

Conditional Excluded Middle without the Limit Assumption

ERIC SWANSON

University of Michigan, Ann Arbor

David Lewis and Robert Stalnaker have some sharp disagreements over how we talk about modality. For example, Stalnaker endorses

The limit assumption for counterfactuals: “for every possible world i and non-empty proposition A , there is at least one A -world minimally different from i .” (Stalnaker 1980, 89; see also Pollock 1976, 18–20, Herzberger 1979, and Warmbröd 1982.)

But Lewis famously objects that counterfactually supposing that a given line had been more than an inch long will not yield an A -world minimally different from i . “Just as there is no shortest possible length above 1’,” he writes, “so there is no closest world to ours among the worlds with lines more than an inch long” (1973a, 20–21; see also 1981b, 228–230).

Lewis and Stalnaker also agree about many features of our talk about modality. This paper starts by exposing one subtle, easily overlooked area of agreement, between Stalnaker’s theory of counterfactuals and Lewis’s theory of counterparts. In particular, I argue in §1 that despite his strenuous objections to the limit assumption for counterfactuals, Lewis endorses an analogue of the limit assumption in his counterpart theory. In §2 I transpose Lewis’s objection to the limit

For helpful discussion, thanks to John Collins, Bas van Fraassen, Jim Joyce, David Manley, Sarah Moss, Graham Oddie, Bob Pasnau, Robert van Rooij, Jason Stanley, and Rich Thomason; to audiences at Rutgers University, University of Colorado, Boulder, and University of Michigan, Ann Arbor; and to an anonymous referee for *Philosophy and Phenomenological Research*.

assumption for counterfactuals into the key of counterparts. This helps clarify the debate over the limit assumption: *many* theories go awry when analogues of the limit assumption are violated, suggesting that we need a general treatment of violations of limit assumptions. In §3 I introduce such a treatment, which I call *ordering supervaluationism*, and I apply it to limit violations in counterpart theory.

Although I use counterpart theory as my point of departure, the limit assumption for counterfactuals—“the principal vice of Stalnaker’s theory,” in Lewis’s estimation (1973a, 79)—is my real focus here. The limit assumption for counterfactuals is important because it together with the uniqueness assumption—the assumption that there is at most one *A*-world minimally different from *i*—secures the validity of

Conditional excluded middle: Either ‘If it had been that φ , it would have been that ψ ’ is true or ‘If it had been that φ , it would have been that $\neg\psi$ ’ is true. (Lewis 1973a, 79–81)¹

And the validity of conditional excluded middle is the “principal virtue” of Stalnaker’s theory of counterfactuals (Lewis 1973a, 79).

The limit and uniqueness assumptions are hefty commitments. Stalnaker himself says that the uniqueness assumption is “a grossly implausible assumption to make about the kind of similarity relation we use to interpret conditionals” (1980, 89). Because of this he deploys traditional supervaluationism as a post-semantic mechanism to “reconcile the determinacy of abstract semantic theory with the indeterminacy of realistic application” (89). In particular, Stalnaker gets the supervalidity of conditional excluded middle even when the uniqueness assumption is violated.

The limit assumption is not so easily handled in post-semantically: many philosophers have observed that traditional supervaluationism does not help at all. In §4 I explain how ordering supervaluationism *does* let us reconcile Stalnaker’s “abstract semantic theory” with failures of the limit assumption. Because ordering supervaluationism subsumes traditional supervaluationism as a special case, I am in effect extending Stalnaker’s approach so that Stalnakerians can handle both failures of uniqueness and failures of the limit assumption. Ordering supervaluationism thus yields the principle virtue of Stalnaker’s theory—conditional excluded middle—without its putative principal vice.

¹ For other perspectives on the costs and benefits of conditional excluded middle, see (e.g.) Adams 1977, Thomason & Gupta 1980, Skyrms 1982, Cross 1985 and 2009, Forster 1986, Gaskin 1993, Barker 1994, Morton 1997, DeRose 1999, McDermott 1999, Bennett 2003, Higginbotham 2003, Schlenker 2004, Hawthorne 2005, Pizzi & Williamson 2005, Williams 2008 and 2010, and Hájek 2009.

Finally, in §5, I put this approach to work in a related domain, by showing how it can be used to improve some of the standard formulations of causal decision theory.

1. Counterparts and counterfactuals: Variations on a theme

On Lewis's view, your counterparts are not you; they are people "you *would have been*, had the world been otherwise" (1968, 115). Lewis defines the counterpart relation as follows:

...something has for *counterparts* at a given world those things existing there that resemble in closely enough in important respects of intrinsic quality and extrinsic relations, and that resemble it no less closely than do other things existing there. Ordinarily something will have one counterpart or none at a world, but ties in similarity may give it multiple counterparts. (1973a, 39, italics in original)

In other words, in a given world w , an object is a counterpart of you if and only if it meets both of the following conditions:

1. it resembles you well enough;
2. it resembles you best of the objects in w , where 'best' permits ties.²

It will soon be important that, on Lewis's definition, any given counterpart of you resembles you no less closely than anything else in that counterpart's possible world.

Now consider Stalnaker's theory of counterfactuals. Stalnaker holds that a counterfactual "If A , then B " is a statement that targets "... a possible world in which A is true, and which otherwise differs minimally from the actual world. [It] is true (false) just in case B is true (false) in that possible world" (45). There are two significant differences between these theories, abstracting away from their subject matter. First, Lewis's counterpart theory allows that an object may have more than one counterpart in a given world, whereas Stalnaker's official semantics for counterfactuals targets a unique counterfactual

² Lewis's first paper on counterpart theory (1968) poses some interpretive challenges. Lewis does not note any change in his views between 1968 and 1973, and I think it is fair to read him as having this meaning in mind all along. This is the standard understanding in the literature. See, e.g., Feldman 1971, 406; Hazen 1979, 331; Johnston 1989, 383; Heller 1998, 301; and Cresswell 2004, 29.

world, and so depends on supervaluationism to handle the “indeterminacy of realistic application” (Stalnaker 1980, 89).³ Second, because Lewis’s counterpart theory requires that an object’s counterparts resemble it “well enough,” some objects have no counterparts at some worlds. Stalnaker, on the other hand, holds that the selection function is total (1968, 45). So whatever its antecedent, a counterfactual statement targets some counterfactual world (or targets the “absurd world” (45)).

