

The Historical Necessity of Life for Cognition

Orlin Vakarelov

Department of Philosophy

University of Arizona

Tucson, Arizona

okv@u.arizona.edu

January 12, 2009

Abstract

I argue that life — metabolic systems satisfying the property of autopoiesis — is *structurally historically necessary* for cognition. That is: (1) the first place where cognition is needed and emerges naturally is within living (metabolic) agents; (2) non-metabolic cognitive agents can emerge only as artifacts of other cognitive agents, the earliest of which must be living. In order to make this precise I define the notion of *priority hierarchy of complexity*, which tracks the construction dependencies of structure formation. This last concept is based on a theory of propagation for order due to Stuart Kauffman.

Keywords: Autopoiesis, Origin of Cognition, Propagation of Order, Hierarchy of Complexity

1 Introduction

With the maturing of cognitive science there has been a renewed interest in the connection between cognition and life. Similar connection has been recognized within the later phenomenological tradition, most explicitly in the work of Hans Jonas [9]. The renewed interest follows the work of Maturana and Varela [13] who attempt to provide a precise characterization of life, using the notion of autopoiesis¹, and argue that cognition is implied by autopoiesis. The basic claim is that the principle of life — the notion of autopoiesis — is sufficient for cognition. A stronger claim is also made. It is claimed that life is necessary for cognition, and that in fact life and cognition are coextensive — cognition and life are two aspects of the same phenomenon. As stated, both the sufficiency and the necessity claims are implausible as general claims about cognition. About sufficiency, it has been argued quite convincingly by Di Paolo [7], for example, that the notion of autopoiesis does not imply adaptability, which seems to be required even for minimal notions of cognition. ([4, 5] provide alternative criticisms of the sufficiency claim.) About necessity, it can be observed that the necessity claim excludes the possibility of cognitive systems in (e.g.) non-metabolic robots. It is not clear why an advanced robot constructed out of metal and wire would not possess at least simple cognition, if a bacterium would. For the purposes of this paper I will grant that non-metabolic systems can be cognitive. Thus, I will assume that the necessity claim is not correct. Still there exists a kind of necessity relation — what I call *structural historical necessity* — between life and cognition which makes the existence of life essential for the existence of

¹A system is *autopoietic* if:

1. it has a semi-permeable boundary,
2. the boundary is produced from within the system, and
3. it encompasses reactions that regenerate the components of the system. ([16], via [12], and [5])

cognition. The central claim of this paper is that *life² is structurally historically necessary for cognition*.

The intuitive idea behind this claim is: (1) the first place where cognition is needed and emerges naturally is within living (metabolic) agents; (2) non-metabolic cognitive agents can emerge only as artifacts of other cognitive agents, the earliest of which must be living. Making this intuition precise demands a theory of the emergence and propagation of organized structure, which allows us to say what it means for an organized system to emerge first, before another organized system; and, what it means for one system to be an artifact of another. This theory must define the needed notion of structural historical necessity. While fully developed theories of propagation of order have not been created, there are embryonic theories that attempt to provide understanding of the natural origin and increase of order in the universe, special cases of which are both the phenomena of life and of cognition. Here I exploit a theory proposed by Stuart Kauffman [11, 10].

The paper provides only an argument sketch of the central claim. A full argument demands a careful definition of the notion of cognition. Currently there are unresolved philosophical/scientific debates about the precise scope of cognition. Providing a definition of cognition is beyond the scope of this paper, I will resort to an intuitive conception only. Also, a full argument demands a more detailed account of how cognition (and various cognitive capacities) may emerge — especially cognitive capacities needed for an agent or a community of agents to be able to create an artificial cognitive system. Such details must come from cognitive science itself, and are also beyond the scope of this paper.

²Whether *life* is properly characterized by a notion similar to autopoiesis is controversial, especially within the biology community. My goal here is not to defend that life is characterized in such a way (although I hold that such a characterization is on the right track). Thus, for the purpose of this paper it is sufficient to claim that *autopoiesis is structurally necessary for cognition*.

