Bio-physics manifesto —for the future of physics and biology—[1]

Y. Oono
Department of Physics and Institute for Genomic Biology,
University of Illinois at Urbana-Champaign
Urbana, Il 61801, USA
yoono@uiuc.edu

Abstract

The Newtonian revolution taught us how to dissect phenomena into contingencies (e.g., initial conditions) and fundamental laws (e.g., equations of motion). Since then, 'fundamental physics' has been pursuing purer and leaner fundamental laws. Consequently, to explain real phenomena a lot of auxiliary conditions become required. Isn't it now the time to start studying 'auxiliary conditions' seriously?

The study of biological systems has a possibility of shedding light on this neglected side of phenomena in physics, because we organisms were constructed by our parents who supplied indispensable auxiliary conditions; we never self-organize. Thus, studying the systems lacking self-organizing capability (such as complex systems) may indicate new directions to physics and biology (bio-physics).

There have been attempts to construct a 'general theoretical framework' of biology, but most of them never seriously looked at the actual biological world. Every serious natural science must start with establishing a phenomenological framework. Therefore, this must be the main part of bio-physics. However, this article is addressed mainly to theoretical physicists and discusses only certain theoretical aspects (with real illustrative examples).

1 Introduction

It is said that this is the century of biology. Many physicists are working on problems apparently related to biology; biophysics is a fashionable branch of physics. I believe physics is a discipline not defined by what it studies, but by how it studies the world. We physicists should not be confined to the conventional interpretation of physics as the study of 'physical' world mostly excluding animated objects and their epiphenomena including the humanities. The so-called complex systems study is supposedly a discipline to take care of this 'unphysical' world. However, have complex systems study and biophysics been really useful to understand our world? This is the question behind this article. Advocating physics imperialism in its ultimate form might be a hidden agenda.

In Section 2, I will discuss what is wrong with the complex systems study practice [1] and then, in Section 3, look at what physics has not been doing. After these preliminary considerations, in Section 4, I point out the important characteristic of genuinely complex systems including organisms. This tells us where to focus our attention to build a phenomenology of biological world (Section 5). Section 6 discusses a necessary condition for complex systems, and Section 7 discusses two qualitative consequences of the phenomenological survey of the world of organisms.

---

[1]Invited talk given in Tokyo, Sept, 2007.

2   What is wrong with complex systems study?

To understand the real difference between simple systems and (genuinely) complex systems let us compare (A) a droplet of saturated salt solution with a floating small salt crystal and (B) a droplet of water containing a single cell of *Escherichia coli*. It is easy to write down the equations of motion for these systems; we have only to write down the Schrödinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = \left\{ -\sum_i \frac{\hbar^2}{2m_i}\frac{\partial^2}{\partial \boldsymbol{r}_i^2} + \sum_{i>j}\frac{q_i q_j}{4\pi\epsilon_0|\boldsymbol{r}_i - \boldsymbol{r}_j|}\right\}\psi. \tag{2.1}$$
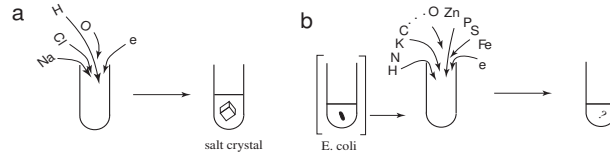

Fig. 1.1 Analogue experiments to distinguish simple and complex systems.

To this end we need only elemental analyses: we need the numbers, mass m_i and the charge q_i of various species of nuclei and the electron charge and mass. Of course, we cannot actually write the equations down, but it is clear that between (A) and (B) exists no fundamental difference thanks to the linearity of the fundamental laws. There is every reason to believe that both (A) and (B) satisfy (2.1). To confirm this we have 'only' to solve the equations. There is no hope of doing this digitally in the foreseeable future, but we can do quantum computation: analogue computation or actual experiments.

Can we obtain (A)? Since (2.1) is a partial differential equation, we need auxiliary conditions: a boundary and an initial condition. In this case, the boundary condition is not so important. Since we can expect that a cell of *E. coli* stays alive at least for a short time in a completely isolated water droplet, we may assume that the boundary conditions are homogeneous Dirichlet conditions. In case (A) we expect that 'almost all' the initial conditions with appropriate energy give a salt water droplet containing a small salt crystal. This is exactly the reason why equilibrium statistical mechanics works without any particular specification of the initial condition.

The case (B) is a futuristic version of Pasteur's famous experiment refuting the spontaneous emergence of life; we cannot do well with a generic initial condition. Since we cannot revive a mechanically destroyed *E. coli* cell, it is clear that structural (geometric) information is crucial. For example, it is well known that the bacterial cell wall cannot be constructed spontaneously. It is very unlikely that ribosome can be constructed spontaneously from its parts. Notice that to fold a protein numerous chaperones (folding catalysts) are usually required [2].