For present purposes the most important similarity between Lewis’s counterpart theory and Stalnaker’s theory of counterfactuals is that they both presuppose that there are objects/worlds that best resemble a given object/world. In Stalnaker’s case this presupposition just is the limit assumption for counterfactuals. Again:

The limit assumption for counterfactuals: “for every possible world i and non-empty proposition A , there is at least one A -world minimally different from i .” (Stalnaker 1980, 89)

In Lewis’s case let us dub the relevant presupposition

The limit assumption for counterparts: for every object o with a counterpart in a possible world w , there is at least one object in w minimally different from o .

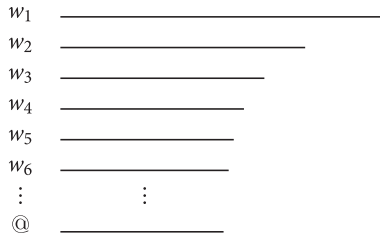
Both Stalnaker and Lewis are committed to the well-definedness of ‘minimally different,’ ‘most closely resembling,’ and the like. They are just committed to their well-definedness with respect to different things. In Stalnaker’s case counterfactual worlds are minimally different from the actual world (and, with appropriate shifting of the world of evaluation, are minimally different from merely possible worlds). In Lewis’s case counterparts are minimally different from actual objects (and, again with appropriate shifting of the world of evaluation, are minimally different from merely possible objects).

2. Limit violations in counterpart theory

Lewis objects to the limit assumption for counterfactuals as follows (1973a, 20; 1973b, 63). Consider a counterfactual with the antecedent “If line L had been more than an inch long.” In some possible world— w_1 , let us say—there is a line very similar to L that is 2 inches long. If there were no worlds in which a very similar line was *shorter*

³ For further discussion see Stalnaker & Thomason 1970; Lewis 1973a; 81–83; van Fraassen 1974; and Stalnaker 1984, 134–135, 140.

than 2 inches long, then on Stalnaker’s theory the counterfactual would target w_1 , and the counterfactual would be true (false) just in case the counterfactual’s consequent is true (false) in w_1 (Stalnaker 1968, 46). But of course there are infinitely many possible worlds with very similar lines that get ever closer to being 1 inch long. For example, consider w_2 , in which the very similar line is $1\frac{1}{2}$ inches long, or w_3 , in which the very similar line is $1\frac{1}{4}$ inches long. There are possible worlds in which the length of the very similar line in that world gets as close as we like to 1 inch without ever equalling 1 inch:

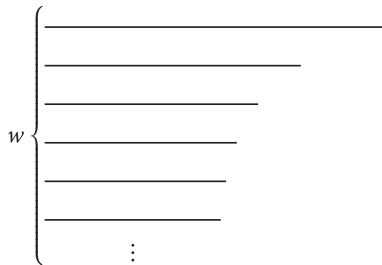


Lewis concludes that Stalnaker has “no right to assume that there always are a smallest antecedent-permitting sphere and, within it, a set of closest antecedent worlds” (1973a, 20). In other words, Lewis thinks Stalnaker is wrong to presuppose that there are possible worlds in which the antecedent is true that differ “minimally” from the actual world. For some antecedents, there will not be a minimal change from the actual world, because for any change making the antecedent true there is another that is ‘more minimal.’

Lewis later steps back from this conclusion, in response to pressure from Pollock 1976, Herzberger 1979, and Stalnaker 1980, among others. In particular, he allows that “some sort of coarse-graining” might allow us to “imitate the finite case by ignoring ... respects of difference that make the possible worlds infinite in number” (1981b, 230). (And Lewis himself is explicit that “some similarities ... may count for nothing”; they “may have zero weight” in the similarity ordering relevant to the evaluation of a counterfactual (1979, 42).) For example, it might be that the miniscule differences between the lengths of the lines at $w_{10^{100}}$, $w_{10^{100}+1}$, $w_{10^{100}+2}$, and so on simply don’t matter for purposes of evaluating counterfactuals. The relevant similarity ordering would then ‘flatten out,’ so that some worlds with lines of different lengths are treated as equally similar to the world of evaluation. After briefly discussing this kind of strategy Lewis offers a fairly complicated example that poses problems for the limit assumption “even if we stick to atomistic, all-or-nothing respect[s] of similarity and difference” (1981b, 229). In the end Lewis says that he considers coarse-graining to be a

“formal option,” but insists that it is “an open question” whether such a strategy “can be built into an intuitively adequate analysis of counterfactuals.” He thinks it is “best to remain neutral on the Limit Assumption” and to develop theories of counterfactuals and modals that “do not need the Limit Assumption to work properly” (1981b, 230). At this point in the dialectic I simply want to follow Lewis. It is also an open question whether coarse-graining “can be built into an intuitively adequate analysis” of counterparts. So it is “best to remain neutral” on the limit assumption for counterparts. Ordering supervenientism gives us a way to do this.

As I argued in §1, Lewis is not neutral with respect to the limit assumption for counterparts; he is committed to the well-definedness of a notion of minimality in his counterpart theory. To see how this gets him into trouble, consider a possible world w that contains lines just like all those in the non-actual worlds represented above.



If L has a counterpart in w , then Lewis is wrong—by his own lights—to hold that its counterpart resembles the line “no less closely than do other things existing there” (Lewis 1973a, 39).

