Synthese (2012) 185:171–186
DOI 10.1007/s11229-010-9719-1

Emergence: logical, functional and dynamical

Sandra D. Mitchell

Received: 21 December 2009 / Accepted: 2 February 2010 / Published online: 4 March 2010

Abstract Philosophical accounts of emergence have been explicated in terms of logical relationships between statements (derivation) or static properties (function and realization). Jaegwon Kim is a modern proponent. A property is emergent if it is not explainable by (or reducible to) the properties of lower level components. This approach, I will argue, is unable to make sense of the kinds of emergence that are widespread in scientific explanations of complex systems. The standard philosophical notion of emergence posits the wrong dichotomies, confuses compositional physicalism with explanatory physicalism, and is unable to represent the type of dynamic processes (self-organizing feedback) that both generate emergent properties and express downward causation.

Keywords Emergence · Downward causation · Reduction · Self-organization · Chaos · Feedback

1 Introduction

The concept of "emergence" has a long history. Aristotle has been attributed with saying: "The whole is more than the sum of its parts,"[1] from his discussion of part-whole causation in the Metaphysics (Annas 1976). Since then, philosophers have worried

[1] A better translation is "the whole is over and above its parts, and not just the sum of them all."

Some of the arguments in this paper are also developed in my 2008 *Komplexitaten: Warum wir erst anfangen die welt zu Verstehen*, Suhrkamp Verlag and a revised English version of that book, *Unsimple Truths: Science, complexity and Policy*, University of Chicago Press, 2009.

S. D. Mitchell (✉)
Department of History and Philosophy of Science, University of Pittsburgh,
1017 Cathedral of Learning, Pittsburgh, PA 15260, USA
e-mail: smitchel@pitt.edu

much about what "is more than" means and what "sum" means. Some have advocated expanding our metaphysical horizons by positing wholly different and unique substances above and beyond the physical to explain biological development (e.g., Driesch 1914; see also Bechtel and Richardson 1998; McLaughlin 2003). In modern times, appeal to entelechies or *élan vital* are taken to be antithetical to physicalism, the claim that there is one substance of which all entities in the world are composed. Yet there are other interpretations of emergence that do not posit additional substances. Such views agree with the substance monism assumed by contemporary scientific explanations, and are candidate explanations for the increase in appeal to emergence in contemporary sciences.

I will argue here that the reductionist presumption that all compositionally complex structures and systems can be explained, without remainder, by appeal only to the properties of their simplest components cannot be sustained. In the nineteenth century British philosophers defined emergence as that which is strictly non-reducible and *therefore* non-explainable by means of the laws governing its component parts. On their account, espoused by Mill (1843), Broad (1925) and others, emergence is identified with the epistemic marks of non-explainability and non-reducibility (see McLaughlin 1992). For Mill, qualitative features of water (fluidity, wetness, turbulence) were deemed emergent because there were no explanations of them in terms of the constituent molecules of oxygen and hydrogen. The atomic elements could not be said to be fluid or wet, or to exhibit turbulence. Similarly, emergent biological properties were taken to be inaccessible to explanations by chemical properties, psychological properties inaccessible to biological explanation and so forth. The nineteenth century "emergentists" equated explanation with reduction. For them, properties that could not be explained, by the properties and interactions of constituent parts, were emergent. Higher-level properties could not be inferred, but only observed, and regularities in and among emergent properties would be built from observations at the higher level, not from laws governing the properties of their component, lower-level parts. On this view, if putatively emergent properties could be explained by a reductive strategy they would cease to be considered emergent.

In fact, twentieth-century science succeeded in providing what are arguably successful reductive explanations of what thinkers like Mill and Broad identified as emergent properties. With the advent of both quantum mechanical accounts of chemical bonding[2] and the explanations of biological phenomena of inheritance by the biochemistry of DNA (see Schaffner 1993), the properties of chemical compounds, like water, and many features of living organisms have arguably been explained in terms of the properties of their constituent parts.[3] The standard examples of emergence appealed to in the nineteenth and early twentieth century were rendered non-emergent. As a result, there was very little use of the term "emergence" in science from the 1920s until the

[2] Dirac, one of the founders of quantum mechanics, stated "The underlying laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that exact applications of these laws lead to equations which are too complicated to be soluble" (1929, p. 714).

[3] For a recent argument that chemistry is not even approximately reduced to physics see Scerri (1994).

1960s.[4] Bertrand Russell claimed that emergent qualities were merely epiphenomena and of no scientific significance, saying that analysis "enables us to arrive at a structure such that the properties of the complex can be inferred from those of the parts" (1927, pp. 285–286). Talk of emergence died out.

The revival in the 1970s of the use of the term "emergence" in scientific literature coincided with renewed interest in chemical and neurological complexity (see Sperry 1969, 1991; Campbell 1974). It became widespread with the rise of what is now known as complexity science (Prigogine 1997; Bak 1996; Amaral and Ottino 2004; Lewin 1992). In response to the increased scientific interest in emergence, philosophers have developed a number of new accounts of what it is to be emergent. Most of these have focused on the emergence of consciousness, and the reducibility or non-reducibility of mental states like beliefs and desires to neurological, physical properties of the brain (Beckermann et al. 1992; Blitz 1992; Anderson et al. 2000; Clayton and Davies 2006; Humphreys 1997; O'Connor 1994; O'Connor and Wong 2005; Rueger 2000; Silberstein and McGeever 1999; Stephan 1997). The key features of emergence for both philosophical treatments and scientific applications are novelty, unpredictability and the causal efficacy of emergent properties or structures, sometimes referred to as downward causation.[5]

