Dynamic Thoughts on Ifs and Oughts

Malte Willer

1 The Plot

Iffy oughts figure prominently in a variety of paradoxes of deontic logic: Chisholm's (1963) *Contrary-to-Duty Paradox*, Forrester's (1984) *Gentle Murder Paradox*, as well as the recent *Miners Paradox* from Kolodny and MacFarlane (2010). A satisfying response to these paradoxes is a non-negotiable component of any adequate semantic story about conditionals and deontic modals. I show how such a story can be told if we supplement a semantics that pays proper attention to the sensitivity of *ifs* and *oughts* to contextual information with a dynamic conception of logical consequence. The resulting framework naturally leads to a nonmonotonic logic for conditionals and deontic modals that distinguishes itself from alternative approaches currently on the market in that it preserves factual as well as deontic detachment and exploits a single natural language phenomenon—persistence failures—in its solution to the paradoxes under consideration.

The remainder of this section says a bit more about the puzzles surrounding deontic conditionals and explains why the world can use yet another discussion of these paradoxes. §2 sketches a simple semantic analysis of conditionals and deontic *ought* that is in the spirit (but not the letter) of Kratzer's seminal work on the meaning of modal expressions. The key goal of this exercise is to identify a single semantic feature of deontic *ought* that explains why the premises of our paradoxes may be jointly true. There are various notions of logical consequence that we may choose as a supplement to the resulting semantic proposal, but not all of them are created equal. In §3, I argue that a dynamic conception of validity is a particularly attractive option since it naturally gives rise to a nonmonotic logic that resolves all paradoxes without sacrificing the intuitive inference rules of factual and deontic detachment. §4 extends the dynamic proposal developed so far to account for a few additional challenges. §5 offers a few general concluding remarks.

1.1   The Miners Paradox

Kolodny and MacFarlane consider the following scenario.[1] Ten miners are trapped either in shaft A or in shaft B, but we do not know which one. Water threatens to flood the shafts. We only have enough sandbags to block one shaft but not both. If one shaft is blocked, all of the water will go into the other shaft, killing every miner inside. If we block neither shaft, both will be partially flooded, killing one miner.

Action	if miners in A	if miners in B
Block A	All saved	All drowned
Block B	All drowned	All saved
Block neither shaft	One drowned	One drowned

Lacking any information about the miners' exact whereabouts, it seems right to say that

(1)   We ought to block neither shaft.

However, we also accept that

(2)   If the miners are in shaft A, we ought to block shaft A,

(3)   If the miners are in shaft B, we ought to block shaft B.

But we also know that

(4)   Either the miners are in shaft A or they are in shaft B.

And (2)-(4) seem to entail

(5)   Either we ought to block shaft A or we ought to block shaft B,

which contradicts (1).[2] Thus we have a paradox.

Kolodny and MacFarlane argue that an adequate response to the puzzle must reject the validity of modus ponens for deontic conditionals. Without it, the conditional obligations articulated by (2) and (3) do not detach and thus we block the classical derivation of the paradoxical conclusion (5) from (2)-(4). There is some reason to think that rejecting modus ponens is not necessary even if one agrees with Kolodny and MacFarlane's general strategy of blocking the argument on its obvious logical form, but that aside it is uncontroversial that at least some semantic creativity will be required to make iffy oughts resistant to the miners paradox.

---

[1]See also the discussion from Parfit (1988, 2011), who credits Reagan (1980). A similar puzzle involving "better" instead of "ought" is discussed by Dreier (2009).

[2]Kolodny and MacFarlane evidently set aside the possibility of inconsistent deontic obligations. Here and throughout it will be harmless to adopt this simplification since none of the paradoxes discussed in this paper can be resolved by invoking the possibility of genuine moral dilemmas.

1.2   The Gentle Murder Paradox

Forrester observes that (6) and (7) are jointly consistent:

(6)   Jones ought not murder Smith,

(7)   If Jones murders Smith, he ought to murder Smith gently.

(6) articulates Jones's primary obligation not to murder Smith. (7) says what Jones ought to do conditional on his violating his primary obligation. Intuitively, (6) and (7) are not only jointly consistent but also compatible with the possibility that Jones in fact murders Smith. But suppose that

(8)   Jones murders Smith.

(7) and (8) entail

(9)   Jones ought to murder Smith gently,

which hardly seems co-tenable with (6). Where did we go wrong?

Forrester thinks of the paradox as a problem for the Inheritance principle that *ought* is closed under logical entailment. Inheritance allows us to derive (10), which contradicts (6) and follows from (9) since even the gentlest of murders must count as a murder:

(10) Jones ought to murder Smith.

Not everyone thinks that Forrester has a conclusive case against Inheritance here: perhaps we can retain the principle and simply point to ambiguities in the logical structure of (9) to block to inference of (10).[3] But regardless of what one wants to say about this issue, blocking the step from (9) to (10) does not resolve the conflict between (6) and (9). Forrester must admit that (6) forbids Jones to murder Smith *gently or otherwise*: even if one thinks that *ought* is not reliably closed under logical entailment, it should be uncontroversial that (6) entails

(11) Jones ought not murder Smith gently,

which now contradicts (9). The upshot: Forrester's paradox is first and foremost a puzzle about the inferences we are inclined to draw from iffy oughts.

