

“SUPERVENIENT AND YET NOT DEDUCIBLE”: IS THERE A COHERENT CONCEPT OF ONTOLOGICAL EMERGENCE?*

JAEGWON KIM
Brown University

Abstract

Formulating a concept of emergence that is intelligible and *prima facie* coherent is a significant issue not only because emergence concepts continue to proliferate, attracting a great deal of positive attention from scientists and philosophers, but also because the idea of emergence is closely related to some of the concepts of central importance in the current debates on the mind-body problem. Most early emergence theorists, like C.D. Broad and C. Lloyd Morgan, intended emergence to be an objective phenomenon in the world and considered emergent properties as real and causally potent characteristics of objects and events of this world. This classic conception of emergence, now called “ontological” or “metaphysical”, or “strong”, is standardly contrasted with an “epistemological”, or “weak”, conception according to which properties are emergent in case they are “surprising” or “unexpected”, or unpredictable and unknowable from information concerning base-level phenomena. But what is ontological emergence? On Broad’s characterization, shared by a number of other writers, ontologically emergent properties are properties that are *determined by, or supervenient on, their base-level conditions and yet not deducible from them*. This paper explores some issues arising from the notion of ontological emergence so conceived, and uncovers what appears to be a possibly damaging incoherence. This raises the question whether there is a workable notion of ontological emergence.

I

C. Lloyd Morgan, one of the leading British emergentists of the early 20th century, describes the “emergent evolution” of the world, or how we got where we are and where we are headed from here, in these words:

“From [the ultimate basal phenomenon, space-time] first emerged ‘matter’ with its primary, and, at a later stage, its secondary qualities. Here new relations, other than those which are spatio-temporal supervene.¹ So far, thus supervenient on spatio-temporal events, we have also physical and chemical events in progressively ascending grades. Later in evolutionary sequence life emerges – a new ‘quality’ of certain material or physico-chemical systems with supervenient vital relations hitherto not in being. Here again there are progressively ascending grades. Then within this organic matrix, or some highly differentiated part thereof, already ‘qualified’ ... by life, there emerges the higher quality of consciousness or mind. Here once more, there are progressively ascending grades. ... As mental evolution runs its course, there emerge ... ‘tertiary qualities’ – ideals of truth, of beauty, and of the ethically right... And beyond this, at or near the apex of the evolutionary pyramid of which space-time is the base, the quality of deity – the highest of all – emerges in us ...”²

It is plain that Morgan is presenting “emergent evolution” as an actual history of the world, though perhaps more than a little speculative. The sequential emergence of matter from space-time, life in inorganic systems, and mind from biological processes is a historical fact about this world. For Morgan, as well as many other emergentists, emergent phenomena like life and mind are genuinely novel features of reality which make their distinctive causal contributions to the subsequent evolution of the world. There are bridges and building, works of art and electronic gadgets, nuclear bombs and ozone holes, because minds and consciousness have emerged. Emergence is an objective feature of the world, with powers to change,

¹ As Brian McLaughlin and Karen Bennett note in their entry “Supervenience” in the *Stanford Encyclopedia of Philosophy* (<http://plato.stanford.edu/>), Morgan is using “supervene” and “supervenient” in this paragraph in their vernacular sense, something like “occur later” or “follow upon”, rather than in the philosophical sense now commonly associated with these terms (as in the title of this paper).

² C. Lloyd Morgan, *Emergent Evolution* (London: Williams and Norgate, 1923), pp.9–10. Here Morgan represents himself as reporting Samuel Alexander’s views in *Space, Time, and Deity* (London: Macmillan, 1920). But there is little doubt that Morgan accepts the picture he presents.

create, and destroy. In the fullness of time, Morgan assures us, if things continue to go right, we humans will achieve divineness.³

In his classic *The Mind and Its Place in Nature*,⁴ C.D. Broad distinguished three types of theories concerning biological (“vital”) phenomena: “Biological Mechanism” (what we would now call physical reductionism), “Substantial Vitalism” (Broad had in mind Hans Driesch’s neo-vitalism which posited “entelechies” to account for biological phenomena), and his own “Emergent Vitalism” (which takes biological phenomena as emergent from, but not reducible to, physicochemical phenomena). These were clearly intended as three possible views making contrasting claims about the nature of biological organisms and their distinctively biological capacities, functions, and activities. These theories stake out disparate and mutually exclusionary positions on the nature of biological entities and processes. Unquestionably, the emergent nature of biological phenomena, for Broad, is an objective fact about them; it does not concern what anyone knows or believes about them.

This conception of emergence as an objectively real fact about the world is now standardly called “metaphysical” or “ontological” emergence; some call it “strong” emergence. This is contrasted with an “epistemological”, or “weak”, conception which seems more common among the burgeoning ranks of emergentists on the current scene, especially those from scientific fields. Unlike metaphysical emergence, the epistemological conception focuses on certain supposed epistemic aspects of emergent properties and phenomena, emphasizing such features as their novelty, unpredictability, and our inability to calculate, or “compute”, them from information concerning the basal conditions from which they emerge. The main point then is that we, as cognizers, cannot get there from here – that is, get to higher-level emergent phenomena from information about the lower-level base phenomena. On the metaphysical conception, it isn’t just that emergent phenomena are unpredictable for us; that may be more of a commentary on our cognitive powers than the phenomena themselves. More importantly, the point is that they are objectively new, extra additions to the ontology of the world. Their newness, or novelty, does not consist in their ability to “surprise” us, or our inability to “predict” them;

³ Most emergentists seem incorrigible optimists; they are blissfully unmindful of the glaring fact that if good things have emerged, so have many unspeakably bad and evil things!

