John Collier

Emergence in Dynamical Systems

Analiza i Egzystencja 24, 17-42

2013

„Analiza i Egzystencja" 24 (2013)
ISSN 1734-9923

JOHN COLLIER[*]

EMERGENCE IN DYNAMICAL SYSTEMS

Abstract

Emergence is a term used in many contexts in current science; it has become fashionable. It has a traditional usage in philosophy that started in 1875 and was expanded by J. S. Mill (earlier, under a different term) and C. D. Broad. It is this form of emergence that I am concerned with here. I distinguish it from uses like 'computational emergence,' which can be reduced to combinations of program steps, or its application to merely surprising new features that appear in complex combinations of parts. I will be concerned specifically with ontological emergence that has the logical properties required by Mill and Broad (though there might be some quibbling about the details of their views). I restrict myself to dynamical systems that are embodied in processes. Everything that we can interact with through sensation or action is either dynamical or can be understood in dynamical terms, so this covers all comprehensible forms of emergence in the strong (nonreducible) sense I use. I will give general dynamical conditions that underlie the logical conditions traditionally assigned to emergence in nature.

[*] John Collier—Philosophy, University of KwaZulu-Natal Durban, South Africa. John Collier works in the Foundations of Systems Theory, Information, Evolution and Development, and Pragmatics. He also dabbles in metaphysics, epistemology and naturalistic ethics. He has worked on five continents. Fortunately, he likes to travel. He has visited Poland for meetings almost every year since 2001, and loves the country. E-mail collierj@ukzn.ac.za.

The advantage of this is that, though we cannot test logical conditions directly, we can test dynamical conditions. This gives us an empirical and realistic form of emergence, contrary those who say it is a matter of perspective.

Keywords: emergence, dynamical systems, reduction, cohesion, predictability

Introduction

"Emergence" has become somewhat of a buzzword in the past few decades. Inevitably it has been used in ways that do not reflect its intellectual roots. Before I go on to explain the traditional philosophical notion of emergence, I want to make clear what that notion is not. It is necessary to do this because some authors have rejected the very idea of emergence based on misuses of the word that often draw for their apparent significance on the philosophical notion. For example, we might talk of something new and unexpected as emergent, such as "the emergence of the Internet" or of "the emergence of a new scientific discipline", or even of a new political party. Characteristic of this sort of case is that something unexpected but integrated evolves that is importantly different from what has come before. What is not implied is that this novelty could not have been in principle predicted, nor that it cannot be fully understood by understanding what it emerged from together with the processes by which it emerged. Sometimes this sort of emergence is called "weak emergence" (Bedau 1997; Collier and Muller 1998; Chalmers 2006), though characterizations of emergence, and weak emergence in particular, are usually done in terms of a relation between levels at the same time (synchronic), whereas the sort of usage I gave examples of above is typically the result of a process and across time (diachronic). Although I don't think this difference is metaphysically significant for the traditional idea of emergence, it can matter in practice. In particular, it creates potential problems for determining when exactly emergence arises. Later I will argue that fully understanding and detecting strong emergence, which unlike weak emergence is not reducible in principle,[1] is enhanced by a process-oriented, diachronic view. One consequence is that emergence is

[1] Chalmers (2006) defines weak emergence so that strong emergence may be a variety of weak emergence. I think this is confusing, and prefer to define them so that they are mutually exclusive. See (Collier and Muller 1998 and Bedau 1997) for more details.

seen as a process, and emerging can reasonably be seen as something that happens by degrees at a fine scale, though at a larger scale it appears sudden.

One example of a group that focuses on weak emergence (though they often give the impression of dealing with something more) is the New England Complex Systems Institute (Bar-Yam 2011). They say such things as:

> In describing collective behaviors, emergence refers to how collective properties arise from the properties of parts, how behavior at a larger **scale** arises from the detailed structure, behavior and relationships at a finer scale. For example, cells that make up a muscle display the emergent property of working together to produce the muscle's overall structure and movement. A water molecule has emergent properties that arise out of the properties of oxygen and hydrogen atoms. Many water molecules together form river flows and ocean waves. Trees, other plants and animals form a forest.
>
> When we think about emergence we are, in our mind's eye, moving among views at different scales. We see the trees and the forest at the same time, in order to see how the trees and the forest are related to each other. We might consider particularly those details of the trees that are important in giving rise to the behavior of the forest.
>
> In conventional views the observer considers either the trees or the forest. Those who consider the trees consider the details to be essential and do not see the patterns that arise when considering trees in the context of the forest. Those who consider the forest do not see the details. When one can shift back and forth between seeing the trees and the forest one also sees which aspects of the trees are relevant to the description of the forest. Understanding this relationship in general is the study of emergence.

They also refer to function as relevant to emergence, but as in interaction of something with its environment:

> Consider a key. A description of a key's structure is not enough to show us that it can open a door. To know whether the key can open a door, we need descriptions of both the structure of the key and the structure of the lock. However, we can tell someone that the function of the key is to unlock the door without providing a detailed description of either.

Their restriction to weak emergence comes out clearly in their non-standard use of "reductionist":

> The perspective that considers emergence is often contrasted with a **reductionist** perspective, which thinks about parts in isolation. Reductionism is the often vilified "anti-complex systems" view of the world. The concept of a system is itself based upon a limited form of reductionism that distinguishes the system from its environment, and the parts of a system from each other. The key difference is that the non-reductionist approach considers the relationships among them.