1.3   Chisholm's Paradox

Forrester's paradox is at its heart a puzzle about deontic conditionals that articulate contrary-to-duty obligations. Chisholm's paradox crucially involves such conditionals as well. Consider the following:

(12) Jones ought to go to the aid of his neighbors,

---

[3]This alternative to Forrester's diagnosis is explored by Castañeda (1985, 1986) and Sinnott-Armstrong (1985). Goble (1991) shows, conclusively I think, that their strategy does not generalize to cover variants of the gentle murder paradox, and endorses Forrester's solution.

(13) If Jones goes to the aid of his neighbors, then he ought to tell them he is coming,

(14) If Jones does not go to the aid of his neighbors, then he ought not tell them he is coming,

(15) Jones does not go to the aid of his neighbors.

It seems that (12)-(15) are consistent and that none of these statements logically implies any other one. Chisholm observes that von Wright's (1951) classical deontic logic violates at least one of these constraints regardless of how we formalize (12)-(15) but there is a more general reason for being interested in the case he describes. (12) and (13) intuitively imply:

(16) Jones ought to tell his neighbors that he is coming.

But it also makes perfect sense to think that (14) and (15) imply

(17) Jones ought not tell his neighbors that he is coming,

which contradicts (16). So it seems that we need to deny at least one of the two types of inferences—often labeled "deontic detachment" and "factual detachment", respectively—which is too bad since both are very plausible.[4]

1.4 Interlude

The literature on deontic paradoxes is, to say the least, extensive. In particular, Chisholm's paradox is blessed with a large variety of proposed solutions, and there is a rapidly growing literature on Kolodny and MacFarlane's puzzle about deontic conditionals. The resulting proposals are too sophisticated to be effectively dismantled here, but let me articulate two assumptions that drive the upcoming proposal and, taken together, distinguish it from the alternatives that are currently on the market. None of these assumptions is beyond dispute but everybody, I think, can agree that they gain some support from general considerations about explanatory power and theoretical simplicity.

First, I assume that factual and deontic detachment are worth preserving. We commonly rely on factual detachment to arrive at practical conclusions from hypothetical imperatives; without it, it is hard to see how conditional obligations could have any force in everyday practical reasoning. Deontic detachment is important as well, since it allows us to derive certain obligations from others. The best explanation of our everyday inferential practices, I submit, is to supplement whatever turns out to be our best semantics for deontic conditionals with a notion of logical consequence that predicts the validity of factual and deontic detachment. Frameworks relying on notions of logical consequence that fail to do so may still illuminate something very important about the meaning and logic of deontic conditionals, but they face an explanatory burden—accounting for the way we actually

---

[4]See Loewer and Belzer (1983) and references therein for discussion. The labels "factual detachment" and "deontic detachment" go back to Greenspan (1975).

proceed in practical reasoning—that is avoided if the validity of factual and deontic detachment is a direct outcome of how our semantics for *ifs* and *oughts* interacts with the underlying notion of logical consequence.

Second, I assume that, all things being equal, it is a strength of a theory if it can offer a uniform solution to the paradoxes under consideration, in the sense that it exploits a single natural language phenomenon in resolving our problems. It is conceivable that the puzzles about conditional obligations have not much in common: after all, both Chisholm's and Forrester's paradox differ from the miners paradox in that they center on how we reason about contrary-to-duty obligations, and so it is possible that the right strategy for resolving the former does not generalize to address what goes wrong in the latter case. But while this possibility cannot be dismissed out of hand, it would be more attractive if we could identify a deeper commonality between the paradoxes under consideration.

The general remarks I have made will conspire in the upcoming story about deontic conditionals. What ties the paradoxes under consideration together, on my view, is the semantic fact that deontic *ought* fails to be persistent. Very roughly, deontic *ought* is sensitive to what is taken for granted in discourse and reasoning, and specifically deontic truths may be defeated through additional information. Kratzer (1991) exploits persistence failures in her analysis of deontic conditionals articulating contrary-to-duty obligations, but her framework is not quite flexible enough to extend this strategy to the miners paradox. I will later explain why doing so does not threaten the core ideas of Kratzer's proposal. But even if this is right, there remains the question whether we can resolve the puzzles about deontic conditionals while doing justice to our everyday inferential practices. I demonstrate that this is possible indeed if we supplement our semantics for *ifs* and *oughts* with a dynamic conception of validity. On this conception, persistence failures immediately translate into monotonicity failures at the level of logical consequence, and this allows us to preserve the validity of factual and deontic detachment while at the same time avoiding any unwelcome contradictions. Let me conclude this section with a few additional remarks about the upcoming proposal.