Now, we clearly understand why all life is from life. Organisms lack self-organizing capability. We should recognize that self-organization is a telltale sign of simplicity. Something can happen spontaneously, because there are virtually not many ways to unfold the system or phenomenon. Unfortunately, however, often self-organizing property has been regarded as an important characteristic of complex systems. For example, Levine [3] says, "By self-organization I mean simply that not all the details, or "instructions" are specified

in the *development* of a complex system." That is, he emphasizes that complex systems are characterized by the non-necessity of all the details to develop. Our emphasis point is fairly different. Needless to say, there are many details that are not required to be specified, but the existence of an indispensable core of (numerous) conditions that must be specified in detail is an important key feature of complex phenomena and systems.

If we ignore this distinction between (A) and (B), we will never understand the crucial nature of organisms. This point is completely ignored by Prigogine and Nicolis [4]. They emphasized that the difference between life and nonlife was not so large as had been thought. Thus, we physicists could relatively easily redirect our energy without any serious possibility of danger. However, as the readers have already sensed, this is a fundamental error that has misled complex systems study.

'Complex systems' require nontrivial auxiliary conditions. However, even if we impose a nontrivial initial condition, not all the systems can exhibit nontrivial results. For example, if we have a very large 2D Ising model, we can draw a very detailed gray-scale Mona Lisa on it (assuming, say, up spins are black pixels and down white). If the system evolves according to the Glauber dynamics, the masterpiece would disappear before very long. This indicates that systems like (B) must be able to store the initial information. That is, if a system can be called a complex system at all, it must be able to behave as a memory device. This observation is rather important, because memory is usually carried by broken symmetry. A complex system must be able to respond effectively to many initial conditions, so it must be a system with a lot of broken symmetries. Organisms must be full of such symmetry breaking processes whose outcomes must be specified carefully. This is, however, not at all a new point of view. Waddington clearly recognized this as illustrated by his epigenetic landscape (Fig. 2.1).
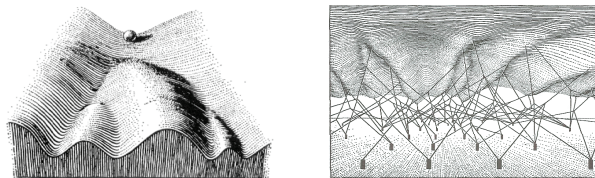

Fig. 2.1 Waddington's picture of developmental process as a cascade of symmetry breakings. Incidentally, the right depicts how genes support this landscape. Both illustrate deep ideas [5].

The so-called emergent property is often an epiphenomenon of symmetry breaking. Prototypical examples are the position and the direction of the salt crystal in (A). It is correct that emergent properties are crucial for complex systems just as self-organization is. However, again the recognized reason is wrong. Emergent properties are important because they are indeterminate and can be specified by the initial conditions.

It should be clear that the so-called complex systems study has failed to recognize the most important prerequisite for complex systems. This failure has, however, a deep root in the history of the modern physics (after Newton).

3   What has not physics done?

What is the structure of the Newtonian Revolution? It may be schematically written as

Phenomenon = Auxiliary Conditions + Universal Laws.

This is well illustrated by the Newton's equation of motion for the Kepler problem:

$$m\frac{d^2\boldsymbol{r}}{dt^2} = -G\frac{M_\odot m}{r^2}\frac{\boldsymbol{r}}{r}, \tag{3.2}$$

where $M_\odot$ is the mass of the central star and m that of a planet. $G, t, \boldsymbol{r}$ have the usual meaning. Since this is an ordinary differential equation (ODE), we need an initial condition. In this case it is not hard to obtain accurate auxiliary conditions, so the discovery of the fundamental law (the ODE) was received with amazement. Also this resounding success seems to have determined the general direction of physics: to find fundamental laws (FL). The ODE (3.2) still contains contingent terms such as $M_\odot$, G, etc. Let us reduce contingent elements from FL. This is the movement illustrated in Fig. 3.1. The modern high energy physics advocates the ultimate version of this philosophy.
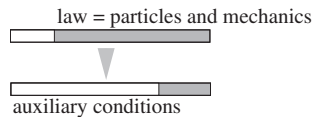
law = particles and mechanics

auxiliary conditions

Fig. 3.1 The history of modern physics is to make the fundamental laws as pure as possible.

String theory wishes to squeeze out all the contingencies from FL. Thus, it is a popular idea that the progress of physics is in the direction shown in Fig. 3.1. What is the consequence? If we push the white-black boundary to the right, we need more 'contingencies.' Even to explain the four fundamental forces we need symmetry breaking processes. That is, to explain the particular world we live in, we must add contingencies (a lot). The growing white portion in Fig. 3.1 will be left intact by (cutting-edge) physics.

We have already noted in the last section that there are very interesting systems including ourselves that require a vast amount of auxiliary conditions. Not only the amount is vast, they are also much harder to obtain or to specify than FL as (B) exemplifies. Up to now contingencies are simply ignored because it does not seem to give universal features physicists love to find. However, there might be universal features there; at least to find universal features in the 'white' region could be an important future direction of natural science, because this region is almost everything in the future.

