

# On the monotonicity of desire ascriptions

Kai von Fintel

Incomplete notes from late 2018

## Introduction

An attractively straightforward possible worlds semantics for desire ascriptions is routinely criticized for its apparently problematic predictions about the monotonicity of such ascriptions. It is time to set the record straight: desire ascriptions are in fact monotonic and accounts that say they aren't are inadequate. This short note will not pull together all the strands of argument (for some of that see [Crnić 2011](#)) but will focus on a few points that have perhaps not been appreciated.

The question before us is whether there is an entailment from (1a) to (1b):

- (1) Assuming that  $p$  entails  $q$ ,
  - a.  $x$  wants  $p$
  - b.  $x$  wants  $q$

We will begin with making a *prima facie* argument that this entailment is indeed present.

## Prima facie

Three scenarios where a monotonic entailment seems exactly correct.

## Fish

Alex and Billy are out for dinner at Oleana. Alex studies the menu and orders the Striper with Sour Cherry Kisir, Pistachio Muhammara, Chard, and Tahini. Billy is astonished:

- (2) B: Really? You want fish?  
A: Yes. The Striper sounds good.  
B: I thought you were vegetarian.  
A: Mostly, but once in a while I have some seafood.

We will leave Alex and Billy to their date. They have much to learn about each other. Let's focus on the fact that they are agreed that wanting the Striper means that Alex wants fish for dinner. It's hard to see how one could deny that Billy is drawing a licit inference on the basis of Alex's stated desire for the Striper.

### **Christmas**

Here's an example fashioned after one from [Asher 1987](#):

- (3) A: Alexis wants a puppet and a puzzle for Christmas.  
B: *goes out and buys a puppet*  
A: ??What were you thinking? I didn't say she wanted a puppet.

Having said that Alexis wants a puppet and a puzzle, it is astonishingly weird for A to deny thereby having said that Alexis wants a puppet.

### **Weekend**

Finally, An example suggested by Paul Crowley (pc):

- (4) A: I want to leave on Saturday.  
B: Really, you want to leave on the weekend? Why not wait for Monday?

Again, B's inference that A wants to leave on the weekend since she wants to leave on Saturday seems straightforward.

### **von Fintel vs. Heim**

Having seen prima facie reasons for a monotonic semantics for *want*, we cast around for relevant analyses on the market. As it turns out, monotonic analyses are thin on the ground.

As stand-ins for the monotonic and the non-monotonic perspective, I will take the account by [von Fintel 1999](#) and that by [Heim 1992](#), respectively. There are plenty of other non-monotonic accounts on the market, especially those by [Villalta 2008](#) and the probabilistic ones developed by [Levinson 2003](#) and [Lassiter 2017](#), but I will leave them aside and leave their assessment vis-a-vis the arguments in this note to the reader as an exercise. In what follows, I'm using simplified versions of von Fintel and Heim's proposals that are

boiled down to their essence, abstracting from complications not germane to the question of monotonicity. Again, it's an exercise to consider the full versions of the proposals vis-a-vis the arguments in this note.

The account by [von Fintel 1999](#) is a refinement of a Hintikka-style semantics for *want*. It is what is now sometimes called a “best worlds” semantics. The gist is that a desire ascription claims that the best worlds according to the agent's preferences are all worlds in which the prejacent is true. There are two additional ingredients: (i) the worlds up for comparison are constituted by the agent's doxastic set, and (ii) the prejacent needs to not already be settled by the agent's doxastic set. The first addition is there to account for the fact that many of our desires are “realistic” in the sense that they make the best of what is perhaps not an ideal situation. The second addition is there to explain why the first addition doesn't lead to us wanting anything that we believe to be true (a desire version of the Samaritan Paradox). Here then is the formulation we will use:

$$(5) \quad \llbracket \text{want} \rrbracket = \lambda p. \lambda x. \lambda w: p \not\subseteq \text{DOX}_{x,w} \ \& \ p \cap \text{DOX}_{x,w} \neq \emptyset. \\ \forall w' \in \text{BEST}_{x,w}(\text{DOX}_{x,w}): p(w') = 1.$$

In words: among  $x$ 's belief worlds in  $w$ , the ones that are best by the lights of  $x$ 's preferences in  $w$  are all  $p$ -worlds.

This semantics is *almost* monotonic on the prejacent. If  $x$  wants (it to be true that she has) striper, then  $x$  wants fish. If all the best worlds are striper worlds, since all striper worlds are fish worlds, they will all be fish worlds. There is just one proviso: the entailment fails if  $x$  believes that fish is the only choice, that is that there are only fish worlds in  $x$ 's doxastic set. In that case,  $x$  can't be said to want fish, since there's no choice. But if fish is an open choice (the diversity presupposition is satisfied), then the entailment holds. In other words, the account is Strawson-monotone on the prejacent.

The account by [Heim 1992](#) is officially a dynamic account but we can extract a static account (based in part on an intermediate formulation in her paper). The idea is based on an intuition articulated by [Stalnaker 1984](#). In Heim's words:

The analysis of desire verbs I want to pursue here is sketched in Stalnaker 1984, p.89: ‘wanting something is preferring it to certain relevant alternatives, the relevant alternatives being those possibilities that the agent believes will be realized if he does not get what he wants.’ An important feature of this analysis is

that it sees a hidden conditional in every desire report. A little more explicitly, the leading intuition is that *John wants you to leave* means that John thinks that if you leave he will be in a more desirable world than if you don't leave.