A puzzle arises in light of the two tracks—one in philosophy and one in science—of the development of ideas about emergence. On the one hand, philosophical analyses have typically followed Bertrand Russell's dismissal of the explanatory, causal and thus potential scientific role of so-called emergent properties. On the other, there has been a revival of interest in emergence in science. As a crude measure of the latter, an Internet search for the combined terms "emergence," "properties," and "science" in Google Scholar, as of this writing, yields close to 1,000,000 "hits," many of which on a cursory inspection appear to be about the kinds of emergent properties that Russell deemed to be of no scientific significance. Indeed the first two entries are articles from *Science* and *Nature* (Barabási and Albert 1999; Olshausen and Field 1996). If the philosophical analyses that dismiss the reality of emergent properties are correct, then why have descriptions of emergent properties in science become so widespread? In what follows I will argue that there are both faulty assumptions and an impoverished conceptual framework that prevents the character of emergent properties referenced by science to be adequately represented in some cases of philosophical analysis. One thing both philosophers of science and scientists agree on: There is only one type of substance—physical matter—that is constituent of the entities and properties that science engages. If there are emergent properties, they have to be made of the same physical substance as non-emergent properties.

[4] For a thorough discussion, see Corning (2002).

[5] The introduction of the notion of "downward causation" marks a distinct divide between nineteenth century emergentist thought and the use of the term in modern science, since emergent properties for the nineteenth-century philosophers had no causal influence on "lower-level" phenomena that yielded to scientific explanation. Broad thought there might be trans-ordinal laws that connected lower levels with emergent levels, but did not discuss what we would now call downward causation. Sperry (1969) discussed what amounts to downward causation in terms of configurational forces, but the term "downward causation" was introduced by Campbell (1974).

2 Kim's analysis

I will consider in some detail arguments of Jaegwon Kim, a contemporary philosopher whose views on emergence have set a high standard for clarity in the philosophical community. Kim's philosophical argument against emergence instructively reveals the role of common assumptions among many who reject the attribution of emergence to the kinds of properties that scientists commonly denote as emergent. In addition to an unwarranted assumption, I will argue that there is a structural problem with Kim's strategy of argumentation.

In his 1999 article on emergence, Kim attempts to explicate what emergence could be. He suggests the following formulation: "[C]omplex systems aggregated out of these material particles begin to exhibit genuinely novel properties that are irreducible to, and neither predictable nor explainable in terms of, the properties of their constituents" (4). The questions Kim poses are clear. First, how can a higher-level property like liquidity be grounded in its physical constituents (e.g., water is H_2O) and yet be both novel and not predictable from the properties of those constituents alone? Second, even if sense is made of such an attribute, how could an emergent property not only be causally autonomous from its physical basis, but actually causally change the properties of its physical constituents? As Kim puts it:

> …[H]ow is it possible for the whole to causally affect its constituent parts on which its very existence and nature depend? If causation or determination is transitive, doesn't this ultimately imply a kind of self-causation, or self-determination – an apparent absurdity? (1999, p. 28)

Kim uses a standard example from the literature of a phenomenon that allegedly cannot be explained by reduction: pain. To be emergent is to be non-reducible. But then, to determine whether or not a property is emergent, one first must determine what is to be reduced. Kim suggests a functional strategy. He argues that a higher order property, like pain, is reducible to its physical components if one can describe the higher property functionally and then show the identity of what is referenced by that function to something at the lower, material level. A functional description of a property is in terms of its causes and consequences instead of its structure. Take a straightforward example: A chair is describable as an artifact designed to function as something suitable for humans to sit on. It is also the case that chairs are always made out of something material: wood, iron, plastic, etc. So a particular chair could also be described in terms of its material components. Chairs in general are described by their capacity to function as something humans use for sitting. Chair-function is *realized by* an individual entity's physical components and structure. Kim applies a functional decomposition to the property of "being in pain." Pain is "being in some state…caused by tissue damage and causing winces and groans" (1999, p. 13). Kim *assumes* that whatever stands in the place of that higher level property, e.g., whatever the term "pain" picks out, is also, at the same time, describable in terms of some physical constituents, namely some physical neurological state. Pain is that which is caused by tissue damage and which in turn causes winces and groans. Pain is reducible if it is *nothing more than* what is also referred to by its neurological realization.

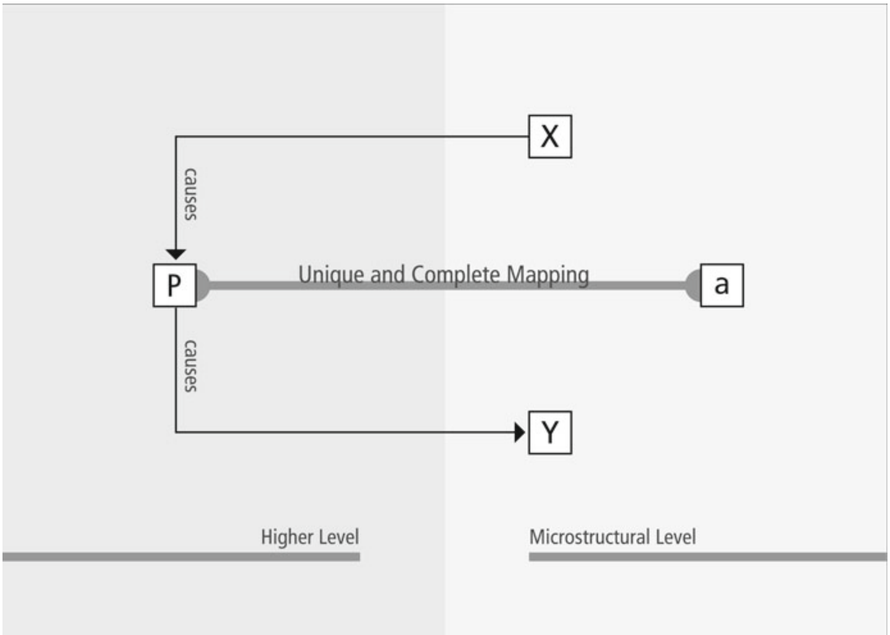

Fig. 1 Static functional mapping

Kim argues that, on this account of reduction, for any property to qualify as emergent is extremely difficult. If a higher level property is just whatever it is that is picked out functionally as what is caused by X (e.g., tissue damage) and in turn causes Y (e.g., winces and groans), then there will always be some configuration of physical components at a lower level that can be identified as realizing the functional property of being caused by X and in turn causing Y.[6] However, to reach this conclusion, Kim makes a strong assumption that "every material object has a unique complete microstructural description" (1999, p. 6). If one can map the functional property at the higher system level into the appropriate part of the *unique, complete* lower level description then any explanation can be run at both levels, every prediction could in principle be generated at both levels (whether we currently have the means to do so or not) and nothing at the higher level would be novel or unpredictable given the micro-level physical description (see Fig. 1).