⁴ C.D. Broad, *The Mind and Its Place in Nature* (London: Routledge and Kegan Paul, 1925), p.58.

rather, their newness is meant to be metaphysical in import: before these emergents came on the scene, there had been nothing like them in the world and they represent net additions to the world's furniture. And what could be the point of these extra entities if it isn't their bringing with them *new causal powers*, powers that go beyond the powers of the lower-level conditions in their emergence base?

I don't know when the distinction between the two types of emergence was explicitly recognized, although of course there have been weak and strong emergence theories and theorists for a long time. In their entry "Emergent Properties" in the *Stanford Encyclopedia of Philosophy* (<http://plato.stanford.edu/>), Timothy O'Connor and Hong Yu Wong say that the emergence concept in J.S. Mill and C.D. Broad was the strong ontological variety, whereas Samuel Alexander's emergence concept, though still metaphysical, had considerably weaker ontological significance⁵ (they liken Alexander's position to contemporary nonreductive physicalism). In his "Weak Emergence" (1997), Mark Bedau introduces an epistemological notion of emergence; and a distinction between ontological and epistemological emergence is explicitly drawn in Silberstein and McGeever's "In Search for Ontological Emergence" (1999). Following Silberstein and McGeever, Van Gulick recognizes a similar distinction between epistemic and metaphysical emergence in his useful 2001 survey article on reduction and emergence. In "Strong and Weak Emergence" (2006), Chalmers' main theme, unsurprisingly, is the difference between the two kinds of emergence he distinguishes.⁶ This list is not intended to be exhaustive.⁷

⁵ See also Philip Clayton, "Conceptual Foundations of Emergence Theory", in *The Re-Emergence of Emergence* (Oxford: Oxford University Press, 2006), ed. Philip Clayton and Paul Davies.

⁶ Mark Bedau, "Weak Emergence", *Philosophical Perspectives* 11, pp.375–399; Michael Silberstein and John McGeever, "In Search for Ontological Emergence", *Philosophical Quarterly* 49 (1999): pp.182–200; Robert Van Gulick, "Reduction, Emergence and Other Recent Options on the Mind-Body Problem: A Philosophical Overview", *Journal of Consciousness Studies* 8 (2001) pp.1–34; David J. Chalmers, "Strong and Weak Emergence", in *The Re-Emergence of Emergence*, ed. Philip Clayton and Paul Davies.

⁷ Carl Gillett distinguishes three concepts of emergence, "strong", "weak", and "ontological", in his interesting paper "The Varieties of Emergence: Their Purposes, Obligations and Importance", *Grazer Philosophische Studien* 65 (2002): pp.95–121. All of these seem to be metaphysical notions of emergence in our sense, and Gillett's distinction does not straightforwardly relate to the standard ontological/epistemological distinction.

II

We will assume that the metaphysical emergentist would want to include supervenience as a component of his emergence concept; that is, if a property emerges from a set of basal properties, it supervenes on the latter. To put it another way, if the same basal conditions recur, the emergent property will necessarily recur as well (we will return to the question what the sort of necessity is involved). C.D. Broad clearly recognized this; he writes:

“No doubt the properties of silver-chloride are completely determined by those of silver and of chlorine; in the sense that whenever you have a whole composed of these two elements in certain proportions and relations you have something with the characteristic properties of silver-chloride ... But the law connecting the properties of silver-chloride with those of silver and chlorine and with the structure of the compound is, so far as we know, an *unique* and *ultimate* law.”⁸

And again:

“And no amount of knowledge about how the constituents of a living body behave in isolation or in other and non-living wholes might suffice to enable us to predict the characteristic behavior of a living organism. This possibility is perfectly compatible with the view that the characteristic behaviour of a living body is completely determined by the nature and arrangement of the chemical compounds which compose it, in the sense that any whole which is composed of such compounds in such an arrangement will show vital behaviour ...”⁹

Supervenience, or “upward determination”, may ultimately turn out to be detrimental to the emergence program; and yet, without supervenience, it would be difficult to explain the “from” in “property P emerges *from* basal conditions C”.¹⁰ David Chalmers says that an emergent phenomenon “arises (in some sense) from”¹¹ lower-level phenomena, but says nothing further about just what sense of “arising from” is involved here. Superven-

⁸ *The Mind and Its Place in Nature*, pp.64–65.

⁹ *The Mind and Its Place in Nature*, pp.67–68.

