

## Chapter 20: Support Theories

Kenny Easwaran

December 7, 2004

### 119: Worlds and consequences

There are two main types of theories for counterfactual conditionals. We have already seen Worlds Theories, which say that ' $A > C$ ' means 'All worlds (in class  $K$ ) satisfying  $A$  also satisfy  $C$ .' Support Theories like Goodman's say something like ' $A > C$ ' means 'There is a true proposition *Support* (satisfying certain conditions) such that  $A \& \textit{Support} \& \textit{Laws}$  entails  $C$ , where *Laws* is the conjunction of all actual causal laws.' The discussion in previous chapters have been attempts at spelling out just what class  $K$  is (e.g., the worlds should be identical to the actual world up to some time  $t$  and should have very few violations of causal laws, etc.).

One advantage of Support Theories is that they account for many independent conditionals in the same way as dependent ones. However, Worlds Theories allow us to consider counterlegals in addition to standard dependent conditionals. Paraconsistency and how to tweak the Worlds Theory... But since Bennett considers both counterlegals and independent conditionals to be somewhat marginal, neither of these considerations will carry much weight.

Bennett suggests that this distinction might not be as clear as it seems, since each of these theories can be expressed in the terms of the other. Bennett says that the Support Theory can be expressed in terms of Worlds Theory, just by letting the condition  $K$  on worlds be that the world satisfies all actual causal laws, as well as *Support*, which we should be able to spell out from our Support Theory. However, he later emphasizes the fact that *Support* is merely quantified over, rather than specified, so the Worlds formulation would need to quantify over  $K$  to be equivalent. In addition, in order for this to be equivalent to the Support Theory it replaces, we would need to consider all *logically* possible worlds, while Worlds Theories discussed earlier seem to only consider all *metaphysically* possible worlds, which may be a reduced class. (This is a further worry that may be addressed to many of Bennett's metaphysical claims about Worlds Theories. If he doesn't subscribe to Lewis' full-fledged views that all conceivable worlds actually exist, then the truth-values of some counterfactuals will turn on whether or not some describable and conceivable situation actually holds in any metaphysically possible world or not.)

Bennett's argument that Worlds Theories can be expressed in terms of Support Theories seems to me substantially weaker. He claims that for any class

$K$  of worlds, there is a proposition that is true exactly of the worlds in that class. We can leave aside the question of whether distinct worlds might satisfy exactly the same propositions, because it doesn't matter whether one or all of a collection of such equivalent worlds are in  $K$ , because either all will satisfy  $C$  or none will. I will also ignore the fact that Support Theories use *Laws* in addition to  $A$  and *Support*, even though most accounts Bennett has given so far allow us to consider some worlds that violate *Laws*. However, it seems to me that the only way to guarantee the existence of a proposition defining the class  $K$  is to allow infinite conjunctions and disjunctions, if we can't assume anything about how the class in question is defined. (There may be infinitely many propositions in *Laws* or *Support* as well, but it is safe to treat those merely as lists of propositions, rather than actual conjunctions, because they only turn up as premises in an entailment. It is impossible to take the disjunction of two infinite lists of propositions, as Bennett would be required to do to describe a class of worlds in terms of the propositions satisfied by each.) In practice though, Bennett may be right, since the sorts of conditions defining the class  $K$  in previous chapters are often ones that can be expressed by propositions (though describing similarity to the actual world may still require an infinite conjunction).

A further objection to the intertranslatability of the two types of theories is based on Bennett's own theory. It seems plausible that at least some of *Laws* (or possibly *Support*) will be indicative conditionals, and thus have no truth values. Thus, although they can still be used in probabilistic entailments of various sorts, they don't define a class of worlds.

## 120: A mistake of Chisholm's

Chisholm suggested that *Support* could be any statement whatsoever, provided that it is one that the speaker had in mind while uttering  $A > C$ . Bennett argues both against the claim that *Support* could be unconstrained, and that the speaker must have it in mind.

