

CONDITIONALS LECTURE 4

1. Lewis's theory of subjunctive conditionals has the outline form: **$p > q$ is true iff the closest p-worlds are q-worlds**: or, equivalently, if some p.q-world is closer than any p.~q-world. We have already seen how a theory with this structure can in principle account for the failures of monotonicity, contraposition and transitivity that seem to play a role in the logic of subjunctive conditionals. But the details of how it does that will depend on what counts as closeness.

2. In his earlier theory (1973) Lewis treated closeness as a matter of overall similarity, in some loose and intuitive sense. For instance, if p is the proposition that Jones was at the meeting yesterday (a meeting that he actually missed), then the overall closest p-worlds are worlds at which he attends the meeting, but things are otherwise much as they actually are. So a p-world at which (say) he gets to the meeting on his bicycle is closer than a p-world at which he gets there by spaceship, notwithstanding that in Lewis's metaphysical system both worlds are equally *real*.

3. Two issues arise from this proposal. The first is that there may be a tie for the closest p-world; and this will make a difference to the truth-value of some counterfactuals. E.g. suppose that this fair coin was never tossed and let p be the proposition that it *was* tossed yesterday at noon. Then it seems fair by Lewis's criteria to suppose that the closest p-worlds include worlds at which it lands heads *and* worlds at which it doesn't. Letting q be the proposition that it lands heads then, it looks as though $p > q$ and $p > \sim q$ are both false, and so we have to reject the principle of **counterfactual excluded middle**, which is that $(p > q) \vee (p > \sim q)$ is true for every p, q. This may not be a decisive objection to Lewis's approach.

4. But the loose, intuitive concept of similarity creates other, more serious difficulties. The **future similarity objection** first appeared in Fine's critical notice (*Mind* 1975). "If Nixon had pressed the button there would have been a nuclear holocaust" is true or can be imagined to be so. Now suppose that there never will be a nuclear holocaust. Then that counterfactual is, on Lewis's analysis, very likely false. For given any world in which antecedent and consequent are both true it will be easy to imagine a closer world in which the antecedent is true and the consequent false. For we need only imagine a change that prevents the holocaust but that does not require such a great divergence from reality.'

5. In part as a response to this objection, Lewis proposed to replace the criteria for closeness in his 1979 paper 'Counterfactual dependence and time's arrow' (*Nous* 1979). There he used the following method to set out criteria for closeness: consider our intuitions about which subjunctive conditionals are true and then fit the criteria around that. The result of this was that he came up with the following procedure for deciding, of any two worlds w1 and w2, which of them is closer to the actual world:

(5.1) It is of the first importance to avoid big, widespread, diverse violations of law.

(5.2) It is of the second importance to maximize the spatio-temporal region throughout which *perfect* match of particular fact prevails.

(5.3) It is of the third importance to avoid even small, localized, simple violations of law.

(5.4) It is of little or no importance to secure *approximate* similarity of particular fact, even in matters that concern us greatly.

If we have two worlds w1 and w2 then the idea is that to tell whether w1 is closer than w2 we go down the criteria, and w1 is closer than w2 iff the highest-priority criterion on which w1 beats w2 beats the highest criterion on which w2 beats w1.

CONDITIONALS LECTURE 4

6. To see why Lewis ranks (5.1) so highly, consider that if our world is deterministic, then *any* world that matches our past and in which things diverge at some point from us, perhaps on account of different choices, must contain some miracle (i.e. some violation of our laws). Such worlds might be very close to ours though: even if determinism is true, we'd agree (Lewis thinks) that if e.g. you had turned left instead of right no *other* laws would have been broken. It isn't true e.g. that if you'd turned left then animals would have started talking etc.

7. To see how Lewis's ranking helps with the future similarity objection, where p is the proposition that Nixon presses it and q is the proposition that there is a holocaust; and suppose to keep things simple that only a miracle would prevent q from being true, if p is true. Then the closest $p.q$ -world matches the closest $p.\sim q$ -world on criteria (5.1)-(5.2) but wins on criterion (5.3); whereas the closest $p.\sim q$ -world wins on criterion (5.4). But since (5.3) has higher priority this means that $p > q$ comes out true as Lewis desired.

8. According to Lewis these criteria also explain the temporal asymmetry of counterfactual dependence: the fact that later times counterfactually depend on earlier times but not the other way around: the past doesn't, or doesn't very much, depend counterfactually on the future. Thus it is true e.g. that if you had caught an earlier train than you actually did then you would have got there sooner than actually you did, and things would then have been different from how they will actually be the next day, and the next day... for the rest of time; but it isn't true that if you had caught an earlier flight then things would have been different not only the day before but also at all earlier times e.g. in neolithic times.

9. The explanation for this phenomenon, which for Lewis also underlies the fact that causes always precede their effects in time, is that criterion (5.1) ranks higher than (5.2) and (5.2) ranks higher than (5.3). To see this, suppose the world is deterministic, let p be the proposition that you catch the earlier train and now consider these three p -worlds:

w1: The same as the actual world until shortly before you leave. Then a tiny miracle occurs – say, a neuron fires in your brain – and this causes you to leave a bit earlier. The world then unfolds in accordance with the actual laws.

w2: The same as the actual world until shortly before you leave. Then a tiny miracle occurs – say, a neuron fires in your brain – and this causes you to leave a bit earlier. Then a huge miracle occurs that wipes out all the post-decision *traces* of your choice to leave earlier, and the future of the world matches the actual future.

w3: The same laws as the actual world, and therefore different from it at all times. You catch the earlier train but also the dinosaurs were wiped out much earlier, the French Revolution never occurred etc.

Because (5.1) ranks higher than (5.2), w1 is closer to the actual world than w2. Because (5.2) ranks higher than (5.3), w1 is closer to the actual world than w3. If w2 and w3 had ranked closer then counterfactual dependence would not have been asymmetric; so it is this ranking that is (according to Lewis) responsible for that version of time's arrow. Note that the criteria (5.1)-(5.3) contain no temporal asymmetry in themselves. What causes the asymmetry is a contingent fact about our world, the fact that Lewis calls asymmetry of overdetermination.