Phenomenal Emergence

Achim Stephan

achim.stephan@uos.de
Institute of Cognitive Science University of Osnabrück (Germany)

1 Introduction

A few years ago Dave Chalmers started a book on consciousness with the following words:

> Consciousness is the biggest mystery. It may be the largest outstanding obstacle in our quest for a scientific understanding of the universe. The science of physics is not yet complete, but it is well understood; the science of biology has removed many ancient mysteries surrounding the nature of life. There are gaps in our understanding of these fields, but they do not seem intractable. We have a sense of what a solution to these problems might look like; we just need to get the details right.
>
> Even in the science of the mind, much progress has been made. Recent work in cognitive science and neuroscience is leading us to a better understanding of human behavior and of the processes that drive it. …
>
> Consciousness, however, is as perplexing as it ever was. It still seems utterly mysterious that the causation of behavior should be accompanied by a subjective inner life. We have good reason to believe that consciousness arises from physical systems such as brains, but we have little idea how it arises, or why it exists at all. How could a physical system such as a brain also be an *experiencer*? Why should there be *something it is like* to be such a system? Present-day scientific theories hardly touch the really difficult questions about consciousness. We do not just lack a detailed theory; we are entirely in the dark about how consciousness fits into the natural order … (1996, xi).

Among cognitive scientists and philosophers, two opposite reactions are common to deal with Dave Chalmers' diagnosis: optimists claim that in the end even the problem of phenomenal consciousness will find an adequate scientific explanation. Science has always succeeded in finding solutions and it will do for consciousness, too, or so they say. In contrast, the pessimistic reaction acknowledges a failure in principle – we will never succeed in finding a reductive explanation of consciousness. In the following, I will concentrate on the pessimistic view. This view interprets consciousness (and phenomenal qualities) as *emergent* phenomena; particularly, as phenomena that are *irreducible* and therefore emergent in a very strong sense.

Since it is highly controversial by what criteria emergent properties should be distinguished from non-emergent properties, I examine several notions of emergence having different strength. In particular, I will distinguish weak, diachronic and synchronic theories of emergence. These distinctions help to clarify the type of emergence that should be attributed to phenomenal consciousness in contrast, for example,

http://www.dif.unige.it/epi/networks

***Networks* 3-4**: 91-102, 2004
© SWIF - ISSN 1126-4780
http://www.swif.uniba.it/lei/ai/networks/

to the type of emergence that could be attributed to some properties of connectionist networks or robots. In addition, we will see what consequences go along with what notion of emergence. Thus, it will become clear that some weaker versions of emergence are compatible with property reductionism, while the stronger versions are not.[1]

Weak, diachronic, and synchronic emergentism

Three theories among the different varieties of emergentism deserve particular interest: *synchronic* emergentism, *diachronic* emergentism, and a form of *weak* emergentism. In synchronic emergentism, the relationship between a system's properties and its microstructure, i.e., the arrangement and properties of the system's parts, is at the center of interest. A property of a system is taken to be emergent if it is *irreducible* or, what I take to be the same, if it is *not* reductively explainable. In contrast, diachronic emergentism is mainly interested in *predictability* of novel properties. Those properties are taken to be emergent that could not have been predicted, in principle, before their first instantiation. Diachronic and synchronic forms of emergentism are not independent of each other, since irreducible properties are *eo ipso* unpredictable before their first appearance. Hence, synchronically emergent properties are also diachronically emergent, but not conversely.

Both of these stronger versions of emergentism are based on a common weak theory from which they can be developed by adding further theses. Currently, *weak emergentism* pervades emergentist theorizing in both theories of self-organization and various approaches to cognitive science, *e.g.*, connectionism, robotics, and artificial life. It is compatible with several reductionist approaches; and some champions of weak emergentism cite this compatibility as one of its merits compared with stronger versions of emergentism.

2 Weak emergentism

The first thesis of current theories of emergence – the thesis of *physical monism* – concerns the nature of systems that have emergent properties or structures. It says that the bearers of emergent features consist of physical entities only. It denies the existence of any supernatural component responsible for a system's having emergent properties. Thus, all substance-dualistic positions are rejected; for they base properties such as being alive or having cognitive states on supernatural bearers, such as an *entelechy* or a *res cogitans*, respectively. Instead, it is claimed that living beings and cognitive systems consist of the same basic entities that make up inanimate nature: it is nothing but specific sequels of highly complex physico-chemical states that realize their vital behavior or their mental states.

