Consequences of conditional excluded middle*

Jeremy Goodman

February 25, 2015

Abstract

Conditional excluded middle (CEM) is the following principe of counterfactual logic: either, if it were the case that φ, it would be the case that ψ, or, if it were the case that φ, it would be the case that not-ψ. I will first show that CEM entails the identity of indiscernibles, the falsity of physicalism, and the failure of the modal to supervene on the categorical and of the vague to supervene on the precise. I will then argue that we should accept these startling conclusions, since CEM is valid.

1 The Argument

Consider the following version of the principle of the identity of indiscernibles:

> PII
> Necessarily, for all x, y: If, for all qualitative properties F, Fx iff Fy, then $x = y$.[1]

*Thanks to Andrew Bacon, Dave Chalmers, Cian Dorr, Hartry Field, Kit Fine, Peter Fritz, John Hawthorne, Jim Pryor, Bob Stalnaker, Tim Williamson, and Juhani Yli-Vakkuri for comments on previous versions of this paper, and to audiences at Oxford and NYU.

[1]Roughly, a property F is qualitative just in case for no relation R and individual x is it the case that to have F is to bear R to x. Qualitative properties, in this sense, can be highly extrinsic and gerrymandered. The restriction to qualitative properties is needed to avoid trivializing the PII, since it is trivial that x is identical to any y with which it shares the property *being identical to x*.

Max Black (1952) is generally thought to have refuted PII. His argument was simple. It is possible that the universe contain nothing more than two duplicate iron spheres orbiting each other in empty space. Were this to happen, distinct spheres would have exactly the same qualitative properties. PII is therefore false.

Those who have resisted Black's argument have resisted its first premise: that it is possible that the universe contain nothing more than two duplicate iron spheres orbiting each other in empty space. The general strategy has been to argue that, although Black was imagining a genuine possibility, he may have been mistaken in taking it to be one in which the universe contains two duplicate iron spheres. Perhaps he was instead imagining a single multiply located sphere, or a single sphere in a non-Euclidean spacetime.[2] I hope these hypotheses strike you as bizarre. It is not at all plausible that we are making any such mistake when we find Black's argument compelling. I will take for granted that the universe could have contained nothing more than two duplicate iron spheres orbiting each other in empty space.

Nevertheless, Black's argument is unsound. Had the universe contained nothing more than two duplicate iron spheres orbiting each other in empty space, these spheres would not have had exactly the same qualitative properties. For only one of them would have had the following qualitative property: being a duplicate of an iron sphere that it is orbiting such that, had one of these two spheres been heavier than the other, it would have been the heavier one. It is hard to express this condition without scope ambiguity in English, so let us move to a formal language. Let Dxy abbreviate "x and y are duplicate iron spheres orbiting each other in otherwise empty space", Hxy abbreviate "x is heavier than y", $x \not\approx y$ abbreviate 'either Hxy or Hyx', and $\Box\!\!\rightarrow$ abbreviate the counterfactual conditional. Were the universe to contain nothing more than two duplicate iron spheres orbiting each other in otherwise empty space, exactly one of the spheres would have the property

$$\Pi x =_{df} \exists y(Dxy \wedge (x \not\approx y \Box\!\!\rightarrow Hxy))$$

Π is qualitative, since it is defined in terms of qualitative relations without making reference to any particular individuals. The argument that exactly one of the two spheres would have Π appeals to an instance of conditional excluded middle, the following controversial schema of counterfactual logic:

[2] See, respectively, Hawthorne (1995) and Hacking (1975).

CEM
$(\varphi \mathrel{\Box\!\!\rightarrow} \psi) \vee (\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$

We count the result of prefixing an instance of CEM by any number of universal quantifiers and necessity operators to also be an instance of CEM. The instance of CEM needed for the present argument is the following:

(1) $\Box\forall x\forall y : (x \not\approx y \mathrel{\Box\!\!\rightarrow} Hxy) \vee (x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg Hxy)$

The argument also requires the following two uncontroversial premises:

(2) $\Box\forall x\forall y : Dxy \rightarrow \neg((x \not\approx y \mathrel{\Box\!\!\rightarrow} Hxy) \wedge (x \not\approx y \mathrel{\Box\!\!\rightarrow} Hyx))$

(3) $\Box\forall x\forall y : Dxy \rightarrow \neg((x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg Hxy) \wedge (x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg Hyx))$

(2) says that, necessarily, for any two duplicate iron spheres x and y orbiting each other, it is not both the case that [had x and y differed in mass, x would have been heavier than y] and [had x and y differed in mass, y would have been heavier than x]. (3) says that, necessarily, for any two duplicate iron spheres x and y orbiting each other, it is not both the case that [had x and y differed in mass, x would not have been heavier than y] and [had x and y differed in mass, y would not have been heavier than x]. These claims should be obvious.

Finally, the argument appeals to the following consequence of Black's description of the scenario:

(4) $\Box\forall x\forall y : Dxy \rightarrow \forall x'\forall y'(Dx'y' \leftrightarrow ((x' = x \wedge y' = y) \vee (x' = y \wedge y' = x)))$

(4) says that, necessarily, if there are two duplicate iron spheres orbiting each other in otherwise empty space, then there is only one such pair.