Against the speaker being required to have it in mind, he puts forwards the conditionals 'If I had pressed the button again, the red light would have gone on again' and 'If I had taken aspirin, my headache would have gone by now.' He says these can be true even if I have no knowledge of what other factors of wiring and body chemistry contribute to these facts. Presumably, causal laws together with button-pushing only entail that the light will go on assuming that all the wires are connected in a certain way, and if I don't know anything about electronics, I might have no idea how to describe the sorts of connections required, and such a description is necessary in *Support*.

However, it seems to me that I can be aware of the fact 'Pushing this button is causally sufficient for the light to turn on', which is true, and is probably just the fact I had in mind when asserting the conditional. Of course, this fact brings in causal claims to the analysis, but Bennett doesn't seem too afraid of doing so as we shall see later on. Another possible value of *Support* that makes the logical entailment clearer might be: 'This box is wired in a way such that

whenever the button is pressed the light is on.’

The only remaining constraint is that *Support* must be true, assuming the agent knows whether or not it entails *C*. Thus, assertibility of the subjunctive is the same as justified belief in *Support*, so that Chisholm’s view may just conflate subjunctive conditionals with the *Support* statements on whose grounds they are asserted. This view has some plausibility, but also seems to make such conditionals superfluous. A similar conflation of assertibility and truth is what Bennett wanted for indicatives. As a result, Support Theories are epistemically much nicer than Worlds Theories, because one can believe and know things about propositions much more easily than about worlds. This seems to be at least part of the reason why Bennett turned to his NTV theory of indicatives.

To argue that *Support* should be constrained, Bennett uses an example where Chisholm’s *Support* contradicts *A&Laws*. One sentence says ‘If that were gold, then some gold would not be malleable’, and the other says ‘If that were gold, then some things would both be malleable and not malleable.’ In the first instance, *Support* is ‘That is not malleable, and nothing is both malleable and not malleable’, while in the second *Support* is ‘That is not malleable, and all gold is malleable.’ Chisholm’s argument for the potential truth of either of these sentences (assuming that the speaker really intends *Support* to be as indicated), is that the speaker intends a contradictory antecedent, which thus entails the consequent. But Bennett suggests that this makes it too easy for someone to (accidentally) intend a counterlegal or even counterlogical conditional. In addition, the fact that either of these types of conditionals automatically come out true in Chisholm’s theory (as long as the underlying logic isn’t paraconsistent) seems problematic, but perhaps a relevance logic would deal with these cases better. Because he uses classical logic though, Bennett argues that *Support* should at very least be consistent with *A&Laws*.

## 121: Cotenability *en banc*

A first consideration for further constraints on *Support* are that  $A > \neg \text{Support}$  shouldn’t be true. However, if this is the only condition on *Support*, then we see that if *C* is true and  $A > \neg C$  is false, then *C* is an acceptable value of support, so that  $A > C$ . The converse applies if  $\neg C$  is true, so we have Conditional Excluded Middle, which Bennett rightly thinks is problematic for counterfactuals. (Bennett does this through the claim PF\* from before.)

Thus, this constraint cannot be sufficient, because it very often allows *C* to be its own support. A stronger condition is that in particular  $A > \text{Support}$  should be true, though this leads to infinite regress (worse if *C* is the only potential value of *Support* that is not otherwise ruled out). Bennett claims that a worse problem is that it eliminates too many conditionals. In a truly non-deterministic case I might like to say that ‘If I hadn’t gone into hospital, I would have taken part in the Dunkirk evacuation.’ But since the Dunkirk evacuation must occur in order for me to take part in it, and its occurrence is not entailed by my lack of a hospital trip, it seems that ‘The Dunkirk evacuation occurred’

must be part of *Support* to make this conditional true. But then for this to be a valid part of *Support*, it must be the case that ‘(Even) If I hadn’t gone into hospital, the Dunkirk evacuation (still) would have occurred,’ but this would require some *Support*<sup>\*</sup> that could entail that the evacuation occurs. Bennett worries that nothing could causally entail that the evacuation occurs, because of the indeterminacy of the world. I don’t share Bennett’s objection, because I think that these two conditionals stand or fall together, and there may be yet another indeterministic event (say the one he mentions in Hitler’s brain in an earlier chapter) that leads to the rescue happening. However, I do agree that the infinite regress in case these conditionals are judged true is somewhat troubling. He hasn’t established that this theory makes all these conditionals false - it just establishes that there is an infinite chain of them that all have the same truth value, which we can’t establish to be ‘true’ on the basis of this argument alone.