2 The propagation of order

I take it to be an uncontroversial observation that, however they are related, both the phenomenon of life and the phenomenon of cognition exist only within highly organized complex systems. A rock, a crystal, a gas cloud can be neither living nor possess cognition. Physical system can be categorized according to a degree of their *complexity* and their *order*. [6] Intuitively, complexity can be viewed as a measure of how much information is necessary to completely describe the “bitmap” structure of the system. A complete description of an ideal gas of N molecules may require N^6 parameters. If N is a large number, as it is usually, then an ideal gas has relatively high complexity. A planetary system has only a few parameters in its description, it is a system of relatively low complexity. Order, on the other hand, measures how efficiently the description of the system may be compressed relative to its complexity. An ideal gas is essentially random, thus its description cannot be compressed, it is a low order system.³ The same is true for a planetary system. A periodic crystal has a very high level of order in relation to its relatively lower complexity. Some system, however, have both high complexity and high order — they are highly organized complex systems. Understanding the relation between life and cognition demands understanding the formation and propagation of organized complexity, or *order* for short.

Kaufman suggests a model for a process of increasing ordered complexity of a system. The central idea of the proposal is the dual relation between physical work and the existence of order. The presence of order in a system can be redescribed as the presence of dynamical *constraints* within the system. Con-

³Note that an ideal gas may be redescribed with macroscopic, aggregate parameters such as temperature, pressure etc. There are alternative measures of complexity that take into account the possibility of such re-description. [8] Under those measures gas is not a complex system. Such measures combine order and complexity under a unified scale whereby complex system are those that have both high complexity and high order. For some purposes such alternative measures of complexity are more useful, however they are more technically demanding to define.

straints are structures within the system that have the effect of constraining the dynamical behavior of the system. For example, if two gas molecules are connected with a rigid rod, then their motion is mutually constrained. A properly constrained physical system with available free energy can perform work. In fact, the physical notion of work can be defined as the *constrained release of energy*. [2] A piston can extract work from the exploding gas because of the structure of a half closed cylinder, channeling the pressure of the explosion into the lateral motion of the piston, which in virtue of the fact that it is attached off center of the rotor, can produce rotation that can drive the automobile. The physical organization of the engine, the constraints that the structure imposes on the dynamics, is essential for the performance of the work. Thus, if a system possesses minimal appropriate organization and free energy, the system can perform work. The effect of work, however can be more constraints, more structure, more order.

One central example of this phenomenon is found within (bio)chemical reactions. In a soup of chemicals some reactions can occur spontaneously (exergonic reactions), while others require energy to occur (endergonic reactions). Some compounds, *catalysts*, can lower the necessary energy of a reaction, making it many times more likely to occur. A catalyst acts as a constraint in the system that channels free energy into an energy hungry reaction.⁴ A catalyst allows, therefore, the system to perform work whose effect is the products of the reaction. The products may be novel compounds, including novel catalysts that could allow other reactions to take place. Another central example of this phenomenon is the development of technology. Instruments are physical (and informational) devices that organize and constrain the systems dynamics to channel energy in a specific way as to perform some work. A hammer is

⁴Some catalysts act a bit differently, not by lowering the needed energy but by bringing together reactants that are very unlikely to meet otherwise.

structured in such a way as to allow the channeling of a great amount of force on the small head of a nail, which is structured to penetrate material easily. A lathe is structured in such a way as to allow the precise positioning of a cutting tip on the surface of a rotating metal object, so that it can carve a cylindrically symmetrical profile. The product of such devices can be other novel and more advanced devices. Technology can propagate itself into more sophisticated technology following the simple principle that constraints can perform work, that can create novel constraints, that can perform more work, etc.

One of the central insights of Kauffman is that the recursive, bootstrapping process generated by the duality of work and constraints can occur spontaneously. Provided appropriate minimal conditions are satisfied — sufficient sources of free energy, relatively stable environment, sufficient initial concentrations of diverse compounds allowing sufficiently high initial probabilities of reactions, etc. — the system will naturally begin increasing the diversity and complexity of its organization.⁵ The process can be described as a sequential stochastic process of generation of novel structure.