In principle Lewis could tollens my ponens here, by simply denying that L has counterparts in w . But this strategy is extremely counterintuitive. w has no lack of good candidates to be L 's counterpart—the problem is rather that it has a surfeit of such candidates. For example, if the sequence of lines in w terminated at some point as the lines come closer and closer to being 1 inch long, then the line at the termination point would be L 's counterpart. *Any* of the lines in w would do perfectly well at being L 's counterpart, if the sequence terminated in that line. It is very odd to think that solely because an object has *too many* good potential counterparts in a world, every de re claim about how things are with the object at that world is semantically incomplete. And intuitively, claims about how things would be with L had things been as they are at w may be perfectly fine, and may be true or false. For example, it seems that if all the relevant lines at w are red, then if things had been as they are at w , L would have been

red. If Lewis simply denies that L has counterparts in w , then he cannot explain this judgment.

3. Ordering supervaluationism

Perhaps Lewis would be willing to concede these costs in the interest of preserving his otherwise elegant and fruitful counterpart theory.⁴ But as I now explain, there is a more conservative tack: ordering supervaluationism lets us reconcile Lewis's counterpart theory with failures of the limit assumption for counterparts, just as it lets us reconcile Stalnaker's theory of counterfactuals with failures of the limit assumption for counterfactuals. Ordering supervaluationism is not an ad hoc patch but a generalization of traditional supervaluationism. We can deploy it both in cases that would usually be thought to call for traditional supervaluationism—cases in which multiple interpretations are equally good, and no interpretations are better than those—and in cases in which for every interpretation of an expression, another is better. Such cases are common enough that ordering supervaluationism has a broad range of applications (Swanson 2010b).

On the traditional supervaluationist picture, a sentence is supertrue iff it is true on every admissible interpretation (Mehlberg 1958, 257–258; van Fraassen 1966, 486–487; Fine 1975, 278). Ordering supervaluationism enriches traditional supervaluationism as follows. Whether a sentence is ordering supertrue is a matter of an ordering of interpretations and the sentence's truth value relative to the ordered interpretations.⁵ A sentence is **ordering supertrue** iff there is some bound on the set of interpretations such that that sentence is true according to every interpretation that is at least as good as that bound. Traditional supervaluationism is subsumed by ordering supervaluationism because if there are best interpretations, then a sentence is ordering supertrue iff it is true according to all the best interpretations. In such a case traditional supervaluationism will treat the best interpretations as the admissible interpretations.

Exactly what it is for one interpretation to be “best” or “better” than another or “at least as good as” another depends on the domain to which we are applying ordering supervaluationism. Consider a Lewisian counterpart theorist using ordering supervaluationism to help with violations of the limit assumption for counterparts. Following

⁴ For a nice sketch of Lewis's cost/benefit approach, see his 1986, 3–5.

⁵ Throughout I follow common usage among philosophers in using ‘order’ to refer to what are, strictly speaking, preorders (see, e.g., Lewis 1973a, 48). Preorders are reflexive and transitive relations; mathematicians often use ‘order’ to refer to anti-symmetric preorders in particular.

Stalnaker 1986, each interpretation over which such a theorist supervaluates treats a particular potential counterpart as if it were the unique counterpart at the relevant world. For example, if world w has two potential counterparts of o that resemble o equally well, then there are two interpretations of ‘ o ’s counterpart at w ’: one for each of the potential counterparts. Now suppose that we have two interpretations of the de re modal claim

(1) Line L exhibits feature F at world w .

On interpretation a , (1) treats the $1\frac{1}{4}$ inch line as L ’s counterpart. On interpretation b , (1) treats the $1\frac{1}{8}$ inch line as L ’s counterpart. In this case interpretation b is better than interpretation a . More generally: interpretation i of a de re modal claim about object o is better than interpretation i' iff the potential counterpart of o targeted by i better resembles o than the potential counterpart targeted by i' .

(1) is ordering supertrue iff there is some bound on the set of interpretations such that (1) is true according to every interpretation at least as good as that bound. For example, it is ordering supertrue that line L is red at world w if all the lines no longer than the $1\frac{1}{4}$ inch long line (or the $1\frac{1}{8}$ inch long line, or the $1\frac{1}{16}$ inch long line, or ...) are red. More generally: a de re modal claim of the form

(2) Object o exhibits feature F at world w .

is ordering supertrue iff there is some bound on the set of potential counterparts of o in w such that all the potential counterparts that resemble o at least as closely as the potential counterparts in that bound are F .⁶

What is a “bound” on the set of interpretations of a sentence? If the ordering of interpretations is total, then any given interpretation can serve as a bound on the set of interpretations, because every interpretation is comparable to every other interpretation. If the

⁶ This approach will yield the supervalidities associated with Stalnaker’s supervaluated counterpart theory (1986, 136–137). For example, the necessity of identity

$$\forall x\forall y(x = y \rightarrow \Box x = y)$$

will be supervalid, but its inner necessitation

$$\forall x\forall y\Box(x = y \rightarrow \Box x = y)$$

will not be. Thanks to a referee for *Philosophy and Phenomenological Research* for posing questions about the logic of ordering supervaluated counterpart theory.

ordering of interpretations is merely partial, in virtue of there being interpretations that are incomparable to each other, then it is less obvious what a bound is. Nevertheless the differences between competing answers to this question, discussed at length in Swanson 2010a, are not important for present purposes. We can simply use standard order-theoretic objects called ‘cutsets’ as bounds on partially ordered sets of interpretations. (A **cutset** of an order is a set that includes an element from each maximal chain of that order. A **chain** is a totally ordered subset of an order. A chain is **maximal** iff it is not a proper subset of any other chains.⁷) Intuitively, a maximal chain of a given order is a maximally specified ‘way of improving’ by the lights of that order. And, again intuitively, a cutset provides a way to slice across a partial order and thereby produce a subset of the order that represents each of the possible ‘ways of improving’ in the order (Swanson 2010a).

Here is a toy example. Some possible world contains infinitely many potential counterparts of me, all of which are smarter than I am, some of which are simply more ‘street smart’ and some of which are simply more ‘book smart.’ Consider some such world in which my potential counterparts come closer and closer to matching my actual level of street smarts and book smarts. Suppose for sake of argument that all the street smarter potential counterparts of me are incomparable (with respect to which more closely resembles me) to all the book smarter counterparts of me.⁸ Then a bound on the ordering of interpretations of a *de re* sentence about how things are with me at that world will have to include one interpretation of the sentence on which my counterpart is street smarter than I am, and one interpretation of the sentence on which my counterpart is book smarter than I am. Any pair of such interpretations would constitute a cutset in this case: one element would represent the street smarter potential counterparts and the other element would represent the book smarter potential counterparts. If all the potential counterparts that resemble me at least as closely as the elements in such a cutset share some feature, then it will be ordering supertrue that I have that feature in that world.

