Polanyiana 2019 (28) 1-2.

A Criticism of Weak Emergence

Mihály Héder and Dániel Paksi

Budapest University of Technology and Economics

e-mail: mihaly.heder@filozofia.bme.hu; daniel.paksi@filozofia.bme.hu

Abstract: This article presents an internal criticism of Mark A. Bedau's concept of *Weak Emergence (WE)*. WE is based on the concept of *explanatory incompressibility:* it states that a phenomenon is weakly emergent if it cannot be explained without crawling the *micro-causal web* of the underlying processes. Bedau's main goal is to show that WE is not merely "Just in the Mind" but is something that can be objectively analyzed. This paper shows that explanatory incompressibility and thus the WE property depends on the observing person's a priori choice of operators for the explanation. The paper demonstrates that Bedau's examples for WE from Game of Life are not weakly emergent if alternative sets of premises are used. Without a non-arbitrary method of picking operators, WE is completely relative to the observer and is not useful for settle debates in the manner that is expected from an objective measure.

Keywords: Mark A. Bedau, weak emergence, computational emergence, reduction, Game of Life.

1. Introduction

Bedau's aim with developing the concept of Weak Emergence (WE) was to overcome a dilemma between the *complete denial of emergence (or claiming that it is just merely in one's mind)* and the position of *strong emergence*. He finds both of these positions to be too extreme in one sense or the other and proposes WE as a middle ground.

In reviewing all the different concepts of emergence, we identify two main types: *ontological* and *epistemological*. The former, sometimes also called *strong emergence,* means that even though there is only one substance that the world consists of – matter – there are higher, non-substantive ontological levels of existence created from, but not reducible to, matter[1]. This

[1] McLaughlin, 1992.

avoids both monisms, for instance, materialism, *substance dualism* of matter, and some other thing, e.g., an immaterial spirit. The proponents of strong emergence usually aim at the improvement of empirical science – very often physics and biology – by including ontologically emergent phenomena they claim to be observable if not approached with the wrong monist or dualist mindset.

The latter position, sometimes also called *weak* emergence (where this is not meant in Bedau's specific sense), only accepts that there are emergent *properties* of phenomena that cannot be reduced *because of epistemic constraints* to the fundamental properties of the substance of which they are composed - usually matter. This position rejects the possibility of ontologically emergent objects, usually again on scientific grounds, for instance, by citing the completeness of physics[2]. Thus with Bedau's words, this kind of emergence is merely in the observers' mind.

In Section 2 of this paper, a recent, analytic account of Mark A. Bedau's *Weak Emergence (WE)* is presented. In Section 3, detailed criticism of WE, is given, followed by the conclusion (Section 4).

2. Bedau's Concept of Weak Emergence

Before providing his definition of WE, Bedau first strongly distances himself from strong emergence. This is important for him because he also wants to distance himself from strict materialism by his WE concept. He proposes two main reasons for distancing himself from strong emergence. One is the problem of the independent causal effectiveness of the higher ontological levels that are a common feature of strong emergence theories. This originates from the British emergentists, who held that causal autonomy is a necessary condition for ontologically emergent entities (Alexander's dictum by Jaegwon Kim's[3] words).

Bedau, however, accepts Kim's argument that the causal autonomy of the higher levels means that there is *downward causation* – that is, an ontologically higher level of existence can have

---

[2] A word of caution: this whole debate is absolutely different from the one of the *emergent computing* researched by Cariani, 1991 or Hordijk, Crutchfield and Mitchell, 1998. While these works also rely on Game of Life and are contemporary with Bedau's efforts they investigate how global coordination can be achieved by simple building blocks in computer algorithms. Philosophically, they can be categorized as reductionist, in the sense that the emergent properties just appear in the eyes of the beholder. See Brunner, 2002.

[3] Jaegwon Kim,1998.

an effect on the underlying material level. This, in turn violates the causal closure of the material world[4], which contradicts the state of science and hence is unacceptable. "Although strong emergence is logically possible, it is uncomfortably like magic." [5]

Bedau's other problem is that in his evaluation strong emergence appears to have no scientific value. On the possible programs of strong emergence he quotes O'Connor, who in turn cites Roger W. Sperry and Michael Polanyi. In O'Connor's interpretation these authors use strong emergence for explaining consciousness and phenomena of life, respectively.[6]

Bedau thinks that these many decades-old works give no good argument for the usefulness of strong emergence. On the contrary, he remarks that strong emergence features the symptoms of a degenerative research program, as all of the proponents cite the same outdated Sperry and Polanyi works but fail to connect to contemporary research topics or to show novel results. Bedau thus concludes that "To judge from the available evidence, strong emergence is one mystery which we don't need".[7] However, Bedau claims that *weak* emergence, particularly as defined by him, "(…) is metaphysically innocent, consistent with materialism, and scientifically useful".[8]

According to his arguments, contemporary science shows that all higher-level phenomena can be reduced to lower-level material parts, so Bedau's main idea is that emergence is not the contrary concept of reduction as it was originally defined by Lloyd Morgan[9] and C. D. Broad[10] but is based on a *special kind of reduction* between the (weakly) emergent level and the fundamental level. However, it follows from this fact that it may be tempting to categorize WE as a theory of *epistemic* emergence. Indeed, a recent paper by Guay and Sartenaer[11] proposes a useful three-dimensional category system for theories of emergence; Bedau's theory is categorized as such. Nonetheless, it is clearly not Bedau's original intention.

