Università degli Studi di Milano

Minimal change conditionals

Stalnaker and Lewis

Sandro Zucchi

2022-23

Too strong a condition

- According to the view that natural language conditionals are strict conditionals, a sentence of the form ⌜If A, then B⌝ is true if and only if the consequent B is true in all the possible worlds in which the antecedent A is true.
- However, as we saw, conditional (1) is clearly true (since in 1778 Kant had not finished writing the *Critique of Pure Reason*), although there are worlds in which the antecedent is true and the consequent is false (any world in which Kant had finished writing the *Critique of Pure Reason* by 1778):

  (1) If Kant had died in 1778, the *Critique of Pure Reason* would have remained unfinished.

- Thus, the view that natural language conditionals are strict conditionals predicts that many conditionals that are clearly true are, in fact, false. That's a problem.

Too weak a condition

- Notice: we cannot avoid the problem by requiring that, for a conditional to be true, the consequent be true in at least one possible world in which the antecedent is true.
- Clearly, this condition is too weak: it predicts that many conditionals that are clearly false are, in fact, true.
- For example, it predicts that (2) is true (since there is at least one world in which Marlene Dietrich became a nun and Kant failed to complete the *Critique of Pure Reason*):

  (2) If Marlene Dietrich had become a nun, the *Critique of Pure Reason* would have remained unfinished.

Stalnaker's idea

- Stalnaker (1968) suggests to solve the problem in this way (as a first approximation):

  Consider a possible world in which A is true, and which otherwise differs minimally from the actual world. *"If A, then B" is true (false) just in case B is true (false) in that possible world.*

- Before we see how Stalnaker makes this intuitive idea more precise, let's see how it works vis-à-vis the previous examples.

Clearly true conditionals

- Conditional (1) is clearly true:

(1) If Kant had died in 1778, the *Critique of Pure Reason* would have remained unfinished.

- Stalnaker's theory accounts for the truth of (1) in this way:
  - other things being equal, a world in which Kant dies in 1778 and the *Critique of Pure Reason* remains unfinished differs less from the actual world than a world in which Kant dies in 1778 and the *Critique of Pure Reason* is completed.
  - Thus, the consequent of (1) is true in the possible world in which the antecedent is true that differs minimally from the actual world.
  - Thus, (1) is true.

Clearly false conditionals

- Conditional (2) is clearly false:

(2) If Marlene Dietrich had become a nun, the *Critique of Pure Reason* would have remained unfinished.

- Stalnaker's theory accounts for the truth of (1) in this way:
  - other things being equal, a world in which Marlene Dietrich becomes a nun and the *Critique of Pure Reason* is completed differs less from the actual world than a world in which Marlene Dietrich becomes a nun and the *Critique of Pure Reason* remains unfinished.
  - Thus, the consequent of (2) is false in the possible world in which the antecedent is true that differs minimally from the actual world.
  - Thus, (1) is false.

The language CS

the symbols

- Now that we have an intuitive idea of how Stalnaker's theory works, we give a more precise formulation (based on Stalnaker 1968 and Stalnaker & Thomason 1970).
- Let's define a language CS, obtained by enriching LS5 with a new connective to represent natural language conditionals: "$>$".

The language CS

the well-formed formulae

- The well-formed formulae of CS are the well-formed formulae of LS5 plus expressions of this form: $\ulcorner(\varphi > \psi)\urcorner$ (where φ and ψ are well-formed formulae of CS).

Selection functions

- Before we define what a model of CS is, let's introduce the notion *selection function*.
- One way of stating the solution proposed by Stalnaker is this: saying that a conditional is true amounts to saying that the consequent is true in a *selected world*, namely the possible world in which the antecedent is true which differs minimally from the actual world.
- Informally, a *selection function* is a function f which, for every formula φ and world w, yields the possible world in which φ is true which differs minimally from w.

The language CS

models

A model for CS is a quadruple <W, R, f, ν>, where

1. W is a non empty set of possible worlds,
2. R is an accessibility relation on W which is universal.
3. f is a *selection function*, which, for every formula φ and possible world w, assigns to $< \varphi, w >$ the possible world w' in which φ is true which differs minimally from w. If there is no possible world in which φ is true, $f(\varphi, w)$ is undefined.
4. ν is a valuation function which assigns a truth value to the well-formed formulae of CS at a world w in the following way:
   (a) for the propositional letters and formulae $\ulcorner\sim\varphi\urcorner$, $\ulcorner\varphi\wedge\psi\urcorner$, $\ulcorner\varphi\vee\psi\urcorner$, $\ulcorner\varphi\supset\psi\urcorner$, $\ulcorner\varphi\equiv\psi\urcorner$, $\ulcorner\Box\varphi\urcorner$, e $\ulcorner\Diamond\varphi\urcorner$, ν is the same as for LS5;
   (b) if there is no world at which φ is true, then $\nu(\varphi > \psi, w) = 1$;
   (c) if there is a world at which φ is true, then: $\nu(\varphi > \psi, w) = 1$ if $\nu(\psi, f(\varphi, w)) = 1$, otherwise $\nu(\varphi > \psi, w) = 0$;

How to read the truth conditions for conditionals

- Condition (c) in the definition of model is read as (c'):
  (c) if there is a world at which φ is true, then: $\nu(\varphi > \psi, w) = 1$ if $\nu(\psi, f(\varphi, w)) = 1$, otherwise $\nu(\varphi > \psi, w) = 0$;
  (c') if there is a world at which φ is true, then $\ulcorner\varphi > \psi\urcorner$ is true at the world w if and only if the consequent ψ is true at the world that the selection function f assigns to the antecedent φ and the world w.

