John F. Horty, <u>Reasons as Defaults</u> (Oxford University Press, 2012) Central APA Author Meets Critics Session Chicago, 2014

# Default Logic and <u>Hafta</u>

Anthony S. Gillies *Rutgers University* 

#### 1 A Confession and Two Caveats

Jeff Horty has written a very good book: Horty 2012.<sup>1</sup> I like more of it than I don't and agree with more in it than I disagree with. That would seem to make for an uninteresting critic and I can't really dispute that. But I'll press on. My focus here will be on the accounts of obligation — all-things-considered-oughts — based on underlying default theories. Two related caveats.

First caveat: Horty develops these accounts as philosophical analyses of obligation not as linguistic theories of the meaning of certain natural language expressions that give voice to those obligations. Nevertheless I will approach what I have to say from the perspective of the semantics of those constructions.

One reason is that I want to stick up for the toolkit that is intensional semantics. There are, if not punches thrown, at least some sharp elbows raised in its direction in a few spots in the book. An example:

They [the disjunctive and conflict accounts] cannot, therefore, be articulated in any simple way within the modal, or intensional, framework that is so often appealed to as a foundation for deontic logic .... Horty 2012: 81

I think the distance here is exaggerated, so I want to press some issues to show that. But there are also differences between what an account based on worlds and orderings and things like that can deliver and what an account based on default logic can deliver. So I want to highlight those, too.

<sup>1</sup> Based on a lot of earlier, also very good, papers — notably Horty 2007. I don't really like books, so I mean this as really high praise.

Another reason I want to approach things from the perspective of the semantics of obligation-describing constructions is that this can be a source for data that might push us one way or another. (More on just what I mean later.) So it can rescue us from the worst sort of table-pounding and footstomping.

Second caveat: Horty follows philosophical convention in labeling these obligations as *oughts*. I will, for the most part, not be following along. Why not? Well because I suspect that most of the time what philosophers mean to be talking about when they talk about *ought* is something more like *hafta/have to*. The thing is that *ought* and *should* seem to be <u>weak necessity</u> modals.<sup>2</sup> That is: (deontic) *have to* seems to asymmetrically entail (deontic) *ought*.

What you ought to do doesn't entail what you have to do.<sup>3</sup>

- (1) a. You ought to be home by ten, but you don't hafta be.
  - b. Everyone ought to wash their hands; employees have to.

If *ought* entailed *have to* then (1a) would be a contradiction and the second conjunct of (1b) would be weirdly non-informative. It isn't and it isn't so it doesn't. Similarly, if *have to* didn't entail *ought* then it should be easy to find coherent stretches where a *have to* is sensibly and informatively continued with an *ought*.

(2) a. ??You have to be home by ten, and moreover you ought to be.b. ??Everyone has to wash their hands; employees ought to.

I take it that the modal expression that gives voice to what we are obliged to do is the stronger, asymmetrically entailing *have to* (and its cognates).

## 2 Something Like the Canon

I want to briefly sketch a picture of the semantics of obligation-describing modals like *hafta* using the toolkit of intensional semantics. In many ways it draws on what is the canon but is re-framed just a bit.<sup>4</sup>

<sup>2</sup> See von Fintel & Iatridou 2008 and the references therein.

<sup>3</sup> For this to be probative, we have to fix on a single flavor and resolution for the modals involved.

<sup>4</sup> The basic framework is developed by Kratzer (1981, 1991, 2010) and now has textbook status (von Fintel & Heim 2011).

The accounts of obligation in Horty 2012 run roughly as follows.<sup>5</sup>

**Definition 1.** Fix a default theory  $\Delta = \langle W, D \rangle$  and let  $\Delta^X = \langle W \cup \{X\}, D \rangle$ . Then:

- i.  $X \parallel \frac{d}{\Delta} \bigcirc Y$  iff for every *E* that is an extension of  $\Delta^X$ :  $E \models Y$ .
- ii.  $X \parallel_{\Delta}^{c} \bigcirc Y$  iff for some *E* that is an extension of  $\Delta^{X}$ :  $E \models Y$ .