Bedau gives two different formal definitions of WE[12,13]. The two definitions are claimed to be

---

[4] Jaegwon Kim,1998.

[5] Bedau, 1997, 377.

[6] O'Connor, 1994.

[7] Bedau, 1997, 377.

[8] Bedau, 1997, 376.

[9] Lloyd Morgan, 1923.

[10] C. D. Broad, 1925.

[11] Guay and Sartenaer, 2016.

[12] Bedau, 1997.

"essentially equivalent". This paper discusses the more recent definition.

So, the distinguishing feature of Bedau's WE concept in contrast to merely epistemic emergence is that, according to it, emergence is not just in the mind, but "is an objective phenomenon that exists in nature"[14]. This means that Bedau tries to create a middle ground and break the epistemic-ontological dichotomy. He tries to show an account of emergence that can be as objectively evaluated as physical features, so it is not just epistemic and, at the same time, does not break with materialism entirely.[15,16]

It can be said that Bedau's project is very similar to those of the early emergentists. Not in its content, but in its aim to break dilemmas. The early emergentists sought for a middle ground between substance monism, that is, strict materialism and dualism, whereas Bedau tries to find a way between strict materialism and strong emergence. Perhaps it is worth to note that his ideas are not entirely new because Broad himself defines his position as "emergent materialism" between strict materialism and such a strong kind of emergentism that was held by Samuel Alexander[17] and Morgan[18]. The idea that emergence involves a special kind of reduction is also not entirely new; for example, Broad distances himself from Alexander just because, according to Alexander, mind and other emergent phenomena in a sense can be reduced[19].

---

[13] Bedau, 2008.

[14] Bedau, 2008, 457.

[15] Because of this, in the Guay-Sartenaer system mentioned above WE is labeled as “strong”. But this is in a different terminological framework than Bedau’s, who labels his system as “weak” in the sense that it is not about ontological emergence.

[16] Not everyone agrees with this interpretation of Bedau’s project which suggests that Bedau’s definition of WE is not unambiguous. Pexton (2016) argues that Bedau, by proposing an objective measure for emergence, creates a non-epistemic account. This is because Pexton associates objectivity with the Laplace demon and concludes that if Bedau’s criteria is objective then it should not rely on an observer’s point of view. While this is a reasonable definition of objectivity, it would make Bedau’s account inconsistent (because it implies an observer). By applying the principle of charity we do not associate objectivity with the Laplace demon while analyzing Bedau’s account.

[17] Samuel Alexander, 1920.

[18] Morgan, 1923.

[19] Broad 1925, 639.

Bedau's definition of WE is as follows:
"If P is a macro-property of some system S, then P is weakly emergent if and only if P is generatively explainable from all of S's prior micro-facts but only in an *incompressible* way"[20].

By generative explanation Bedau means that the macro-property in question can be "exactly and correctly" shown to be produced based on the "micro-causal dynamics" and "earlier micro-states and boundary conditions" of the system in either a discrete or a continuous temporal process. Bedau also points out that this kind of explanation implies that the macro-property in question at any given point of time is the aggregate or sum of the micro-properties of the system that exist at the same time. This is what he calls *synchronic* reduction of macro to micro. In other words, there is *no emergence* of macro-properties over the micro-properties, realizing them; rather, emergence occurs in the process of state changes in the system.

In Sartenaer's[21] account of synchronic vs. diachronic emergence, Bedau's is a *diachronic* theory of emergence. It is indeed categorized as such by Guay and Sartenaer[22]. It is worth noting that even this part of the WE concept presupposes there is no strong emergence, in line with Bedau's intentions, of course. Moreover, this means a major departure from an old terminology of emergence, as it is strongly associated with irreducibility – either because the entity in question is more than its parts (strong version) or at least because of epistemic constraints (weak version). The fact that everything *is* reducible in Bedau's framework makes WE similar to *emergent computing*, a concept that also does not usually entail any kind of irreducibility.

Indeed, from Bedau's two articles on the subject, it is evident that this explanation is easiest to imagine as a computer simulation. He emphasizes, though, that it *does not have to be* a computer simulation (as with continuous/analog phenomena, it cannot be, because computers work in discrete steps).

Bedau states that by iterating over a system's micro-interactions we are "crawling the micro-causal web". Now, if a system's state in T can be explained without loss only if we do not skip any in-between states, then the explanation is *incompressible*, otherwise, it is *compressible*.
The most authentic way to illustrate WE is using the Game of Life (GOL), just like Bedau

[20] Bedau 2008, 445, emphasis added.

[21] Sartenaer’s, 2015.

[22] Guay and Sartenaer, 2016.

does[23].