I belabour this because NECSI claims to deal with complex systems in general, which might lead one to deal with strongly emergent systems, but their focus and definitions make this dubious at best. Their notion of reduction is not the one used in traditional philosophical approaches to emergence, as I will explain in the next section. Furthermore, there is more than a hint of the idea that emergence depends on the viewpoint or stance of an observer.[2]

Bedau (1997) claims that weak emergence was developed in the context of complexity theory (also Chalmers 2006). I think this is right, with an important caveat. Something was needed to explain the breakdown and reappearance of order in chaotic processes like the "period doubling route to chaos."[3] These are often surprising, but need not be unpredictable in principle, although they can be computationally intractable to model due to combinatoric explosions requiring more computational capacity than we can provide. Bedau also points out that what is sometimes called "emergent computation" (e.g., Forrest 1991) at best models strong emergence, and is weakly emergent (examples are the "gliders" in cellular automata). Similar confusion exists over notions like complexity itself and self-organization. It turns out that there are two notions of complexity in the complexity literature, and also two distinct notions of self-organization. Some authors, like Robert Rosen (1991), connect emergence with notions of computability; indeed, I do myself (Collier 2008a). It turns out, furthermore, there are two notions of computability that must be distinguished (Collier 2012a). I will return to these points at the end of this section.

---

[2] Ernst Nagel (1961, p. 367) argues that emergence is relative to a theory, based on his positivist principles. Basically, this makes emergence a language-dependent property. The observer-dependence of emergence (and complexity) is also advanced by Allen and Hoekstra (1991), among others, though their text seems to me to often presuppose the opposite.

[3] Contrary to some opinion, the state of onset of "chaos" via the period doubling route is completely analysable by classical mathematical techniques (Smith 1999).

Some emergentists invoke notions of separate substances, causal independence and teleology that border on obscurantism. They believe that emergent properties must "arise from", but not be causally dependent on, underlying or prior properties. Whether or not this position is coherent, it is certainly mysterious, and evidence in its favour is lacking. Nagel (1961, p. 377) points out that although emergence is sometimes associated with radical indeterminism and/or teleological causation, this association is not essential. Let us assume that his usage, which follows C. D. Broad (1925), is authoritative. I will therefore assume that emergent entities and properties supervene on the level that they emerge from; that is, given the lower-level properties in total, at most one set of supervenient properties is possible. Kim (1978) bases the principle of supervenience on a general metaphysical position that the world is determined by its physical structure, whereas Kincaid (1987) suggests that the principle is empirically based, given that we have no testable reasons to believe that there is anything nonphysical. I believe that the metaphysical and empirical reasons are each sufficient independently, but combined they are stronger than either alone. Each answers certain doubts otherwise left open by the other. If emergence, weak or strong, entails radical indeterminism, the principle of supervenience rules it out. Whereas weak emergence is too weak, the separate substance view is too strong.

For now it is important that, while weak emergence is interesting, it is not nearly as interesting as strong emergence, nor as controversial.[4] Furthermore, it is strong emergence that corresponds to the traditional philosophical notion of Mill and Broad.[5] A final point is that though there is nothing especially dynamically unusual about weak emergence, strong emergence is very special dynamically and violates longstanding assumptions about dynamical systems that have held up until very recently and are still assumed in much of science, let alone amongst philosophers, though the assumptions were questioned as early as 18th Century criticisms of Laplace's work (Collier 2012b). I will make the violation of these assumptions clear

[4] Chalmers (2006) thinks that only consciousness is strongly emergent; I will argue that this is false. Chalmers seems to me to flirt with the two-substance view, or at least two realms of some sort. There are also emergent views about naturalized religion that seem to me to flirt with the two-substance view. Examples are Stuart Kauffman (1995) and a number of the articles in Davies and Gregersen (2010).

[5] Chalmers agrees with this.

below. Clarifying this will require making the distinctions between weak and strong emergence, two notions of complexity and two notions of computation as clear as possible.

The philosophical notion of emergence

The term 'emergent' was introduced in its modern philosophical version by G. H. Lewes in 1875 (Blitz 1992), though the idea arguably appeared as early as Aristotle. Lewes said that the emergent is incommensurable with its components and cannot be reduced to their sum or their difference. This notion is basically the same as Bill Wimsatt's notion of non-aggregativity (Wimsatt 1995; 2007, pp. 174–177). A system is aggregative just in case its properties are determined fully by some sum of the properties of its parts. J. S. Mill had earlier developed a very similar notion in his *System of Logic* (1843, Book III, Ch. 6), according to which a living body cannot be understood as a mere summing up of the separate actions of its components. He accepted that basic physical laws were not violated, but what we would now call strongly-emergent laws could impose further restrictions.

The next major step was taken by C. D. Broad (1925), who expanded in some detail on the idea of in-principle irreducibility, in the sense of non-derivability of emergent dynamics from any composition of the lower level dynamics (compare with the weak version of non-reduction proposed by the NECSI group). Broad's notion of emergence makes any emergent property unpredictable from its basis dynamics (Broad 1925, p. 61). Broad, like his predecessors, also held that emergent properties are novel: they are not properties of their underlying basis, except when *fused* (Humphreys 1997; Wong 2006) to produce the emergent properties.[6]

Strong emergence constitutes a cluster of properties that are tied together by a logical notion of non-derivability and its counterpart, non-reducibility. Corning (2002) lists the properties, following Goldstein (1999) as follows:

---

[6] See Collier (2008a) for an explanation of how fusion defeats Kim's (2005) mistaken view that if supervenience holds, the supervenient system has no casual power. Humphreys views fusion as replacing the underlying properties, whereas my cohesion (Collier and Muller 1998; Collier 2003) retains the lower level properties but restricts their range. The difference is not large. Both, I think, explain Campbell's (1974) rather mysterious "downward causation" (shudder quotes in the original title).

> (1) radical novelty (features not previously observed in systems); (2) coherence or correlation (integrated wholes that maintain themselves over some period of time); (3) A global or macro "level" (i.e., there is some property of "wholeness"); (4) it is the product of a dynamical process (it evolves); and (5) it is "ostensive" (it can be perceived). For good measure, Goldstein throws in supervenience—downward causation.

