Level 8

From Building Blocks to *Theories of Everything*

When studying the cognitive system in the neocybernetic perspective, as was done in the previous chapter, some interesting results can be found. However, when the semantic grounding is left floating, so that concepts are defined contextually only in terms of other ones, there is the annoying question haunting: "So what?" The resulting models are just computational constructs; intuitively they cannot have very much to do with "real" intelligence that has to be grounded on flesh and bones. This problem does not only apply to the cognitive systems: Generally, one has problems if trying to capture behaviors in complex domains where semantics is detached. The lower-level system one is studying, the more there is need to attack the deep coupling between the system and the real world. Only if the domain-area semantics is to some extent captured, the computing machinery can reveal something relevant and non-trivial that has not been programmed in the code in the beginning.

What are these unexpected results — and what does this capturing of semantics mean? The problem here is that the essence of domain-area semantics assumedly is different in different domains, and no generic approaches perhaps can be determined. However, specific examples can be illuminating, giving hints of what cybernetics is all about after all.

Here, a very detailed case is studied, the environment with its semantics being implemented; the environment is that of complex molecules, and the application domain is modeling the sequences of amino acid sequences corresponding to the translated genetic code. The sequences of atoms in the molecules are codes themselves, and the objective here is to interpret that "language of molecules" that is interpreted as functionalities in the chemical environment. From the point of view of understanding living systems, the case of amino acids is crucial, because they determine how the final functional proteins are folded, thus dictating their structure and function. The protein structures are the basis for all biochemical processes, and they are the basic structures in all living organisms. Understanding such codes is the key to real "bioinformatics".

And if the principles of life can be reduced back to quantum physics, perhaps

there is life in large-scale physical systems, too? The wild speculations in the end reveal the value of analogues as a source of intuition.

8.1 Computationalism cybernetized

Ludwig Wittgenstein observed that the language determines the limits of one's thinking: If there are no appropriate concepts, there is no way to express oneself. Similarly, one cannot discuss complex systems without appropriate concepts. However, in the adopted evolutionary framework, this thinking goes much deeper: No humans are needed there to use the language; the language is there for interaction with the environment, it is the environment that reads and interprets the code. There is a code, based on a special language, to make it possible for the system to "express itself" — or, indeed, in very concrete terms, the system is defined in that language. Without domain-oriented languages appropriate constructs and interpretations cannot be defined, and the systems do not become alive in their environments.

8.1.1 Formal and less formal languages

The essence of evolution, or any developmental processes, can be represented as code — or, anyway, different kinds of codes is the way how nature does it. The brain is not a unique medium of decoding languages — perhaps it is the most versatile, but qualitatively the mental system is by no means alone in its aspirations towards capturing the complexity of evolutionary systems. The deep structures of the language [18] are different in different domains. What do we know about such representations?

In addition to the natural languages, there exists a wealth of formal languages being used today. It is instructive to study the special case of *programming languages,* as there seems to be evolution among them, too, from clumsy towards more natural ones. The larger the programs have become, the more structured the formalisms have become, being better capable of representing different substructures in the observed world. In modern programming languages, the computation has been "packaged" tight, the control of computation being distributed; the programming languages have evolved through procedural languages to the today's object-oriented ones. The development of programming languages seems to lead towards representations that more and more match the mental structures: In modern languages, there are "classes" that stand for categories, "objects" being individual instances or examples representing the class. The implementations of the attributes, or the "methods", however, are rather different. To make the programming languages more useful, different kinds of sequential control structures are employed; parallelity or fuzziness in data processing, however, is not heavily addressed, reflecting the dominance of centralized thinking. Of course, there are also the pragmatic reasons for the shortcomings, the processors in today's computers operating one instruction at a time in an all-or-nothing fashion.

The programming languages, as well as the other code systems, are used to describe the desired functioning of the world. In the general-purpose program-

ming languages, all this functioning needs to be implemented explicitly in the code. In the other extreme, one can think of a code that only determines the structures, the functioning being supplied by the language interpreter: The environment can carry out the functionalizing, or "waking up", of the non-living structures, if it supplies ready-to-use hooks to existing functionalities, or dynamic attractors just waiting to be activated — that is, if the code matches with the semantics supported by the environment. Codes just select among the candidate attractors to put up the system.

By definition, however, formal languages are formal or syntactic, missing semantic content. Just as in formal logic, the syntax is separated from the details of the domain field to keep the structures general. In the other extreme, there are the spoken natural languages that are overfull of semantics, being loaded with non-formalizable nuances. A speech act consists of not only the actual utterance but there are, for example, the facial expressions and gestures, and there can be the spices of humor and irony accompanied. The (collective) mental system has developed natural language for communication face to face, not for losslessly storing and transmitting information in text form. Are there other kinds of "somewhat natural" languages that would be appropriate for representing the coordinated developmental processes in other phenospheres? Such a cybernetic language should be some kind of a compromise between formal and natural, capturing a narrow domain with restricted semantics, offering a window to the attractors of dynamic low-level processes in the environment.

To understand the challenges being faced, another aspect about truly natural languages needs to be pointed out. As studied in chapter 3, the key observation there was that everything is implemented locally by uncoordinated actors. This fact applies here, too: The exact-looking representations are not so exact, they are just the emergent nominal patterns. The same applies to codes — meaning that one needs to master the regions within and between the codes. In principle, all diversions from the exact code are errors, but, in practice, there are "degrees of impossibility". As there is normally just one way to interpret the codes, new innovations cannot be studied: Adding noise just breaks the structures, making codes completely undecipherable. One needs (and nature needs) a possibility for "domain-oriented noise", where alternative routes of evolution can be taken in reasonable directions to escape the current stasis. It has been observed that random mutations practically never result in enhancement of the genetic code — it is like adding typing errors in a book: If the text can still be read, there are no meaningful changes in the contents, there are no new memetic structures. Even though the code is represented as a linear list, its structure is hierarchic, and there is huge difference in the effects depending on whether the mutations are in the "leaves" of the hierarchy, or in the "root".

To have more understanding of languages that are not so fixed to the distinct symbols, and to make the topology among the constructs visible and analyzable, wider perspectives are necessary. One needs to implements also "metainformation" for reading the information in the actual code. Only looking at the language is not enough as this metainformation concerning the code does not exist in the code itself — it is in the interpreter mechanism, or in the environment.

To determine an appropriate interpretation of the code, the semantics of the environment has to accompany the language. Actually the key point is not

the language but this interpreter, or compact representation of the application domain. How to determine the attractors that are relevant in the domain to be employed by the systems? If this can be done explicitly, computers can be employed, and the "seminatural" language changes to a programming language, facilitating analyses of complex system *in silico.* The computing power alone does not help if the computer does not do the correct things, but if the environment is implemented appropriately, computationally keeping the dynamics running, supplying the domain-specific dynamic attractors, *simulation* of the development processes can be carried out.