(1)-(4) entail that, necessarily, if there are two duplicate spheres such that the universe consists of nothing more than them orbiting each other in empty space, then exactly one of the spheres has Π:

(5) $\Box\forall x\forall y : Dxy \rightarrow (\Pi x \leftrightarrow \neg\Pi y)$

Since it is not contingent that Π is a qualitative property, (5) entails the advertised conclusion: Had the universe contained nothing more than two duplicate iron spheres orbiting each other in empty space, these spheres would not have shared all qualitative properties.

The argument generalizes. Suppose the universe had contained nothing more than three duplicate iron spheres symmetrically orbiting their center of mass in empty space. Only one of them would have the qualitative property of being one of three duplicate iron spheres symmetrically orbiting their center of mass in empty space such that, had one of the three spheres been heavier than the others, it would have been the heavier one. Call this property Π_1. Of the two remaining spheres, only one would have the qualitative property of being one of three duplicate iron spheres symmetrically orbiting their center of mass in empty space such that, had one of the three spheres not having Π_1 been heavier than the others, it would have been the one. Call this property Π_2. Π_1 and Π_2 distinguish all three spheres. In this way counterfactuals give us a recipe for qualitatively distinguishing any finite number of duplicate iron spheres.

The argument generalizes still further. It turns out that CEM entails the *necessity of discernibility*: necessarily, if two things are qualitatively indiscernible, then there is no qualitative property that even possibly divides them. (Since the argument is somewhat more involved, I have relegated it to an appendix.) I have no proof that distinct objects cannot be necessarily qualitatively indiscernible, and so no proof of PII. But it is hard to imagine what such pairs of objects would be like, and such extreme necessary connections between distinct existents is hard to fathom.[3] We therefore have strong

[3]Suppose we follow Hawthorne (2006, 241) and accept the following principle of plenitude regarding material objects: "for every [function from worlds to filled regions of spacetime] there is an object whose modal pattern of spatiotemporal occupation is correctly described by that [function]." One might think that pairs of 'world-bound' objects – objects that occupy non-empty regions of spacetime in only one world – can be necessary qualitatively indiscernible. The thought is that, were there a pair of world-bound objects respectively coincident with two duplicate iron spheres, then, unlike the spheres, these objects could not fail to be duplicates. But while it is true that we cannot simply replace "iron sphere" with "world-bound object coincident with an iron sphere" in (1)-(4) to establish such objects' qualitative discerniblility, such objects would nevertheless be qualitatively indiscernible, since only one will have the qualitative property of coinciding with something that has Π.

The two square roots of -1 are perhaps more promising candidates for a pair of necessarily qualitatively indiscernible objects. But even here there are some reasons for doubt. For one thing, being signified by a symbol with the qualitative orthographic features of "i" might be thought to be a qualitative property that distinguishes them. Moreover, if we accept a set-theoretic foundation for mathematics, then all mathematical objects can be qualitatively distinguished in terms of $\in$. At any rate, the example does not suggest any way for pairs of *concrete* objects to be necessarily qualitatively indiscernible. See Shapiro reason to accept PII.

CEM has other striking consequences. Consider the following statement of physicalism:

> PHYSICALISM
> For all qualitative properties F, it is nomologically necessary that, for all concrete individuals x: If Fx, then, for some physical property G: Gx and it is nomologically necessary that, for all concrete y: If Gy, then Fy.

Physical properties should be understood broadly so as to be closed under logical operations. A proposition is nomologically necessary just in case it is true in all worlds 'like ours' (to borrow Lewis' (1994) phrase), where a world is like ours just in case it shares our laws and instantiates no fundamental properties alien to our world. So understood, PHYSICALISM enjoys wide support. It is compatible, for example, with the metaphysical (though not nomological) possibility of Chalmers' (1996) notorious 'zombie worlds'.

But wait. It seems to be nomologically possible that the universe contain nothing more than two duplicate iron spheres orbiting each other in empty space.[4] Such spheres would share all physical properties. But they would be divided by the qualitative property Π. PHYSICALISM is therefore false.

2 Objections

In this section I consider ways in which the above arguments might be resisted without rejecting CEM. In the next section, I will consider whether the above arguments give us reason to reject CEM.

I have already considered and rejected denying the possibility of a universe containing nothing more than two duplicate iron spheres orbiting each other in empty space. One might deny (2) and (3) by claiming that Black's spheres would have their matter essentially and so neither could have been heavier than the other (which would make the counterfactuals in (2) and (3) vacuously true). But since any qualitative difference will do for the purpose of the above argument – e.g., difference in shape rather than difference in mass – the objection is insufficiently general to avoid the argument's conclusion.

(2012) for discussion.

[4]Less controversially: It is nomologically possible that the physical world be complex and symmetric enough to run the above style of argument.