GOL is a well-known "game" – actually a cellular automaton – that takes place in an infinite two-dimensional universe of a regular grid, in which each cell is either "dead" or "alive". There are four simple rules of state transformation in this universe:

1. Any alive cell with fewer than two alive neighbors dies, as if caused by underpopulation.
2. Any alive cell with more than three alive neighbors dies, as if by overcrowding.
3. Any alive cell with two or three alive neighbors lives, unchanged, to the next generation.
4. Any dead cell with exactly three alive neighbor cells will come to life.[24]

Started in certain initial configurations, GOL can feature interesting, ever-changing patterns, based only on these four simple rules. What does it mean for a state in GOL to be an instance of WE? It means that the state in question cannot be explained without calculating all the intermediate states. As was shown above, Bedau proposes that this is an objectively evaluable property of the GOL states and other phenomena. Phrasing it in another way, the claim is that there is no more compact explanation to these examples than just simulating all of their steps from the beginning or "crawling the micro-causal web" – which in this case is equivalent to playing GOL. Bedau even adds – as a means to emphasizing the objectivity of WE – that the Laplace demon would have to do the same[25].

3. The Weaknesses of Incompressible Explanation

3.1. Problems with Incompressibility

The main weakness of an incompressible explanation is that its "generative" incompressibility depends not only on the "available micro-facts" of S but also on the operators that can be used in the simulation or explanation process. Bedau's definition presupposes a given set of operators for a generative explanation, which seems to be assumed to be equal to the rules of GOL itself. However, as it will be explained below, this assumption leads to the absurd situation in which all new developments in a GOL world can count as a WE macro-property. However, in the other horn of the dilemma, lifting this restriction on the usable operators by

---

[23] Bedau, 1997.

[24] The rules are quoted from http://conwaylife.com/, where additional information and examples can be found.

[25] Bedau, 1997.

allowing non-GOL operators in the explanation will lead to the situation in which different operator bases yield different judgments about whether or not some phenomenon is weakly emergent. That, in turn, is in contradiction with the claim that WE is scientifically useful.

To illustrate this, let us bring here Bedau's examples of GOL patterns called "R pentomino" and "gliders"[26]. The R pentomino is capable of featuring 1103 non-repeating states of growth and change, after which it stabilizes. Gliders are patterns capable of traveling on the GOL grid. A simple glider is shown in Figure 1.

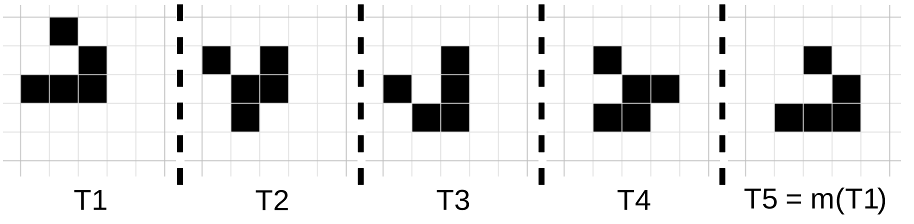


Fig. 1 A full cycle of the simple "Glider" Game of Life pattern. All figures were created using Golly and Inkscape.

Bedau claims that the stable end state of an R pentomino is weakly emergent because there is no compressible explanation for how it emerged from its initial configuration. Gliders up to the point where they do not repeat their earlier states should also be considered as weakly emergent.

To explain the problem with the latter claim, the following example shows how any state of the glider can be explained from any previous state in one step (that is, in a compressed manner) by using equations that can calculate the progress and shape changes of the glider. Waiting for the first repetition is not necessary.

The following equation system, together with the Glider[T] lookup table on Fig 2. can be used to calculate any previous or future state of a Game of Life glider when it is observed at step T in any of its four states. These equations, if they are known a priori, enable Laplace demon-like calculations for the world of Game of Life, just like the complete knowledge of the laws of nature enable the Laplace demon to calculate the past and future of nature at any time by simply relying on arithmetic. In other words, these equations make it possible to explain *without* the crawling of the causal web and thus indicating a non-WE case.

The equations are formulated to calculate any past state leading to a current state - as this is

[26] Bedau, 1997.

the usual direction of explanation. When encountering a pattern of the glider at iteration T, we can get any previous state of that glider in arbitrary GOL iterations ago with a single arithmetic calculation. To get the state 10000 GOL iterations before the current state, we assume the current state to be T=10000 and then Pattern_0, Bounding box x_0, and Bounding box y_0 will yield the pattern and its relative x and y positions (respectively) 10000 cycles ago. Naturally, with trivial rearrangements, these equations can be used to predict the future of a glider as well.

$$\text{Pattern}_0 = \text{Glider}[T mod 4]$$

$$\text{Bounding box } x_0 = x_T -$$

$$\text{Bounding box } y_0 = y_T - \left\{ \lfloor \frac{T}{4} \rfloor + \lceil \frac{\lfloor T mod 4 \rfloor}{3} \rceil \right\}$$