  Given the way we described the selection function f, clause (c') amounts to saying that $\ulcorner\varphi > \psi\urcorner$ is true at a possible world w (when φ is true at some world) if and only if the consequent ψ is true in the world in which the antecedent φ is true that differs minimally from w.

Formal constraints the selection function

- In the definition of model for CS, we described the selection function informally by saying that it assign to every world w and formula φ the world in which φ is true which differs minimally from w.
- According to Stalnaker, what counts as a world in which the antecedent is true that differs minimally from a given world depends to some extent on the context in which the conditional is uttered.
- Yet, it is possible to articulate some formal constraints on the selection function. For example, whenever $f(\varphi, w)$ is defined, we may require the following:
  1. $\nu(\varphi, f(\varphi, w)) = 1$;
  2. if $\nu(\varphi, w) = 1$, then $f(\varphi, w) = w$.
- Let's examine these constraints.

Formal constraints the selection function

truth of the antecedent in the selected world

- Constraint 1 says that the selection function f, applied to a formula φ and a world w, must select a world in which φ is true (when f is defined):
  1. $\nu(\varphi, f(\varphi, w)) = 1$.
- Clearly, this constraint follows from the intuitive remark by which the selection function assign to every formula φ and world w the world *at which φ is true* which differs minimally from w.

Formal constraints the selection function

true antecedents

- Constraint 2 says that, if φ is true at a world w, the selection function f, applied to φ and w, must select w (when f is defined):
  2. if $\nu(\varphi, w) = 1$, then $f(\varphi, w) = w$.
- Again, this constraint follows from the intuitive remark by which the selection function assign to every formula φ and world w the world at which φ is true *which differs minimally from w*.
- Indeed, as Stalnaker observes, whatever criterion we assume for minimal difference, there is clearly no possible world that differs less from a world w than w itself.

Validity in CS

- The notions *valid argument* and *valid formula* in CS are defined thus:
  - $\{\varphi_1, \ldots, \varphi_n\} \models_{CS} \psi$ if and only if there is no model <W, R, f, ν> of CS and world $w \in$ W such that $\nu(\varphi_1, w) = 1, \ldots, \nu(\varphi_n, w) = 1$ and $\nu(\psi, w) = 0$.
  - $\models_{CS} \varphi$ if and only if there is no model <W, R, f, ν> of CS and world $w \in$ W such that $\nu(\varphi$, w)=0.

Some consequences

- Let's now consider some consequences of Stalnaker's semantics for conditionals.

Impossible antecedents

- Clause (b) in the definition of valuation says that, if the antecedent of a conditional is impossible (namely false at all possible worlds), the conditional is true:
  (b) if there is no world at which φ is true, then $v(\varphi > \psi, w) = 1$;
- *Prima facie*, this assumption is problematic if we want to use ">" to represent natural language conditionals (see our discussion of strict conditionals).
- We'll come back to this issue later on.

The uniqueness assumption

- For every model of CS, the following is the case:
  Uniqueness: for every world w and formula φ, there is at most one world w' in which φ is true which differs minimally from w.
- This consequence depends on the fact that f is a *function* from sentence-world pairs to worlds: this means that f cannot yield more than one value for each world-sentence pair.

The limit assumption

- Moreover, for every model of CS, the following is the case:
  Limit: for every world w and formula φ (which is true at some world), there is at least one world w' in which φ is true which differs minimally from w.
- This consequence depends on the fact that f is defined for every pair consisting of a formula (which true at some world) and a world.

Stalnaker's conditionals and natural language conditionals

- Now that we have seen how the semantics of $\ulcorner\varphi > \psi\urcorner$ works, let's see how the view that natural language conditionals are of the form $\ulcorner\varphi > \psi\urcorner$ fares.
- In particular, let's try to figure out to what extent this theory is an improvement over other theories of conditionals that we have examined.

Indicative conditionals and counterfactual conditionals

- Stalnaker's view is that indicative conditionals and counterfactual conditionals have *the same core semantics*, namely for Stalnaker both indicative and counterfactual conditionals are represented by means of the connective ">".
- But then how can we explain the fact that (3) is clearly true while (4) isn't?

(3) If Oswald did not shoot Kennedy, someone else did.

(4) If Oswald had not shot Kennedy, someone else would have.

- To see what Stalnaker's answer is, we have to introduce some notions first.

Common ground

In any conversation, there is a common ground of propositions that are accepted within that conversation. (The participants may not actually believe these propositions, since one can accept a proposition, in the framework of a conversation, without believing it.)

The common ground is common in the sense that there is common knowledge about what is mutually accepted. If I am in doubt about whether you accept p, then p is not part of the common ground, even if in fact we all do accept it. Indeed, even if we all accept p, and we all know that the others accept p, p will fail to be in the common ground if we suspect that the others might not know that we accept p.

(J. MacFarlane 2021, *Philosophical Logic*).