¹⁰ For more on this, see my “Emergence: Core Ideas and Issues”, *Synthese* 151 (2006): pp.547–559. In his “From Supervenience to Superdupervenience” (*Mind* 102 (1993): pp.555–586), Terence Horgan has an interesting footnote (fn.7) in which he quotes Arthur Lovejoy as countenancing a form of emergence which excludes supervenience.

¹¹ “Strong and Weak Emergence”, p.244.

ience supplies a clear and robust sense to Chalmers' "arises from". We can explore whether some weaker relation might be able to serve the purpose,¹² but it is clear that if the emergent phenomenon occurs randomly when its purported basal conditions are realized, it would be difficult to make sense of the claim that the phenomenon emerges "from" these conditions, or that the lower-level phenomena are its "basal conditions". Besides, this would likely allow cases of emergence to proliferate beyond what even the most lavish and bountiful emergentist would want. In any case, it is quite certain that major early emergentists accepted the supervenience of the emergents as a condition of emergence.¹³

What else do we need to characterize ontological emergence? Some writers (for example, Silberstein¹⁴) cite the capacity for "downward causation" – that is, the power to causally affect the events at the basal level – as a condition of ontological emergence. But I think this is not a wise move; downward causation is highly controversial and building it into the very concept of emergence will make it more difficult to defend the claim that there are any real cases of emergence. A better course would be to define emergence, or ontological emergence, in less contentious terms, postponing the issue of downward causation to be threshed out another day. The most important and widely used strategy, which goes back to Morgan and

¹² The idea of probabilistic supervenience is clearly coherent and deserves consideration. But we will presumably need stable lawlike probabilities grounded in the laws at the basal level. Details of this approach need to be worked out. I hope it is obvious that the condition that the basal conditions are merely "necessary" for the occurrence of an emergent property will not do.

¹³ It is interesting to note that several current philosophical advocates of emergence reject supervenience, for apparently different reasons, as a component of emergence; see, e.g., Paul Humphreys, "How Properties Emerge", *Philosophy of Science* 64 (1997): pp.1–17; Timothy O'Connor and Hong Yu Wong, "The Metaphysics of Emergence", *Noûs* 39 (2005): pp.658–678; Michael Silberstein, "In Defense of Ontological Emergence and Mental Causation", in *The Re-Emergence of Emergence*. In the current post-classical period of neo-emergentism, the idea of emergence has become very fluid, plastic, and variegated. Since "emergence" is a term of art, one is free to define it as one wishes, the only constraint being that the resulting concept is a philosophically or scientifically useful one. However, in rejecting supervenience, these writers are making a radical (and, in my view, unwise) departure from the core concept held by the classic British emergentists like C.D. Broad and C. Lloyd Morgan. I believe that one should stay at least in the vicinity of these writers in order to justify the claim that one is dealing with "emergence".

¹⁴ Michael Silberstein, "In Defense of Ontological Emergence and Mental Causation", p.203.

Broad, has been to add the condition that an emergent, even though supervenient on, and determined by, its basal conditions, is *not deducible* from them. When Broad first introduces the idea of emergence in *The Mind and Its Place in Nature*, this is what he says:

“Put in abstract terms the emergence theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in 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 of 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). The mechanistic theory rejects the last clause of this assertion.”¹⁵

So the difference between emergentism and its principal rival, mechanism/reductionism, consists precisely in that, on reductionism, “the characteristic properties” of a whole are *deducible* from the facts about its parts and their relationships, whereas emergentism holds them to be *not so deducible*. Or, to put it another way, let *F* be a property of a whole which is determined by, or supervenient on, properties and structural relations characterizing its constituents: *F* is emergent if and only if *F* is not deducible from these facts about its constituents.¹⁶ Emergentism, in its broadest form, would be the claim that there are properties like *F*, while reductionism would deny that such exist. (There are also specific emergentist theses concerning selected classes of phenomena; notably, biological phenomena in relation to physicochemical phenomena, and mentality and consciousness vis-à-vis the domain of the neural/biological sciences.) It is widely assumed that there is an important connection between logical deduction on one hand and explanation and reduction on the other. Thus, we might add: *F* is mechanistically, or reductively, explainable, or reducible, just in case *F* is deducible from facts about the constituents of the whole. In the quote above, Broad speaks of “the most complete knowledge” of the facts about the whole’s constituents, but this seemingly epistemological aspect of his

¹⁵ p.61. Emphasis added.

¹⁶ In speaking of “deducing” properties we follow the usual practice. When we say “*F* is deducible from conditions *C*” what is meant is that the fact, or proposition, that something has *F*, or that *F* is instantiated in something, is deducible from the proposition that conditions *C* hold for that thing. Similarly, when we speak about deducing a property from other properties.

characterization is easily eliminated: instead of “the most complete knowledge”, we can refer to “all the facts” or “a complete set of truths”. In fact, this replacement would be appropriate because whether or not anyone “knows” these facts about the constituents is irrelevant. Further, we should understand “all the facts”, or “the most complete knowledge”, to include *all the laws* operative at the basal level. So, when we say biological phenomena are emergent from physicochemical phenomena, the latter is understood to include physicochemical laws as well as individual facts at this level. In any case, a property’s nondeducibility from base-level facts is what separates emergent from nonemergent properties.