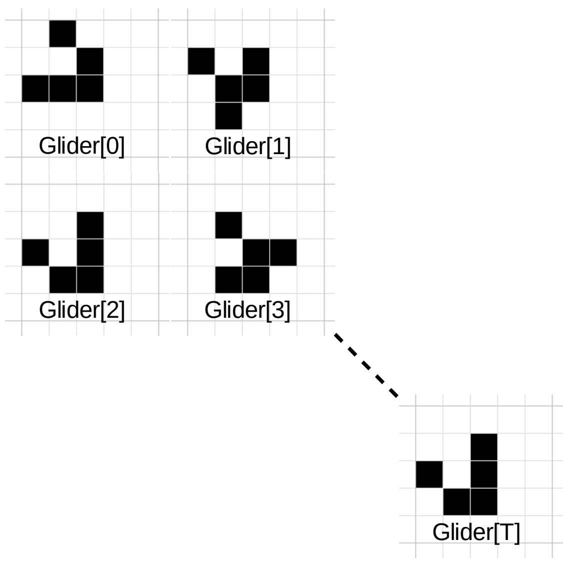


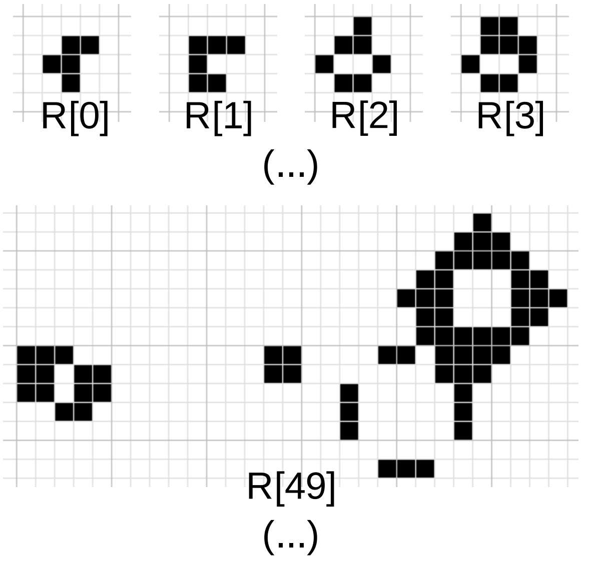
Fig. 2 By knowing the transformation cycle of the glider, any instance of the glider pattern at time T_x can be explained without "crawling the micro-causal web".

Again, this explanation does not require the "crawling of the micro-causal web".

There is no fundamental difference between an R pentomino and a glider in this respect. Since R pentomino is a known pattern, its evolution can be recorded as a sequence of 1000-odd GOL states. Using this array as a lookup table (see Figure 3), we can always identify an R pentomino if we encounter its initial state or any of its intermediate states. This knowledge, in turn, means that we do not have to run a simulation to find out how the particular GOL world evolved into its current state. In fact, we do not have to crawl the micro-causal web to show how the state came about from an initial configuration because it can be calculated similarly as was shown in the simpler case of the glider.[27]

[27] This solution for the explanation of R-pentomino is not avoidable by adjusting the technicalities of the definition of WE. As Paul Hovda, 2008, pointed out after analyzing Bedau's WE concept, "(…) s-derivability is relative to a derivation system. For example, one can always concoct, ad hoc, a derivation system that will derive anything in one step: add it as an axiom." (S-derivability is a term Bedau used before he switched to *incompressible explanation* in his 2008 *Minds and Machines* article). In Hovda's terminological framework our solutions above for explaining the glider and the R-pentomino count as "adding an axiom" (Hovda, 2008, 463) to the knowledge base. Hovda then goes on to propose a formalism for GOL that excludes such axioms, and at the same time makes the quantification of emergence possible. In his solution, weak emergence depends on the "amount of simulation" (p. 468). While Hovda's work provides valuable contribution to an important problem, it does not provide a profound argument about why the given formalism should be employed and not another. Hovda points out various advantages of his formalism, i.e., that it provides a canonical description for any GOL state, it is "fairly minimal" (p. 464), and that it makes possible the measurement of WE on a 1-to-infinite scale. But it fails to reinforce Bedau's framework because it is particular to GOL, and Bedau's WE concept is meant to be more general. However, the paper serves as a great reminder about the amount of work required to establish a formalism for actually defining what "crawling the causal web" means in a single case application. Bedau does not have a proposal about the knowledge base an observer should possess before judging whether or not something is WE.

Fig. 3 An "album" of the states of R pentomino (denoted as R, parts of the sequence is omitted).



An argument against our criticism of the WE condition above could be to insist that our equations are mere trickery – that is, no arbitrary operators or functions should be allowed to be added to the derivation system. In the case of GOL, only the state-change rules of GOL itself should be used. However, that leads to an even bigger problem: namely, that this way any GOL pattern will qualify as WE at any iteration, because within the world of GOL simply there is no way to compress their generation from earlier states. For instance, one of the simplest GOL patterns is 2x2 tile-size rectangle of alive cells surrounded by dead cells. This is a completely stable, non-changing pattern that is constant over any number of iterations. When armed only with the rules of GOL however, there is no way to explain the pattern of the 2x2 block from any previous pattern without "crawling the micro-causal web", that is, actually running a simulation due to the four simple rules of GOL. Therefore, the block and any other pattern qualify as WE, which makes WE a completely non-discriminating criterion.

