Chapter 19 Subjunctive Conditionals
And Time's Arrow
(Plus Antecedent Relativity)

§ 114 Explaining the Arrow of Time

Bennett's account of subjunctive conditionals is based on Lewis's. However, there are differences. This is the topic of Chapter 19.

Lewis wants his account of subjunctive conditionals to help explain temporal asymmetry, the difference between past and future. Although Bennett uses the blanket term "time's arrow" for the direction of time, there are a few different categories of temporal asymmetry, e.g. entropic processes, expanding light spheres, consciousness, and the dependency of the future on the past. Lewis's account deals with the last of these.

Lewis wants his analysis of subjunctive conditionals to explain our sense that the past is fixed and that there are many possible futures, which he terms the Asymmetry of Openness (AO). He also wants it to explain the Asymmetry of Causation (AC), the fact that causes predate their effects.

Neither the future nor the past can change. Logically, whatever happens at a future time Tf will be whatever happens at time Tf. However, we can say that the past affects the future.

A affects C : A is true, C is true, (A>(C is true.

Lewis then says that C counter-factually depends on A.

Earlier times affect later times because most true counterfactuals have an antecedent which concerns a time which is earlier than the time of the consequent. This is all there is to openness. A causes C if there is a chain of counter-factual dependencies of this sort which connects A and C. Bennett adds that for Lewis, causation only holds between events.

Lewis thinks that it makes sense to speak of the future affecting the past in exotic situations such as sci-fi time travel, precognition, Godel topologies, etc. Presumably, in such scenarios there is a true counterfactual (A>(C, where A is later than C, and, intuitively, we judge that A causes C.

The problem is that Lewis's analysis seems to imply that the future affects the past in ordinary, everyday situations. There are true, backwards counterfactuals which we do not think involve backwards causation. These are conditionals which say that if it were the case that A, then C would have to have obtained. We covered these last week.

Bennett's example:

> If Adlai Stevenson had been the undisputed President in February 1953, then he would have to have been elected in November 1952.

On Lewis's analysis this implies that Stevenson's not being president in 1953 caused him to lose the election.

(Note that the consequent must be on the ramp. If the consequent is before the fork, then this is a non-interference conditional, which does not display counter-factual dependency.)

Lewis's answer is that in backwards conditionals the relationship between A and C is not definite and detailed.

Bennett replies that in this example, being President implies being elected in as definite and detailed a way as one could imagine. However, Bennett goes on, perhaps what Lewis means is that there is no relationship between events here. Lewis recognizes causal relationships between events, but not between facts. Stevenson's not being President in February 1953 is a fact, not an event.

However, Bennett produces another counter-example, which he thinks does involve events. Suppose a runner, Lucy, nearly wins a race, but stumbles and falls at the last minute, losing. In order to have won, the race, she would have to have not fallen. Losses and falls are events. So,

((L) (Lucy lost the race > (Lucy fell.

If so, then Lewis's account must judge that her losing the race caused her fall.

However, Bennett may be equivocating here between winning and not losing. Even if we grant

she won the race > (she fell,

((L) does not follow. And while winning is an event, losing looks less like an event under examination. The actual event of her loss is identical with a specific competitors win.

Say she has two competitors, Anne and Beth. Anne won. Now, we don't have any reason to suppose that if Anne had not won, then Lucy would not have stumbled. Perhaps, if Anne had not won, then Beth would have won. But the event of Anne winning is the same as the event of Lucy losing.

Or suppose Lucy's only competitor is Anne. If a small miracle can prevent Lucy from falling, then it seems reasonable that a small miracle can lead to Anne's also falling. In which case, the world in which Lucy does not lose because Anne also fell is as close as the world in which Lucy does not lose because she didn't fall. So, ((L) is not true when evaluated between events. It only looks plausible when "Lucy loses" is considered as a fact.

And it only looks plausible if one takes the lateness of the fork criterion very strictly. But remember that Bennett says lateness of fork should be taken loosely. Well, the earlier you allow the fork, the less plausible ((L) appears. Professor Fitelson points that if Lucy had not lost it may have been the case that she had never entered the race at all; or that if she had not lost, it would have to have been the case that the other runners were running more slowly, perhaps.

