Ellen Bales
November 9, 2004

15. Truth at the Actual World

89. Similarity in Matters of Particular Fact

On Bennett's account there are at least three things we know we must attend to in evaluating the forward subjunctive conditional in a possible worlds model. To wit, we must go to A-worlds that:

1) sufficiently resemble the actual world until shortly before $T_{A.}$
2) fork in an admissible way—that is, through a different outcome of an indeterminacy, a small miracle, or an exploding difference—in a direction that makes A true at them
3) conform to the laws of α from the fork onwards

Bennett asks if this is the whole banana or do states of α after T_A come into the analysis? Here, he says, we encounter a problem we must look into further.

To try to solve it, he thinks, it will first help to simplify our thoughts about A>C by pretending that there can only be one admissible fork when we evaluate the conditional and that its date and nature are the same at all candidates for the title of "closest A-world." So up to T_F, these worlds are exactly the same and all we have to consider is how they differ in times thereafter and how such differences affect closeness. Bennett asks if we have any basis for preferring some of these worlds over others.

Taking as givens that these worlds are exactly alike at T_F and are law-abiding from then on, they can only later become unalike through indeterminacy. We have a problem then if α is not strictly deterministic because in that case, the worlds in question could develop differently after T_F. Then we would want to know how these post-T_F differences bear on the closeness of worlds issue.

Particular Fact:

Bennett gives us what he calls a "tempting answer," a rhetorical clue that it isn't going to work or be satisfying (at least on his account). This is the "particular fact" theory:

(PF) If (1) w_1 exactly resembles w_2 up to T_F and (2) both conform to the laws of α thereafter, and (3) they first become unalike in respect of one particular matter of post-T_F fact that obtains at α and w_1 but not at w_2, then w_1 is closer to α than w_2 is.

Lewis discussed this issue in his 1979 paper, but according to Bennett, Lewis dismissed particular facts post-T_A as being "of little or no importance" in determining closeness. Bennett seems to imply that Lewis tosses away the argument by saying that when we try to decide between "little" and "none" in these cases, "Different facts come out differently."

Bennett wants to pursue the particulars of this phenomenon because it will be a starting point for his own theory of the truth conditions of subjunctive conditionals, as we will see in subsequent sections. He makes two arguments against using PF as a further determinant of closeness following Lewis's main theory (which determines closeness on the basis of exact likeness up to T_F, smooth parting of worlds at T_F, and strict legality thereafter). One is logical and the other is intuitive.

The logical argument against PF:

Following Charles Cross's "plainly valid" argument in support of CEM, Bennett gives us the argument against PF:

PF has the consequence that if C is true at α, and is not ruled out by A, then it is ruled in by A. In brief, PF implies PF* (below) in this way: If we start with C being true at α, then add w_1 at which C and w_2 at which ¬C, at T_F (time of the antecedent) w_1 will *always* be closer to actual than w_2, so that A>C. It is true at all closest A-worlds. Therefore,

PF*: C & ¬(A>¬C) entails A>C.
What is interesting about Bennett's account is that he completely elides the (not obvious) steps that get us from PF to PF*, leaving us to reconstruct them. We might interpret this elision as Bennett's way of ducking a couple of difficult matters, which are that (1) getting from PF to PF* is not an uncontested issue AND (2) that by this sleight of hand, he can build his proof on PF*, claiming universal validity for it (rather than for PF). So is this an argument against PF or, in the final analysis, against PF*?

In any event, Bennett is clear enough that he is building the proof on PF* as follows:

PF* entails Conditional Excluded Middle (CEM) by this proof, starting with PF*:

(1) C & ¬(A>¬C) entails A>C.

This is a schema with universal validity, so we can go ahead and replace ¬C for C in two:

(2) ¬C & ¬(A>¬¬C) entails A>¬C.

This has the form 'P&Q entails R' which is classically logically equivalent to 'P&¬R entails ¬Q' By the substitutivity of classical logical equivalents we get:

(3) ¬C & ¬(A>¬C) entails ¬¬(A>¬¬C).

Now ¬¬P is classically logically equivalent to P, and by substitutivity of classical logical equivalents, we get:

(4) ¬C & ¬(A>¬C) entails A>C.

Since C & ¬(A>¬C) entails A>C, and ¬C & ¬(A>¬C) entails A>C, therefore by a classically valid form of inference, we get:

¬(A>¬C) entails A>C,

which clearly entails CEM, which is (A>C) ∨ (A>¬C). So PF* entails CEM and if PF is stronger than PF* then it too entails CEM.