What room is left for emergence on Kim's functional replacement account of reduction? Not much. To really be unexplainable, novel, and unpredictable from lower-level

[6] There is an enormous literature on reduction and multiple realizability that stems from the seminal papers of Putnam (1967) and Fodor (1974). Some (e.g., Pylyshyn 1984) take the fact that instances of "higher-level" properties, like pain, can be realized in many different kinds of lower-level properties (different neurological configurations, for example) as sufficient to argue that reduction of types of higher-level properties can never be reduced to types of lower-level properties. Others have argued that this feature is not enough to derail reductionism (see Horgan 1993; Bickle 1998). See Sober (1999) for a discussion of its merits. My argument does not address this particular thread of argumentation.

constituents, a higher-level property would have to be in some sense intrinsic, and not describable in extrinsic, physical terms of what causes it and what it in turn causes.

But *scientific* access to properties and entities is entirely by means of what causes them to occur or change and what they in turn cause to occur or change. Anything measurable, for example, could be re-described functionally in the manner prescribed by Kim. Anything that is functionally describable could then be associated with a lower-level "unique and complete" description. Thus explanations and predictions based on higher-level properties lose their claim to emergence. None of the currently scientifically-identified emergent properties (e.g., color patterns on mammals, flocking behavior of birds, division of labor in social insects, etc.) can qualify as emergent on Kim's account. The only plausible candidates for emergence that Kim acknowledges are subjective properties of consciousness, e.g., what it feels like to be in pain, not the causes and consequences of being in pain. Such subjective feelings are, as such, not within the purview of scientific study.

Kim clearly specifies conditions a property would have to meet to be genuinely novel and unpredictable from the properties of its physical constituents. For Kim,

> [i]f emergent properties exist, they are causally and hence explanatorily, inert and therefore largely useless for the purpose of causal/explanatory theories. If these considerations are correct, higher-level properties can serve as causes in downward causal relations only if they are reducible to lower-level properties. The paradox is that if they are so reducible, they are not really "higher-level" any longer. (1999, p. 33)

For Kim there is nothing independent of the lower level to be found at the higher level that is even a candidate for causing lower-level behavior. Thus on his account, nothing available for scientific study would count as emergent. It is not therefore surprising that the main interesting feature of emergent properties in contemporary science, their autonomous causal capacity, turns out also to be impossible to defend on Kim's analysis. However, such a conclusion is based on a strong assumption that, I believe, should be rejected. I reject the assumption that reducibility is a necessary consequence of there being a physical realization of higher-level causal relations and will defend an alternative view, a broader notion of emergent properties that makes sense of scientific use and is part of a more adequate naturalized epistemology.

One problem with Kim's account of reduction and emergence is that, while appearing to be merely preserving physicalism, the view that there is no new substance at the higher level that is somehow mysteriously unlike the physical substance from which all things are constructed, he actually imports into the argument a much stronger assumption, namely that there is always a unique and complete *description* of the higher-level phenomena in terms of the lower level. If we are concerned with *types* of higher-level phenomena (rather than particular instances) then his uniqueness claim is not satisfied. Consider the well-known problem of multiple realizability. Using Kim's example of pain, if we consider pain in general, then each instance of pain will be realized in some neurological microstructure, but that structure may vary radically between instances within and between individuals experiencing the pain. Headaches are a type of pain and instances of headaches can occur by very different physical realizations.

For some, the fact of multiple realizability is sufficient to render a property at the higher functional level different in kind from the perhaps open-ended disjunction of various physical configurations that realize it (see footnote 6). For others, it indicates only that there may be one actual worldly event that can be captured by different vocabularies at different levels of abstraction (Fodor 1974; Putnam 1967; Pylyshyn 1984). On this second view, epistemological reasons can be given for preferring non-reductive explanations, but ontologically there is nothing more than one event which in principle could be described at the lower level. Regardless of the conclusion one draws from multiple realizability, there is a shared assumption on how one represents the relationship between higher-level properties and lower-level properties. The common perspective takes a property at a micro-level, compares it to a property at a higher level, and then attempts to determine if there is one real property (perhaps with two descriptions) or two properties that are related by some reductive type of mapping function.

What the philosophical arguments structured statically as a mapping function miss entirely is a question at the center of much scientific concern with emergence, namely: *How* is the property at the higher level produced, and what are the differences among the many kinds of relationship between higher- and lower-level properties that occur in nature? The logical analyses of the kind Kim and others adopt are both static and at a level of abstraction inadequate to represent the dynamic and concrete realities that are the concerns of practicing scientists.

