

The linguistics of desire ⁽⁶⁾

Kai von Fintel & Sabine Iatridou (MIT)

CSSL18 — July 16–27, 2018

Today

1. A bit of mop-up from yesterday
2. Anankastic conditionals

Conditional wants

- (1) If there's desert, I want you to stop me.
- A 1L analysis seems most appropriate.
 - This is possible if *want* has a restrictable modal base component.

Other conditional attitudes?

- (2) If Bill is on a plane to Cuba, then I am surprised that he left without saying goodbye.

Blumberg & Holguín 2018, “Embedded attitudes”

Additional ordering effects

Recall Kaufmann & Schwager's argument:

- (3) If you lose your job, take a lower-paying one!
- (4) But if you lose your job and have a comparable offer, don't take a lowerpaying one.

Lewis on Sobel Sequences

If we take any one counterfactual, this will do nicely. But trouble may come if we consider several counterfactuals together. (1) *'If I (or you, or anyone else) walked on the lawn, no harm at all would come of it; but if everyone did that, the lawn would be ruined.'* (2) *'If the USA threw its weapons into the sea tomorrow, there would be war; but if the USA and the other nuclear powers all threw their weapons into the sea tomorrow there would be peace; but if they did so without sufficient precautions against polluting the world's fisheries there would be war; but if, after doing so, they immediately offered generous reparations for the pollution there would be peace;'** (3) *'If Otto had come, it would have been a lively party; but if both Otto and Anna had come it would have been a dreary party; but if Waldo had come as well, it would have been lively; but. . . .'*

Sobelized conditional desires

[There can only be one kind of music at the party.]

If Otto comes, I want rock'n'roll music. If Otto and Anna come, I want reggae. If Otto and Anna and Waldo come, I want jazz.

Plus, imagine that if any of these people come, I will be so stressed out that I will not have any relevant desires. Which is why I need to spell out my current actual desires for the various scenarios.

NB: if you are interested in the fact that these sequences only work smoothly in one direction, please go to Google Scholar and search for “Reverse Sobel Sequences”.

The conundrum

We want a 1L restricted reading (since we're talking about actual preferences), but we also want stereotypicality.

1.5L?

Can we have a semantics for *if* that

- introduces stereotypicality
- but doesn't move the priority operators's evaluation point away from the actual world?

Yes, 1.5L!

Use the plural definite analysis and restrict the overt operator to those worlds.

if p , want q

x wants $(f \cap [\text{if } f'g'p])(g)(q)$

- f : DOX (in w)
- g : x 's desires (in w)
- f' : W or DOX (in w)
- g' : similarity (to w)

But maybe we can get this another way

The stereotypicality effect seems to arise without an outer layer:

(5) A: I lost my job.

B: You should take a lower-paying one, ... but of course, if you have a comparable offer, you should take that one

Huh

Skipped (see slides from yesterday that we didn't get to)

- Q: How to restrict attitudes when attitudes are not quantifiers?
- A: Restrict the (c)overt harmonic modals that spell out the content of the attitudes!

Anankastic conditionals:
An unsolved problem

The history

Sæbø 2001

von Fintel & Iatridou 2005/2006

von Stechow, Krasikova, Penka 2004-2006

Huitink 2005/2008

Nissenbaum 2005

Werner 2006

Condoravdi & Lauer 2016

Phillips-Brown 2018

Non-anankastic conditionals with *want* in the antecedent

- (6) If you want sugar in your soup, you must/might be diabetic.
- (7) If you want sugar in your soup, you should see a doctor.

Anankastic conditionals

- (8) If you want sugar in your soup, you should ask the waiter.
- (9) If you want to go to Harlem, you have to take the A train.

Problem I: conditioning on goals

How can the antecedent ensure that the advice modal *should* in the consequent values the goal of having sugar in you soup?

Solution I: The 2L analysis of anankastics

Consensus solution: the *if*-clause is part of a higher layer of modal quantification.

\approx in all worlds w' close to w where you want sugar in your soup, all of the worlds w'' where you best achieve your goals in w' are worlds where you ask the waiter

Problem II: The Hoboken scenario (incompatible goals)

- (10)
 - a. You want to go to Hoboken.
 - b. Harlem and Hoboken are de facto incompatible goals, e.g. for time reasons you can't visit both places on one day.
 - c. The PATH train goes to Hoboken.
 - d. The A train goes to Harlem.
- (11) If you want to go to Harlem, you have to take the A train.