More radically, one might reject to (4) by claiming that, in response to familiar puzzles about soritical modal variation in the matter out of which material objects are composed, we should think that each of Black's spheres coincides with a great many duplicate spheres each differing slightly in which portions of their matter they have essentially.[5] Setting aside the implausibility of such views, this objection is also insufficiently general. Any two coincident spheres with different modal profiles are *possibly* qualitatively discernible, since they are obviously discernible in worlds where only one of them exists. They are hence *necessarily* qualitatively indiscernible, by the argument of the appendix. So, for any two coincident spheres x and y, there is a qualitative property Q_x^y that x has and y lacks. Let Q_x be the conjunction of all such properties for all spheres y coincident with x. Q_x is qualitative, since it is the conjunction of qualitative properties. And it distinguishes x from every sphere with which it coincides. So we can simply modify the original argument by replacing being an iron sphere with being a Q_x iron sphere. By construction, there is at most one pair of Q_x spheres, securing (4).[6]

One might instead try to downplay the interest of PII by claiming that we should understand the identity of indiscernibles in a different way that *is* refuted by Black's argument, or that is at least not supported by my argument using CEM.[7] Two alternative principles suggest themselves:

> PII$_\Delta$
> Necessarily, for all x, y: If, for all qualitative properties F, *determinately* Fx iff *determinately* Fy, then $x = y$.
>
> PII$_\Box$
> Necessarily, for all x, there is a qualitative property F such that, necessarily, for all y: Fy iff $x = y$.

I will consider each in turn.

[5] Leslie (2011) defends such a view; see also Williamson (1986) and Hawthorne and McGonigal (2008).

[6] Another response to the objection is that, unlike artifacts and organisms, portions of matter have all of their matter essentially, so the line of reasoning supporting a plenitude of coincident but modally discernible artifacts and organisms, even if accepted, gives us no reason to believe in distinct coincident hunks of iron. See Goodman (unpublished a).

[7] For example, Leibniz's version of the identity of indiscernibles *is* arguably refuted by the possibility of Black's spheres, since he seemed to think that distinct objects cannot share all *intrinsic* properties.

PII_Δ says that, necessarily, no two objects have exactly the same qualitative properties *determinately*. It presupposes a notion of determinacy metaphysically robust enough to allow us to sensibly quantify into the operator's scope. For those who recognize such a notion, it is natural to think that, had the universe contained nothing more than two duplicate iron spheres orbiting each other in empty space, it would have been indeterminate which sphere had Π.[8] Given the further assumption that the spheres would have exactly the same qualitative properties determinately, Black's thought experiment refutes PII_Δ.

$\mathrm{PII}_\Box$ says that, necessarily, every object has a qualitative *haecceity* – a qualitative property necessarily equivalent to being identical to that object. This principle is strictly stronger than PII. PII says that, necessarily, for every object, there is a qualitative property had by it and by nothing else. $\mathrm{PII}_\Box$ says that, necessarily, for every object, there is a qualitative property *necessarily* had by it (if it exists) and by nothing else. (For this reason, $\mathrm{PII}_\Box$, unlike PII, entails *anti-haecceitism*: the view that all facts supervene on the qualitative facts.) For all I have said, Black's two spheres would be a counterexample to $\mathrm{PII}_\Box$. At any rate, there is no straightforward argument from CEM to $\mathrm{PII}_\Box$.

PII_Δ and $\mathrm{PII}_\Box$ are both interesting metaphysical theses, and it is a welcome consequence of the argument from CEM to PII that it forces us to isolate principles that we might have otherwise confated. But merely isolating PII from PII_Δ and $\mathrm{PII}_\Box$ should bring little comfort to those who were initially inclined to resist the former. After all, it is extremely pre-theoretically plausible that there could be a complex yet qualitatively symmetric universe. If my argument from CEM is sound, this intuitive view is simply mistaken.

Similar points apply to the argument against PHYSICALISM. One could try to understand physicalism in some weaker way, such as one of the following:

> PHYSICALISM$_\Delta$
> For all qualitative properties F, it is nomologically necessary that, for all concrete individuals x: If determinately Fx, then, for some physical property G: Gx and it is nomologically necessary that, for all concrete y: If Gy, then Fy.
>
> PHYSICALISM$^-$
> For all qualitative propositions p, it is nomologically necessary that: If p,

[8]See Stalnaker (1981).

then, for some proposition q concerning only how things are physically, it is nomologically necessary that: If q is true, p is true.[9]

PHYSICALISM$_\Delta$ stands to PHYSICALISM as PII$_\Delta$ stands to PII. PHYSICALISM$^-$ in effect claims that the qualitative *weakly globally supervenes* on the physical, whereas PHYSICALISM claims that the qualitative *strongly individually supervenes* on the physical.[10] While my argument against CEM brings into focus the difference between these three versions of physicalism, it seems unlikely to weaken the pull of PHYSICALISM for those tempted by the thought that, in a world like ours, there are no qualitative goings on over and above the physical goings on.[11] That two duplicate iron spheres orbiting each other in otherwise empty space could be divided by a qualitative property is simply inconsistent with that world view. Perhaps it would be indeterminate *which* sphere had Π, but it would still be determinate *that* Π divides them. The fact that the world's global qualitative profile may still be necessitated by its global physical profile will be cold comfort. For example, consider the albeit bizarre view that there is a nomologically possible physically symmetric universe in which it is definitely the case that only one of two physically identical collections of particles composes a chair, although it is indeterminate which one. I take it such a view is inconsistent with physicalism about chairs.[12]

Another way to resist my argument against PHYSICALISM would be to claim that Π is a physical property. I don't want to enter into the debate about which properties count as physical (although it is worth noting that it is implausible that Π can be expressed by any open sentence of a language

[9]A propositions p is qualitative just in case for no property F and individual x is it the case that for p to be true is for x to have F.

[10]See McLaughlin and Bennett (2005) for a survey of different kinds of supervenience.

[11]Compare Field (1992).