If we take a snapshot view of the higher and lower levels, then the dynamics of *how* the higher level is constituted and stabilized is lost. Contemporary sciences show us that there are processes, often involving negative and positive feedback or self-organization, that are responsible for generating higher-level stable properties, and these processes are not captured by a static mapping. Kim's attempt to clarify the philosophical conception of emergence has stripped it of any scientifically interesting features, and hence it fails to adequately engage the properties that scientists have identified as emergent, properties like division of labor in social insect colonies. Division of labor is multiply realized in different species of ants, bees and termites,[7] but may display the same dynamics of *how* the higher-level properties are generated. Furthermore, a type of downward causation is in evidence when higher-level properties initially emerge by means of self-organization then place constraints on the behavior of their constituent parts.[8]

The philosophical problem with the argument that Kim developed is the conflation of compositional physicalism (there is one kind of substance from which all things are created) with descriptive fundamentalism (there is a privileged, complete

[7] Division of labor refers to the distribution of a suite of tasks to different individuals, rather than their all being performed by each individual. For example, solitary insects reproduce offspring, build nests, feed larvae, defend their nest and forage for food. In eusocial insects these tasks are divided among members of the colony. The queen reproduces offspring and the workers, for the most part, are functionally sterile. Younger individuals tend to the queen and feed the larvae while older individuals guard and forage outside the colony.

[8] Bedau (1997) develops an account of weak emergence that is sensitive to the complicated, rather than simple, way in which emergent properties are generated from their component parts. For other criticisms of the Kim-style view of emergence and reduction see Loewer (2001) and Shoemaker (2002).

description of the world in terms of fundamental components). Why is this a problem? All descriptions are abstractions or idealizations. They do not stand in a one-to-one mapping relationship with the entirety of the undescribed world. To think that our language (or any human artifact intended for representation including mathematics or computer simulation) captures the physical world exactly is simply misconceived. Descriptions are always partial. The metaphysical claim that, at the physical level, there are unique processes that bring about physical results is inescapable, unless one is either a dualist or believes there are uncaused events. However, *representing* these processes in a language, whether that is the vocabulary and syntax of formal logic or of fundamental physics, is an entirely different matter.

It is likely that all the factors contributing to the complete cause of some physical event, say a window breaking when hit by a rock, cannot be represented by any single theory in the syntax of logic or even the language of physics. The local, contingent components of a causal process are just not included in the scope of physical theory or its abstract language (see Cartwright 1994, 2000; Cartwright et al. 1995), and these will always be part of the complete cause. It may well be that the complete causal process is enacted by *physical* entities; what else could there be? But at the same time there will not be a representation that completely captures this process in terms of *physics* entities. Thus Kim's functional strategy will fail to deliver. If there is something in the world that can be isolated by the functional description (caused by X and causing Y) there is no reason to think that a physical description of that piece of the world, partial as it is, will be identical with a higher-level description of that piece of the world, partial as it is (see also Dupré 1993). Without the requirement of a unique, complete description at the lower level, the functional identification will not go through, and, as I have argued, there are compelling reasons to reject the assumption of completeness of our representations.

Furthermore, although all properties, events or structures are physical (what else could they be?),[9] not all physical properties, events or structures are biological. Compare a conglomeration of molecules constituting a rock with an organization of molecules making up a baby monkey. What's the difference? It is not to be found by looking at what they share, namely physical composition, but only by looking at how they differ (see also Wimsatt 2007, Figure 9.1, p. 183). Some form of difference is what is being picked out by identifying some higher-level structures as emergent. There is an important, explanatory difference that cannot be captured by the emphasis on the reasonable claim that any simple or complex whole is made up of simpler, physical parts.

3 Scientific emergence

A new understanding of emergence has become widespread in scientific accounts of biological and social phenomena that does not share either the strong epistemic

[9] What is means to be physical is not always clear; see Markosian (2009). For my argument all that is required is the claim that what is physical (metaphysically) is not one-to-one identical with what is represented by any given theory of physics.

assumptions of the philosophical view Kim advocates or a non-physicalist metaphysics, i.e., which does not preclude explanation from below or above. On this understanding, emergence is identified with certain types of non-aggregative compositional structures, including self-organization (see Wimsatt 1986, 2007; Bedau 1997; Kauffman 1984, 1993, 1995; Camazine et al. 2001). Aggregation is a particularly simple kind of compositional relationship between component parts and the whole. The weight of a pile of rocks being the aggregate of the weight of each component rock is an example. However, there are a number of ways in which composition can fail to be aggregate and hence fail to be reducible.

One way is found in the complexity of chaotic systems represented by nonlinear dynamics.[10] Poincare in the early 1900s discovered dynamical instability in which a physical system could end up in wildly different end states depending on very small differences in its initial state. This sensitivity to initial conditions is often identified with chaos. Nonlinearity, on the other hand, refers to a behavior that cannot be modeled by a linear equation, i.e., it cannot be solved by treating the variables in it as the sum of independent contributions. Feedback refers to a particular way in which processes in a complex system interact, which may lead to chaotic outcomes, adaptive outcomes, equilibrium states, etc. Many natural phenomena display these various aspects of complexity.

Simple additive relations and simple linear equations, while adequate to explain some simple behaviors, will fail to make sense out of much of the complexity that we find in nature even though patterns and structures emerge from the simple interactions of the constituents. The vee pattern that emerges in a flock of geese or the more complex patterns of flocking starlings are not predictable by an aggregation of behaviors of individuals in solo flight, but only from the non-aggregative interaction or self-organizing that derives from the local rules of motion plus feedback among the individuals in group flight (see Couzin 2007; see Rosen 2007 for photos of the starling patterns). Ontologically, there are just physical birds; there is no new substance, no director at a higher level choreographing the artistic patterns the flocks make. Nevertheless, this type of behavior is emergent.

