Weak Emergence and Complexity

Henrik Thorén[1] and Philip Gerlee[2]

[1]Lund University, Department of Philosophy, Kungshuset, Lundagård, 222 22, Sweden, henrik.thoren@fil.lu.se
[2]Center for Models of Life, Niels Bohr Institute, Blegdamsvej 17, 2100, Copenhagen, Denmark , gerlee@nbi.dk

Abstract

In this paper we consider Mark Bedau's notion of *weak emergence* (WE) and relate it to various attempts to objectively construe complexity. We argue that the heavy reliance on a specific notion of complexity risks rendering the concept superfluous. Furthermore we discuss what sort of systems might reasonably be understood as exhibiting emergence at all and point out that the macro-level needs to be at least minimally structured. A worry may thus be formed that macro-level generalisations provide the sort of short-cut that is explicitly excluded from WE thus potentially making the concept apply only to chaotic systems of limited interest (in this context).

Introduction

Artificial life research can in many instances be characterised as a search for the surprising. A very general question posed by researchers in the field is: what type of behaviour can we expect from a system with the following dynamics? If the answer is obvious or expected the system is often neglected or simply not classified as Alife because it is not life-like enough. Biological life is full of surprises and therefore ALife should be as well.

Fortunately systems with interesting and often surprising behaviour are not difficult to find. Classical examples include cellular automata of class IV (Wolfram, 2002), evolving systems such as Tierra (Ray, 1992), Avida (Ofria and Wilke, 2004) and more recently systems investigating chemical interactions such as Urdar (Gerlee and Lundh, 2010) and the Organic Builder (Hutton, 2009).

This notion of surprise or appearance of higher-order structure such as universal computation in CA or the evolution of parasites in Tierra is often in the literature labelled with the term *emergence*. The notion of emergence is however originally a philosophical term, with many precise albeit disparate definitions. In order to bring the concept more formally into the ALife-community Bedau (1997) recently introduced the notion of *weak emergence*, which takes a simulation-based approach to the definition of emergence. Roughly put, the idea being that a property P of a system S is weakly emergent iff the only procedure for deciding if S will have P at some later time is to simulate the system.

His approach has however been met with critique from several philosophers, e.g. for being too broad (Stephan, 2006). A defense of the thesis has been presented on several occasions (Bedau, 2003, 2008), clarifying his intentions and arguing for the merits of WE.

In this paper we will argue that Bedau's definition of weak emergence relies so heavily on a notion of *complexity* it risks conflating into it. Further we note that complex systems often exhibit higher-order structures, which can be described by law-like generalisations on that level, but this contradicts the very notion of weak emergence, suggesting that it misses the point all together. Whatever the outcome of this debate is we also note that established measures of complexity can lead to a quantification of weak emergence applicable to both real and artificial systems.

Emergence

The concept of *emergence* is usually traced back to a handful of British thinkers active during the second half of the 19th century among them figuring names such as John Stuart Mill, Samuel Alexander and C.D. Broad. They considered themselves as inhabiting a moderate position in which both dualism in the form of vitalism and mechanism could be avoided (Kim, 1999, 4). At its intuitive base the idea is that a whole can be more than the sum of its parts. Complexes may have properties not analysable in terms of the properties of their constituent parts. At the time this thought was very much empirically justifiable. The special sciences—chemistry was a favourite example—seemed to be hopelessly irreducible to ontologically more fundamental sciences, such as for instance physics.

Despite its appeal the idea withered to the onslaught of the unity of science movement and fell out of vogue from the 30s and onwards, not to be considered seriously again until the ultimate demise of that tradition in the early 70s.[1] Since

[1]Quantum mechanical explanations of chemical bonds is often blamed, chemistry being a favourite example of emergence for these philosophers and scientists.

then emergentism has experienced a small renaissance, not least within the scientific community. The interest in complexity as of the past couple of decades seem to have ushered its return.[2] In philosophical quarters emergentism or similar positions found new defenders among non-reductive materialists.