It may seem that this defense of Lewis's postion relies on being overly demanding on what constitutes a true conditional; more demanding than for other conditionals. Well, it may be the case that evaluating conditionals between events for the purposes of determining causal relations calls for more rigorous evaluation. Combining this demanding assessment of conditionals with a requirement that events be characterized in their minute details brings the asymmetry of event conditionals into focus. Apply Lewis's concern with "definite and detailed" relations to the descriptions of the events. Then, causal relations hold between events which have specific precise characteristics and take place at specific precise times and places. So, if any of the events which played a causal role in Lucy's loss had not occured, then she would not have lost in that precise manner. The more precise the specification, the easier it is for "If she had not fallen, she would not have lost" to be true. If she had not fallen, then she would certainly not have lost in the specific manner in which she did lose. On the other hand, if she had not lost in the way she did, it might have been because she lost by Anne coming in a second later, or a second earlier, or even by Lucy not falling. Each of the causal factors that play a role in a later event are necessary for that later event to occur. If any one of those factors had not occured, then the later event would not have occurred. So, the non-occurrence of that event only requires that some one of those factors did not occur, and does not require a specific one to not have occurred. For any given event, this yields many true forward counter-factuals, but no true backward counter-factuals. This is related to what Lewis terms the asymmetry of over-determination. Each of an events effects determine that that event occurred. But the event does not determine all or any of its effects.

Leaving aside the racer and events, the Adlai Stevenson example still suggests the following problems:
Bennett would like an account of the asymmetrical dependency of facts.
This account still seems to imply that the recent past is open and can be affected by a subsequent time.
Lately Lewis has admitted that although there is not fact-causation, there can be causation without events, "causation with no causal relata at all." I don't know what to say about this.

The moral: Lewis's account of subjunctive conditionals does not capture the temporal asymmetries of openness and causation. (Note: these "morals" are just a quick gloss on Bennett's main points in each section. They may contradict what I suggest in the notes.)

§ 115 Keeping Temporal Order Out of the Analysis

Bennett and Lewis's accounts are essentially equivalent. The difference is that Lewis's account does not use the notions of earlier and later. Bennett says these are concepts of temporal order, or temporal direction.

Bennett is correct that these concepts are both temporally ordered and temporally directed. However, being ordered and being directed are distinct properties. One can say that point B is between points A and C without saying whether A is to left of B, or vice versa. The betweeness relation establishes ordering, but not directedness. Lewis's account avoids both temporal directedness and temporal order, but it may not need to.

Because Bennett just writes of "time's arrow" without distinguishing types of asymmetries it may sound like Lewis is worried about explanatory circularity. This would be a concern if Lewis wanted to show that the Asymmetries of Openness and Causation resulted from an asymmetrical analysis of subjunctive counterfactuals. Actually, Lewis wants to show that AO and AC result from the asymmetry of counter-factual dependency, and this asymmetry results from a symmetrical analysis of conditionals when applied to underlying asymmetries in the world. But I suspect that Lewis is being unnecessarily abstemious in denying himself both temporal order and temporal direction.
He might avoid the difficulty of § 114 if he simply stipulated that there is no causation when the consequent of a conditional lies between the fork and the antecedent.

We'll assume for the purpose of argument that in each case (A > (C. I and II involves forks from (. III and IV involve forks to (. Bennett judges I and III to be cases of causation because A is earlier than C. On my proposal, I and IV are cases of causation because A is between the fork and C. These evaluations disagree regarding III and IV, but that shouldn't bother Lewis, since he thinks it is just an empirical fact that we do not have to be concerned with these sorts of forks. The important point is that it is possible to get the same judgement on I and II without using direction in the analysis.

Perhaps this is all too ad hoc. Is being between a fork and some other point in time conceptually related to causation? It would be interesting would be to see if we are more inclined to call IV an instance of causation than we are to say that III is.

In any case, Lewis's analysis differs from Bennett's in its use of four ranked criteria for similarity:

If w1 and w2 are two worlds, w1 is closer to (than w2 if

(1) w2 contains a large miracle and w1 does not.