But in the end, the logical argument against PF is only as strong as the independent case against CEM, a matter on which there is legitimate disagreement, a lesson from preceding sections. Bennett of course rejects CEM, a point that will become important later.

So we turn to…
The intuitive argument against PF:

The problem with PF, according to Bennett, is that it gives wrong truth values in certain cases. An example is forthcoming:

We have an objectively random coin-tossing machine that is "deeply" so because the mechanism is as follows: you press a button that activates a device that fires a single photon towards two slits. If the photon goes through one slit, the coin is tossed so that it comes down heads. If it goes through the other, then the coin is tossed so that it comes down tails.

So a guy named George W. Bush activates this mechanism at T_A and the coin comes down heads. Now we want to analyze the statement, "If John Kerry had pressed the button at T_A, the coin would have come down heads."

On Lewis's theory, this conditional is false; on Stalnaker's theory, it is indeterminate. Neither theory calls it true, but PF does. Bennett says that this reveals a problem with PF because the majority of people will have the intuition that the Bush world and the Kerry world are worlds at which different causal chains occur. Bennett suggests that PF can be rescued. This will be the topic of the next section and an important step in Bennett's own case, but also where he gets into considerable trouble.

90. Solving the Particular-Fact Problem

What Bennett concerns himself with here is unknotting the problem of what to do with PF, which raises the question of what we need with it anyhow. Bennett says if we're going to stick with the Lewis possible worlds model, we aren't going to be able to get where we are going with JUST Lewis's theory of the particular-fact subjunctive conditional. So it may be useful to see what we can get out of some variation on PF.

Why isn't Lewis's account sufficient? Well, Bennett says, it give us wrong truth values for certain kinds of particular-fact conditionals, including the following:

"Imagine a completely undetermined coin-tossing device. Your friend offers you good odds that it will not come down heads next time; you decline the bet, he activates the device, and the coin comes down heads. He then says: 'You see; if you had bet heads, you would have won.'"

Bennett says that this is intuitively true, but that Lewis's theory makes it false because the device is random and at some of the worlds the coin will come down tails, which means that at these worlds, you will lose your bet.

Bennett then says, well, if you want to oppose me and say, "Hey, that conditional is false," you'll get into hot water fast. Here he uses the example of Gone-Hospital>Rescue:

A not fully causally determined event in Hitler's brain at T_2 *led to his strange decision not to wipe out the British Expeditionary Force at Dunkirk. Much of the army was rescued from the beach at* T_3*, by small private craft from England. Mr. Miniver had such a boat, but couldn't take part in the rescue because at* T_1 *he had gone into the hospital for elective surgery. He says later, "If I hadn't gone into the hospital, I would have taken part in the Dunkirk rescue."*

The problem with this kind of example in a larger sense has to do with our intuitions. For Bennett, these are important to heed. The reactions we have are ultimately guided by our sense of what causal chain of events is being followed and whether it is the "same" causal chain at the closest worlds.

In a typical Bennett move, he then says that if our thoughts are going in that direction we should be able to put that into theory.

The theory Bennett offers is…
Causal Chain PF:

If (1) w_1 *exactly resembles* w_2 *up to* T_F*, and (2) both conform to the laws of* α *thereafter, and (3) they first become unalike in respect of one particular matter of post-*T_F *fact that obtains at* α *and* w_1*—**through the same causal chain at both**—but not at* w_2*, then* w_1 *is closer to* α *than* w_2 *is.*

Where does this get us exactly? For one thing, it allows us to bring in the concept of an indeterministic causal chain, a critical element for Bennett in analyzing how particular post-T_A facts relate to closeness. Bennett clarifies his definition of a causal chain as a sequence of token events, rather than as a relation between facts.

Bennett's intuition is that the pushing of the button in the coin toss example is part of the salient causal chain, but that the betting on the coin toss is independent of the salient causal chain. The question then becomes, what does Bennett mean here by independence? In an indeterministic world *no* token event is causally sufficient or necessary for any *other* token event. So this is not a distinguishing factor. Yet Bennett wants some events to be "causally relevant" to the coin toss outcome and others to be "causally irrelevant."

Where does this take him? Here we'd have to look at probabilistic causation (indeterministic causation). This is a probability raising idea:

$\Pr(E/C) > \Pr(E/\neg C)$

Where C_1 = button pushing
and C_2 = betting:

$\Pr(E/C_1) > \Pr(E/\neg C_1)$ but…
$\Pr(E/C_2) = \Pr(E/\neg C_2)$

But Bennett has run into a problem, which is that it is generally accepted that purely probabilistic accounts don't work because of preemption. This is an example of preemption that was discussed in the seminar:

Suppose that Pam and Bob each aim a brick at a window. Pam throws and shatters the window, while Bob hold his throw on seeing Pam in action (i.e., because she aims). It seems that Pam's throw caused the window to shatter—her brick is what crashes through the glass. But it doesn't need to be the case that Pam's throw raised the probability of the shattering—if Bob is a more reliable vandal, then Pam's throw might even have made the shattering less likely.