3.2. An Alternative Interpretation of Weak Emergence

Our criticism might be countered by an adjustment of Bedau's theory that replaces the role of the incompressible explanation of phenomena with computational complexity. One example of this is Huneman's[28] work. In his article in Philosophy of Science Huneman, starting from

[28] Huneman, 2008.

Bedau's 1997 formulation of WE ("*a state of a computation process is weakly emergent [if and only if] there is no shorthand to get to it except by running the simulation*") proposes an alternative description of the emergence problem: "*weak emergence defined as inaccessibility [of a future state of an automaton] except by simulation*" (599). By this, he means that the general problem of predicting the future states of a subset of automata is in the PSPACE complexity class. According to the current state of mathematics[29], to predict these systems, there is no practical way other than actually running them since *"[this kind of] the problem cannot be solved in polynomial time*".

Huneman appears to be unaware of the fact that GOL is no such automaton: it has been proved that the computation of the future state of GOL cells is P-complete, hence polynomial[30]. That is not a serious issue for his argument, however, since it only means that Bedau was mistaken about his original GOL examples, but other systems could be WE. Alternatively, it could be claimed that P-completeness is just as an objective, mathematical property, like being PSPACE complete.

A more serious problem with Huneman's argument is about the claim that these systems can only be predicted by simulation. Shortcuts are still available by adding the kind of formulation to the operators of explanation we presented for the glider above. Of course, this amounts to building in a priori knowledge into our explanatory apparatus that works only for a certain subset of cases. But disallowing such moves is worse because of the following.

Huneman's claim about incompressibility is only true if we forbid these kinds of axioms and rely only on the automaton's own rules. This is the only sense in which Buss et al. show that certain automatons are PSPACE complete. Of course, this assumption is commonplace in complexity theory since it allows a characterization of a whole class of problems. But the result of applying this constraint to weak emergence is that any future state, even those of the GOL 2x2 block, are emergent, which again makes WE completely non-discriminative.

Huneman argues that his complexity-focused interpretation of WE meets "scientific-adequacy requirements" quite well. This seems to be false. To see this, we need not forget what the original intent of WE was: to meaningfully discriminate emergent and non-emergent natural phenomena, thus are useful for natural science. Huneman's approach could only say that

---

[29] Buss et al., 1992.

[30] Greenlaw et. al., 1995. This also means that for any given n cells it is the same order of complexity, as the number of single-cell calculations necessary is directly proportional with the input size.

nature *as a whole*, if to be simulated by a computer, only fits a certain complexity class, let's say, NP-complete. Game of Life and its rules was meant to be analogous with nature and its laws, after all. But a claim about the computational complexity assignable to nature cannot discriminate between *particular* emergent and non-emergent phenomena *within nature*.

3.3. Measures of Complexity and Weak Emergence

In yet another approach – not compatible with Huneman's – one might propose that the simulation apparatus should "learn" the already simulated state changes so that encountering them the second time, simulation can be spared. The result of this is contra-intuitive, as for example, the second cycle of a glider will not be WE since from the point when the glider returns to its initial state, the observer has all the necessary information to explain any other states of the lone glider without running a simulation (see Figure 4).

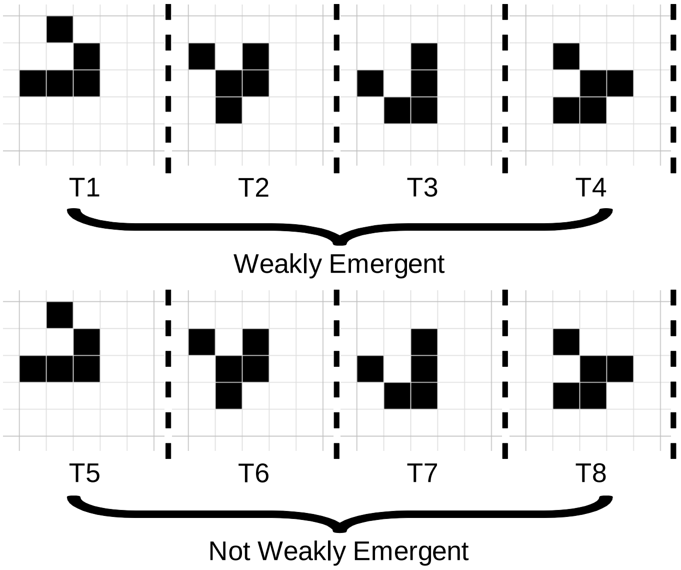


Fig. 4 Having zero a priori knowledge, even a simple glider qualifies as WE. But after the 5th step, there is enough information to explain any later states without crawling the micro-causal web; therefore, these states are not WE.

