Lewis and Stalnaker on Subjunctive Conditionals

02.13.12

William Starr

Phil 6710 / Ling 6634 – Spring 2012

Announcements:

- There is a semantics group at Cornell
  - To get involved enroll in the Blackboard site: "Language-Murray-Fall2010: Cornell Semantics Group"
  - You can also contact this year's organizer Ed Cormany (esc53@cornell.edu)
- 1st meeting: Sarah Murray "Two Kinds of Indexicals" this Friday at 3pm, Morrill 110
- Adam Bjorndahl (math) will be giving an informal workshop on topological semantics
  - Provides a more general approach to modal semantics than Kripke structures!
  - Contact him this week if you are interested in attending (afm53@cornell.edu)

1 Subjunctive Conditionals and Strict Conditionals

Our goal is a semantic analysis of subjunctive conditionals, but what are they?

1.1 Subjunctive Conditionals

- Indicative:
  (1) If Oswald didn't shoot Kennedy, someone else did
- Subjunctive: modal in consequent + past(ish) antecedent
  (2) If Oswald **had**n't shot Kennedy, someone else **would** have
  (3) If Antarctica possess**ed** nuclear weapons, no one **would** know
  (4) If I **were** to drop this pen, it **could** break
- Counterfactual:
  - Subjunctive conditional w/false 'antecedent'
    (5) If David Lewis were alive, he would live on the moon
      - Strictly speaking, *David Lewis were alive* isn't false, it's ungrammatical!

Email: will.starr@cornell.edu.
URL: http://williamstarr.net.

      - What's false: *David Lewis is alive*
- Do all subjunctive conditionals presuppose or entail that their antecedents are false? No!
  (6) If Jones had taken the arsenic, he would have just exactly those symptoms which he does in fact show (Anderson 1951: 37)
  (7) If the butler had done it, we would have found blood on the kitchen knife. The knife was clean; therefore, the butler did not do it. (Stalnaker 1975: 71)
- Basic problem for analysis in classical logic:
  - (5) is clearly false, but the corresponding material conditional is true!
    - Material conditional: $\mathsf{A} \supset \mathsf{B} := \neg\mathsf{A} \vee \mathsf{B}$
    - Material conditional analysis: *if A then B* means $\mathsf{A} \supset \mathsf{B}$
- So the classical analysis predicts that all counterfactuals are true!
- **Notation**: represent a subjunctive conditional as $\mathsf{A} > \mathsf{B}$

1.2 Strict Conditionals

Maybe modal logic can help?

Strict Conditional

- $\Box(\phi \supset \psi)$ is true in w just in case $\phi \supset \psi$ is true at all worlds accessible from w
  - Accessible according to R
  - $R(w, w')$ means: w' is possible with respect to w
- $$\begin{aligned}[\![\Box(\phi \supset \psi)]\!]^R &= \{w \mid \forall w' : R(w, w'), w' \in [\![\phi \supset \psi]\!]^R\} \\ &= \{w \mid (R(w) \cap [\![\phi]\!]^R) \subseteq [\![\psi]\!]^R\} \qquad (R(w) := \{w' \mid R(w, w')\})\end{aligned}$$

Does $\phi > \psi = \Box(\phi \supset \psi)$?

1.3 Problem 1: Antecedents that Couldn't be True

- JFK couldn't have passed universal healthcare (true)
  - $\neg\Diamond\mathsf{H}$ (equivalently: $\Box\neg\mathsf{H}$)
- But if he had, he would have granted insects coverage as well (false)
  - $\Box(\mathsf{H} \supset \mathsf{C})$
- But strict conditional is true!
  - If $\neg\Diamond\mathsf{H}$, then every accessible world is a $\neg\mathsf{H}$-world
  - Then every accessible world is a $\mathsf{H} \supset \mathsf{C}$-worlds
  - So $\Box(\mathsf{H} \supset \mathsf{C})$ is true!