As I indicated before, my goal is to arrive at a framework that is general enough to take care of all our other paradoxes and preserves Kratzer's basic insights into the semantics of modality. I will thus set aside proposals that are based on a wide scope analysis of deontic conditionals or follow von Wright (1956) in using a primitive binary conditional connective to represent conditional obligations.[5] I will also set aside the possibility of challenging one or more premises of the arguments under consideration since this strategy is not promising for the miners paradox and it is hard to deny that (6)-(8) and (12)-(15) are at least consistent.[6]

---

[5]Wide scope analyses treat the deontic modal as taking scope over a material conditional in a sentence like "If Jones goes to the aid of his neighbors, then he ought to tell them he is coming" and are thus hard to square with the insight that conditional antecedents are best understood as restrictors of modals. Notice also that wide-scoping does not help with Chisholm's paradox; furthermore, it fails to be a promising strategy for resolving the miners paradox (see Kolodny and MacFarlane (2010, Sect. III.1)). Thomason (1981a) is an early critic of the use of primitive conditional connectives in representing conditional obligations.

[6]The possibility of rejecting soundness while preserving validity in the case of the miners paradox is critically discussed by Kolodny and MacFarlane (2010, Sect. I.1) and also Charlow

It is sometimes suggested that many deontic paradoxes can be resolved by paying close attention to the times of the obligations. The basic idea from Thomason (1981a,b) is that obligations depend on what possibilities are presumed to be open in context and since such possibilities change over time, so does what ought to be done.[7] The first bit of the suggestion—that what ought to be done varies with what counts as settled in deliberation—sounds right and will in fact be the starting point for the upcoming story about *ifs* and *oughts*. Jones ought to go help his neighbors and also tell them that he is coming, and thus ought not tell them that he is not coming. But assume that Jones does not go: then in light of *that* information, it is no longer true that he ought to tell them that he is coming; rather, he should at least let them know that he is not coming. The use of deontic *ought* that is prominent here is what Thomason calls *deliberative*—the question is what ought to be done taking the facts as given—and it contrasts with *non-deliberative* uses in which the facts themselves may be subject to evaluation.[8] I shall work here under the assumption that Thomason's distinction is real, and the hypothesis that I want to pursue here is that all the paradoxes we have considered above involve deliberative uses of deontic *ought*. A semantic solution to these paradoxes must then be sensitive to the flow of information in discourse and reasoning but it will not do, as Thomason suggests, to tie changes in what is settled to progression in time: possibilities in deliberation may shift without corresponding shifts in matters of fact, as in hypothetical reasoning. Taking a close look at the interaction between tense and modality is important but what matters for a solution to our paradoxes is the dynamics of information flow rather than the dynamics of time.[9,10]

2 Basics

The goal of this paper is to demonstrate that a dynamic perspective on discourse and reasoning offers an attractive solution to our puzzles about *ifs* and *oughts*. But such a solution must start with a reasonable semantics for conditionals and deontic modals that explains why the premises of the paradoxes under consideration

---

(forthcoming, Sect. 2). There is a broad consensus in the literature that (6)-(8) and (12)-(15) are consistent, but see Arregui (2010) for a dissenting voice.

[7]See also, among others, Åqvist and Hoepelman (1981), van Eck (1982), Feldman (1986, 1990), and Loewer and Belzer (1983, 1986) for combinations of deontic and tense logic.

[8]Schroeder (2011) argues that one must distinguish between a deliberative sense of *ought* that relates agents to actions and an evaluative sense that does not. The distinction he has in mind does not coincide with Thomason's and so the question whether Schroeder is right, though important, need not detain us here.

[9]The idea that deontic *ought* is sensitive to what is taken for granted in discourse and reasoning is also present in the discussion from Prakken and Sergot (1996, 1997), who talk about obligations as contextual obligations. They also present a timeless Chisholm scenario—their "fence example"—that puts pressure on views treating tense as the silver bullet against all paradoxes deontic.

[10]For reasons of space, I set aside the question what the upcoming proposal has to say about non-deliberative interpretations of deontic *ought*. Dealing with this issue requires, I think, that the proposal is extended so that deontic modals may quantify over sets of possible other than those that are compatible with what is taken for granted in discourse and reasoning. While this can be done, it raises a number of challenges that would need to be addressed in a longer version of the present draft.

are consistent. I follow here Yalcin (2007, 2011) in assigning to epistemic modals truth-values relative to a possible world and a separate informational parameter that keeps track of what is taken for granted in discourse and reasoning (a set of possible worlds). Kolodny and MacFarlane (2010) generalize this strategy to cover deliberative interpretations of deontic *ought* by treating them as specifications of informational modals:

$$\llbracket \Box_f \phi \rrbracket^{w,i} \text{ is true iff for all } w' \in f(i), \llbracket \phi \rrbracket^{w',i} \text{ is true}$$
$$\llbracket \Diamond_f \phi \rrbracket^{w,i} \text{ is true iff for some } w' \in f(i), \llbracket \phi \rrbracket^{w',i} \text{ is true}$$

Here f is a selection function mapping information states to modal quantifier domains. Different modals require different selection functions. The modal quantifier domain of epistemic modals, for instance, is determined by an epistemic selection function e, which is just the identity function. Deontic *ought* is a universal quantifier ranging over a set of possible worlds that is provided by a deontic selection function d.