Broad is not alone in this. When C. Lloyd Morgan talked about nonemergent properties as “additive and subtractive only, and predictable”,¹⁷ he is naturally taken to be referring to something similar to deducibility. In a paper published in 1926, shortly after Morgan (1923) and Broad (1924), Stephen C. Pepper describes emergentism thus:

“The theory of emergence involves three propositions: (1) that there are levels of existence defined in terms of degree of integration; (2) that there are marks which distinguish these levels from one another over and above the degree of integration; (3) that it is *impossible to deduce* the marks of a higher level from those of a lower level ...”¹⁸

Later writers who have invoked nondeducibility as the pivotal criterion of emergence include Carl G. Hempel, Ernest Nagel, James Van Cleve, and David Chalmers.¹⁹

Broad, like many other writers, often resorts to epistemological terms to explain emergence. We have noted this in regard to Broad’s statement that “no amount of knowledge” about the micro-constitution of a living thing is sufficient for the “prediction” of the biological features of the organism. Similarly, an emergentist about consciousness would sometimes put his emergentist claim by saying that a complete physical, physiological, and computational knowledge of our brain at a given time does not

¹⁷ *Emergent Evolution*, p.3.

¹⁸ “Emergence”, *Journal of Philosophy* 23 (1926): pp.241–245. (The quote is from p.241; emphasis added).

¹⁹ Carl G. Hempel and Paul Oppenheim, “Studies in the Logic of Explanation”, *Philosophy of Science* 15 (1948): pp.135–175. Ernest Nagel, *The Structure of Science* (New York: Harcourt, Brace, and World, 1961). James Van Cleve, “Mind-Dust or Magic? Panpsychism versus Emergence”, *Philosophical Perspectives* 4 (1990): pp.215–226. Chalmers, “Strong and Weak Supervenience”.

suffice to give us any knowledge about our consciousness – whether or not we are conscious at the time and if we are, what sort of consciousness is being experienced. There is also Frank Jackson’s Mary,²⁰ the famous superstar vision scientist confined to a black-and-white room: we are to suppose that she has complete physical information about the physical/neural processes involved in the workings of our visual systems but, before her release from the room, she has no knowledge of color qualia. One could take such talk as referring to an epistemological conception of emergence, but I believe that would be premature. I think that the epistemological relationships being talked about are best explained in terms of deducibility. If, as Broad thought, physicochemical knowledge of an organism doesn’t yield knowledge of its biology, that would be so because biological truths are not deducible from physicochemical truths. If only truths about visual qualia were deducible from physical/neural truths, Mary could know, before her release, what it would be like to see a ripe tomato. In this way, metaphysical emergence characterized in terms of nondeducibility would appear to offer an explanation of epistemological emergence. We will soon see that the notion of deducibility is itself fraught with problems but, at least on a first pass, it seems like just what we need to characterize metaphysical emergence.

Deducibility can fail on two levels. First, there is the idea that the novelty of an emergent property consists precisely in the fact that the property is beyond our conceptual reach before it makes its first appearance and we have a chance to observe or experience it. The pre-release Mary seems often taken to be in a situation of that sort in regard to color qualia (though I don’t believe she has to be, for Jackson’s purposes). According to Thomas Nagel,²¹ our epistemic position vis-à-vis the experiences of a bat is precisely like that: we have no idea, no conception, of what a bat’s phenomenal experiences are like, and as a result we cannot even entertain propositions about their qualitative character. This would mean that we don’t even know just what propositions we should try to deduce about bat phenomenology from truths about bat physiology. Second, deducibility can fail even though we know what the supposed emergent properties are and know just what propositions are being considered for deduction. I believe a situation of this kind is what figures primarily in the emergentists’ claim that propositions involving emergent properties are not deducible from truths about the base-level processes. This is the sort of situation Broad

²⁰ “Epiphenomenal Qualia”, *Philosophical Quarterly* 32 (1982): pp.126–136.

²¹ “What Is It Like to Be a Bat?”, *Philosophical Review* 83 (1974): pp.435–450.

considers in regard to the deducibility of biological truths from physicochemical truths. The availability of biological concepts is not at issue.

To sum up, then, deducibility, or the absence thereof, is the key to the standard traditional conception of metaphysical emergence. Properties of a whole are emergent just in case they are not deducible from properties and relations characterizing its constituent parts, even though they are determined by and supervenient on them.

III

So then, do we now have a properly characterized ontological concept of emergence, a conception that makes emergence something objective in the world, not a phenomenon that has to do with our cognitive resources and powers? We have reached a concept that takes nondeducibility as the mark that distinguishes emergent properties from the rest. Now the critical question arises: Is deducibility, or nondeducibility, itself a wholly nonepistemic concept? If biological properties are emergent from physicochemical properties, we cannot deduce truths involving the former from those that only involve the latter – that is, we cannot deduce biological truths from physicochemical truths. But whom does this “we” refer to? Who is doing the deduction? How adept a logician is Jackson’s Mary supposed to be? No human person, we may assume, has unlimited logical powers. Although “we” cannot deduce biological truths from physical truths, why couldn’t a cognizer with vastly greater logical and mathematical powers produce the required deductions? Mustn’t we fix the level of deductive or logical competence we have in mind to give a clear meaning to “deducible”? If so, is there a “right” level to pick, and what makes it “right”? But, more importantly, don’t these questions show that the idea of deducibility, or nondeducibility, threatens to turn into an epistemic notion, making Broad’s emergence epistemic rather than metaphysical?