Strangely, they leave out the non-reducibility/non-derivability/non-aggregativity requirement, which leaves open the possibility of weak emergence. From talks and my reading of both Corning's and Goldstein's other works, I am pretty sure that they intended to capture something both stronger and objective. The non-reducibility requirement is crucial to the traditional view, and from it Goldstein's' other properties can be derived without ambiguity (see Collier and Muller 1998; Collier 2008a). They also don't include the traditional property of emergent systems being non-predictable. As we shall see, this is tied up with non-reducibility. Instead, they have the slippery condition of "radical novelty" and tie it to observation, suggesting that they have not moved beyond weak emergence. Broad's conditions are four: *Unpredictability*—The higher level cannot be predicted in principle; *Non-reducibility*—The whole is logically more than the sum of its parts; *Holistic*—The system cannot be decomposed into its parts without loss; and *Novelty*—The tricky one because there is no clear definition, but not merely surprise, implying a new kind of property. I will argue below that the first three of Broad's conditions have the same dynamical source, and then I will argue that this implies a specific sort of novelty.

The problem with accounts of strong emergence in terms of non-reducibility, non-predictability, holism and novelty is that these are logical conditions that cannot be observed directly, because they cannot be interacted with directly. For this reason, I advocate a dynamical account of emergence from which the logical conditions for emergence can be derived. The dynamical conditions can be tested for directly, confirming the system has the required properties for emergence. I will give these conditions after a couple of brief excursions through related ideas.

A Note on Systems

A system in the sense of Systems Theory is a coordinated set of elements or components that combine through their relationships to form a unity. Systems are distinguished from each other either by their parts or their components (including qualities and especially relations), or both. For any dynamically embodied system, the relations are causal interactions (or at least grounded in causal interactions). The causal interactions can go in both directions between elements, allowing for positive and negative feedback, and non-linearity in general. These relations are constituted of forces and flows among nodes, making them dynamical in a physical sense. Dynamical relations among combinations of three or more components that cannot be reduced to pairwise dynamical relations are not excluded. If the relations are stable, then their totality can be called the structure of the system. We can also speak of the boundary conditions of a system, which are a set of constraints on its behaviour, and the laws of a system, which refer to regularities in its behaviour and possible behaviours (usually allowing for a wide range of possible boundary conditions). Note that system laws can be peculiar to a particular system and are not the same as natural laws, though they may be instantiations of natural laws. If the nodes change (are added, subtracted or merge), the system is much more complex and the system structure is

	Equation:	Algebraic	Ordinary Differential	Partial Differential
Linear Equations	One Parameter	Trivial	Easy	Difficult
	Several Parameters	Easy	Difficult	Intractable
	Many Parameters	Intractable	Intractable	Impossible
Nonlinear Equations	One Parameter	Very Difficult	Very Difficult	Impossible
	Several Parameters	Very Difficult	Impossible	Impossible
	Many Parameters	Impossible	Impossible	Impossible

itself dynamical. Without going into details, we can say that the constraints on the system are not static and the constraints and system laws cannot be separated. I will go into more detail below, as this is a necessary condition for dynamical emergence.

'Dynamical system' is a mathematical concept that describes how a system evolves from one state to another. Technically, it is smooth mapping of either the reals or integers of a manifold (state space) onto itself. The mathematical concept is the accepted way of describing systems, at least as an ideal to be approximated, and as a regulating principle in any case for descriptions of natural systems. This ideal can be easy to satisfy in some cases, difficult in others, and cannot be obtained in practice for many systems. For other mathematical kinds of systems, no general solutions are possible. The full set of possibilities is given in the table above (after Bertalanffy 1968, p. 20). The impossible cases have no analytical solution in principle. This is where we should look for strong emergence.

Computability and Predictability

A system can be predicted across time if and only if its trajectory can be calculated from its initial and boundary conditions specified within some region of its phase (state) space, together with its equations of motion, to be within some region of phase space at some arbitrary later time. Specifically, the trajectory of a system is predictable if and only if there is a region η constraining the initial conditions at t_0 such that the equations of motion ensure that the trajectory of the system passes within some region ε at some time t_1, where the region η is chosen to satisfy ε. Indeterministic systems have probabilistic predictability. Predictability applies in principle to all closed Hamiltonian systems (specifically, conservative of energy and holonomic; i.e., roughly, that the parameters of the system are a function of its energy and positions only), including those without exact analytical solutions, such as the three-body case.[7] The systems without exact analytical solutions can be numerically calculated in principle for any finite time, if we

[7] Laplace was able to show that the orbits of the major bodies of the Solar System were stable for at least 100 million years, no mean accomplishment for a many-bodied system (Collier 2012b; also Gillespie 1997).

have a large enough computer. We might call this stepwise computability. All computations are stepwise computable, but some computations do not terminate. These computations, however, are stepwise computable and allow, in principle—the required computer might have to be larger than the known universe—the arbitrarily exact computation of a finite later state. The macrostate of a microsystem can be predicted similarly by composing the trajectories of the microcomponents and averaging to get the expected macrovalues. This makes higher levels synchronically predictable. Stochastic (chance) systems are predictable within the bounds of their probabilities in general, and introduce no special problems: higher levels can again be predicted statistically as long as all of the low level laws are holonomic.

To undermine predictability, at least one of the assumptions must go. The assumptions are: 1) the system is closed, 2) the system is a traditional Hamiltonian system with holonomic or at least near-holonomic laws, and 3) there exist sufficient computational resources. The third condition is a shorthand way of saying that the information in all properties of the system can be computed from some set of boundary conditions and physical laws, where information is understood as an objective measure of asymmetry as in (Muller 2007, and less rigorously in Collier 1996). I will demonstrate later that all three of these assumptions are violated for some simple physical systems, including some in the solar system.[8] There are some systems that violate (3) because no computer could, even in principle, have sufficient power, let alone one connected to the system under study, yet the systems evolve to more ordered states. I will give an example below in terms of Newtonian body mechanics with gravity and friction that serves as an exemplar of this sort of case. More complex cases may be more nuanced, but there are systems that go beyond just unpredictability and the formation of ordered properties.