Kratzer's treatment of conditional antecedents as modifying the modal base in light of which we evaluate the consequent is captured as follows: in order to determine the truth value of a conditional we first *strengthen* the information in light of which we reason with the antecedent, and then evaluate the consequent in that light. The process of strengthening is defined as follows:

Strengthening The result of strengthening i with ϕ, $i + \phi$, is defined as the intersection of i and $\llbracket \phi \rrbracket^i$, i.e. $i + \phi = i \cap \{w \colon \llbracket \phi \rrbracket^{w,i} \text{ is true}\}$

An informational parameter i affects the proposition expressed by ϕ, and ϕ in turn strengthens i by ruling out all possibilities that are incompatible with the proposition expressed by ϕ in i. The semantic analysis of conditionals is then:

$$\llbracket \phi \Rightarrow \psi \rrbracket^{w,i} \text{ is true iff } \llbracket \psi \rrbracket^{w,i+\phi} \text{ is true}$$

Here the assumption is that conditional consequents contain a modal operator. By default, the modal is a (perhaps implicit) epistemic necessity operator but in iffy oughts the modal is deontic.

The semantics presented here is in a position to explain why the premises in our paradoxes are consistent but doing so requires some assumptions about the role that information strengthening plays for truth at a point of evaluation. Borrowing some useful terminology from Gillies (2010), the question is to what extent truth at a point is persistent:

Persistence

1. ϕ is *t-persistent* iff for all w, i and $i' \subseteq i$: if $\llbracket \phi \rrbracket^{w,i}$ is true then $\llbracket \phi \rrbracket^{w,i'}$ is true

2. ϕ is *f-persistent* iff for all w, i and $i' \subseteq i$: if $\llbracket \phi \rrbracket^{w,i}$ is false then $\llbracket \phi \rrbracket^{w,i'}$ is false

Sentences whose truth-values do not depend at all on what information is taken for granted in discourse and reasoning are, of course, both t-persistent and f-persistent.

Epistemic *might* is f-persistent but not t-persistent: additional information may eliminate all p-worlds from the relevant informational parameter, thus switching the truth-value of $\Diamond_e p$ from true to false. Epistemic *must* is t-persistent but not f-persistent: additional information may eliminate all $\neg p$-worlds from the relevant informational parameter, thus switching the truth-value of $\Box_e p$ from false to true. And of course, this is all intimately tied to the fact that our epistemic selection function is *monotonic*: for all i, $i' \subseteq i$, $e(i') \subseteq e(i)$, which is not surprising since e is just the identity function.

The hypothesis that I want to pursue here is that deontic *ought* is neither t-persistent nor f-persistent and, moreover, that we can exploit this phenomenon to explain why the premises of the paradoxes under consideration are consistent. In the framework I have outlined, we can think of an informational parameter i as a modal base and of d as selecting, from the modal base, the set of possible worlds that count as deontically ideal. Kratzer suggests that the set of deontically ideal possible worlds in the modal base is determined by an ordering source, which is modeled as a function from possible worlds to a set of propositions. Here it will do no harm to simplify things slightly and treat the ordering source as constant across possible worlds. The general hypothesis then is that $d(i)$ yields the set of elements of i that are ranked 'best' in light of the set of propositions provided by an ordering source OS.[11] The only wrinkle we must add to this story is that we now assign to sentences truth-values relative to a possible world and a separate informational parameter, and so the details of the proposal are as follows:

Deontically Ideal Worlds Consider any w, i and ordering source OS:

1. $\langle w, i \rangle \leqslant_{OS} \langle w', i' \rangle$ iff for all $[\![\phi]\!] \in OS$: if $[\![\phi]\!]^{w',i'} = 1$, then $[\![\phi]\!]^{w,i} = 1$

2. $w \in d(i)$ iff $w \in i$ and for all $v \in i$: $\langle w, i \rangle \leqslant_{OS} \langle v, i \rangle$

As in Kratzer's original framework, an ordering source imposes a ranking on the sets of points of evaluation, but points of evaluation are now world-information state pairs. Given some information state i in its role of modal base, a world w is deontically ideal, $w \in d(i)$, just in case $w \in i$ and there is no other world $v \in i$ so that $\langle v, i \rangle$ outranks $\langle w, i \rangle$ in light of the relevant ordering source.

The framework outlined here can model the key ideas of Kratzer's treatment of contrary-to-duty conditionals, and a bit more. Start with deontic conditionals articulating contrary-to-duty obligations, like the one that figures prominently in Forrester's paradox. If m stands short for "Jones murders Smith" and g stands short for "Jones murders Smith gently," we can think of the ordering source OS as follows:

$$OS = \{\neg m, m \supset g\}$$

This immediately predicts that (6) and (7) can both be true:

(6)   Jones ought not murder Smith,

---

[11]The usual disclaimer: I rely here for simplicity on the Limit Assumption from Lewis 1973.

