Aaron Leung
Philosophy 290-6
Fall 2004
Week 9 Handout

12
Unrolling from the Antecedent Time

74. CLOSENESS AND SIMILARITY

Consider this claim about the closeness of worlds, along with a corresponding logical principle:

C_1: No world is closer to α than α is to itself. That is, the class of worlds closest to α includes α.

P_1: A>C entails A⊃C

C_1 stands or falls with P_1; P_1 implies C_1 because if C_1 is false, then there is some world w is closer to α than α is to itself; pick an A that is true in w and α, and pick a C that is true in w and false in α; then A>C is true while A⊃C is false, which falsifies P_1. Now consider a stronger claim about closeness, along with another logical principle:

C_2: No world is as close to α as α is to itself. That is, the class of worlds closest to α contains *only* α.

P_2: A&C entails A>C

C_2 implies P_2 (in addition to P_1); if A and C are true at α, then C is true at all the closest A-worlds because α is the only closest A-world according to C_2.

Everyone in the literature agrees that closeness involves similarity; when we think about A>C, we think about worlds in which A is true and everything else is pretty much the same. Our intuitive notions of similarity support C_1; certainly α should be one of the worlds most similar to α. But what about C_2? Is α the only world most similar to α? What about the idea that there can be differences that make no difference? Similarity can be made to give us C_2 if we talk about similarity down to every last detail—'all-in' similarity, as Bennett puts it. Then any given world is closer to itself than any other world is to it. But it isn't clear whether we want this result, or P_2 which follows from it.

75. A > BIG-DIFFERENCE

It turns out that all-in similarity doesn't work so well for the closeness relation because it requires us to declare false all subjunctives that have the form A>Big-difference, when many such subjunctives are intuitively true. For example: "If on July 20 1944 Stauffenberg had placed the bomb on the other side of the trestle, Hitler would have been killed." This is almost certainly true, but all-in similarity arguably tells us to declare it false because any A-world in which Hitler survives the blast is more similar to α than any A-world in which Hitler dies early.

Basically, we believe in many conditionals of the form A>Big-difference because it seems plausible that small differences in the vicinity of the antecedent can amplify and become big differences in the consequent. This suggests that we only want to consider similarity up to the time of the antecedent. So we start with:

(1) Worlds should be compared with respect to their states up to the time of the antecedent—call that time T_A. Do not take into account differences between possible worlds and the actual world after T_A. So, if we're comparing two worlds in which Stauffenberg plants the bomb a little to the right, with Hitler dying one and not in the other, then the difference between Hitler's death and survival doesn't affect closeness to α because the difference pertains to a post-antecedent time.

However, this isn't enough. By itself, the state of a world at a given time says nothing about what we should expect that world to be like in the future, and given two different worlds which respectively satisfy A>C and A>~C, nothing so far tells us which to prefer. Therefore:

(2) We also need to consider causation. The closest A-worlds, in addition to being similar to α up to T_A, must also share the same causal laws as α. On this account, two worlds that are exactly alike up to T_A may not be exactly alike at all future times if the causal laws are indeterministic. To put it in more detail: if C obtains in α at T_C but the state of affairs at T_A isn't sufficient to guarantee C, then there exist two *causally* possible worlds which are alike at T_A but differ at T_C because C obtains at one and not the other. So we now limit our considerations to causally possible worlds; all others are worlds in which miracles occur.

So, according to this account, the truth of A>C depends on whether C is true at the A-worlds that (1) are similar to α with respect to states of affairs up to the time of A (or perhaps exactly like α in the case of indeterministic worlds), and (2) and exactly like α with respect to causal laws. According to Bennett, pretty much everyone accepts this.

[I have elected to skip section 76 (objective indeterminateness) for the discussion.]

13
Forks

77. BEFORE THE ANTECEDENT TIME

Let's look more closely at the two clauses in our account of closeness.

(1) Causally, it should make no difference whether we consider the world at T_A, or at all the times preceding T_A. If two worlds are exactly alike at T_A and the laws of α are deterministic, then those worlds will be exactly alike forever before and after T_A; on the other hand, if the laws of α are indeterministic, then those worlds may diverge. What we need not worry about is the case in which two worlds are unalike before T_A, then become exactly alike at T_A, and then become unalike again after T_A. Therefore we only need to consider similarity *at* T_A.