Novelty does not necessarily follow from mathematical unpredictability, since there may be no new properties formed in unpredictable systems, but novelty is impossible with unpredictability, except in the trivial sense that a pile of blocks is novel with respect to the block components scattered

---

[8] I have argued that causation can be understood as computation (Collier 1999; 2012a). I resolve the problem of non-computable processes by making the distinction between Turing computability and stepwise computability. Only the former allows information entailments to be calculated. It is quite possible for information in one state to entail the information in another without the relation being Turing computable.

about in a child's toy box. The predictability of the macrostate of a system from its microstate (states of the components of its substrate) is just the condition of reducibility, so unpredictability is also required for and sufficient for irreducibility. The advantage of the mathematical rendition of the characteristics of emergence is that we have reduced three to one, except for possible additional requirements to ensure novel properties. Now the problem is to determine when computability fails in dynamical systems. That is when emergence begins.

Complexity and Organization

Complexity can mean complicated, that is, having many factors and many components, but it also is used in complexity theory to refer to organized complexity. Collier and Hooker (1999) note that if complexity in the complicated sense and organization are put on orthogonal axes, there are four quadrants of the following kinds:

- Type I: Simple with low organization
- Type II: Complex with low organization
- Type III: Simple with high organization
- Type IV: Complex with high organization

The first type would be typified by single-particle, conservative and decomposable (linearizable) multi-particle systems. The second would be exemplified by statistically-specified systems at or near equilibrium (e.g., gases and fluids). The third type would be sufficiently well-constrained but non-linearizable multi-particle systems; e.g., many machines and some electromagnetic systems. All three of these types of systems have analytic solutions for their dynamics, or convergent higher order additive approximations to these. This makes them tractable in the sense of predictability in the previous section. Type IV systems, however, are not tractable and can produce new organization through time. Known living systems fit into this category; so do weather systems, stream eddies and solar systems. Unlike systems of types I and III, whose organization, if any, is fully determined by their initial internal and boundary conditions, some type IV systems can produce new organization through time. And, unlike type II systems, whose organization, if any, is entirely imposed by initial and boundary conditions (think of a personal computer running a program), type IV systems contribute internally to maintaining their organization and, where it increases, to increasing it. I will explain how this works in more detail below; my main point here is to clearly distinguish complexly organized systems from merely complex systems. The latter can at best be weakly emergent; I will argue that type IV systems are typically emergent. They achieve this through dynamics that modify their own boundary conditions. This is self-organization.

I said earlier that two types of self-organization need to be distinguished. The first type is widespread and results from characteristics of components that allow them to form patterns as energy in the system is dissipated. Collier and Hooker (1999) call these self-reorganizing systems. They give the unambiguous example of a coin-sorting machine:

> Coins are often sorted by being placed in a sorting box, above a series of graduated meshes ordered by mesh size (biggest size on top) and chosen so that each mesh size is intermediate between the corresponding coin sizes. The sorter is randomly jiggled horizontally and the coins are automatically sorted by size as each eventually falls to its appropriate mesh, the meshes acting as passive filters. The coins have acquired a small increase in organization (von Foerster's order-from-noise principle, with gravity the ordering principle) but the coin+sorter system has simply re-organized.

Note that in this device the extra energy of the coins must be dissipated, otherwise they would just continue to bounce around. All self-re-organization involves energy dissipation and leads to a local minimum of energy. Self-re-organization is often called self-assembly.

Consider hydrophobic-hydrophilic (ambipathic) molecules, such as fatty acids, that self-assemble to form a membrane through rearrangement of their hydrophobic and hydrophilic ends. In water, the hydrophobic ends tend to bunch together, whereas the hydrophilic ends tend to stick towards water. This leads to a double molecule layer that naturally forms into a nearly spherical vesicle. It has been proposed that such vesicles could allow amino acids and other molecules to reach a concentration inside that allows (with very large numbers of simultaneous "experiments") eventually life-forming chemistry within the relatively protected regime inside the vesicle. Although the exact process of vesicle formation is not fully understood, dissipation is surely at work, or the molecules would just bounce back rather than bond. The vesicle forms a lower-energy state than other arrangements.

In both these cases of self-re-organization, there is no new information created in the system, despite the appearance of what we deem order. If anything, information is lost as dissipation occurs. There is no compensating information produced. The order in the different sizes of the coins and the interaction potentials of water with fatty acids is implicit, and the dissipation process eliminates what is essentially noise, which is analogous to friction and the production of heat.

It used to be thought that the Moon always facing one side to the Earth was the result of another example of such an energy minimizing process in which there is a single final state determined by the properties of the components. Laplace was able to show that the 1-1 ratio of the Moon's rotation to its revolution around the Earth was stable. If the Moon were to move slightly faster, the gravity of the Earth would pull it back and similarly if it were to go slightly slower: the lowest energy condition is the 1:1 ratio. This didn't explain how the Moon got into the 1-1 ratio, however. George Darwin, one of Charles' sons, was able to explain the approach to the ratio through the dissipation of tidal torque that is exerted if the ratio is either greater or less than 1-1. Laplace also explained the complex resonances of the four largest moons of Jupiter by the same dynamics. He answered objections that his theory did not fit the known observations by developing probability theory to show that it was more probable that his theory was right than that the measurements were right. He was right, and the addition of dissipation led to a theory that was accepted for a long time. The assumption was that dissipation leads to resonances that are at the lowest energy level, essentially the same mechanism as self-assembly.