Context set

- The *common ground* of a conversation is the set of propositions accepted by the participants in the conversations (with the qualification spelled out in the above passage).
- We may represent the common ground of a conversation as a set of possible worlds, the possible worlds in which all the propositions accepted by the participants in the conversations are true. Stalnaker calls this set of worlds *context set*.
- Thus, a proposition is *accepted* in a conversation if and only if it is true in all the worlds of the context set.
- Intuitively, we may think of the *context set* of a conversation as the set of possibilities that are *open* for the participants in the conversation, namely the possibilities that the participants in the conversation are not in a position to exclude, on the basis of the propositions they accept.
- The worlds of the context set are worlds that might be the actual world, given the propositions that the participants in the conversation accept.

A further constraint on the selection function

- Now, Stalnaker suggests that in the case of indicative conditionals the selection function is subject to a further constraint:

  3. If w is a world of the context set and φ is the antecedent of an indicative conditional, $f(\varphi, w)$ must be a world of the context set.

- In other words, if it is applied to the antecedent of an indicative conditional and a world of the context set, the selection function must select a world in which the propositions accepted by the participants in the conversation are true.
- For subjunctive conditionals, this constraint does not hold: the selection function may select a world that does not belong to the *context set*, namely a world in which (some of) the propositions accepted by the participants in the conversation are not true.

A case of acceptance

- Let's make an example to see how the constraint on the selection function for indicative conditionals works.
- Suppose that the participants in the conversation accept that (a) either the butler is the killer or the gamekeeper is the killer (but they don't know which of the two is the killer). Suppose, moreover, that they accept that (b) the killer, whoever he is, is left-handed.
- Given that the participants in the conversation accept (a) and (b), they will accept (5):

(5) If the butler did it, he is left-handed.

- This is expected by Stalnaker's constraint on the selection function for indicative conditionals.
- Let's see why.

The constraint at work

- Given that the participants in the conversation accept that (a) either the butler is the killer or the gamekeeper is the killer, the context set will contain both worlds in which the killer is the butler and worlds in which the killer is the gamekeeper.
- However, given that they accept that (b) the killer is left-handed, it follows that in every world of the context set the killer is left-handed.
- Now, let w be a world of the context set. By the constraint on indicative conditionals, the selection function, applied to the antecedent of (5) and w, must select the world of the context set in which the butler did it which differs minimally from w:

(5) If the butler did it, he is left-handed.

- But in the worlds of the context set the killer is left-handed. Thus, the butler is left-handed in the world of the context set in which the butler did it which differs minimally from w. Thus, (5) is true in w.
- Since w was an arbitrarily chosen world of the context set, it follows that (5) is true in every world of the context set, namely (5) is accepted in the conversation.

Indicative conditionals with antecedents accepted as false

- Stalnaker's constraint on the selection function for indicative conditionals correctly predicts that indicative conditionals, unlike subjunctive conditionals, should be anomalous when the participants in the conversation accept that the antecedent is false.
- For example, suppose that we all know that I have a car (and know that we all know...etc.). In this case, I may felicitously assert (7), but not (6) (the example is by MacFarlane):

(6) If I don't have a car, I'll take the bus.

(7) if I didn't have a car, I'd take the bus.

- Stalnaker constraint on indicative conditionals leads us to expect this contrast.
- Indeed, by this constraint, the selection function applied to the antecedent of (6) and a world of the context set must select a world of the context set in which I don't have a car. However, the context set contains no such world, since we all accept that I have a car. Thus, (6) is infelicitous.
- On the other hand, no such problem arises for (7), since in this case the selection function may select a world outside the context set.

Back to Adams' problem

- Now let's go back to the contrast observed by Adams (1970):

(3) If Oswald did not shoot Kennedy, someone else did.

(4) If Oswald had not shot Kennedy, someone else would have.

- We accept indicative conditional (3), but not subjunctive conditional (4). Here's Stalnaker's explanation:
  - It is common knowledge that someone shot Kennedy. Thus, in the worlds of the context set someone shot Kennedy.
  - Since (3) is an indicative conditional, the selection function, applied to the antecedent and a world of the context set, must select a world of the context set, namely a world in which Oswald did not shoot Kennedy, but someone did. Thus, (3) is true in any world of the context set (since the consequent of (3) is true in any world picked out by the selection function). Thus, we accept (3).
  - In the case of (4), the selection function, applied to the antecedent and a world of the context set, may select a world which is not in the context set, for example a world in which nobody shot Kennedy. Thus, (4) may turn out to be false in some worlds of the context set (since the consequent of (4) may not be true in a world picked out by the selection function). Thus, we don't accept (4).

Comparisons

- Let's now see how Stalnaker's theory fares with respect to some of the problems other theories of conditionals run into.

Transitivity

- As we saw, the thesis that indicative conditionals are material conditionals incorrectly predicts that argument (8) is valid, since "$\supset$" is transitive:

(8) If Trump will not run for the 2024 elections, he will flee the country. If Trump is in jail, he will not run for the 2024 elections. Thus, if Trump is in jail, he will flee the country.

- The same problem arises for the view that natural language conditionals are strict conditionals, since "$⥽$" is transitive.