Think about how we go about making deductions – how we reason from premises to a conclusion in practice. I believe we have something like the following picture in mind: we start off with a list of premises, and proceed from there, step by step, where each step is seen as obviously and directly implied by selected earlier steps (in the best cases, in accordance with simple formal rules known to guarantee implication, like *modus ponens*), and, with luck, finally reach the proposition to be deduced. In his *Rules for the Direction of the Mind*, Descartes described deduction as “a

continuous and uninterrupted movement of thought in which each individual proposition is clearly intuited”.²² Further, he says, “the self-evidence and certainty of intuition is required not only for apprehending single propositions, but also for any train of reasoning whatever.”²³

More recently, Gilbert Harman has given a similar characterization of reasoning. According to him, we have a fundamental disposition, or power, to recognize “immediate implications” and “immediate inconsistencies”, and this capacity is what guides us through a process of reasoning. Harman recognizes that all this is relative to individual cognizers, saying “I suggested that certain implications and inconsistencies are ‘immediate’ for a given person.”²⁴

It is clear that the problem with this picture of deduction, or deducibility, is that it makes the notion of deducibility relative to the cognitive powers and dispositions of the deducers and thereby makes the concept of emergence defined in its terms both epistemic and relative, whereas what we are seeking is an objective, ontological conception of emergence. To his credit, Broad was well aware of this problem. Observant readers have surely noticed that in the last quoted passage above, he says that the emergent properties of a whole “cannot, *even in theory*, be deduced” (emphasis added) from those of its constituents. Clearly in a similar spirit, Chalmers writes that “truths concerning [emergent phenomena] are not *deducible even in principle*” from truths about lower-level phenomena.²⁵ But what does this mean? Here is where Broad summons his “mathematical archangel”:

“If the emergent theory of chemical compounds be true, a mathematical archangel, gifted with the further power of perceiving the microscopic structure of atoms as easily as we can perceive hay-stacks, could no more predict the behavior of silver or of chloride or the properties of silver-chloride without having observed samples of those substances than we can at present. And he could no more deduce the rest of the properties of a chemical element or compound from a selection of its properties than we can.”²⁶

²² *The Philosophical Writings of Descartes*, vol.1, tr. John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge: Cambridge University Press, 1985), p.15.

²³ *Ibid.* pp.14–15.

²⁴ *Change in View* (Cambridge: MIT Press, 1986), p.19.

²⁵ “Strong and Weak Emergence”, p.244. Italics added to “even in principle”.

²⁶ *The Mind and Its Place in Nature*, p.71.

The mathematical archangel,²⁷ we may presume, is logically and mathematically omniscient. If a proposition is deducible, “in principle” or “in theory”, from a set of premises, it will know that it is, and be able to construct a step-by-step proof. If the archangel cannot produce a proof, it’s because there is no proof and the proposition is not deducible, in an absolute sense, from the premises. Thus, if a property is emergent from a set of basal conditions, there is no deduction of it from those conditions; and our cognitive limitations have nothing to do with it. This idealization of deduction is Broad’s attempt to purge any epistemic and relativistic aspects from the notion of deduction and thereby objectify deducibility, or nondeducibility. For him, there being no deduction of an emergent property from its basal conditions is not an epistemological fact. It is not because we are not smart enough, or don’t have enough time or inclination, that we cannot deduce, say, geological truths from the truths of macroeconomics, or facts about the surface composition of the moon from facts about neurotransmitters in the human brain; not even the mathematical archangel can do that, and that is because there are no deductions between these sets of truths. This surely seems like an objective fact about the relationships between sets of truths, or facts. Broad’s idealization strategy appears to remove from the notion of deducibility an apparent epistemic relativity, and the characterization of emergence as supervenience *plus* nondeducibility appears to stand as an ontological conception of emergence. At least, so it may seem at this point.

IV

Thus, there being a deduction or proof is an objective matter independent of epistemological facts about us or anyone else. Does this solve the problem with Broad’s conception of emergence? There remains one more issue to deal with: deduction or proof makes sense only relative to a specific set of rules of inference, or a proof system, which specifies permissible trans-

²⁷ Achim Stephan calls the mathematical archangel “a colleague of the Laplacian demon” in his “Emergence – A Systematic View on its Historical Facets”, in *Emergence or Reduction?*, ed. Ansgar Beckermann, Hans Flohr, and Jaegwon Kim (Berlin: de Gruyter, 1992). For interesting discussion of some issues formulated in terms of Laplace’s demon that are relevant to our concerns in this paper, see Terence Horgan, “Supervenience and Cosmic Hermeneutics”, *Southern Journal of Philosophy* 22, Spindel Supplement (1984): pp.19–38.

formations of sentences in constructing proofs. That is to say, a sequence of sentences is a proof only relative to a system of deduction. Should we say, following Descartes and Harman, that the relevant system must include only those rules that give us “immediate implications”? No; that would bring back epistemic relativity. Evidently, what the mathematical archangel sees when it recognizes a sentence as “deducible” from a given set of premises is a proof in some proof system. But what proof system does the archangel use? Does it matter?