A central tenet of British emergentism was that emergents were entirely unpredictable from knowledge of their emergent base Kim (2006). The early Emergentists considered the appearance of emergent properties as metaphysically contingent, brute facts of nature. No amount of knowledge about the underlying structure allows one to predict the emergent. But since supervenience was thought to hold, appearance of emergent properties were considered to be lawful. Given that one had observed some emergent property in connection with some specific microstructure an "emergence law" (transordinal law on Broad's terminology) could be formulated. Such a law would be a fundamental law of nature. 'Prediction' should hence be understood as *theoretical* prediction, or derivation, and not as what one may call *inductive* prediction. Broad e.g. writes "[i]f emergence be true they [the emergent properties] *could* not have been deduced from any amount of of reflexion on the properties of these constituents taken separately or in non-living wholes..." (Broad, 1925, 75) Mill seem to have held a view very similar to this.[3] Properties of wholes that could be deduced straight-forwardly from the properties of their constituent parts were referred to as *resultant* properties. Oft cited C. Lloyd Morgan (1923) writes concerning the distinction between resultant and emergent properties.[4]

> ...both distinguish those properties (a) which are additive or subtractive only, and predictable, from those (b) which are new and unpredictable; both insist on the claim that the latter no less than the former fall under the rubric of uniform causation. (Morgan, 1923)

As Kim (1999) has pointed out there is reason not to take the 'additivity and subtractivity' requirement literally. The idea was to pick out properties that could be predicted by means of *some* compositional principle, as e.g. additivity or subtractivity. Other principles however were clearly acceptable; the law of composition of forces being a favourite example.[5]

So a resultant property is such that it can be calculated from knowledge of the basal properties by means of some compositional principle. Emergent properties of some whole were understood in contrast to this as properties that: 1), supervene on some basal property; and 2), is not predictable by means of such a compositional principle (and knowledge of properties of the parts).

But this is clearly not enough to make the distinction lucid. As the early Emergentists well understood given one is to combine a few quantities it is logically contingent what sort of principle one should use. Physics is riddled with straight-forward compositional principles and it seems that faced with a new case it is an entirely empirical matter which one is appropriate. Thus this would render cases like weight addition, composition of forces etc. cases of emergence which is clearly not right and definitely not what the early Emergentists had in mind. Broad and Mill solved this dilemma by putting restrictions on these principles disallowing principles working for properties of parts in other combinations. As McLaughin (2008, 92f) has pointed out the problem with such an approach is that almost nothing counts as emergent.[6]

An alternative strategy involves prohibiting what Van Gluick (2001) calls *specific value emergence*. Strictly speaking specific value emergence is not a form of emergence at all, but rather the most trivial form of resultance. Suppose we have a whole consisting of two proper parts a kilogram each in weight. The whole will weigh two kilograms despite none of the parts having that *specific* weight. We will return to this idea in the section below as this is part of Bedau's strategy.

Conclusively what is sometimes called *strong* emergence has been offered significant attention in the philosophical debate in the past twenty or so years and it has been found to suffer from serious problems. A lot of these problems stem from the difficulty to get the emergence/resultance distinction just right. Either too much or too little counts as emergent. Contemporary accounts typically strive for weaker formulations trying to salvage some part of the concept whilst giving others up. Mitchell (2009) does this by means of defending a form of downward causation deploying a multiple realisation argument. A different strategy is put to work by Bedau that defends a notion of emergence that tries to find objective criteria for a form of unpredictability that seems to fit the purposes.

[2]A search on Google Scholar combining the keywords complexity and emergence generates over a million hits. A quick browse through the philosophical literature will also reveal a connection between the terms 'emergence' and 'complex' that seems deeper than the connection warranted by taking 'complex' to denote an object that has parts.

[3]Mill never used the term 'emergence' but discussed what he called *heteropathic effects*, effects to which the causes do not abide by any principle of composition of causes. See McLaughin (1997) for a thorough discussion of Mill's views on this matter.

[4]The "both" here refer to the thinkers to which Morgan claims to owe this distinction; John Stuart Mill and George Henry Lewes.

[5]See e.g. (Mill, 1869, 210ff)

[6]Interestingly Kim (2006) has voiced critique seemingly pointing in the opposite direction claiming that emergence an accounts such as the above is under-characterised. The problem is that both supervenience and (in this case) non-derivability are negatively defined. Though not a decisive argument it raises the problem that the phenomena *emergence* might not be a genuine category.

Weak Emergence

Within the field of Artificial Life philosopher Mark Bedau has over a number of years developed and defended a variety of emergence he calls *weak emergence* (henceforth WE). WE may be characterised as a *strong* form of epistemological emergence since it does not rely on psychological or logical limitations of human cognition but rather an objective notion of complexity.