If neither contains a large miracle then w1 is closer if

(2) w1 exactly resembles the actual world for more time than w2 does.

If they are equal by (1) and (2) , then w1 is closer if

(3) w2 contains more small miracles than w1 does.

If they are equal by (1), (2), and (3), then w1 is closer if

(4) w1 has a greater degree of imperfect similarity to the actual world than w2 does.

Criterion (2) does the work of ensuring that the fork is a late fork without using the temporally directed concept of lateness. Criterion (1) rules out bumps.

Bennett raises a problem here. Bennett's analysis says that the closest A-worlds must be legal after the fork. This is supposed to be taken care of by (3).
Now consider:

If Nixon had pressed a certain button, then a nuclear war would have ensued.

World w1 at which the button is pressed diverges from (by a small miracle and nuclear war ensues. W2 diverges from (by a small miracle, then perfectly reconverges by a second small miracle. W1 is closer by (3), but w2 is closer according to the higher priority (2). This implies

Big Deal: If Nixon had pressed the button, no one would have noticed.

In response, Lewis argues that small miracles cannot make worlds converge to perfect likeness. In fact, he believes that this is the deeper asymmetry which underlies the phenomenal asymmetries of openness and causation.

§ 116 The Metaphysics of World Convergence

Lewis argues that a single difference rapidly ramifies in various, numerous fashions. One might say that it explodes. Undoing all these repercussions takes a big miracle. To get the Big Deal consequent, you have to erase light waves, Nixon's memory, an electrical current, fingerprints, etc.

Bennett: If the interval between antecedent and the reconverging miracle is on the order of a nano-second, the repercussions won't be too varied for a small miracle to undo them.
Lewis: Yes, they will. And if the interval is even smaller, there may not be enough time for the antecedent to happen at all.

Apparently Tooley argues that there are worlds where an antecedent will produce a single consequence for a period of time, after which it branches to form numerous consequences. A miraculous halting of that consequence before the branch would only be a small miracle.

Lewis says that a world with only one atom in a void does not have asymmetry of miracles, but is too exotic to concern us. If Tooley's world is like this, then Lewis is unconcerned.

This objection is reconvergence, i.e. a fork away followed by a fork back. What about convergence simpliciter?

The Bennett Worlds

Bennett showed that a small miracle can make one world converge onto another. When w1 forks away from (, the real world, it forks onto another world, lets call it (1. (is the perfectly legal world that shares all of the pre-fork history of w1. (1 is the perfectly legal world that shares all of the post-fork history of w1. From our perspective, the small miracle pushes w1 away. From the perspective of an inhabitant of (1, the small miracle pulls w1 in. For every fork there is a corresponding (, or Bennett World.

It is not clear that there is a legal (corresponding to each fork. Kenny Easwaran points out that for cellular automata that are certain states which cannot be arrived at by evolving earlier states. Perhaps a miracle leads to a state which cannot be legally reached from an earlier state. However, I believe that the rules governing cellular automata are not generally time-reflexive. More importantly, the temporal reverse rules, whatever they may be, are not deterministic. Some cellular states may be arrived at by more than one starting point. This contrasts with the suppositions we are making regarding the laws of the actual world. From the stand point of a deterministic Newtonian physics, all the fundamental laws are time symmetric. So, the temporal reverse laws are also deterministic. Take some moment soon after a fork. Take the temporal inverse of the state description of that moment, reversing all the velocities, etc. Some series of events is going to ensue. Let the world run on for as long as you like, preferably eternally. Take the temporal reverse of this half-history, slap it on to the history of the w after the fork. This should give you a legal history for (. Of course, I may be wrong about this.

Assuming for now that there are legal (worlds, most ordinary subjunctives which do not have Earth-shattering effects involve forking worlds that are nearly identical to our world for at least a short time after the fork. Outside the light cone of the miracle they are identical. Hence, the Bennett worlds are nearly identical to our world just after the fork. By and large, they contain the same records, clues, traces of earlier times that our world does.

But Lewis maintains

> Amplify: Two deterministic worlds that are somewhat like ours and are slightly different at some time, become greatly unalike at later times.