Given a collection of actual structures in a system (e.g. chemical compounds, foldings of proteins, etc.) one can define the collection of possible novel structures that can be generated by the system with a single reaction. Kauffman calls the collection of novel structures the *adjacent possible* of the system.⁶ With this terminology the process can be described as a move into the adjacent possible,

⁵The precise nature of this process is fairly complicated. Kauffman spends considerable effort to characterize the nature of the process, which develops in a self-organized way, with explosions and partial collapses of diversity, where the very speed of the process becomes controlled by structures in the system. The details of the process are not important for the argument.

⁶Because the underlying process of structure formation is stochastic, what is included in the adjacent possible depends on a probability threshold. However, because constraints (e.g. catalyst) can increase the probability of a reaction from practically insignificant to very likely, there exist a (vague) collection of structures that are very likely to occur at the next stage. Moreover, because structures can decay probabilistically, there may be natural thresholds (depending in the particular compounds) below which the formation of a particular novel structure is too slow for it to make a difference in subsequent states of the process.

whereby at every stage of the process some novel order is generated, which order helps previously inaccessible order to be generated.⁷ The process has the effect of organizing the space of possible structures into a hierarchy — which I call the *priority hierarchy of complexity* — where structures at the bottom of the hierarchy can form spontaneously from the initial conditions of the system, and structures at consecutive levels require earlier stages of the hierarchy to have occurred.

The definition of the priority hierarchy of complexity over the space of possible structures gives me the necessary machinery to define dependence relations between structures and types of structures. For example, if we consider the collection of structures that are cognitive and the collection of cognitive structures that are living, and if it is observed that all cognitive structures that appear earlier in the hierarchy are living, then we can say that the *first* cognitive systems are living.

With the structure defined by the priority hierarchy of complexity I can define the notion of *structural historical necessity* (or *historical necessity*⁸ for short):

A type of structure A is *structurally historically necessary* for a type of structure B iff, in every possible priority hierarchy of complexity, structures of type A occur before structures of type B

The definition of structural historical necessity captures the idea that the formation of structures of type B requires the previous formation of structures of type A — every sequence of work steps needed to generate a structure of type

⁷The novelty in the hierarchy concentrates on structures that require work to achieve. Some structures, however, can self-organize from other available structures without the need of endergonic reactions. For example, if at a stage of the process lipids form, the lipids can self-organize into membranes spontaneously. Such structures can be included into the adjacent possible because no new “technology” is needed for their formation.

⁸The notion of *historical necessity* is not to be confused with the notion with similar name from temporal modal logic stating that the history of an even is uniquely determined. [15]To avoid this confusion I include the word ‘structural’ in the name.

B must have an essential use of structures of type A in it. If this is not the case, there exist a hierarchy of complexity where B gets generated without A , where A is not generated at all, or after B . A being historically necessary for B does not imply that A partakes in the immediate construction of an instance of B . A may have been required at an earlier stage for the production of a constraint needed for the production of a constraint, etc. needed for the generation of B . For example, while the technology for making of a transistor may not use a steam engine, it may be that the steam engine has been necessary for the initiation of the technology leading to the construction of the transistor — in fact, in our actual hierarchy of complexity the steam engine was essential for the sequence of technological steps leading to the transistor. As a final note on the definition, the word necessity is used in a weaker sense than in typical modal logic. The process generating the priority hierarchy of complexity is stochastic. Therefore, there is non-zero probability that a type B structure may appear in an alternative way. Similarly, there is non-zero probability that if the parts of a Jumbo Jet are thrown into the air they will spontaneously assemble into a complete plane; or, that a lightning striking a swamp will spontaneously generate a person. However, such probabilities are so hyper-astronomically small that they cannot occur in a scientific explanation of a phenomenon. The necessity in the notion of *structural historical necessity* relies on plausible sequences of work steps, each of which can occur with physically significant probability.