- In general: $\Box\neg\phi \vDash \Box(\phi \supset \psi)$

1.4 Problem 2: Antecedent Strengthening

- If $\Box(\phi_1 \supset \psi)$ is true, then $\Box((\phi_1 \wedge \phi_2) \supset \psi)$ is true
  - Suppose $\Box(\phi_1 \supset \psi)$ is true in w
  - Then $(R(w) \cap [\![\phi_1]\!]^R) \subseteq [\![\psi]\!]^R$
  - But then $(R(w) \cap ([\![\phi_1]\!]^R \cap [\![\phi_2]\!]^R)) \subseteq [\![\psi]\!]^R$
    - Since $([\![\phi_1]\!]^R \cap [\![\phi_2]\!]^R) \subseteq [\![\phi_1]\!]^R$
  - So $(R(w) \cap ([\![\phi_1 \wedge \phi_2]\!]^R)) \subseteq [\![\psi]\!]^R$
    - Since $\wedge$ is $\cap$
  - Hence $\Box((\phi_1 \wedge \phi_2) \supset \psi)$ is true in w
- Subjunctive conditionals don't seem to go in for this inference!
- Goodman (1947) noticed that (8) seems true, but (9) false

(8) If I had struck this match, it would have lit

(9) If I had struck this match and done so in a room without oxygen, it would have lit

- Lewis (1973b) sequences:

(10) If I had shirked my duty, no harm would have ensued
$\mathsf{I} > \neg\mathsf{H}$

(11) Though, if I had shirked my duty and you had too, harm would have ensued
$(\mathsf{I} \wedge \mathsf{U}) > \mathsf{H}$

(12) Yet, if I had shirked my duty, you had shirked your duty and a third person done more than their duty, then no harm would have ensued
$(\mathsf{I} \wedge \mathsf{U} \wedge \mathsf{T}) > \neg\mathsf{H}$

⋮

  - Strict analysis says every other conditional is inconsistent w/previous one
    - Yet the Lewis sequence sounds fine!
- What went wrong?
- One diagnosis (Lewis):
  - When you evaluate $\mathsf{A} > \mathsf{C}$, you consider some alternative worlds where A holds
    - And express a *generalization* about them
  - When you evaluate $(\mathsf{A} \wedge \mathsf{B}) > \mathsf{C}$, it may be appropriate to consider entirely different A-worlds
  - But all strict conditionals express generalizations about the same space of worlds $R(w)$
    - E.g. All the A-worlds are B-worlds

- What to do?
  - Insead, we should think of subjunctives as expressing generalizations about varying sets of possibilities
    - 'Variably-strict' conditionals, as Lewis calls them

2 Goodman on the Problem of Counterfactual Conditionals

- Goodman (1947) asked:
  - $\phi > \psi$ is true just in case ψ follows from ϕ, together with some other information
    - What information, exactly?
  - An anachronistic restatement:
    - $\phi > \psi$ is true just in case $([\![\phi]\!] + I) \subseteq [\![\psi]\!]$
    - Exactly which ϕ-worlds make a given counterfactual true?
      - What operation is $+$ and what information is I?
- It is hard to answer this question:
  - First, I can't be all truths
    - Often $\neg\phi$ will be among those, and everything follows from ϕ and $\neg\phi$
  - What about all truths logically compatible with ϕ?
    - There are a lot of truths logically compatible with ϕ but not logically compatible with each other
      - Both *Bob is happy* and *Bob is not happy* are logically compatible with *It is 2pm*
  - What about all truths which are co-tenable with ϕ?
    - All truths one can consistently maintain alongside ϕ
    - Problem: logical compatibility isn't enough
      - Sandra and Lynn are on the teeter-totter. Sandra is up and Lynn is down. But if Sandra were down, Lynn would be up.
      - *Sandra is down* and *Lynn is down* are **logically** compatible, but physically incompatible
    - This is really just antecedent strengthening all over again
      - *The match was struck* and *The match was soaked in water* are logically compatible
      - So why don't we include the latter in I when we evaluate *If the match were struck it would light*?

- This makes clear that + and I need to appeal to **laws**
    - Goodman: But what are laws? True generalizations? No!
        - Humeanism: 'Causation' is unscientific!
    - Hopefully not: true subjunctive conditionals
    - Then circularity threatens
- Goodman: what makes a generalization law-like?
    - It's acceptance does not depend on any one of its instances
    - We accept *Everything in my pocket is silver* by looking at the current contents of my pocket
    - We don't accept *Ravens are black* by looking at all actual ravens
- But what exactly is dependence?
    - This idea that some truths (laws) don't depend on others (particular facts)
- Goodman: I have no idea

3 The Lewis-Stalnaker Theory

The Strategy We can say some things about how I and + work which capture the logic of subjunctive conditionals without proposing a solution to Goodman's problem. The basic idea is that $[\![\phi]\!] + I$ gives the worlds most **similar** to the actual world.