It turns out these assumptions were wrong. This was discovered when it was discovered in the 1960s that Mercury turns on its axis three times for each two times that it goes around the Sun (Collier 2012b). Without going into great detail, the state space of the Mercury-Sun system has a number of stable basins (or attractors) that, while not the lowest energy situations, are lower energy than any place nearby in space, so that if the system gets close to one of these basins, dissipation is more likely to result in capture in that basin. The boundaries between the basins are almost certainly fractal, meaning that for any two points in one attractor basin there is a point between them in another basin. In fact, there are many basins in the system corresponding to ever higher resonances. The chances of a 1-1 capture are about ½. For a 3-2 ratio the odds are about ⅓. The rest of the probabilities are taken up by higher order resonances. What does this mean for Darwin's explanation of the 1-1 ratio of the Moon's orbit? Basically, it means that it is only a partial explanation: It explains why the 1-1 ratio is possible, and even likely, but it does not explain why the ratio is 1-1 rather than some higher resonance.[9]

In the planetary resonance cases, unlike in self-re-organization, there is arguably new information produced in the system as one or another of the possible resonances is selected. One might ask why this is not true in the self-assembly case. The answer is that the parts (coins or molecules) are interchangeable, so although there are many ways to get to the final result, they are all equivalent to each other.[10] Noise dissipation creates a more ordered state, but the way it happens does not matter, since the noise (shaking in the case of the coin-sorting machine and random molecular translations, and vibrations in the case of the molecules) is random. In the planetary resonance case, however, minor differences that are effectively random are enough to put the system in one attractor rather than another. The result is not fully predictable (or retrodictable), even in principle. Unlike the unsolvable three-body case that can be predicted in principle for any finite time and for a specific margin of error, the result in this case is not predictable in finite time. The reason is that the dissipation allows the whole unsolvable process to be carried out in finite time, which never happens in systems without dissipation.

So, are planetary resonances emergent? They certainly are unpredictable, at least in their details, and these details are significantly different. It might be argued, however, that they fail the condition of novelty. After all, we have no radically new properties: we start with orbits and rotations, and we end with orbits and rotations, albeit synchronized. I have argued that while such systems are not reducible, they can be explained, insofar as they can be explained, by reductive explanation using our understanding of the underlying dynamics (gravity and dissipation through tidal torques). If there

[9] Critics in his own time were right in saying that he had not given a full explanation of how the resonances came about. Such an explanation is in principle impossible.

[10] But it is worth noting that the location of formed fatty acid vesicles might be intrinsically unpredictable because their nuclei of condensation might depend on chaotic fluctuations in the densities of their distribution in water.

is a novel property, it is synchronization. This is certainly not radically novel like life or mind, but perhaps it is a beginning.

It is worth noting that despite their unpredictability, planetary resonances are diagnosable in the sense that once they are established, we can treat them as type I or III systems. I use "diagnosable" technically to mean that given the system laws – gravity and tidal dissipation, in this case—states of the system can be calculated to an arbitrarily high degree of accuracy. The resonant state is diagnosable, but its ancestral states are not fully diagnosable. In particular, planetary resonances do not appear to be complexly organized. In fact, they do not seem to be complex at all. Furthermore, there is no clear sense in which there is anything that we might call levels or wholes that emerge out of some underlying basis. So on a scale of emergence they are not weakly emergent, because they are not predictable, but they are not strongly emergent either, because they do not form wholes, and aggregativity is not relevant because they do not even seem to aggregate (or not to aggregate) anything. I think, however, that despite their simplicity, they can be worked into the same category as other, more paradigmatic type IV systems.

There are fairly simple systems that do clearly have the properties of non-aggregativity and of forming wholes. Self-organizing systems, after Prigogine, are dissipative systems that use some of their dissipated energy to form order within themselves, where this order is at a larger scale than any order that existed before. Examples are eddies and cyclones, but classic examples are Bénard cells and slime molds. In all of these cases, small fluctuations are promoted to larger scale order at what is clearly a higher level, one that contains and constrains motions of the parts at a lower level. The higher level properties that form cannot be predicted knowing only the lower level properties. These lower level properties constrain what can happen at the higher level, but the higher level also constrains what happens at the lower level. Because of these higher-level constraints, possibilities become available that would not be available without them.

It is worthwhile looking in some detail at the Bénard cell system. This is a fluid in a container with a lid (to avoid surface effects) heated gently from below. As the heat differential from the bottom to the top is increased, there is a relatively sudden change from conduction of heat through the fluid to convection. It is possible to calculate the transition point from the properties of the fluid by comparing the equations of motion of the molecules of the fluid and their pairwise (local) interactions with the equations of motion of the fluid in the convecting state. The two are equal at the transition point. The convecting state cannot be computed (in principle) from the equations of motion for the molecules (though interesting computational models can be made). We have to know that there is a convecting state in order to compare the two states and derive the transition point.[11]

There is no known way to derive a convecting state from the microscopic movements of molecules or local fluid dynamics, but we can give an intuitive explanation. It is very likely impossible due to (a) the impossibility of solving the equations of motion and (b) likely chaos at the molecular level. Nonetheless, we can get a pretty good understanding of what is going on. There are fluctuations in density in the system. These fluctuations are larger the higher the temperature. There is also viscosity that creates cohesion among the molecules of the fluid (or we may think of the fluid as intrinsically cohesive). As the temperature gradient is increased, there are larger regions of greater and lesser density. These regions are either buoyant or the opposite, respectively. The viscosity of the fluid holds these regions together against their tendency to disperse thermodynamically. As the regions grow larger, this tendency overcomes the dispersive tendency and the buoyant regions float upwards, while the denser regions sink. Because of the close constraints on the experimental conditions, regular cells form. The transition is thermodynamically stable, because it minimizes entropy production, basically following the path of least resistance. Viscosity, buoyancy and gravity make convection possible. Thermodynamic stability maintains it.

It is tempting to call Bénard cells emergent except for one thing, and that is that they are not unexpected. From our past knowledge of fluids we expect to see whorls and convection cells, so it is not surprising to find convection in Bénard cells. In fact, they were designed and studied just because we expected them to convect. Unlike planetary resonances, there is only one final state possible, given the design of the system. On the other hand, we are concerned with ontological emergence, not merely the surprisingly or unexpectedly new. Perhaps Bénard cells have the right dynamical properties to be emergent by sharing the required properties with other emergent systems. There is a clear sense in which Bénard cells are like

---

[11] For mathematical and intuitive details in more depth, see (Collier and Banerjee 2000).

type IV systems: They are organized at a higher level, and their organization is deep in the sense that computing their surface structure from the dynamics of their components is at best highly non-trivial. It is worth noting at this point that my description of the Mercury-Sun resonance was over-simplified. In fact, it is a very weakly dissipative system. Gravitational influences from other planets cause perturbations from the perfect 3-2 resonance; however, tidal dissipation brings the resonance back (the perturbation times are so slow compared to the relaxation back to resonance, that Mercury basically maintains the resonance continuously, but the principle of continuous dissipation to maintain the resonant state holds). All resonances are subject to external perturbations.

