Nicolas Gomes
11/23/04

Chapter 18:
Backwards Subjunctive Conditionals

§108. Banning Backward Conditionals

Downing Denied:

P. B. Downing chose to ban backward conditionals in response to his scare story (§78), rephrased as follows:

> Generous Jim is mad at timid Tom. *Thus, if Tom had asked Jim for a favor, Jim would have refused.* However, Tom isn't so cavalier or oblivious to ask a favor from someone who is mad at him. **So if Tom had asked Jim for a favor, Jim wouldn't have been mad at Tom.** *Thus if Tom had asked Jim for a favor, Jim would not have refused.*

The conflicting forwards conditionals (*italicized*) are dealt with by Downing though the rejection of the backwards subjunctive conditional (**bold**) used to derive the latter forward subjunctive conditional. It is unclear, then, what Downing thinks of the latter forward conditional.

Given Bennett's rejection of the scare story and support of backward conditionals' common and natural usage, he demands proper treatment of them, even if one hopes to ultimately explain them away as Davis does…

Davis' Deficiency:

Bennett considers Davis attempt to reduce backwards conditionals (BC) into terms of forward conditionals (FC).

The distinction is made between:

a) If Kerry had won the election, he would have won Iowa.
b) If Kerry had won the election, he would *have to* have won Iowa.

Davis sees (b) as the natural and correct way of what we mean by saying (a). He contends that the consequent of conditional (b) is not a reference to state of affairs before the election, but rather a post-election modalized-past-tense state of affairs. Namely, the closest Kerry-election worlds merely share the post-election historical fact that Kerry won Iowa.

Bennett first argues Davis' account is deficient in explaining the "modal element in the consequent of a backwards conditional." Bennett argues that "have to" expresses our inability to understand any Kerry-elected world other than by won-Iowa, and in this way calls it the "best explanation" element.

Bennett then argues that to say of the (closest) post-Kerry-elected worlds "Kerry won Iowa is a modal fact" is just to say of the closest Kerry-elected worlds "At an earlier time Kerry did actually win Iowa." Thus, BCs still say something about the past of (closest) Kerry-elected worlds.

This line of reasoning can turn FCs back into BCs, thus corrupting an already confused type of conditional. (I'm not sure if Bennett means this line of reasoning creates even more BCs than would otherwise come about.) This argument is said to rely on "the principle that if A>C and C entails P, then A>P."

More generally, any world at which:
It is the case at T_2 that: necessarily that P at T_1.
Is a world which:
It is the case at T_1 that: P.

This is shown symbolically as follows:
$A>\Box P$
$\Box P>P$

$\therefore A>P$

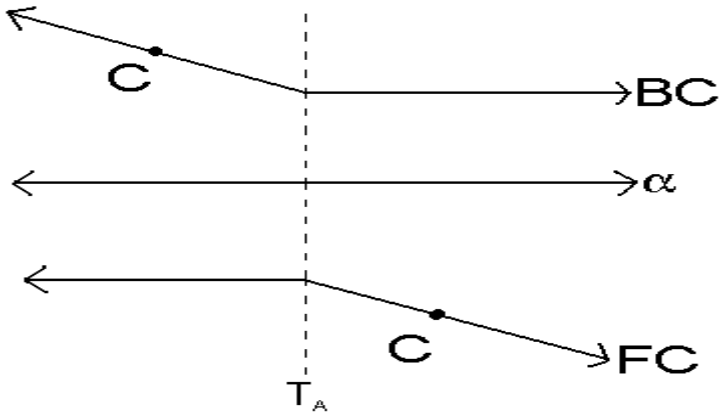
Thus, the "have to have been P" conditionals Davis accepts entail conditionals which say "P at some earlier time". In other words, (a) type conditionals naturally and unavoidably brings about (b) type conditionals.

§109. Backward Conditionals and the Future

Jackson Mirroring BC from FC:

Jackson allows BCs, albeit cautiously. Reminiscent of his FC analysis, Jackson says the following:

> If T_A is later than T_C, then A>C is true just in case C is true at (A-)worlds that are exactly like α after T_A and obedient to the laws of α before T_A.



Branden describes them as duals (as opposed to opposites) to one another, an apt title given the following table:

$A >_b C$	$A >_f C$
• **T_C comes before T_A.**	• **T_A comes before T_C.**
• **C is true at A-Worlds similar to α after T_A.**	• **C is true at A-Worlds similar to α before T_A.**
• **Before T_A, α is legal.**	• **After T_A, α is legal.**

Note: This does **not** mean that $X >_b Y \equiv Y >_f X$ – a property that, while not altogether unreasonable, does not seem very plausible after prolonged consideration.