[12]Some self-proclaimed 'physicalists' reject PHYSICALISM on the grounds that coincident physical objects, such as a statue and the lump of clay composing it, differ qualitatively without differing physically. They will presumably accept only the weaker principle:

> PHYSICALISM*
> For all qualitative properties F, it is nomologically necessary that, for all concrete individuals x: If something coincident with x has F, then, for some physical property G: Gx and it is nomologically necessary that, for all y: If Gy, then something coincident with y has F.

Note that Π is also a counterexample to PHYSICALISM*. See Goodman (unpublished b) for discussion.

containing only logical vocabulary and the sorts of vocabulary one find in physicists' formal statements of their theories). The interest of my argument is independent of such terminological decisions. Its interest lies in the fact that, if sound, it refutes a family of widely held supervenience theses, many of which happen to be advanced under the banner of "physicalism". I have in mind the sort of supervenience theses captured by slogans like "the modal supervenes on the categorical" and "there is no primitive modality."[13] Indeed, many who reject physicalism (about consciousness, for example) nevertheless accept these supervenience theses. If there is 'primitive modality,' its interest is not hostage to whether we dub it "physical". I should add that, insofar as we can make sense of quantification into the scope of a "definitely" operator, the present argument also arguably refutes the popular view that the vague supervenes on the precise, for the reasons discussed in the previous paragraph.

One might resist both of my arguments by denying that Π is qualitative. Since I have expressed Π using only qualitative predicates and without making reference to any particular individuals, this line of resistance seems committed to there being something non-qualitative about the counterfactual conditional itself. There are a number of things to note about this proposal. First, to reiterate: for philosophers concerned to reject 'primitive modality,' it is bad enough that CEM threatens the supervenience of Π on 'non-modal' properties, irrespective of whether Π is qualitative. An analogous point applies to those concerned with the supervenience of the vague on the precise. Second, although the distinction between qualitative and non-qualitative ideology is at times murky, we are pretheoretically inclined to judge Π to be qualitative. To simply deny this in light of my argument from CEM would be *ad hoc*; it offers no helpful picture of *how* counterfactual ideology might fail to be qualitative. (What individuals is it about, and how so?) Finally,

[13]See Stalnaker (1984) and Williamson (2013, 8.2) for skepticism about the categorical/modal distinction. On the basis of this skepticism, Stalnaker arrives at a conclusion in a similar spirit to my own:

> [T]he introduction of counterfactuals allows for finer discriminations between possible worlds than could be made without counterfactuals. The selection functions relative to which counterfactuals are interpreted do not simply select on the basis of facts and criteria of similarity that are intelligible independently of counterfactuals. Rather, the claim is, the fact of selection gives rise to new ways of cutting up the space of possibilities, and so to a richer conception of the way the world is. (Stalnaker, 1984, 169)

we are strongly inclined to judge that necessity and causation are qualitative notions, not to mention a host of central notions characterized in terms of them. The intimate connections between necessity, causation and counterfactuals give us reason to think that they agree as regards qualitativeness.[14] If all of these notions turned out to be non-qualitative, this would make PII and PHYSICALISM hollow doctrines with few domains of non-trivial application. The failure of Π to supervene on the physical, if true, is a surprising and interesting discovery however it is glossed. And, if CEM is valid, it is a discovery we must accept.[15]

3 Whither CEM?

I now turn to a defense of CEM. Most of the following arguments are not original to me, so I will be brief.

First, CEM simply strikes us as valid. Even David Lewis, its principal detractor, admits as much.[16] CEM explains why it is so difficult to hear a difference between negated counterfactuals and (non-vacuous) opposite counterfactuals: for example, between "It is false that I would have been happy had it rained" and "Had it rained, I would not have been happy." It is also needed to explain the validity of many inferences involving quantified counterfactuals, such as inferences of the form: No F would G were it to H; therefore, every F would fail to G were it to H.[17]

[14] Another argument: practical rationality is a qualitative notion, and being practically rational is a matter of coordinating one's intentions with one's beliefs about various counterfactuals (about what would happen if one were to do various things), and hence is a qualitative notion only if the counterfactual conditional is.

[15] In Goodman (unpublished b) I argue that puzzles about the possible material origins of artifacts generate counterexamples to PHYSICALISM superficially similar to the the present one. Since the argument of that paper does not appeal to CEM, it provides abductive support for the present conclusion.

[16] See Lewis (1973, 80).

[17] Consider the following example, which Williams (2010) adapts from Higginbotham (1986): "No student would have passed if they had goofed off; therefore, every student would have failed to pass if they had goofed off". While the behavior of quantified conditionals is admittedly complicated, the emerging consensus seems to be that CEM is at least part of its explanation. See Klinedinst (2011), Kratzer (forthcoming), and references therein; for a dissenting voice, see Leslie (2008). von Fintel and Iatridou (2002) are often cited as the source of this view, but this attribution is complicated by the role played by presupposition in their proposal; see Klinedinst (2011) for discussion.