> [I]t is usually not possible to predict how the interactions among a large number of components within a system result in population-level properties. Such systems often exhibit a recursive, nonlinear relationship between the individual behavior and collective ('higher-order') properties generated by these interactions; the individual interactions create a larger-scale structure, which influences the behavior of individuals, which changes the higher-order structure, and so on. (Couzin and Krause 2003, p. 2)

What the contemporary scientific sense of emergence has in common with the nineteenth-century views is that interaction among the parts generates properties which none of the individual components possess, and these higher-order properties in turn can have causal efficacy, i.e., novelty. What is different is that for new scientific emergence there are concrete accounts of how and why rules of interaction among

[10] It is important to distinguish between nonlinearity, chaos, and feedback systems. While these can all be present in a single system at the same time, they need not be. See Bak (1996) and Sole and Goodwin (2001).

components produce difficult-to-predict emergent behaviors, i.e., explanation may be possible. Self-organization is grounded in multiple forward and reverse feedback loops, which can be radically sensitive to initial and evolving boundary conditions. This type of emergence is dynamic and thus something important is lost in modeling it using Kim's explication of emergence in terms of static property descriptions and an implicit feed-forward linear type of causation.

There are a number of further complex causal interactions that attention to emergent dynamics can reveal. If the feedback system stabilizes a property in the face of fluctuating external conditions, for example, the form of the flock remains dense with a certain percentage of individuals in the interior regions, then this stable property or structure can be a target of natural selection. If some birds organize their behavior in such a way and others do not, and if the structure emerging in the first case presents an advantage against larger bird predation, for example, then this higher-level property will exhibit causal saliency (see Carere et al. 2009). It will causally explain why one population of birds is more adapted to an environment with lots of large predators compared to another. If the individual behaviors which give rise to the higher level property are heritable, then evolutionary consequences will follow, depending of course on the trade-off of other adaptive considerations and the force of non-selective processes on the future states of the populations.

Rob Page and I (Page and Mitchell 1991, 1998) developed simulations of group living for social insects to investigate what features at the colony level may be the result of self-organization and what features might have varied in the evolutionary past and thus could be candidates for explanation by natural selection. We attributed to our computer bees minimal features, namely that individual behavior varies in response to different levels of a stimuli and that the stimuli itself is constituted and changed by the behavior of the individuals. These features are consistent with those found in solitary insects and thus, arguably, in the evolutionary ancestors of social bees. We found that a crude form of division of labor was the result of group living itself; indeed it is inescapable when random individuals interact over time with each other. Thus while further refinements of division of labor, like specific caste ratios, may be the result of natural selection in response to differing environments, once a group is formed, some form of division of labor emerges. Fewell and Page (1999, see also Fewell and Jeanson 2008) forced normally solitary ants to interact and experimentally showed the same result, i.e., the emergence of division of labor. The inference to draw is that emergence of higher-level properties from interaction among components via self-organization is not antithetical to features of that property being tuned by natural selection. Rather the two types of processes may work collaboratively to generate the forms and features that we find in evolved complex systems. Selection operating on heritable variation of the components may reductively explain some of the features of a higher-level property like division of labor. Emergence and self-organization may explain other features of the higher-level property.

Kim's aim was to give an explication of what it is to be emergent in general. The limitations I have outlined show that his analysis fails to cover typical examples of emergence in the contemporary scientific literature. To be clear, it may well be claimed that Kim's aim and my own in providing a philosophical analysis of emergence differ. That is, the constraints each of us acknowledges for an appropriate account, and the evidence each of us admits as relevant to the success or failure of an account may diverge. While one cannot fault the logic of Kim's argument, the assumptions made are subject to criticism. My criticism comes from a naturalistic perspective, i.e., that our philosophical understanding of concepts (like emergence) should track not just logical consistency, but also empirical adequacy. Thus how we come to represent nature has its limitations given by the actual processes of representation. To assume otherwise is to acknowledge only the constraints of logic on our reasoning. To assume in an argument what we might know at "the end of science," for example some posit of unique, complete descriptions of phenomena in terms of a privileged language, is to ignore the facts of the history of science and the state of current science. The constraints I choose to impose on what is an adequate account of emergence acknowledge both the naturalistic constraints and the current state of knowledge. Indeed if my analysis makes sense of contemporary scientific practice, though not being designed as a mere description of it, then that constitutes support for the analysis.

There are significant consequences when we expand our notion of causality to include the types of complex interactions that I have suggested are common for biological systems. Novelty, unpredictability and downward causation, the conditions of emergence that appeared as unattainable, are more easily met. Determinism no longer entails predictability. Even if a behavior, described at a higher level of organization, is determined by the interactions of entities at a lower level of organization, if the dynamics are nonlinear, the behavior will not be predictable. The first aspect of emergence identified epistemically with unpredictability or ontologically as novelty automatically assigns the label of emergence to some behavioral outcomes of nonlinear systems.

The nonlinearity that constitutes dynamical complexity of some systems carries with it methodological consequences for understanding such systems. Such behaviors, while deterministic, are unpredictable given their extreme sensitivity to immeasurably small variations in initial and evolving boundary conditions. Many biological systems display features of dynamical complexity including bifurcation, amplification, and a type of phase change (Nicolis and Prigogine 1989; Beninca et al. 2008). Systems exhibiting these behaviors are known variously as "complex," "chaotic," or "nonlinear" systems. A bifurcation is a period doubling, quadrupling, etc. It marks the sudden appearance of a qualitatively different solution for a nonlinear system as some parameter is varied. Amplification refers to the nonlinear dramatic increase in response to small increases in the value of some parameters (see May and Oster 1976). Two points in a system may move in vastly different trajectories, even if the difference in their initial configurations is very small. Examples include weather phenomena, fluid turbulence, and crystal growth. Edward Lorenz, a meteorologist, metaphorically termed this type of complexity "the butterfly effect" from the possibility that the flap of a butterfly's wings in Brazil might cause a tornado in Texas (Lorenz 1996). In the slime mold *Dictyostelium discoideum*, the transformation from single-cell populations to multi-cellular organism is by means of cell–cell signaling and chaotic feedback mechanisms that induce a dramatic macroscopic phase change in the system. Positive feedback can create a situation in which the cellular response changes abruptly and irreversibly once the signal magnitude crosses a critical threshold (Tyson et al. 2003; Goldbeter 1997; Goldbeter et al. 2001; Kessin 2001; Mitchell 2003).