Wolfram [91] says that tomorrow's science is based on simulation. Even though his basic hypotheses is not true (cellular automata are not the only available model family), perhaps he is not totally wrong. Simulation is the way to escape the formal rules and formulas in the theories — but it is not simulation of the system itself, it is simulation of the environment that is needed. Traditional theories are still needed — first to implement the domain-area semantics, and after simulations for compiling the results. Simulation is the method of creating data, fresh processes supplying alternatives or local solutions, the final cybernetic model then being composed over the map of alternatives in the environment.

8.1.2 Simulators of evolution

Evolution is more that mere adaptation in an environment; it involves structural change taking place after the initiation of the system. If this kind of evolution takes place within a single individual it can be called development — in any case, modeling such processes is crucial when trying to capture the essence of biological systems. Unfortunately, there is no strong topology for systems where structures evolve. When seen in the mathematical perspective, the changing structures can be characterized as the processes being highly nonlinear; there are typically no closed-form mathematical solutions to them. A robust and generic approach "homogenizing" the details of nonlinearities is iterative: The models can always be simulated.

Just as the nonlinear representations, codes, too, can have very different outlooks — for example, there are many natural languages that still span the same mental view. There do not necessarily exist any one-to-one mappings between the surface forms of languages — their correspondences can only be evaluated in terms of semantics, through the deep structures, or the activated attractors in the appropriate environment. Only in this sense uniformity among representations can be reached: Codes are identical if they result in the same system of internal dynamics. General analysis of codes and their correspondences can only be carried out through the simulation of the environment — perhaps this is the way towards a general theory of "natural linguistics". To make the syntax and the semantics go together, the point of view needs to be changed. It is not the code that is simulated; it is the environment that runs, and when codes are put in, they are interpreted in that process, so that the operational systems emerge from the simulation. Simulator of the environment remains intact, the systems changing; the special challenge is caused by the fact that changing of the systems can also change the environment possibly introducing new attractors and

modifying the old ones. Of course, the codes can be put in the potion only after the connection between the language constructs and the existing dynamic attractors are defined.

As is evident, the implementation of the simulators in different environments are different, but there are some general ideas. All complex environments share common characteristics: One of the challenges — nonlinearity — was discussed in chapter 6; The other challenge — time-varying nature — was studied in chapter 7, and also in this chapter. To support nonlinearity, one needs to support a population of individual processes, where each of such submodels takes care of its personal local attractor, or local minimum of the cost criterion; to support varying in time, the simulator has to decode the instructions changing the system structures in a coordinated way. To summarize, the simulator has to simultaneously (fractally!) host different types of processes:

- To tackle with the nonlinearity of the world, there have to be *parallel* processes.
- To tackle with the time-variability of the world, there have to be *sequential* processes.

The contribution of the cybernetic thinking here is to offer intuitions. For example, how the population of competing submodels can be maintained and kept stable — of course, this control is localized, coordination being implemented implicitly through the environmental feedbacks, and the neocybernetic frameworks offers an off-the-shelf framework. The other cybernetic contribution is the emphasis on balance pursuit: At each level, dynamic equilibria are searched for — it is these balances that are the dynamic attractors, or the domain-area "concepts". For example, in the case below, the domain-area semantics is implemented based on the neocybernetic model. One can concentrate on such important issues rather than details in simulations.

In short, the role of the simulation machinery is to host the proto-systems in an appropriate environment, shuffle this container, and see the spectrum of outcomes — just as one would do with *in vivo* experiments. It is the computer that now supplies for the mindless signal carriers that operate on the structural building blocks to construct the final systems. The population of surviving systems characterize the local solutions, representing the distribution to be analyzed. The process of "running the programs" in such an environment is a form of *Markov Chain Monte Carlo* simulation. As compared to today's approaches for doing this kind of first-principles simulation, there are only minor differences: For example, the models based on the basic theories of quantum theory are too involved to deliver enough relevant information as the simulations are very computationally demanding; now one does not need such detailed calculations, as it is not only shuffling and cumulating of numeric errors. When one concentrates on the relevant attractors only, one has convergent rather than divergent models, and errors due to simplifications fade away. But what is this "emergent-level quantum physics"?

All this sounds magnificent — but can such mechanization of semantics ever be carried out in any interesting domain? In what follows, the basics of physical chemistry are studied. Study what kind of consequences it has if a molecule is

regarded as a (truly) cybernetic population of charge fields; try to constitute an absolute grounding of semantics in such domain, define appropriate data structures that seamlessly reflect the underlying realm, and define the rules for combining the data structures. The pursuit for semantics must be extended to the very kernel — indeed, when studying physico-chemical systems, it is *quantum theory* that is addressed. From the point of view of understanding biological systems, this problem setting is very relevant: For example, if the code is the sequence of amino acids, its interpretation is the three-dimensional, folded and functioning protein molecule — and this is the nature's way of implementing structures that are defined in the genes. In this environment, the translation of a sequential code into a high-dimensional structure is implemented in very concrete terms. This has to be seen as a (very) preliminary sketch of what the actual implementation of a "proteomic simulator" could look like.

8.2 Emergence in a physical system

Erwin Schrödinger, one of the pioneers in modern physics expected and hoped to find new physics through a study of life [69]. He saw a deep connection between the high-level systems and the low-level principles. Also since him, similar more or less well-grounded intuitions have been exploited: For example, the mysteries of cybernetic processes — not only life itself, but cognition, etc. — have been reduced into the twilight of elementary particles, hoping that "free will" emerges from the unpredictability of quanta. Perhaps there exist more concrete contributions, too — fundamental physics is after all the realm where the living chemical systems reside, and it is here where the system semantics, and "grammar" of the codes, has to be based on.

8.2.1 Cybernetic view of electrons

There is no central control among individual electrons, but the electron systems — atoms, molecules — still seem to be stable and organized. Either there is some yet unknown mechanism that is capable of maintaining the stability and the structures — or, it is the neocybernetic model that applies. The latter assumption is now applied, and the consequences are studied. The starting point (set of electrons) and the goal (cybernetic model) are given, and the steps in between need to be motivated; of course, this is a risky way to proceed, as everything is interpreted in a predetermined way. Results are not conclusive, the goal is just to present an idea and an approach that is offered by the adopted neocybernetic framework as there can be new useful intuitions and interpretations available. Perhaps these studies can be motivated using the words of Max Born:

> All great discoveries in experimental physics have been made due to the intuition of men who made free use of models which for them were not products of the imagination but representations of real things.