The answer is that it does matter – and matters very much. The proof system must be a *correct* system in the following sense: If there is a proof of Q from P_1, \dots, P_n , then P_1, \dots, P_n must logically imply, or entail, Q . That is, the premises of a proof in a correct system must guarantee the truth of the conclusion proved. We are assuming that the language in which the issues of emergence are considered is provided with a semantics and that semantic notions like validity and implication are available for sentences of the language. What matters from a metaphysical point of view is semantics, more specifically logical implication, not syntax – proofs are relevant only if they are proofs in a correct system, and that is so because that guarantees these proofs will have the right semantic property, that of preserving truth from premises to conclusions. Does it matter which correct proof system is used by the archangel? Some proof systems are more intuitive and perspicacious than others; the usual deduction systems we find in logic textbooks are formulated, we may assume, with an eye toward simplicity and perspicuity – that is, for the typical student. But that is epistemology; the archangel is hardly a typical student, and any system is as perspicuous to it as any other. Remember: the archangel was expressly called into service in order to cancel out epistemological considerations and get unalloyed metaphysics reinstated, untainted by cognitive limitations and relativity. There also is a formal reason why for first-order logic (with identity), arguably the core of what we call “logic”, the choice of a system does not matter – because this is a “complete” system, that is, there are formalizations of this logic that are complete. A proof system is complete in the technical sense just in case there is a proof of a sentence from a set of premises *if and only if* the premises imply, or entail, that sentence. In consequence, it doesn’t matter which one of these complete deductive systems we use to define deducibility; for they are all equivalent in that something is provable in one system from a given set of premises if and only if it is so provable in each of the rest. The only thing that matters is the fact that there being a proof guarantees implication, or logical entailment. And if

there is no proof in the system, then there is no logical implication; that is, nondeducibility amounts to the absence of implication in languages whose logic is amenable to complete formalization.

So the relativity of proofs to proof systems cancels itself out as cause for concern for us. Does this mean that we are finally home free with Broad's metaphysical conception of emergence as supervenience plus nondeducibility?

Unfortunately, the answer is no, at least not yet; in truth, we are now in far worse trouble than before. For look where we are: we have just seen that in speaking of deducibility and nondeducibility, what matters turns out to be logical implication, or entailment, or the absence of thereof. Deducibility in a complete system of proof, or deducibility as the mathematical archangel sees it, is nothing but logical entailment; similarly, nondeducibility amounts to logical nonentailment. With this in mind, look again at Broad's attempt to combine supervenience with nondeducibility to obtain a metaphysical relation of emergence. An emergent property supervenes on basal-level conditions (including basal-level laws); this means that certain basal conditions, namely those that constitute a supervenience base for it, necessitate, or entail, the emergent property. We have just seen that the net effect of nondeducibility comes to the absence of logical necessitation. Consequently, metaphysical emergence, as conceived by Broad and others, comes to this: Emergent properties are entailed but not logically entailed by their basal conditions. Or, to put it another way, emergent properties supervene, but do not logically supervene, on their basal conditions.

On the face of it, this seems like a coherent conception, because we think that logical supervenience is not the only kind of supervenience. There is another recognized variety that is weaker than logical supervenience, what is called "nomological", or "natural", supervenience: the bases of supervenient properties do not logically entail them but when combined with prevailing laws of nature, the entailment holds. Thus, the necessitation is nomological, not logical or metaphysical. As nomological supervenience appears to be the only alternative to logical supervenience, the Broad-style conception of emergence has now taken the following form: Emergent properties supervene nomologically, but not logically, on their basal conditions. And this may seem to fit in well with certain current forms of dualism, in particular, Chalmers' "naturalistic dualism", which consists in the claim that consciousness is naturally (that is, nomologi-

cally), but not logically, supervenient on physical phenomena.²⁸ On the present construal, therefore, the classical British emergentism of Broad and others might seem to be an almost exact anticipation of Chalmers' naturalistic dualism.

But not so fast! We must at this point look a bit deeper into the nature of nomological supervenience that may be involved in the Broad-style metaphysical emergence. When we refer to the "basal conditions" of an emergent property, as we have noted more than once, these conditions are taken to include not only particular facts – events, states, and processes – at the base level but also *laws operative at that level*. So as regards the emergence of mental properties on biological/physical properties, the appropriate basal conditions include all biological/physical laws – all laws applicable to biological, neurological, and physicochemical systems and phenomena. Recall the iconic emergentist question: Knowing all about the biological, neural, and physicochemical facts about the brain, can we predict, or know, what conscious experience, if any, will be present in that brain? Obviously, "knowing all about what goes on at the neural level" should be taken to include knowing the laws that hold at that level.²⁹