Goodman’s attempt to deal with this regress points out that the new conditional is a ‘semifactual’ with a false antecedent and true consequent, while the first is fully counterfactual. Bennett claims that this division of conditionals is not useful, but this is far from obvious to me. Perhaps one can use a *Support* Theory of this type for full counterfactuals, and something weaker for semifactuals?

More troubling for either of these ideas, I see no real argument for choosing one of these conditions rather than merely the constraint mentioned at the end of the previous section, that *Support* must be consistent with *A&Laws*. Just because these conditions are stronger doesn’t suggest that they strengthen the constraint in the right way.

## 122: The causal contrapositive problem

This problem was first mentioned by Goodman, saying (truly) about a match *M* that ‘If *M* had been struck, it would have lit.’ One conjunct in *Support* seems in particular to be ‘*M* was dry.’ However, if ‘*M* didn’t light’ is also an acceptable conjunct in *Support*, then the theory judges it true that ‘If *M* had been struck, it wouldn’t have been dry,’ which seems obviously false. (This inference depends on modus tollens, which not all logics endorse.) Thus, our constraint on possible values of *Support* should rule in ‘*M* was dry’ but rule out ‘*M* didn’t light’.

Rescher suggests that the problem is that there is some law that allows one to derive ‘*M* lights’ from ‘*M* is struck’ and ‘*M* is dry’, which is logically equivalent to its contrapositive, but not equivalent in whatever sense is needed here. This is supposed to be the analogy with the puzzle of the ravens, in which a law seems to be confirmed by different evidence from its contrapositive. However, most current resolutions of this paradox shy away from separating the two logically equivalent laws, instead focusing on the issues of how confirmation works, and what evidence might confirm either law, and to what degree. The role played by the laws here is that of being an eligible value for *Support* or *Laws* in this theory of counterfactuals, but since the requirements to be eligible haven’t been fully

spelled out yet, it isn't clear that it might be conceivable that logical equivalents could have different eligibilities. In addition, since the role they play in finally verifying a claim is merely a logical entailment, allowing either one will allow the entailment to go through. So unless Rescher would like to propose another role for these statements to play (perhaps some sort of causal entailment?) his solution seems implausible.

Another possible interpretation of Rescher's attempt arises if one considers the law in question to be a material conditional. In such a case, these contrapositives aren't even equivalent! Thus, the worry doesn't arise. This would also destroy Worlds Theories, because there would no longer be a criterion to judge whether or not a particular world satisfied a law, because Bennett argues that such laws have no truth values. This doesn't seem to be suggested by any of the Rescher quotes, but it would reconstruct his partial solution in a relevant way.

Sellars tried to solve the problem by suggesting that counterfactual conditionals should always be interpreted with the initiation of the apparent consequence as the statement that must be entailed, so that the two conditionals would be read 'If  $M$  had been struck, it would have started to be alight' and 'If  $M$  had been struck, it would have started to be wet'. In this case it seems clear that no true values of *Support* could actually support the second conditional. However, Bennett points out that this can't explain conditionals whose consequence is a preservation, like 'If that egg hadn't been washed it would have stayed fresh longer.'

## 123: Causal direction

To sharpen this condition to solve the causal contrapositive problem while allowing for preservation conditionals, Bennett suggests that in addition to being logically compatible with *A&Laws*, any conjunct in *Support* should be such that it isn't caused by  $\neg A$ . Thus, since the not lighting of the match was caused by its not being struck, it is ruled out as a value of support, so that the second conditional comes out false. Since its dryness isn't caused by not being struck, the first conditional is acceptable.