(7)   If Jones murders Smith, he ought to murder Smith gently.

Take any i that leaves it open whether Jones murders Smith: then $[\![\Box_d\neg m]\!]^{w,i} = 1$ but also $[\![\Box_d g]\!]^{w,i+m} = 1$ and thus $[\![m \Rightarrow \Box_d g]\!]^{w,i} = 1$. This is equivalent to the classical treatment of contrary-to-duty obligations from Kratzer 1991, but it also establishes that deontic *ought* is neither t-persistent nor f-persistent: simply remember that according to what we just said, "Jones ought not murder Smith" is true given i but false (or at least not true) given $i + m$. Similarly, "Jones ought to murder Smith gently" is false given i but true given $i + m$. And of course, the persistence failures exhibited here are completely benign: they are just what we expect since our deontic selection function is defined so that it is guaranteed to be realistic and non-empty.

Realism   For all i, $d(i) \subseteq i$

Non-Emptiness   For all $i \neq \varnothing$, $d(i) \neq \varnothing$

Start with the observation that $\ulcorner\Box_d\phi\urcorner$ does not entail $\ulcorner\Box_e\neg\phi\urcorner$. Non-emptiness requires that information change leaves us with some deontically ideal worlds as long as the additional bit of information is not incompatible with what is already taken for granted; realism requires that these worlds cannot be ϕ-worlds if that additional bit of information entails $\ulcorner\neg\phi\urcorner$. This is just what is needed to make sense of the persistence failures in Forrester's paradox.

A bit more is required to maintain that persistence failures also stand behind the fact that (1) and (2) from the miners may both be true (repeated):

(1)   We ought to block neither shaft,

(2)   If the miners are in shaft A, we ought to block shaft A.

Maintaining that the *if*-clause in (2) restricts a modal with the same meaning as the one in (1) demands a deontic selection function that is "seriously information dependent" in the following sense (cf. Kolodny and MacFarlane 2010):

Serious Information Dependence   For some $i' \subseteq i$, there is a $w \in i'$ such that $w \in d(i)$ but $w \notin d(i')$

Serious information dependence requires that, on some occasions, a world is deontically ideal with respect to some state of information but fails to be so with respect to some strengthened information state that contains it. In other words, if our deontic selection function leaves no room for serious information independence, deontic truths are preserved under strengthening as long as the prejacent is compatible with the additional bit of information, which is just to say that "We ought to block neither shaft"—if true in the absence of any assumptions about the miners' whereabouts—remains true under the assumption that the miners are in shaft A.

Kolodny and MacFarlane (2010, fn. 29) suggest that contextually provided ordering sources do not allow for serious information dependence, but several authors also observe such ordering sources could in principle reflect what information is available in the scenario under consideration.[12] The specific suggestion from the literature is that the ordering source in the miners scenario may very well rank worlds in which we do not know the miners' whereabouts and block neither shaft as equal to those in which we know where the miners are and block the right shaft. This immediately predicts that "We ought to block neither shaft" is true in the miners scenario, but it does not explain why "We ought to block shaft A" seems true under the assumption that the miners are in shaft A: after all, the setup is such that we do not know where the miners are, and so the proposal predicts that we ought not block shaft A even if they are in shaft A.

In response to the previous problem, it is tempting to suggest that the deontic modals in (1) and (2) are interpreted with respect to different modal bases or ordering sources (see Dowell 2011 and von Fintel (2012); for critical discussion of this strategy, see Silk 2012). But this is not necessary once we allow for information dependence to be built into the ordering source. The following proposal does the trick:

$$OS = \{[\![\Box_e inA \supset blA]\!], [\![\Box_e inB \supset blB]\!], [\![(\Diamond_e inA \wedge \Diamond_e inB) \supset (\neg blA \wedge \neg blB)]\!] \}$$

To see what is going on here, recall that our ordering source imposes a ranking on world-information state pairs. Suppose then that in w_1 we block neither shaft, in w_2 we block shaft A, and in w_3 we block shaft B. Furthermore, assume that i_1 contains the information we have in miners scenario, let i_2 be the result of strenghtening i_1 with the information that the miners are in shaft A, and i_3 be the result of strenghtening i_1 with the information that the miners are in shaft B. It is straightforward to verify that the OS ranks the following world-information state pairs as deontically ideal: $\langle w_1, i_1 \rangle, \langle w_2, i_2 \rangle, \langle w_3, i_3 \rangle$ (recall that epistemic *might* fails to be t-persistent while epistemic *must* fails to be f-persistent). Consequently, $d(i_1) = \{w_1\}$ but $d(i_2) = \{w_2\}$ and $d(i_3) = \{w_3\}$. It follows that d is seriously information dependent in the way required to predict that the premises of the miners scenario are all true.