Another way out of this problem would be suggesting that the explanation method proposed above for the R pentomino or the glider is just as complex as running the simulation itself. However, this is not the case. Mathematically, our construct is just a function that returns different pre-defined values for the integers 1-1103, and then the same state for the 1104-infinity interval. Its description might be large *quantitatively*, but then it is questionable how that should make a difference for the WE condition. This invokes the fact that is relying on algorithmic complexity that was devised by Kolmogorov[31], Solomonoff[32] and Chaitin[33]. Let us rely on the specific formulation of the concept called Kolmogorov complexity which might be a way forward for the WE program. Kolmogorov complexity of an object (in our case a GOL pattern) is the length of the shortest program that produces that object in a given universal machine language. This mathematical tool appears to be slightly more useful than the complexity of formal languages that represent problem classes discussed before (these classes are designated with PSPACE, NP, etc.), because Kolmogorov complexity can easily be applied to *single instances* of a problem not only to the whole class. Relying on this measure, it might be required from a phenomenon that any kind of explanation should be no less Kolmogorov complex than its "causal web". In the case of GOL, that would mean that complexity of the causal generation of the phenomena by the four rules of the "world" is already Kolmogorov-minimal, and there is no less complex solution available (which is not the case with GOL as shown above). This would be in line with Bedau's intuition that the Laplace demon would have to do the same (as long as it is restricted to the Universal Turing Machine or the Universal Turing Machine itself is proven to be the most efficient concept for computing).

Weinstein[34] interprets Dennett's[35] idea of "Real Patterns" as an application of the idea of algorithmic complexity. While Dennett, who precedes Bedau in using GOL as a means of explaining his philosophical ideas, does not directly cite Kolmogorov or Solomonoff, but he does cite a related Chaitin[36] paper on randomness. This makes Weinstein's Dennett interpretation very plausible since the kind of randomness is used in connection with algorithms only. In Dennett's related definition of real patterns "[a] pattern (…)—is real—if there is a description of the data that is more efficient than the bit map (…)" which in terms of Kolmogorov complexity means that the description of the given pattern is longer than its minimal description, the length of which being the Kolmogorov-complexity of the pattern. Or simply put, a pattern is real if it is *compressible*. As pointed out above, the R-pentonimo and the glider are compressible. A piece of evidence is our construct for the glider above. In the case of the R-pentomino, a simple program simulating a GOL world plus the initial

---

[31] Kolmogorov, 1963.

[32] Solomonoff, 1964.

[33] Chaitin, 1966.

[34] Weinstein, 2003.

[35] Dennett, 1991.

[36] Chaitin, 1975.

configuration of the pattern would be way shorter than a sequential enumeration of its 1000-odd states. This means that both are *real* by Dennett's definition, the intuition behind which is that *being real is being non-random*. This seems to be the opposite intuition of incompressibility, and therefore Dennett's *real patterns* are exactly the ones that are not *weakly emergent* in Bedau's framework, which shows just how arbitrary incompressibility as an emergence criterion is.

Also, it is questionable whether algorithmic complexity would be a meaningful condition at all, since it would just show that some phenomena could have been generated in a less complex manner than they actually were. At any rate, nothing can be said about non-discrete time events or non-deterministic events – that is, anything not digital. And this is a problem since the whole effort was to facilitate science, which currently deals with a large number of phenomena of interest that is analog.

3.4. On the Level of Explanation

The problems around WE in part originate from the usage of the term *explanation* in the definition. Explanation involves a priori knowledge base and a context, both of which help with the right interpretation of the phenomenon the observer encounters. In WE, *explanation* is used in a special way: *closer to computation or generation* than *interpretation and understanding*. Yet, as we have seen above, this explanation concept is still not independent from the observer's prior knowledge.

This problem is not limited to the Game of Life patterns. Let us take a simple chemical reaction that can be explained by relying on natural laws. Knowing these laws, the phenomenon is not WE since the result of the reaction can be derived from the laws and the initial parameters. For an observer *not* knowing the laws, that is not the case. Again, the intuition behind WE dictates that it does not matter if the observer *knows* the laws or not. It only matters that there *is* such a law; therefore, it is an objectively evaluable condition. But this interpretation of WE leads to the following, even deeper problem.

The problem is that any regularly reproducible phenomenon can be explained by some shortcut, and this violates the formal condition of WE. For instance, a chemical reaction can be explained by pointing out that the phenomenon is an example of the same reaction we experienced before. This explanation does not draw on any knowledge of the underlying processes. This, of course, is a shallow, local explanation. But if the *level* of explanation matters to WE, then computers running GOL or GOL patterns are to be explained on the *physical* level. As Bedau states, macro-properties are the sum of the micro-properties of the underlying system.

Thus, a GOL pattern is eventually a sum of material particles, and that raises the question of whether or not it can be WE with regards to the micro-causal web of the underlying *physical* world. It is possible that while a GOL pattern might be WE with regards to the GOL rules and initial states, it is not WE if we explain it with the laws of physics to which the actual computer producing the pattern is claimed to be reducible by the theory of WE.

If synchronic reduction of a GOL pattern to material particles would not be possible, then a GOL pattern would exceed the limits of the material world, and that would point towards an account of strong emergence. This, however, is in conflict with the goals of Bedau's WE program. Therefore, WE, in its current formulation, does not provide a firm middle ground between strict monism and strong emergence. Instead, it is rather similar to the earlier concepts of epistemic(-only) emergence.