Say that  $\bigcirc$  (*Y*|*X*) (w.r.t.  $\triangle$ ) iff *X*  $\parallel_{\overline{\Delta}} \bigcirc$  *Y*.

Obviously this requires something to be said about extensions. Horty does this by saying what the <u>stable/proper</u> sets of defaults of a theory  $\Delta$  are. They are the sets of defaults from *D* such that the defaults that are triggered and not conflicted given that set are just that set of defaults.

Given this set-up, built as it is on a nonmontonic logic, it is unsurprising that obligations are predicted to be nonmonotonic as well. That is:

**Fact 1.**  $X \parallel_{\Delta} \bigcirc Y \neq Z \parallel_{\Delta} \bigcirc Y$  even though  $X \subset Z$ .

And similarly if we enrich not the factual premises but the default rules.

**Fact 2.**  $X \parallel_{\Delta} \bigcirc Y \neq X \parallel_{\Delta'} \bigcirc Y$  even though  $D \subset D'$ .

That seems right on both scores. There is a strict attendance policy, but it tolerates exceptions if you have a note from the doctor.

- (3) a. Given it's monday, you hafta go to work.
  - b. Given it's Monday and you have a note from the doctor, you don't hafta go to work.

There's another sense in which *hafta* is monotonic and Horty's accounts deliver that, too. The sense is this one: what you have to do is closed under entailment.

**Fact 3.** If  $Y \models Z$  then  $X \parallel_{\overline{\Delta}} \bigcirc Y \Rightarrow X \parallel_{\overline{\Delta}} \bigcirc Z$ .

In particular this kind of monotonicity says that  $\bigcirc$  is <u>upward entailing</u> (hence the alias of this property in semantics: upward monotonicity).

<sup>5</sup> What I have to say can be said—I think—with respect to fixed-priority default theories where the ordering is empty so I'll be assuming that throughout.

There is an allegation that this is a mistake.<sup>6</sup> But the allegation that this is a mistake is a mistake. For one thing, as Horty points out: if I hafta do this and doing this entails doing that, it's hard to see how my obligations stop and this and don't go to that. We can do more than just pound that table, though. If our obligation-describing modal vocabulary weren't upward entailing then we'd find plenty of consistent conjunctions like these:

- (4) a. #You hafta go to the department and moreover you do not hafta go to campus.
  - b. #You hafta agree to write the review and do it on time but of course you don't hafta do it on time.

These are terrible, and not contingently so.

There's more: assuming upward entailingness of *hafta* gives a nice prediction of why negating such obligation-describing talk licenses negative polarity items (NPIs). Here's the short version. A robust generalization is that NPIs like *any* are only permitted in downward entailing environments.

- (5) a. #Sophie left <u>any</u> later than 3pm.
  - b. #Some student left <u>any</u> later than 3pm.
  - c. No student left <u>any</u> later than 3pm.
  - d. Every student who left any later than 3pm missed Pedro.
- (6) a. #Sophie has to leave <u>any</u> later than 3pm.
  - b. Sophie doesn't hafta leave any later than 3pm.

This pattern is expected if  $\bigcirc$  is upward entailing since then negating it produces a downward entailing environment.

So what about the canon? The idea is that *hafta*, like modal expressions generally, is a context-dependent quantifier over possibilities. I'll just treat contexts as determining a body of information (the modal base, typically what the relevant agent knows) together with a set of normative constraints (e.g., what the law says). Normative constraints will be <u>local preferences</u>: that is, statements saying that given *X*, *Y* is better than not-*Y*.<sup>7</sup>

<sup>6</sup> The allegations tend to rely on either Ross's Paradox or professors who are too optimistic about whether they are chronic procrastinators. See von Fintel 2012 for a discussion of this.

<sup>7</sup> In Kratzer's framework, the set of constraints is a set of propositions that induces an ordering; what you hafta do is what is best in the modal base given the ordering. I'll depart in presentation from this a bit. This "normative constraints" talk is meant to be a neutral way of talking about defaults, preferences, propositions in a an ordering source, etc. all in one breath.

**Definition 2.** A context *c* determines both a set *K* of propositions characterizing what we know and a set *P* of constraints.