(1) *Physical monism*. Entities existing or coming into being in the universe consist solely of physical components. Likewise, properties, dispositions, behaviors, or structures classified as emergent are instantiated by systems consisting exclusively of physical entities.

[1] For more details on emergence and emergentism, see my book (Stephan 1999).

While the first thesis places emergent properties and structures within the framework of a physicalistic naturalism, the second thesis – the thesis of *systemic properties* – delimits the type of properties that are possible candidates for emergents. It is based on the idea that the general properties of complex systems fall into two different classes, those that some of the systems' parts also have, and those that none of the systems' parts have. These latter properties are called *systemic* or *collective* properties. General properties are properties of a general type, such as having a weight; they are not specific properties, such as having a weight of 85,4 kilogram.

(2) *Systemic properties*. Emergent properties are systemic properties. A property of a system is systemic if and only if the system possesses it but no proper part of the system possesses it.

It should be uncontroversial that both artificial and natural systems with systemic properties exist. Those who would deny their existence would have to claim that all of a system's properties are instantiated already by some of the system's parts. Countless examples refute such a claim; e.g., it is among the properties of a frog to jump or to croak, but no part of it (head, leg, heart, nor any cell assembly) can jump or croak; and it is among the properties of a connectionist network to generalize or to recognize patterns, but no single part of it (unit, etc.) has these properties.

While the first thesis restricts the type of parts out of which systems having emergent properties may be built up, and while the second thesis characterizes in more detail the type of properties that might be emergent, the third thesis specifies the type of relationship that holds between a system's microstructure and its emergent properties:

(3) *Synchronic determination*. A system's properties and dispositions to behave depend nomologically on its microstructure, that is to say, on the properties and arrangement of its parts. There can be no difference in a system's systemic properties without some difference in the properties or arrangement of its parts.

In recent debate, the thesis of *synchronic determination* is sometimes stated in a less stronger version as the thesis of *mereological supervenience*, which claims that a system's intrinsic properties supervene on its parts' properties and their arrangement. Then, too, there is no difference in the systemic properties without differences in the parts' properties or their arrangement. The thesis of mereological supervenience, however, is weaker than the thesis of synchronic determination, since it does not imply, strictly speaking, the *dependence* of the system's properties on its microstructure, it only entails their *covariance.*

Anyone who denies the thesis of synchronic determination has either to admit properties of a system that are not bound to the properties and arrangement of its parts, or to suppose that some other factors, in this case non-natural factors, are responsible for the different dispositions of systems that are identical in their microstructure. One may have to admit, for example, that there may exist objects that have the same parts in the same arrangement as diamonds, but that lack the diamond's hardness. This seems implausible. Equally implausible is the idea that there may exist two physically identical organisms, one viable and the other not. In the case of mental phenomena, opinions may be more divided; but one thing seems to be clear: anyone who believes, for example, that two physically identical creatures could be such that one is colorblind while the other is not, is not a physicalist. Similar considerations hold for propositional attitudes only as long as one does not subscribe to externalism, that is to say, if one does not claim that, *e.g*., the content of a belief depends essentially on the nature of the referents of the believer's thoughts and concepts.

Weak emergentism as sketched so far specifies the minimal criteria for emergent properties. It is *weak* because it is compatible with nearly all reductionist approaches. It is also the common base for all stronger theories of emergence. Moreover – and this is a reason for distinguishing it as a theory in its own right – it is held by both philosophers (*e.g.* Bunge 1977) and cognitive scientists (*e.g.* Rumelhart and McClelland 1986, Varela et al. 1991) in exactly its weak form.

3 Diachronic emergentism

All diachronic theories of emergence are based on a thesis about the occurrence of genuine novelties in evolution. This thesis excludes all preformationist positions.

(4) *Novelty*. In the course of evolution exemplifications of genuine novelties occur again and again. Already existing entities form new constellations that produce new structures which may constitute new entities with new properties and behaviors.

However, the thesis of novelty does not by itself turn a weak theory of emergence into a strong one, since reductive physicalism remains compatible with such a variant of emergentism. Only the addition of the thesis of *unpredictability*, in principle, will lead to stronger forms of *diachronic* emergentism.