Simulations in addition to laboratory or field experiments are needed to explore the space of outcomes that a chaotic system can visit. The differential equation models that work so well for Newtonian systems are completely inadequate for chaotic systems. Bedau (1997) identifies the need for simulation as a requirement of emergence. That a system ends up in one state rather than another will have to do with the immeasurably small differences in the initial conditions under which the very same deterministic rules are applied. It is still the case that the properties and behavior of the component parts cause the ensuing behavior of the system, but there is a shift of emphasis to the features of the history and context that the system experiences to understand why one outcome occurred rather than another. Knowing just the function that describes the causal structure of the parts and our most precise account of its initial state will not tell you in what later state the system will find itself. Unpredictability is not ruled out by physical determinism.

Unlike some well-known cases of simple physical causation, often in complex systems more than one, or a few dominant causes are responsible for the behavior we wish to explain. A billiard ball moves in the direction and with the velocity it does because of the impact on it from the cue ball. Of course there are slight perturbations in the trajectory, due to spin and friction, that skew the behavior from what it would be if only the single impact were operating, but for the most part, such behavior is explained by a single dominant cause. Not so in the world of the complex. Of course there are multiple kinds of "complexity." Some systems are closer to the billiard ball, just with a greater number of causal influences. The multiplicity of factors is not particularly problematic, especially if there are simple rules of interaction, such as additivity. The set of forces on an airplane (a more complicated version of the billiard ball) are identified as the causes of its trajectory and can be summed to predict where it is going. However, complex systems often involve feedback mechanisms resulting in amplification or damping of the results or of nonlinear chaotic behavior, and under these conditions, causal explanations by additivity will fail.

Our understanding of causation expands when we pay close attention to the way in which negative and positive feedback both stabilize phenomena at a higher level and constrain the behavior of the components at the lower level. Feedback provides an operational understanding of one type of downward causation where system level properties constrain and direct the behavior of the components. Thus a second condition for emergence is met by the complexity studied by contemporary science: There is causal influence of higher-level properties on lower-level behaviors. Feedback loops are operating in these cases. What is new is the interaction of clearly individual behaviors and emergent, higher-level properties of the system as a whole. An example from division of labor in social insects will illustrate.[11]

What causes a honey bee to forage for nectar at a given time, rather than rest comfortably in the hive? Genetic differences among individuals account for some of the differences in foraging frequency. Although the specific pathways are unknown, the correlation between genes and foraging frequency is robust and, in the absence of other interacting factors, explains differences within and between colonies. However, it is

[11] Division of labor has also been identified in mammals (Bennet and Faulkes 2000) and crustaceans (Duffy 1996).

also known that individuals change their foraging frequency as a result of environmental stimuli. In a genetically homogeneous population, variation in foraging frequency depends on the different environmental factors encountered by individuals. In ideal situations, where only one factor is active, a systematic study can reveal the strength of genes or the strength of an environmental stimulus in producing the effects. In natural settings both are operating simultaneously. The result of their joint operation is not a simple, aggregative, linear compound of their individual contributions.

Self-organized systems are ones in which feedback interactions among simple behaviors of individual components of a system produce what appears to be an organized group-level effect. Honey bees collect nectar from flowers, which is then processed by digestive enzymes and evaporation into the honey that is found in the combs in the hive. It is the major source of carbohydrates for worker bees and is part of the nutritional substance, along with pollen, that is fed to larvae. Older bees form the caste that flies out of the hive to collect nectar, pollen and water. An individual forager sucks up the nectar from flowers through her proboscis and returning to the hive, she unloads it to a younger worker, whose job is to store the nectar in an empty cell. For an individual honey bee forager the probability that she will continue to forage for nectar may be affected by how long she waits to unload what she has collected upon returning to the hive. The outcome of each individual bee experiencing waiting time collectively generates a system that "tunes" the number of foragers to the "need for nectar" in the hive (Seeley 1989).

The mechanism by which the system level property—how much nectar is stored—controls the behavior of individual foragers does not involve any mysterious forces or substances beyond the physical makeup of the bees and nectar and hive. But clearly the amount of nectar stored in a hive is not a property of any of the individual bees, although it is the sum of the results of their individual behavior. It is significant for understanding the complexity of this system to see how an adaptive structure at the higher level (nectar supply) emerges from individual behaviors *and* how the higher-level structure *causally influences* in a feedback loop adaptation at the lower level. Clearly, if there is already a lot of nectar stored, then it is more adaptive for foragers to stop collecting, and if there is not much nectar stored, foragers should go out and collect more. Whether a bee continues to forage or begins to rest is at least partially caused by the surplus or surfeit of nectar.

When a forager lands with her crop full, she must have the nectar unloaded by another worker bee and carried to an empty cell to be stored. The bee who unloads then returns to the next forager waiting to be unloaded. If many of the cells in the hive are already full, it will take the unloading bee longer to find an empty cell in which to deposit the nectar from the first bee, and correlatively if there are many empty cells it will be quicker. How long a bee waits to unload is a measure of the amount of nectar already stored, and this waiting time is a trigger for continuing to forage or stopping.

What supports the view that the amount of nectar stored is an emergent property is not just that interactions among individuals generate it. In addition the higher-level property has causal efficacy. It can be the target of selection, where colonies that store more rather than less may be better adapted in an environment of unpredictable supply of nutritional resources. The specific behaviors of the individuals are not selected—they might develop various systems for storing nectar—but rather, the amount stored is the causally relevant variable in the selection scenario. In addition, the higher-level property, though generated through the actions of the individuals, not only is a variable in the function governing the individual behavior (in addition to genetic factors and experience) but can *change* the behavioral repertoire of the individual, and even the expression pattern of the genes that contribute to individual probability to forage (see Robinson et al. 2005).

