lancet. Gardobes: Primordial cell nano-precursors with organic catalysis, compositional genome and capacity to evolve. *Instruments, Methods, and Missions for Astrobiology II*, 3755:144–162, 1999.

[79] Barak Shenav, Arren Bar-Even, Ran Kafri, and Doron Lancet. Polymer gard: Computer simulation of covalent bond formation in reproducing molecular assemblies. *Origins of Life and Evolution of Biospheres*, 35:111 – 133, 2005.

[80] Daniel Segre and Doron Lancet. Mutually catalytic amphiphiles: simulated chemical evolution and implications to exobiology. In J.Chela-Flores and F.Raulin, editors, *Exobiology: Matter, Energy and Information in the Origin and*

*Evolution of Life in the Universe*, pages 123 – 131. Kluwer Academic Publishers, The Netherlands, 1998.

[81] D. Segre, D. Ben-Eli, and D. Lancet. Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. *Proceedings of the National Academy of Sciences of the United States of America*, 97(8):4112–4117, April 2000.

[82] Carlos Gershenson. Introduction to random boolean networks. *Workshop and Tutorial Proceedings, Ninth International Conference on the Simulation and Synthesis of Living Systems (ALife IX).*, IX:160–173, August 2004.

[83] Stuart Kauffman, Carsten Peterson, Björn Samuelsson, and Carl Troein. Random boolean network models and the yeast transcriptional network. *Proceedings of the National Academy of Sciences of the United States of America*, 100(25):14796–14799, December 2003.

[84] Stefan Bornholdt. Boolean network models of cellular regulation: prospects and limitations. *Journal of the Royal Society*, 5 Suppl 1:85–94, August 2008.

[85] Graham J. Hickman and T. Charlie Hodgman. Inference of gene regulatory networks using boolean-network inference methods. *Journal of Bioinformatics and Computational Biology*, 7(6):1013–1029, December 2009.

[86] J. E. S. Socolar and S. A. Kauffman. Scaling in ordered and critical random boolean networks. *Physical Review Letters*, 90(6):68–72, February 2003.

[87] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. OUP USA, September 1993.

[88] S. A. Kauffman. Antichaos and adaptation. *Scientific American*, 265(2):78–84, August 1991.

[89] Thomas Junker. *Die Evolution des Menschen*. Beck, 2nd edition, January 2009.

[90] Heinz von Foerster. Entdecken oder erfinden: Wie läßt sich verstehen verstehen? In *Einführung in den Konstruktivismus*, pages 41–88. Piper München, 1992.

[91] H.R. Maturana. Cognitive strategies. In *Unity and Diversity of Man*. Le Seuil, 1974.

[92] Humberto R. Maturana and Francisco J. Varela. *Der Baum der Erkenntnis. Die biologischen Wurzeln des menschlichen Erkennens*. Fischer, 3rd edition, April 2010.

[93] H.R. Maturana and F.J. Varela. *Autopoiesis and Cognition: The Realization of the Living*. Springer, 1st edition, August 1991.

[94] H.R. Maturana. Biology of language: The epistemology of reality. In *Psychology and Biology of Language and Thought*. Academic Press, 1978.

[95] H.R. Maturana. Ontology of observing, the biological foundations of self-consciousness and the physical domain of existence. *Conference Workbook: Texts in Cybernetics, American Society For Cybernetics Conference*, 1988.

[96] William Bechtel. Biological mechanisms: organized to maintain autonomy. In Fred Boogerd, Frank J. Bruggeman, Jan-Hendrik S. Hofmeyr, and H.V. Westerhoff, editors, *Systems Biology: Philosophical Foundations*, pages 269—302. Elsevier Science, 1st edition, May 2007.

[97] Alwyn Scott. Biological reductionism. In *Alife VIII Workshop Proceedings*, pages 73–81, 2002.

[98] John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. Oxford University Press, USA, February 1998.

[99] Erzsebet Ravasz and Albert-Laszlo Barabasi. Hierarchical organization in complex networks. *Physical Review E*, 67(2):026112, February 2003.