Even for organisms not all the auxiliary conditions required by eq.(2.1) need be specified uniquely. Obviously, we need not specify the position of every atom. However, we must clearly specify the auxiliary conditions that specify sectors after symmetry breaking. Let us call such indispensable auxiliary conditions Fundamental Conditions (FC). To understand a phenomenon from physicists' point of view is to understand FL and FC (see Fig. 3.2).
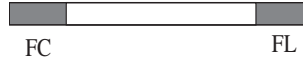
4

Fig. 3.2 We wish to understand Fundamental Conditions (FC) and Fundamental Laws (FL).

4 Fundamental Conditions

If we are interested in genuinely complex systems, we should concentrate our attention to FC. Therefore, in this section let us exhibit preliminary considerations on FC.

In the preceding section auxiliary conditions may be entitled to be called FC that specify sectors after symmetry breakings. In this sense, even the system (A) has a room to accommodate some FC (to specify the position of the crystal and its orientation). In this case to find out FC is not very hard. Therefore, from now on we pay due attention to FC that contains numerous conditions (i.e., ideally, we take a sort of 'thermodynamic limit'). Thus, auxiliary conditions satisfying the following two conditions are FC:[2]
(FC1) FC must be uniquely specified to realize system's characteristic features; especially they must specify the fate of the system after symmetry breaking processes.
(FC2) FC cannot emerge spontaneously (within the characteristic time of the system).
The second condition implies that history and tradition are crucial.[3] Often physicists hate history. However, we should listen to Ortega stressing the importance of universality.[4] If he is right, we may believe in universal features in biology.

First, we assume a principle that there is no information not carried by material structures (microscopic no-ghost principle).[5] Classification of FC is an important topic, but no systematic consideration has been given yet, so here we will be contented with a rather informal classification: If a structure itself must be specified in the auxiliary condition and if the specified structure itself functions as information (say, as a template), we call the FC a structural FC (SFC). If a specified structure is used as a symbol,[6] we call it a symbolic or instructional FC (IFC). Although it must be stressed that FC for an organism is not exhausted by the genome (the totality of DNA in a cell),[7] let us take a genome to illustrate SFC and IFC. A genome consists of
>   C: coding part
>>   Cs: structural — e.g., structural proteins, enzymes

---

[2]I have no intention to confine FC to be characterized by these two conditions only.

[3]To respect the complexity of our society is to respect tradition as Hayek stresses. We must not forget to pursue the consequences of (corrected) complex systems study in the humanities as well.

[4]"It is true that it is only possible to anticipate the general structure of the future, but that is all that we in truth understand of the past or of the present. Accordingly, if you want a good view of your own age, look at it from far off. From what distance? The answer is simple. Just far enough to prevent you seeing Cleopatra's nose." (Ortega, *La rebelión de las masas* (1930), p55). This is nothing but an expression of his belief in universality.

[5]However, there are informations carried by emergent structures; they look rather like ghosts, so the adjective 'microscopic' is attached. It may well be the case that the general chaperone atmosphere or that of the genomewide methylation condition can collectively carry important cues. In this paper this important topic will not be discussed.

[6]in the sense of C. S. Peirce

[7]Recall even irradiation damage that kills a bacterial cell is not on its DNA but on its proteins.[6]

Cr: regulatory — e.g., transcription factors

NC: non-coding part

I mean Cr + NC is IFC and Cs is SFC. Probably, the genome is the easiest part to handle among all the FC required by an organism as can be seen from a recent whole genome replacement experiment [7].

Informally, an organism requires FC, but this FC cannot be produced *de novo* (cf. FC2). If constructing an organism in this world is analogized as Biology solving a problem posed by Nature, without FC Biology cannot solve the problem. Thus, we must regard the problem posed by Nature very hard. To solve it within a short time FC is required as a sort of an oracle set in the sense used in the theory of computation.

Let us tentatively characterize a complex system as a system at least requiring FC. Then, it is likely that

'Theorem' D. A complex system can die and never be resurrected.

'Theorem' E. If a complex system has existed for a sufficiently long time, it must be maintained by Darwinism.

'Theorem' C. A complex system that requires SFC must be spatially locally segregated.

Here, quotation marks imply that these propositions are hoped to be theorems eventually in the true sense of this word. At this level of insufficient formalization, however, they are pseudotheorems summarizing informal arguments.

We need several basic assumptions about the world:

A1. Ubiquity of Almighty Noise that damages/degrades everything including FC.

A2. Available information is always space-time local and rather small (incomplete).

A2 implies that the law of large numbers (LLN) cannot be used fully to beat noise. A1 means that FC will be damaged sooner or later. Then, A2 + FC2 implies that damaged FC cannot be repaired in the long run. Hence, 'Theorem' D.

