Conditionals: between language and reasoning

Class 5 - Limit assumption and conditional excluded middle

November 24, 2017

Two choice-points for minimal change semantics:
the limit assumption and the uniqueness assumption.

Assumption	Limit	Uniqueness
For φ entertainable	$\mathrm{min}_w(\varphi) \neq \emptyset$	$\mathrm{min}_w(\varphi)$ is a singleton $\mathrm{min}_w(\varphi) = \{F(w, \varphi)\}$
Simplified clause: $w \Vdash \varphi \,\Box\!\!\rightarrow \psi$ iff	$\mathrm{min}_w(\varphi) \subseteq \|\psi\|$	$F(w, \varphi) \Vdash \psi$
Repercussions on logic:	No effect on propositional logic	CEM valid: $(\varphi \,\Box\!\!\rightarrow \psi) \vee (\varphi \,\Box\!\!\rightarrow \neg\psi)$
	Consistency of counter-factual consequences	Negation as opposite $\neg(\varphi \,\Box\!\!\rightarrow \psi) \equiv \varphi \,\Box\!\!\rightarrow \neg\psi$

Part I

The Limit Assumption

- The limit assumption is the assumption that, if φ is entertainable, there are always some closest φ-worlds.
- In Stalnaker's theory, this assumption is made:
indeed, in this theory there is always a single closest φ-world.
- By contrast, Lewis allows for cases in which we can get closer and closer φ-worlds with no end.
- Can such a situation occurs in making a counterfactual assumption?

Lewis's attempt

(1) If this line was more than a meter long, . . .

- Among the worlds where this is true, we find some where the line is
  - 2m
  - 1.1m
  - 1.01m
  - 1.001m
  - . . .
- Presumably, these count as more and more similar to our own world.
- Among the lines smaller than 1m, there is no smallest one.
- So, the limit assumption is violated.

Lewis's attempt

(2) If this line was more than a meter long, ...

- Among the worlds where this is true, we find some where the line is
  - 2m
  - 1.1m
  - 1.01m
  - 1.001m
  - ...
- Presumably, these count as more and more similar to our own world.
- But do they really?

- Certainly they do if similarity has its intuitive meaning.
  But this assumption is untenable (more in two weeks).

  Jones is possessed of the following dispositions as regards wearing his hat. Bad weather invariably induces him to wear his hat. Fine weather, on the other hand, affects him neither way: on fine days he puts his hat on or leaves it on the peg, completely at random. Suppose moreover that actually the weather is bad, so Jones is wearing his hat. *(Tichy 1976)*

(3) If the weather had been fine, Jones would still be wearing his hat.

- (3) does not seem true.
- However, fine-weather worlds where Jones wears his hat are intuitively more similar to our world than fine-weather worlds where he is not.
- So if similarity has its ordinary meaning, (3) is predicted true.
- Thus, the notion of similarity that matters for counterfactuals is not (always) the intuitive one.

- In fact, assuming the intuitive ordering in the line example leads to paradoxical conclusions.

(4) If this line was more than a meter long, ...

- With that ordering, the following sentences would all be true:

(5)
a. it would be more than a meter long.
b. it would be less than 2m long.
c. it would be less than 1.1m long.
d. it would be less than 1.01m long.
e. ...

- In fact, assuming the intuitive ordering in the line example leads to paradoxical conclusions.

(4) If this line was more than a meter long, . . .

- With that ordering, the following sentences would all be true:

(5)
a. it would be more than a meter long.
b. it would be less than 2m long.
c. it would be less than 1.1m long.
d. it would be less than 1.01m long.
e. . . .

- The line would be longer than 1m, but shorter than 1+ε m for each $\varepsilon > 0$.

- In fact, assuming the intuitive ordering in the line example leads to paradoxical conclusions.

  (4) If this line was more than a meter long, . . .

- With that ordering, the following sentences would all be true:

  (5)
  a. it would be more than a meter long.
  b. it would be less than 2m long.
  c. it would be less than 1.1m long.
  d. it would be less than 1.01m long.
  e. . . .

- The line would be longer than 1m, but shorter than 1+ε m for each $\varepsilon > 0$.
- But this is impossible!
- The counterfactual consequences of (4) would be inconsistent.

This problem can be turned into a general argument for the limit assumption. To see how, consider the following desideratum.

Counterfactual Consistency Condition (Herzberger 79):
the set of counterfactual consequences of an entertainable supposition must be consistent.