The connective ">" is not transitive

- The problem of transitivity does not arise for Stalnaker's theory.
- Indeed, the connective ">" is not transitive, namely:

(9) $p > q, q > r \nvDash_{CS} p > r$

A counter-model

- We can show (9) by observing that any model of CS that meets conditions 1-10 makes the premises in (9) true at $w0$ and the conclusion false at $w0$:

(9) $p > q, q > r \nvDash_{CS} p > r$

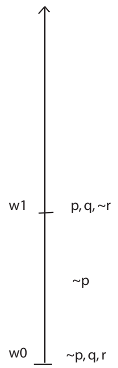
1. $W = \{w0, w1\}$
2. $w0Rw0$, $w0Rw1$, $w1Rw1$, $w1Rw0$
3. $f(p, w0) = w1$
4. $f(q, w0) = w0$
5. $v(p, w0) = 0$
6. $v(q, w0) = 1$
7. $v(r, w0) = 1$
8. $v(p, w1) = 1$
9. $v(q, w1) = 1$
10. $v(r, w1) = 0$

- In a model of this kind, "$p > q$" is true at $w0$, since $f(p, w0) = w1$ and "q" is true at $w1$. Moreover, $f(q, w0) = w0$ and "r" is true at $w0$, thus "$q > r$" is true at $w0$. However "$p > r$" is false at $w0$, since $f(p, w0) = w1$ and "r" is false at $w1$.

A visual representation of the counter-model

The arrow indicates the degree of similarity to the base world $w0$: the further away a world is from $w0$, the less similar the world is to $w0$.

(9) $p > q, q > r \nvDash_{CS} p > r$



Contraposition

- As we saw, the thesis that indicative conditionals are material conditionals incorrectly predicts that argument (10) is valid:

  (10) If it rains, it is not the case that will rain a lot. Therefore, if it rains a lot, it is not the case that it will rain.

- The same problem arises for the view that natural language conditionals are strict conditionals, since, as we have seen, contraposition is valid for "⥽".

Invalidity of contraposition

- The problem posed by contraposition does not arise in Stalnaker's theory of conditionals.
- Indeed, contraposition is invalid for ">":

  (11) $p > q \nvDash_{CS} \sim q > \sim p$

A counter-model

- We can show (11) by observing that any model of CS that meets conditions 1-8 makes the premises in (11) true at $w0$ and the conclusion false at $w0$:

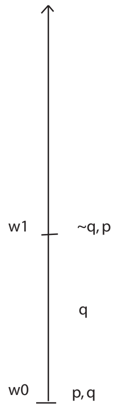
  (11) $p > q \nvDash_{CS} \sim q > \sim p$

  1. $W = \{w0, w1\}$
  2. $w0Rw0$, $w0Rw1$, $w1Rw1$, $w1Rw0$
  3. $f(p, w0) = w0$
  4. $f(\sim q, w0) = w1$
  5. $v(p, w0) = 1$
  6. $v(q, w0) = 1$
  7. $v(p, w1) = 1$
  8. $v(q, w1) = 0$

- In a model of this kind,, "$p > q$" is true at $w0$, since $f(p, w0) = w0$ and "q" is true at $w0$. However, "$\sim q > \sim p$" is false at $w0$, since $f(\sim q, w0) = w1$ and "$\sim p$" is false at $w1$.

A visual representation of the counter-model

(11) $p > q \not\models_{CS} \sim q > \sim p$



Strengthening of the antecedent

- The thesis that indicative conditionals are material conditionals incorrectly predicts that argument (12) is valid:

(12) If Holmes accepted the case, the case will be solved. Thus, if Holmes accepted the case and gave it up right after, the case will be solved.

- The same problem arises for the view that natural language conditionals are strict conditionals, since, as we have seen, strengthening of the antecedent is valid for "$\dashv$".

Invalidity of strengthening of the antecedent

- The problem posed by strengthening of the antecedent does not arise for Stalnaker's theory of conditionals.
- Indeed, strengthening of the antecedent is invalid in CS:

(13) $p > q \not\models_{CS} (p \wedge r) > q$

A counter-model

- We can show (13) by observing that any model of CS that meets conditions 1-10 makes the premises in (13) true at $w0$ and the conclusion false at $w0$:

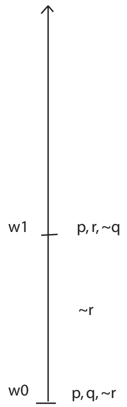
(13) $p > q \not\models_{CS} (p \wedge r) > q$

1. $W = \{w0, w1\}$
2. $w0Rw0$, $w0Rw1$, $w1Rw1$, $w1Rw0$
3. $f(p, w0) = w0$
4. $f(p \wedge r, w0) = w1$
5. $v(p, w0) = 1$
6. $v(q, w0) = 1$
7. $v(r, w0) = 0$
8. $v(p, w1) = 1$
9. $v(q, w1) = 0$
10. $v(r, w1) = 1$

- In a model of this type, "$p > q$" is true at $w0$, since $f(p, w0) = w0$ and "q" is true at $w0$. But "$(p \wedge r) > q$" is false at $w0$, since $f(p \wedge r, w0) = w1$ and "q" is false at $w1$.

A visual representation of the counter-model

(13) $p > q \not\models_{CS} (p \wedge r) > q$



Disjunction

- The thesis that indicative conditionals are material conditionals predicts that sentence (14) is valid:

(14) If I am right, you are right or if you are right I am right.

- The problem is that disjunction (14) is false if you and I hold incompatible views.