Dynamical Emergence

As I have said, self-reorganizing systems are not emergent in any but a weak sense, since they are reducible, and any novelty, such as it is, is predictable in principle (at least statistically). Self-organizing systems are different. Yet the steady state of convection in Bénard cells is diagnosable (the equations of motion for convection are solvable), as are the steady states of planetary resonances. The main differences are that Bénard cells have only one final state, whereas planetary resonances have multiple possible final states, and that Bénard cell convection cannot be computed from only the molecular dynamics, whereas resonant motion can. In one case, we have diachronic predictability, but not form the lower level; in the other case, exactly the opposite. On the other hand, in both cases we have truly novel properties appearing, albeit relatively simple. This leads to the question of what is the same across the two sorts of cases. Although neither case seems to be like paradigmatic cases of emergence, like life or mind, perhaps a common property is essential to emergent systems.

Hamiltonian systems as they are commonly understood have an overall force function that is holonomic; i.e., dependent only on the position coordinates and time if and only if the force is conservative, an example being particles in a gravitational field. It is possible that component forces are not conservative, but their combination must be. Energy being constant is also holonomic, as it depends only trivially on position coordinates and time. In general, if a system is holonomic, it can do no virtual work, because all virtual displacements are perpendicular to the forces of the constraints, so there is (or would be) no force on them. This is really just another way of saying that the Hamiltonian of holonomic systems depends only on (appropriately chosen) generalized coordinates and energy. A result of this, as I indicated in the section above on predictability, is that holonomic systems are predictable, at least by numerical methods.[12] This is of central importance to the theory of dynamical emergence, as I will argue below.

The Bénard cell and resonance cases, being essentially dissipative in their dynamics, do not have constant system energy, since some leaves the system through dissipation. They are thus non-holonomic and cannot be dealt with by standard methods. This is true of both Bénard cells and the origin of the evolution of planetary resonances, simple as they are. For this reason, I include them with type IV systems, since they are more like them than any of the other three types of system. The characteristic of type IV systems, then, is not so much the amount of organization, but its irreducibly larger scale.

An important characteristic of non-holonomic systems is that their equations of motion cannot be separated from their boundary conditions. Conrad and Matsuno (1990) make clear the consequences for dynamical systems:

> Differential equations provide the major means of describing the dynamics of physical systems in both quantum and classical mechanics. The indubitable success of this scheme suggests, on the surface, that in principle it could be extended to a universal program covering all of nature. The problem is that the essence of a differential equation description is a separation of itself from the boundary conditions, which are regarded as arbitrary.

Conrad and Matsuno draw conclusions for the whole universe (they claim the method breaks down, but it must be compatible with "no boundary conditions" constraints on cosmological theories). More significant for present purposes is the breakdown of the separation of differential equations and boundary conditions in non-integrable systems, exactly the ones that are non-holonomic (in which constraints like boundary conditions cannot

---

[12] The argument is given in a more technical form in (Collier 2008a). See also Hooker (2012).

be separated from their dynamics). In these systems, computation from partwise interactions fails, and the system is in a sense holistic. In any case, its dynamics cannot be reduced to the dynamics and partwise relations. This is one of the conditions for emergence. Non-integrability also implies some sort of unpredictability as I have defined it. Note that inseparability applies in both the Bénard cell and planetary resonance cases. The origin of the cells in the first case cannot be understood in terms of boundary conditions (gravity, heat differential) alone; the formation of the cells creates new boundary conditions on the cells (a macro constraint) that cannot be computed from relations of the parts (molecules) alone. In the planetary resonance case, the formation of the resonance changes the boundary conditions by adding a new constraint (resonance). This is the common feature.

It is generally recognized that standard Hamiltonian systems are mechanical (mechanistic). This idea is summed up by their holonomic character, in an engineer's sense that all their constraints can be expressed algebraically and are basically geometric. Non-holonomic systems, on the other hand, must have a constraint that is expressed as a rate of change, so their form cannot be integrated into an algebraic form and they cannot be understood geometrically. Some examples of non-holonomic systems are a rolling wheel (friction matters) or whirlwinds. A wheel can undergo a sudden change of state when it goes into a skid; a whirlwind dies if it loses its sustaining energy differential. These indicate two central features of emergent systems: relatively sudden changes of the state space, and the necessity of dissipation in understanding the dynamics. These two characteristics also hold of the examples of planetary resonance (resonance arises rapidly in astronomical time) and Bénard cells.

Some systems are non-Hamiltonian, but are nearly Hamiltonian. We can deal with such systems with approximations. This is a common method, using a version of perturbation theory. Other non-Hamiltonian systems step rapidly from one state to another (rapidity is relative here). The dynamics of these systems can be analyzed by comparing the micro- and known macro-mechanics, along with knowledge of the transition, and by treating the transformation as a step function. This is how Bénard cells were in fact analyzed. However, there is a large range of systems that are not close to step functions or close to smooth Hamiltonian systems. I conjecture that this sort of radically non-Hamiltonian behavior underlies all emergence. In particular:

1. The system must be non-holonomic, implying that the system is non-integrable (this ensures non-reducibility).
2. The system is energetically (and/or informationally) open (boundary conditions are dynamic).
3. The system has multiple attractors (the Bénard cell has two).
4. The characteristic rate of at least one property of the system is of the same order as the rate of the non-holonomic constraint (radically non-Hamiltonian).
5. If at least one of the properties is an essential property of the system, the system is essentially non-reducible; it is thus an emergent system.