- i.  $\prec_P$  is based on *P* iff:  $w \prec_P v$  iff
  - a. for every  $X \in P$ : if  $v \in X$  then  $w \in X$ ; and
  - b. for some  $X \in P$ :  $w \in X$  and  $v \notin X$ .
- ii. Let best( $K, \prec_P$ ) be the non-bettered worlds (w.r.t.  $\prec_P$ ) compatible with K.
- iii.  $K \parallel_{\overline{P}} \bigcirc X$  iff if  $w \in \text{best}(K, \prec)$  the *w* is an *X*-world for the contextually determined  $\prec$  faithful to *P*.

The idea here is simple: preferences-plus-context determine an ordering; what you hafta do is what is true in the best worlds compatible with what you know given those preferences.

Whatever its other virtues and vices are, this sort of account has three properties worth noticing.

First: it is up to its eyeballs in the apparatus of modal/conditional semantics.

Second: it predicts exactly the same nonmonotonicity that the account(s) based on default logic predicts. That is: what we hafta do is nonmonotnic both in what we know (the best worlds compatible with *K* needn't be among the best worlds compatible with *K*-plus-some-stuff) and in what our normative constraints are.

**Fact 4.**  $K \parallel_{\overline{P}} \bigcirc X \neq K' \parallel_{\overline{P}} \bigcirc X$  even though  $K \subset K'$ .

**Fact 5.**  $K \parallel_{\overline{P}} \bigcirc X \neq K \parallel_{\overline{P'}} \bigcirc X$  even though  $P \subset P'$ .

Third:  $\bigcirc$  is upward entailing.

**Fact 6.** If  $X \models Y$  then  $K \parallel_{\overline{P}} \bigcirc X \Rightarrow K \parallel_{\overline{P}} \bigcirc Y$ .

None of this is accidental, of course. So I think the closeness of the modal/intensional framework and the default logic framework is underemphasized.

## 3 Disjunctive Account

For now let's focus on the disjunctive account. The canon in semantics actually encodes something very much like this account. But the way we glossed it made it seem like this wasn't a decision so much as a forced-choice. I want to remedy that.

The first thing to notice is how we've been hopping back and forth between a set of constraints (propositions in this case) and an ordering. This back-and-forth is possible because, as Lewis (1981) showed, talk of premises/propositions and talk of orderings turn out to be the same thing. Not only that, but talk of <u>determinate</u> partial orderings and talk of <u>sets</u> of total orderings come to the same thing. So that means we can re-state our gloss on the canon and in so doing make it clear that the disjunctive account isn't a hardwired part of the canon as we may have thought. Tinkering with the option makes it clear that a conflict account in this framework is also possible. (This will be helpful later. I hope.)

**Definition 3.** Let *P* be a set of constraints and suppose  $\prec$  is based on it. A total ordering  $\prec^*$  a refinement of  $\prec$  iff: if  $w \prec v$  then  $w \prec^* v$ .

- i. <u>Disjunctive</u>:  $K \parallel_{\overline{P}}^{d} \bigcirc X$  iff for every  $\prec^*$ : if  $w \in best(K, \prec^*)$  then w is an *X*-world.
- ii. <u>Conflict</u>:  $K \parallel_{p}^{c} \bigcirc X$  iff for some  $\prec^{*}$ : if  $w \in best(K, \prec^{*})$  then w is an *X*-world.

Intuitively: refinements collect up ways of deciding incomparabilities that preserve the original ordering.<sup>8</sup> Incomparability in this set-up corresponds to being faithful to sets of constraints that pull in opposite directions. So the resolving of incomparabilities represents adjudicating between such constraints. The disjunctive account says that what you hafta do is what is true in the best worlds compatible with the facts no matter how those decisions are made. The conflict account says you hafta do whatever is true in the best worlds on some resolution.

The disjunctive account — though not called that — is the default position in semantics. Kratzer (1981) discusses an example:

**Example 1.** A would-be politician wants (exactly) two things. One: to become mayor. Two: to avoid going to the pub. The trouble is that you have to go to

<sup>8</sup> Ditto for breaking ties if the refinements are strict total orders.

the pub regularly to become mayor. "This," she says "is a horrible story of someone who wants something but rejects the necessary means leading to fulfillment of her desires" (Kratzer 1981: xx).