This means that when we say that an emergent property is not logically implied, or necessitated, by its basal conditions, the latter include not only particular basal facts but also all basal laws, laws holding at the base level. So in a sense the logical supervenience we are talking about is a form of nomological supervenience, where the laws involved are base-level laws (we may assume that these laws include all laws holding at still lower levels). It is crucially important to keep in mind that *the laws assumed to be included in the basal conditions are not all the laws of nature; they are only laws that hold at the basal level*.³⁰ Thus, when we say emergent properties are not logically but only nomologically necessitated by the basal conditions, the *additional* laws needed to yield the necessitation are not

²⁸ *The Conscious Mind* (New York and Oxford: Oxford University Press, 1996). Chalmers says: "In general, B-properties supervene naturally on A-properties if any two *naturally possible* situations with the same A-properties have the same B-properties", where a situation is naturally possible if "it could actually occur without violating any natural laws", p.36. On naturalistic dualism, see pp.168ff.

²⁹ Chalmers' notion of logical supervenience on physical facts is similar; the physical facts are stipulated to include all physical laws. See *The Conscious Mind*, p.33.

³⁰ I believe it should be assumed – and this is what I assume – that laws operative at a given level include all laws that hold at the lower levels as well. Thus, for example, laws at the biological level include all physical and chemical laws.

base-level laws (we have them already) but laws that connect the emergents with specific basal conditions, namely those that Broad calls “trans-ordinal laws”, laws connecting events and states at different levels, or “orders”. These are also sometimes called “laws of emergence” or “supervenience laws”,³¹ and take the following form:

When conditions *C* at the base level hold for system *S* at *t*, *S* instantiates, at *t*, emergent property *E*.

These laws are the auxiliary premises we need to deduce statements about facts involving emergent properties from statements about their basal conditions. Trans-ordinal laws, like the bridge laws in Nagelian theory reduction,³² are not exclusively about the base level; they concern the relationships between the base-level phenomena and the phenomena at a higher level. In the case of the mind-body relation, these are psychophysical laws telling us under what neural/biological conditions, a given type of conscious experience occurs. On emergentism,³³ every conscious state will be connected by such an inter-level law with its underlying neural basal conditions; and there can be multiple emergence bases for a given type of conscious state. As these laws involve higher-level phenomena, they cannot be part of the basal conditions from which the deducibility of an emergent is considered.

These considerations bring to light a deep difficulty – in fact, what may well be an incoherence – in Broad’s characterization of emergence as supervenience *plus* nondeducibility. For this combination threatens to turn into an outright contradiction: supervenience says that basal conditions entail, or imply, the emergent phenomenon; however, nondeducibility, which, as we saw, comes to nonentailment, or nonimplication, says that the basal conditions do not entail or imply the emergent phenomenon. Or equivalently, the emergent phenomenon both supervenes and does not supervene on its basal conditions. We tried to defuse this potentially disastrous situation by construing the combination as nomological supervenience *plus* the denial of logical supervenience; that is, emergent properties

³¹ Chalmers’ term in *The Conscious Mind*, p.127.

³² See Ernest Nagel, *The Structure of Science* (New York: Harcourt, Brace, and World, 1961), chapter 11.

³³ Obviously, an assumption of this kind is shared by many physicalist theories of the mind.

are nomologically, but not logically, supervenient on their basal conditions. We can now see that this rescue strategy fails. The reason is that nomological supervenience as it is usually understood, for example, Chalmers's "natural" supervenience, fails to capture Broad's concept of supervenience, or determination. Look at an earlier quote from Broad again:

"No doubt the properties of silver-chloride are completely determined by those of silver and of chlorine; in the sense that whenever you have a whole composed of these two elements in certain proportions and relations you have something with the characteristic properties of silver-chloride ... But connecting the properties of silver-chloride with those of silver and chlorine and with the structure of the compound is, so far as we know, an unique and ultimate law."³⁴

What Broad is claiming is that when we have fixed the base-level conditions and laws, that determines what phenomena will, or will not, emerge, even though the latter are not deducible from the former. So an emergent property supervene on its basal conditions plus basal laws but not deducible from them. It is crucial here that *the supervenience base and deduction base are held identical*. When we speak about nomological supervenience, we usually have in mind *all laws* prevailing in this world, which will include Broad's trans-ordinal laws, or Chalmers's "supervenient laws", or Nagel's "bridge laws" – that is, laws connecting phenomena at different levels. Broad will cheerfully admit that if trans-ordinal laws are admitted as part of the deduction base (that is, among the premises), emergent properties or phenomena are easily deduced from the basal conditions. His point is that when we are limited to base-level laws and conditions, no deduction is possible, since there is none. And yet he maintains that the emergents supervene on, or are determined by, the base-level phenomena and laws. And this supervenience is logical or metaphysical supervenience. There is a nomological aspect to that because the supervenience base includes basal laws. But once these laws are considered as part of the base, the supervenience relation becomes the logical/metaphysical variety.