To solve this, most people bring in counterfactuals, but here we are trying to *determine* the truth conditions for counterfactuals so such a solution would be circular and therefore no solution at all.

What Bennett seems to really be talking about it not causal relevance, but rather explanatory relevance. In a deep sense, what he is really getting at is not causality, but correlation. A quote from Jeffrey on the toss of a genuinely indeterministic coin that is strongly biased toward heads elucidates this point about explanatory relevance:

…The fact that the coin has been tossed is the only factor relevant to either outcome (heads or tails) and that factor is common to both outcomes once we have cited the toss…we left nothing out that influences the outcome.

91. Non-Interference Subjunctive Conditionals

In this important section, Bennett codifies his principles. Bennett reminds us that the general idea behind Causal Chain PF is this:

If at the actual world, A is not the case, C is the case, and these two facts are causally unrelated to one another at α, then C also obtains at the closest A-worlds.

But there are other kinds of subjunctive conditionals, in fact more prominent as a type, where A's being true helps to explain somehow C's being true at the world in question.

There are two types in this category, the making-true sort and the non-interference sort. Bennett takes up the non-interference subjunctive conditionals because they are the hairier and more obscure of the two kinds, and he says that these are "appropriate and natural as support for other conditionals."

Pollock's description, however, gives us a sense that in fact it is not quite correct to speak of "non-interference" as a *kind* of conditional. The label instead directs us to the source of its truth. A non-interference conditional is believable because A would *allow* C to obtain. A making-true conditional is believable because A would *make* C obtain. Here is another, and perhaps clearer way, of talking about the distinction Bennett wants to make with the coin toss example. In the button pushing instance, we have a making-true conditional and in the betting instance, we have a non-interference conditional.

Here Bennett illustrates with the Cut > Safe example ("If the channels were cut, the village would be safe"), a conditional spoken first by an agronomist and then by an engineer that gets its truth value from two completely different sets of presuppositions held by the two different speakers, despite being—on the surface of things—the exact same conditional statement.

Still, the conditional they both assert is true, in spite of the fact that from the point of view of the agronomist, it is true in the non-interference fashion and from the point of view of the engineer, it is true in the making-true manner. Bennett claims that it follows that one man says something true that he derives from an error, but that this is not a problem because we often believe truth on the basis of errors. But this is not an indeterministic case.

Bennett concludes by saying that Davis uses the word "then" in a construction to excluded a conditional from being of the non-interference sort, and Lycan explains why this "then" can work in this way, but Bennett says that he finds this device a "thin and narrow basis for the edifice of theory that Davis and Lycan have erected upon it."

92. Does A & C entail A > C?

Bennett wonders here if A>C is automatically true if A and C are both true. The project is to problematize and unpack a matter that all the main analyses of subjunctive conditionals have as a theorem and along the way, Bennett will adopt a couple of principles that will ultimately allow him to say something more general about subjunctives and how they should be evaluated at closest A-worlds.

So what are the issues here? The theorem certainly works, Bennett says, if the actual world is deterministic. We can fall back on the weaker **Sufficiency Principle**:

If A and C are true, and at T_A *sufficient conditions exist for the obtaining of C at* T_C*, then A>C is true.*

Now Bennett admits that many conditionals that are true by this principle 'feel false' when asserted at the actual world, but you can "provide a suitable context" to dissolve your feelings of falseness, as we did with the non-interference conditionals whose antecedents were false. Since Bennett has brought contextualization into the picture, this might be a good moment to consider what we are doing here. Are we talking about metaphysics? Causation? Explanation?

Of course, the real challenge to the condition occurs in the presence of indeterminism. Bennett asks if the conditional, "If you were to toss the coin, it would land heads," is automatically true if the addressee does toss the coin and it does fall heads. He leans toward no because this does not give you the right kind of causal connection in an indeterministic world, although he suggests that an equally weak argument could be made for yes. So how does he argue for no? He appeals to the general line of thought underlying Causal Chain PF, and bases his idea on the **Irrelevance Principle**:

When A and C are both true, and the causation of C is indeterministic, A>C is true just in case A is irrelevant to the causal chain that leads to C's obtaining.

Where does this get us? This is an articulation of causal irrelevance, but what we really need is explanatory irrelevance.

