Metaphysics // Fall 2025

Handout 21

Counterfactuals: Lewis

The core idea. The general form of counterfactuals is just this:

(1) If it were the case that ϕ, then it would be the case that ψ.

This subjunctive conditional is not meant (it seems so anyway!) to describe our actual world. It is meant to describe some other, merely possible world. Lewis' approach is to give an extensional analysis of (1) that uses quantification over possible worlds. The main insight is this:

(2) ' If kangaroos had no tails, they would topple over ' means: in any possible world where kangaroos lack tails, which *resembles* our actual world as closely as possible, given that lack of tails, the kangaroos topple over.

This sets up the discussion: we need to explicate the idea of similarity. But some important clarifications first.

True antecedents. Suppose I told you:

(3) If Trump were the US President in 2025, the inflation would be low.

You should be right to complain:

(4) But Trump *is/was* the US President in 2025!

The complaint is just, but it doesn't show that what I said was false, only that it was inappropriate. In fact, that's just the beginning of the complaint: what are you really complaining about?

Suppose that Abe the amnesiac doesn't know that Trump is the President. He is talking to two 27
doctors who play along. Then consider the following scraps of a dialogue:

(5) (Abe) If Trump were the US President in 2025, the inflation would be low. [$\phi \mathrel{\Box\!\!\rightarrow} \psi$]

a. (Ben) This is false [$\sim(\phi \mathrel{\Box\!\!\rightarrow} \psi)$]: for Trump *is* the President in 2025, and the inflation is high. [$\phi \,\&\, {\sim}\psi$]

b. (Cyd) This is true [$\phi \mathrel{\Box\!\!\rightarrow} \psi$]: for Trump *is* the US President in 2025, and the inflation is low. [$\phi \,\&\, \psi$]

Ben disagrees with Abe's assertion, not because of the true antecedent, but because of the conditional itself. The falsity of the conditional he derives from the truth of the antecedent and falsity of the consequent. Cyd agrees with Abe's assertion, not because of the true antecedent, but because of the conditional itself. The truth of the conditional he derives from the truth of the antecedent and truth of the consequent.

Now Lewis uses these facts to mount a general argument in support of his theory. But presently for us, we only need this evidence to show that counterfactuals with true antecedents may be rejected *and* accepted.

Remark 1. I am not sure why Lewis doesn't consider the possibility that Cyd should and would regard Abe's utterance as an indicative conditional, by analogy with the Laos example in page 4. I.e. ' that ' refers to the indicative conditional that Abe *should have* asserted. Compare also Lewis' remarks on ' belief ' in page 28.

Would and might. Lewis introduces two operators corresponding to ordinary language constructions:

(6) Would: $\phi \mathrel{\Box\!\!\rightarrow} \psi$: If it were the case that ϕ, then it would be the case that ψ.
Might: $\phi \mathrel{\Diamond\!\!\rightarrow} \psi$: If it were the case that ϕ, then it might be the case that ψ.

The two operators are defined in terms of each other, meaning we only need to define the truth conditions for one:

(7)
$$\phi \mathrel{\Diamond\!\!\rightarrow} \psi = {\sim}(\phi \mathrel{\Box\!\!\rightarrow} {\sim}\psi)$$
$$\phi \mathrel{\Box\!\!\rightarrow} \psi = {\sim}(\phi \mathrel{\Diamond\!\!\rightarrow} {\sim}\psi).$$

Question 2. Test the plausibility of the stipulations in (7) with the following examples:

(8) a. If Otto behaved himself he would be ignored.

b. If Otto were ignored, he might behave himself.

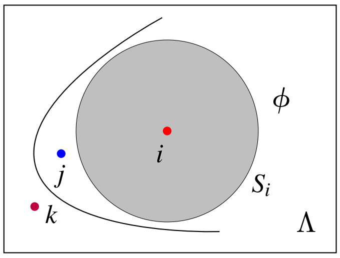


Figure 1: $\Box\phi$

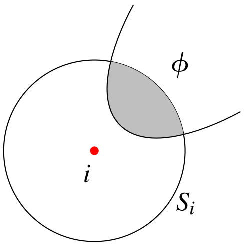


Figure 2: $\Diamond\phi$

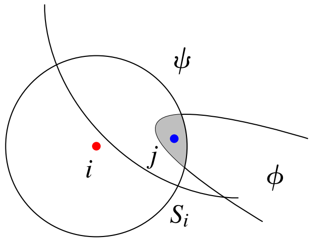


Figure 3: $\Box(\phi \supset \psi)$ 6

Ptolemaic astronomy. We are interested in showing that a counterfactual $\phi \mathrel{\Box\!\!\rightarrow} \psi$ is not the strict conditional $\Box(\phi \supset \psi)$. But first, we need to introduce some useful concepts and tricks.

Although we have interpreted the necessity operator as truth in all possible worlds, that's a simplification. We rather need to say that it is truth in all accessible possible worlds, i.e. accessible from the actual world.

We represent accessibility as a sphere S_i around the actual world i: namely, S_i is the set of all possible worlds accessible from i. Hence we get the diagrams for necessity, possibility, and strict implication (Figures 1, 2, 3).

Question 3. Spell out the truth conditions of strict implication using the Figure 3.

Correspondingly, we can interpret various degrees of strictness. Here are some examples of the accessibility relation that determines the scope of the sphere S_i:

(9) Logical necessity: $S_i = \Lambda$. 7

Physical necessity: S_i includes all possible worlds with the physical laws the same as in i.

Fatalistic necessity: $S_i = \{i\}$ when we require that each member of S_i is exactly like i in every regard. 8

Deontic necessity: Interestingly, we can have: $i \notin S_i$. That's the case when the actual world is not itself a morally perfect world.

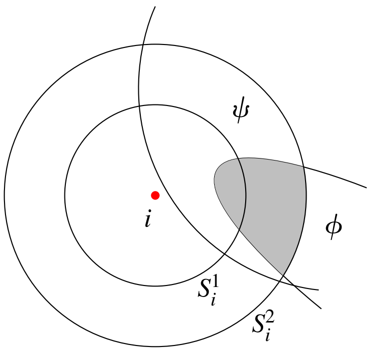
Vacuous necessity: $S_i = \varnothing$. Then for every ϕ, $\Box\phi$ is true at i.

Question 4. Draw explanatory diagrams for each of these cases of necessity.

Question 5. Draw a diagram for strict implication where the box operator is interpreted as logical necessity.

We can similarly illustrate stricter implications in Figure 4. As we have seen in the case of logical necessity, for example, the stricter the box, the larger the sphere S_i. The reason is that the greater strictness requires to consider more worlds where $\Box p$ is true at. Another way of putting it is to say that the stricter box entails the laxer one.

Question 6. Draw a diagram for a material conditional and for a vacuous conditional.



8

Figure 4: $\Box_2(\phi \supset \psi)$ is stricter than $\Box_1(\phi \supset \psi)$