Bennett says that Amplify implies that the intermediate past of any (is very different from the past of (. This is because a small difference in the distant past would become amplified in the more recent past before becoming very much like our world at the fork.

I find this argument puzzling and I'm not sure Bennett has a firm grip on Amplify. I certainly don't. Do the differences amplify quickly? If not, then (can have a past like ours. It just becomes unalike after the time of the fork. Lewis must intend amplification to be rapid. But do all differences large and small get amplified? If so, then how does (manage to become like our world by the time of the fork? Perhaps this serves as a reductio of the hypothesis that (is a world like ours. But then we lose our basis for saying that the past of (is unlike our past. I think what Lewis wants here is just a flat denial that exploding differences are possible.

Moving on...
The upshot is that the records, clues, traces at (are not the result of the events that caused them at (. So, Lewis says, (is deceptive. Since (is legal, these records must be the result of a possible but extremely improbable decrease in entropy, or backwards entropic-like process. A footprint in the sand results from a chance confluence of sand grains moved into position by precise winds and their internal Brownian motion. Likewise, apparent dinosaur fossils are not the result of dinosaurs, and books are not the result of an act of writing. And so on.

Lewis says a world like this is so unlike ours that we should not even consider it when evaluating conditionals. Perhaps, we should change (1).

(1*) w1 is closer to (than w2 if w2 contains a large miracle or large areas in which the change in entropy runs opposite the change in entropy in adjacent areas and w1 doesn't.

What are the implications of perfect world convergence for Lewis's theory of conditionals? Elga has a paper that describes a world in which a small miracle causes it to converge on (. Furthermore, he argues that this world shows that Lewis's analysis provides the wrong truth value for some conditionals. Consider a scenario in which Greta cracks an egg into a hot frying pan at 8:00. So,

If Greta hadn't cracked the egg, then a 8:05 there wouldn't have been a cracked egg in the pan.

Because, at the closest worlds, one small miracle prevents her cracking the egg, and there is no egg in the pan at 8:05.

However, there is a world, which is just as close, in which a decayed egg coalesces out of the particles floating in the air. This egg sits in the frying pan, becoming fresher and fresher until sometime between 8:00 and 8:05 when a small miracle occurs, after which the history of this world perfectly matches that of (. This world is just as close as the above world since there is no large miracle at it, it exactly matches (on one side of the fork, and it only contains one small miracle. This world falsifies the conditional, because no one cracks the egg in it, but there is an egg in the pan at 8:05.

However, this world has an indefinitely large area of decreasing entropy within which time appears to run backwards. Outside this area everything appears normal. This anomalous area decreases over time until it vanishes at the fork. Henceforth, the world is indiscernible from ours. It is legal, but I don't see that this world is overall "as 'like ours' as we are entitled to demand," as Bennett claims. (1*) judges it to be very dissimilar.

The moral of §§ 115, 116: Perfect similarity at (2) may not yield the correct truth-values for subjunctives.

§ 117 Further Problems for Lewis's Analysis

1. Are exploding difference forks legitimate? If so, Lewis must make (2) weaker:
(2*) w1 resembles the actual world in every discernible detail for more time than w2 does.
If he does this, then Lewis has to argue that not only that perfect convergence is impossible, but that convergence within discernible detail is impossible.

Also, Bennett thinks that exploding difference forks are incompatible with Amplify.

2. So far we have been assuming determinism. Dropping this assumption leads to the problem we might call Particular Fact Reconvergence (PFR). This problem is similar to the problem of perfect reconvergence, but works at criterion level (4), instead of (2).

At w1 a miracle causes Nixon to decide to push the button. We get the desired nuclear war.

At w2 an indeterministic outcome causes Nixon to decide to push the button. But, while his finger is on the button, a small miracle causes him to change his mind. This is uncontroversially small because the other consequences of Nixon's decision remain. His fingerprint is on the button, any observer's heart is racing, whatever. But no war.

w1 and w2 each have only one small miracle, so they're not distinguished by criterion (3). But w2 has more imperfect similarity to (than w1 since there is no war at w2. So (4) judges w2 as closer, giving the wrong truth value to the conditional.

Bennett does not want to drop (4) which would amount to dropping PF(Particular Fact). Instead, he suggests that Lewis might change criterion (3) so that small miracles are only allowed to occur at forks.