So, assume that the nuclei are fixed (according to the Born-Oppenheimer approximation), drop the electrons in the system to freely search their places, and see what happens.

When studying the elementary particles, traditional thinking has to be turned upside down: For example, it seems that in that scale the discrete becomes continuous, and the continuous becomes discrete. Distinct electrons have to be seen as delocalized, continuous charge distributions; however, their interactions have to be seen not as continuous but discrete, being based on stochastic photons being transmitted by the interacting charge fields. This view needs to be functionalized.

Assume that there are two (non-measurable) charge fields i and j, variables x_i and u_j representing their momentary intensities. These fields are manifested through the photons emitted by them; the probability for a photon to be transmitted is proportional to the field intensity. For two fields to interact, the photons need to meet — assuming that the photon transmission processes are independent, this interaction probability is proportional to the product of the intensities, or $x_i u_j$. However, such momentary elementary interactions cannot be detected; the macroscopic phenomena are emergent and become analyzable only through statistical considerations. It is *electric potential* that is such an emergent phenomenon, assumedly being a longer-term average of interactions over time.

To estimate the energy that is stored in the potential fields, one can calculate $p_{ij}\, x_i u_j$, where p_{ij} is the overall probability of the two fields to overlap. Because the field has a dual interpretation, also representing the probability for a charge to be located there, one can estimate the probability of coexistence as $p_{ij} = \mathrm{E}\{\bar{x}_i \bar{x}_j\}$ when the two charge fields are assumed independent. Energy is a scalar quantity; when there are various overlapping charges, their total potential can be expressed as a sum $\sum_{i,j} p_{ij} x_i u_j$, or when expressed in the matrix form, $x^T \mathrm{E}\{\bar{x}_i \bar{x}_j\} u$. However, there are different kinds of charge fields, attractive and repulsive. Assuming that the vector x contains the negative fields, representing the electrons, and u represents the positive charges, one can write for the total energy

$$J = \frac{1}{2} x^T \mathrm{E}\left\{\bar{x}\bar{x}^T\right\} x - x^T \mathrm{E}\left\{\bar{x}\bar{u}^T\right\} u. \tag{8.1}$$

In the former term there are the self-referential structures (there is potential energy stored also when a single charge field is put together), and its outlook can be motivated as in 3.1.2. The key point here is that when appropriate interpretations are employed, it is the neocybernetic cost criterion that is found, meaning that the solutions for electron configurations also implement the same neocybernetic structures. If the assumptions hold, there is self-regulation and self-organization among the electrons, emerging through local attempts to reach potential minimum. Not all electrons can go to the lowest energy levels, and "electronic diversity" emerges automatically. Surprisingly, because of their delocalization, "overall presence" and mutual repulsion, the electron fields implement explicit feedback, following the model of "smart cybernetic agents" (see 5.2.2).

It is interesting to note that there are no kinetic energies involved in the energy criterion, and no velocities or accelerations are involved. As seen from the system perspective, the charges are just static "clouds". This means that some theoretical problems are now avoided: As there are no accelerating charges, there are no electrodynamic issues to be explained as no energy needs to be emitted, and the system can be in equilibrium. In contrast, such electrodynamic inconsistencies plagued the traditional atom models where it was assumed that the electrons revolved around the nucleus, experiencing constant centripetal acceleration, so that radiation of energy should take place.

Whereas the electrons are delocalized, the heavier nuclei can be assumed to be better localized. The key observation here is that the analysis of the continuous space — modeling of the charge distribution of electrons — changes into an analysis of a discrete, finite set of variables, or the nuclei. The idea of "mirror images" is essentially employed here — rather than studying the system itself, the electrons, its environment is analyzed: In this special case it is the environment that happens to be simpler to operate on. Assuming that the interactions among the distinct nuclei can be represented in terms of a covariance matrix $\mathrm{E}\{\bar{u}\bar{u}^T\}$, the charge distributions of electrons are revealed by its eigenstructure. When the eigenvectors are denoted as ϕ_i, one can rename the constructs: The "orbits" of electrons, determined by the eigenvectors, are discretized (molecular) orbitals. Following the Max Born's formalism, the "squares" of orbitals represent "real" probabilities, or actual charges — and, indeed, the basic quantum theoretical assumptions seem to hold: For example, the integrals, when changed to summations, again equal

$$\sum_{j=1}^{n} |\phi_{ij}|^2 = \phi_i^T \phi_i = 1. \tag{8.2}$$

Because of the properties of eigenvectors, the discrete orbitals are mutually orthogonal. Traditionally, it is assumed that there is just room for a unique electron in one orbit (or, indeed, for a pair of electrons with opposite spins). However, now there can be many electrons in the same orbital, and there is no need to employ external constraints about the structures, like assumptions of spins, etc. The charge field can be expressed as $\psi_i = \sqrt{\lambda_i}\,\phi_i$, so that the overall charge becomes $\psi_i^T \psi_i = \lambda_i$. The "variance" λ_i is the emergent measurable total charge in that field. This means that there are some conditions for the charge fields to match with the assumption of existence of distinct charge packets:

1. The eigenvalue λ_i has to be an integer times the elementary charge, this integer representing the number of electrons in that orbital.

2. The sum of all these integers has to equal the number of valence electrons, sum of all free electrons in the system.

These constraints give tools to determine the balance configuration among the nuclei (see later). As studied in 3.2.3, the energy drop in the orbital is related to λ_i^2.

How to quantize the continuous fields, and how to characterize the effects in the form $\mathrm{E}\{\bar{u}\bar{u}^T\}$, and how to determine the parameters? And how is this all

related to established quantum theory? In short, how are the above discussions related to real physical systems? These issues are studied next.

8.2.2 Molecular orbitals

Atoms are already enough well understood — at least what comes to the hydrogen atom (!). The contemporary theory of atom orbitals can explain their properties to sufficient degree. However, it seems that one needs new approaches to understand the emergent level — or the level of molecules. Molecular orbitals are interesting because the chemical properties of compounds are determined by their charge distribution — essentially these orbitals reveal how the molecule is seen by the outside world.

The molecules have been a challenge for modern physics for a long time, and different kinds of frameworks have been proposed to tackle with them: First, there are the *valence bond theories,* where the individual atoms with their orbitals are seen as a construction kit for building up the molecules, molecule orbitals being just combinations of atom orbitals; later, different kinds of more ambitious *molecule orbital theories* have been proposed to explain the emergent properties of molecules. In both cases it is still the ideas of atom orbitals that have been extended to the molecules. Unfortunately it seems that very often some extra tricks are needed: for example, to explain the four identical bonds that carbon can have, peculiar "hybridizations" need to be employed; and still there are problems, a notorious example being *benzene* (and other aromatic compounds) where the "bottom up" combinations of atom orbitals simply seem to fail. And, unluckily, it is exactly carbon and its properties that one has to tackle with when trying to explain living systems and their building blocks.