Here is a brief review of the dialectic. Lewis objects to the limit assumption for counterfactuals. But Lewis is subject to an analogous objection to the limit assumption for counterparts. Ordering

⁷ For a good introduction to the relevant order theory, see (e.g.) Schröder 2002.

⁸ This sort of ‘street smarts’ and ‘book smarts’ example is traditional, but nothing turns on whether it is good example of incomparability. For suggestive reasons to think that many natural language comparatives do not induce total orderings see McConnell-Ginet 1973, 135–137; Kamp 1975; Cresswell 1976, 266; and Klein 1980.

supervaluationism gives Lewis a straightforward, low-cost response to this *tu quoque* objection. And as I am about to argue, ordering supervaluationism also lets Stalnaker respond to Lewis's objection to the limit assumption for counterfactuals, while holding on to his original semantics for counterfactuals.

4. Ordering supervaluationism and counterfactuals

It will be helpful first to see why traditional supervaluationism doesn't help with violations of limit assumptions. As I mentioned earlier, Stalnaker applies traditional supervaluationism to handle failures of the uniqueness assumption.

When the uniqueness assumption fails to hold for a comparative similarity relation among possible worlds, then the selection function in terms of which conditionals are interpreted ... is left underdetermined by that relation. Many selection functions may be compatible with the comparative similarity relation, and it would be arbitrary to choose one over the others. (Stalnaker 1980, 96)

It is natural to apply traditional supervaluationism here because it is natural, when we are confronted with underdetermination, to ask whether any ways of resolving the underdetermination will produce results that differ significantly from the others ways of resolving the underdetermination. If all the ways of resolving the underdetermination would have what are, for our purposes, the same final results, then it is safe to ignore the underdetermination and focus on those final results. This is the guiding thought of traditional supervaluationism.

The guiding thought also contains the seed of the reason why traditional supervaluationism does not help with limit assumption violations. In particular, when there is a limit assumption violation there are really no ways of *resolving* the underdetermination, insofar as every potential way of resolving the underdetermination is imperfect. That is, for any potential way of resolving the underdetermination, another does better at resolving it. As Stalnaker puts the point,

... if the limit assumption were to fail, there would be too few candidates to be the selection function rather than too many. Any selection function would be forced to choose worlds which were less similar to the actual world than other eligible worlds. This is why the supervaluation method does not provide a way

to avoid making the limit assumption. (1980, 96; see also Lewis 1973a, 82 and Stalnaker 1984, 140.)

And note that to supervaluate over worlds that are less similar to the actual world than other eligible worlds is in effect to use supervaluationism to emulate a strict semantics for counterfactuals, with all the problems such semantics bring.⁹

The guiding thought of ordering supervaluationism, by contrast, is that if there is some resolution of ‘good enough’ such that all the good enough ways of resolving some underdetermination have the same final results, then it is safe to ignore the underdetermination and focus on those final results. The fact that the ‘resolutions’ are imperfect is mitigated by the fact that we ask not about all the imperfect resolutions, but only about the resolutions that are good enough, again for some resolution of ‘good enough.’ And to connect this intuitive motivation to the more formal characterization I gave earlier: a bound on the ordering of interpretations represents a resolution of ‘good enough.’ So one way to ask about imperfect interpretations that are good enough is to ask about interpretations that are at least as good as the interpretations in a suitably good bound, exactly as ordering supervaluationism does.

When we apply ordering supervaluationism to counterfactuals, we thereby renounce the putative “vice” (Lewis 1973a, 79) involved in using Stalnaker’s semantics. The semantics has a selection function that yields, for each possible world i and non-empty proposition A , exactly one A -world minimally different from i (Stalnaker 1980, 89). We then use ordering supervaluationism as a post-semantic mechanism to handle *both* failures of the uniqueness assumption and failures of the limit assumption, just as Stalnaker uses traditional supervaluationism to handle failures of the uniqueness assumption. One interpretation of given counterfactual will be “better” than another, for purposes of ordering supervaluationism, iff the first interpretation’s selection function yields a possible world that is closer to the world of evaluation than the possible world yielded by second interpretation’s selection function. If the relevant ordering is total, then it is trivial to show that Stalnaker’s semantics supplemented with ordering

⁹ For influential discussions of the problems see Stalnaker 1968, 48–50 and Lewis 1973a, 4–13. Recently von Fintel 2001 and Gillies 2007 have defended strict analyses of counterfactuals but—tellingly—their approaches both make the limit assumption. It is hard to make sense of strict analyses of counterfactuals without the limit assumption (or a substitute for it, like ordering supervaluationism) because when the limit assumption fails it is hard to see how such analyses could avoid quantifying over worlds that are irrelevant to the truth or falsity of the counterfactual.

supervaluationism in the post-semantics emulates Lewis's original semantics for counterfactuals in the following sense: a counterfactual is ordering supertrue given Stalnaker's semantics iff it is true given Lewis's original semantics.¹⁰

If we do not want to assume that the relevant ordering is total, then Stalnaker's semantics with ordering supervaluationism in the post-semantics can emulate the semantics for modals and conditionals developed by Angelika Kratzer (1976, 1977) and Frank Veltman (1976), and later adopted by Lewis (1981b). Swanson 2010a shows that the Kratzer/Veltman semantics in effect uses maximal antichains as bounds on the orders relative to which modals and counterfactuals are evaluated. (An **antichain** of an order is a set of elements each of which is incomparable to the others. An antichain is **maximal** iff it is not a proper subset of any other antichains.) To emulate the Kratzer/Veltman semantics with ordering supervaluationism, then, we too need to use maximal antichains as bounds. If we do this, then a counterfactual is ordering supertrue given Stalnaker's semantics iff it is true given the Kratzer/Veltman semantics. Because the Kratzer/Veltman semantics can make counterintuitive predictions when there are infinitely many incomparabilities between worlds, we might prefer to return to the view of bounds that I sketched earlier, on which they are cutsets. The total theory would then emulate the revised ordering semantics for counterfactuals laid out in Swanson 2010a: a counterfactual is ordering supertrue given Stalnaker's semantics iff it is true given Swanson's semantics. Either way a counterfactual will be ordering supertrue iff there is some bound on the relevant order such that the counterfactual is true according to every interpretation that is at least as good as one of the interpretations in the bound.