The first occurrence of a systemic property can be unpredictable for at least two different reasons. Firstly, it is unpredictable, in principle, if it irreducible. This does not mean, however, that further occurrences of the property might not be predicted adequately. Secondly, it can be unpredictable because the microstructure of the system that exemplifies the property for the first time in evolution is unpredictable. For, if the microstructure of a newly emerging system is unpredictable, so are the properties which depend nomologically on it. Since in the first case the criteria for being unpredictable are identical with those for being irreducible, I postpone the analysis of those properties to the next section and confine myself here to the second case, *unpredictability of structure*. This version gains considerable significance in the teeth of strong interest in dynamical systems and chaotic processes.

The structure of a newly formed system can itself be unpredictable for two reasons. If the universe is indeterministic, then its novel structures will be unpredictable. However, from an emergentist perspective, it is of no interest if a new structure's appearance is unpredictable only as a result of its indeterminacy – most emergentists claim that the development of new structures is governed by deterministic laws. Nevertheless, deterministic formings of new structures can be *unpredictable in principle* if they are governed by laws of deterministic chaos.

An essential outcome of the theory of chaos is that there exist – even very simple – mathematical functions, whose own "behavior" cannot be predicted. Only the rise of "experimental mathematics" on highly efficient computers has revealed, for example, the properties of various logistic functions. Their intra-mathematical unpredictability has to do with an aperiodic behavior of these functions, by which marginally different initial values of some variable can lead to radically distinct trajectories of the functions.

A standard example is the logistic function $f(x) = \mu x(1\text{-}x)$ for $0 \leq x \leq 1$. For a parameter μ with $0 \leq \mu \leq 4$ the logistic function maps the interval [0,1] onto itself. Of particular interest is, how parameter μ exercises an influence on the long term behavior of the function when iterated repeatedly. For $0 \leq \mu \leq 1$ the situation is obvious. All initial values of the variable x let the function $f(x)$ approximate the value 0 after sufficiently many iterations, thus, the origin is the attractor. For $1 < \mu < 3$ exists exactly one attractor A of value $A = 1\text{-}1/\mu$: the function balances out on a stable value. If μ equals 3, the fixed point of the function is "marginally stable"; convergence is decidedly slowly – an indication for fundamental change in the function's behavior. For larger values dynamic becomes considerably complex. In the case of $3 < \mu < 1 + \sqrt{6}$ values oscillate between two fixed points. By increasing μ the attractors of period two will become instable, too. We get a cycle of period four (*i.e.*, after four iterations the values of the function approach in each case the four fixed points). At 3.56 the period doubles again and becomes eight, at 3.567 it becomes sixteen, and then we get a quickly rising sequence of periods to 32, 64, 128, etc. – vividly one speaks of cascades. At about 3.58, this sequence comes to an end. The period has doubled itself infinitely many times. Hereafter, predictions do not seem to be possible. Marginally different initial values x lead to radically different trajectories of the iterated function. Values jump pell-mell, convergence and divergence are not discernible: chaos dominates.

Thus, it looks as if just the most exact science of all has led us back to one of the starting points of emergentism. Whereas – after pioneering successes in chemistry and physics – we today do not count properties and dispositions of chemical compounds any more among emergent phenomena, examinations of deterministic chaos suggest the existence of systems that might develop structures that are unpredictable in principle and thus might show *structure-emergent* behavior.

Of course, one might argue that a Laplacean calculator could predict even chaotic processes correctly. Whether or not this "actually" could be the case is not yet settled. It depends mainly on the question of what kind of information we allow such a creature to have. For example, in Alexander's considerations (cf. 1920, ii, 72 f., 328) Laplace's calculator knows several earlier states of the whole world and, in addition, all natural laws that govern changes in the world. He seems to be able to extrapolate from his knowledge of all events that have occurred in the universe so far even the course of chaotic processes. But on what basis could he do that? Since chaotic processes are aperiodic, one can not determine definitely from the processes that have occurred up to a certain time the exact formula which would describe their further course. Even if the further course of the world is governed by deterministic laws, it does not follow from the earlier events and states alone, by *which* laws it is governed. Entirely different continuations seem to be compatible with the earlier course of the world. Therefore, even a Laplacean calculator could fail in his predictions. If one grants, however, that he knows *all* details of earlier world states – up to infinitely many digits –, and if one grants that he knows *a priori* which processes are governed by which *specific* chaotic laws, then, of course, he would be able to predict the forming of structures that are governed by these laws. I will leave it open whether or not it is plausible to ascribe such a knowledge to such a fabulous creature. At least we can be sure that creatures of our mental capacities do not have these forecasting abilities, and thus, we can legitimately suppose that where chaos exists, structures exist that are unpredictable in principle.