3.5. On Alternative Interpretations of WE

A possible objection to our arguments is that it is possible to interpret the original definition "*If P is a macro-property of some system S, then P is weakly emergent if and only if P is generatively explainable from all of S's prior micro-facts but only in an incompressible way*," in a way that those micro-facts to themselves constitute a fundamental set of operators.[37]

If this stricter approach, which equates the micro-facts with the fundamental set of operators, were reasonable, our argument would be rendered irrelevant. We select our operators seemingly arbitrarily and therefore make incompressibility dependent on the construction of the system. The counter-argument is that in doing this we simply abandon how the actual system works on the micro-level, and this unfaithful model then proves nothing about WE.

From this strict operator-interpretation then it is possible to derive that Bedau's version of weak emergence is a dynamic version that shows how a large number of complex micro-entities generate macro-entities dynamically. This requires us to assume we have a "correct" model of the phenomenon, based on which we can establish the operators.

This interpretation, however, requires such ontological claims that defeat the purpose of WE. Hence, this is neither charitable nor a reasonable understanding of the definition of the WE.

The question of operators, which is not mentioned in the WE definition itself, arises from the

[37] An earlier version of this paper received a couple of reviews those pointed to this issue, helping us further clarify our objection to WE. We are immensely grateful for all of these comments.

fact that "incompressible generative explanation" requires simulation, or in more general terms, computation to be carried out. The WE definition claims that running such a simulation will establish whether the phenomena is weakly emergent or not. And since one of the stated motivations of WE is to create an "ontologically innocent" that is a non-assuming account of emergence, this very test is supposed to establish whether we can talk about a "micro-level" and a "macro-level" of actual phenomena in the first place. If the test passes, we make this distinction.

Therefore, any reference to a presupposed arrangement between micro-and macro-levels of a given phenomenon is a serious blow to WE. It makes the whole argument circular: the test of incompressibility should be carried out using only those fundamental operators that express the causal relations of the micro-level faithfully, and its result will tell us whether there is a micro-level to speak about. As the existence of ontological levels is ruled out from the outset – this is what makes WE ontologically innocent – any separation of micro and macro is in an epistemic sense only. The stricter interpretation still demands one given understanding of what is micro and macro while not allowing for these to be real in an ontological sense.

We admittedly put a large burden on WE as we expect it to be a way of establishing that we can fruitfully talk about micro-level and macro-level in the context of a given phenomenon. But not putting this burden on WE is worse in our view, since then the correct understanding of the micro-and macro-levels must come from a source that is independent of the WE test. However, barring a reference to some real, self-evident leveling of the natural phenomenon, this always will be a question of interpretation. Our Glider and Pentomino generating time-dependent functions are not arbitrary: they generate the exact same patterns as the four-game of life rules would. And they are not "higher-level" or more "aggregate" either, in that they are all functions over a grid unless we can fill these claims with content.

An approach for doing this in the context of WE is a reference to non-linearity. Non-linearity was introduced into the emergence debate by the British Emergentists and was used as a telltale sign of metaphysical emergence (Wilson 2013). In itself, non-linearity is a shape of a function between two variables, and for the British Emergentists, it meant a general non-additivity of some feature. The issue with non-linearity is that using the mathematical definition, we will merely find any models that have many linear relations in them. For instance, GOL is not linear. Moreover, the four basic rules of GOL are just as non-linear as our Glider and Pentonimo generating functions; the state of a cell, in general, is not a linear function of time (or any other variable that naturally lends itself). And even worse, it is just a matter of interpretation: we could pick an individual cell that is not in the path of a glider, and the status of that cell is constant "dead," which is in linear relation with time. Or, arguably, our Glider function is linear in the number of alive cells vs. time (a constant 5, which would be a horizontal line on a chart). Unless linearity is redefined as something more than its mathematical meaning, it is unclear how it may help in WE.

4. Conclusion

The purpose of this paper was to provide a mainly internal criticism for Bedau's concept of Weak Emergence. WE is based on the concept of incompressible explanation. Several problems were shown with this concept.

One problem stems from the meaning of *incompressibility.* Even though it is not apparent from Bedau's definition, a priori knowledge or the selection of operators by the observer cannot be eliminated. It has been shown in the previous section that the crawling of the micro-causal – in the sense as proposed by Bedau – can always be avoided, as long as the process is deterministic[38] – like in the case of Game of Life – by creating a knowledge base that contains a record of state-changes.

This means that WE is dependent upon such a knowledge base or choice of operators and thus completely relative; if we allow a priori knowledge, different investigators will qualify the same phenomena differently. The process might be performed objectively in the sense that the investigators follow explicit rules, but the starting assumptions are subjective.
If to prevent the above problem, WE is to be understood without such a priori knowledge – in the GOL world; this could mean that only the basic rules of GoL are to be used – then the first encounter of very simple – intuitively, not weakly emergent – phenomena will be qualified as WE, and that signifies a serious problem because in contrast to Bedau's intention every macro-phenomenon becomes emergent at first sight but then some of them becomes non-emergent. A third option would be to establish some middle ground between these two extreme positions with regard to the applicable operators of an explanation – there appears to be no objective way right now.