Then, how can complex systems continue to exist despite noise? Damaged FC must be restored to the level of being capable of forming a viable organism. However, there is no correct FC posted anywhere. Since no large scale information collection is possible, LLN cannot work fully and sooner or later noise destroys the original FC. Thus, comparing several FC does not guarantee to restore the uncorrupted FC. Furthermore, the comparison of FC is actually not practical, either, as can been seen from the following consideration. We must recognize that the analogy between making an organism and solving a hard question posed by Nature is rather deep. As is well known, hard problems are represented by NP complete problems. For such problems finding a solution is hard, but to check the correctness of a solution is easy. In our context to check that FC gives the right solution is equivalent to forming a viable organism. Of course, we know developmental process is 'easy' (far easier than the evolution process that created FC). Thus, comparing FC without forming actual organisms is computationally inefficient.

Suppose a corrupted FC is known to fail to give a viable organism. How can an uncorrupted FC be restored? Correcting corrupted solution to a hard problem is again a hard problem.[8] Therefore, given a corrupted FC, whether 'supposedly repaired versions' are really repaired or not can be checked only through using them to form organisms. Notice that this procedure is non-deterministic in the sense of the theory of computation. We must pro-

---

[8]Notice that there is no incremental (or recursive) way to solve NP hard problems.

duce (supposedly) repaired FC randomly (or with a stochastic process whose Kolmogorov complexity is at least comparable to that of the noise in the world; as we know the actual biological systems use the *same* noise that can destroy them). Then, we 'run' them to form corresponding viable organisms. How can we find the corrected outcome? By comparison, but comparison with what?

We need another assumption as to the finiteness of the world, or as to the finiteness of the available resources:

A3. We cannot sustain indefinitely many organisms.

Inevitably, comparison becomes competition and entails selection.

Our argument up to this point may be summarized as follows. If we assume that complex systems exist for sufficiently long time, then FC is maintained against noise by non-deterministic computation. This implies reproduction and mutation are absolutely needed to *maintain* complex systems for a long time. Because of the bounded resources this process entails selection. Thus, if there has been a complex system for a sufficiently long time, it must have been sustained by Darwinism. This is 'Theorem' E. However, Darwinism does not imply the presence of complex systems as was demonstrated by Spiegelman's monster [8].

FC must contain SFC, if we consider a complex system that arises from the microscopic molecular scale. On this basis can symbolic or IFC work.

(1) SFC provides templates/nuclei to guide self-organizable materials. If repetitive structures are barred and if there is no IFC, the size of the resultant system is comparable to the size of SFC.

(2) Structure itself is the information for SFC; the components of SFC must keep well-defined spatial relations, so SFC must be localized in space.

(3) The SFC for a system and SFC for other systems must be clearly distinguished.

Therefore, complex systems requiring SFC must be locally segregated in space. If it is not tethered (e.g., as a single polymer), it must be enclosed in a well-defined domain. In this sense cellularity is required by complex systems. This is 'Theorem' C.

According to Barbieri [9] the two pillars of the modern biology are Darwinism and Cell theory. Notice that complexity implies both. There are two versions of cell theory, weak and strong. The weak version asserts that the organisms we know are made of cells. Here, 'Theorem C' asserts the strong version: all organisms are made of cells.

5   How to obtain Phenomenology of Complex Systems

We might be able to push such a deductive study further, but our imagination and logical power are very limited. Every empirical science must start with phenomenology. Here, I mean by this word a summary of what we observe free of any ontological bias. We should clearly recall that without thermodynamics, equilibrium statistical mechanics could not be created/cannot be formulated; because there is no well defined phenomenology nonequilibrium statistical mechanics does not exist. If we wish to understand a class of phenomena or systems in general terms (in a non-fetish fashion, so to speak), we need good phenomenology. To have a phenomenological summary we must collect facts. Therefore, I must start with surveying the world of organisms. Most physicists may call this stamp collecting, following

Rutherford, but the hope here is that even stamps may tell us important lessons if sufficiently accumulated. However, already G. G. Simpson warned in *The Meaning of Evolution*: "Indeed facts are elusive and you usually have to know what you are looking for before you can find one." Thus, my strategy is to organize observed facts around FC. This project does not aim at contributing biology but primarily at enriching physics (and perhaps mathematics), but as a physics imperialist I firmly believe that this is the only way to understand biology properly.

Fundamental questions about FC include the following:

(1) How is FC organized? e.g., organization of genome.

(2) How is FC used?

development, aging, — at the organismal time scale.

behavior, defense, homeostasis — at the time scale of sec to min.

(3) How has FC been changing?

History—paleontology and evolution.

Resultant diversity—taxonomy.

Collective phenomena— ecology, sociobiology.

We deliberately ignore the aspects directly connected to materials. This implies to ignore a large portion of biophysics as irrelevant. We have seen that complex systems require a lot of symmetry breaking processes that are prepared by self-organizing properties of materials. Therefore, naturally there are two research directions: (i) emphasizing self-organizing capability of materials, (ii) emphasizing indeterminate aspects produced by self-organization. Biophysics stresses (i); it is largely molecular and materials science of biological matter; dead bodies, albeit fresh, are enough; physics of minced meat. In contrast, I stress (ii), because how to utilize the emergent indeterminacy is the key to complex systems.