Bedau has written extensively on the subject but here we are going focus on two more recent works, Bedau (2003) and Bedau (2008) respectively. In these texts one find several characterisations, in the first article WE is defined in terms of a requirement of simulation, in the second an appeal to *explanatory incompressibility* is voiced. Bedau himself however views these two varieties as essentially one, "[t]hese two definitions are similarly indirect, and they are essentially equivalent" (2008, 444). We shall also treat them as such. Hence we believe that the following reflects Bedau's idea well. For a macro-property M of a system S to be WE the following two criteria should be met;

1. M is nominally emergent.

2. There is a derivation from P to M but that derivation can only be generated through simulation.

Nominal emergence is understood as the "...notion of a macro property that is the kind of property that cannot be a micro-property." (Bedau, 2003, 158) Notably this is equivalent to what Van Gluick (2001) calls *modest kind emergence*, at least taken in the stronger modal version. The necessity claim here is not further specified though the name suggest nominal necessity. In that case this qualification taken by itself includes a host of phenomena on both sides of the resultant/emergent divide. Bedau seems well aware of this (Bedau, 2003, 158).

This second criteria is a little more difficult. Importantly Bedau accepts (for the systems under scrutiny anyway) what he calls *causal fundamentalism*, the thesis that "...macro causal powers supervene on and are determined by micro causal powers" (Bedau, 2003, 159). So strictly speaking WE properties are only resultant, as there exists a derivation from micro to macro. Bedau's idea however is to pick out a certain kind of derivation. In Bedau (2003) this is to be thought of as "derivation by simulation," and this in turn should be interpreted in the strongest possible sense. Bedau writes:

> A derivation by simulation involves the temporal iteration of the spatial aggregation of local causal interactions among micro elements. (Bedau, 2003, 164)

What Bedau seems to be saying is that a simulation here is a process that produces or reproduces the *actual* mechanism in question.[7] Hence WE phenomena appear in accurate computer simulations and natural systems alike.[8] A central feature of such a derivation is that it must be done stepwise so that the further into the future one is interested in making predictions, the longer the derivation will be. In Bedau (2008) WE is thought of in terms of *incompressible generative explanations* connecting micro-state P with emergent M. Bedau writes:

> An explanation is generative just in case it exactly and correctly explains how macro-events unfold over time, how they are generated dynamically. (Bedau, 2008, 445)

This characterisation also requires the 'explanation' to follow the *actual* procedure (crawling the causal web) and 'short-cuts' are explicitly prohibited.

> If an explanation of some macro-property of some system is incompressible, then there is no short-cut generative explanation of that macro-property that is true, complete, accurate, and can avoid crawling the causal web. (Bedau, 2008, 446)

Let is try to construe this in a more formal fashion.[9] Suppose we have a micro-P (an initial condition) and a macro-M (at some later time) that stand in a WE relation to each other.[10] Then there is some sequence $P_1, P_2, ..., P_n$ connecting P and M, let us call this sequence D. There is no other sequence connecting P and M that is shorter than D and also satisfies the criteria of being true, complete, accurate and avoids crawling the causal web. We take it that if it is true and complete it must also be accurate and "crawling the causal web" entails that for every other derivation E that is exactly as long as D, then E is identical to D.

What about false derivations that are shorter but none-the-less accurately predict M from P? It seems that this characterisation is much too strong. Truth, completeness, accuracy and causal web-crawling trivially homes in on just these micro-sequences, regardless of the system at hand. If it is the dynamics one is interested in, then broad and approximative statistical models that essentially leap-frogs the bowels of whatever process one is studying, just won't do. But that is

[7]On a weaker understanding one would only require from the simulation that it be sufficiently similar with respect to some characteristics of the original process. The mechanism driving the simulation however would not have to be qualitatively identical to process which it mimics.

[8]Of course inaccurate simulations could also exhibit WE, put perhaps with other emergents than the ones belonging to the system they are mimicking.

[9]In the below section we use 'derivation' instead of 'explanation,' we do not however think it matters. The explanation Bedau seem to have in mind are derivational. Besides 'derivation' is the preferred term in Bedau (2003).

[10]Bedau interchangeably talks about objects, properties, states and facts so let us give this a neutral account.

so regardless of whether it is possible to do so or not. So it seem Bedau would have to opt for some more inclusive idea of what exactly amounts to a short-cut. Perhaps the idea that derivations concerning states further away requires more computational power is more important and promising. Our worry however is that in order to avoid making a characterisation that is non-trivial Bedau would have to accept that there can be no regularities at all elsewhere in the system, and this in turn warrants the question whether the system at hand has any macro-level at all. We will however return to this topic in our discussion.