Solution IIa: designated goals

vF&I 2005: the goal raised to salience by the antecedent becomes the “designated” goal, overriding any possibly incompatible goal

Nice intuition, but no compositional implementation

Solution IIb: von Fintel's *want*

Roger Schwarzschild, pc to Kai, Sept 17, 2005

What if you adopt a “hard-nosed” semantics for want, such as the one in von Fintel 1999?

If both Harlem and Hoboken are in your ordering source, then it's false that you want to go to Harlem and false that you want to go to Hoboken. You're indifferent.

Solution IIc: effective preference *want*

Condoravdi & Lauer 2016 propose a different hard-nosed semantics for *want* that seems to work just as well as von Steinhilber's semantics as far as anankastics are concerned.

“An agent is generally subject to a large number of constraints and attitudes that influence his actions: desires, inclinations, personal moral codes, and obligations, to name but a few. All of these come in varying degrees of importance.”

“at any given world, any agent has a set of such structures representing the various sources of his preferences”

- (12) Given a set of worlds W , a *preference structure* is a pair $\langle \mathbf{P}, \prec \rangle$, where $\mathbf{P} \subseteq \wp(W)$ and \prec is a strict partial order on \mathbf{P} .

“a rational agent a in w will have a distinguished preference structure that he uses to guide action choice, which we call a ’s effective preference structure at w ”

- (13) **Consistency.** A preference structure $\langle \mathbf{P}, \prec \rangle$ is *consistent* with respect to an information state B iff for any $X \subseteq \mathbf{P}$, if $B \cap \bigcap X = \emptyset$, there are $p, q \in X$ such that $p \prec q$.

A function to pick out the highest-ranked preference in a preference structure:

$$(14) \quad \max[\langle \mathbf{P}, \prec \rangle] := \{p \in \mathbf{P} \mid \neg \exists q \in \mathbf{P}: p \prec q\}$$

The meaning of *want*:

(15) $want_P(a, \phi)$ is true in w iff $\llbracket \phi \rrbracket \in \max[P(a, w)]$

The return of conflicting goals

Phillips-Brown 2018:

Just having a conflict-free semantics is not enough, since the sense in which these semantics are conflict-free is that they don't allow conflicting desires that the agent believes are conflicting. Nothing prevents desires that the agent believes are consistent but that are actually conflicting.

The New Hoboken Scenario

- the only way to get to Harlem from here is the A train, there's no bus
- you want to never take the subway anywhere (but you haven't told anyone)
- you (falsely) believe that there might be a bus to Harlem from here

Case 1

You form a desire to go to Harlem. This is compatible with your other desire (to never take the subway), as long as there's a bus, which is compatible with your beliefs.

von Fintel: *I want to go to Harlem* is true if said by you.

Condoravdi & Lauer: *I want to go to Harlem* is true as a statement about your effective preferences (going to Harlem and never taking the train are consistent with each other and your belief state).

(16) A: I want to go to Harlem.

B: You have to take the A train.

Both accounts incorrectly predict B's response to be on shaky grounds. Since B can't know that A doesn't have another desire that unbeknownst to A conflicts with the desire to go to Harlem, B can't be sure that all of the worlds where A's desires are satisfied as best as can be are A train worlds.

But B's response seems entirely correct. A does have to take the A train to go to Harlem.

Case 2

B is wondering what A wants to do. B says:

(17) If you want to go to Harlem, you have to take the A train.

Both accounts incorrectly predict this to be false. B can't rule out a scenario where A wants to never take trains and wants to go to Harlem and thinks those might be consistent, but in such a scenario not all the best worlds by A's lights are worlds where A goes to Harlem by train.

Conclusion: We need to knock out *all* inconsistent goals, not just the ones believed to be inconsistent by the agent. The nested analysis plus any semantics that only rules out conflicts that the agent believes exist will not do the trick.

Back to the drawing board!

Last but not least

X-marking

in conditionals, in modals, and in desires