(5) *Structure-unpredictability*. The rise of a novel structure is unpredictable in principle if its formation is governed by laws of deterministic chaos. Likewise, any properties that is instantiated by the novel structure is unpredictable in principle.

Although diachronic emergentism with the thesis of structure unpredictability implies the unpredictability of all properties instantiated by systems that emerge from chaotic processes, it does not thereby imply their irreducibility. The unpredictability, in principle, of systemic properties is entirely compatible with their being reducible to the microstructure of the system that instantiates them. However, systemic properties are also unpredictable in principle, if they are irreducible. It is this feature that is at the core of synchronic emergentism, the doctrine which is most important for the interpretation of phenomenal consciousness, and which I am going to examine next.

4 Synchronic emergentism

The notion of *irreducibility* lies at the heart of all strong versions of emergentism. Although we might think of introducing it in contrast to the concept of *reductive explanation* as defined by Levine and Kim, it is still intriguing first to follow C. D. Broad's attempt to explicate a strong (synchronic) notion of emergence. In his book *The Mind and its Place in Nature* we find a passage that nowadays may count as downright classical; it reads:

> Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents *A*, *B*, and *C* in a relation *R* to each other; that all wholes composed of constituents of the same kind as *A*, *B*, and *C* in relations of the same kind as *R* have certain characteristic properties; that *A*, *B*, and *C* are capable of occurring in other kinds of complex where the relation is not the same kind as *R*; and that the characteristic properties of the whole *R(A,B,C)* cannot, even in theory, be *deduced* from the most complete knowledge of the properties of *A*, *B*, and *C* in isolation or in other wholes which are not of the form *R(A,B,C)* (1925, 61).

According to Broad's definition, a systemic property which is supposed to be nomologically dependent on its system's microstructure (by the thesis of synchronic determination), is called *irreducible* and therefore *emergent*, if and only if it cannot be deduced from the arrangement of its system's parts and the properties they have "in isolation" or in other (more simple) systems.

Although, prima facie, it looks as if Broad's proposal gives us a clear and distinct explication of what it is for a systemic property to be irreducible (or non-deducible), a further look reveals that two different kinds of irreducibility are concealed that have quite different consequences. As we will see, one type of irreducibility seems to imply *downward causation* while the other seems to imply *epiphenomenalism*. The failure to keep apart the two kinds of irreducibility has muddled the recent debate about the emergence of properties.

To make things clearer, I shall first discuss when a systemic property is *reducible*. For this to be the case, two conditions must be fulfilled: The first is that from the behavior of the system's parts alone it must follow that the system has some property *P*. The second condition demands that the behavior the system's parts exhibit when they are part of the system follows from the behavior they show in isolation or in simpler systems than the system in question. If both conditions are fulfilled, the behavior of the system's parts in *other* contexts reveals what systemic properties the actual system has. That is to say, those properties are reducible. Since both conditions are independent from each other, two different possibilities for the occurrence of *irreducible* systemic properties result: (a) a systemic property *P* of a system *S* is *irreducible* if it does *not* follow, even in principle, from the behavior of the system's parts that *S* has property *P*; and (b) a systemic property *P* of a system *S* is *irreducible* if it does *not* follow, even in principle, from the behavior of the system's parts in simpler constellations than *S* how they will behave in *S*.

Thus, a necessary requirement for a systemic property to be reducible is that its instantiation has to follow from the behavior of its bearer's parts. In other words: From the behavior of the system's parts it should follow that the system has all characteristic features that are essential for having the systemic property. Broad, for example, takes this condition, which is enclosed in the first criterion for reducibility, to be always fulfilled in the case of the characteristic properties of chemical compounds and viable organisms. Their properties might be irreducible only by violation of the second criterion, what means that from the behavior of the system's parts in other (simpler) systems it would not follow how they will behave in the actual system.