Second, the chances of counterfactuals often pattern with the conditional chances, at some contextually salient time, of their consequents conditional on their antecedents.[18] For example, the chance that this coin would land heads if it were flipped is approximately .5. If genuine, these data immediately entail (by the probability calculus) that the corresponding instances of CEM have chance 1, which ought to be sufficient grounds for accepting them. Since counterfactual chancy processes are often thought to be the greatest obstacle to CEM, its validity concerning such cases strongly confirms the general principle. Moreover, even if CEM is guaranteed to hold only in cases where the corresponding conditional chances are well-defined, that would be enough for my argument, since we can modify Black's case by stipulating that both spheres contain a radioactive particle that never decays, but each is such that, conditional on exactly one of them decaying, the chance of it decaying is .5 – simply substitute 'x but not y undergoes radioactive decay' for 'x is heavier than y' in the argument.[19]

Third, our everyday practical deliberations seem to reveal a tacit commitment to CEM. In deciding whether to perform an action, I ask myself "What would happen if I were to do it? Would it be a good thing that I did it, or would it be a bad thing that I did it?" This intuitive thought finds formal expression in the original versions of causal decision theory.[20] A commitment to CEM also manifests itself in conversation. If someone asked me how this coin would have landed had it been flipped (perhaps because they don't understand that coin flipping is a chancy process), the most natural reply would be "I don't know". Without CEM it is hard to make sense of such professed ignorance.

The main arguments in the literature against CEM are weak. I will mention three of them. The first appeals to David Lewis' 'closest-worlds' semantics for counterfactuals together with his similarity-based theory of closeness between worlds.[21] In order for CEM to be valid on this semantics, it must

[18]See Skyrms (1981 [1980]), Williams (2012), and especially Moss (2013).

[19]Those who accept this restricted version of CEM can agree with Quine (1950, 15) that 'If Bizet and and Verdi had been compatriots, they would have been Italian, or, if Bizet and and Verdi had been compatriots, they would not have been Italian' is a counterexample to unrestricted CEM, on the grounds that the corresponding conditional chances are not well-defined.

[20]See Stalnaker (1981 [1972]) and Gibbard and Harper (1981 [1978]). Lewis (1981) develops a version of causal decision theory that does not assume CEM.

[21]Lewis (1973, 1979).

be case that, for any world w and possibly true proposition p, some world w' at which p is true is more similar to w in the relevant respects than any other world at which p is true. If similarity is understood in the way Lewis recommends, this condition will be false and so CEM will be invalid – indeed, its failure will be commonplace. But this is the least of this theory's revisionist consequences. Given plausible assumptions about statistical mechanical and quantum mechanical objective chances, the Lewisian semantics seems to entail that almost all ordinary counterfactuals are false![22] I take this consequence to be a reductio of the view; at any rate, it renders the view a dialectically weak way to challenge CEM.

A second argument against CEM appeals to Lewis' claimed duality of "would"-counterfactuals and "might"-counterfactuals, according to which "If it were the case that φ, it would be the case that ψ" is logically equivalent to "It is not the case that, if it were the case that φ, it might be the case that not-ψ." Since "If I were to flip the coin, it might land heads" and "If I were to flip the coin, it might not land heads" are both true, duality entails the falsity of CEM. But duality is not well motivated. The behavior of "might"-counterfactual is best explained compositionally in terms of the interaction of modals and conditionals.[23] Such explanations do not threaten CEM.

A third complaint against CEM is that it generates a kind of objectionable arbitrariness. Suppose I have a coin that I never flip. It seems arbitrary to say that, had I flipped it, it would have landed head, and also arbitrary to say that, had I flipped it, it would not have landed heads. But notice that this objection is structurally parallel to familiar arguments from vagueness against the law of excluded middle (LEM): in exactly the same way, it seems arbitrary to say of a borderline bald man either that he is bald or that he is not bald. So it is natural to diagnose the coin example and similar apparent failures of CEM simply as cases of vagueness. (I already hinted at such vagueness in discussion of PII_{Δ}.) Assuming we have learned to live with LEM in the face of vagueness, this diagnosis neutralizes the threat to CEM from arbitrariness.

(Suppose one agreed that CEM and LEM stand or fall together, but thought that we should reject both of them on account of vagueness. What

[22] See Bennett (2003) and Hawthorne (2005); Hajek (manuscript 2011) embraces this conclusion.

[23] For replies to Lewis (1973) along these lines, see Stalnaker (1981), Hawthorne (2005) and Williams (2010); the latter is an instructive rejoinder to a proposal by Bennett (2003). On the interaction of modals and conditionals more generally, see Kratzer (2012).

should such a person think about my argument for PII? The issue is subtle. Even those who reject LEM should acknowledge the existence of *penumbral connections*. For example, a borderline red/orange paint chip is red if and only if it isn't orange. Field (2008) reconciles such penumbral connections with the rejection of LEM by denying the equivalence of $\varphi \rightarrow \psi$ and $\neg\varphi \vee \psi$. Roughly, when a classical logician accepts a disjunction of vague but penumbrally connected disjuncts, Field will reject the disjunction but accept the corresponding conditional. So, if we model our rejection of CEM on Field's rejection of LEM (which I take to be the best developed view in the vicinity), we should still accept:[24]

$\text{CEM}^{\rightarrow}$
$\neg(\varphi \mathrel{\Box\!\!\rightarrow} \psi) \rightarrow (\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$

$\text{CEM}^{\rightarrow}$ is weaker than CEM in Field's logic, although they are obviously classically equivalent. Nevertheless, $\text{CEM}^{\rightarrow}$ is strong enough to run a version of my argument for (5) in Field's system.[25] We do not escape the conclusion that the spheres would be such that one had Π if and only if the other did not.