Lewis-Stalnaker Theory (Stalnaker 1968; Stalnaker & Thomason 1970; Lewis 1973a)

- $\phi > \psi$ is true at w just in case all of the ϕ-worlds most **similar** to w are ψ-worlds
    - Most similar according to the selection function f
    - f takes a proposition p and a world w and returns the p-worlds most similar to w
- $[\![\phi > \psi]\!]^f = \{w \mid f(w, [\![\phi]\!]^f) \subseteq [\![\psi]\!]^f\}$

(Momentarily making 'Limit Assumption': there are most similar worlds)

- Logic of $>$ is determined by constraints on f (where $p, q \subseteq W$ and $w \in W$):

(a)	$f(w,p) \subseteq p$	**success**
(b)	$f(w,p) = \{w\}$, if $w \in p$	**strong centering**
(c)	$f(w,p) \subseteq q$ & $f(w,q) \subseteq p \implies f(w,p) = f(w,q)$	**uniformity**
(d)	$f(w,p)$ contains *at most* one world	**uniqueness**

Stalnaker Constraints (a)-(d)

- Since Stalnaker adopts (d), he actually takes f to return a world, rather than a singleton set of worlds
- Problem: which world is $f(w, [\![\phi \wedge \neg\phi]\!])$?
    - Stalnaker: λ, the 'absurd world'
- Our solution is more elegant: $f(w, [\![\phi \wedge \neg\phi]\!]) = \varnothing$, no need for λ

Limited Lewis Constraints (a)-(c)

- 'Limited Lewis': Lewis if he had accepted the Limit Assumption

- Success:
    - The most similar p worlds must be worlds where p is true!
- Strong Centering:
    - If w is an p-world, the p-world most similar to w is w itself!
    - Without this principle, modus ponens would not be valid:
        - $\phi > \psi$ could be true in a $\phi \wedge \neg\psi$-world w because ϕ and ψ are true at these other most similar worlds
    - Relatedly, it captures the idea that if $\phi \wedge \neg\psi$ is true, then $\phi > \psi$ is false
    - Actually, **weak centering** would suffice for both purposes:

      Weak Centering $w \in f(w,p)$ if $w \in p$
    - Why strong rather than weak centering?
    - Strong centering ensures that $\phi \wedge \psi$ entails $\phi > \psi$
        - Lewis (1973a: 27-8): inferences of this form sound odd, but that is just because it is weird to assert a subjunctive conditional when you've already asserted the antecedent.
            - Further, you say: *if Caspar had come it would have been a good party.* I can say: *That's true, for he did and it was a good party. You didn't see him because you spent the whole time in the kitchen, missing the fun.*
- Uniformity:
    - No justification in terms of similarity, logical justification instead:
        - $\mathsf{A} > \mathsf{B}, \mathsf{B} > \mathsf{A}, \mathsf{A} > \mathsf{C}$ should entail $\mathsf{B} > \mathsf{C}$
            - The premises say that A and B are counterfactually equivalent, and that $\mathsf{A} > \mathsf{C}$. So $\mathsf{B} > \mathsf{C}$ must follow!
            - If I were to eat an apple, I would eat a banana. If I were to eat a banana, I would eat an apple. Also, if I were to eat an apple, I would eat a cherry.
            - So if I were to eat a banana, I would eat a cherry.

- Uniqueness:
  - Again, no justification in terms of similarity
  - Logical consequences:

    Conditional Negation $\neg(\phi > \psi) \vDash \phi > \neg\psi$

    - Predicts that (13) seems to entail (14):

      (13) It's false that Vietnam would have been just as bad if Kennedy had lived.

      (14) If Kennedy had lived, Vietnam would not have been just as bad

    Conditional Excluded Middle $(\phi > \psi) \vee (\phi > \neg\psi)$

    - Predicts one of the following must be true:

      (15) If Stalnaker were a farmer, he would grow corn

      (16) If Stalnaker were a farmer, he wouldn't grow corn

3.1 Antecedent Strengthening Revisited

- How does this semantics block the inference from $\mathsf{A} > \mathsf{C}$ to $(\mathsf{A} \wedge \mathsf{B}) > \mathsf{C}$?