When thinking of alternative approaches, it is encouraging that molecules have been studied applying discretized eigenvalues and eigenvectors before: For example, Erich Hückel proposed an approach that is known as *Huckel's method,* also reducing the analysis of energy levels in molecules into essentially an eigenvalue problem [8]. However, this method is still based on combinations of atom orbitals, and being based on crude simplifications, it is regarded as an approximation. It is also quite commonplace that linear additivity of orbitals is assumed on the molecular level — normally it is atomic orbitals that are added together, now it is molecular orbitals directly. Indeed, basic physics *is* linear; the problems are normally caused by the huge dimensionality of the problems.

Now it is assumed that all of the molecular orbitals extend over the whole molecule, and it is assumed that (8.1) characterizes the electrons; the challenge is to combine this with current theories and models. It is the *time-independent Schrödinger equation* that offers a solid basis for all quantum-level analyses [10]. It can be assumed to always hold, and it applies also to molecules (h is the Planck's constant, and m_e is the mass of an electron):

$$-\frac{h^2}{8\pi^2 m_\mathrm{e}}\frac{d^2}{dx^2}\psi(x) + V(x)\psi(x) = E\psi(x). \tag{8.3}$$

Here, $V(x)$ is the potential energy, and E is the *energy eigenvalue* corresponding to the *eigenfunction* $\psi(x)$ characterizing the orbital. As $\psi(x)$ is continuous,

Schrödinger equation defines an infinite-dimensional problem, and as x is the spatial coordinate, in higher dimensions this becomes a partial differential equation. Normally this expression is far too complex to be solved explicitly, and different kinds of simplifications are needed. Traditional methods are based on reductionistically studying the complex system one part at a time, resulting in approaches based on the atom orbitals. Now, start from the top: As studied in the previous section, assume it is simply a non-controlled play among identical electrons that is taking place in a molecule. It is all "free" electrons that are on the outermost shell that are available for contributing in the orbitals, that is, for a carbon atom the number of valence electrons is increased by the number $v_{\mathrm{C}} = 4$, for hydrogen $v_{\mathrm{H}} = 1$, and for oxygen $v_{\mathrm{O}} = 6$ (!). What kind of simplifications to (8.3) are motivated?

The time-independent *discrete* Schrödinger equation that is effectively being studied is defined now as

$$-V_0\phi_i + V\phi_i = E_i\phi_i, \tag{8.4}$$

where ϕ_i are now vectors, $1 \leq i \leq n$, dimensions equaling the number of atoms in the molecule n; because of the structure of the expression, these are the eigenvectors of the matrix $V - V_0$ corresponding to the eigenvalues E_i. Comparing to the discussions in the previous section, there holds $E_i = \lambda_i^2$, the eigenvectors being the same. Rather than analyzing the infinite dimensional distribution of electrons study the finite-dimensional distribution of nuclei; one only needs to determine the $x \times n$ elements of the potential matrix $V - V_0$ to be able to calculate the orbitals (or the negative charge fields around the positive nuclei).

To determine the matrix of potential energies among the nuclei, the challenge is to determine the terms corresponding to the first term in (8.3). The diagonal entries of $V - V_0$ are easy: Because the "local potential" is assumedly not essentially affected by the other nuclei, the atoms can be thought to be driven completely apart, so that the non-diagonal entries vanish; the diagonal entries then represent free separate atoms, so that the electron count must equal the number of available valence electrons, that is, the i'th diagonal entry is proportional to v_i^2, where v_i presents the number of valence electrons in that atom. For non-diagonal entries, the sensitivity to changes to distant nuclei becomes small, so that the term with the second derivative practically vanishes, and the corresponding entry in the potential energy matrix is according to basic electrostatics approximately proportional to $\frac{v_i v_j}{|r_{ij}|}$ without normalization. Here, $|r_{ij}|$ stands for the distance between the nuclei i and j. When the preliminary potential matrix has been constructed, elements of the matrix $V - V_0$ have to be justified so that the eigenvalues of the matrix become squares of integers, and the sum of those integers equals the total number of valence electrons.

So, given the physical outlook of the molecule in equilibrium, one simply carries out principal component analysis for the matrix $V - V_0$, finding the set of "discrete orbitals", or orbital vectors ψ_i and the corresponding eigenvalues E_i and electron counts λ_i. The elements of the vectors ψ_i reveal around which nuclei the orbital mostly resides; the overlap probability p_{ij} is spatial rather than temporal. For illustration, study the benzene molecule: Benzene is the prototype of aromatic compounds, consisting of six carbon atoms and six hydrogen atoms in a carbon-ring. Altogether there are 30 valence electrons (6 times 4

for carbon, and 6 times 1 for hydrogen). The results are shown in Fig. 8.1 — compare this to Fig. 8.2: It seems that the three first orbitals have essentially the same outlook in both cases. Now there are altogether 7 electrons on the lowest energy level! All orbitals extend over the whole molecule; the hydrogen orbitals are also delocalized — such delocalization applies to all molecules, not only benzene. Note that the orbitals having the same energy levels are not unique, but any orthogonal linear combinations of them can be selected; such behavior is typical to symmetric molecules. The "bonding energy" is the drop in total energy, or the difference between the energies in the molecule as compared to the free atoms; possible values of this energy are discretized, now it (without scaling) is $1 \cdot 7^2 + 2 \cdot 4^2 + 3 \cdot 3^2 + 6 \cdot 1^2 - (6 \cdot 4^2 + 6 \cdot 1^2) = 12$.

The presented approach is general and robust: For example, the unsaturated double and triple bonds as well as aromatic structures are automatically taken care of as the emerging orbitals only depend on the balance distances between nuclei: If the nuclei remain nearer to each other than what is normally the case, there also must exist more electrons around them. Spin considerations are not needed now, as there is no need for external structures (orbitals of "two-only capacity") to keep the system stable and organized. However, no exhaustive testing has been carried out for evaluating the fit with reality. When different molecules were experimented with, the results were not fully satisfactory. Anyhow, the objective here is to illustrate the new horizons there can be available when employing non-centralized model structures.