There are two consequences of this, Bennett says and they can be illustrated with recourse to our deeply random coin tossing machine:

Consequence 1) I say, "If you were to activate that device, the coin would come up heads." You do, indeed, activate it, and the coin comes up heads. I have said something false, despite the truth of both the antecedent and the consequent.

Consequence 2) I say, "If you were to bet on heads on the next toss of that device, you would win." You bet on heads, the coin comes down heads. Here I have spoken truly, even though I was just guessing and not being rational in any discernable way.

Ultimately, when we accept the Irrelevance Principle, we deny that A&C entails A>C. Something shocking happens here, which is that we must abandon our received idea that the falseness of A>C and the falseness of C at some closest A-world are linked. We are then in a situation where it can occur that A and C are both true at α, and C is false at some closest A-world. Gasp! Now we have to accept that some other worlds are as close to α as it is to itself, which seems peculiar when closeness is thought of in terms of similarity.

In 1973, Lewis made a stab at dealing with this conundrum. "Perhaps," he wrote, "our discriminations of similarity are rather coarse and some worlds different from α are enough like α so that such small differences as there are fail to register."

The problem with this form of consolation is that the cases Bennett brings up are not concerned with such small differences. But Bennett says that in the intervening thirty years, we've made progress and now we pretty well agree that closeness is *not* like simple untailored similarity.

This means that our needs are simpler and all we need to get out of the unwanted entailment is for some respects of (dis)similarity to be irrelevant to closeness.

How would this work? What would define what respects were relevant or irrelevant? We would certainly need a theory of causal relevance which, as we have seen, Bennett doesn't have (and no one really has).

As a sidebar, an important application of this in epistemology is Nozick's truth-tracking account of knowledge:

S knows that p just in case (roughly):
(1) S believes p
(2) p is true
(3) If p were true, S would believe p. (p > S believes p)
(4) If p were false, S would believe ¬p. (¬p > S believes ¬p)

But if A&C entails A>C, the (3) is redundant. But it doesn't SEEM redundant to the notion of truth-tracking. This forces Nozick to abandon Lewis' theory of ">", which has this consequence, and adopt his own theory of what can be described as a neighborhood of closest worlds.

93. Home Thoughts from Abroad

Now Bennett wants to say something more global about evaluating subjunctive conditionals at the actual world in light of how they should be evaluated at certain other worlds. On his account, this result will follow from his two stands (the Sufficiency Principle and the Irrelevance Principle) on the truth conditions for A>C when A and C are both true and what those imply about how A>C should be evaluated at closest ¬A-worlds.

Combining these concerns, Bennett comes up with two consequences:

1) Following the Sufficiency Principle and reading back into the situation from the viewpoint of the closest ¬A-world, we get the following:

When A and C are both true and C is deterministically caused, A>C is true at α just in case it is true at the closest ¬A-world.

2) When we take the Irrelevance Principle and step across to the closest ¬A-world and look back we get:

When A and C are both true and C is not deterministically caused, A>C is true at α just in case it is true at the closest ¬A-world.

When we combine these two conclusions, we conveniently get what Bennett dubs "**Home from Abroad**":

Whether or not C is deterministically caused, if A and C are both true then A>C is true (at α) just in case it is true at the closest ¬A-worlds.

Bennett then confesses that part of his reason for adopting the Irrelevance Principle is that is gives us this nice result (and it doesn't violently conflict with any of our intuitions, he says).

94. Stand or Fall

In this section, Bennett will begin to build the argument that Stand or Fall, the thesis that certain indicative conditionals stand or fall with the corresponding subjunctives is "grandly and globally false." He finds it utterly uncompelling as a basis for arguing that Does-will indicative conditionals belong in the same category with subjunctives. This shattering indictment will be further explained in section 135, but in the meantime, Bennett thinks it will be worthwhile to note Stand or Fall's "more local" failures.

There are three that Bennett asks us to consider.

1) The first we can call the **Stand-Off Caution**. Bennett admits that it doesn't cut very deeply into the Stand or Fall thesis, mainly because there aren't many stand-offs between pairs of Does-will indicatives, but nonetheless, they are possible, so Bennett feels it is justified in spending a page and a half discussing the matter.

The basic idea here is this:

We cannot infer from Stand or Fall that subjunctive conditionals can enter into Gibbardian stand-offs such as the ones we saw in section 34 with Does-will indicatives. *In other words, we are persuaded in these Gibbardian stand-offs that one person be legitimately entitled to accept A→C and another to legitimately accept A→¬C. What won't work, according to Bennett, is for one person to legitimately believe A>C and another to accept A>¬C.*

Since it may not be intuitively obvious why this is so (and in fact there is trouble ahead), Bennett takes us back to the Top Gate example.