World	A	B	C
w_0	1	1	1
w_1	1	1	0
w_2	1	0	1
w_3	1	0	0
w_4	0	1	1
w_5	0	1	0
w_6	0	0	1
w_7	0	0	0

Fig. 1. The space of worlds W

  - Let's evaluate $\mathsf{A} > \mathsf{C}$ and $(\mathsf{A} \wedge \mathsf{B}) > \mathsf{C}$ in w_5
  - A valid f:
    - $f(w_5, [\![\mathsf{A}]\!]^f) = \{w_2\}$
    - $f(w_5, [\![\mathsf{A} \wedge \mathsf{B}]\!]^f) = \{w_1\}$
  - Since C is true in w_2, $\mathsf{A} > \mathsf{C}$ is true in w_5
  - Since C is false in w_1, $(\mathsf{A} \wedge \mathsf{B}) > \mathsf{C}$ is false in w_5

3.2 Surprising Consequences

- Transitivity is invalid: $\phi_1 > \phi_2, \phi_2 > \psi \nvDash \phi_1 > \psi$
  - Surprising, since this sounds good:

    (17) If I were to kick the door, my foot would hurt.

    (18) If my foot were to hurt, I would be sad.

    (19) So, if I were to kick the door, I would be sad.
  - Stalnaker's counterexample:

    (20) If J. Edgar Hoover were today a communist, then he would be a traitor.

    (21) If J. Edgar Hoover had been born a Russian, then he would today be a communist.

    (22) So, if J. Edgar Hoover had been born a Russian, he would be a traitor.
- Actually, transitivity can't be valid if antecedent strengthening is invalid:
  - $(\mathsf{A} \wedge \mathsf{B}) > \mathsf{A}$ is a tautology
  - But from this and $\mathsf{A} > \mathsf{C}$ it would follow by transitivity that $(\mathsf{A} \wedge \mathsf{B}) > \mathsf{C}$
- Contraposition is invalid: $\phi > \psi \not\dashv\nvDash \neg\psi > \neg\phi$
  - Counterexample:

    (23) If I were to kick the door, my foot would hurt (more).

    (24) So, if my foot weren't hurting, I wouldn't have kicked the door.
- Contraposition can't be valid if antecedent strengthening is invalid:
  - Fact: $\mathsf{A} > \mathsf{C}$ if $\mathsf{A} > \mathsf{B}$ and $\mathsf{B} \vDash \mathsf{C}$.
  - Suppose $\mathsf{A} > \mathsf{C}$ and infer $\neg\mathsf{C} > \neg\mathsf{A}$ by contraposition.
  - $\neg\mathsf{A} \vDash \neg(\mathsf{A} \wedge \mathsf{B})$, so by our fact: $\neg\mathsf{C} > \neg(\mathsf{A} \wedge \mathsf{B})$
  - By contraposition again: $(\mathsf{A} \wedge \mathsf{B}) > \mathsf{C}$
- Import-Export is invalid: $\phi_1 > (\phi_2 > \psi) \not\dashv\nvDash (\phi_1 \wedge \phi_2) > \psi$
  - This generally sounds good:

    (25) If Ann had been happy, then Carl would have been happy if Bob had been happy.

    (26) So, if Ann and Bob had been happy, then Carl would have been happy.

3.3 Difference 1: The Uniqueness Assumption

- Lewis: obviously there are ties in similarity and in these cases conditional excluded middle seems to fail
  - It seems false that if I were older, I would be 35.

- But it also seems false that I were older, I wouldn't be 35.
- Stalnaker: where you see ties, I see vagueness in which f we are using
  - His strategy: apply standard technology for vague language to conditionals
  - In particular, supervaluations (van Frassen).
  - I'm neither 35 nor not 35 in the closest world, because the context underdetermines the selection function.
- Furthermore, as Lewis (1973a: 80) acknowledges, this sounds inconsistent:

(27) It isn't the case that if Bizet and Verdi were compatriots, Bizet would be Italian
$\neg(\mathsf{C} > \mathsf{B})$

(28) It isn't the case that if Bizet and Verdi were compatriots, Bizet wouldn't be Italian
$\neg(\mathsf{C} > \neg\mathsf{B})$

(29) If Bizet and Verdi were compatriots, Bizet either would or would not be Italian
$\mathsf{C} > (\mathsf{B} \vee \neg\mathsf{B})$