Note: One need not worry about how to actually verify or go about knowing the truth of a conditional (or other epistemic issues); this is discussion about the metaphysics of conditionals' truth conditions.

And a Counter-Example…

While graciously bypassing the returning ills of bumps, Bennett attempts a simple application of Jackson's theory on a BC such as:

> If Adlai Stevenson had been the undisputed president of the USA in February 1953, he would (have to) have been elected in November 1952.

This conditional satisfies each criterion of Jackson's true BCs, which is not a problem, considering the conditional sounds perfectly reasonable.

However,

> If Adlai Stevenson had been the undisputed president of the USA in February 1953, then in November 1952 he would (have to) have celebrated a short-lived success.

…too satisfies Jackson's criteria, due to the suspicious requirement of having A-worlds look similar to α after T_A (a time in which Eisenhower is president). This is unfortunate, considering how untrue this conditional sounds. Even if you pressed the issue, Bennett would respond by saying the consequent has lost its explanatory power over the antecedent - a basic element of BCs.

This objection not only pretty clearly destroys hopes of BC-analysis mirroring that of FCs, but also more generally, destroys hopes of BC-analysis being based on matching other world's futures with α's future.

§110. A Better Theory of Backward Conditionals

Lewis Wins:

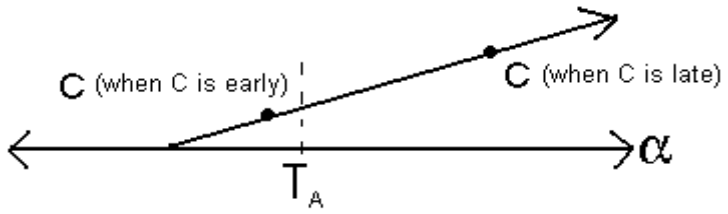
Bennett sees his "better theory" as "something everyone will accept after reflection" as one which mimes Lewisian analysis of forward conditionals, while dodging objections with notions of legality. Bennett claims BC analysis is similar to Lewis's (annotated) claim:

> A>C ≡ C obtains at the A-worlds that most resemble α at T_A, out of all the A-worlds that become unlike α for the first time at a late modest fork and are legal from then onward.

Bennett then adds to this old interpretation of A>C in order to accommodate BCs:

> A>C ≡ C obtains post-T_A at the A-worlds that most resemble α at T_A, out of all the A-worlds that become unlike α for the first time at a late modest fork and are legal from then onward. *Or*, if C obtains pre-T_A, in which case C belongs to every legal ramp to A which starts with a latest fork from actuality.

The merit of including BC's into Lewis' analysis of FC is that of unification and very similar ramp considerations, as seen by capturing both theories in one graph:



Bennett, however, is worried about the lopsided requirements of lateness for FCs and BCs. The fact is, there many more distant futures that could be legally brought about from a tiny difference now, than distinct distant pasts that legally lead to tiny differences right now.

In other words, it is not the case that many new and exciting worlds in the distant past are found when a small change is made now. However, a small change made now may easily blow up (legally) into new and exciting worlds where many strange C's can be satisfied.

Does this lopsided requirement oppose intuition? Do we expect counterfactuals with minor antecedent changes to satisfy more interesting consequents set in the distant future?

Is > reflexive?

All this talk about symmetry leads us to consider whether FCs and BCs are symmetrical under this theory, i.e. whether A>C ≡ C>A. Bennett assures us not, as we should expect by the following.

Considering A>earlier: If I made it to my 8 am class, then I had to have woken up at least by 7 am.

However the converse, Earlier>A, is false: If I had woken up at least by 7 am, then I made it to my 8 am class.

Bennett explains this phenomenon by saying that "a necessary condition for the slide need not be sufficient for it." Ergo, while the analysis for backwards and forward conditionals uses the consideration of the same worlds, true BCs can't be flipped around to make true FCs (or vice-versa).

§111. Looking for a More Generous Theory

Doing without Forks?

Bennett applauds the Latest Fork theory of BCs but investigates "less restrictive" analyses. Bennett claims that a theory of backwards conditionals may be adapted to any particular view on haves about what are acceptable forks (e.g., asserting the indeterminacy of human behavior).