What to say about what the would-be politician has to do? Kratzer's verdict: she doesn't hafta go to the pub and she doesn't hafta not go. But she does hafta do one or the other.

Worlds faithful to the become-mayor constraint at the expense of the avoid-the-pub constraint don't better and aren't bettered by the worlds faithful to the avoid-the-pub constraint at the expense of the become-mayor constraint. So some refinements decide in favor of one of the constraints; others in the other direction. So we end up quantifying over both kinds of worlds, knocking out each individual *hafta* claim but verifying the disjunctive *hafta*.

And, of course, things work out in pretty much the same way for the disjunctive account based on default logic. As we saw, extensions aren't guaranteed to be unique and the disjunctive account says that  $\bigcirc Y$  w.r.t.  $\Delta$  iff  $Y \in E$  for every extension E of  $\Delta$ .

**Example 2.** Take the two defaults to be  $\top \rightarrow M$  and  $\top \not\rightarrow P$  and the factual information to be just  $\neg M \lor P$ . A little checking verifies that we have two stable scenarios, each one containing just one of the defaults. So we have two extensions:  $E_1 = Cn(\{\neg M, \neg P\})$  and  $E_2 = Cn(\{M, P\})$ . So we have that  $M \not\models_{\Delta} \bigcirc P$  and  $M \not\models_{\Delta} \bigcirc \neg P$ . But of course  $M \not\models_{\Delta} \bigcirc (P \lor \neg P)$ .

So there's quite a lot of agreement here: both frameworks have a natural way of expressing both the conflict and disjunctive accounts, and the disjunctive accounts in particular look pretty similar. Having exaggerated some similarities, I now want to exaggerate some differences.

The underlying apparatus for the canon isn't standard deontic logic but stuff imported from the logical investigation of conditionals. Which makes sense: (simple) selection functions and accessibility relations merely bipartition the set of worlds and what we want is some comparative thing to capture how close a world is to satisfying a bunch of constraints. (This is what Kratzer calls "approaching ideals".) In fact it's not too hard to show that more or less what we've done is reproduce the minimal conditional logic in the guise of  $\bigcirc$ . That provides some clues to where we might see some divergence. I want to mention two such spots.

**Example 3.** Both Alex and Billy like the same coffee shop: so meeting Alex

requires that I go to the coffee shop and meeting Billy requires that I go to the coffee shop.

Given all this:

- (7) a. I have a meeting with either Alex or Billy (but can't remember which one). So I hafta go to the coffee shop.
  - b. #I have a meeting with either Alex or Billy (but can't remember which one). Still, I don't hafta go to the coffee shop.

The disjunctive account based on default logic doesn't straightforwardly predict this. The reason is that since  $A \lor B, C \not\models A$  and  $A \lor B, C \not\models B$  neither default gets triggered in any scenario.

**Fact 7.** Let  $D = \{A \to C, B \to C\}$  and  $W = \emptyset$ . The theory  $\Delta^{A \lor B}$  has a unique extension  $E = Cn(\{A \lor B\})$ . So  $A \lor B \not\models_{\Delta} \bigcirc C$ .

Things are actually a little worse in that (of course): since  $\bigcirc C$  doesn't follow  $\neg \bigcirc C$  does.

In the intensional framework however the disjunctive account does easily predict this pattern. The reason: worlds where I am supposed to meet Alex and go and worlds where I meet Billy aren't comparable (they both satisfy both constraints). Given that I am supposed to meet one or the other, such worlds are also the best. But since I go to the coffee shop in all such worlds regardless of how we resolve the incomparability, I hafta go.

**Fact 8.** Let  $P = \{A \to C, B \to C\}$  and  $K = \{A \lor B\}$ . Then  $K \mid_{\overline{P}} \bigcirc C$ .

In fact, in this set-up this pattern for "or" is equivalent to giving gold-star status to reasoning by cases. That's probably a credit.

The other property I want to mention: what you hafta do accumulates. If you hafta do *Y* and given that you do *Y* you hafta *Z*, then you hafta do *Z*. This property — susually called <u>cautious monotonicity</u> — is a kind of restricted monotonicity of obligation: it says that so long as obligations don't get in each other's way, then conditional on the same set of facts, what you hafta do is monotonic.