What sort of systems might this be true of then? Bedau relates this to systems that are *complex*. Emergents, on Bedau's take, is not epistemological in the sense that emergents are dependent on "human frailty." To the contrary not even infinite knowers could avoid using this type derivation in making successful predictions regarding these systems.

> Incompressibility of explanations is a consequence of the objective complexity of the local micro-causal interactions that are ultimately generating the emergent behavior being explained (Bedau, 2008, 453).

Thus Bedau means to move the 'ontological burden' away from the notion of emergence, where it has shown to be problematic, to the notion of complexity. We will now move on to discuss the notion of complexity introducing a few of formal complexity measures, and propose a link between WE and the complexity of a system.

Complexity

An intuitive understanding of the predicate 'complex' with regards to some object (process or pattern) entails that the object is structured in such a way that it is very difficult (or perhaps impossible) to describe.[11] In recent years the study of complex systems have enjoyed some popularity, especially within biology and ecology but also within e.g. statistical mechanics where the aim often have been to provide formal definitions or objective criteria. A quantitative measure has however turned out to be difficult to find. This is at least partially due to disparate use of the term in various disciplines; complexity is often thought to be salient in structures such as the human brain, weather and climate systems, but also in single-celled organisms. In the scientific community it has been in use since the rise of systems theory and cybernetics in the 40s and 50s, and has the last 20 years experienced a revival. On some construals the notion seems to approximate the concept of emergence. Consider for example the definition by Simon (1962):

> Roughly, by a complex system I mean one made up of a large number of parts that interact in a nonsimple way. In such systems, the whole is more than the sum of the parts, not in an ultimate, metaphysical sense, but in the important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole.

This definition falls close to the weak sense of emergence, but of course depends on how we interpret 'not a trivial matter'. A more recent remark by physicist Nigel Goldenfeld (Editorial, 2009) states that:

> Complexity starts where causality breaks down.

This claim is even stronger, and might put complexity on par with stronger notions emergence. However, independent of the exact interpretation of these statements our point is that the notions of emergence and complexity are intertwined, and that Bedau's notion in fact lies close to well-developed quantitative measures of complexity. Before we proceed with this thesis, let us look more closely into what we mean by complexity and how to measure it.

The concept of complexity has a relatively short history in the natural sciences. Before the 20th century the physical sciences were confined to the study of *simplicity*, while biology and the medical sciences, unable to explain the omnipresence of complex form and function, were concerned with collection and classification of living systems. It is here important to distinguish between systems which are complex and those which are merely complicated, or as put by Weaver (1948): complex in a organised vs. disorganised way. By complicated systems we refer to those which consist of large number of interacting parts with many degrees of freedom, such as an ideal gas, which yield to a statistical description, while complex systems are those which tend to organise themselves and exhibit structure despite being governed by local microscopic rules of interaction.

Intuitively we would like to class objects as being complex if they lie somewhere in between complete order and randomness. The human eye and the organisation of a colony of termites are things typically considered complex, while a crystal structure with its endless repetition, or an unstructured gas both fall outside our notion of complexity. To capture this intuition into a quantitative measure has however turned out to be immensely difficult. Many attempts have been made at defining complexity, either from a *structural* or *functional* point of view (McShea, 1996; Wimsatt, 1972), although none fully satisfactory, and the most successful route has instead been to consider the complexity of strings, called *sequence* complexity.

The first attempt along these lines was made by Kolmogorov (1968) (and later Chaitin (1975)) and quantifies the complexity of a sequence as the shortest possible description of that sequence. This is done by considering the shortest computer program or algorithm which when executed will reproduce the sequence in question, and from this

[11] One may thus note that already on this early stage there is some tension between ontological and epistemological aspects of the concept.

complexity measure has gained its name Algorithmic Complexity (AC). It is also related to the amount of information contained in the sequence as defined by Shannon entropy (Shannon, 1948). The problem with this measure is that it assigns maximal complexity to sequences that are completely random, and also assigns low complexity to intricate objects that can be generated with simple rules. A prime example of this is the Mandelbrot set, which because it can be generated with a very short algorithm has a low AC, although its structure suggests otherwise. AC therefore deviates from our intuitive notion of complexity, at least in some instances.

By measuring the running time of the shortest computer program generating the sequence, instead of its length, Bennett (1988) was able to overcome the problem of assigning low complexity to seemingly complex mathematical objects. This approach was motivated by the fact that complex objects often have a long causal history, and by equating the history with running time a quantitative measure can be defined. These attempts are nevertheless intractable because the length of the shortest program is provably non-computable, and we have no way of *a priori* telling which program is the most plausible.