I do not claim that the current proposal has an *empirical* edge over the many alternative approaches to the miners paradox that are currently on the market. I am also sympathetic to the suggestion that what is deontically ideal is not determined by a simple ordering source but derived from more foundational considerations about practical reasons or rational decision-making (see the discussions by Charlow (forthcoming) and Cariani et al. (2011)). The main point of the current exercise is this. We have independent reason to believe that deontic *ought* fails to be t-persistent or f-persistent, and this feature already figures prominently in classical Kratzer-style analyses of deontical conditionals articulating contrary-to-duty conditionals. Maintaining that persistence failures are also relevant for the analysis of the miners paradox requires a seriously information dependent deontic

---

[12]The point is made by Kratzer in unpublished notes as well as by von Fintel (2012) and Silk (2012).

selection function, but it is straightforward to find a single contextually provided ordering source that delivers the result. The resulting proposal is, at a minimum, of genuine methodological interest since it makes sense of the idea that the paradoxes under consideration have something important in common: the consistency of their premises is rooted in the non-persistence of deontic *ought*.

The outlined strategy departs from Kratzer's classical analysis of deontic conditionals in that points of evaluation are world-information state pairs. There is, as far as I can tell, nothing in her proposal that prevents us from increasing the complexity of points of evaluation, and in any case we have strong independent evidence from the discussion of epistemic modals that some natural language constructions have truth-values relative to a separate informational parameter (see Yalcin 2007, 2011).[13] Most importantly, we can preserve the idea that deontically ideal possible worlds are selected from the modal base on the basis of a contextually provided ordering source. While this ordering source may contain information-sensitive propositions to provide for a seriously information dependent deontic selection function, the ordering *itself* does not vary across informational parameters and is thus is information-insensitive. Accordingly, the uniform perspective on our deontic paradoxes that I have advertised can be afforded without departing from the key insights of Kratzer's classical analysis of deontic conditionals.

Putting persistence failures into the spotlight does not only identify a commonality between the paradoxes under consideration but will also allow us to resolve the puzzles while preserving the validity of factual and deontic detachment without running into unwelcome contradictions. Unpacking this idea takes a bit of stage setting, so let me prepare the upcoming discussion with a few basic remarks about how one might think about logical consequence once we assign to sentences truth-conditions relative to a possible world and an informational parameter. Kolody and MacFarlane define logical consequence *neoclassically* as necessary preservation of truth as a possible world and with respect to some fixed informational parameter:

Neoclassical Logical Consequence $\phi_1, \ldots, \phi_n \vDash_N \psi$ iff for all w and i such that $w \in i$: if $w \in [\![\phi_1]\!]^i$ and...and $w \in [\![\phi_n]\!]^i$, then $w \in [\![\psi]\!]^i$

The proviso that $w \in i$ enforces that we only consider "proper" points of evaluation: this is to require that the informational parameter is properly *epistemic* in that it does not carry any misinformation about the index of evaluation. Logical consequence thus defined blocks the inference of (5) from (2)-(4): given some w and i such that $w \in i$ and $w \in [\![inA \Rightarrow \Box_d blA]\!]^i$, $w \in [\![inB \Rightarrow \Box_d blB]\!]^i$, and $w \in [\![inA \lor inB]\!]^i$, we know that both $w \in [\![\Box_d blA]\!]^{i+inA}$ and $w \in [\![\Box_d blB]\!]^{i+inB}$. But this, given the lack of f-persistence of deontic *ought*, is compatible with both $w \notin [\![\Box_d blA]\!]^i$ and $w \notin [\![\Box_d blB]\!]^i$ (and even if $i \subseteq [\![inA \lor inB]\!]^i$).

A neoclassical supplement to our semantics for *ifs* and *oughts* successfully blocks the miners paradox on its obvious form. Notice, however, that modus ponens fails

---

[13]I prefer a dynamic treatment of epistemic modals in the spirit of Veltman (1996), but the issue does not matter for current purposes. Yalcin's proposal is, from an abstract perspective, a static version of Veltman's and thus the best that truth-conditional semantics has to offer.

11

to be valid for deontic conditionals. For suppose that $w \in [\![\phi]\!]^i$ and $w \in [\![\phi \Rightarrow \psi]\!]^i$ are true. It follows that $w \in [\![\psi]\!]^{i+\phi}$ is true, but that is compatible with $w \notin [\![\psi]\!]^i$ in case ψ fails to be f-persistent, which is just what we have if ψ is a deontically modalized sentence. As a result, neoclassical consequence *by itself* fails to justify our everyday inferential practice of relying on factual detachment. This is no knock-down argument against thinking of validity in a neoclassical fashion, but it raises the question whether we can find an alternative that combines the virtues of neoclassical consequence with an additional bit of explanatory power.

To see what such an alternative could look like, start with the observation that we can distinguish the notion of truth at a point of evaluation from the one of *acceptance*:

Acceptance ϕ is accepted in information state i iff for all $w \in i$: $[\![\phi]\!]^{w,i}$ is true.