Invalidity of "$(p > q) \vee (q > p)$"

- The problem posed by disjunction does not arise for Stalnaker's theory of conditionals.
- Indeed, this formula is not valid in CS:

(15) $\not\models_{CS} (p > q) \vee (q > p)$

A counter-model

- We can show (15) by observing that any model of CS that meets conditions 1-8 makes the formula in (15) false at $w0$:

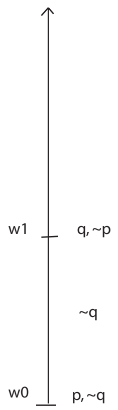
(15) $\not\models_{CS} (p > q) \vee (q > p)$

1. $W = \{w0, w1, \lambda\}$
2. $w0Rw0$, $w0Rw1$, $w1Rw1$, $w1Rw0$
3. $f(p, w0) = w0$
4. $f(q, w0) = w1$
5. $v(p, w0) = 1$
6. $v(q, w0) = 0$
7. $v(p, w1) = 0$
8. $v(q, w1) = 1$

- In a model of this kind, "$p > q$" is false at $w0$, since $f(p, w0) = w0$ e "q" is false at $w0$. Moreover, $f(q, w0) = w1$ and "p" is falsa at $w1$, thus "$q > p$" is false at $w0$. Thus, "$(p > q) \vee (q > p)$" is false at $w0$ since both disjuncts are false at $w0$.

A visual representation of the counter-model

(15) $\not\models_{CS}\ (p > q) \vee (q > p)$



Falsity of the antecedent

- The thesis that indicative conditionals are material conditionals predicts that the falsity of the antecedent should be sufficient to make the conditional true.
- Thus, if indicative conditionals are material conditionals, sentence (16) is true:

(16) If New York is in New Zealand, Rome is in France.

- *Prima facie*, this prediction is incorrect.

Falsity of the antecedent in Stalnaker's theory

- In Stalnaker's theory, the falsity of the antecedent is not sufficient to make the conditional true:

(17) $\sim p \not\models_{CS}\ p > q$

A counter-model

- We can show (17) by observing that any model of CS that meets conditions 1-6 makes the premises in (17) true at $w0$ and the conclusion false at $w0$:

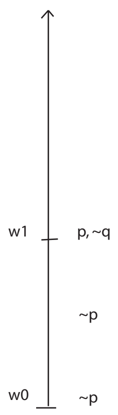
(17) $\sim p \not\models_{CS}\ p > q$

1. $W = \{w0, w1\}$
2. $w0Rw0$, $w0Rw1$, $w1Rw1$, $w1Rw0$
3. $f(p, w0) = w1$
4. $v(p, w0) = 0$
5. $v(p, w1) = 1$
6. $v(q, w1) = 0$

- In a model di of this kind, "$\sim p$" is true at $w0$. However, "$p > q$" is false at $w0$, since $f(p, w0) = w1$ and "q" is false at $w1$.

A visual representation of the counter-model

(17) $\sim p \nvDash_{CS} \; p > q$



Truth of the consequent

- The thesis that indicative conditionals are material conditionals predicts that the truth of the consequent should be sufficient to make the conditional true.
- Thus, if indicative conditionals are material conditionals, sentence (18) is true:

(18) If New York is in New Zealand, Rome is in Italy.

- *Prima facie*, this prediction is incorrect.

Truth of the consequent in Stalnaker's theory

- In Stalnaker's theory, the truth of the consequent is not sufficient to make the conditional true:

(19) $q \nvDash_{CS} \; p > q$

A counter-model

- We can show (19) by observing that any model of CS that meets conditions 1-7 makes the premises in (19) true at $w0$ and the conclusion false at $w0$:

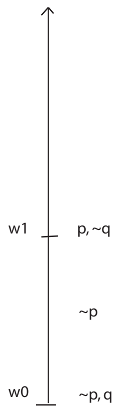
(19) $q \nvDash_{CS} \; p > q$

1. $W = \{w0, w1, \lambda\}$
2. $w0Rw0$, $w0Rw1$, $w1Rw1$, $w1Rw0$
3. $f(p, w0) = w1$
4. $v(p, w0) = 0$
5. $v(q, w0) = 1$
6. $v(p, w1) = 1$
7. $v(q, w1) = 0$

- In a model of this kind, "$\sim q$" is true at $w0$. However, "$(p > q) \vee (q > p)$" is false at $w0$, since $f(p, w0) = w1$ and "$\sim q$" is false at $w1$.

A visual representation of the counter-model

(19) $q \not\models_{CS} \; p > q$



True antecedents and consequents

- Stalnaker's theory predicts, however, that a conditional is true if both antecedent and consequent are true.
- We may show that this is a prediction of the theory with the following reasoning:
  - if the antecedent φ is true in the real world, then, by constraint 2 on the selection function, the selection function applied to φ and the real world selects the real world.

    2. if $v(\varphi, w) = 1$, then $f(\varphi, w) = w$.
  - Thus, if the consequent ψ is true in the real world, it follows by the semantics of $>$ that $\ulcorner\varphi > \psi\urcorner$ is true in the real world.

A problem

- As Read (1995) points out, one consequence of the above feature of Stalnaker's theory is that, if John is in Alaska, conditional (20) is true:

  (20) If John is not in Turkey, he is not in Paris.
- The problem is that, even if John is in Alaska, (20) doesn't seem to be true.

A pragmatic explanation

- Stalnaker does not discuss the problem raised for his theory by (20):

  (20) If John is not in Turkey, he is not in Paris.
- It seems that appeal to pragmatic considerations is needed here to explain why we should not assert sentences like (20).
- Here is a try. If the speaker knows that John is in Alaska he should say so, instead of making the weaker statement in (20). On the other hand, if the speaker does not know that John is in Alaska (or doesn't know that John is neither in Turkey not in Paris), it is not clear on what basis she could assert (20).
- Does this explanation hold water? (Exercise for the reader).