Now we find ourselves back at the gates and the dam and this time both levers are down. This being the case, Top Gate cannot open, but later there may be truths in the form of "If Top Gate had opened…" Suppose now that we find several A-worlds that are about equally close to actuality—each diverges inconspicuously from exact likeness to α shortly before T_A, the time to which A pertains. Now if C obtains at all those worlds, the conditional is true. The issue is that there might be several suitably close worlds with C obtaining at only some of them. In this case, neither A>C nor A>¬C is true. In contrast to the Gibbardian stand-offs in indicative conditionals, Bennett says that it is never the case that both can be true or fully acceptable.

While it is the case that both cannot be true, Bennett overextends his case by saying that they both cannot be fully acceptable. In fact, evidence very often supports two incompatible theories (this happens all the time in science) and the evidence can be equally strong for both. On the other hand, it is possible that Bennett is assuming different standards for acceptability in indicatives vs. subjunctives. He does not, however, make this clear.

Bennett proceed to say that in order for Stand or Fall to say that because Gibbardian stand-offs occur with indicatives they can occur with subjunctives, the thesis would have to isolate the following items as standing or falling together:

1) There being facts that would make it proper for someone who knew them to accept A→C at T_1*, and 2) the truth of A>C at* T_2.

This is incredible, according to Bennett. In the Top Gate example, Wesla may later *believe* that if Top Gate had been open the water would have flowed to the west, and Esther may *believe* that if Top Gate had been open the water would have flowed to the east, but if so then at least one of them is in error.

Again, this just seems wrong, unless Bennett has left something unstated in his account.

2) Bennett's second warning can be called the **Multiple Utensil Problematic** and was suggested by Mark Lance. Multiple Utensil occurs when there are two recent, inconspicuous forks from actuality to A, the slightly later one leading to C while the earlier does not, which means that we do not consider A>C true. The problem for Stand or Fall is that in the moment between the two forks there is a sound basis for accepting A→C.

It works like this:

Sheep are checked first for weight and then for health; if they pass both they go to the slaughter house. However if they fail for weight they go to the meadow, while if they fail for health, they go into the barn. Think now upon a sheep who passes on both counts. We would not say that if it hadn't been picked for slaughter it would have gone to the barn, yet during the minute between the two checks there is a sound basis for saying that if it isn't picked for slaughter it will go to the barn, and no basis for saying that if it isn't it won't.

3) Warning number three about Stand or Fall is the **Laplace's Archangel Issue**. There are innumerable examples of A>C being true even though at no time was A→C even remotely plausible for any reasonable person. For example, no one can reasonably say, "If you get on that plane, you will be killed," but after the plane has crashed, you can say, "If you had gotten on that plane, you would have been killed." In these cases, the subjunctive conditional are true non-interference conditionals, but because A is not connected with C, there may be nothing to support the corresponding indicative conditional.

95. Independent Subjunctives

Bennett writes this section almost as a footnote to the chapter. Although he notes that his main treatment of subjunctive conditionals is complete, he says we need to decide what to do with independent subjunctives. These really have to be considered separately because the truth of an independent subjunctive does not depend on any particular facts about how things stand at the actual world.

In brief, if A>C is logically true, its truth comes from C's *obtaining* at all the A-worlds; if causally true, on C's obtaining at all the *causally possible* A-worlds; if morally true, on C's being judged to *apply* at all the A-worlds. In none of these three cases does the truth of the conditional involve its being variably strict.

We can therefore conclude that independent subjunctives do not do interesting things like run counter to logical principles like Antecedent Strengthening and Transitivity and do not give us any interesting problems about the closeness of worlds. So what kind of analysis will help us here? Does the closest-A-world type of analysis not apply? If so, doesn't that open the door to disunity in our account of subjunctive conditionals.

If we theorize that independent subjunctives have a special meaning, then we will have to choose between three options:
A>C means that A entails C if (1) the speaker thinks that A entails C, or (2) A does entail C, or (3) both of the above.

Bennett thinks that this is a waste of time to commit ourselves to either the metaphysical or epistemic option (or both) because we don't get into any special difficulties if we analyze independent subjunctives the same way we analyze subjunctive conditionals generally. In the case of logically and morally independent subjunctives, every A-world will be a C-world, so the question of closeness just doesn't have to be sorted out. With causally independent subjunctives, we look to all worlds that obey the same causal laws as actual, which again eliminates the issue of determining closeness.