Instead of thinking about validity neoclassically we might say that an inference is valid just in case whenever its premises are accepted in light of some informational parameter i, so is its conclusion. An *informational* conception of logical consequence as proposed by Yalcin (2007) captures this idea:

Informational Consequence $\phi_1, \ldots, \phi_n \vDash_I \psi$ iff for all i: if ϕ_1 is accepted in i and... and ϕ_n is accepted in i, then ψ is accepted in i

Yalcin shows that informational consequence is not without appeal and so it is reasonable to ask whether this notion of validity also has something useful to say about the miners paradox. To streamline the upcoming discussion, it is helpful to establish that informational consequence is equivalent to what Kolodny and MacFarlane label "quasi-validity:"

Quasi-Validity $\phi_1, \ldots, \phi_n \vDash_Q \psi$ iff $\Box_e\phi_1, \ldots, \Box_e\phi_n \vDash_N \psi$

An argument is quasi-valid just in case its conclusion follows neoclassically whenever its premises are epistemically necessary.

Fact 1 $\phi_1, \ldots, \phi_n \vDash_I \psi$ iff $\phi_1, \ldots, \phi_n \vDash_Q \psi$[14]

From left to right: suppose that $\phi_1, \ldots, \phi_n \nvDash_Q \psi$, then for some w and i such that $w \in i$: $w \in [\![\Box_e\phi_1]\!]^i$ and...and $w \in [\![\Box_e\phi_n]\!]^i$ but $w \notin [\![\psi]\!]^i$. So for all $w' \in i$: $[\![\phi_1]\!]^{w',i}$ is true and...and $[\![\phi_n]\!]^{w',i}$ is true. But since $w \in i$, there is some $w' \in i$ such that $[\![\psi]\!]^{w',i}$ is false. So there is some i such that ϕ_1 is accepted in i and... and ϕ_n is accepted in i but ψ is not accepted in i, and thus $\phi_1, \ldots, \phi_n \nvDash_I \psi$. From right to left: suppose that $\phi_1, \ldots, \phi_n \nvDash_I \psi$, then there is some i such that ϕ_1 is accepted in

---

[14]I follow here Yalcin in thinking of informational consequence and thus quasi-validity as a serious alternative to the neoclassical approach. In contrast, Kolodny and MacFarlane treat quasi-validity as a tool for explaining how certain inferences that are strictly speaking invalid may nonetheless be reasonable in certain situations. Their idea is also present in Stalnaker (1975) and I will come back to it at a later stage.

i and. . .and ϕ_n is accepted in i but ψ is not accepted in i. So for all w, $w \in [\![\Box_e\phi_1]\!]^i$ and. . .and $w \in [\![\Box_e\phi_n]\!]^i$. Moreover, since ψ is not accepted in i, it follows that for some $w \in i$, $w \notin [\![\psi]\!]^i$. So for some w and i such that $w \in i$: $w \in [\![\Box_e\phi_1]\!]^i$ and. . .and $w \in [\![\Box_e\phi_n]\!]^i$ but $w \notin [\![\psi]\!]^i$, and thus $\phi_1, \ldots, \phi_n \nvDash_Q \psi$.

As Kolodny and MacFarlane observe, modus ponens is quasi-valid. Assume that $w \in [\![\Box_e\phi]\!]^i$ and $w \in [\![\Box_e(\phi \Rightarrow \psi)]\!]^i$. Then $w \in [\![\phi \Rightarrow \psi]\!]^i$ and thus $w \in [\![\psi]\!]^{i+\phi}$. But since $w \in [\![\Box_e\phi]\!]^i$, $i + \phi = i$ and thus $w \in [\![\psi]\!]^i$, as required. But the inference of (5) from (2)-(4) is *not* quasi-valid. To see this, notice that the truth-value of informational modals and conditionals entirely depends on the informational parameter. Such sentences are thus locally invariant in the following sense:

Local Invariance ϕ is locally invariant iff for all i: $[\![\phi]\!]^i = \varnothing$ or $[\![\phi]\!]^i = W$

Whenever ϕ is locally invariant, $w \in [\![\phi]\!]^i$ if and only if $w \in [\![\Box_e\phi]\!]^i$. In such cases strengthening i with ϕ idles, i.e. $i + \phi = i$ and thus it is guaranteed that $w \in [\![\phi]\!]^{i+\phi}$ as well. The simple observation is that the inference of (5) from (2)-(4) is quasi-valid just in case $\Box_e(inA \vee inB), inA \Rightarrow \Box_d blA, inB \Rightarrow \Box_d blA \vDash_N \Box_d blA \vee \Box_d blB$. But we already know that given the lack of f-persistence of deontic *ought*, the falsity of (5) is compatible with the truth of (2) and (3) even in case (4) is an epistemic necessity.

It follows from this and **Fact 1** that informational consequence avoids the miners paradox while preserving the validity of modus ponens. This is important because it shows that the first exhibit from our triad of puzzles about iffy ought does not present a compelling case against factual detachment: rejecting modus ponens for deontic conditionals is not necessary to block the derivation of our paradoxical conclusion. Factual detachment may or may not be a reliable rule of inference, but it takes more than the miners paradox to show that it is not.[15] More importantly, however, is the general lesson that deciding on a semantics for conditionals and deontic modals does not settle the question which notion of logical consequence makes best sense of the role of iffy oughts in discourse and reasoning. Neoclassical consequence is not the only game in town: at a minimum, we know that informational consequence offers an alternative to defining validity as necessary preservation of truth at a proper index of evaluation.