Impossible antecedents

- We saw that by clause (b) of the semantics of >, a conditional $\ulcorner\varphi > \psi\urcorner$ is true if φ is impossible (false in all worlds).
- Thus, Stalnaker's theory predicts that any natural language conditional with an impossible antecedent is true.

Counterfactuals and impossible antecedents

D. Lewis

Lewis (1972) difende così la scelta di rendere veri i controfattuali con antecedenti impossibili:

> *There is at least some intuitive justification for the decision to make a 'would' counterfactual with an impossible antecedent come out vacuously true. Confronted by an antecedent that is not really an entertainable supposition, one may react by saying, with a shrug: If that were so, anything you like would be true! . . .*
>
> *Moreover, one sometimes asserts counterfactuals by way of reductio in philosophy, mathematics, and even logic. . . . These counterfactuals are asserted in argument, and must therefore be thought true; but their antecedents deny what are thought to be philosophical, mathematical, or even logical truths, and must therefore be thought not only false but impossible. These asserted counterphilosophicals, countermathematicals, and counterlogicals look like examples of vacuously true counterfactuals.*

Counterfactuals and impossible antecedents

D. Lewis -cont.

> *Of course there are some we would assert and some we would not:*
>
> *If there were a largest prime p, p!+1 would be prime.*
>
> *If there were a largest prime p, p!+1 would be composite.*
>
> *are both sensible things to say, but*
>
> *If there were a largest prime p, there would be six regular solids.*
>
> *If there were a largest prime p, pigs would have wings.*
>
> *are not. But what does that prove? . . . We have plenty of cases in which we do not want to assert counterfactuals with impossible antecedents, but so far as I know we do not want to assert their negations either. Therefore they do not have to be made false by a correct account of truth conditions; they can be truths which (for good conversational reasons) it would always be pointless to assert.*
>
> D. Lewis *Counterfactuals*, pp. 24-25.

Lewis's argument

- In the above passage, Lewis claims that, when a conditional with an impossible antecedent doesn't seem true to us, this depends on the fact that the conditional is not assertable, not on the fact that it's false.
- Lewis's argument is this:

  If conditional (21) were false, then its negation should be assertable. But the negation of (21) is not assertable (it doesn't seem reasonable to assert (21) nor to assert its negation). Thus, (21) is not false.

  (21) If there were a largest prime p, pigs would have wings.

- But is Lewis right in claiming that the negation of (21) is not assertable? In any case, Priest would be willing to assert (22), although the antecedent is impossible:

(22) It is false that, if you had squared the circle, I would have given you my life's savings.

Necessary consequents

- For any world w, a conditional of the form $\ulcorner\varphi > \psi\urcorner$ is true at w, if the consequent ψ is necessarily true at w.
- Indeed, if ψ is true in every possible world, it is also true in the world singled out by the selection function at the world $f(\varphi, w)$. Thus, if ψ is true in every possible world, $\ulcorner\varphi > \psi\urcorner$ is true at w (for any w).
- Thus, the problem raised by necessary consequents also arises for Stalnaker's theory. The theory predicts that (23) is true:

(23) If Rome is in Italy, there is an infinity of natural numbers.

- Presumably, Stalnaker must claim that (23) is true but not assertable: since it is true in all worlds that there is an infinity of natural numbers, the speaker should assert (24), rather than the longer (23):

(24) There is an infinity of natural numbers.

The distribution of labor between semantic and pragmatics

- Let's pause briefly to reflect on the difference between Stalnaker's theory and the other theories of conditionals we have seen.
- Stalnaker's semantics for conditionals, unlike the material analysis and the strict analysis, predicts that contraposition, strengthening of the antecedent, and transitivity are all invalid argument forms.
- Yet, to deal with conditionals with true antecedents and consequents, conditionals with impossible antecedents, and conditionals with necessary consequents, Stalnaker must presumably resort the to the distinction between truth and assertability, as other theories of conditionals do to deal with problematic cases.
- Thus, Stalnaker's theory and the other theories we have seen propose *different ways of distributing labor* between semantics and pragmatics.

Stalnaker and Lewis

- We are now going see some problems for Stalnaker's theory raised by David Lewis (1973).
- We present the theory of counterfactuals proposed by David Lewis, which avoids these problems.
- Like Stalnaker's semantics, Lewis's semantics of counterfactuals is based on the idea that possible worlds are ordered with respect to how similar they are to (how much they differ from) a given world.
- Finally, we are going to see how Stalnaker suggests to deal with the problems raised by Lewis.

The problem with Uniqueness

- Consider these examples proposed by Quine (1959):

(25) If Bizet and Verdi had been compatriots, Bizet would have been Italian.

(26) If Bizet and Verdi had been compatriots, Bizet would have been French.

- Which of these conditionals is true? One plausible answer is that neither of them is true. If Bizet and Verdi had been compatriots, Bizet might have been Italian or French. However, it does not seem plausible to assert that he would have been Italian and it does not seem plausible to assert that he would have been French.
- The problem is that in Stalnaker's theory, as we saw, the selection function associates one and only one world to each antecedent-world pair. In the selected world, either Bizet is Italian or he is French.
- Thus, by Stalnaker's theory, one of the conditionals in (25)-(26) is true, contrary to intuition.