This shortcoming was addressed by Grassberger (1986) who suggested an Effective Measure Complexity, which measures the complexity of a sequence as the value of having observed all previous symbols in the sequence when guessing the next. A similar measure termed Statistical Complexity was developed by Crutchfield and Young (1989), and measures the minimum amount of information required to make optimal guesses of the symbols in the sequences at an error rate h, where h is the Shannon entropy of the sequence. One drawback with these two measures is that they cannot measure the complexity the of a single sequence, but only of the ensemble from which sequences are drawn, although one can argue that complexity in fact is a property of an ensemble and not of a single object.

Applying these measures to dynamical processes can be accomplished by mapping the trajectory of the system, by a partition of the state space, into a symbol sequence which can then be analysed. For example the trajectory of the logistic map can be mapped to a binary alphabet and the corresponding binary sequence then reflects the complexity of the underlying dynamical systems, which turns out to be maximal at the period-doubling accumulation (Crutchfield and Young, 1989; Crutchfield, 1994). However, the structure of objects such as living organisms are currently impossible to capture by the dynamics of their underlying processes, which means that the above measures still fall short of a satisfactory account of complexity.

Systems which exhibit a high degree of complexity (in the sense of EMC and SC) have the interesting property that they exhibit structure (i.e. they are not maximally random) but at the same time the future state of the system is difficult to predict. This property has been termed "computational irreducibility" (Wolfram, 2002) and more precisely means that there is no way of predicting how the system will behave except by explicit simulation. Please note that this also holds for chaotic systems[12], but is of less interest as it is the combination of structure and unpredictability which we usually find interesting.

Precisely which systems qualify as computationally irreducible is currently unclear, but one sufficient condition is computational universality (i.e. Turing completeness). This condition is met by a few surprisingly simple systems such as Wolframs one-dimensional CA rule 110 (Cook, 2004), and the Game of Life (Berlekamp et al., 1982), which for some specific initial conditions instantiate a Universal Turing Machine. At least for a subset of these initial conditions the system is computationally irreducible, otherwise it would violate the halting problem. This suggests a link between universality and complexity which led Wolfram (2002) to formulate the Principle of Computational Equivalence, which states that all processes in nature (that are not obviously simple) can be considered as computations, and are of such complexity that they attain computational irreducibility. The human brain, an ant colony and a weather system, are according to the principle of the same computational sophistication, and instantiate computations which are irreducible. This is an intriguing and very bold statement, which if it is true, clearly has bearing on the ontological status of these objects.

Returning to WE several connections should become clear. Obviously unpredictability plays an integral part. Moreover incompressibility as Bedau thinks of it is very similar to computational irreducibility. Systems which are computationally irreducible and thus in principle impossible to forecast (and do not exhibit chaos) are precisely those of high complexity. This was already noted by Bedau (2003), but he did not follow through on the connection, which in the end leads to an interesting conclusion. In avoiding the metaphysical pitfall of the otherwise attractive idea of ontological emergence by appealing to complexity one find similar questions can be stated yet again, is *complexity* to be understood in ontological or epistemological terms? Wolfram's claim is that computational irreducibility and thus ontological complexity is ubiquitous in nature, and possibly the only one worth considering, although both concepts could clearly coexist.

Although the question of ontological complexity might be impossible to answer the link established between weak emergence and complexity might allow for quantification of the emergence a system exhibits. Systems with low complexity are easy to forecast, while those with high complexity might be impossible to predict the future of without ac-

[12]The relation between WE and deterministic chaos will be discussed below.

tually iterating the dynamics. This might provide a different route to quantifying weak emergence than the one suggested by Hovda (2008), which measures the degree of emergence as the length of a formal derivation of property P from the initial conditions, and instead focuses on the amount of information needed to make optimal predictions about the future of the system with respect to some property P.

It is also worth mentioning that complexity has previously been suggested as a route to defining emergence, by considering the predictive efficiency of a set of causal variables describing a system (Shalizi and Moore, 2003). The predictive efficiency can be quantified as the ratio between EMC and SC, and a set of variables are considered emergent from another set if 1) one is a coarse-graining of the other and 2) the coarse-grained variables can be predicted with higher efficiency. The prototypical example for this type emergence is the relation between statistical mechanics and thermodynamics.