> *All the things that would have been true if Bizet and Verdi had been compatriots should form a coherent if somewhat sparse picture of a possible state of affairs.* (Herzberger 79)

Let's make this formally precise.

- Let's refer to a set of possible worlds p as a proposition.
- $w \Vdash p \iff w \in p$.
- All notions we gave for sentences can be given for propositions.
- p is an entertainable supposition at w iff p overlaps a sphere around w
- $w \Vdash p \mathrel{\Box\!\!\rightarrow} q \iff$ either of the following holds:
  1. p is not entertainable at w
  2. for some p-permitting sphere S around w: $p \cap S \subseteq q$
- q is a counterfactual consequence of p at $w \iff w \Vdash p \mathrel{\Box\!\!\rightarrow} q$
- Let $\Theta_w(p)$ be the set of counterfactual consequences of p at w:

$$\Theta_w(p) = \{q \mid w \Vdash p \mathrel{\Box\!\!\rightarrow} q\}$$

Counterfactual Consistency:
If p is entertainable at w, then $\Theta_w(p)$ is a consistent set.

Counterfactual Consistency:
If p is entertainable at w, then $\Theta_w(p)$ is a consistent set.

Theorem:
Counterfactual Consistency implies the Limit Assumption.

Counterfactual Consistency:
If p is entertainable at w, then $\Theta_w(p)$ is a consistent set.

Theorem:
Counterfactual Consistency implies the Limit Assumption.

Proof:

- Suppose the limit assumption fails at w.
- Then around w there is an infinite descending chain of spheres $S_1 \supset S_2 \supset S_3 \supset \dots$
- Let $S = \bigcap_{i \in \mathbb{N}} S_i$.
- Obviously $w \Vdash \neg S \;\Box\!\!\rightarrow \neg S$
- For each $i \in \mathbb{N}$, $w \Vdash \neg S \;\Box\!\!\rightarrow S_i$
- $\Theta_w(\neg S)$ includes $\neg S$ along with all the S_i for $i \in \mathbb{N}$
- But if all the S_i are true, then S is true. So $\neg S$ cannot be true.
- Thus, counterfactual consistency fails.

Part II

The Uniqueness Assumption

The uniqueness assumption is the assumption that, if φ is entertainable, there is a unique closest φ-world.

The uniqueness assumption is the assumption that, if φ is entertainable, there is a unique closest φ-world.

Negation as opposite

- An appealing consequence is that, for an entertainable φ, we have:

$$\neg(\varphi \,\Box\!\!\rightarrow \psi) \equiv \varphi \,\Box\!\!\rightarrow \neg\psi$$

- This accounts for the apparent equivalence between (6) and (7).

(6) It is not true that if you had opened the trunk you would have found a treasure. $\neg(\varphi \,\Box\!\!\rightarrow \psi)$

(7) If you had opened the trunk, you would not have found a treasure. $\varphi \,\Box\!\!\rightarrow \neg\psi$

More evidence for the equivalence $\neg(\varphi \mathrel{\Box\!\!\rightarrow} \psi) \equiv \varphi \mathrel{\Box\!\!\rightarrow} \neg\psi$.
The following pairs sound equivalent.

(8) a. I doubt that if you had opened the box you would have found a treasure. $B\neg(\varphi \mathrel{\Box\!\!\rightarrow} \psi)$

b. I believe that if you had opened the box, you would not have found a treasure. $B(\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$

(9) a. No student would have passed the test if they had goofed off. $\forall x\neg(Gx \mathrel{\Box\!\!\rightarrow} Sx)$

b. All students would have failed the test if they had goofed off. $\forall x(Gx \mathrel{\Box\!\!\rightarrow} \neg Sx)$

This is a puzzle for Lewis:

- in his theory, $\neg(\varphi \mathrel{\Box\!\!\rightarrow} \psi)$ is consistent with $\neg(\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$
- so, the following should sound consistent:

(10) I doubt that if you had opened the box you'd have found a treasure; but I also doubt that if you had opened the box you would not have found a treasure.

$$B\neg(\varphi \mathrel{\Box\!\!\rightarrow} \psi) \wedge B\neg(\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$$

(11) No student would have passed the test if they had goofed off, but no student would have failed if they had goofed off.

$$\forall x\neg(Gx \mathrel{\Box\!\!\rightarrow} Sx) \wedge \forall x\neg(Gx \mathrel{\Box\!\!\rightarrow} \neg Sx)$$

- Stalnaker also gives another argument for his view that conditionals do not involve a universal quantification over a set of worlds.
- To appreciate this argument, consider first a modal statement like:

(12) President Carter has to appoint someone to the Supreme Court.