Now let us turn to the logic of counterfactuals. In his defense of conditional excluded middle, Stalnaker deploys traditional supervaluationism to handle cases in which "the selection functions that are actually used in making and interpreting counterfactual conditional statements correspond to orderings of possible worlds that admit ties and incomparabilities" (Stalnaker 1980, 90). This is in part because

... in the conditional logic C2 (the logic of the theory I am defending), the principle of conditional excluded middle, $(A \Box \rightarrow B) \vee (A \Box \rightarrow \neg B)$, remains valid when supervaluations are added, even though there may be cases where neither

¹⁰ On the assumption that the ordering is total see Lewis 1971, 76; 1973a, 48–50, 123–124, and 127–128; and 1973b, 64 and 80–81. For early skepticism about the plausibility of this assumption see Thomason 1970, 411–412.

$(A \Box \rightarrow B)$ nor $(A \Box \rightarrow \neg B)$ is [super]true. It may be that neither disjunct is made true by every arbitrary extension of a given partial interpretation, but it will always be that each arbitrary extension makes true one disjunct or the other. (1980, 90–91)¹¹

Stalnaker then draws an analogy between conditional excluded middle and ordinary excluded middle. By the supervaluationist's lights, the following are all jointly consistent:

1. it is not supertrue that a given color patch is yellow (since it does not count as yellow on some admissible interpretations of 'yellow');
2. it is not supertrue that the color patch is not yellow (since it does not count as not yellow on some admissible interpretations of 'yellow');
3. it is supertrue that the color patch is either yellow or not yellow (since it counts as either yellow or not yellow on every admissible interpretation of 'yellow').

As Stalnaker puts it,

The theory of supervaluations, applied to this logic of conditionals, gives the principle of conditional excluded middle the same status as it gives the simple principle of excluded middle. $(B \vee \neg B)$ is logically true [i.e., it is supertrue] even though sometimes neither B nor $\neg B$ is true. (91)

In short, the 'supervalidities' just are the validities relative to all the admissible interpretations. Thus a supervaluationist can get the supervalidity of (conditional) excluded middle without having to endorse (conditional) bivalence.

Analogous reasoning applies with ordering supervaluationism. A sentence that is true according to every interpretation will be ordering supertrue, because for any bound the sentence will be true

¹¹ The strategy of using supervaluations to handle failures of the uniqueness assumption was first articulated in print by Stalnaker & Thomason 1970, 27–28. (See also Lewis 1973a 81–83; 1973b, 61–63; and 1981a, 330; and van Fraassen 1974.) One aspect of the technique that Stalnaker sketches here—handling partial orders by supervaluating over their total extensions—goes awry when the partial order involves infinitely many incomparabilities. Swanson 2010b explains how ordering supervaluationism treats such cases successfully.

according to all the interpretations that are at least as good as the interpretations in that bound. So validities according to every interpretation will be ordering supervalidities. In particular, whether or not the limit assumption holds, conditional excluded middle will be ordering supervalid.

5. Causal decision theory without the limit assumption

In a 1972 letter to Lewis, Stalnaker sketches a causal decision theory as follows:

... if P is a probability distribution, and if for any A and B , $P_B(A) = P(B \Box \rightarrow A)$, then P_B is a probability distribution too ... What it is good for ... is deliberation—the calculation of expected utilities.

Let S_1, \dots, S_n be an exhaustive set of mutually exclusive propositions characterizing the alternative possible outcomes of some contemplated action. Let A be the proposition that I perform the action. My suggestion is that expected utility should be defined as follows: $u(A) = P(A \Box \rightarrow S_1) \times u(S_1) + \dots + P(A \Box \rightarrow S_n) \times u(S_n)$. (Stalnaker 1972, 151)

This decision theory presupposes that conditional excluded middle is valid (see also Lewis 1981a, 333). As Jim Joyce puts it, “The Stalnaker formula makes no sense without Conditional Excluded Middle since the function $P(A \Box \rightarrow \bullet)$ is not additive unless $\neg(A \Box \rightarrow E)$ and $A \Box \rightarrow \neg E$ are equivalent” (1999, 169; see also Lewis 1976, 141–142). To spell this out a bit: Stalnaker’s P_{BS} are not guaranteed to be probability distributions unless conditional excluded middle is valid, and if they are not probability distributions, Stalnaker’s derivation of expected utility will lead us astray.

For example, suppose that the A worlds alternate between $A \wedge S$ worlds and $A \wedge \neg S$ worlds, as they more and more closely resemble the world of evaluation, and suppose that the A worlds do not ever culminate in a world that most closely resembles the world of evaluation. Then $A \Box \rightarrow S$ and $A \Box \rightarrow \neg S$ are both false, on Lewis’s semantics. Suppose that the deliberating agent believes that $A \Box \rightarrow S$ and $A \Box \rightarrow \neg S$ are both false. In particular, $P(A \Box \rightarrow S) = 0$ and $P(A \Box \rightarrow \neg S) = 0$. Suppose also that $u(S) > u(\neg S) > 0$. Intuitively, it should be that $u(A) > 0$: after all, $P(A \Box \rightarrow (S \vee \neg S)) = 1$, and $u(S \vee \neg S) > 0$. But Stalnaker’s formula would have the counterintuitive result that $u(A) = 0$, relative to the partition $\{S, \neg S\}$, because Stalnaker’s formula would sum the product of $u(S)$ and 0 and the product of $u(\neg S)$ and 0.