Discussion

Complexity is usually thought to relate to emergence by causing it, or giving rise to it. Once a system reaches a certain degree of complexity emergent properties will start to appear. The relationship is more curious however. The reason is that complexity itself is an obvious systemic property that, at least in the systems under scrutiny here, spring from micro-structures that do not exhibit it. Quite to the contrary, at their ontological bottom they are notoriously simple. One the other hand the opposite might be true. A system may have a microstructure that is beyond description whilst being highly predictable on the macro-level. In that case we would perhaps talk of the emergence of simplicity. Given of course we deploy a weaker version of the concept. In the previous section we established a link between WE and complexity as measured by statistical complexity or effective measure complexity. We will now elaborate on this and the implications it has.

Interestingly it is often in complex systems that we find higher-level structure that behaves lawfully with respect to some higher-order dynamics. This is precisely the domain of the special sciences. Let us consider two examples of this lawfulness: In the Game of Life (GOL) (Berlekamp et al., 1982) there is a configuration known as a 'glider'. It consists of five active cells and has the peculiar property of moving across the lattice in a diagonal fashion. Now if we know that a glider is moving in a particular direction and at a given time is located at position x, then if it does not collide with any other cells predicting its position for all future times is easy, and does not require that we simulate the entire system. Next consider the dynamics of an ant colony. Without knowing the exact details of the anatomy of a particular ant, we can by coarse-graining it into what type of ant it is (queen, soldier etc.) get a good picture of what duties it will have in the colony. The system clearly exhibits regularities which allows us to formulate higher-order laws (or at least law-like generalisations), which in turn allow for prediction of the dynamics.

Although these systems, might be computationally irreducible on the micro-level they are still amenable to a coarse-grained description which can make reasonable predictions about the future state of the system. There is thus a clear tension in the link between WE and complexity that was presented above. Complex systems are possibly computationally irreducible and thus WE, but at the same time a WE system does not allow any short-cut derivations, which is precisely what higher-order structure allow. But again picking out systems with no higher-level structure at all seems to exclude precisely the kind of systems about which talk of emergence is the most appropriate.

Higher-order descriptions are typically coarse-grained in more than one respect; firstly by individuating the system differently (e.g. by using functional definitions), and secondly that they may imply some loss of accuracy in the predictions. This can happen in two ways, either as a consequence of noise, or as consequence of abstraction to more general terms.

The loss of accuracy is dependent on the level of coarse-graining one applies to the system. At the level of no coarse-graining we have to, assuming that the system is computationally irreducible, iterate the dynamics explicitly to make predictions about the future state of the system, e.g. if it will have a certain property P at time t. Now if we move one level up in the coarse-graining, e.g. in GOL we start talking about gliders and blinkers, we might be able to formulate laws at this level which faithfully describe the system, such as the fact the gliders move diagonally at the speed of light. These laws allows us to circumvent the actual simulation, but on the other hand introduces inaccuracy in the description. It also denies us any knowledge about the micro-state of the system at future times, as coarse-graining procedures by definition are non-invertible.

In the above example of the ant colony, knowing the type of ant only gives us a better than null prediction as to its behaviour, obviously not a perfect prediction of the future actions of the ant in question. For every coarse-grained description of the system we thus have an error rate of prediction. What we save in terms of not having to simulate the system at the 'basal' level is lost in the power of prediction. The rate at which this error increases varies between different systems depending on their regularity. Now, one way to read Bedau is to say that a WE occurs when the error rate of prediction on all coarse-grained levels is sufficiently high. To reliably forecast the dynamics it is necessary to revert to an explicit simulation of the system.

This discussion can in fact be couched in terms of Cruthfields ϵ-machine reconstruction (Crutchfield, 1994), where automata with different 'causal' states are able to predict the future state of a system with varying accuracy. Viewing different levels of description as different ϵ-machines, we can make a formal comparison of both their complexity[13] and their accuracy. A similar approach to different levels of description has been pursued by Dennett (1991) in his discussion on the reality of patterns and ultimately beliefs in nature. He also notices the inherent trade-off between an accurate and complicated description versus a simple one with a higher error rate, and that this leads to a multitude of possible 'patterns' in the same data.