It seems one might even analyze these conditions without forks at all (as was done with Jackson's analysis and Bennett's Simple Theory). This would (supposedly) result in a theory of BCs as follows:

A>Earlier ≡ Earlier obtains at all the legal A-worlds that most resemble α at T_A.

Doesn't Work:

This claim is quickly discounted by Bennett since the explicit antecedents of many BCs are mixed undesirably with α. For example, someone saying "If dinosaurs had been roaming the Earth today, the signs of global warming would still have existed in the early twentieth century" wouldn't expect the listener to consider cases in which humans also roam the Earth.

Note: This objection is not unique to BCs. This can be seen, for instance, by changing "have existed in the early twentieth century" into "exist well into the twenty-first century" from the previous example. The Simple theory again has us considering worlds in which humans coexist with dinosaurs.

Removing the condition of A-worlds' resemblance α at T_A will rescue this analysis (how could it, since we are only adding to the number of worlds to consider and the objection above was one about considering too many worlds?). Bennett says such a reduction would turn BCs into a statement about mere nomological possibility, which doesn't seem to be enough for our analysis.

Capturing the proper relation between α and earlier - that the conditions of A relate to α in such a way as to *explain* earlier - seems to be a achievement Bennett only ascribes to the Latest Fork theory.

§112. Counterparts

Keeping Track of Objects through worlds:

Lewis's extreme realism about possible worlds may create problems for counterfactuals which reference to particular objects. However, even non-extreme realists face difficulties of how to trace though alternate worlds the identity of an object and the properties it should satisfy when referred to in a counterfactual.

Note: This is a concern for FCs as well as BCs. In fact, this issue is so important that without its resolution no theory of counterfactuals is complete.

Here is an example:

> If I had four fingers on my right hand, my handprint would look different than it does now.

Here there needs to be a world in which "I" have four fingers on my hand. Given *indiscernability of indenticals* (widely accepted), which says that for all a's and b's if a=b, then a and b have all the same properties (like having four fingers), there is difficulty finding an identical "I" which has a different property than *me*. This is the fundamental problem of *de re* modal claims (modal claims about individuals).

To resolve this dilemma of speaking of corresponding objects in different worlds, Lewis introduced the notion of counterparts. Counterparts are described as objects which play near or exactly the same role in their respective worlds, are sufficiently similar to the actual objects, etc. So now, Lewis holds, counterfactuals quantify over all counterparts (like ones who have four fingers).

The alternative to making claims about counterparts (who are admittedly different than me, thus seemingly misplaced in a counterfactual about *me*), is to insist that the guy with four fingers really is me but somehow the property of having four fingers acts as the counterpart, i.e. as some *other* property which is sufficiently like my having four fingers in the actual world.

Either way, one must either say that it is some other *guy* we are talking about in the other worlds, or it is some other *property* we are talking about in the other worlds.

Determining the criteria for relating particular objects is a delicate and crucial element of evaluating counterfactuals, as seen by the conditional, "If I had put a little more arch into that shot, this basketball would have gone in the hoop." The truth value of this conditional might be very much influenced by which changes to the basketball are acceptable in alternate worlds (since alternate worlds allowing more arch to sink the ball might require alteration to the ball's qualities).

Linking Objects by Past:

Bennett says that we don't have to worry about the similarity of an alternate world's object to the corresponding object in α after time T. As he says, "We can simply read off the identities of the particulars at (an alternate world) from those at α (before they diverge)." Thus a common past can act as a link between objects of distinct worlds.

This issue, however, does not seem to be resolved in the event that the latest ramp extends earlier than the object's coming-to-be. In this case, Bennett offers an example that allows Bennett to identify objects based on having similar causes at similar times and having the "same shape, size, and pattern":

If the potter used more aluminum in the clay body of his ceramic bowl, it wouldn't have cracked when he shipped it.

Establishing identities of counterparts such as these, where the object in question can't at some point before the fork begins be linked to the actual bowl, must be established by reference to the bowl's cause, the potter, and the intrinsic, important qualities of that bowl. At some point, the A-world and α will be similar enough to relate objects by similar properties. This theory of counterparts Bennett holds as a virtue of his Latest Fork Theory of BCs. As Branden says, "It provides a natural counterpart relation without appealing explicitly to similarity comparisons of objects across worlds."

I am not satisfied with this response because I don't see all disputes over counterparts being so clear-cut.

For example, the truth of "If this basketball had been made bouncier, then it would have gone in the hoop" must be analyzed by determining which is the "*this* basketball" in the alternate world with the latest fork from α. However, one could identify the balls based on similar *production* time and location or by which one was *bought* by the same player at the same time. Assume the latest fork gives us a factory which uses the bouncy rubber at the cost of slower production, thus entailing that the similarly *bought* basketball is different from the similarly *produced* basketball and potentially giving two different results for the same conditional.