8.2.3 Characterizing molecules

The time-independent Schrödinger equation (8.3) is not the whole story. As explained, for example, in [10], the complete wave equation consists of *two* parts, the other being time-dependent (and location-independent), these two parts being connected through the energy eigenvalue E. The complete solution has the form

$$\psi(x,t) = \psi(x)\mathrm{e}^{\sqrt{-1}\,2\pi Et/h}. \tag{8.5}$$

Because of the imaginary exponent, the time-independent part oscillates at a frequency that is determined by the energy level of the orbital. Now in the case of discretized orbitals, one can write for the orbital vectors

$$\psi_i(t) = \psi_i \sin\frac{2\pi E_i t}{h}. \tag{8.6}$$

Each energy level also oscillates with unique frequency. This means that the orbitals cannot interact: Because the potentials are assumed to be related to integrals (averages) over the charge fields, there is zero interaction if the fields consist of sinusoids of different frequencies. On the other hand, if the frequencies are equal, the time-dependent part does not affect the results. This way, it seems that each energy level defines an independent interaction mode, and these modes together characterize the molecule — and also each of the individual atoms within the molecule. Thus, define the matrix Ψ where each of the columns represents one of the atoms, from 1 to n, the column elements denoting the

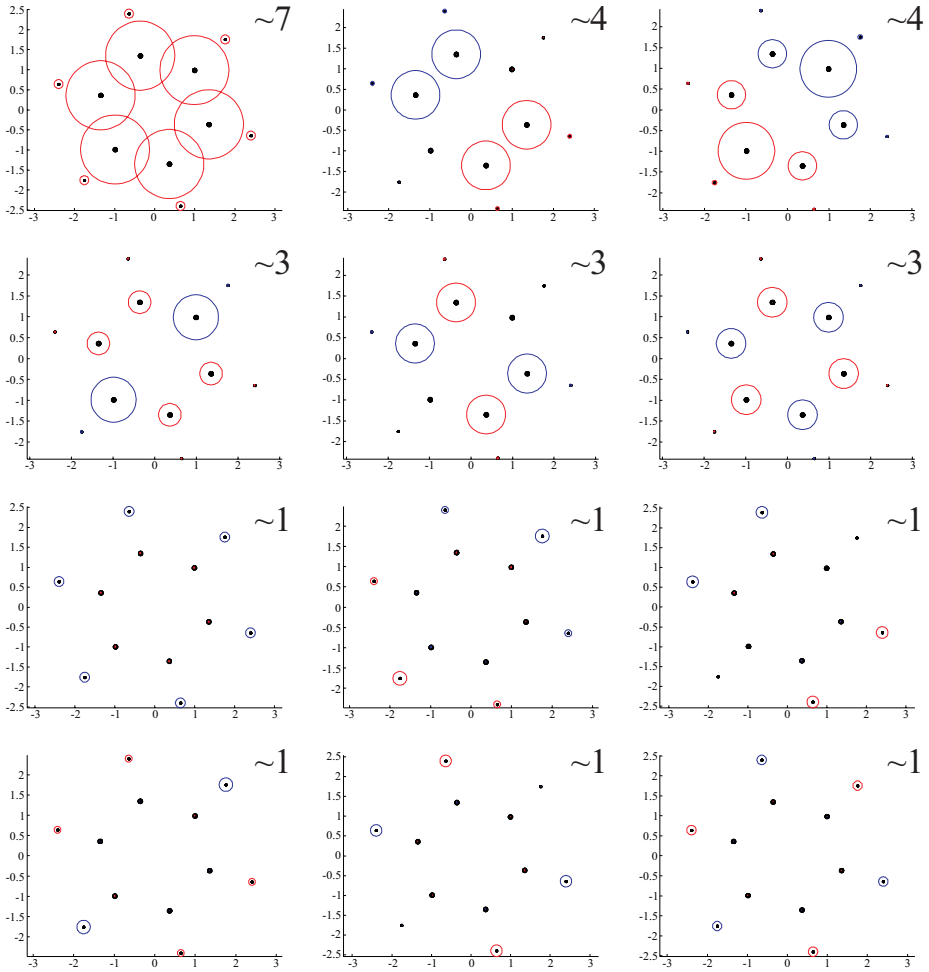

Figure 8.1: "Cybernetic orbitals" ψ_i in the benzene molecule (see text). The larger dots denote carbon nuclei and the smaller ones hydrogen nuclei, distances shown in Ångströms (1 Å$= 10^{-10}$ m). The orbitals, shown as circles about the nuclei, have been scaled by the corresponding λ_i to visualize their relevances. The circle colors (red or blue) illustrate the correlation structures of electron occurrences among the nuclei (the color differences are to be compared only within a single orbital at a time)

contribution of each of the orbitals, from 1 to n, to the total field in that atom:

$$\Psi(t) = \begin{pmatrix} \psi_1^T(t) \\ \vdots \\ \psi_n^T(t) \end{pmatrix} = \begin{pmatrix} \Psi_1(t) \,\big|\, \cdots \,\big|\, \Psi_n(t) \end{pmatrix}. \tag{8.7}$$

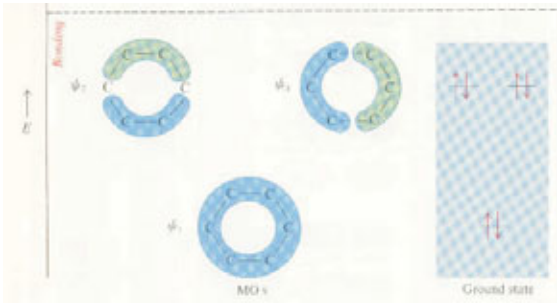

Figure 8.2: Benzene orbitals as proposed in literature (see [56]). Compare to Fig. 8.1

So, rather than characterizing an orbital, Ψ_j represents the properties of an atom j within the molecule. The key point here is that the elements in these vectors reveal the mutual forces between the atoms: If the other of the atoms always has excess field when the other has deficit — the orbitals being "red" and "blue", respectively, as in Fig. 8.1 — the atoms have opposite average occupation by electrons, and the positive attracts the negative. On the other hand, in the inverse case there is repulsion among similar charges. These forces determine whether the atoms can get enough near each other to react; indeed, this force is closely related to the concept of *activation energy* that is needed to overcome the repulsion among atoms. In the adopted framework, this activation energy can be formulated as

$$\Psi_i \Lambda^2 \Psi_j, \tag{8.8}$$

where the total energy is received by weighting the attractive and repulsive components by the appropriate orbital energies (Λ being a diagonal matrix containing the electron counts on the orbitals).

There are only some 100 different atom types, but it seems that there are no bounds for molecule types and behaviors. The above discussion gives guidelines to understanding how this can be explained and how this understanding can be functionalized. A sequential molecule is like a "string" whose vibrations are modulated by the additional "masses" that are attached along it. Because of the linear structure of the protein chain, it is clear that the interaction covariance matrix is diagonally dominant. it is an interesting question how different amino acids in different locations are reflected in the frequency structure of the final molecule. In the proposed framework, it is also possible that the oscillating charge fields interact with the nuclei; this gives rise to extra complexity in the characterization of molecules as the mechanical vibrations can also affect the chemical properties of the compounds — for example, different isotopes having different masses can have differing behaviors.