3 The Historical Necessity of Life for Cognition

It remains to explore two points: (1) Cognition must emerge first in living, metabolic systems; (2) non-metabolic cognition may exist only as an artifact of cognitive systems.⁹ Of the two, the second is sketchier, relying on intuitive

⁹These two points are a bit stronger than what is required for the central claim of the paper that *life is structurally historically necessary for cognition*. See the end of the paper

arguments. The arguments for both points are only defeasible, but I hope sufficiently compelling, at least at the level of a typical scientific argument.

The argument for (1) has two parts: (a) the principle of life, e.g. autopoiesis, provides the only way for sustaining the kind of ordered complexity necessary for cognition; (b) cognition, as a capacity, is itself needed by living systems in order to increase their complexity — to climb the priority hierarchy of complexity, so to say. Let us tackle (1a) first.

Structure requires a principle of persistence if it is to play any role in the process of generation of novel order. Thus, if we are interested in structures responsible for high complexity, we must consider what principles of persistence there may be. One such principle is *static rigidity*. A rock is a statically rigid structure. It is very stable, i.e. it can persist through large environmental fluctuations, because its internal connections are very strong. Another principle is *dynamic stability*. This principle applies only to the so called dissipative systems — systems that exist within a constant flow and exchange of matter and energy with the environment. Some such systems, fluid eddies, Benard cells in heated oil, etc. can remain stable for a considerable periods of time, however such systems depend exclusively, and quite sensitively so, on boundary conditions. Let us call such systems *heteropoietic stable dynamical systems*. An alternative dynamically stable dissipative system would be one where stability depends on the system itself — the system is such that it constructs and maintains itself with resources from the environment. Such a system satisfies what is referred to as *operational closure* — the constraints needed to do the work of constructing the system are contained within the system, and are themselves built by the system. Thus a system that is operationally closed is a system that can maintain its stability in light of larger environmental fluctuations than other dynamically stable systems. The idea of operational closure is central for any

for discussion.

general conception of life: it is central for the conception of autopoiesis (13); or, it is related to the notion of autocatalytic system, which is central for Kauffman's definition of life; or, it is connected to the notion of a metabolic-repair system [14], or to self-maintenant systems [3]. Life, therefore, is not merely a curious novelty of dynamics, it is based on a principle providing a novel, radical solution to the problem of persistence. Moreover, and most importantly, life allows a mode of stability that does not constrain, but in fact takes advantage of the increase of ordered complexity. In contrast, both rigid systems, and heteropoietic stable dynamical systems reduce their stability when their complexity increases — they either weaken their internal structure, in the case of rigid systems, or transition into chaotic systems, in the case of heteropoietic systems¹⁰. Moreover, such systems, especially rigid systems, provide little opportunity for novel order (outside of their participation in larger organized systems) because of their inertness. Without operationally closed systems that can support (and produce) high ordered complexity and remain sufficiently stable, the priority hierarchy of complexity would remain relatively low. It would not be able to achieve sufficient complexity to support the subtle control operations of a cognitive system.

To (1b) now: Even if life is needed for the priority hierarchy to get high enough for cognition, it is not immediately clear why cognition is anything more than an odd occurrence. Developing more and more complex life forms demands more efficient and subtle uses of environmental resources, more complex and sensitive constraints that can perform subtle work tasks. This demands more subtle and complex interaction of the system with its environment. If some task demands a rare chemical compound (or a technology demands a rare ore,

¹⁰This is a problem that all dynamically stable systems have, but systems with operational closure are much more resilient to collapse due to increasing of complexity. Operationally closed systems overcome this problem by organizing in multilevel structures, like in multicellular organisms. [11]

like uranium) the system must proactively seek, find and incorporate it. An operationally closed system of high complexity demands ability to adapt to its immediate (and future) environment. It must be able to enter perception/action control cycles. It must be able to discover and utilize patterns in its history to anticipate its future, etc. These are all capacities that are associated with (at least simple) cognition. Cognition is not a odd occurrence but an essential strategy for the increase of complexity (and success) of living systems.