In contrast, he claims that the irreducibility of secondary qualities and phenomenal qualities results already from a violation of the first condition, since they were neither adequately characterizable by the macroscopic nor by the microscopic behavior of the systems' parts, even in principle. For, when we say that a certain object is red or a chemical substance has the smell of liquid ammonia, we do not mean that the corresponding system's parts *behave* or *move* in a certain way. No progress in the sciences could change this state of affairs in any way.[2] Broad has illustrated the fundamental distinction between (behaviorally) *analyzable* and *un-*

[2] However, whether reference to linguistic usage might answer questions concerning reducibility in a definite way is controversial. Particularly, P. Churchland has opposed arguments of the Broadian style (see 1988, 29 ff.).

analyzable properties by pointing to characteristic properties of organisms and secondary qualities, respectively.[3] If secondary qualities and phenomenal qualities are not analyzable, even in principle, then there is no prospect that an increase of scientific knowledge will close the gap between physical processes and secondary qualities or between physiological processes and phenomenal states of consciousness, respectively.

We can now specify in more detail the feature of irreducibility that is central for *synchronic* emergence. Its first variant is based on the behavioral unanalyzability of systemic properties:

(6) *Unanalyzability*. Systemic properties which are not behaviorally analyzable – be it micro- or macroscopically – are (necessarily) irreducible.

However, even if secondary and phenomenal qualities are unanalyzable properties it does not follow that the specific behavior of the system's parts upon which they supervene is itself not deducible from the behavior these parts exhibit in simpler systems. The irreducibility which results from a violation of the first criterion of reducibility does not imply, by itself, a violation of the second criterion of reducibility.

On the other hand, however, even analyzable systemic properties can be irreducible and therefore emergent. This is the case when the second criterion of reducibility will be violated, *i.e.*, when the behavior of the system's parts does not follow from their behavior in other (simpler) constellations. Broad thinks that such examples of irreducible behavior might occur in chemical compounds and also in organisms.[4] His central idea is that the parts of a genuinely novel structure, such as, *e.g.*, an organism in comparison to any inorganic compound, might behave in a way that is not deducible from the part's behavior in other structures. Implicitly, that means that the actual behavior of parts that interact in wholes does not result from their behavior in pairs. If the behavior of some system's parts is irreducible in this respect, then all properties that depend nomologically on the behavior of the system's parts (for example, reproduction) are irreducible too.

Thus, we can specify more precisely the second variant of a systemic property's irreducibility. It is based on the non-deducibility of the behavior of the system's parts:

(7) *Irreducibility of the components' behavior*. The specific behavior of a system's components within the system is irreducible if it does not follow from the components' behavior in isolation or in other (simpler) constellations.

3 Properties that are called "unanalyzable" for simplicity here, might be analyzable in other ways than by behavioral features. A certain smell, for example, might be analyzed as a mixture of the smells of chanel 5 and fish-meal. This, however, would not be an analysis based on concepts of motion and behavior.

4 In a recent paper (cf. Boogerd et al., forthcoming) we argue that cell biology offers many examples of strong emergence of this type.

A violation of the second criterion of reducibility, which is manifested in the irreducibility of the component's behavior, does not imply, however, a violation of the first criterion of reducibility. Systemic properties that cannot be reduced because the behavior of the system's parts is irreducible might nevertheless be behaviorally analyzable. Hence, the two necessary conditions of reducibility as well as those irreducibilities that are based on the violation of these conditions are independent of each other. Summarizing, we obtain from (6) and (7) the following modified version of systemic property irreducibility:

(8) (vi) *Irreducibility*. A systemic property is irreducible if (a) it is neither micro- nor macroscopically behaviorally analyzable, or if (b) the specific behavior of the system's components, on which the systemic property supervenes, does not follow from the component's behavior in isolation or in other (simpler) constellations.

Thus, we have to distinguish two completely different types of irreducibility of systemic properties. Equally different seem to be the consequences that go along with them. If a systemic property is irreducible because of the irreducibility of the parts' behavior on which the property supervenes, we seem to have a case of "downward causation". For, if the components' behavior is not reducible to their arrangement and the behavior they show in simpler systems or in isolation there seems to exist some "downward" causal influence exerted by the system itself (or by its structure) on the behavior of the system's parts. To be sure, if there existed such instances of "downward causation" it would not amount to a violation of some widely held assumptions such as, for example, the principle of the causal closure of the physical domain. Within the physical domain, we would just have to accept additional types of causal influences besides the already known basal types of mutual interactions.