- Depending on the scope of “someone”, this sentence has two readings:
  - narrow scope: $\Box \exists x A x$
  - wide scope: $\exists x \Box A x$
- This ambiguity is brought out in the following dialogue.

A: President Carter has to appoint someone to the Supreme Court.
B: Who do you think he has to appoint?
A: He doesn’t have to appoint any particular person;
he just has to appoint some person or other.

- B takes A’s claim in the wide scope reading; A corrects him.
- Notice that A’s correction, seems perfectly coherent, since the narrow scope reading does not entail the wide scope reading.

- Now consider an indefinite in a counterfactual consequent:

(13) President Carter would have appointed someone to the Supreme Court if there had been a vacancy.

- Depending on the scope of “someone”, Lewis predicts two readings:
  - narrow scope: $V \mathrel{\Box\!\!\rightarrow} \exists x A x$
  - wide scope: $\exists x (V \mathrel{\Box\!\!\rightarrow} A x)$

- The following dialogue should bring out the ambiguity:

A: Predident Carter would have appointed someone to the Supreme Court if there had been a vacancy last year.

B: Who do you think he would have appointed?

A: He would not have appointed any particular person;
he just would have appointed some person or other.

- Now consider an indefinite in a counterfactual consequent:

(13) President Carter would have appointed someone to the Supreme Court if there had been a vacancy.

- Depending on the scope of “someone”, Lewis predicts two readings:
  - narrow scope: $V \mathbin{\Box\!\!\rightarrow} \exists x A x$
  - wide scope: $\exists x (V \mathbin{\Box\!\!\rightarrow} A x)$

- The following dialogue should bring out the ambiguity:

  A: Predident Carter would have appointed someone to the Supreme Court if there had been a vacancy last year.

  B: Who do you think he would have appointed?

  A: He would not have appointed any particular person; he just would have appointed some person or other.

- However, here B’s reply actually sounds contradictory.

- For Lewis, this is unexpected: it should be possible to deny the wide-scope reading while defending the narrow-scope reading.

This is predicted in Stalnaker's account:
(assuming, plausibly, that in making the assumption we don't change the set of people)

- Suppose it is true that $V \,\Box\!\!\rightarrow \exists x Ax$.
- Then in the closest vacancy-world, there is somone who is appointed.
- So, someone is s.t. she/he is appointed in the closest vacancy-world.
- Thus it is true that $\exists x(V \,\Box\!\!\rightarrow Ax)$.

This shows that, at least when V does not change the domain:

$$V \,\Box\!\!\rightarrow \exists x Ax \equiv \exists x(V \,\Box\!\!\rightarrow Ax)$$

Thus, the uniqueness assumption predicts that it is inconsistent to assert the narrow scope reading while denying the wide-scope reading.

- So, the uniqueness assumption has very attractive repercussions.
- However, on the face of it, the assumption seems grossly implausible. Consider again:

(14) If Verdi and Bizet were compatriots, . . .

- Would they be Italian or French?
- It seems that there is a perfect tie between worlds of both kinds. How can one kind be closer than the other?

- Put another way, the uniqueness assumption implies the validity of CEM:

$$(\varphi \mathrel{\Box\!\!\rightarrow} \psi) \vee (\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$$

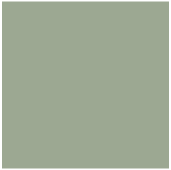
- As an instance of CEM consider (15).

(15) Either if V. and B. were compatriots, they would be Italian, or if they were compatriots, they would not be Italian.

- How can this be true, given that neither disjuncts seems true?

- Stalnaker's response to this criticism: this is a case of vagueness.
- The above instance of CEM is similar to the following instance of plain EM, where the relevant patch of color is in between

(16) Either this patch is green, or it is not green.



- We need to resort to a general modeling of vagueness.

Supervaluations (van Fraassen 66)

A valuation v is the semantic determinant relative to which a sentence receives a definite truth-value.

- E.g., a valuation for the predicate green is a set of objects, the set of objects which are green under a fixed demarcation.

A supervaluation S is a set of valuations.

- E.g., a super-valuation for green is the set of all possible extensions of the predicate under different demarcations.
- Our patch will count as green under some valuations in S, and as not green under others.