Given the conventional wisdom that without the limit assumption we lose the validity of conditional excluded middle, it is no surprise to see Allan Gibbard and Bill Harper fleshing out Stalnaker's sketch with the following caveat: "we have imposed the Stalnaker-like constraint that there is a unique world W_a which would eventuate from performing a at t " (1981, 189). They go on to conjecture that "In circumstances where these assumptions break down, it would seem to us that using conditionals to compute expected utility is inappropriate. A more general approach is needed to handle such cases" (189).

On the contrary, with the help of ordering supervaluationism we can use conditionals to compute expected utility even when the "Stalnaker-like constraint" is not met.¹² There are many ways in which one might deploy ordering supervaluationism here. The following is very straightforward: do not ask about the value of $u(A)$ directly; rather, ask which claims about the value of $u(A)$ are ordering supertrue. If the limit and uniqueness assumptions for counterfactuals are satisfied, then these questions obviously will get the same answers. But suppose the limit assumption is not satisfied. Then the post-semantics looks at claims about the value of $u(A)$ relative to interpretations of the various ' $A \Box \rightarrow S_x$ ' counterfactuals that treat antecedent worlds closer and closer to the world of evaluation as the world yielded by the counterfactual's selection function. As before, a claim about the value of $u(A)$ will be ordering supertrue iff there is some bound B on interpretations of that claim with the following property: the claim is true according to all the interpretations the selection functions of which yield worlds at least as close to the world of evaluation as do the interpretations in B .

It will be worth considering a few different examples. The first two involve similarity orderings that are total; the third generalizes to a merely partial order. In each example we suppose that the limit assumption is violated: for every A world, another better resembles the actual world. Without ordering supervaluationism, the Stalnaker semantics cannot handle a counterfactual with A as its antecedent, so Stalnaker's definition of expected utility will crash. But with ordering supervaluationism, we get intuitively plausible results.

¹² Some philosophers argue that the Stalnaker-like constraint might fail in a special way because of indeterminacy or objective chance in the world. Lewis, for example, writes that "If the world is the chancy way I mostly think it is, there's nothing at all arbitrary or indeterminate about the counterfactuals in the full patterns [of causal dependence]. They are flatly, determinately false" (1981a, 329; see also Hájek & Hall 1994, Joyce 1999, and Hájek 2009). To keep my project in this paper tractable, I bracket the issues raised by this kind of indeterminacy.

Example 1

In addition to supposing that for every A world, another better resembles the actual world, suppose that all the A worlds are S worlds, and suppose that the deliberating agent knows all this. Then treating any given A world as the image of the world of evaluation under the selection function will make ' $A \Box \rightarrow S$ ' true, and will make ' $P(A \Box \rightarrow S) = 1$ ' true. So ' $A \Box \rightarrow S$ ' and ' $P(A \Box \rightarrow S) = 1$ ' will both be ordering supertrue. Moreover, suppose that $u(S) = 10$ and $u(\neg S) = 2$. Then relative to the partition $\{S, \neg S\}$, treating any given A world as the image of the world of evaluation under the selection function will yield

$$\begin{aligned} u(A) &= P(A \Box \rightarrow S) \times u(S) + P(A \Box \rightarrow \neg S) \times u(\neg S) \\ &= 1 \times 10 + 0 \times 2 \\ &= 10 \end{aligned}$$

So ' $u(A) = 10$ ' will be ordering supertrue.

Example 2

Return to the example on which the A worlds alternate between $A \wedge S$ worlds and $A \wedge \neg S$ worlds as they more and more closely resemble the world of evaluation, and $u(S) > 0$ and $u(\neg S) > 0$. In particular, suppose that $u(S) = 10$ and $u(\neg S) = 2$. Suppose that the deliberating agent knows all this. Then treating any given A world as the image of the world of evaluation under the selection function will either make ' $A \Box \rightarrow S$ ' and ' $P(A \Box \rightarrow S) = 1$ ' true or make ' $A \Box \rightarrow \neg S$ ' and ' $P(A \Box \rightarrow \neg S) = 1$ ' true. So treating any given A world as the image of the world of evaluation under the selection function will either yield

$$\begin{aligned} u(A) &= P(A \Box \rightarrow S) \times u(S) + P(A \Box \rightarrow \neg S) \times u(\neg S) \\ &= 1 \times 10 + 0 \times 2 \\ &= 10 \end{aligned}$$

or yield

$$\begin{aligned} u(A) &= P(A \Box \rightarrow S) \times u(S) + P(A \Box \rightarrow \neg S) \times u(\neg S) \\ &= 0 \times 10 + 1 \times 2 \\ &= 2 \end{aligned}$$

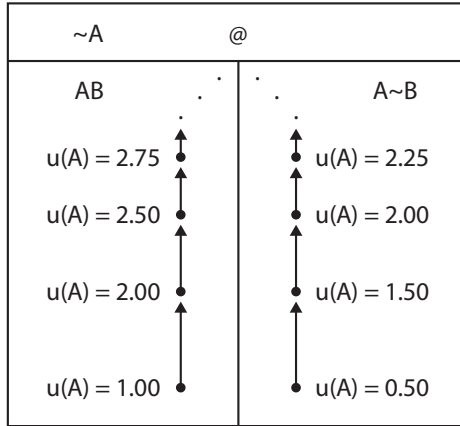
So treating any given A world as the image of the world of evaluation under the selection function will yield

$$P(A \Box \rightarrow S) \times u(S) + P(A \Box \rightarrow \neg S) \times u(\neg S) \in \{2, 10\}$$

So this will be ordering supertrue, and thus ‘ $u(A) \in \{2, 10\}$ ’ will be ordering supertrue. Despite its imprecision, this kind of verdict can often be helpful in decision making.¹³

Example 3

In addition to supposing that for every A world, another better resembles the actual world, suppose that the similarity ordering of the A worlds is merely partial: the B worlds and the $\neg B$ worlds are incomparable with respect to their similarity to the actual world. Suppose that the values of $u(A)$ treating particular worlds (represented as solid dots) as the image of the world of evaluation under the selection function are as below.¹⁴



¹³ Note that in this example—unlike example 1—all we can say about the value of $P(A \Box \rightarrow S)$ is that

$$P(A \Box \rightarrow S) \in \{0, 1\}$$

is ordering supertrue. This is meager guidance for decision making. So this example illustrates the importance of applying ordering supervaluationism to claims about the value of $u(A)$. Applying ordering supervaluationism to claims about the value of $P(A \Box \rightarrow S)$ and calculating $u(A)$ on the basis of the result gives the wrong results in cases like this one.