But there is a catch: In Field's logic, (5) is consistent with the claim that the spheres would be such that one had Π if and only if the other *did too*. Compare the case of the liar sentence. According to Field, the liar sentence is both true if and only if not true, and true if and only if true. Insofar as we are willing to think of such a situation as one in which the liar sentence is indiscernible from itself in truth value, perhaps we can likewise think of

[24]Field (2000, 6) seems to accept the parity of CEM and LEM, given that in a discussion of vagueness and LEM he writes that "there is no objective fact of the matter as to whether Bizet and Verdi would have been French rather than Italian had they been compatriots".

[25]By $\text{CEM}^{\rightarrow}$ we have:

$(1_a^{\rightarrow})$ $\Box\forall x\forall y : \neg(x \not\approx y \mathrel{\Box\!\!\rightarrow} Hxy) \rightarrow (x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg Hxy)$

$(1_b^{\rightarrow})$ $\Box\forall x\forall y : \neg(x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg Hxy) \rightarrow (x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg\neg Hxy)$

Since Field will accept the corresponding conditional whenever a classical logician accepts a negated conjunction with penumbrally connected conjuncts, we may assume:

$(2^{\rightarrow})$ $\Box\forall x\forall y : Dxy \rightarrow ((x \not\approx y \mathrel{\Box\!\!\rightarrow} Hxy) \rightarrow \neg(x \not\approx y \mathrel{\Box\!\!\rightarrow} Hyx))$

$(3^{\rightarrow})$ $\Box\forall x\forall y : Dxy \rightarrow ((x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg Hxy) \rightarrow \neg(x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg Hyx))$

Together with (4), the above principles entail (5) in Field's system.

the Field-inspired proposal as one according to which the two spheres would have been qualitatively indiscernible *despite* the truth of (5).)[26]

Of course, the committed opponent of PII or lover of PHYSICALISM may reject CEM on those grounds. This would be to set the considerations against PII and in favor of PHYSICALISM against the above considerations in favor of CEM. What are these considerations? The standard argument against PII is Black's, which is dialectically ineffective as objection to CEM; general arguments for PHYSICALISM (as opposed to arguments for physicalism about particular domains of facts) are hard to come by. I find as attractive as anyone the metaphysical vision of physical truth as the ground of all qualitative truth, but to reject CEM on such grounds smacks of wishful thinking. After all, (1) appears no less valid than any other instance of CEM. (Imagine being one of two duplicate people orbiting each other alone in the void considering the disjunction: Either, had one of us been heavier than the other, it would

[26] Here is one last argument against CEM. Suppose that, following Fine (1977) and Stalnaker (2012), we think that it is contingent what propositions there are, and hence contingent what distinctions among possibilities there are to be drawn. Fritz and Goodman (unpublished a) show that we can no longer accept (the full modal closure of) CEM. But they also show that we may still accept the weaker schema:

CEM^{-}
$\forall p((\varphi \mathrel{\Box\!\!\rightarrow} p) \vee (\varphi \mathrel{\Box\!\!\rightarrow} \neg p))$

As with CEM, we count the result of prefixing CEM^{-} with arbitrary strings of necessity operators and universal quantifiers to also be an instance of the schema. Now consider the instance:

(1^{-}) $\Box\forall x\forall y\forall p : (x \not\approx y \mathrel{\Box\!\!\rightarrow} p) \vee (x \not\approx y \mathrel{\Box\!\!\rightarrow} \neg p)$

(1) follows from (1^{-}) together with the assumption:

(6) $\Box\forall x\forall y\exists p\Box(p \leftrightarrow Hxy)$

(6) says that, necessarily, for any two things there is a proposition necessarily equivalent to the first thing being heavier than the second. This assumptions should be uncontroversial even by the lights of those who think it is contingent what propositions there are. So such 'propositional contingentism' does not block my argument for (5). More generally: assuming that, necessarily, there exist all distinctions among possibilities that can be drawn in terms of the qualitative properties of and relations between existing individuals, we will be able to use CEM^{-} in place of CEM in the argument that, necessarily, no two individuals are contingently qualitatively indiscernible. For more on the motivation for and technical implementation of this conception of propositional contingentism, see Fine (1977) and Fritz and Goodman (unpublished b).

have been me, or, had one of us been heavier than the other, it would have been my duplicate. I would find this proposition extremely hard to deny.)[27]

4 Conclusion

You might feel that there is something incongruous about the way I have drawn controversial metaphysical conclusions from a principle I have defended on linguistic, probabilistic, and psychological grounds (although also on grounds of apparent validity). Some philosophers will have an impulse to use high-powered metametaphysical resources to diagnose this perceived incongruence. They might claim that discourse involving counterfactuals fails to be 'factual', to concern how things are 'in reality', or to 'carve nature at its joints'.[28] Since I have said nothing at all about factuality, reality, or cosmic arthrology, such claim in no obvious way threaten the conclusions of this paper. And given the close connection between counterfactuals and a wide range of notions central to our most fruitful everyday and scientific theorizing, any aspersions we cast on counterfactuals are in danger of generalizing to much of what we hold dear.