**Fact 9.**  $X \parallel_{\overline{\Delta}} \bigcirc Y$  and  $X \parallel_{\overline{\Delta}} \bigcirc Z \neq X, Y \parallel_{\overline{\Delta}} \bigcirc Z$ 

Counterexamples showing that Reiter's (1980) default logic doesn't satisfy cautious monotony turn out to be resilient: they also suffice to show that

various tweaks on the concept of extension won't solve the problem.<sup>9</sup> They also suffice for the deontic version built on default logic in Horty's work.

**Example 4.** Let  $\Delta = \langle W, D \rangle$  where  $W = \emptyset$  and  $D = \{\top \to A, A \lor B \to \neg A\}$ . Then  $\emptyset \parallel_{\Delta} \bigcirc A$  and  $\emptyset \parallel_{\Delta} \bigcirc (A \lor B)$  even though  $A \lor B \parallel_{\Delta} \neg \bigcirc A$ .

The reason why is that although  $\Delta$  has a unique extension  $E_{\Delta} = Cn(A)$  the default theory  $\Delta^{A \vee B}$  has two.<sup>10</sup>

You might suspect that I (probably) wouldn't have brought up cautious monotonicity if the disjunctive account based on the canon <u>also</u> failed to satisfy it. You'd be right: it does.

**Fact 10.** 
$$X \parallel_{\overline{P}} \bigcirc Y$$
 and  $X \parallel_{\overline{P}} \bigcirc Z \Rightarrow X, Y \parallel_{\overline{P}} \bigcirc Z$ 

My aim here wasn't (and isn't) to counterexample (or not, as the case may be) the account of *hafta* based on default logic. It is to press on the true points of divergence between that account and the kinds of accounts found in the intensional/modal family. There are, as I said, lots of points of agreement (for better or worse).

Fact 11. Both accounts satisfy the following principles:

i. 
$$X \models \bigcirc X$$
  
ii.  $(X \Leftrightarrow Y) \Rightarrow (X \models \bigcirc Z \Leftrightarrow Y \models \bigcirc Z)$   
iii.  $(X \models \bigcirc Y \text{ and } Y \models Z) \Rightarrow X \models \bigcirc Z$   
iv.  $X \models \bigcirc Y \text{ and } X \models \bigcirc Z \Rightarrow X \models \bigcirc (Y \land Z)$ 

Nothing particularly exotic: (iii) is upward entailingness and (iv) is agglomeration but might be better called adjudication.<sup>11</sup>

...

- b. If there is a fence up then there has to be.
- c. If the Mayor's car is parked there then it is allowed to be parked there.

<sup>9</sup> See Makinson 1994.

<sup>10</sup> I admit that it looks like a funny disjunction thing happening here, but that's not really so. Other variants in the simple counterexample family involve three chaining defaults with conclusions that can't all be true together.

<sup>11</sup> Reflexivity — (i) on the list — isn't innocent. It flatfootedly predicts that there should be no contingently true deontic conditionals with the same antecedent and obligation. But that's not true:

<sup>(</sup>i) a. If Britney orders a coke, then she has to order a coke.

We've seen two properties that the two approaches differ on. Add cautious monotony and antecedent disjunction to the list of properties and the result is a set of properties that characterize the core of conditional logic (and once we lift to the consequence relation the core of nonmonotonic reasoning). That logic is often called  $\mathbb{P}$ , as in preferences. That is because it characterizes what can be represented in a semantics where the main apparatus is a set of worlds together with a well-behaved preference relation on that set.<sup>12</sup>

**Fact 12.**  $\mathbb{P}$  is sound and complete w.r.t. the disjunctive account based on the canon.

Upshot(s)? There's some intuitive pull toward the idea that the language of obligation is somehow preference-based. The normative constraints in a situation somehow translate into a global ranking of possibilities, declaring some as being better than others. What we hafta do is what is best-given-thecircumstances. If that's right then it's hard to see how an account based on default logic can be pulled around the phenomena.