Another problem is the *level* of explanation. In Bedau's system, weak emergence depends on whether or not the micro-causal web has to be crawled in order to explain the macro-state.

[38] In the case if the process is *not* deterministic, e.g., there is a quantum effect that shapes the output of the process, then Bedau proposes that WE is only applicable if complete information is known about the indeterministic micro-state changes. That in effect converts the problem back to deterministic, but it is unclear how such complete information could be attainable.

However, this micro-causal web cannot be identified unambiguously at all. What we see as the relevant micro-causal web depends on our knowledge and choices. In the case of Game of Life, it might be the level of the grid and the four rules of GoL, but also the level of physics of the material world in which it all takes place. In Bedau's intention, weak emergence is consistent with materialism; therefore, all phenomena must be material.

The goal of Bedau's weak emergence program was to provide an objective, scientifically useful emergence concept that does not carry the controversy of strong emergence accounts that propose the emergent phenomenon is more than the sum of its parts in an ontological sense. But he also wanted emergence to be more than a merely epistemic property. Weak emergence, while its aim is valid, does not achieve these goals in its current state for the reasons above.

Acknowledgments

This paper was funded by the Bolyai János scholarship of the Hungarian Academy of Sciences.

References

Alexander, S. (1920), *Space Time and Deity*. London: MacMillan and Co.

Bedau, M. A. (1997), "Weak Emergence", *Philosophical Perspectives 11*. 375–379.

Bedau, M. A. (2008), "Is Weak Emergence Just in the Mind?" *Minds and Machines*, special issue on dynamical emergence and computation. 18. 443–459.

Broad, C. D. (1925), *The Mind and its Place in Nature*. New York: Routledge.

Brunner, K. (2002), "What's Emergent in Emergent Computing?" in: Trappl, R.: EMCSR 2002, *Proc. European Meeting on Cybernetics and Systems Research*, Vienna, April 2002, pp. 189–192.

Buss, S., C. Papadimitriou, and J. Tsisiklis (1992), "On the Predictability of Coupled Automata: An Allegory about Chaos", *Complex Systems* 5(5): 525–539.

Cariani, P. (1991), emergence and artificial life. *Artificial life* II, 10, 775-798.

Chaitin, G. (1966), "On the Length of Programs for Computing Binary Sequences", *Journal of the ACM* 13: 547 – 569.

Chaitin, G. J. (1975), "Randomness and mathematical proof". *Scientific American*, 232(5), 47-52.

Dennett, Daniel. (1991), "Real Patterns", *Journal of Philosophy* 88: 27 – 51.

Greenlaw, R., Hoover, H. J., & Ruzzo, W. L. (1995), *Limits to parallel computation: P-completeness theory*. Oxford University Press.

Guay, A., & Sartenaer, O. (2016), "A new look at emergence. Or when after is different." *European Journal for Philosophy of Science*, 6(2), 297-322.

Hordijk, W., Crutchfield, J. P., & Mitchell, M. (1998), "Mechanisms of emergent computation in cellular automata", In *International Conference on Parallel Problem Solving from Nature* (pp. 613-622). Springer Berlin Heidelberg.

Hovda, Paul. (2008), "Quantifying Weak Emergence", *Minds and Machines* special issue on dynamical emergence and computation 18. 461-473.

Huneman, P. (2008), "Emergence made ontological? computational versus combinatorial approaches", *Philosophy of Science*, 75(5), 595-607.

Kim, J. (1998), *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press.

Kim, J. (2002), "Mental Causation and Consciousness: The Two Mind-Body Problem for the Physicalist", In *Physicalism and Its Discontents*. Eds:. Carl Gillet and Barry Loewer. 271–283. Cambridge: Cambridge University Press.

Kolmogorov, A. (1963), "On tables of random numbers", *Sankhyā,* 25. 369–376.

McLaughlin, Brian P. (1992), "The Rise and Fall of British Emergentism", In *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Eds.: Ansgar Beckermann, Hans Flohr and Jaegwon Kim. Berlin, New York, Walter de Guryter. 49-93.

Morgan, C. L. (1923), *Emergent Evolution*. London: Williams and Norgate.

O'Connor, T. (1994), "Emergent Properties", *American Philosophical Quarterly,* 31. 91–104.

Pexton, M. (2016), "Emergence and interacting hierarchies in shock physics", *European Journal for Philosophy of Science*, 6(1), 91-122.

Sartenaer, O. (2015), "Synchronic vs. diachronic emergence: a reappraisal", *European Journal for Philosophy of Science*, 5(1), 31-54.

Solomonoff, Ray (1964), "A Formal Theory of Inductive Inference", *Information and Control* 7: 1 – 22, 224 – 254.

Weinstein, S. (2003), "Objectivity, information, and Maxwell's demon", *Philosophy of Science*, 70(5), 1245-1255.

Wilson, J. M. (2013). Non-linearity and metaphysical emergence. In Stephen Mumford & Matthew Tugby (eds.), Metaphysics and Science (2013).