This apparently puts Broad and his like-minded colleagues in an untenable position. He has been reduced to saying that emergent properties logically supervene on basal conditions and laws and yet they are not deducible from them. To repeat, nondeducibility, on Broad's idealization, turns into the absence of logical entailment. So Broad's conception of emergence turns into an apparent incoherence: logically supervenient on

³⁴ *The Mind and Its Place in Nature*, pp.64–65.

basal facts but not logically entailed by them – that is, to be brute, both logically supervenient and not logically supervenient on basal facts! I think the source of the problem is rather obvious: to avoid an epistemic interpretation of “deducible”, Broad tries to idealize it in terms of the mathematical archangel, as Laplace tried with his logically all-powerful “demon” (who, like the mathematical archangel, is also microphysically omniscient), but idealized deducibility turns into entailment, contradicting the first component of his metaphysical emergence, that is, supervenience. Is there a way out for Broad from this apparently incoherent situation?

V

One point to consider is the following. We have been assuming that the language in which emergence issues are considered is a first-order language with a complete proof system. Without completeness, we cannot equate nondeducibility with nonentailment, or the absence of implication. So what about languages for which there are no complete proof systems? Elementary number theory, or arithmetic, is famously incomplete; that is, there is no consistent formalization of number theory in which all and only number-theoretic truths are provable. Surely, it might be said, number theory must be taken to be part of the deductive system when we consider the deducibility of emergent phenomena.

However, it doesn't seem to me that the emergentist can exploit the incompleteness of arithmetic to dispute our claim that nondeducibility boils down to nonentailment. The reason is that we can happily let the emergentist add *all mathematical truths* (including of course arithmetic truths) to her deductive system and ask her whether this would help her deduce the emergents from their basal conditions. I believe the emergentist would have to say no. After all, the mathematical archangel is mathematically omniscient and it has at its disposal all mathematical truths. Its deductive system has to be complete. If a truth is not deducible from another in such a system, it must be because the former is not implied by the latter.

There is another concern we should address briefly. One might point out that “*x* is a male” is not deducible in first-order logic from “*x* is a bachelor” and yet obviously “*x* is a bachelor” logically implies “*x* is a male”. The reply is that definitions are free in deductions and don't count as additional premises; one is entitled to use them at any point in a proof.

So the deduction goes through with the definition “ x is a bachelor iff x is an unmarried adult male”.

But this may not put away the concern entirely. For one might continue: Even so, “ x is water” is not deducible from “ x is H_2O ” since there is no conceptual definition linking “water” and “ H_2O ”. And yet the proposition that Lake Michigan is filled with H_2O entails that it is filled with water. Therefore, we cannot say that nondeducibility, even with the proviso concerning definitions, amounts to nonentailment. This point touches on various controversial issues currently debated in the philosophy of mind, philosophy of language, and metaphysics – issues arising from (supposed) cases of a posteriori necessities and entailments. I will simply state here my own response without detailed explanation or justification (this response derives from – and, I believe, is consistent with – the views of philosophers like David Chalmers and Frank Jackson³⁵). True, there are no direct conceptual links between “water” and “ H_2O ”, but this doesn’t mean appropriate deductive links cannot be forged between statements about water and statements about H_2O . What do we mean by “water”? We reply: it means something like “the local watery stuff”, where “watery” is short for a conjunction of predicates designating the observable properties by which we ordinarily recognize and identify water – properties like transparency, the power to quench thirst, the characteristic viscosity (the way it “flows”), its freezing and boiling temperatures, its power to dissolve sugar and salt but not butter, and so on. We take this to be a conceptual fact grounded in meanings. Now consider the following array of statements:

- (i) Lake Michigan is filled with H_2O
- (ii) The local watery stuff = H_2O
- (iii) Water = the local watery stuff
- (iv) Therefore, Lake Michigan is filled with water

We assume that (ii) is deducible from physical facts – physical laws as well as particular facts. Physical theory, we may assume, can show that the local stuff that is transparent, can dissolve sugar, and flows in a certain way is made up of H_2O molecules; and it can explain why quantities of

³⁵ See, e.g., Chalmers, *The Conscious Mind*; Chalmers and Frank Jackson, “Conceptual Analysis and Reductive Explanation”, *Philosophical Review* 110 (2001): pp.315–360. For views critical of the present approach, see Ned Block and Robert Stalnaker, “Conceptual Analysis, Dualism, and the Explanatory Gap”, *Philosophical Review*, 108 (1999): pp.1–46.

H₂O molecules behave in the way water behaves. Step (iii) is a meaning-based definition; so it comes free. And (iv) can be deduced from the preceding steps in any first-order logic. Note that the use of physical laws is appropriate since, as we have repeatedly emphasized, the base of deduction for the emergence debate is taken to include base-level laws as well as particular conditions. In any case, other cases of a posteriori entailments hopefully can be dealt with in a similar fashion.

There are more issues and problems to be discussed, but that would have to wait for another occasion. However, our provisional conclusion stands: We are still in need of an intelligible and coherent metaphysical characterization of emergence, one that does not involve ineliminable references to our, or anyone's, cognitive situations or powers. So there is work to do – not only for the emergentists bent on defending emergentism as a metaphysical thesis about the world but also for those reductionist physicalists who think that although metaphysical emergentism is false and refutable, it is an intelligible thesis.