Getting closer and closer

- Now, let's examine the following case:

  In the real world Leo is 180 cm tall. Consider the counterfactual worlds in which Leo is taller than 180 cm, and suppose we want to order them according to their degree of similarity to the real world (according to how little they differ to the real world). Other things being equal, the following is true: a world in which Leo is 185 cm tall differs less from the real world than a world in which Leo is 190 cm tall, a world in which Leo is 183 cm tall differs less from the real world than a world in which Leo is 185 cm tall, a world in which Leo is 180.5 cm tall differs less from to the real world than a world in which Leo is 1.83 cm tall, and so on. Clearly, given any world w in which Leo is taller than than 180 cm, it is always possible to find a world that, other things being equal, differs less from the real world than w (assuming that a spatial dimension like tallness is dense).

- Ok, now we are ready to raise a further problem for Stalnaker's semantics.

The problem with Limit

- Suppose Leo would like to join a basketball team. He could, if only he were taller than he actually is.
- Now, consider sentence (27) (from Read 1995):

(27) If Leo were over 2 m tall, he would join a basketball team.

- Clearly, (27) is true.
- Yet, according to Stalnaker's theory, in order for (27) to be true, it must be the case that Leo plays in a basketball team in the world in which he is over 2 m tall that differs less than any other world from the real world.
- The problem, as we just saw, is that there is no such world.
- As Lewis observes, this example raises a problem for the Limit assumption, namely the assumption that there is a world in which the antecedent is true which differs less than any other world from the real world.

Lewis's theory

- Lewis (1973) proposes that the semantics of counterfactual conditionals be stated as follows (where $\Box\!\!\rightarrow$ is the connective Lewis uses to represent counterfactuals):
  1. if there is no accessible world at which φ is true, then $v(\varphi \Box\!\!\rightarrow \psi, w) = 1$;
  2. if there is a world at which φ is true, then: $v(\varphi \Box\!\!\rightarrow \psi, w) = 1$ if some accessibile world in which φ and ψ are true differs less from w than any world in which φ is true and ψ is false.
- We may read $\ulcorner\varphi \Box\!\!\rightarrow \psi\urcorner$ as $\ulcorner$if it had been the case that φ, it would have been the case that $\psi\urcorner$.
- (Keep in mind that Lewis regards indicative conditionals as material implications, so his analysis in terms of possible worlds in only meant to apply to counterfactuals).

An illustration

- Lewis's analysis correctly predicts that (1) is true and (2) is false:

(1) If Kant had died in 1778, the *Critique of Pure Reason* would have remained unfinished.

(2) If Marlene Dietrich had become a nun, the *Critique of Pure Reason* would have remained unfinished.

- It predicts that (1) is true, since in the real world the *Critique of Pure Reason* is not yet finished in 1778, thus some accessible world in which Kant dies in 1778 and the *Critique of Pure Reason* is unfinished differs less from the real world than any world in which Kant dies in 1778 and the *Critique of Pure Reason* is finished.
- It predicts that (2) is false, since in the real world the *Critique of Pure Reason* is finished in 1881, thus some accessible world in which Marlene Dietrich becomes a nun and the *Critique of Pure Reason* is finished differs less from the real world than any world in which Marlene Dietrich becomes a nun and the *Critique of Pure Reason* is unfinished.

Back to the problem with Uniqueness

- Moreover, Lewis's analysis predicts that both (25) and (26) are false:

  (25) If Bizet and Verdi had been compatriots, Bizet would have been Italian.

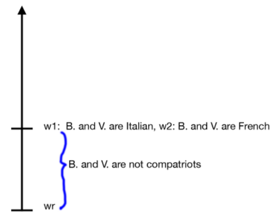
  (26) If Bizet and Verdi had been compatriots, Bizet would have been French.

- It predicts that (25) is false, since it is false that some accessible world in which Bizet and Verdi are compatriots and Bizet is Italian differs less from the real world than any world in which Bizet and Verdi are compatriots and Bizet is French.
- It predicts that (26) is false, since it is false that some accessible world in which Bizet and Verdi are compatriots and Bizet is French differs less from the real world than any world in which Bizet and Verdi are compatriots and Bizet is Italian.
- (In other words, Lewis's analysis, unlike Stalnaker's, allows the possibility of *ties*: there may be more than one world at which a conditional antecedent is true that minimally differs from the real world. The conditionals in (25)-(26) are exactly a case of this kind).

The tie

(25) If Bizet and Verdi had been compatriots, Bizet would have been Italian.

(26) If Bizet and Verdi had been compatriots, Bizet would have been French.



- In Lewis's theory, for (25) to be true, all the worlds in which the antecedent is true that differ minimally from *wr* must be worlds in which Bizet is Italian.
- For (26) to be true, all the worlds in which the antecedent is true that differ minimally from *wr* must be worlds in which Bizet is French.
- Clearly, neither is the case: the world *w*1 and *w*2 are worlds in which Bizet and Verdi are compatriots which differ minimally from *wr*, but in *w*1 Bizet is Italian and in w2 Bizet is French.

Truth of the disjunctive consequent

- Notice that Lewis's analysis, although it predicts that (25)-(26) are false, it also predicts that (28) is true:

  (25) If Bizet and Verdi had been compatriots, Bizet would have been Italian.

  (26) If Bizet and Verdi had been compatriots, Bizet would have been French.

  (28) If Bizet and Verdi had been compatriots, either Bizet would have been Italian or Verdi would have been French.