But maybe there is reason to think that a global ranking of possibilities isn't the right tool for representing the kind of preferences that normative constraints induce anyway. That is because there is reason to think that  $\mathbb{P}$  is too cautious: it under-generates.

**Example 5** (Tucson Running Problem). It is invariably sunny and almost invariably hot in Tucson. So it is better to run early rather than late and better to wear your sunglasses than not. In January it is better to run late than early. Given that you are going for a run in January, you hafta wear your sunglasses.<sup>13</sup>

This pattern is beyond what  $\mathbb{P}$  can deliver and there are few options for adding strength to  $\mathbb{P}$  to get it. So maybe we need a better framework for moving from local ceteris paribus constraints to all-things-considered obligations: a global ranking won't get us where we need to go.

Each of expresses (in the relevant circumstances) a non-trivial claim about what's obligated and permitted. See Frank 1996; Zvolenszky 2002; Kratzer 2010 for a discussion of the issues involved here.

<sup>12</sup> The "preferences" really only have to be asymmetric and transitive (ties and incomparabilities permitted) so not "well-behaved" in the economists' sense. Just not unruly relations.

<sup>13</sup> This sort of example plays a role in Gillies 2012.

## 4 Conflict Account

One of the great things Horty has done is present a clean and elegant framework within which we can at least explore an account of <u>hafta</u> that renders conflicting obligations coherent.

How? Well as we saw earlier: extensions aren't guaranteed to be unique and the conflict account says that  $X \parallel_{\Delta} \bigcirc Y$  iff  $Y \in E$  for some extension E of  $\Delta^X$ .

We also saw that this can be mimicked in the intensional framework as well. Since  $\prec_P$  may have different refinements the conflict account says  $X \parallel_{\overline{P}} \bigcirc Y$  iff the best $(K, \prec^*) \subseteq Y$  for some refinement  $\prec^*$  of  $\prec_P$ .

So both frameworks have the expressive capability to formulate the conflict account. That's good. The conflict account shouldn't be ruled out a priori. We should instead go where the data says we should.

One place to look for data is how *hafta* interacts with permission modals like *may*. We have been treating *hafta* as a (strong) necessity modal. Certainly the natural thought is that permission modals are duals to this.

**Constraint 1.**  $P(X) \Leftrightarrow \neg \bigcirc \neg X$ 

There are two good reasons for this. One: <u>epistemic</u> strong necessity modals and epistemic possibility are duals.

- (8) [The marble is either in the box or under the couch.]
  - a. Given what I know, it has to be in the box.
  - b. Given what I know, it can't be under the couch.

It would be surprising — and something in need of explaining — if epistemic modals had this relationship but their deontic counterparts didn't. That would mean that the <u>flavor</u> of a modal influenced its logical relationships with other modals of the same flavor and that's weird.

Two: if (deontic) *hafta* and *may* weren't duals, then we should find acceptable instances along these lines:

- (9) a. *#*I may submit my grades one day late and I hafta submit them on time.
  - b. #I hafta submit my grades on time and I may submit them one day late.

These don't sound coherent, and I don't think that's either accidental or has

to do with the particulars of grade-submitting.

Similarly: if obligation is a (strong) universal modal and permission its dual we would expect the former to asymmetrically entail the latter.

### **Constraint 2** (Weak Permission). $\bigcirc X \Rightarrow P(X)$

Again: that relationship certainly holds for the epistemic *hafta* and *may* so it'd be weird if it didn't hold good here.

But as Horty notes we can't merely append this constraint to the conflict account. Doing so immediately renders that account inconsistent — not surprising since the two constraints together are straightforwardly entail that there can be no moral conflicts. One way to see this: on (either incarnation of) the conflict account, the presence of a deontic conflict is sufficient to render false the corresponding permission claim, and such instances counter-example the constraint that *hafta* asymmetrically entails *may*. Horty is exactly right when he says that pounding the table about intuitions of the viability of this constraint is little more than (loud) question begging.

Maybe we can get some independent traction. To be clear, the conflict account <u>isn't</u> signed up for instances of Weak Permission <u>never</u> holding. (Some of the literature Horty is responding to seems to inadvertently take that position.) Indeed:

**Fact 13.** Weak Permission fails only when the underlying default theory has multiple extensions/the underlying ordering has multiple refinements.