The inductive part of phenomenology consists of two parts: (PI1) Compiling facts and (PI2) Distilling phenomenology from the facts. The main part of PI1 is, for physicists, to develop new (methods and devices to aid) experiments and field work. If you are interested in biology seriously, you should have a taxonomic group you are familiar with. Natural history is very important, because we are interested in universality. We can find universality only through comparative studies. In this article I do not discuss any experimental aspects, but I wish to emphasize the importance of developing high-throughput phenomic studies (in contrast to the genomic studies already in bloom). We are interested in organisms, not in molecules *per se*; molecules make sense only in the light of natural history. This is why I must stress phenomics. PI2 is the data analysis, text-mining, etc. This is also crucial because we are inundated with numbers from high-throughput experiments.

We should look for exact phenomenologies like thermodynamics that allows us to make precise predictions. I do not have such a well defined phenomenology yet. However, some general observations I have may already be of some use. As an example, in the next section, I outline presumably the most common complexification process.

6 Basic observations about FC and complexification

What is complexity? Our provisional necessary condition for a complex system is that it requires FC for its construction. Therefore, a certain quantitative measure of FC might characterize the complexity of a system. However, probably it is the consensus that complexity has many facets, so organisms are not well ordered with respect to complexity [10]. Therefore, it may be expected that no single important complexity measure exists. However, Fig. 6.1 is an interesting observation about the non-coding DNA [11].

A remarkable message of Fig. 6.1 is that the usual anthropocentric viewpoint detested
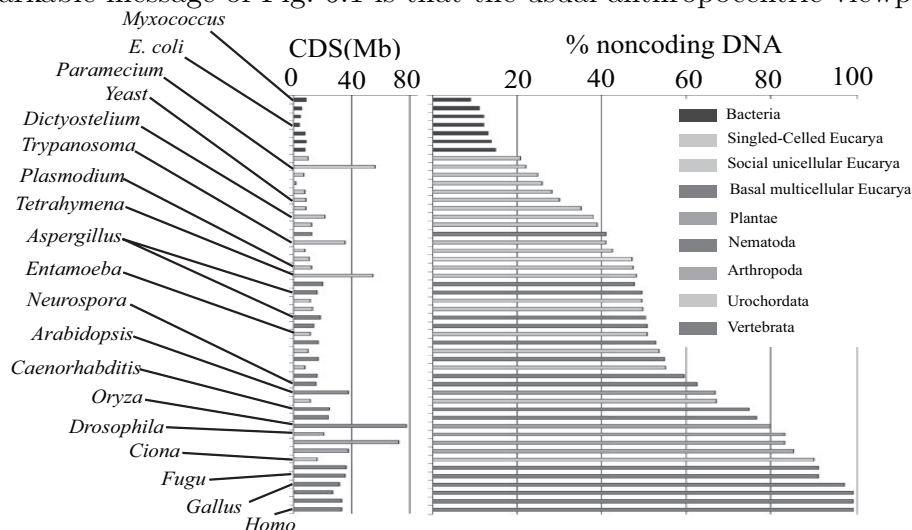

Fig. 6.1 If we pay attention to the amount of non-coding DNA in the genome, organisms are ordered naturally in the usual 'anthropocentric' order which Gould detested. CDS is the amount of protein coding DNA in megabase. The figure is due to Taft et al.[11]

by Gould [12] (and by Woese [13]) seems vindicated. From the point of view of FC, the amount of IFC is a good measure of complexity. This viewpoint is consistent with the evolution of microRNA [14]. A natural logical consequence is, as can be seen from Fig. 6.1, that study of complexity or biocomplexity must be the study of Eucarya.

Complex systems may be classified into two major classes; one that mainly relies on SFC alone, and the rest that relies on IFC as well.[9] The Procarya/Eucarya dichotomy roughly corresponds to this distinction. Thus, even though Procarya is a paraphyletic group, it may be mathematically a well defined natural group. It should be recognized that spontaneous formation of a 'large' system is impossible with Brownian motion + SFC alone.[10] If a large

---

[9][Footnote added on Nov 21, 2008] Viruses may be regarded as complex systems mainly relying on IFC. The proposed classification of all organisms into two major classes: ribosome-encoding organisms and capsid-encoding organisms suggests this point of view. See D. Raoult and P. Forterre, "Redefining viruses: lessons from Mimivirus," Nature Rev. Microbiol., **6**, 315 (2008).