Self-organization and feedback make scientific sense of emergent features of complex systems (see Camazine et al. 2001; Bonabeau et al. 1997). Higher-level properties, the pattern of division of labor, and the amount of nectar stored in the hive are caused by the interactions of the components, but not in a simple aggregative way. Interactions are often chaotic, displaying both positive and negative feedback, which can generate novelty in the overall response which is not predictable from the intrinsic properties of the individual components. And in turn, the higher-level properties can influence the behavior of the components whose actions themselves determine the higher-level properties, as is the case in nectar foraging behavior and amount of nectar stored. The "self-causation" in a system that Kim had alluded to as an "apparent absurdity" turns out to be a common feature of the types of complex systems investigated and explained by modern science. What is required to make sense of scientific emergence is a richer conceptual framework that accommodates the dynamics by which the behavior of the component parts of a complex system generate a higher-level property, as well as how that higher-level property in turn can causally influence the very components from which it has been generated. Not all higher-level properties are emergent, but some are. Taking account of nonlinear dynamics and feedback causal processes in addition to static and linear representations, which may be adequate for simpler domains of nature, is one of the ingredients of understanding complexity. Emergence identifies an important class of phenomena, whose analysis permits a reframing of some of the standard approaches to understanding scientific explanation.

References

Amaral, L. A. N., & Ottino, J. M. (2004). Complex networks: Augmenting the framework for the study of complex systems. *European Physical Journal B, 38*, 147–162.

Anderson, P., Emmeche, C., Finnemann, N., & Christiansen, P. (Eds.). (2000). *Downward causation: Minds, bodies, and matter*. Aarhus: Aarhus University Press.

Annas, J. (1976). *Aristotle: Metaphysics books M and N*. Oxford: Clarendon Press.

Bak, P. (1996). *How nature works: The science of self-organized criticality*. New York: Copernicus, USA

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*(5439), 509–512.

Bechtel, W., & Richardson, R. C. (1998). Vitalism. In E. Craig (Ed.), *Routledge encyclopedia of philosophy* (pp. 639–643). London: Routledge.

Beckermann, A., Flohr, H., & Kimeds, J. (1992). *Emergence or reduction?*. Berlin: Walter de Gruyter.

Bedau, M. (1997). Weak emergence. In J. Tomberlin (Ed.), *Philosophical perspectives: Mind, causation, and world* (pp. 375–399). Oxford: Blackwell.

Beninca, E., Huisman, J., Heerkloss, R., Johnk, K. D., Branco, P., & Van Nes, E. H., et al. (2008). Chaos in a long-term experiment with a plankton community. *Nature, 451*, 822–825.

Bennet, N. C., & Faulkes, C. G. (2000). *African mole-rats ecology and eusociality*. Cambridge: University of Cambridge Press.

Bickle, J. (1998). *Psychoneural reduction: The new wave*. Cambridge, MA: MIT Press.

Blitz, D. (1992). *Emergent evolution: Qualitative novelty and the levels of reality*. Dordrecht: Kluwer Academic Publishers.

Bonabeau, E., Theraulaz, G., Deneubourg, J., Aron, S., & Camazine, S. (1997). Self-organization in social insects. *TREE, 12*(5), 188–193.

Broad, C. D. (1925). *The mind and its place in nature* (1st ed.). London: Routledge & Kegan Paul.

Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2001). *Self-organization in biological systems*. Princeton: Princeton University Press.

Campbell, D. (1974). Downward causation in hierarchically organised biological systems. In F. J. Ayala & T. Dobzhansky (Eds.), *Studies in the philosophy of biology: Reduction and related problems* (pp. 179–186). London/Bastingstoke: Macmillan.

Carere, C., Montanino, S., Moreschini, F., Zoratto, F., Chiarotti, F., & Santucci, D., et al. (2009). Aerial flocking patterns of wintering starlings, *Sturnus vulgaris*, under different predation risk. *Animal Behaviour, 77*, 101–107.

Cartwright, N. C. (1994). Fundamentalism vs. the patchwork of laws. *Proceedings of the Aristotelian Society, XCIV*, 279–292.

Cartwright, N. C. (2000). *The dappled world: A study of the boundaries of science*. Cambridge: Cambridge University Press.

Cartwright, N. D., Cat, J., Fleck, L., & Uebel, T. E. (1995). *Philosophy between science and politics*. (Cambridge: Cambridge University Press, Ideas in Context Series).

Clayton, P., & Davies, P. (Eds.). (2006). *The re-emergence of emergence*. Oxford: Oxford University Press.

Corning, P. (2002). The re-emergence of emergence. *Complexity, 7*(6), 18–30.

Couzin, I. D. (2007). Collective minds. *Nature, 445*, 4.

Couzin, I. D., & Krause, J. (2003). Self-organization and collective behavior in vertebrates. *Advances in the Study of Behavior, 33*, 1–75.

Dirac, P. A. M. (1929). Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London A, 123*, 714–733.

Driesch, H. (1914). *The history and theory of vitalism*. London: Macmillan.

Duffy, J. E. (1996). Eusociality in a coral-reef shrimp. *Nature, 381*, 512–514.

Dupré, J. (1993). *The disorder of things*. Cambridge, MA: Harvard University Press.

Fewell, H., & Jeanson, R. (2008). Influence of the social context on division of labor in ant foundress associations. *Behavioral Ecology, 19*, 567–574.

Fewell, J. H., & Page, R. E., Jr. (1999). The emergence of division of labour in forced associations of normally solitary ant queens. *Evolutionary Ecology Research, 1*(5), 537–548.

Fodor, J. (1974). Special sciences: Or the disunity of science as a working hypothesis. *Synthese, 28*, 97–115.

Goldbeter, A. (1997). *Biochemical oscillations and cellular rhythms: The molecular bases of periodic and chaotic behaviour*. Cambridge: Cambridge University Press.