The account is easiest to grasp if one uses Stalnaker's semantics for conditionals, which is based on a function that for any pair of a world  $w$  and a proposition  $p$  selects from among the worlds where  $p$  is true the unique world that is maximally similar to  $w$ . This will be  $w$  itself when the proposition is true of  $w$ . (As before, variants are left to the reader to ponder.)

Heim's account is then this:

$$(6) \quad \llbracket \text{want} \rrbracket = \lambda p. \lambda x. \lambda w. \forall w' \in \text{DOX}_{x,w} : \\ \text{SIM}_{w'}(\text{DOX}_{x,w} \cap p) <_{x,w} \text{SIM}_{w'}(\text{DOX}_{x,w} - p).$$

In words: every one of  $x$ 's belief worlds that is a  $p$ -world is better than the most similar non- $p$ -world and every one of  $x$ 's belief worlds that is a non- $p$ -world is worse than the most similar  $p$ -world.

Some notes:

- The comparison is a pairwise comparison of  $p$ /not- $p$  pairs of worlds.
- As we will see, the semantics is very demanding. But one finds misrepresentations of it in the literature, according to which it makes an even stronger demand: that all  $p$ -worlds are better than any non- $p$ -worlds. This is not so.
- Note also that the diversity condition is built in by assuming that the SIM-function presupposes that it is being given a non-empty proposition.
- It doesn't follow from  $w'$  being the most similar non- $p$ -world to the  $p$ -world  $w$  that  $w$  is the most similar  $p$ -world to  $w'$ .

Heim's semantics is non-monotonic and proudly so. The basic feature that makes non-monotonic analyses non-monotonic is that when we move to the weaker prejacent, which is true in a superset of worlds compared to the original prejacent, more worlds come into play and affect the *want*-ascription.

So, consider Alex who wants the striper. For Heim, this means that each striper world is better than the most similar non-striper world. But to want fish, much more would need to hold: every fish world, including the swordfish worlds (which Alex explicitly did not choose when looking at the menu),

needs to be better than the most similar non-fish-world. Nothing about the first *want*-ascription would guarantee this, so we have non-monotonicity.

The stark difference between von Fintel and Heim can be brought out with the following scenario, a variant of the dinner date scenario:<sup>1</sup>

Imagine you fancy some fish for dinner. If you don't order fish, you'd order a green salad. There's a belief world where the fish is undercooked; after all, you may not be so sure about the restaurant. There's no way even this place will mess up a green salad. So, there is a *p*-world among your doxastic alternatives (undercooked fish) that is worse than the nearest non-*p*-world (acceptable green salad). So, according to Heim, it's false that you want fish. It may also be false that you want green salad. It's hard to have desires in Heim's world.

### **Non-monotonicity, hypersensitivity, and hyperspecificity**

A Heimian semantics for desire ascriptions, and perhaps any other reasonable non-monotonic semantics, has much in common with a very plausible idea about *desires*. The idea is that we can find out what a desire is by investigating what it takes for the desire to be satisfied (and of course also, unsatisfied, and perhaps moot). And if we ask what the satisfaction conditions are for Alex's desire for fish at the date with Billy, the answer is that Alex's desire won't be satisfied by just any fish dish. It goes without saying that the dish has to be well-prepared, and Alex may have other relevant preferences that modulate her desire: sustainably fished, no oregano, what have you. This much I completely agree with.

So one way to motivate the Heimian semantics is that it reflects the usually quite specific satisfaction conditions of desires. To desire something means to prefer any doxastically accessible way of it being realized to the most similar way of its not being realized.

But carrying this true insight into the nature of *desire* over to the semantics of desire *ascriptions* is a confusion. *x wants p* does not mean that *if p, x's desire for p will be satisfied* or anything like that. Desire ascriptions do not report what the satisfaction conditions for a desire are. Rather, desire ascriptions report a property of the worlds in which the agent's desires are as best satisfied as they can be given their beliefs. *x wants p* means *given x's beliefs, the worlds that best satisfy x's desires are all p-worlds*.

---

<sup>1</sup> The fish scenarios in this paper are all inspired by similar experiments in [Fara 2013](#).

In a slogan: **desires are non-monotonic, desire ascriptions are monotonic.**

## References

- Asher, Nicholas. 1987. A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy* 10(2). 125–197. <https://doi.org/10.1007/BF00584317>.
- Crnić, Luka. 2011. *Getting even*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis. <http://pluto.huji.ac.il/~crnic/crnic-diss-11.pdf>.
- Fara, Delia Graff. 2013. Specifying desires. *Noûs* 47(2). 250–272. <https://doi.org/10.1111/j.1468-0068.2012.00856.x>.
- von Stechow, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97–148. <https://doi.org/10.1093/jos/16.2.97>.
- Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221. <https://doi.org/10.1093/jos/9.3.183>.
- Lassiter, Daniel. 2017. *Graded modality: Qualitative and quantitative perspectives* (Oxford Scholarship Online). Oxford University Press. <https://doi.org/10.1093/oso/9780198701347.001.0001>.
- Levinson, Dmitry. 2003. Probabilistic model-theoretic semantics for *want*. *Semantics and Linguistic Theory (SALT)* 13. 222–239. <https://doi.org/10.3765/salt.v13i0.2888>.
- Stalnaker, Robert. 1984. *Inquiry*. MIT Press.
- Villalta, Elisabeth. 2008. Mood and gradability: An investigation of the subjunctive mood in Spanish. *Linguistics and Philosophy* 31(4). 467–522. <https://doi.org/10.1007/s10988-008-9046-x>.