Now that is reasonable to believe that life is needed to allow the priority hierarchy of complexity to grow high enough to allow the complexity needed for cognition, and that cognition itself provides constraints that can do the work of raising complexity higher, it remains to be convinced that it is reasonable to believe that cognition is necessary for the existence of non-metabolic cognitive systems. Here are two unsatisfactory moves that one may make, explicitly or by mistake. (1) We cannot rely on the trivial observation that because a cognitive system is the target of an AI project it must be an artifact of a cognitive system, namely the AI researcher. (2) We should be suspicious of arguments that identify some (what appears to be a) necessary feature of cognition that may appear in living creatures naturally, but must be transferred, like a virus, to a non-metabolic system. For example, one may argue that non-derived content can be grounded only in operationally closed systems because such and only such systems (e.g.) possess natural teleology [1]. Any other cognitive system must operate with content that have been derived ultimately from some grounded source, *ergo* the cognitive system must trace its origin to one such a system, similar to the way a viral infection may be traced to a source. Such an argument is good provided one can show that indeed the only way of getting non-derived content is the so-and-so and that such content indeed is needed for a system to count as cognitive. Usually, this part of the story remains mysterious. It

is likely that if it didn't remain mysterious, if one were to provide a detailed theory of non-derived content¹¹, one would be able to construct a non-metabolic system that can have non-derived content.

We end with an argument sketch for (2): A full argument for a claim of historical necessity of cognition for non-metabolic cognition must explore the structure of the priority hierarchy of complexity — the possible construction steps for the formation of a type of system. What does it mean that a cognitive system is non-metabolic? A system is metabolic if in sufficiently stable conditions the system can maintain an operational closure without explicit use of cognition, even though cognition may assist the operation of the system — finding food, avoiding predators, etc. A non-metabolic system, in contrast, can maintain its integrity without operational closure, even though it is possible to regard the system together with its cognitive capacity as, in a sense, operationally closed — the system may order spare parts from eBay and repair itself. In other words, the system must rely for its integrity on rigid constraints. The question then becomes: What can produce a rigid system that is sufficiently flexible in its interaction with the environment, i.e. having sufficiently high complexity, that is a part of a cognitive system?

The only reasonable possibility that does not rely clearly on a design mechanism is if the system is produced by a process of emergent self-organization, similar to the process of ontogeny. Two possibilities exist: the participants in the process are simple, non-metabolic components; or the participants are alive. The subtle organization of the non-metabolic components and the boundary conditions needed to systematically organize the formation of a rigid cognitive agent makes it reasonable to assume that such a system will exist only if it is externally organized by a complex, i.e. cognitive, system. This is the weakest link in the argument. It relies on the intuitive implausibility of the organization

¹¹A theory without the infamous implicit "... and here a miracle happens".

of such a process. It resembles Paley’s argument from design. In itself the argument from design is not a bad argument (it is a form of abduction); Paley’s intuition that a watch found in the desert is evidence for intelligence (or cognition) was warranted. The problem with Paley’s argument was that systems like human beings differ from watches in that a construction procedure for them exists that does not require design. Thus, while there are systems that may serve as evidence for design, there need not be an ultimate designer. Note something interesting about watches: while human beings are orders of magnitudes more complex than watches, they are much more resilient. The organization of a watch is too sensitive. A small deviation from the structure makes the watch non-functional. Watches don’t exhibit graceful degradation. For this reason, probably, it is difficult to imagine a messy natural self-organizing process that produces a watch, or worse, a rigid cognitive system. I don’t know how to make this argument more precise, however; although probably it could be made more precise.