This causation of course requires not just that  $A$  and the causal laws are compatible with *Support* (as already required), but rather that  $\neg A$  is not 'causally sufficient' for *Support* in the context of certain background facts. This is what Bennett means when he says that  $\neg A$  could be a 'NS' condition for *Support*. There exist some actually obtaining background conditions such that  $\neg A \& \text{Conditions}$  (causally) entail *Support*, but *Conditions* doesn't entail *Support*. This is remarkably similar to the Support Theory analysis of subjunctive conditionals.

This is a troubling idea because most useful analyses of causation seem to involve counterfactuals, but Bennett dismisses these by saying "although unanalysed, [causation] works pretty well for us; we have fairly firm and agreed intuitions about how it applies in particular cases." I think I could say the same

about counterfactuals if I were trying to use them to analyze causation.

One might try to use facts about temporal order to avoid mention of causation, but this would ignore problems of backward causation. In addition, the way to rule out a conjunct in *Support* that might have been caused by  $\neg A$  is to rule out all propositions about times later than  $t_A$ , but this rules out too much, making false most counterfactuals where the consequence occurs some time after the antecedent. So temporal accounts don't seem to be able to replace accounts involving causation, problematic as it might be.

## 124: The logical cleansing problem

Although we have already required that *Support* be logically consistent with *A&Laws*, and may require that it not be caused by  $\neg A$ , Bennett argues that this is not enough.

He considers a conditional suggested by Goodman ( $T_1$  represents a time at the height of the Korean war, and  $T_2$  is slightly later), saying 'If Jones had gone to Korea at  $T_1$ , he would have been a prisoner at  $T_2$ ', which is judged true if *Support* includes the fact that Jones did not in fact go to *South* Korea at  $T_1$ , while leaving out any statement about whether or not Jones went to *North* Korea at  $T_1$ . By a slight change of *Support*, we get the statement 'If Jones had gone to Korea at  $T_1$ , he would have been conscripted into the US Army at  $T_2$ '. Clearly not both should be true, and in particular, it seems that the second should be true, because if Jones had gone to Korea at  $T_1$ , we can pretty safely say that he would have gone to South Korea (unless he was a particularly odd character).

Goodman tried to eliminate both conditionals by adding to the Support Theory analysis the idea that for  $A > C$  to be true, it should be the case not just that there is *Support* such that *A&Support&Laws* entails  $C$ , but also that there not be a competing *Support'* such that *A&Support'&Laws* entails  $\neg C$ . However, this ends up dismissing both Korea conditionals, when we only wanted to eliminate the first. And Parry showed that for any non-independent conditional, such *Support* and *Support'* can always be found (consider the material conditionals  $A \supset C$  and  $A \supset \neg C$ , which are both true), so that this theory would validate only independent conditionals.

However, I think that the causal condition applied in the previous section already takes care of this problem. Although  $A \supset C$  and  $A \supset \neg C$  are both consistent with *A&Laws*, they are also entailed by  $\neg A$ , possible in a causal way. In the example at hand, it seems that Jones' not going to Korea *is* the cause of his not going to South Korea, while his not going to Korea doesn't cause his not going to North Korea, because that is determined by more general facts about the war and his character. (Perhaps more accurately, his not going to South Korea is what caused his not going to Korea, and using an element of *Support* that causes  $\neg A$  is certainly problematic in evaluating  $A > C$ .)

## 125: Parry's repairs

Parry wanted to add the requirement that no conjunct in *Support* be a logical consequence of  $\neg A$ , which would outlaw all the above examples, although it seems that we might be able to find a true and independent value of *Support* that explains why Jones didn't go to North Korea. Since we already require that such conjuncts be consistent with  $A$  and that *Support* be true (so that any conjunct is consistent with  $\neg A$ ), and any conjunct that is a logical consequence of  $A$  is redundant, this proposal ends up being a requirement that each conjunct in *Support* be logically independent of  $A$ .