Finally, let me return to the objection that there must be *something* wrong with the argument of this paper since its conclusion is simply beyond belief. While it is hard to argue philosophers out of their incredulity, one can at least try to diagnose its source. My suspicion is that, in many cases, such incredulity stems from a prior commitment to a metaphysical vision according to which all qualitative truth bottoms out in the pattern of instantiation of some small list of fundamental properties and relations. And this reductive vision is simply a prejudice – one we ought to reject. Counterfactuals impose their own structure on modal space, and that structure projects down onto the material world. For all I have said, this structure is highly constrained by

[27] Of course, when one runs out of objections, there is always the incredulous stare, such as the following from Black's original paper:

> You might just as well say "By 'a' I mean the sphere which would be the first to be marked by a red mark if anyone were to arrive and were to proceed to make a red mark!" You might just as well ask me to consider the first daisy in my lawn that would be picked by a child, if a child were to come along and do the picking. This doesn't now distinguish any daisy from the others. (Black, 1952, 157)

[28] See, e.g., Field (1994), Fine (2001) and Sider (2011).

the pattern of instantiation of certain fundamental properties and relations. Indeed, maybe all of the *determinate* structure of modal space is explicable in more fundamental terms. But counterfactuals nevertheless have a life of their own, and it is because they have such a life that reasoning about them is so central to our lives.

A The necessity of discernibility

Let $\hat{x}$ rigidly designate the set objects qualitatively indiscernible from x; think of it as a meta-linguistic abbreviation for the wide-scope definite description "$\{y : y$ is qualitatively indiscernible from $x\}$". (If we were worried that the things qualitatively indiscernible from x might not form a set, we could instead run the argument using plural quantification.) Let

$$\Lambda^+ x =_{df} \hat{x}/F \mathrel{\Box\!\!\rightarrow} Fx$$

$$\Lambda^- x =_{df} \hat{x}/F \mathrel{\Box\!\!\rightarrow} \neg Fx$$

where F is some qualitative property and

$$x/F =_{df} \neg\forall y(y \in x \rightarrow Fy) \wedge \neg\forall y(y \in x \rightarrow \neg Fy).$$

Intuitively, something has Λ^+ just in case, had the set of things qualitatively indiscernible from it been divided by F, it would have been F. Λ^+ is qualitative, since it is defined in terms of qualitative properties without reference to any particular individual. Likewise for Λ^-. Hence

(A) $\forall y\forall z : (y \in \hat{x} \wedge z \in \hat{x}) \rightarrow ((\Lambda^+ y \leftrightarrow \Lambda^+ z) \wedge (\Lambda^- y \leftrightarrow \Lambda^- z))$

By CEM we have:

(B) $\forall y((\hat{x}/F \mathrel{\Box\!\!\rightarrow} Fy) \vee (\hat{x}/F \mathrel{\Box\!\!\rightarrow} \neg Fy))$

Since qualitative indiscernibility is an equivalence relation, we have:

(C) $\forall y(y \in \hat{x} \rightarrow \hat{y} = \hat{x})$

By (A)-(C) and the definitions of Λ^+ and Λ^-, we have:

(D) $\forall y(y \in \hat{x} \rightarrow (\hat{x}/F \mathrel{\Box\!\!\rightarrow} Fy)) \vee \forall y(y \in \hat{x} \rightarrow (\hat{x}/F \mathrel{\Box\!\!\rightarrow} \neg Fy))$

Now consider the schema

$(\star)$ $\forall y(y \in x \to (\varphi \,\Box\!\!\to \psi)) \to (\varphi \,\Box\!\!\to \forall y(y \in x \to \psi))$, where φ is free for y

$(\star)$ should be uncontroversial: If φ counterfactually implies $\psi(y)$ for all $y \in x$, then it counterfactually implies the conjunction of those instances. (Note that even if it is contingent what things there are, no set could have had members that it does not actually have.) In particular, we should accept the following two of its instances:

(E) $\forall y(y \in \hat{x} \to (\hat{x}/F \,\Box\!\!\to Fy)) \to (\hat{x}/F \,\Box\!\!\to \forall y(y \in \hat{x} \to Fy))$

(F) $\forall y(y \in \hat{x} \to (\hat{x}/F \,\Box\!\!\to \neg Fy)) \to (\hat{x}/F \,\Box\!\!\to \forall y(y \in \hat{x} \to \neg Fy))$

By (D)-(F), we have:

(G) $(\hat{x}/F \,\Box\!\!\to \forall y(y \in \hat{x} \to Fy)) \vee (\hat{x}/F \,\Box\!\!\to \forall y(y \in \hat{x} \to \neg Fy))$

Since nothing possibly true counterfactually implies something with which it is inconsistent, we have

(H) $(\hat{x}/F \,\Box\!\!\to \forall y(y \in \hat{x} \to Fy)) \to \neg\Diamond\hat{x}/F$

(I) $(\hat{x}/F \,\Box\!\!\to \forall y(y \in x \to \neg Fy)) \to \neg\Diamond\hat{x}/F$

By (G)-(I) we have:

(J) $\neg\Diamond\hat{x}/F$

In other words, for any x and qualitative F, the set of things qualitatively indiscernible from x is not possibly divided by F. We may generalize and necessitate to establish the advertised conclusion:

> THE NECESSITY OF DISCERNIBILITY
> Necessarily, if x and y are qualitatively indiscernible, then for all qualitative F, necessarily, Fx iff Fy.

References

Jonathan Bennett. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press, 2003.

Max Black. The identity of indiscernibles. *Mind*, 61:153–64, 1952.

David J. Chalmers. *The Conscious Mind*. Oxford: Oxford University Press, 1996.