Because of the universal quantization of the energy levels, the repulsions and attractions are, in principle, comparable among different molecules — but there is the question of synchronization of the oscillating fields. What is more, the proposed characterization of the molecules is not exhaustive: For example, "handedness" is beyond this kind of analysis, and optical isomers have identical representation even though they have differing chemical properties. Because of these shortcomings, the representation is not a general-purpose one — but within a single molecule there may be new tools available.

8.2.4 Folding of proteins and splicing of RNA

All genetic programs are manifested as proteins being products of a complex process of DNA transcription and RNA translation. The proteins are used either as building blocks themselves or as enzymes catalyzing further reactions. Assumedly the proteins are as diverse as the genes themselves — how can proteins have so different properties, being composed of the very basic atoms? It has been proposed that it is the physical outlook, or *folding* of the proteins that is largely responsible for the properties.

The DNA, and after that RNA, only determines the linear sequence of amino acids, the formation of the three-dimensional structures taking place afterwards. But, of course, it is the sequence of amino acids, as being interpreted by the environment, that determines also the final outlook of the molecule: There is affinity among far-apart atoms in the molecule as determined by the structure. Because of its importance, this folding process has been studied extensively, mostly applying computational approaches. But no matter how heavy supercomputation is applied the long-range interactions cannot be revealed or exploited when these long-range effects are abstracted away to begin with in the standard molecular models.

This protein folding seems to be an example of a wider class of phenomena: Intra-molecular affinities have to be understood to master many different kinds of processes. For example, study *RNA splicing.*

In eukaryotic cells, the gene sequences in DNA contain non-coding fractions, or *introns,* in addition to the coding ones, or *exons.* During the processing of pre-mRNA into the actual messenger-RNA, the non-coding portions are excluded in the process of splicing where the exons are connected to form a seamless sequence. What makes this process specially interesting is that the splicing process does not always produce identical messenger-RNA's, but there are alternative ways — sequences can be interpreted as introns or as exons in different environments. It has been observed, for example, that a single mouse gene can in theory produce more different kinds of proteins than what is the size of the whole genome. No doubt understanding the splicing mechanisms will become very important, as nature has found this mechanism because it offers a flexible way to alter the gene expression results without having to go through the highly inefficient route of evolving the whole genome. However, today these mechanisms are still less understood than what protein folding is — and it seems that the real essence of RNA splicing cannot even be explicated yet. Because there is no central control, it is evident that the locations that are to be reconnected need to attract each other. Again, it would be invaluable to master the attractions and repulsions among the atoms in the molecule.

When analyzing reactions, it is often the energy levels before and after a reaction that are studied. However, when studying reaction probabilities, analysis of the final energy levels is not enough: The key point is to see whether the reaction can ever take place. It needs to be recognized that carbon is very reactive, and it forms a bond whenever two atoms are enough near each other — the total energy seems to go down as there are more atoms in the molecule. The most important thing is the activation energy, or the energy that is needed to bring the atoms near each other. Indeed, if the activation energy is low — or, specially, if the

reacting components attract each other — the reactions probably take place.

Understanding the underlying principles of attraction and repulsion among atoms gives tools to understand not only the folding processes, but also catalytic (or enzymatic) reactions; and it is enzymes that are responsible of most of the biological reactions. How is it possible that there seems to exist an infinite number of catalysts even though the number of alternative form for "keys" and "locks" seems to be so limited? The new view explains that there can exist an infinite number of energy levels, and thus there can exist an infinite number of attraction patterns. When applying the "holistic" view of molecules as electron systems, orbitals extend over the whole molecule. All atoms count, and it becomes understandable how atom groups far apart can alter the chemical properties of the whole molecule. Specially, in this framework it can be explained how the coordinated-looking very long reaction chains of transcription and translation can exist. The sequential reading of codes can be locally controlled when there is "chained catalysis": Only when the previous piece of code is processed, the next step is catalyzed by the previous reaction result.

Now there are enough tools to implement an "emergent-level simulation" for modeling the protein folding: Rather than doing the extensive quantum theoretical ab-initio simulations, represent the chemicals in terms of their emergent affinity structures Ψ, and shuffle the potion, producing a distribution of chemicals and their structures in equilibrium. That is, put in the "codes" — linear amino acid sequences — and let the environment interpret and process them into a "system". Note that the representations Ψ change as the physical appearance changes, the atoms traversing in the force fields, and these also need to be adapted; what is more the physical outlook also affects the possibilities of the atoms to approach each other, and this also needs to be taken care of in the simulations. Whenever orbitals come near enough each other, they automatically merge, thus forming a bond[1]. The simulation is also simulation of nuclei movements: Given a nucleus combination, the orbitals can be determined, and after the charges are found around the nuclei, and when the physical constraints are taken into account, the total forces affecting the nuclei can be calculated. When the nuclei are moved slightly in the directions of their attraction, the orbitals need to be calculated again, and the whole loop is repeated, until a balance is reached so that no residue attraction remains.

The interpretation of one-dimensional code into a high-dimensional operational representation (see 7.3) is very literal in the case of protein folding — the chain of amino acids becomes a many-folded structure. The end result is the folded protein; and it has been claimed that it is this structure that mainly determines the functions of the molecule. Is there anything more to be said here?

After all, one is interested in the functions, the tissues or catalytic effects of enzymes, not in structures themselves. Is there any possibility of analyzing how the function of a folded protein emerges from its structure in general terms? It should be clear that it cannot be the physical structure alone that would offer the complete answer: Molecules cannot perceive physical dimensions; neither is it some three-dimensional jigsaw puzzle in the tissues. It is tempting to

[1] As soon as there is *any* interaction among atoms, the potential matrix has non-zero non-diagonal elements, and a mixture of orbitals is found; indeed, this model predicts the *tunneling* of electrons also among far-apart atoms

hypothesize that it is not only the physical structure, but again the cybernetic structure that plays a role here: As the energy levels of the molecule specify its oscillatory structure in the quantum level, perhaps neighboring molecules find synchronization and become aligned. There emerges resonance, and the molecule-level structure is repeated and magnified, being manifested as a special type of homogeneous tissue, or — why not — as a crystal lattice in the inorganic case. The resonances define a Pythagorean "Harmony of the Phenospheres", cybernetic balance of vibrations ... is this the key to the next emergent level where the molecular components are combined into a next-level structures?