Bennett suggests that this is insufficient because we can replace any banned conjunct  $S$  in *Support* with  $p$  and  $p \supset S$  where  $p$  is some irrelevant contingently true statement, and effectively smuggle  $S$  back in. We can't require that no consequence of *Support* be a consequence of  $\neg A$ , because  $\neg A \vee \text{Support}$  is always a consequence of both, even for good values of *Support*.

Parry suggests that we require that no truth-functional component of *Support* be a consequence of  $\neg A$ , since  $S$  is a truth-functional component in the above problem case. However, as Bennett points out, the particular truth-functional structure of *Support* can be disguised by writing it using odd language. Considering the logical structure underlying a statement can lead one to conclude that any of its consequences is a conjunct in it (if conjuncts in all of its logical equivalents are considered), bringing  $\neg A \vee \text{Support}$  back in. Thus Bennett says that we are still in trouble.

However, I still don't see why the causal consequence relation used above doesn't save us even without having to move to logical consequences.

## 126: Simple propositions

Bennett tries to save Parry's repairs by considering only 'simple propositions' that appear as conjuncts in *Support*, rather than material conditionals or disjunctions. He points out that propositions can't be analyzed merely as sets of possible worlds, because then either all are simple or none are. In addition, they can't be analyzed in terms of how they are constructed from concepts, because Frege failed to do so, and all others have as well, unless they make them purely linguistic entities, which brings us back to the problem of expressing *Support* entirely in neologisms. Similarly, a simple proposition can't be analyzed either as one that can't be expressed with a logical connective (because all can) or as one that can be expressed without logical connectives (because all can, again through neologisms).

The one suggestion Bennett sees as plausible is by Pollock, who suggests an epistemic analysis, according to which a proposition is simple if an agent can come to know it in a way that is neither inductive nor deductive. (Direct observation is taken to be neither.) However, Bennett points out that this suggestion must be relativized only to actual human epistemic capacities, and therefore can't be fully universal. The idea of restricting conjuncts in *Support*

to ‘simple propositions’ is so that none can be a conjunction of two others. But since the idea of being simple has been expressed in epistemic terms, it seems to me that the following three propositions can be taken to be simple, even though the first has sometimes been analyzed as the conjunction of the other two: ‘ $x$  is human’, ‘ $x$  is featherless’ and ‘ $x$  is a biped’. So maybe resorting to simples in this epistemic sense doesn’t work at all.

More problematically for Bennett, he wanted a metaphysical theory of subjunctive conditionals, rather than an epistemic one. However, it seems to me that one main reason to prefer Support Theories over World Theories is that they are far more epistemically grounded. It is plausible to think that one can come to know (or at least have beliefs) about true statements, causal laws, and logical entailments, than to think one can have knowledge about possible non-actual worlds (if such exist). Bennett takes a relatively epistemic view towards indicative conditionals, and therefore it shouldn’t trouble him too much to take one towards subjunctives as well. By analyzing both in terms of sentences rather than worlds, one would get an analysis that is more ‘Y-shaped’ than Bennett’s actual analysis. Making the entailment in the subjunctive case probabilistic would unify them even further. The main truly problematic point is that the analysis makes use of causation in the first criterion for being in *Support*, which at least seems to be a notion that depends on counterfactuals.

If we start with a theory like Chisholm’s, we can take subjunctive conditionals to be just as subjective as Bennett takes indicatives to be. We may even be able to express the Ramsey condition that Bennett uses to define indicative conditionals in terms of some sort of probabilistic entailment of the consequence from a conjunction of the antecedent and some facts known to the speaker beforehand. It would be nice to strengthen Chisholm’s theory slightly so that the conditionals about matches and Korea end up having their intuitive values not just by coincidence about which values of *Support* people actually tend to intend. This strengthening can be achieved using the theories that Bennett endorses here for less subjective accounts of indicatives. This account of the relation between subjunctives and indicatives would be far more Y-shaped than the one Bennett actually gives.