[Bennett seems to assume here that causal laws are Markovian; that is, there is no nonlocal causal influence, and causal laws don't need to appeal to anything before T_A in their determination of what happens afterward. Another issue to keep in mind is what counts as a state of affairs for a world. If we think of states in the classical mechanical sense, as just the positions and velocities of particles, then what Bennett says in (1) about convergence and divergence is wrong given indeterminism. But if we think of states in the quantum mechanical sense, as including the probability distributions of token events, then what Bennett says is fine. This will come up again.]

While causal inference only requires us to consider the world at T_A, logic may require us to consider times before T_A. For example, we may have a conditional whose consequent simply refers to a pre-T_A time: "If it had rained on the night of October 25 2004, then Aaron would have made up an example referring to the 1944 Stauffenberg bomb incident." The truth of this conditional depends on there having been such an incident sixty years before T_A, but this excursion into pre-antecedent time is motivated by semantics or logic, not causality.

(2) A world is *legal* if it conforms to the causal laws of α at all times. The term 'Legal' names the theory that closest A-worlds must be legal. Notice that if α is deterministic, then the closest A-worlds that are not α itself are unlike α at all times preceding and following T_A. Even if these differences are small in the vicinity of T_A, they may become very large with temporal distance (in fact, they almost certainly will, given the complexity of worlds).

Determinism seems to pose problems for counterfactuals (how could the antecedent be true when it was determined to be false?) and therefore for subjunctives. According to Bennett, however, indeterminism is neither necessary nor sufficient to rescue subjunctives from this kind of difficulty.

Bennett says that indeterminism isn't sufficient because we don't know which parts of the world are indeterministic; even with quantum theory, it isn't clear that indeterminacies at the quantum level manifest themselves as indeterminacies at the macroscopic level. A great deal of the world of 'medium-sized dry goods' may actually be deterministic. [There's a big problem with this. According to quantum mechanics, there is in fact indeterminism at the macroscopic level; it's just that certain events, such as the spontaneous disappearance of cars, have an extremely small probability. But even an event with a probability of zero isn't nomologically impossible; there's still some world at which it occurs, and this is enough to make or break a subjunctive.]

Indeterminism isn't necessary because, given a world which is exactly like α except for the occasional small miracle, it's plausible that such a world is closer to α than is a world which conforms exactly to α's causal laws but is otherwise unlike α with respect to states of affairs. (This is, in other words, a rejection of Legal.) The basic idea is that enough differences in matters of fact can yield greater dissimilarity than small, localized miracles.

Note that Legal may require us to consider, as candidates for closest, worlds which are unlike α for all times before T_A. Many seem to find this disagreeable, and it has motivated the 'scare stories' in the following section.

78. TWO SCARE STORIES ABOUT BACKWARD CONDITIONALS

One feature of Legal is that it allows backward causal inference and thus backward subjunctives, ones of the form "If A had obtained, then C would (have to) have obtained earlier." Additionally, it doesn't prevent combinations of forward and backward subjunctives: "If A had obtained, the C would have obtained later and B would have obtained earlier." However, this gets us into trouble according to Downing. Consider the following scare story:

> Aaron has never read *Pride and Prejudice*, in addition to which he procrastinates and wouldn't have read it even if it had been assigned, so *if he had turned in a paper on it, he would have gotten a bad grade*. On the other hand, Aaron doesn't like getting bad grades, is too proud to submit a bad paper, and has a Blockbuster™ card, so *if he had turned in a paper on* Pride and Prejudice*, he would have rented and watched the A&E film adaptation (starring Colin Firth) the night before*. Armed thusly with knowledge of the book, *if Aaron had turned in a paper on it, he would have gotten a not-bad grade*.

The two conditionals Submit>Bad-grade and Submit>Not-bad-grade can't both be true. According to Downing, the mistake is to allow the backward subjunctive Submit>Rent-movie-the-night-before. His conclusion is that we must do without backward subjunctives as well as Legal which permits them, and that a correct account of subjunctives will not imply that closest worlds are unlike α before T_A.