8.3 Towards "cosmic cybernetics"?

The previous example raises the question whether the physical systems also are cybernetic. If the cybernetic principles are rooted so deep in the structure of matter, perhaps there are reflections also on the wider scale? After all, the systems in the large also are locally controlled only, and there is balance and non-trivial structures in the universe. The advocates of "theories of everything" claim that after combining the different kinds of basic forces, everything can be explained; however, problems of cognition, for example, do not belong to this "everything". The real GUT theory has to explain emergence; does the cybernetic thinking have something to offer here?

8.3.1 Formation of stellar structures

For example, when looking at complex physical systems, like galaxies, etc., where there also are individual "agents", the stars, one can see that together the bodies constitute seemingly self-organized, long-living constellations. As compared to the neocybernetic studies, and when studying the possibilities of extending such considerations to stellar systems, there are many problems. First, because of the geometries and the outlook of the gravity law, the formulas are highly nonlinear; second, as the forces are determined by the mutual orientations among the stars, and as these orientations constantly change, it seems that there cannot cumulate any structure among the stellar bodies. The main problem, however, is that there seems not to exist a repulsive force: Gravity attracts bodies, and if there exist no balancing mechanisms, no cybernetic self-regulation or self-organization can emerge.

Still, the solar systems and galaxies *are* rather stable. Originally there were only clouds of stellar dust, but, based on local interactions only, different kinds of structures have emerged. There evidently exist balancing effects — and, indeed, as the attraction increases speed, the experienced centrifugal forces try to drive the bodies apart.

How can the cybernetic ideas of elasticity, optimality, and adaptation be motivated when it is mindless stellar bodies that implement all functionalities? However counterintuitive it may sound, all mechanical systems in the global scale for some reason implement optimization. In *Lagrangian mechanics,* and later in *Hamiltonian mechanics,* it is observed that the Newtonian laws of motion can be reformulated as optimization problems: Along the motion trajectory, the

time integral of the quantity $L = W_{\text{kin}}W_{\text{pot}}$ reaches its minimum value, where W_{kin} is the kinetic energy of the system, and W_{pot} is the potential energy. This starting point has been applied a long time for deriving dynamic models for mechanical systems, but it is not applicable for analysis from the point of view of neocybernetics.

Since the Newtonian times, the emphasis in dynamic modeling has been on accelerations induced by forces. However, in a galaxy in balance, such accelerations — even though everything is after all based on them — seem to be rather irrelevant. As in the case of electrons before, now try to eliminate the accelerations. Acceleration means change in velocity; assume that the velocities remain constant in the large, and the neocybernetic steady state prevails. Study the possibilities of finding an "emergent model" for a galaxy where the component-level interactions are abstracted away. What would such a model perhaps look like? What are the local-level adaptations that can be justified in such a non-adaptive environment where the laws of mechanics and gravity cannot be adjusted?

What kind of alternative "stationary" models there are for central motion? A characteristic model family is offered by the Navier-Stokes equations of fluid dynamics: Vortex structures in flows are commonplace. Indeed, the vortex model offers intuitively appropriate behavioral patterns also for representing gravitational fields, down until the black hole like singularities. Correspondingly, in an combined electric/magnetic field such rotors also exist. The common denominator characterizing such models are *vector products* between some vector-form quantities. Such mathematical representations can be seen as the *emergent evolutionary goals* that are easier visible in systems with vortices of faster time scales than in slow gravitational systems that still are in their transitory state towards that final balance. In each case, the underlying actors (photons vs. mass units) assumedly only obey tensions determined by their local environments, and the challenge now is to reinterpret the variables to implement the observed behaviors in terms of physically relevant, locally observable quantities.

Of course, this all is very vague, but some bold (yet schematic) conclusions are here possible. The hypothesis here is that the *eventual stationary gravitational vortices can be modeled using vector products between the velocities of bodies and the forces acting on them.* One motivation for this selection is that the product of (scalar) velocity and (scalar) force has the dimension of power — and this quantity sounds like a reasonable emergy measure to be pursued in a mechanical domain. Assuming that the momentary velocity vector of a particle j is denoted as v_j and the total force acting upon it is F_j, the size of the corresponding vector product is $|v_j|\,|F_j|\sin(\alpha_j)$, where α_j is the angle between the vectors, and the average of it is $\mathrm{E}\{|v_j|\,|F_j|\sin(\alpha_j)\}$. When approaching the steady state, the average gets nearer to the momentary value.

But what is the assumed adaptation scheme among mindless bodies, why should they try to do maximization of some quantity and how could they implement that? It needs to be observed that orbits that maximize the cross product criterion (under converged central motion) are *circular orbits* where the velocity and acceleration remain perpendicular. It can be assumed that if such circularity constraint is fulfilled by all stellar bodies, their *orbits minimally interact* and remain intact, whereas non-circular behaviors are more probable to die

out through collisions. This means that circularity is evolutionarily beneficial: Circularity of orbits serves as an emergent fitness criterion — even though the actors never see the chaos around them in such a wider perspective.

If the gravitational system is interpreted as an elastic one (stronger force meaning shorter radius and faster motion when the angular momentum is preserved), the product $|F_j| \sin(\alpha_j)$ can be seen as the external input, and $|v_j|$ can be seen as an internal variable, one can express the local-level aspirations in the neocybernetic framework when one defines the system state as $x_i = |v_j|$ and the inputs as $u_j = |F_j| \sin(\alpha_j)$. When the local maximization takes place, global-level emergence assumedly takes place: The final state after convergence under elasticity assumption minimizes the familiar cost criterion

$$\min_x \left\{ \mathrm{E}\{\frac{1}{2} x^T \mathrm{E}\left\{xx^T\right\} x - x^T \mathrm{E}\left\{xu^T\right\} u\} \right\}. \tag{8.9}$$

Looking closer at (8.9), one can see that the former term represents *kinetic energy,* being essentially based on products of velocities; the diagonal elements represent translational components, whereas the non-diagonal entries are inertial components. Following the neocybernetic spirit, the state vector x can be compressed in the PCA spirit so that the essence of the system is minimally affected. It is here where the added value can be seen: As the state space is compressed, the huge number of individual mass entities becomes represented by a simpler model. This essence of the emergent model is in the lower-dimensional *inertia matrix* that can be determined through observations as $\mathrm{E}\left\{\mathrm{xx^T}\right\}$. It turns out that *the uncoordinated stellar units can be seen as a more or less three-dimensional rigid body.* Still, before the final convergence in infinity, there is still noise in the system: The local whirls among the stellar bodies are manifested in the model as non-vanishing variations in the data. In principle, the model does not only capture the movements of the galaxies and stars, but also stars and planets, and planets and moons in its statistical structure. The larger model can also be decomposed: Single solar systems can be characterized in terms of their local PCA structures. The evolution towards circular orbits around a single mass center still continues — for example, the tidal effects gradually bind the motions of the moons.