- Indeed, given that in the real world Bizet is French and Verdi is Italian, some possible world in which Bizet and Verdi are compatriots and either Bizet is Italian or Verdi is French differs less from the real world that any world in which Bizet and Verdi are compatriots and neither Bizet is Italian nor Verdi is French.
- In the previous picture, this is illustrated by the fact that the closest worlds in which Bizet and Verdi are compatriots are the worlds *w*1, in which Bizet and Verdi are Italian, and the world *w*2, in which Bizet and Verdi are French. In either world, the disjunction "Bizet is Italian or Verdi" is true.

Back to the problem with Limit

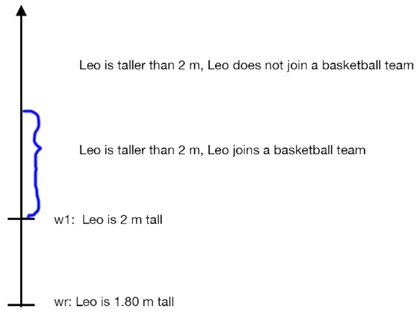
- Lewis's analysis correctly predicts (27) is true in case Leo would like to join a basketball team, but he cannot because he is only 1.80 m tall.

  (27) If Leo were over 2 m tall, he would join a basketball team.

- It predicts that (27) is true because in the real world the only reason why Leo does not join a basketball team is that he is not tall enough, thus some world in which Leo is taller than 2 m and joins a basketball team differs less from the real world than any world in which Leo is taller than 2 m and he does not join a basketball team.
- Notice: this does not require that there is a world in which Leo is 2 m tall which differs less from the real world than any other world.

Truth without Limit

(27) If Leo were over 2 m tall, he would join a basketball team.



- In the case described in the picture, (27) is predicted to be true by Lewis, although there is no world in which Leo is taller than 2 m which is the closest to the real world.

Stalnaker's reaction

- Stalnaker is not convinced by Lewis's solution to the problem with Uniqueness and the problem with Limit.
- First, let's look at the problem with uniqueness raised by (25)-(26):

(25) If Bizet and Verdi had been compatriots, Bizet would have been Italian.

(26) If Bizet and Verdi had been compatriots, Bizet would have been French.

- Since Lewis's theory makes both (25) and (26) false, it predicts that their negation is true and thus that we should be able to deny them both. Yet, as Lewis (1973) admits, (29) sounds contradictory:

(29) It is not the case that if Bizet and Verdi were compatriots, Bizet would be Italian; and it is not the case that if Bizet and Verdi were compatriots, Bizet would not be Italian; nevertheless, if Bizet and Verdi were compatriots, Bizet either would or would not be Italian.

Stalnaker's modified theory

- Stalnaker (1981, 1984) proposes a modified version of his theory based on van Fraassen's (1974) supervaluations. The modified version may be articulated thus:
  - Give up Uniqueness, namely give up the assumption that for any given world w, there is at most one world at which the antecedent is true that differs minimally from w.
  - When there is more than one world at which the antecedent is true that differs minimally from the world of evaluation w, this means that there is more than one admissible selection function. Each admissible selection function singles out a different world when applied to the antecedent and the world of evaluation w.
  - The semantics for conditionals with possible antecedents may now be stated as in 1-3:
    1. $\ulcorner\varphi > \psi\urcorner$ is true at w if $v(\psi, f(\varphi, w)) = 1$ for every admissible selection function f;
    2. $\ulcorner\varphi > \psi\urcorner$ is false at w if $v(\psi, f(\varphi, w)) = 0$ for every admissible selection function f;
    3. otherwise $\ulcorner\varphi > \psi\urcorner$ is neither true nor false.
- (It may be shown that this modified version of Stalnaker's theory validates the same inferences as Stalnaker's original theories).

Bizet and Verdi again

- It is plausible to assume that, other things being equal, a counterfactual world in which Bizet and Verdi are both Italian is as similar to the real world as a world in which Bizet and Verdi are both French.
- Thus, for (25)-(26) there is more than one admissible selection function: one which, applied to the antecedent and the real world, singles out a world in which Bizet and Verdi are both Italian and one which, applied to the antecedent and the real world, singles out a world in which Bizet and Verdi are both French:

(25) If Bizet and Verdi had been compatriots, Bizet would have been Italian.

(26) If Bizet and Verdi had been compatriots, Bizet would have been French.

- Thus, by the modified semantics for conditionals proposed by Stalnaker (25) is neither true nor false and (26) is neither true nor false. This correctly predicts that neither (25)-(26) nor their negation should be assertable.

Minimal difference and context

- Since both Stalnaker's original theory and the modified version assume Limit (there is at least one world at which the antecedent is true that differs minimally from the world of evaluation), they fail to account for the truth of (27):

  (27) If Leo were over 2 m tall, he would join a basketball team.

- Indeed, as we saw, it seems that there is no world at which Leo is over than 2 m tall that differs less than any other world from the real world.
- Stalnaker's reply is that what counts as a world that differs minimally from a given world depends, to some extent, on the context.
- For example, in the context we described for (27), any world in which Leo's height is more than 2 meters and less than 2 meters and one millimeter plausibly counts as a world at which the antecedent of (27) is true that differs minimally form the real world.
- If this is the case, (27) poses no problem for the Limit assumption.

Summing up

- We presented the semantics for conditionals which make use of the notion *minimal change*. In particular,
  - we examined the theory of conditionals proposed by Stalnaker (1968);
  - we examined the variant of the minimal change approach to counterfactuals proposed by Lewis (1973);
  - we examined a modified version of Stalnaker's theory, which discards Uniqueness.