So the only place for traction is precisely in situations where there is conflict.

**Example 6.** Kai's taxes are due tonight by midnight. Unfortunately, so is his overdue referee report for L&P. He simply does not have time to do both.

This is a bad spot to be in, for sure. It looks, first blush anyway, like exactly the sort of scenario in which the conflict account is going to deliver that Kai has to do his taxes and has to do his referee report.<sup>14</sup>

- (10) a. Kai has to do his taxes and has to do his referee report.
  - b. Kai may do his taxes.

I admit to having some sympathy for hearing (10a) as true. That's some good news for the conflict account. But (10b) seems true, too. The problem

<sup>14</sup> The example is discussed in von Fintel 2012. I'm sure its fictional.

is that it is unambiguously and straightforwardly false according to (either incarnation of) the conflict account. We can press the point a bit.

- (11) [Kai has opted for doing his taxes.]
  - a. Thony: What are you doing?
  - b. Kai: My taxes. # I have to, even though I can't.

Again, we expect this to be a clear case where Weak Permission would have to fail and therefore a clear case in which conjunctions like (11b) would be true. But it seems not only not true, but contradictory.

One more wrinkle. Putting the conflict in the past—embedding the relevant modals under a past tense operator—should do nothing to the coherence of putative failures of Weak Permission. But things don't sound that way.

- (12) [It is now April 16.]
  - a. Thony: So, what did you decide to do?
  - b. Kai: # My taxes. I had to even though I couldn't.

This is terrible.

It is odd that we could find ourselves in a situation in which we are obliged to do stuff we are not permitted to do. But, as I said, I agree with Horty that we can't rest on intuition. I've tried to get some independent way in by focusing not on what we think obligation amounts to but by thinking about how the modals that express obligation interact. Things are tricky here but there is evidence that the conflict account isn't going to capture the dominant modal dialect.

I don't expect this settles the (or any) issue. These waters are not unmurky.

#### References

- von Fintel, Kai. 2012. The best we can (expect to) get? challenges to the classic semantics for deontic modals. URL http://mit.edu/fintel/fintel-2012-apa-ought.pdf. Ms, MIT.
- von Fintel, Kai & Irene Heim. 2011. Intensional semantics. MIT.
- von Fintel, Kai & Sabine Iatridou. 2008. How to say ought in foreign: The composition of weak necessity modals. In Jacqueline Guéron & Jacqueline Lecarme (eds.), <u>Time and modality</u>, 115–141. Springer.
- Frank, Anette. 1996. <u>Context Dependence in Modal Constructions</u>. Ph.D. thesis, IMS Stuttgart.
- Gillies, Anthony S. 2012. Expectation modals. Ms., Rutgers University.
- Horty, John F. 2007. Reasons as defaults. Philosophers' Imprint 7(3). 1-28.
- Horty, John F. 2012. <u>Reasons as defaults</u>. Oxford University Press.
- Kratzer, Angelika. 1981. The notional category of modality. In H. J. Eikmeyer & H. Rieser (eds.), <u>Words, worlds, and contexts. new approaches in word</u> <u>semantics</u>, 38–74. Berlin: de Gruyter.
- Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), <u>Semantics: An international handbook of contemporary research</u>, 639–650. Berlin: de Gruyter.
- Kratzer, Angelika. 2010. <u>Papers on Modals and Conditionals</u>. Oxford University Press.
- Lewis, David. 1981. Ordering semantcs and premise semantics for counterfactuals. Journal of Philosophical Logic 10. 217–234.
- Makinson, David. 1994. General patterns in nonmonotonic reasoning. In Dov Gabbay, C. J. Hogger & J. A. Robinson (eds.), <u>Handbook of logic in artifical</u> <u>intelligence and logic programming</u>, vol. 3, 35–110. Oxford University Press.
- Reiter, Raymond. 1980. A Logic for Default Reasoning. <u>Artificial Intelligence</u> 13. 81–132.
- Zvolenszky, Zsófia. 2002. Is a possible-worlds semantics of modality possible? A problem for kratzer's semantics. <u>SALT</u> 12.