Perhaps the Pythagorean "harmony of the celestial spheres" can be defined in terms of principal component modes (stars on the wider scale and planets on the narrower one), modulating dissonances being caused by local rotations (orbiting planets and moons, respectively, etc.) and ellipticity of the orbits.

8.3.2 Everything, and more

The neocybernetic principles can be claimed to span behaviors from the elementary level (orbitals) to the cosmic level (orbits) — but hypotheses can be made even beyond that, towards the most extensive levels of all.

The neocybernetic models pop up in very different cases: Perhaps this is not a coincidence. Many complex systems can be characterized in terms of optimization, models being derived *applying calculus of variations,* the explicit formulas (constraints) being the emergent outcomes of underlying tensions. For example, *Maxwell's electrodynamics* can be formulated in terms of such optimization.

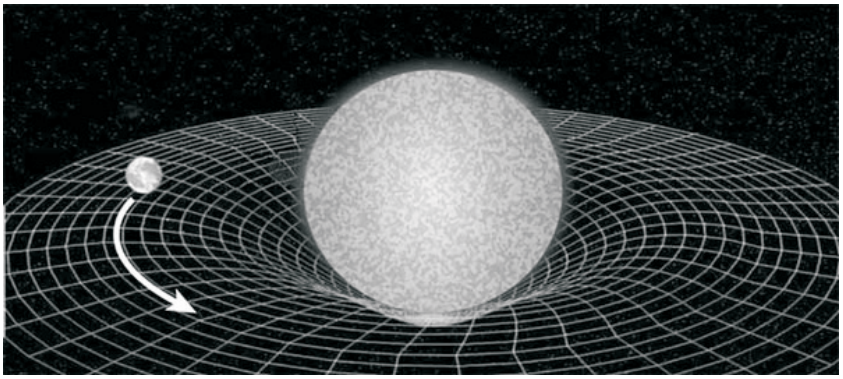

Figure 8.3: Illustration of how the four-dimensional space-time is deformed as a concentration of energy (mass) is found in the universe

What if the idea of distributedness of underlying stochastically behaving actors is explicitly employed, perhaps it is the neocybernetic framework that automatically pops up?

The faith here is that all observable behaviors even in the most universal levels are finally implemented by uncoordinated low-level actions. In addition to the stellar and planetary motions, there are also connections between astrophysics and cybernetics on the more fundamental, less immediate level. It seems that all non-trivial systems that still exist after those millions of years have found their ways of implementing adaptive feedback, otherwise no self-regulation and self-organization could have been reached. And, indeed, it seems that the idea of elasticity is intuitively applicable also in wider scales — for example, the Einsteinian tensors are tools to formalize pressures, or tensions (see Fig. 8.3). There are problems when trying to apply the relativity theory to elementary particle level; in this scale, all variables are quantized and the continuum models collapse. There also exist models based on "cell-structured universes" that are (more or less) compatible with the Einsteinian cosmology — the cybernetic ideas could directly be applied in such models, neighboring cells being interacting subsystems transferring emergy: Compare to Fig. 3.5, where impulses traverse through the "space" as the coupling between "cells" is tight enough.

There are efforts to find the Grand Unifying Theory (GUT) that would combine all basic forces like combine electromagnetics, gravity, weak and strong nuclear forces into the sme framework. Elasticity seems to offer fresh ideas also in the field of basic physics: Beyond the observations, in super string theories, the elementary particles are seen as vibrating springs (or *vibrations* of strings). But regardless of the form of the final theories, it seems that thinking of the universe as an elastic self-balanced shell reacting to external pressures, this "shell" being distributed in matter particles, offers a useful framework for studying matter. The Heisenbergian thinking is to be extended, as it is all interactions (not only measurements) that affect the system, the effective variables being reflections of the emergent balance among the system and the environment. Measurable variables are "interaction channels", each interaction mechanism introducing a spring of its own; individual seemingly static relations of the form $\bar{x}_i = q_i \bar{u}_i$ connect observations through some coupling coefficient q_i. The natural constants

are not predetermined, but they are the visible manifestation of the balance ratio between reaction and action. The modern theories employ some 11 dimensions (some theories necessitating introduction of several dozen dimensions!) where there are some "collapsed dimensions" among them: Now there is no upper limit to the dimensions as they are no actual coordinate axes but they only represent the number of interaction channels into the universe; and it is easy to think of the vanishing degrees of freedom as being only tightly coupled to others through the cybernetic feedback controls. The constants of physics should not be seen as predetermined quantities: There are propositions that the natural constants are gradually changing as the universe gets older. One of such propositions is by Paul Dirac, who claims that cosmology should be based on some dimensionless ratios of constants (known as "large number hypothesis).

If the cybernetic thinking universally applies, one can exploit the understanding concerning such systems: Perhaps universe as a whole is *optimizing* some criterion? This would help to escape some deadlocks one is facing today.

It has been estimated that to have a stable, non-trivial and long-living universe that can maintain life, the natural constants have to be tuned with $1/10^{55}$ accuracy. Such astonishing coincidence has to be explained somehow, and different families of theories have been proposed. First, there are the *anthropic* theories, where it is observed that the world just has to be as it is, otherwise we would not be there to observe it; the other theories are based on the idea of *multiverses,* where it is assumed that there is an infinite number of "proto-universes" in addition to our own where physics is different. However, in each case it seems that physics reduces to metaphysics, where there are never verifiable or falsifiable hypotheses.

If the universe is (neo)cybernetic, each particle maximizes the share of power it receives, resulting in the whole universe becoming structured according to the incoming energy flows. Then there is no need for multiverses, as it is the only the best alternative that really incarnates. It is as it is with simple subsystems: Fermat's principle says that light beams "optimize" selecting the fastest route; it is the group speed that determines the wave propagation, the emerging behavior representing the statistically most relevant alternative. Similarly, the only realized universe is where the optimality of energy transfer is reached.

Again, applying the neocybernetic intuitions concerning adaptive controls (as studied in chapter 5), the reasonably evolving universe must not be too balanced: Perhaps there must exist some level of asymmetry to avoid the final stagnation and subsequent collapse — this collapse perhaps signaling the end of the universe as we know it, giving room for the "next version" to start the cycle from the beginning.

The idea of the evolutionary universe is intriguing. How about the adaptation and evolution mechanisms? The key point is not whether the cybernetic thinking can be applied to modeling non-living physical systems — the most interesting question is whether the universe can be interpreted as being a *living entity* itself. Perhaps this all is not only loose metaphysics: As studied in the next chapter, it seems that there seems to exist a nice connection between the universal physical principles and the neocybernetic systems.