¹⁴ Here is one way to get such values. The $A \wedge B$ worlds are ordered by the positive integers: $(A \wedge B)_1, (A \wedge B)_2, (A \wedge B)_3, \dots$. The $A \wedge \neg B$ worlds are ordered by the negative integers: $(A \wedge \neg B)_{-1}, (A \wedge \neg B)_{-2}, (A \wedge \neg B)_{-3}, \dots$. The greater the absolute value of a world’s associated integer, the closer the world is to the actual world. The outcomes $\{S_1, S_{-1}, S_2, S_{-2}, S_3, S_{-3}, \dots\}$ are such that $(A \wedge B)_n \Box \rightarrow S_n$, and the deliberating agent knows this. $u(S_1) = 1.00, u(S_2) = 2.00, u(S_3) = 2.50, u(S_4) = 2.75, \dots$; and $u(S_{-1}) = 0.50, u(S_{-2}) = 1.50, u(S_{-3}) = 2.00, u(S_{-4}) = 2.25, \dots$

In this example, as in example 2, we clearly do not want to say that it is ordering supertrue that $u(A)$ is *equal* to any particular value. Nevertheless we can say (for example) that it is ordering supertrue that $2.25 \leq u(A) < 3$, because that sentence is true from some bound on up. For example, it remains true as we proceed closer to the actual world from the lower bound consisting of the $A \wedge B$ world relative to which $u(A) = 2.50$ and the $A \wedge \neg B$ world relative to which $u(A) = 2.25$. Again, even imprecise verdicts like this one can often be helpful in decision making.

6. Conclusion

Despite their disagreements over the limit assumption for counterfactuals, Lewis and Stalnaker are both committed to limit assumptions. I argued that Lewis should save his theory by appealing to ordering supervaluationism—but then, so too should Stalnaker. For Stalnaker, the use of ordering supervaluationism as a post-semantic mechanism to handle violations of the limit assumption for counterfactuals naturally extends his use of traditional supervaluationism as a post-semantic mechanism to handle violations of the uniqueness assumption. The total picture of counterfactuals that results is one that blends attractive features of Stalnakerian accounts and attractive features of Lewisian accounts. In particular, we get the ordering supervalidity of conditional excluded middle—and an attractively simple causal decision theory—without having to commit ourselves to either the uniqueness assumption or the limit assumption.

References

- Adams, Robert Merrihew. 1977. "Middle Knowledge and the Problem of Evil." *American Philosophical Quarterly*, 14(2): 109–117.
- Barker, Stephen J. 1994. "Conditional Excluded Middle, Conditional Assertion, and 'Only If'." *Analysis*, 53(4): 254–261.
- Bennett, Jonathan. 2003. *A Philosophical Guide to Conditionals*. Oxford University Press, Oxford.
- Cresswell, Max. 1976. "The Semantics of Degree." In: *Montague Grammar*, Barbara H. Partee, editor, 261–292. Academic Press, New York.
- Cresswell, Max. 2004. "Adequacy Conditions for Counterpart Theory." *Australasian Journal of Philosophy*, 82(1): 28–41.
- Cross, Charles B. 1985. "Jonathan Bennett on 'Even If'." *Linguistics and Philosophy*, 8: 353–357.
- Cross, Charles B. 2009. "Conditional Excluded Middle." *Erkenntnis*, 70: 173–188.

- DeRose, Keith. 1999. "Can It Be That It Would Have Been Even Though It Might Not Have Been?" *Philosophical Perspectives*, 13: 385–413.
- Feldman, Fred. 1971. "Counterparts." *Journal of Philosophy*, 68: 406–409.
- Fine, Kit. 1975. "Vagueness, Truth and Logic." *Synthese*, 265–300.
- von Fintel, Kai. 2001. "Counterfactuals in a Dynamic Context." In *Ken Hale: A Life in Language*, Michael Kenstowicz, editor. MIT Press, Cambridge.
- Forster, Malcolm R. 1986. "Counterfactual Reasoning in the Bell-Epr Paradox." *Philosophy of Science*, 53(1): 133–144.
- van Fraassen, Bas C. 1966. "Singular Terms, Truth-Value Gaps, and Free Logic." *Journal of Philosophy*, 63(17): 481–495.
- van Fraassen, Bas C. 1974. "Hidden Variables in Conditional Logic." *Theoria*, 40: 176–190.
- Gaskin, Richard. 1993. "Conditionals of Freedom and Middle Knowledge." *Philosophical Quarterly*, 43(173): 412–430.
- Gibbard, Allan & William L. Harper. 1981. "Counterfactuals and Two Kinds of Expected Utility." In Harper et al. (1981), 153–190.
- Gillies, Anthony S. 2007. "Counterfactual Scorekeeping." *Linguistics and Philosophy*, 30: 329–360.
- Hájek, Alan. 2009. "Most Counterfactuals Are False." Ms., Australian National University.
- Hájek, Alan & Ned Hall. 1994. "The Hypothesis of the Conditional Construal of Conditional Probability." In *Probability and Conditionals: Belief Revision and Rational Decision*, Ellery Eells & Brian Skyrms, editors, 75–111. Cambridge University Press, Cambridge.
- Harper, William L., Robert Stalnaker & Glenn Pearce, editors. 1981. *Ifs: Conditionals, Belief, Decision, Chance, and Time*. D. Reidel Publishing Company, Dordrecht.
- Hawthorne, John. 2005. "Chance and Counterfactuals." *Philosophy and Phenomenological Research*, 70(2): 396–405.
- Hazen, Allen. 1979. "Counterpart-Theoretic Semantics for Modal Logic." *Journal of Philosophy*, 76(6): 319–338.
- Heller, Mark. 1998. "Property Counterparts in Ersatz Worlds." *Journal of Philosophy*, 95(6): 293–316.
- Herzberger, Hans G. 1979. "Counterfactuals and Consistency." *Journal of Philosophy*, 76(2): 83–88.
- Higginbotham, James. 2003. "Conditionals and Compositionality." *Philosophical Perspectives*, 17: 181–194.
- Johnston, Mark. 1989. "Fission and the Facts." *Philosophical Perspectives*, 3, *Philosophy of Mind and Action Theory*: 369–397.