However, it turns out that the scare story, and therefore Downing's argument, fails. The backward subjunctive by which we obtain the incompatible forward subjunctives is unnecessary. Consider this revised scare story:

> Aaron has never read *Pride and Prejudice*, so *if he had turned in a paper on it, he would have gotten a bad grade*. On the other hand, he doesn't like getting bad grades and is too proud to submit a bad paper, instead preferring to ask for an extension or submit nothing at all (or, heaven forbid, actually do the assigned reading). So we can be sure that *if he had turned in a paper on* Pride and Prejudice*, he would have gotten a not-bad grade*.

The real problem is that for one conditional we focus on Aaron's ignorance and laziness while ignoring his academic conscientiousness, and for the other conditional vice versa. This generates conflicting ideas about what the closest worlds are like, and therefore we derive conflicting conditionals. But it isn't necessary to take a trip into the past.

Bennett then discusses his own scare story for rejecting backward subjunctives. First, the factual background: on June 22 1941 German armies invaded the Soviet Union; on December 6 1941 they reached the suburbs of Moscow and were beaten back by Soviet armies and the Russian winter. Now consider: "If the German army had reached Moscow in August 1941, it would have captured the city." This may be true; suppose that we must infer backward in order to evaluate it. We must consider worlds at which the German army enters Moscow in August, and then consider what could have brought that about, etc. Things must have been different from α in July 1941, and then June 1941, and so on, perhaps back to the beginning of time.

The difficulty that this account is supposed to raise is that it tells us we shouldn't believe August>Capture. If the closest A-worlds are unlike α for a very long time preceding T_A, then we may end up with a situation at T_A that lacks the conditions for securing C. In the case of August>Capture, if we causally infer backward and then forward again, things might turn out differently enough such that the brilliant commander of the German army was never born, without whom we no longer believe in the truth of August>Capture.

However, there's a problem with this scare story as well. The point was to track backward and forward through a possible world in order to undermine Capture. But in order to undermine Capture, we had to combine the backtracking with facts about α, such as the capabilities of the German army's commander. That's where the trouble comes from.

This leads Bennett to present a version of Legal which he calls the Simple Theory: A>C is true iff C obtains at the legal A-worlds that most resemble α in respect of their state at T_A. This says nothing about what goes on before and after T_A, and allows backward and forward inference. But inferences about a particular world don't rely on states of affairs at other worlds, so backtracking is harmless, and therefore Legal is OK.

79. BUMPS

One advantage of the Simple theory is that it allows smooth run-ups to the antecedent. This is unlike the analysis by Jackson in which A>C is true iff C obtains at all the A-worlds which are exactly like α up to T_A and which conform to α's laws from T_A onwards. This analysis requires that the antecedents of many subjunctives come true more or less miraculously—that is, through a sudden 'bump'.

This account is problematic because it assigns incorrect truth values to many subjunctives. Consider Bennett's example in which a dam bursts at 8:47 pm, and as a result nine motorists are killed, with no other deaths occurring. The conditional "If no cars had been on the road just then, then no lives would have been lost." On Jackson's account this is true, as desired. However, suppose that in the closest A-world there were cars on the road up to 8:47, and then at 8:47—bump!—the cars vanish just as the dam bursts. Then Jackson's theory would declare true the conditional "If no cars had been on the road just then, we would be investigating the mystery of where they had all gone to." But intuitively this should be false, because we really imagine a sequence of events in which it smoothly comes about that no cars are on the road at 8:47.

[But then, what exactly does "smooth" mean? If it means accordance with causal laws, then in an indeterministic world, bumps can occur, although they're improbable. The evolution of these probabilities may be 'smooth', but token events don't have to be. Perhaps what Bennett wants is an account of smoothness that allows us to ignore highly improbable token events.]

Jackson denies Bennett's intuitions about smoothness, saying that "When we evaluate [forward subjunctive conditionals] we set aside questions of as to what prior conditions might have led to the antecedent. We proceed as if the attribution had been miraculously realized. That is, we ignore causality prior to T_A." If by 'miraculous' Jackson means 'bumpy', then this statement is almost okay, except that he seems to be saying that we *always* reason by way of miracles. Anyway, this account does get certain cases right, for example "If he had turned left at the end of the lane, then …" If we believe in something like unconstrained free will for which choices have no determining cause, then Jackson's theory gets this example right; if we don't, then surely we can tolerate an infinitesimal miracle inside the driver's head.