I don't claim that these conditions are independent; in fact I think they are not. I choose them because they are relatively easy to argue for in specific dynamical cases, and from that to emergence. I do claim, however, that the conditions are necessary and sufficient dynamical conditions for emergence. All are required for the emergence of systems, and all but the last for emergence of properties. If any condition above is violated (perhaps implying the violation of others), there is no emergence. To show necessity is fairly trivial. If condition 1 is violated, the system is at least numerically computable and hence predictable. If condition 2 is violated, the boundary conditions are fixed rather than dynamic, so they are holonomic. If condition 3 is violated, we can predict a single attractor, as we can for classic complex systems like the Lorenz attractor. If condition 4 is violated, then the system can be treated as approximately Hamiltonian, and it is predictable. If condition 5 is violated, there is no emergent property, perhaps just a chaotic system. Since none of these conditions are specific to the examples I have given, they apply to all cases. Together, I claim, the conditions are sufficient. So I have given necessary and sufficient (but probably not independent) dynamical conditions for non-reducibility and unpredictability.

There might be worries at this point. Typical life and mind have been taken to be emergent. These seem to be novel in a way that my examples are not. Am I missing important properties of emergence? Living systems are functional in a way that nonliving systems are not. Mind has the additional property of consciousness. Although I have argued that Bénard cells and planetary resonance have novel properties (convection and resonance, respectively), perhaps these are not novel enough. They certainly do not seem to be on the same scale. Novelty is a tricky issue with dynamical emergence, since all of the causes are driven in some sense at the lower level. Given that conditions 1-5 are satisfied, the new property is not a sum of the properties of the components, either. There is something genuinely novel. In my own work, I have focused on cases in which the emergent entity is a system, rather than a system property, and I have called the fusion of the dynamical unity property of the system cohesion (Collier and Muller 1998; Collier and Hooker 1999). Both Bénard cells and resonances require cohesion for their central emergent property. Novelty, rather than being hard to get, is rather easy to achieve, and many systems have emergent properties that define them dynamically. This might be reflected in Broad's view that water is emergent from its components (whether or not he was right about this).

The centrality of work

One of the things that is typical of living systems is that they do work on themselves in order to maintain themselves. The major currency for the cell, and consequently for known living systems, is ATP. It allows otherwise thermodynamically impossible chemical processes to occur (typically producing more organized states). Organization in living systems ensures that these processes contribute to the maintenance and growth of the system. Basically, an organism does work on itself in order to maintain the conditions for its continued existence. It does this by producing the boundary conditions required. This is a typical emergent process in which the dynamics of the system interact in both ways with the boundary conditions. I have argued elsewhere (Collier 2004; 2011) that the organization that contributes to the continued existence of the organism, which I call autonomy, is the dynamical identity of the organism. Anything that contributes to autonomy is biologically functional. This approach guarantees the "forward-looking" character of function by incorporating anticipation (Collier 2008b). A full explanation would be too long, but we can ask the question of whether functionality is emergent. Given the non-equilibrium and dissipative quality of life (required for it to do work on itself), life is surely emergent in my sense. Unfortunately, I do not know at this point if functionality itself is emergent in some further sense. I suspect so, since I see it to be a necessary property of autonomy, which emerges from the work done to maintain organization. However, if autonomy itself is a non-dynamic (static or steady state) network, then its dynamics are fully diagnosable as outlined above. But if autonomy is dynamic, with new nodes forming and old ones disappearing, then it is a good candidate for emergence. I think the evidence favours the latter, but it would take me to far afield to recite all the evidence.

The issue of consciousness is much more difficult, especially the notorious "hard problem" (Chalmers 1996). Terrence Deacon (2011), in a masterful work, makes some headway on this problem, but does not claim to have solved it. Obviously the brain itself is supported by dissipative processes and uses a very large percentage of the body's energy consumption by weight. Presumably there is a lot of work done to maintain the nervous system, and this is especially important for maintaining thought. This does not show, however, that thought requires dissipation, let alone consciousness. If thought is strictly computational, as has been widely held,[13] then it is reducible to its underlying processes (like all computation). Deacon argues that this leaves open questions about meaning and intention that are either presupposed, explaining nothing, or dismissed, explaining nothing. I will have to leave the issue here as far as analysis goes, but I do have a few observations.

First, as with function, a steady state model of thought implies reducibility (of thought to its components), ruling out emergence. This position is implied by Fodor's Language of Thought Hypothesis (Fodor 1975), as all the basic nodes are innate. As with function, the emergence in thought requires the formation and/or elimination of nodes. This is an empirical matter for which there is currently only ambiguous evidence. I have argued (Collier 2001) that if thought is not in equilibrium with its external conditions, an unexpected stimulus can cause it to reorganize around the new stimulus, accommodating it and thereafter being prepared for it to some degree. This would represent emergence in thought, though it must be said that the dynamical closure involves both the thinker and the environment, and there is a clear sense in which that whole is the locus of the emergence. Further work requires empirical studies.

---

[13] The arguable view goes back to the Cartesians, like de La Mettrie, but re-emerged in the 20th Century with the invention of digital computers. It held sway for some time, but wilted under attacks concerning its capacity to explain meaning, something that Deacon has taken up. S. Edelman (2008) has revived the approach.

Conclusions

Emergence has two forms, weak and strong. Although the weak version is interesting, it is the strong one that requires revision of many current views of the nature of things. The central properties of strong emergence, irreducibility, unpredictability, holism and novelty stem from the same source in computability. This source can be explained in dynamical terms. Strong emergence turns out not to be rare, but fairly common from physics on up, but clear examples even from biological function require more work, and even more so for psychological, mental and social sciences.

References

Allen, T. F. H. and Thonas, W. H. 1991. *Toward a Unified Ecology*. New York: Columbia University Press.

Bar-Yam, Y. 2011. "Concepts: Emergence." New England Complex Systems Institute: http://www.necsi.edu/guide/concepts/emergence.html (accessed 2012-12-05).

Bedau, M. A. 1997. "Weak Emergence." In: J. Tomberlin (ed.). *Philosophical Perspectives: Mind, Causation and World*. Malden, MA: Blackwell: 375–399.

Bertalanffy, L. von. 1968. *General System Theory: Foundations, Development, Applications*. New York: George Braziller.