In contrast, the occurrence of unanalyzable properties does not imply any kind of downward causation. Systems that have unanalyzable properties that depend nomologically on their bearer's microstructures need not be constituted in a way that amounts to the irreducibility of their components' behavior. Neither is implied that the system's structure has a downward causal influence on the system's parts, nor is there any reason to assume that unanalyzable properties themselves exert a causal influence on the system's parts. Rather it is dubitable, how unanalyzable properties might play any causal role at all. Since they are not behaviorally analyzable – that is to say, they neither seem to correspond to any "mechanism" nor do they seem to result from any "mechanism" –, it is hard to see how they themselves could be causally efficacious. And, if one cannot conceive of *how* unanalyzable properties might play a causal role, it is hard to conceive of them other than being epiphenomena.[5]

It is instructive to see how Kim and Levine's notions of *reductive explanation* (which may also serve to define the notion of synchronic emergence) relate to the foregoing analysis. Both authors note that in a first step we have to work the con-

5 In Stephan (1997) I tried to find a way how we nevertheless can attribute causal powers to unanalyzable properties.

cept of the property to be reduced "into shape" for reduction. Kim also calls this the "priming procedure" in which we must construe, or reconstrue the property to be reduced relationally or extrinsically. He considers a domain B of properties (or phenomena, facts, etc.) serving as the reduction base, and a property *E* that is going to be reduced (cf. 1999, 10-12):[6]

Step1 *E* must be *functionalized* – that is, E must be construed, or reconstrued, as a property defined by its causal/nomic relations to other properties, specifically properties in the reduction base B.

Now, this condition which we may call the *functionalization condition* is nothing but a new guise of Broad's analyzability condition. If the concept of a property cannot be worked "into shape" for reduction, that is, if the property cannot be "functionalized" then reduction necessarily fails. The property will turn out to be irreducible, hence synchronically emergent. But notice, as we just have seen, if irreducible properties exert "downward causation" they do not so *qua* being unanalyzable (or "unfunctionalizable").

As a second step both Kim and Levine mention the empirical task of finding the realizers of a property that could be worked into shape for reduction: "Stage 2 involves the empirical work of discovering just what those underlying mechanisms are" (Levine 1993, 132). This task is not explicitly mentioned by Broad, since he discusses situations where we already have nomological correlations between a system's microstructure and some systemic property, but are not able to deduce the property from our reduction base, *i.e.*, from the behavior of the system's parts in simpler systems or in isolation.

Step 2 Find realizers of *E* in B. If the reduction, or reductive explanation, of a particular instance of *E* in a given system is wanted, find the particular realizing property *P* in virtue of which *E* is instantiated on this occasion in this system (similarly, for classes of systems belonging to the same species or structure types).

Kim, in opposition to Levine, postulates a third step in which he refers to the development of theories that can explain how the realizers of *E*, the property to be reduced, perform the particular functional role that is characteristic for *E*.

Step 3 Find a theory (at the level of B) that explains how realizers of *E* perform the causal task that is constitutive of *E* (i.e., the causal role specified in Step 1). Such a theory may also explain other significant causal/nomic relations in which *E* plays a role.

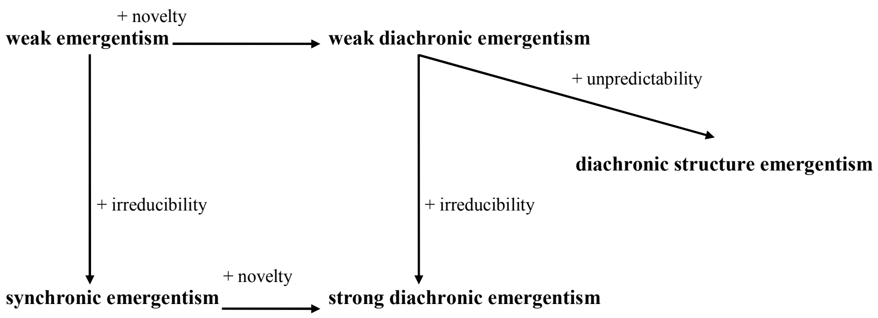
Although *prima facie* Kim's third step seems to correspond to Broad's second criterion, a closer look reveals important differences. Kim only considers theories that connect the reduction base of a systemic property (usually: the system's microstructure) with the property to be reduced. Thus, nothing can be said about any

6 Levine has put it as follows: "Stage 1 involves the (relatively? quasi?) a priori process of working the concept of the property to be reduced 'into shape' for reduction by identifying the causal role for which we are seeking the underlying mechanisms" (1993, 132).