80. HISTORIES FOR ANTECEDENTS

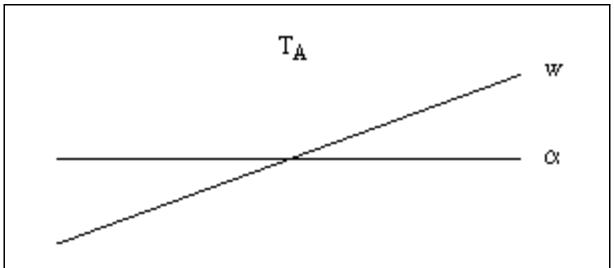
According to Bennett, Simple doesn't suffer from the bump defect. [But why not? After all, simple only talks about similarity at T_A; from that point on it's up to the causal laws to determine how things unfold. If the laws allow bumps, then there may be bumps.] It's also symmetrical with respect to forward and backward inferences, and is simpler than its rivals. Unfortunately, it isn't true.

The problem is that Simple demands as much resemblance as we can get at T_A, and the history from A will be filled in by causal laws alone. But this doesn't match against common usage; it's often the case that A will implicate more dissimilarities with α than it explicitly mentions. For instance, in August>Capture, when the Germans reach Moscow in August, we expect a few other things to be different—such as a higher concentration of Soviet troops in the vicinity of the city, caused by the German's early advance. Thus, an antecedent suggests at least a brief history which precedes it. (An additional difficulty for Simple is the case mentioned earlier, concerning conditionals which logically or semantically depend on the past.)

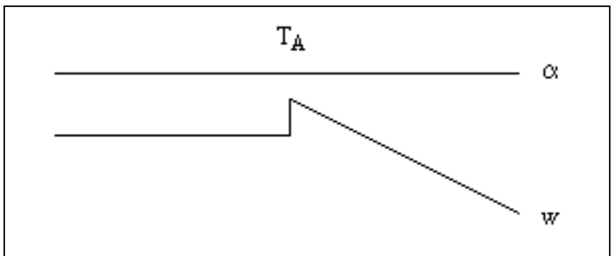
81. THE NEED FOR FORKS

So what we want in an analysis of A>C is a run-up to the antecedent, preferably a smooth 'ramp' from a history that is pretty much like the history of α. The closest A-worlds should (1) resemble α up to some moment shortly before T_A, and then (2) diverge—that is, 'fork'—from α without a sudden or noticeable bump, and finally (3) unfold in a causally coherent way so that A becomes true at them at T_A. [Note that there is still the problem of what we mean by 'smooth'. It seems that the most that Bennett can safely claim is that our desire for smoothness arises from our use of language and our beliefs about the world, but he can't require smoothness for his metaphysics.]

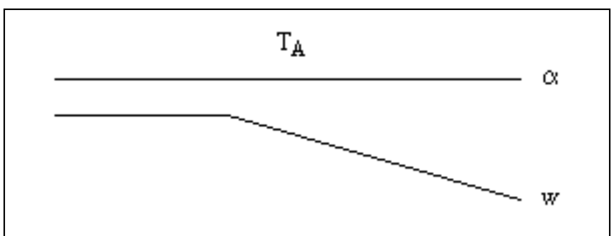
We desire (1) because if the relevant worlds resembled α in some respects but differed in too many others, we would encounter the problem of Bennett's scare story. We desire (2) because if the closest A-worlds forked from α with a noticeable bump, it would be as bad as Jackson's account by which we have a noticeable bump at T_A. We desire (3) because it just makes sense. Notice that we must amend our requirement for legality, pushing it from the time of the antecedent to the time of the fork. The following diagrams illustrate the three accounts discussed so far:



Simple



Bumpy



Forky

One thing to keep in mind is that when we speak of a world as forking off from α, it doesn't mean that one world splits into two. Rather, it means that two worlds, w and α, are alike up to the moment of the fork, and then become unalike in some suitably inconspicuous manner.