[10]One might say dissipative structure may evade such constraints. However, dissipative structures without microscopic materials organization change are too fragile to be relevant to biology. Those with materials bases are essentially equilibrium structures modulated by dissipation. Thus, dissipative structures are basically irrelevant to biology.

system requires FC, it requires IFC. This also implies that Procarya is not really interesting from the complexity point of view.[11]

We know FC can be made only through Darwinism.[12] Then, how does it evolve? A possible question is: Can FC grow gradually? I do not know any such example. Requiring new symmetry breakings may be at odd with continuity. My conclusion is summarized in Fig. 6.2.
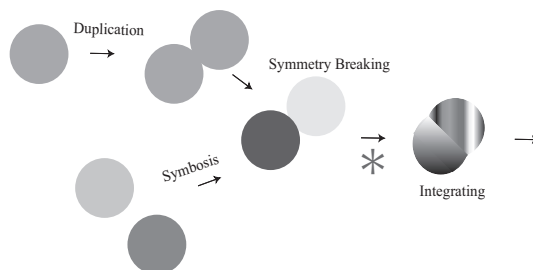

Fig. 6.2 A unit process for complexification: it consists of two steps; the duplication/juxtaposition step before ∗ and the much more important integrating step ∗.

Complexification occurs in two steps. Complexification requires symmetry breaking (or rather numerous symmetry breaking processes). Symmetry breaking requires a stage that can accommodate various sectors created by symmetry breaking processes. Expanding the stage is often due to symbiosis (juxtaposition) or due to duplication/multiplication. The importance of duplication and subsequent subfunctionalization/neofunctionalization was stressed by S. Ohno [15]. The importance of symbiosis has been stressed by Levine [16]. Major historical events other than mass extinction may have been driven by establishment of new symbiosis. For example, the landfall of green plants must have been due to plant-fungi alliance. There are numerous such examples.

This is the first of the two major steps making a unit process of complexification. The second step is integration. There is no good name for this step, because the real importance of this step has not fully been recognized as a general process. In short, what happens is to use all the elements created by the first step in order to realize a higher level organization. Even in the case of endosymbiosis, the process of incorporating endosymbionts into the host cell as new cellular organelles to create a new type of cells may be understood as an example of this second step. Incidentally, most Procarya cannot afford duplication due to its sheer size. This is also a reason why there is not much complexification in Procarya; the most important complexification path is blocked.[13]

---

[11]One might say that from the biodiversity point of view Procarya is crucial. We could say where there is a free energy difference there is a prokaryote exploiting it. However, this is a diversity of organic chemistry; if we change methyl to ethyl to propyl to ⋯, we could make a diverse set of reactions and compounds. Thus, I bet that only in this sense Procarya is diverse, so from physicists' point of view a simple universal picture might be obtainable for the whole Procarya.

[12]Even if God were to exist and to have created organisms, to maintain them against Almighty Noise we need Darwinism. Consequently, even the initial intelligent design is meaningless under Almighty Noise.

[13]One might say that extending the biofilm and other multicellular structures even Procarya could complexify. However, this is highly unlikely due to frequent adaptive sweeps.

The second step of the complexification process is the crucial step. The first step is often a preparatory step. This step can quantitatively increase parts and functions, but qualitative changes may not occur there. The idea is supported by the formation of, e.g., Metazoa and Bilateria. We now know that Choanoflagellata has [17] many signal pathway components and cellular communication molecules that are organized and utilized by Porifera[18]. Even Anthozoa (Cnidaria) has (and probably Porifera had) Hox genes [19]; Hox genes are used to make the bilaterian body plan. Another example is our language. It is highly likely that all the components required by the linguistic capability exist in primates. Therefore, the rate process for the emergence of language could have been the integration step. The lesson is: some sort of 'nucleation process' that starts to integrate preexisting key components is really the crucial step to achieve higher level complexity. Even the evolution of society and civilization could be understood along this line. This is the step marked with $*$ in Fig. 6.2.

It is often said that excessively specialized organisms cannot evolve. Perhaps, our general consideration sheds some light on this folklore. The complexification process consists of two steps. If an organism loses many elements created in the first step, the integration step would be virtually aborted or at best incomplete. In this sense, complexification occurs most likely in the lineage preserving most primitive (or plesiomorphic) features. Loss of features prepared during the first step seems to be the key ingredient of 'specialization.' Furthermore, many examples tell us that an efficient way to lose these features is the sessile and/or filter feeding life style (or the loss of capability to move around [20]). The observation is supported by our position in Deuterostomia.[14] Echinodermata and Hemichordata are specialized branches compared with Chordata. Within Chordata notice that Cephalochordata that can move around is the most primitive to which we (Vertebrata) are close; Urochordata are sessile filter feeders, so they are specialized. Thus, we humans are in the lineage of the least specialized within Deuterostomia. The same may be said about Deuterostomia among Metazoa. We can expect that actively moving creatures were the common ancestors of Calcarea and Eumetazoa, so we came from something like planulae. Porifera are sessile filter feeders, a dead end from the complexification point of view. Notice that the Planulozoa-Porifera relation reminds us of the Cephalochordata-Urochordata relation. Thus, we humans are in the lineage of the least specialized within Animalia. The recent *Nematostella* genome [21] supports this point of view. Where is then Opisthokonta that includes Animalia within Eucarya? It is likely that Unikonta is the basic group. Again, we are in the group basic to Eucarya.