Bennett cautions that this does not solve the problem of perfect world convergence.

The moral: Bennett's analysis avoids these difficulties by banning post fork miracles and stipulating perfect similarity before the fork. Lewis's analysis differs so that it can account for temporal asymmetries, but it doesn't reach that goal anyway. So, Bennett prefers his own account.

§118 Antecedent Relativity

Bennett used to think that Antecedent Relativity was related to asymmetry in the analysis of conditionals and was a difference between his analysis and Lewis's. If either of these things were right, this would be the chapter in which to address the topic. Neither thing is right, but little orphan AR needs a home and this is the only family she has ever known.

Lewis’s analysis of subjunctive conditionals uses absolute criteria of closeness. Lewis’s criteria (1) - (4) allow him to say of any two worlds which is closer to (regardless of the conditional being evaluated.

An analysis that is Antecedent Relative specifies that there are some worlds w1, w2, such that w1 is closer when evaluating one conditional, but w2 is closer when evaluating another conditional with a different antecedent.

Bennett thought his analysis was antecedent relative for two reasons. It specifies that a world is closer to (if the time of the fork was closer to the time of the antecedent. It also specifies that miracles do not occur after the antecedent.

However, he can get the same result while ignoring antecedents. For Bennett, w1 is closer than w2 if the fork in w1 is later than the fork in w2 (allowing for the vagueness caveat from § 83.) Also, w3 is closer than w4 if w3 only has a miracle at the fork and w4 has miracles elsewhere.

Causal chain PF also involves absolute closeness. It refers to facts regarding times after the fork, not the antecedent.

Absolute Closeness entails

> Substitution: A>B, B>A, A>C entail B>C

Bennett calls the biconditional for >, “>-equivalence.” A>B and B>A taken together imply that the set of closest A-worlds is the same as the set of closest B-worlds. A>C means that they are all C-worlds, as well. So, B>C.

This does not follow from Antecedent Relativity. A>B and A>C means that the closest A-worlds are both B and C when we have antecedent A. With antecedent B, however, some distant A and B world that is not also a C-world may come closer. So, we can have B>A without B>C.

Substitution implies

> Limited Transitivity: A>B, (A&B)>C entail A>C

Here’s the derivation:

> A>B entails A>(A&B)
> (A&B)>A is a necessary truth
> So, (A&B) and A are >-equivalent
> (A&B)>C is a premise
> Substitution gives you A>C

Limited Transitivity does not hold under Antecedent Relativity.

Limited Transitivity is useful for arguing against Transitivity:

Transitivity is false because

(1) If I had died yesterday I would be in heaven now.

(2) If I had glutted myself with arsenic-laced eclairs, I would have died yesterday.

but, since gluttony is a deadly sin, it is not the case that

(3) If I had glutted myself with arsenic-laced eclairs, I would be in heaven now.

How do we reconcile the invalidity of Transitivity with the following inference?

(a) If I'd gone to sleep earlier I would have woken up earlier

(b) If I'd woken up earlier I would have seen day -break

hence

(c) If I'd gone to sleep earlier I would have seen day-break.

Isn't this Transitivity at work? No, it's an instance of Limited Transitivity. Because we go through the argument quickly, we read (b) as "If I'd gone to sleep earlier and woken up earlier, I would have seen day-break."

On reflection, it's very easy to construct counter-examples to transitivity. Here's a quasi-algorithm for producing an indefinite number of counter-examples from every non-trivial apparent instance of transitivity.

A>B, B>C (A>C

Find some possible but unlikely D which is consistent with A&B, but inconsistent with C.

(A&D) > B

B>C

but

(((A&D)>C)

Finally, Bennett and Lewis qualify their Absolute Closeness. Gricean considerations can modify the closeness relation. See §71 Consequent as Context. When we interpret a conditional so that it is neither trivially true nor trivially false, the closeness is loosely relative to the conditional as a whole.

Bennett leaves us with an exercise. Find a scenario in which there are three prima facie eligible forks, one of which makes a subjunctive trivially true, one of which makes it trivially false, and one of which makes it interestingly true.