The above discussion covered systems which exhibit structure on some higher level, but there is also an interesting link between WE and deterministic chaos (DC). Chaotic systems are generally governed by local micro-level rules, or non-linear equations of evolution, and their hallmark is their sensitive dependence of initial conditions. This means that trajectories at machine precision distance from each other diverge exponentially, and implies that predictions about the future state of the system are difficult or impossible to make.[14]

These systems do not show regular structure[15], except possibly for some isolated regions of parameter space, and are also highly sensitive to initial conditions. The future state of a chaotic system is difficult to predict without simulation, and for reasonable choices of a property P it thus fulfills the criterion for WE, i.e. there are no short-cuts for predicting if the system will have P, it can only be decided by explicit simulation.

Depending on our rigour when accepting short-cut derivations, based on their accuracy, we naturally get different degrees of overlap between weakly emergent and chaotic systems. If we only accept predictions which are perfectly accurate then the class of WE-systems might incorporate both chaotic and complex systems, while if our criterion for accuracy is lower, and we accept statistical laws, then WE coincides more with systems considered chaotic.

Suppose we consider a form of system of which a concept of emergence does some actual work. As we have noted before the most obvious category consists of systems that have higher levels that are at least minimally structured, i.e. systems that succumb to macro-level generalisation of some form and degree of accuracy.[16] However, as discussed above, these systems seems to be excluded by definition from WE. The reason would be that macro-level regularities plausibly could be understood as exactly the kind of 'short-cut' Bedau dismisses. If this is true it seems WE can only be applicable to systems that are macroscopically unstructured. But it seems systems that lack structured macroscopic levels are usually uninteresting.

In a way this worry seems entirely misguided. The reason is that since these macro-level generalisation are located on the macro-level they themselves constitutes the emergents in this contexts and it is the derivation *of* them rather than *between* them that is under scrutiny. In other words, the rules which govern the higher-order structures (e.g. the collision of two gliders in GOL) are not derivable except by simulation from the micro-level dynamics.

To determine if this objection is genuine it seems one would have to specify what is micro and macro properties for the system under investigation. Though this might seem conceptually trivial it is decidedly less than straight forward in this particular context. We have already hinted at an example; a lot of kinds are functionally defined in GOL, take e.g. spaceships; anything that moves whilst retaining its shape over a relatively short period of time is a spaceship. Thus it makes out a *kind* on some non-basal level of description. But since any number of different micro-level configurations might exhibit this behaviour it seems there won't be a micro structural definition of spaceships. Some specific kinds of spaceships do have micro structural definitions, gliders are an example of that.

Other interesting candidates are more abstract systemic features like *chaos* or *complexity* that both seem to intuitively fit well on at least some conceptions of emergence. These predicates are usually ascribed (in this context at least) to entire systems where microscopical structures typically are very simple. They are thus systemic properties that are genuinely novel—systems with simple microstructures are not always complex—and they apparently aren't trivial in the sense that one can easily find configurations in e.g. GOL that do not exhibit complexity on any technical understanding of the term. Yet another category that might coextend with the one just mentioned concerns questions regarding specific initial states. Suppose one has a certain initial state for GOL and wants to know if it will produce a bounded dynamic or not. For some configurations these questions are computationally irreducible and thus also weakly emergent on Bedau's understanding, but what sort of macro-properties do these future states represent?

These are the types of questions that need to be addressed if we are to get a proper account of the relation between weak emergence, complexity and deterministic chaos.

In this paper we have elaborated on the connection between weak emergence and complexity. We found that WE lies very close to certain measures of complexity, and this might allow for a quantitative measure of WE. Further we noticed that complex systems often exhibit higher-order structure

[13]If the machine is minimal, then its statistical complexity is the amount of memory (in bits) required for the agent to predict the environment at the given level 'ϵ' of accuracy

[14]See Kellert (1993) for an extensive argument of the latter.

[15]Here we disregard from coarse-grained structure such as invariant measures, which can be defined for chaotic systems exhibiting ergodicity.

[16]This needs to further specified but following Fodor (1974) we think that minimally the higher level consists of functional kinds, usually however these kinds will allow for something more, macro-level laws or at least law-like generalisations.

which allows for coarse-grained prediction of the dynamics. This is in possible contradiction to the definition of WE, which implies that the scope of WE is narrow and possibly only covering systems exhibiting deterministic chaos. Instead we propose a different interpretation of the concept which focuses on the derivability of the rules acting on the higher levels in the system.

Acknowledgment

This work was in part supported by Adlerbertska Forskningsstiftelsen.

References

Bedau, M. (1997). *Philosophical Perspectives: Mind, Causation, and World*, chapter Weak Emergence, pages 375–399. Blackwell.