- The uniqueness assumption renders this inconsistent, but without it they should be consistent!
- But, Lewis (1973a: 80-3) thinks the cost is too high. How can Stalnaker capture *might* or *could* counterfactuals?
  - If I had been born in the 60s, I might have done drugs with my parents
  - This is inconsistent with:
    - If I had been born in the 60s, I wouldn't have done drugs with my parents
  - This suggests analyzing them as $\neg(\phi > \neg\psi)$
  - But uniqueness validates conditional negation, so this implies $\phi > \psi$, so *might* counterfactuals come out as equivalent to *would* counterfactuals. Oops.
- Stalnaker (1984: Ch.7) develops an account built on the idea that *might* here is epistemic and takes scope over a *would* counterfactual.
  - This is compositionally implausible
  - And it is just implausible for the case above
  - It seems clear that the *might* is not about my knowledge
  - But in the case of *could* this is even clearer:
    - If I had been born in the 60s, I could have done drugs with my parents
  - It's not that I don't know if I would have. It's that I would have had the opportunity to.

3.4 Difference 2: The Limit Assumption

- One potential consideration in favor of the limit assumption:

  - Everybody agrees this is true:
    - If:
      - $\mathsf{A} > \mathsf{B}$
      - And $\mathsf{B} \vDash \mathsf{C}$

      Then:
      - $\mathsf{A} > \mathsf{C}$
  - And this:
    - If:
      - $\mathsf{A} > \mathsf{B_1}$
        $\vdots$
        $\mathsf{A} > \mathsf{B_n}$
      - And $\mathsf{B_1}, \ldots, \mathsf{B_n} \vDash \mathsf{C}$

      Then:
      - $\mathsf{A} > \mathsf{C}$
  - But limit assumption is required for general version (Pollock 1976; Herzberger 1979):
    - Let $\Gamma = \{\mathsf{B_1}, \mathsf{B_2}, \ldots\}$ (potentially infinite set of premises)
    - If:
      - For all $\mathsf{B_i} \in \Gamma$: $\mathsf{A} > \mathsf{B_i}$
      - And $\Gamma \vDash \mathsf{C}$

      Then:
      - $\mathsf{A} > \mathsf{C}$
- But then Lewis (1973a: 20) brings up the case of a line just under 1 inch long:
  - If it were more than 1 inch long...
  - Which worlds count as best?
    - For each world with n greater than 1, there is a still closer world, where the line is $n - m$ long.
  - Why not say that all worlds with a line greater than 1 inch count as most similar?
  - Because a world where the line is 1.1 inches long is more similar to our world than one where it is 100 feet long!
- More on this next week.

Next week: objections to Lewis-Stalnaker analysis

References

Anderson, AR (1951). 'A Note on Subjunctive and Counterfactual Conditionals.' *Analysis*, **12**: 35–38.

Goodman, N (1947). 'The Problem of Counterfactual Conditionals.' *The Journal of Philosophy*, **44**: 113–118.

Herzberger, H (1979). 'Counterfactuals and Consistency.' *Journal of Philosophy*, **76(2)**: 83–88.

Lewis, DK (1973a). *Counterfactuals*. Cambridge, Massachusetts: Harvard University Press.

Lewis, DK (1973b). 'Counterfactuals and Comparative Possibility.' *Journal of Philosophical Logic*, **2(4)**: 418–446.

Pollock, J (1976). *Subjunctive Reasoning*. Boston: D. Reidel Publishing Co. URL http://oscarhome.soc-sci.arizona.edu/ftp/PAPERS/Pollock_Subjunctive_Reasoning.pdf.

Stalnaker, RC (1968). 'A Theory of Conditionals.' In N Rescher (ed.), *Studies in Logical Theory*, 98–112. Oxford: Basil Blackwell Publishers.

Stalnaker, RC (1975). 'Indicative Conditionals.' *Philosophia*, **5**: 269–286. Page references to reprint in Stalnaker (1999).

Stalnaker, RC (1984). *Inquiry*. Cambridge, Massachusetts: The MIT Press.

Stalnaker, RC (1999). *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford: Oxford University Press.

Stalnaker, RC & Thomason, RH (1970). 'A Semantic Analysis of Conditional Logic.' *Theoria*, **36**: 23–42.