We can then define a partial notion of truth relative to a super-valuation S:

$$[\![\varphi]\!]^S = \begin{cases} \text{true} & \text{if } [\![\varphi]\!]^v = \text{true for all } v \in S \\ \text{false} & \text{if } [\![\varphi]\!]^v = \text{false for all } v \in S \\ \text{undefined} & \text{otherwise} \end{cases}$$

In the above context, both (17-a) and (17-b) would be undefined, since each is true under some but not all admissible valuations.

(17) a. This patch is green.
b. This patch is not green.

However, the corresponding instance of excluded middle (18) is true, because it is true under any admissible demarcation.

(18) Either this patch is green or it is not green.

More generally, the move from valuations to super-valuations does not change the logic of a theory:

- suppose φ is true relative to all valuations;
- given a super-valuation S, φ will be true relative to each $v \in S$;
- so φ will be true at S;
- it follows that φ is true relative to all super-valuations.

Back to conditionals:

- a valuation is a world selection function;
- a super-valuation is a set of such functions.

In the Bizet-Verdi case:

- S will contain functions that select worlds where V. and B. are Italian, as well as functions that select worlds where they are French;
- S will not contain, e.g., functions that make them both Brazilian.

This predicts that, w.r.t. S, (19-a) and (19-b) are both undefined in truth-value, even though (20) is true.

(19) a. If V. and B. were compatriots, they would be Italian.
b. If V. and B. were compatriots, they would not be Italian.

(20) a. If V. and B. were compatriots, they would be Italian, or if they were compatriots, they would not be Italian.

Counterfactual under supervaluations

- Suppose that we have several equally good candidates $w_1, \ldots, w_n$ for the closest φ-world to w.
- Then there will be vagueness as to which of these worlds is selected: S contains selection functions resolving the vagueness in various ways.
- For $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ to be true at S, it must be true under each resolution.
- Thus, $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ must be true whichever of $w_1, \ldots, w_n$ is selected for φ.
- This requires that ψ be true at each of $w_1, \ldots, w_n$.
- Thus for cases with ties in similarity, the supervaluationist view gives us the same truth-conditions as Lewis.

Counterfactual under supervaluations

- However, there is an important difference when it comes to falsity.
- $\varphi \mathbin{\Box\!\!\rightarrow} \psi$ is false at S iff it false under each resolution of vagueness.
- This means that ψ must be false at each of $w_1, \ldots, w_n$.
- Thus, falsity is more demanding than for Lewis.
  This guarantees that $\neg(\varphi \rightarrow \psi) \equiv \varphi \rightarrow \neg\psi$
- If ψ is true at some worlds and false at others,
  then $\varphi \mathbin{\Box\!\!\rightarrow} \psi$ is not false, but undefined in truth-value.
- Thus, we get different predictions for counterfactuals like (21) and (22).

(21) a. If V. and B. were compatriots, they would be Italian.
     b. If V. and B. were compatriots, they would not be Italian.

(22) a. If I had thrown this coin, it would have landed heads.
     b. If I had thrown this coin, it would have landed tails.

(23) a. If V. and B. were compatriots, they would be Italian.
b. If V. and B. were compatriots, they would not be Italian.

On Lewis and Pollock's analysis, these counterfactuals are false. On the analysis that I am defending, both are indeterminate — neither true nor false. It seems to me that the latter conclusion is clearly the more natural one. I think most speakers would be as hesitant to deny as to affirm either of the conditionals, and it seems as clear that one cannot deny them both as it is clear that one cannot affirm them both. *(Stalnaker 80)*

There is also a way to get Stalnaker's truth and falsity conditions without the uniqueness assumption, and without supervaluations.

Homogeneity presupposition (von Fintel 99)

$$[\![\varphi \mathrel{\Box\!\!\rightarrow} \psi]\!]_w = \begin{cases} \text{true} & \text{if } \psi \text{ is true at all the worlds in } \mathrm{min}_w(\varphi) \\ \text{false} & \text{if } \psi \text{ is false at all the worlds in } \mathrm{min}_w(\varphi) \\ \text{undefined} & \text{otherwise} \end{cases}$$

- This is more parsimonious than Stalnaker's implementation.
- Together with the natural treatment of negation, this yields negation as opposite:

$$\neg(\varphi \mathrel{\Box\!\!\rightarrow} \psi) \equiv \varphi \mathrel{\Box\!\!\rightarrow} \neg\psi$$

- But it does not seem to explain the lack of a scope ambiguity.