Bedau, M. A. (2003). Downward causation and the autonomy of weak emergence. *Principia Revista Internacional de Epistemologica*, 6:5–50. Page references are made to Bedau and Humphreys (2008).

Bedau, M. A. (2008). Is weak emergence just in the mind? *Minds & Machines*, 18:443–459.

Bedau, M. A. and Humphreys, P., editors (2008). *Emergence: Contemporary Readings in the Philosophy of Science*. MIT press.

Bennett, C. (1988). *The Universal Turing Machine*, chapter Logical depth and physical complexity, pages 227–257. Oxford University Press.

Berlekamp, E., Conway, J., and Guy, R. (1982). *Winning ways for your mathematical plays*. New York: Academic Press.

Broad, C. D. (1925). *The Mind and Its Place in Nature*. Routledge.

Chaitin, G. (1975). Theory of program size formally identical to information-theory. *J Acm*, 22(3):329–340.

Cook, M. (2004). Universality in elementary cellular automata. *Complex Systems*, 15:1–40.

Crutchfield, J. (1994). The calculi of emergence - computation, dynamics and induction. *Physica D*, 75(1-3):11–54.

Crutchfield, J. and Young, K. (1989). Inferring statistical complexity. *Physical Review Letters*, 63(2):105–108.

Dennett, D. (1991). Real patterns. *The Journal of Philosophy*, 88:27–51.

Editorial (2009). No man is an island. *Nature Physics*, 5:1.

Fodor, J. (1974). Special sciences. *Synthese*, 28:97–115.

Gerlee, P. and Lundh, T. (2010). Productivity and diversity in a cross-feeding population of artificial organisms. *Evolution, In Press.*

Grantham, T. (2006). Is macroevolution more than successive rounds of microevolution. *Paleontology*, 50:75–85.

Grassberger, P. (1986). Toward a quantative theory of self-generated complexity. *Int J Theor Phys*, 25(9):907–938.

Hovda, P. (2008). Quantifying weak emergence. *Minds and Machines*, 18:461–473.

Hutton, T. (2009). The organic builder: A public experiment in artificial chemistries and self-replication. *Artificial Life*, 15:21–28.

Kellert, S. H. (1993). *In the Wake of Chaos*. The University of Chicago Press.

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

Kim, J. (2006). Emergence: core ideas and issues. *Synthese*, 151:547–559.

Kolmogorov, A. N. (1968). Three approaches to the quantitative definition of information. *International Journal of Computer Mathematics*, 2:157–168.

McLaughin, B. (1997). Emergence and supervenience. *Intellectica*, 25:25–43.

McLaughin, B. (2008). The rise and fall of british emergentism. In Bedau, M. A. and Humphreys, P., editors, *Emergence: Contemporary Readings in Philosophy and Science*. MIT press.

McShea, D. (1996). Perspective: Metazoan complexity and evolution: Is there a trend? *Evolution*, pages 477–492.

Mill, J. S. (1869). *System of logic*. Harper & Brothers, New York.

Mitchell, S. D. (2009). *Unsimple Truths: Science, Compexity, Policy*. The University of Chicago Press.

Morgan, C. (1923). *Emergent evolution*. Williams and Norgate, London.

Ofria, C. and Wilke, C. (2004). Avida: A software platform for research in computational evolutionary biology. *Artificial Life*, 10:191–229.

Ray, T. (1992). An approach to the synthesis of life. In Langton, C., Taylor, C., Farmer, J., and Rasmussen, S., editors, *Artificial Life II*, pages 371–408, Redwood City, CA. Addison-Wesley.

Shalizi, C. and Moore, R. (2003). What is a macrostate? subjective observations and objective dynamics. *http://arxiv.org/abs/cond-mat/0303625*.

Shannon, C. E. (1948). A mathematical theory of information. *Bell System Technical Journal*, 27:379–423.

Simon (1962). Architecture of complexity. *Proceedings of the American Philosophical Society*, 106:1–17.

Stephan, A. (2006). The dual role of 'emergence' in the philosophy of mind. *Synthese*, 151:485–498.

Van Gluick, R. (2001). Reduction, emergence and other recent options on the mind-body problem: a philosophical overview. *Journal of Consciousness Studies*, 8:1–34.

Weaver, W. (1948). Science and complexity. *American Scientist*, 36(4):536–544.

Wimsatt, W. (1972). Complexity and organization. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1972:67–86.

Wolfram, S. (2002). *A new kind of science*. Wolfram Media.