82. WHAT HAPPENS AT A FORK?

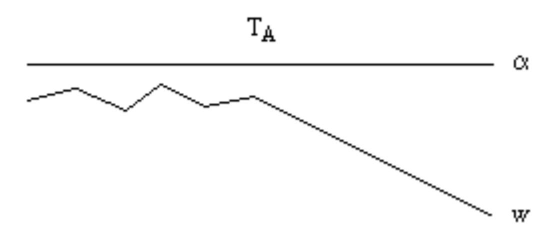
So we now have account that reflects our intuitions; when we think about A>C, we think about worlds that diverge in an initially modest way from α in order to legally yield A, and this divergence should not admit conspicuous bumps as Jackson's theory does. We want to consider worlds which reflect the intuition that A could plausibly have occurred in α. (This account doesn't cover cases of the form "If a miracle had occurred, then …", but it's still a reasonable default. [In fact, it doesn't cover all of ordinary indeterminism either.])

Let's review some of the terminology that's been tossed around. Given a counterfactual A>C, a *fork* is an event at an A-world by virtue of which that world for the first time becomes unlike α. A *ramp* is the segment of that world's history starting at a fork and ending at the obtaining of A. A world or ramp is *legal* if it conforms to the causal laws of α. Now we can say:

To evaluate a counterfactual A>C we must look to worlds at which A obtains by virtue of a legal ramp running from a fork that occurs not long before T_A.

So what happens at a fork? Lewis pointed out two possibilities: (1) an indeterministic causal transaction starts out identically at w and α, and has different outcomes in each world, both of which are legal; (2) a deterministic causal transaction starts out identically at w and α and has one outcome at α and a different outcome at w, the outcome at w being in conflict with α's causal laws—that is, a miracle, albeit a small one.

There's a third possibility not mentioned by Lewis: that of the 'exploding difference' (more precisely, *late* exploding difference). It's conceivable that a legal world w can be unlike α in only minute and imperceptible ways for a very long time, and then eventually the differences amplify and w diverges noticeably from α in some way that satisfies subjunctive that might interest us.



Exploding Difference

On the exploding difference account, a fork is an event through which a world becomes 'less than sufficiently like α'. One thing is 'sufficiently like' another if they are alike in any way that we could ever notice or care about; perfect likeness is an extreme case.

Exploding difference worlds are extremely improbable, but they are still legal, so according to Bennett they ought to be considered.

So now we have a couple of different kinds of closest world, but according to Bennett they're equivalent for the purposes of evaluating the truth conditions of subjunctives. The three accounts (that is, Bumpy, Forky, and Exploding Difference) all agree that the closest A-worlds are all, until shortly before T_A, exactly or pretty much like α, and we can ignore the difference between small miracles and legal bumps.

It may be the case that several different ramps lead up to some A. Sometimes this won't matter, if C is true at all of those differently-ramping A-worlds. However, the truth of a conditional may depend upon which ramp we take, and it may not be possible to evaluate such a conditional at all. This concern leads us to consider constraints on ramps.

83. LOCATIONAL CONSTRAINTS ON FORKS

In analyzing A>C, we naturally need to put some constraints on how early the fork can be. Quantifying over forks won't work; if we say "A>C is true iff C is true at *some* A-world reachable by an inconspicuous fork from α", then A>C will be made true too easily, and there may be forks which secure C in some A-worlds and ~C in others. If we say "A>C is true iff C is true at *all* A-worlds reachable by an inconspicuous fork from α", then it will be too difficult to make A>C true because there will be too many possibilities to consider; we can't know every single contingency that would have led to A's obtaining, and even if we did, we would still have the problem of evaluating whether C obtains as well.

Therefore we need a locational constraint. Lewis' proposal of analyzing subjunctives in terms of 'the latest admissible fork gives us wrong answers'; consider Bennett's example in which his coat could have been stolen from a restaurant earlier or later, and the perpetrator of the later theft is someone who always sells his stuff to Fence. If we only admit *the* latest fork, then we would accept "If my coat had been stolen from the restaurant, then it would now be in Fence's shop", whereas we have no reason to believe this. So Bennett settles on *a* latest admissible fork, where two forks can both count as 'latest' if their times of occurrence are not too far apart. [Unfortunately, this is vague and says nothing about how to adjudicate between a number of latest admissible forks.]

[I have also elected to skip section 84 (doing without forks), but I will include notes about it in the final revision of the handout.]