- Joyce, James M. 1999. *The Foundations of Causal Decision Theory*. Oxford University Press, Oxford.
- Kamp, Hans. 1975. "Two Theories of Adjectives." In *Formal Semantics of Natural Language*, Edward L. Keenan, editor, 123–155. Cambridge University Press, Cambridge.
- Klein, Ewan. 1980. "A Semantics for Positive and Comparative Adjectives." *Linguistics and Philosophy*, 4: 1–45.
- Kratzer, Angelika. 1976. "Was 'Können' und 'Müssen' Bedeuten Können Müssen." *Linguistische Berichte*, 42: 1–28.
- Kratzer, Angelika. 1977. "What *Must* and *Can* Must and Can Mean." *Linguistics and Philosophy*, 1: 337–355.
- Lewis, David K. 1968. "Counterpart Theory and Quantified Modal Logic." *Journal of Philosophy*, 65: 113–126.
- Lewis, David K. 1971. "Completeness and Decidability of Three Logics of Counterfactual Conditionals." *Theoria*, 37: 74–85.
- Lewis, David K. 1973a. *Counterfactuals*. Basil Blackwell Ltd., Malden, MA.
- Lewis, David K. 1973b. "Counterfactuals and Comparative Possibility." In Harper et al. (1981), 57–85.
- Lewis, David K. 1976. "Probabilities of Conditionals and Conditional Probabilities." In Harper et al. (1981), 129–147.
- Lewis, David K. 1979. "Counterfactual Dependence and Time's Arrow." *Nouûs*, 13(4): 455–476.
- Lewis, David K. 1981a. "Causal Decision Theory." In *Philosophical Papers*, 305–339. Oxford University Press, Oxford.
- Lewis, David K. 1981b. "Ordering Semantics and Premise Semantics for Counterfactuals." *Journal of Philosophical Logic*, 10: 217–234.
- Lewis, David K. 1986. *On the Plurality of Worlds*. Blackwell Publishers Inc., Malden, MA.
- McConnell-Ginet, Sally. 1973. *Comparative Constructions in English: A Syntactic and Semantic Analysis*. Ph.D. thesis, University of Rochester.
- McDermott, Michael. 1999. "Counterfactuals and Access Points." *Mind*, 108: 291–334.
- Mehlberg, Henryk. 1958. *The Reach of Science*. University of Toronto Press, Toronto.
- Morton, Adam. 1997. "Can Edgington Gibbard Counterfactuals?" *Mind*, 106(421): 101–105.
- Pizzi, Claudio & Timothy Williamson. 2005. "Conditional Excluded Middle in Systems of Consequential Implication." *Journal of Philosophical Logic*, 34: 333–362.
- Pollock, John L. 1976. *Subjunctive Reasoning*. D. Reidel Publishing Company, Boston.

- Schlenker, Philippe. 2004. "Conditionals as Definite Descriptions (A Referential Analysis)." *Research on Language and Computation*, 2(3): 417–462.
- Schröder, Bernd S. W. 2002. *Ordered Sets: An Introduction*. Birkhäuser, Boston.
- Skyrms, Brian. 1982. "Counterfactual Definiteness and Local Causation." *Philosophy of Science*, 49: 43–50.
- Stalnaker, Robert C. 1968. "A Theory of Conditionals." In Harper et al. (1981), 41–55.
- Stalnaker, Robert C. 1972. "Letter to David Lewis: May 21, 1972." In Harper et al. (1981), 151–152.
- Stalnaker, Robert C. 1980. "A Defense of Conditional Excluded Middle." In Harper et al. (1981), 87–104.
- Stalnaker, Robert C. 1984. *Inquiry*. MIT Press, Cambridge.
- Stalnaker, Robert C. 1986. "Counterparts and Identity." *Midwest Studies in Philosophy*, 11: 121–140.
- Stalnaker, Robert C. & Richmond H. Thomason. 1970. "A Semantic Analysis of Conditional Logic." *Theoria*, 36: 23–42.
- Swanson, Eric. 2010a. "On the Treatment of Incomparability in Ordering Semantics and Premise Semantics." *Journal of Philosophical Logic*. Forthcoming.
- Swanson, Eric. 2010b. "Ordering Supervaluationism, Counterpart Theory, and Ersatz Fundamentality." Ms., University of Michigan Ann Arbor.
- Thomason, Richmond H. 1970. "A Fitch-Style Formulation of Conditional Logic." *Logique et Analyse*, 52: 397–412.
- Thomason, Richmond H. & Anil Gupta. 1980. "A Theory of Conditionals in the Context of Branching Time." *Philosophical Review*, 89(1): 65–90.
- Veltman, Frank. 1976. "Prejudices, Presuppositions and the Theory of Conditionals." In *Amsterdam Papers in Formal Grammar, Volume I*, Jeroen Groenendijk & Martin Stokhof, editors, 248–281. Central Interfaculteit, University of Amsterdam, Amsterdam.
- Warmbröd, Ken. 1982. "A Defense of the Limit Assumption." *Philosophical Studies*, 42: 53–66.
- Williams, J. Robert G. 2008. "Chances, Counterfactuals, and Similarity." *Philosophy and Phenomenological Research*, 77(2): 385–420.
- Williams, J. Robert G. 2010. "Defending Conditional Excluded Middle." *Noûs*, 44(4): 650–668.