Hartry Field. Physicalism. In John Earman, editor, *Inference, Explanation, and Other Frustrations*, pages 271–91. University of California Press, 1992.

Hartry Field. Disquotational truth and factually defective discourse. *Philosophical Review*, 103:405–52, 1994.

Hartry Field. Indeterminacy, degree of belief, and excluded middle. *Noûs*, 34:1–30, 2000.

Hartry Field. *Saving Truth from Paradox*. Oxford University Press, Oxford, 2008.

Kit Fine. Properties, propositions and sets. *Journal of Philosophical Logic*, 6:135–191, 1977.

Kit Fine. The question of realism. *Philosophers' Imprint*, 1:1–30, 2001.

Peter Fritz and Jeremy Goodman. Counterfactuals and propositional contingentism. unpublished a.

Peter Fritz and Jeremy Goodman. Higher-order contingentism, part 1: Closure and generation. unpublished b.

Alan Gibbard and William Harper. Counterfactuals and two kinds of expected utility. In Robert Stalnaker William L. Harper, editor, *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pages 153–190. D. Reidel, 1981 [1978].

Jeremy Goodman. Matter and mereology. unpublished a.

Jeremy Goodman. Williamson's earrings. unpublished b.

Ian Hacking. The identity of indiscernibles. *Journal of Philosophy*, 72(9): 249–256, 1975.

Alan Hajek. Most counterfactuals are false. Unpublished manuscript, manuscript 2011.

John Hawthorne. Chance and counterfactuals. *Philosophy and Phenomenological Research*, 70(2):396–405, 2005.

John Hawthorne. *Metaphysical Essays*. Oxford: Oxford University Press, 2006.

John Hawthorne and Andrew McGonigal. The many minds account of vagueness. *Philosophical Studies*, 138(3):435–40, 2008.

John (O'Leary-) Hawthorne. The bundle theory of substance and the identity of indiscernibles. *Analysis*, 55:191–96, 1995.

James Higginbotham. Linguistic theory and Davidson's program in semantics. In Ernest Lepore, editor, *Truth and Interpretation: perspectives on the philosophy of Donald Davidson*, pages 29–48. Oxford: Basil Blackwell, 1986.

Nathan Klinedinst. Quantified conditionals and conditional excluded middle. *Journal of Semantics*, 28(1):149–170, 2011.

Angelika Kratzer. *Modals and Conditionals*. Oxford: Oxford University Press, 2012.

Angelika Kratzer. Chasing hook. In Lee Walters and John Hawthorne, editors, *Conditionals, probability, and paradox: themes from the philosophy of Dorothy Edgington*. Oxford: Oxford University Press, forthcoming.

Sarah-Jane Leslie. "If", "unless", and quantification. In Robert Stainton and Christopher Viger, editors, *Compositionality, Context, and Semantic Values*, pages 3–30. Springer Netherlands, 2008.

Sarah-Jane Leslie. Essence, plenitude, and paradox. *Philosophical Perspectives*, 25:277–96, 2011.

David Lewis. *Counterfactuals*. Oxford: Basil Blackwell, 1973.

David Lewis. Counterfactual dependence and time's arrow. *Noûs*, 13:455–76, 1979.

David Lewis. Causal decision theory. *Australasian Journal of Philosophy*, 59 (1):5–30, 1981.

David Lewis. Humean supervenience debugged. *Mind*, 103:473–90, 1994.

Brian McLaughlin and Karen Bennett. Supervenience. 2005. Available at `http://plato.stanford.edu/entries/supervenience/`.

Sarah Moss. Subjunctive credence and semantic humility. *Philosophy and Phenomenological Research*, 87(2):251–278, 2013.

W. V. O. Quine. *Methods of Logic*. Holt, Rinehart, and Winston, New York, 1950.

Stewart Shapiro. And 'i' for an i: singular terms, uniqueness, and reference. *Review of Symbolic Logic*, 5(3):380–415, 2012.

Theodore Sider. *Writing the Book of the World*. Oxford: Oxford University Press, 2011.

Brian Skyrms. The prior propensity account of subjunctive conditionals. In Robert Stalnaker William L. Harper, editor, *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pages 259–265. D. Reidel, 1981 [1980].

Robert Stalnaker. A defense of conditional excluded middle. In William L. Harper, Robert Stalnaker, and Glenn Pearce, editors, *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pages 87–104. D. Reidel, 1981.

Robert Stalnaker. Letter to David Lewis. In William L. Harper, Robert Stalnaker, and Glenn Pearce, editors, *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pages 151–152. D. Reidel, 1981 [1972].

Robert Stalnaker. *Inquiry*. Cambridge, MA: MIT Press, 1984.

Robert Stalnaker. *Mere Possibilities*. Princeton: Princeton University Press, 2012.

Kai von Fintel and Sabine Iatridou. If and when 'if'-clauses can restrict quantifiers. manuscript, MIT, 2002.

J. Robert G. Williams. Defending conditional excluded middle. *Noûs*, 44(4): 650–668, 2010.

J. Robert G. Williams. Counterfactual triviality. *Philosophy and Phenomenological Research*, 85(3):648–670, 2012.

Timothy Williamson. Criteria of identity and the axiom of choice. *Journal of Philosophy*, 83(7):380–94, 1986.

Timothy Williamson. *Modal Logic as Metaphysics.* Oxford: Oxford University Press, 2013.