"downward" causal influence by the system or its structure. The detailed analysis of Broad's approach, however, has given us the means to diagnose cases in which we should assume downward causation, namely such where no theory is capable to explain the behavior of a system's parts within this very system by reference to their behavior in simpler systems.

5 Synopsis

The following figure shows the logical relationships that hold between the different versions of emergentism:



Weak diachronic emergentism results from *weak emergentism* by adding a temporal dimension in the form of the thesis of novelty. Both versions are compatible with reductive physicalism. Weak theories of emergence are used today mainly in cognitive science, particularly for characterization of systemic properties of connectionist networks, and in theories of self-organization. *Synchronic emergentism* results from weak emergentism by adding the thesis of irreducibility. This version of emergentism is important for the philosophy of mind, particularly for debating nonreductive physicalism and phenomenal consciousness. It is not compatible with reductive physicalism. *Strong diachronic emergentism* differs from synchronic emergentism because of the temporal dimension in the thesis of novelty. *Structure emergentism* is entirely independent of synchronic emergentism. It results from weak diachronic emergentism by adding the thesis of structure-unpredictability. Although structure emergentism emphasizes the boundaries of prediction within physicalistic approaches, it is compatible with reductive physicalism, and so it is weaker than synchronic emergentism. Theories of deterministic chaos can be considered as a type of structure emergentism. This perspective is important for evolutionary research.

Coda

A case in point concerning irreducible mental phenomena is phenomenal consciousness. Among others, Levine and D. Chalmers claim that we cannot reductively explain it, and not so for empirical reasons. Apparently, phenomenal properties cannot be individuated by their causal roles, or as Levine says: "What seems to be responsible for the explanatory gap, then, is the fact that our concepts of qualitative character do not represent, at least in terms of their psychological contents, causal roles. […] Thus, to the extent that there is an element in our concept of qualitative character that is not captured by features of its causal role, to that extent it will escape the explanatory net of a physicalistic reduction" (1993, p. 134). Kim seems to share this appraisal when he states: "To get to the point without fuss, it seems to me that the felt, phenomenal qualities of experiences, or qualia, are intrinsic properties if anything is" (1998, 102). Therefore, he thinks that the functionalization of qualia won't work. It's for the same reasons that Broad already took them to be unanalyzable. Thus, qualia seem to be perfect candidates for irreducible, *i.e.*, synchronically emergent properties – properties, however, that do not exert downward causal powers *qua* being unanalyzable.

References

Alexander, S. (1920), *Space, Time, and Deity*. Two Volumes. London: Macmillan.

Boogerd, Fred C., Frank J. Bruggeman, Robert C. Richardson, Achim Stephan, and Hans V. Westerhoff (forthcoming), "Emergence and its place in nature. A case study of biochemical networks". *Synthese*.

Broad, Charles D. (1925), *The mind and its place in nature*. London: Kegan Paul, Trench, Trubner & Co.

Bunge, Mario (1977), "Emergence and the mind". *Neuroscience* 2, 501-509.

Chalmers, David J. (1996), *The conscious mind*. Oxford: Oxford University Press.

Churchland, P. (1988), *Matter and consciousness*. Revised Edition. Cambridge, Ma., London: MIT Press.

Levine, J. (1993), "On leaving out what it's like". In: M. Davies and G. W. Humphreys (eds.) *Consciousness*. Oxford: Blackwell, 121-136.

Kim, J. (1998), *Mind in a physical world*. Cambridge: MIT Press.

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

Rumelhart, David E. and James L. McClelland (1986), "PDP models and general issues in cognitive science". In: D. E. Rumelhart, J. L. McClelland and the PDP Research Group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition.* Vol. 1. Cambridge: MIT Press, 110-146.

Stephan, A. (1997), "Armchair arguments against emergentism". *Erkenntnis* 46, 305-314.

Stephan, A. (1999), *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation*. Dresden, München: Dresden University Press.

Varela, Francisco J., Evan Thompson, and Eleanor Rosch (1991), *The embodied mind.* Cambridge: MIT Press.