To simplify, we may say that the first expanding stage of the complexification process prepares a (wide) stage and actors. The second step gives scenarios. IFC is crucial in this step. Specialization implies loss of actors (and a shrinking stage) before any interesting play begins. Sessile life style is an efficient way to decimate actors.

---

[14]For convenience, some classification rudiments are given here. We vertebrates are in Chordata containing Cephalochordata and Urochordata (sea squirt, etc.) as well. Chordata is among Deuterostomia with Xenoturbellida, Echinodermata (sea urchins, sea stars, etc) and Hemichordata. Deuterostomia is among Bilateria (including most of invertebrates). Bilateria and Cnidaria (sea anemone, hydra, etc.) make up Eumetazoa, which with Porifera (sponge) makes the major portion of Metazoa (= Animalia). The closest sister group in Opisthokonta to Metazoa is Choanoflagellata. Opisthokonta includes Fungi and us and is one of a few kingdoms of Eucarya. In Domain Eucarya Opisthokonta is among Unikonta with Amoebae.

7 Potential use of qualitative phenomenology

There have been attempts to construct a 'general theoretical framework' of biology, but most of them never seriously looked at the actual biological world. Thus, these studies are not so interesting to biologists.

The program I am proposing may be called Integrative Natural History that unifies molecular, phenomic and much larger scale observations to understand genuine complex systems in an unified fashion.[15] Its theoretical (deductive) side consists of two parts: (PD1) Constructing phenomenological theory of complex systems as mathematics and (PD2) Formulating many biologically meaningful questions based on the phenomenological summary.

Although I do not have any precise phenomenology, it seems possible to say something on the PD2 side. For example, we have already seen that major historical events other than mass extinctions are likely to be driven by symbiotic relations that open up new (vast) niches. Then, a natural big question is whether there was a major symbiotic relation driving the landfall of animals. Gut symbionts must be paid due attention.

We have also seen that organisms with most plesiomorphic traits complexify most; we are directly related to the most basic group of Eucarya. What is the natural 'analytic continuation' of this idea? The most natural conclusion is that (anaerobic) Eucarya[16] is the basic organism, and Procarya is a specialized group that lost many plesiomorphic features. This idea may sound crazy, but notice that the common notion that Procarya (Bacteria) was earlier than Eucarya has no unambiguous supporting fact. This crazy idea may well be consistent with the evolution of codons [22] and introns [23].

I wish to conclude the article with two more such general observations.
(1) Biological systems lack foresight.
This may be called the 'principle of non-determinism' (in the theory of computation sense). This is well understood and not at all a new observation. Evolution process itself is a great example. Overproduction combined with subsequent thinning is a common strategy in neural development. There is every reason to believe that molecular machines work on this principle: the right outcome is not aimed at. Instead, when the desired outcome is realized, it is stabilized (preserved). Even the translation of a DNA sequence to the corresponding amino acid sequence relies on this strategy. Thus, we may say, molecular machines work on the principle of motional Darwinism. This implies that driving with some potential is not essential for molecular motors. Only some steric hindrance forbidding some class of movements is needed.
(2) Biological systems cut Gordian knot(s), or biology never solves difficult problems (the Gordian Principle?). There are many mathematically difficult problems apparently relevant to the biological world.

A famous example is the protein folding problem: to determine the tertiary structure from the primary sequence. Since there is an astronomical number of conformations possible for a given primary sequence, Levinthal pointed out that folding should be combinatorially very hard. However, we know many proteins are formed very quickly *in vivo*. Has biology

---

[15]This unification has a much more significant implication in biology, because organisms are in a certain sense inflated microscopic systems, quite different from many systems physicists have been studying that have layered structures with separated micro, meso and macroscopic levels.

[16][Added on Nov. 21, 2008] I do not mean extant mitochondrion-less organisms such as *Giardia*.

solved this problem? No. For short chains, Biology picked up special quickly foldable sequences, and then connected them to make longer chains. Thus, proteins consisting of up to 200-300 amino acids are based on special quickly foldable proteins. The problem is not solved generally at all. For longer chains Biology has given up, and relies on chaperones (folding catalysts).

Another famous problem is the DNA entanglement, a topological difficulty. When prokaryote DNAs were found to be ring-shaped, mathematicians were delighted, expecting topological invariants to interfere replications, etc. However, such a topological difficulty is non-existent thanks especially to topoisomerase II.

Statistical physicists discussed vesicle shapes, expecting such problems were biologically relevant (think of ER): how can complicated shapes be formed 'spontaneously'? However, no differential geometrical problem of manifolds arises, because these membrane shapes are in detail controlled even stoichiometrically by membrane associated proteins, e.g., clathrin, caveolin, coatomer complex, etc.

These may be rather disappointing stories for mathematically oriented theoreticians, but we should learn an important lesson. Whenever we can expect difficult problems in biological processes (or rather, processes relevant to complex systems), the real problem for theoreticians is to think how Biology avoids solving or even facing them. I wish to point out three such examples. I believe all are important problems.