Blitz, D. 1992. *Emergent Evolution: Qualitative Novelty and the Levels of Reality*. Dordrecht: Kluwer Academic.

Broad, C. D. 1925. *The Mind and its Place in Nature*. Routledge and Kegan.

Campbell, D. T. 1974. "'Downward Causation' in Hierarchically Organized Biological Systems." In: F. J. Ayala and T. Dobzhansky (eds.). *Studies in the Philosophy of Biology*. New York: Macmillan.

Chalmers, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, D. J. 2006. "Strong and Weak Emergence." In: P. Clayton and P. Davies (eds.). *The Re-emergence of Emergence: The Emergentist Hypothesis from Science to Religion.* Oxford: Oxford University Press: 244–254. Originally published 2002 at http://consc.net/papers/emergence.pdf.

Collier, J. 1996. "Information Originates in Symmetry Breaking." *Symmetry: Culture & Science* 7: 247–256.

_______ 1999. “Causation is the Transfer of Information.” In: H. Sankey (ed.). *Causation and Laws of Nature*. Dordrecht: Kluwer: 279–331.

_______ 2001. “Dealing with the Unexpected.” *Partial Proceedings of CASYS 2000: Fourth International Conference on Computing Anticipatory Systems. International Journal of Computing Anticipatory Systems* 10: 21–30. Leuven: CHAOS.

_______ 2003. “Hierarchical Dynamical Information Systems with a Focus on Biology.” *Entropy* 5: 57–78.

_______ 2004. “Interactively Open Autonomy Unifies Two Approaches to Tunction.” In: *Computing Anticipatory Systems: CASY ’03—Sixth International Conference*. Ed. by D. M. Dubois. American Institute of Physics Melville, New York. AIP Conference Proceedings 718: 228–235.

_______ 2008a. “A Dynamical Account of Emergence.” *Cybernetics and Human Knowing* 15, 3–4: 75–86.

_______ 2008b. “Simulating Autonomous Anticipation: The Importance of Dubois’ Conjecture.” *Biosystems* 91: 346–354.

_______ 2011. “Explaining Biological Functionality: Is Control Theory Enough?” *South African Journal of Philosophy* 30, 1: 53–62.

_______ 2012a. “Information, Causation and Computation.” In: G. Dodig Crnkovic and M. Burgin (eds.). *Information and Computation: Essays on Scientific and Philosophical Understanding of Foundations of Information and Computation*. Singapore: World Scientific.

_______ 2012b. “Holism and Emergence: Dynamical Complexity Defeats Laplace’s Demon.” *South African Journal of Philosophy* 30: 229–243.

Collier, J. and Hooker, C.A. 1999. “Complexly Organized Dynamical Systems.” *Open Systems and Information Dynamics* 6: 241–302.

Collier, J. and Banerjee, M. 2000. “Bénard Cells: A Model Dissipative System.” Unpublished presentation at Emergence, Complexity, Hierarchy, Organization conference. Espoo, Finland, August 2000. http://web.ncf.ca/collier/papers/BC-ECHOIV.pdf.

Collier, J. and Muller, S. 1998. “The Dynamical Basis of Emergence in Natural Hierarchies.” In: G. Farre and T. Oksala (eds.). *Emergence, Complexity, Hierarchy and Organization, Selected and Edited Papers from the ECHO III Conference. Acta Polytechnica Scandinavica, MA91*. Espoo: Finish Academy of Technology.

Conrad, M. and Matsuno, K. 1990. “The Boundary Condition Paradox: A Limit to the Universality of Differential Equations.” *Applied Mathematics and Computation* 37: 67–74.

Davies, P. and Gregersen, N. H. (eds.). 2010. *Information and the Nature of Reality: From Physics to Metaphysics*. Cambridge: Cambridge University Press.

Deacon, T. W. 2011. *Incomplete Nature: How Mind Emerged from Matter.* New York: W.W. Norton & Company.

Edelman, S. 2008. *Computing the Mind: How the Mind Really Works*. Oxford: Oxford University Press.

Fodor, J. A. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Forrest, S. 1991. *Emergent Computation: Self-organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks*. Cambridge, MA: MIT Press.

Gillespie, C. C. with collaboration of Fox, R. and Grattan-Guiness, I. 1997. *Pierre Simon Laplace, 1794–1827: A Life in Exact Science*. Princeton: Princeton University Press.

Hooker, C. A. 2012. “On the Import of Constraints in Complex Dynamical Systems.” *Foundations of Science*. DOI 10.1007/s10699-012-9304-9.

Humphreys, P. 1997. “How Properties Emerge.” *Philosophy of Science* 64: 1–17.

Kauffman, S. 1995. *At Home in the Universe*. New York: Oxford University Press.

Kim, J. 1978. “Supervenience and Nomological Incommensurables.” *American Philosophical Quarterly* 15: 149–156.

_______ 2005. *Physicalism or Something Near Enough*. Princeton: Princeton University Press.

Kincaid, H. 1987. “Supervenience doesn’t Entail Reducibility.” *Southern Journal of Philosophy* 25: 343–356.

Mill, J. S. 1843. *System of Logic*. London: Longmans, Green, Reader, and Dyer. [8th ed., 1872].

Muller, S. J. 2007. *Asymmetry: The Foundation of Information*. Berlin: Springer-Verlag.

Nagel, E. 1961. *The Structure of Science*. New York: Harcourt, Brace and World.

Rosen, R. 1991. *Life Itself: A Comprehensive Inquiry into The Nature, Origin, and Fabrication of Life*. New York: Columbia University Press.

Smith, J. D. H. 1999. “Wreath Products along the Perioddoubling Route to Chaos.” *Ergodic Theory and Dynamical Systems* 19: 1617–1636.

Wimsatt, W. 1995. “The Ontology of Complex Systems.” *Canadian Journal of Philosophy* 20: 564–590.

________ 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.

Wong, Hong Yu. 2006. “Emergents from Fusion.” *Philosophy of Science* 73: 345–367.