Note: It is suggested that this problem may be resolved by considering the counterparts in terms of context and a full fleshing-out of the relevant attributes.

Backwards Causation:

Bennett takes a quick moment to inform the reader of his opinions of backwards causation: he believes it impossible, and has no explanation of it if it is possible. Moreover, his theory of BCs is radically incomplete if backwards causation is allowed.

I refer the interested reader to the Stanford Encyclopedia of Philosophy's enlightening entry on this topic found at: http://plato.stanford.edu/entries/causation-backwards/

§113. Dispensing with Histories

Counterfactuals without Worlds:

There seems to be another class of (backwards) counterfactuals which don't seem to involve consideration of possible worlds. Here are a few:

> If this wasn't a conditional, it wouldn't have sounded like one.
> If the problem was in the carburetor, we'd hear it.
> If this is a hamburger, where's the beef?
> If that was joke, it would have had a punch line.
> He couldn't have fought in Vietnam. If he was alive in 1960 he would now be over 40 years old.

Bennett asserts that "nobody would mean or understand these as involving possible histories — branching-off from an α-like world to events" in which the antecedent is satisfied. Rather, it seems like the speaker maintains the truth of the conditional based upon a fact that he or she believes will hold in a wide-variety of circumstances – a generalization process Bennett calls "Trust," or "direct grounding."

The process seems to be for the speaker, "Y holds for all things properly satisfying X in all circumstances."

When World and No-World Analyses Collide:

The problem comes when a backwards conditional may be interpreted either by direct-grounding or histories-involving analysis. For instance:

> If Nic played for the A's, he'd be their worst player.

Under direct-grounding, one might agree with this conditional, appealing to the fact that no professional baseball player is worse at playing baseball than Nic.
However, using the Lewis-type grounding one might find that the closest world to α in which Nic plays for the A's (perhaps by being placed in little league then training hard in high school and college) is one in which Nic isn't the worst player on the team. And hence, that $A>\neg C$ is true (an apparent stand-off).

Bennett assures the reader that this clear instance of ambiguity is rarely encountered in practice, resolved by context when it is, or just plain inconsequential.

Conflicting Conditionals:

Nevertheless, Bennett is compelled to discuss a similar yet more serious set of ambiguous conditionals, namely those in which both $A>C$ and $A>\neg C$ are true. Bennett introduces some properties of stand-offs:

1) The stand-off depends on two different sets of information.
2) The stand-off involves two different meanings for the conditionals.
3) The stand-off involves two conditionals which could, in theory, be held by one person at one time.

So Gibbardian stand-offs satisfy (1) and never (2) or (3). While the extreme version of the stand-off mentioned earlier (history vs. direct grounding) contrastingly satisfies (2) and (3) and never (1).

Almost-Gibbardian Stand-Offs:

Bennett also mentions a "new" type of stand-off that occurs when conflicting subjunctive conditionals are both true using direct-grounding. One possible template of such a stand-off is provided:

a) Something both F and G cannot be H.
b) What is F and H is High.
c) What is G and H is Low.

A person Trusting (a), (b), and (c) as well as knowledge that object X is both F and medium naturally concludes that X is not H and that:

If it were H, it would be high. (Relying on Trust in (b))

While a new person similarly Trusts (a), (b), and (c) except only knows X is both G and medium naturally concluded that X is not H and that:

If it were H, it would be low. (Relying on Trust in (c))

The problem is, the conditionals are jointly inconsistent (hence not assertible together), while individually they seem perfectly true (and assertible).

Stand-offs of this sort, like those of Gibbard, satisfy (1) and never (2) or (3). I think this stand-off is more than *similar* to Gibbardian stand-offs – especially once one substitutes F with "situations in which the east gate lever is down", G with "situations in which the west gate lever is down", H with "situations in which the top gate is open", translates "High" as "situations in which the water flows through east gate", "Low" as "situations in which the water flows through west gate", and finally, "Medium" as "situations in which the water doesn't flow." Given this, it seems Bennett needs to do more work distinguishing the two. Additionally, I have troubles even seeing Bennett's stand-off involves BCs, given he has only talked about direct grounding in reference to BCs (like my preliminary examples).
None the less, Bennett shrugs these issues away as being irrelevant to the history-involving analyst of BCs.