Goldbeter, A., Gonze, D., Houart, G., Leloup, J. C., Halloy, J., & Dupont, G. (2001). From simple to complex oscillatory behavior in metabolic and genetic control network. *Chaos, 11*, 247–260.

Horgan, T. (1993). From supervenience to superdupervenience: Meeting the demands of a material world. *Mind, 102*, 555–586.

Humphreys, P. (1997). How properties emerge. *Philosophy of Science, 64*, 1–17.

Kauffman, S. A. (1984). Emergent properties in random complex automata. *Physical, 10D*, 145–156.

Kauffman, S. A. (1993). *The origins of order*. Oxford: Oxford University Press.

Kauffman, S. A. (1995). *At home in the universe: The search for laws of self-organization and complexity*. New York: Oxford University Press.

Kessin, R. H. (2001). *Dictyostelium: Evolution, cell biology, and the development of multicellularity*. Cambridge: Cambridge University Press.

Kim, J. (1999). Making sense of emergence. *Philosophical Studies, 95*, 3–36.

Lewin, R. (1992). *Complexity: Life at the edge of chaos*. Chicago: Chicago University Press.

Loewer, B. (2001). Review of J. Kim, mind in a physical world. *Journal of Philosophy, 2001*, 315–324.

Lorenz, E. N. (1996). *The essences of chaos*. Seattle, WA: University of Washington Press.

Markosian, N. (2009). Physical object. In E. Sosa, J. Kim, & G. Rosenkrantz (Eds.), *A companion to metaphysics* (2nd ed., pp. 486–489). Malden, MA: Blackwell Publishing.

May, R. M., & Oster, G. (1976). Bifurcations and dynamic complexity in simple ecological models. *American Naturalist, 110*, 573.

McLaughlin, B. (1992). The rise and fall of British emergentism. In A. Beckermann, J. Kim, & H. Flohr (Eds.), *Emergence or reduction?* (pp. 49–93). Berlin: De Gruyter.
McLaughlin, B. (2003). Vitalism and emergence. In T. Baldwin (Ed.), *The Cambridge history of philosophy 1870–1954* (pp. 631–640). Cambridge: Cambridge University Press.
Mill, J. S. (1843). *A system of logic*. London: Longmans, Green & Co.
Mitchell, S. D. (2003). *Biological complexity and integrative pluralism*. Cambridge: Cambridge University Press.
Mitchell, S. D. (2008). *Komplexitaten: Warum wir erst anfangen die welt zu Verstehen*. Frankfurt: Suhrkamp Verlag.
Mitchell, S. D. (2009). *Unsimple truths: Science, complexity and policy*. Chicago, IL: University of Chicago Press.
Nicolis, G., & Prigogine, I. (1989). *Exploring complexity: An introduction*. New York: W. H. Freeman.
O'Connor, T. (1994). Emergent properties. *American Philosophical Quarterly, 31*, 91–104.
O'Connor, T., & Wong, H. Y. (2005). The metaphysics of emergence. *Noûs, 39*, 658–678.
Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*, 607–609.
Page, R. E., & Mitchell, S. D. (1991). Self organization and adaptation in insect societies. *Proceedings of the Philosophy of Science Association, 2*, 289–298.
Page, R. E., Jr., & Mitchell, S. D. (1998). Self organization and the evolution of division of labor. *Apidologie, 29*, 101–120.
Prigogine, I. (1997). *The end of certainty*. New York: The Free Press.
Putnam, H. (1967). Psychological predicates. In W. H. Capitan & D. D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). Pittsburgh: University of Pittsburgh Press.
Pylyshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.
Robinson, G. E., Grozinger, C. M., & Whitfield, C. W. (2005). Sociogenomics: Social life in molecular terms. *Nature Reviews Genetics, 6*, 257–270.
Rosen, J. (2007, April 22). Flight patterns. *New York Times* (for photos of the starling patterns).
Rueger, A. (2000). Physical emergence, diachronic and synchronic. *Synthese, 124*(3), 297–322.
Russell, B. (1927). *The analysis of matter*. London: Allen and Unwin.
Scerri, E. (1994). Has chemistry been at least approximately reduced to quantum mechanics? In D. Hull, M. Forbes, & R. Burian (Eds.), *PSA: Proceedings of the biennial meeting of the philosophy of science association* (pp. 160–170). East Lansing, MI: Philosphy of Science Association.
Schaffner, K. F. (1993). *Discovery and explanation in biology and medicine*. Chicago, IL: University of Chicago Press.
Seeley, T. (1989). Social foraging in honey bees: How nectar foragers assess their colony's nutritional status. *Behavioral Ecology and Sociobiology, 24*, 181–199.
Shoemaker, S. (2002). Kim on emergence. *Philosophical Studies, 108*, 53–63.
Silberstein, M., & McGeever, J. (1999). The search for ontological emergence. *The Philosophical Quarterly, 49*, 182–200.
Sober, E. (1999). The multiple realizability argument against reductionism. *Philosophy of Science, 66*, 542–564.
Sole, R., & Goodwin, B. (2001). *Signs of life: How complexity pervades biology*. New York: Basic Books.
Sperry, R. W. (1969). A modified concept of consciousness. *Psychological Review, 76*, 532–536.
Sperry, R. W. (1991). In defense of mentalism and emergent interaction. *The Journal of Mind and Behavior, 122*, 221–246.
Stephan, A. (1997). Armchair arguments against emergentism. *Erkenntnis, 46*, 305–314.
Tyson, J. J., Chen, K. C., & Novak, B. (2003). Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology, 15*, 221–231.
Wimsatt, W. C. (1986). Forms of aggregativity. In A. Donagan, A. N. Perovich, Jr., & M. V. Wedin (Eds.), *Human nature and natural knowledge* (pp. 259–291). Boston: Reidel Publishing Company.
Wimsatt, W. C. (2007). *Re-engineering philosophy of limited beings: Piecewise approximations to reality*. Cambridge, MA: Harvard University Press.