How about if the participants are metabolic systems? There are precedents for this phenomenon; in fact, every complex cognitive system is produced by a process of ontogeny, out of small single cell metabolic systems. Could we imagine a process where the cells organize, secrete a hard substance and then die, leaving a complex, rigid cognitive system behind — a nanotechnological production process might operate in this way. Should we expect that such a system may appear without an external “design”? There is one major difference between this case and the case of ontogeny — multicellular organisms are not merely a product of the simple metabolic system, they affect the organization of the simple systems, providing feedback for the modification of the structure of simple systems. This is the connection that exist between phenotype and genotype.¹² Phenotypical characteristics have the capacity to affect the process

¹²When talking about genotype we are moving beyond the concept of autopoiesis, and

of modification of the genotype, and therefore to modulate the adjacent possible. There is, therefore, a mechanism that exist with the priority hierarchy of complexity that modulates the development of structures that self-organize into full blown cognitive agents, like people. This feedback is naturally lacking in the case where the product of the emergent process is a rigid system, which form the point of view of the hierarchy of complexity is mostly an epiphenomenal system.¹³ This makes it also reasonable to conclude that the existence of such a process is highly unlikely without an external intervention and control — without a feedback mechanism channeled through a cognitive system. Even if it turns out that my imagination is too limited, and there could be natural mechanism where simple living communities of organisms can “secrete” robots, the main claim of this paper, that life is historically structurally necessary for cognition will still hold. It is only the secondary claim that all non-metabolic cognition must be the result of metabolic cognition, i.e. a cognitive artifact, must go. Until such burdensome process is discovered or conceived within the hierarchy complexity, however, I think that it is reasonable to believe that non-metabolic cognition can exist only as an artifact of cognition, and therefore, life is structurally historically necessary for cognition.

begin talking about mechanisms of replication and recombination. It is plausible to assume that mechanism for replication and recombination are required for life to move higher in the hierarchy of complexity. In fact, I think that the following two claims are also true (although I will not argue for them here): (1) Autopoiesis is *historically structurally necessary for* living systems with mechanism for replication and recombination. And, (2) living systems with mechanisms for replication and recombination are *historically structurally necessary for* cognition. Further intermediate stages may also be reasonable.

¹³Some cases of emergent rigid systems do exist, termite mounds are an example, but such systems are orders of magnitude simpler than a cognitive system would need to be.

References

- [1] Frederick Adams and Kenneth Aizawa. *The Bounds of Cognition*. Wiley-Blackwell, 2008.
- [2] Peter Atkins. *The Second Law*. W H Freeman and Co, 1986.
- [3] M. H. Bickhard. Representational content in humans and machines. *Journal of Experimental and Theoretical Artificial Intelligence*, 5:285–333, 1993.
- [4] M. Bitbol and P. L. Luisi. Autopoiesis with or without cognition: defining life at its edge. *J. R. Soc. Interface*, 1:99–107, 2004.
- [5] Paul Bourguine and John Stewart. Autopoiesis and cognition. *Artificial Life*, 10:327–345, 2004.
- [6] J. D. Collier and C. A. Hooker. Complexly organised dynamical systems. *Open Systems & Information Dynamics*, 6:241–302, 1999.
- [7] Ezequiel A. Di Paolo. Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4:429–452, 2005.
- [8] H. Haken. *Information and Self-Organization: A Macroscopic Approach to Complex Systems*. Springer, 2nd. edition, 2000.
- [9] Hans Jonas. *The Phenomenon of Life: Toward a Philosophical Biology*. Univ of Chicago Press (1982), 1966.
- [10] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, 1993.
- [11] Stuart A. Kauffman. *Investigations*. Oxford University Press, 2002.
- [12] P.L. Luisi. Autopoiesis: a review and a reappraisal. *Naturwissenschaften*, 90(2):49–59, 2003.

- [13] Humberto R. Maturana and Francisco J. Varela. *Autopoiesis and Cognition: The Realization of the Living*. Springer, 1980.
- [14] R. Rosen. Organisms as causal systems which are not mechanisms: An essay into the nature of complexity. In R. Rosen, editor, *Theoretical Biology and Complexity*, pages 165–203. Academic Press, 1985.
- [15] R. H. Thomason. Combinations of tense and modality. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic II: Extensions of Classical Logic*, volume II. D. Reidel Publishing Company, 1984.
- [16] Francisco Varela. *El fenomeno de la vida*. Santiago de Chile: Dolmen Ediciones, 2000.