(i) We can expect combinatorial difficulties in regulation of genes. Therefore, Biology must have worked hard to avoid unnecessary combinatorial entanglements. For example, after duplication, gene-gene relations that may cause entanglements are selectively suppressed [24].

(ii) We can make extremely highly nested clause structure 'formally,' so this recursiveness has been stressed as an important feature of our natural language. However, as is well known, self-referencing can cause numerous paradoxes such as the liar paradox. How do our brains avoid such logical difficulties? We should reflect on our daily language practice; we usually do not decide whether we decide or not whether we should go to a meeting or not, for example. Thus, it is highly questionable that our brain honestly handles recursiveness as such.

(iii) If you read contemporary ethics textbooks, you will certainly find many serious ethical aporias. However, if we take seriously the lesson Biology teaches us, we ought to realize that the most important problem of ethics is to avoid the situations of ethical aporias; we should work hard to avoid hard problems! Although the current environmental problems are not simply due to the population problem[25], still there is no doubt that the human population problem will be the most important factor that will destabilize our world. How do our fellow creatures cope with this problem? It is known that the actual wild population is far less than the environmental capacity (self-limitation is observed [26]).

Sociopolitical issues should not be avoided by physicists.

Bibliography

[1] Y. Oono, Int. J. Mod. Phys. B, **12**, 245 (1998).

[2] See, for example, J. C. Young, V. R. Agashe, K. Siegers and F. U. Hartl Nature Reviews Molecular Cell Biology **5**, 781 (2004).

[3] S. A. Levine, *Fragile Dominion* (Perseus Publishing, New York, NY, 1999).

[4] G. Nicolis and I. Prigogine, *Exploring Complexity: an introduction*, (W. H. Freeman, New York, 1989).

[5] C. H. Waddington, *The Strategy of the Gene* (Allen and Unwin, London, 1957).

[6] M. J. Daly, E. K. Gaidamakova, V. Y. Matrosova, A. Vasilenko, M. Zhai, R. D. Leapman, B. Lai, B. Ravel, S.-M. W. Li, K. M. Kemner, J. K. Fredrickson, PLoS Biol., **5**, e92 (2007).

[7] C. Lartigue, J. I. Glass, N. Alperovich, R. Pieper, P. P. Parmar, C. A. Hutchison III, H. O. Smith, J. C. Venter, Science, **317**, 632 (2007).

[8] D. L. Kacian, D. R. Mills, F. R. Kramer, and S. Spiegelman, Proc. Natl. Acad. Sci., **69**, 3038 (1972)

[9] M. Barbieri, *The Organic Code, An introduction to semantic biology* (Cambridge, 2002).

[10] D. W. McShea, Evolution **50**, 477 (1996).

[11] R. J. Taft, M. Pheasant, and J. S. Mattick, BioEssays **29**, 288-299 (2007).

[12] S. J. Gould, *The Structure of Evolution Theory* (Harvard UP, 2003), esp., p897 and p1321.

[13] C. R. Woese, Proc. Nat. Acad. Sci., **95**, 11043 (1998).

[14] R. Niwa and F. J. Slack, Curr. Op. Gen. Dev., **17**, 145 (2007); L. F. Sempere, C. N. Cole, M. A. McPeek, and K. J. Peterson, J. Exper. Zool., B **306**, 575 (2007).

[15] S. Ohno, *Evolution by Gene Duplication* (Springer, New York, 1970)

[16] S. A. Levine, PLoS Biol., **4**, e300 (2006)

[17] N. King and S. B. Carroll, Proc. Natl. Acad. Sci., **98**, 15032 (2001); N. King, C. T. Hittinger, and S. B. Carroll, Science **301**, 361 (2003).

[18] W. Dirks, J. S. Pearse, and N. King, Proc. Nat. Acad. Sci., **103**, 12451 (2006).

[19] K. J. Peterson and E. A. Sperling, Evol. Dev., **9**, 405 (2007).

[20] The importance of movement in the compleification process has been stressed by T. Ikegami. See, for example, his recent book, *Movement Creates Life* (Seido-sha, 2007) [In Japanese].

[21] N. H. Putnam, M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, A. Salamov, A. Terry, H. Shapiro, E. Lindquist, V. V. Kapitonov, J. Jurka, G. Genikhovich, I. V. Grigoriev, S. M. Lucas R. E. Steele, J. R. Finnerty, U. Technau, M. Q. Martindale,7 D. S. Rokhsar, Science, **317**, 86 (2007).

[22] S. Itzkovitz and U. Alon, Genome Res., **17**, 405 (2007)

[23] A. M. Poole, D. C. Jeffares and D. Penny, J. Mol. Evol., **46**, 1 (1998).

[24] G. C. Conant and K. H. Wolfe, PLoS Biol., **4**, e109 (2006).

[25] J. Vandermeer and I. Perfecto, *Breakfast of Biodiversity, the political ecology of rain forest destruction*, 2nd Edition (Food First Books, 2005).

[26] J. E. C. Flux, Oikos **92**, 555 (2001).