At this stage, one may worry that there is not much to choose between neo-classical logical consequence and its informational alternative: both offer an escape route from the miners paradox, and the remaining differences are too murky to give rise to a substantial philosophical debate. I agree that there is not much to choose here but this is because none of the notions of logical consequence that we have considered so far fully accounts for how we actually proceed in deontic discourse and reasoning. Let me explain.

---

[15]And it also takes more than the observation that modus tollens is unreliable for deontic conditionals, which Kolodny and MacFarlane take as evidence against modus ponens since it figures prominently in the classical derivation of modus tollens. Informational consequence does not support modus tollens: simply notice that "If the miners are in shaft A, we ought to block shaft A" and "It is not the case that we ought to block shaft A" are accepted at i while "The miners are not in shaft A" is not. Dynamic logical consequence, as I have demonstrated elsewhere, also accounts for the invalidity of modus tollens while preserving the validity of modus ponens.

3   Toward Dynamic Consequence

Neoclassical as well as informational consequence offer an escape route from the miners paradox. This is not surprising since our semantics alone already explains how (1)-(4) can be true while (5) is false. Things get more complicated once we look at Forrester's paradox. The problem, remember, starts with the observation that (6) and (7) are jointly consistent:

(6)   Jones ought not murder Smith $\Box_d \neg m$

(7)   If Jones murders Smith, he ought to murder Smith gently $m \Rightarrow \Box_d g$

Furthermore, (6) and (7) are compatible with the assumption that (8) is true yet (7) and (8) together entail (9), which contradicts (6):

(8)   Jones murders Smith m

(9)   Jones ought to murder Smith gently $\Box_d g$

To avoid a paradox, the neoclassical approach must insist that (9) does not follow from (7) and (8), but I have already explained why this leaves leaves something important about our everyday inferential practices unexplained. Informational consequence, in contrast, avoids the charge of dismissing factual detachment for deontic conditionals: whenever a conditional and its antecedent are accepted in an information state, so is its conclusion, and this is exactly what it takes for modus ponens to be reliable on the informational conception of logical consequence. However, informational consequence does nothing to avoid the gentle murder paradox since it is committed to (6)-(8) being in fact inconsistent: (9) follows from (7) and (8) by modus ponens, but unless Jones is subject to inconsistent obligations we will not find an information state that settles both "Jones ought not murder Smith" and "Jones ought to murder Smith gently" as true. Unlike its neoclassical alternative, informational consequence thus preserves the intuitive inferential properties of deontic conditionals, but it runs into the other horn of the dilemma: the premises of Forrester's paradox are predicted to be inconsistent, which is not the result we want.

Supplementing our semantics for *ifs* and *oughts* with a neoclassical or informational conception of logical consequence works fine for the miners paradox but leaves something to be desired when it comes to the gentle murder paradox. This is interesting but the more important question is *why* the notions of validity that we have considered so far fail to deliver an adequate solution to Forrester's puzzle. Start with the observation that both accounts can be interpreted as working under the assumption that the premises of an argument restrict the set of proper indices in light of which its conclusion is evaluated. Precisely, if Σ is the set of proper indices $\langle w, i \rangle$ such that $w \in i$, we can make the following observations about neoclassical and informational consequence:

Fact 2 $\phi_1, \ldots, \phi_n \vDash_N \psi$ iff $\Sigma \cap [\![\phi_1]\!] \cap \ldots \cap [\![\phi_n]\!] \subseteq [\![\psi]\!]$

Fact 3 $\phi_1, \ldots, \phi_n \vDash_I \psi$ iff $\Sigma \cap [\![\Box_e\phi_1]\!] \cap \ldots \cap [\![\Box_e\phi_n]\!] \subseteq [\![\psi]\!]$

Neoclassical and informational consequence thus have a purely *eliminative* conception of the role of premises in discourse and reasoning. This resonates well with the common conception of logical consequence as necessary preservation of truth at a point of evaluation but also brings in its wake a commitment to logical consequence being *monotonic* in the following sense:

Monotonicity   If $\phi_1, \ldots, \phi_n \vDash \psi$, then $\phi_1, \ldots, \phi_n, \phi_{n+1} \vDash \psi$

A commitment to monotonicity does not leave us much choice when it comes to reasoning about gentle murders: we must either deny that (9) follows from (7) and (8)—which is just what the neoclassicist does—or accept that (6)-(8) are inconsistent, which is just what we have if we adopt an informational approach to logical consequence. At a minimum, this observation shows that the neoclassical and informational problems with Forrester's paradox are not coincidental but stem from a deep shared theoretical commitment. As a bonus, however, it also suggests that there is much to be gained from rejecting this commitment.[16]

Challenging monotonicity promises to remove the air of paradox from Forrester's puzzle since it allows us to say that even if (9) follows from (7) and (8) and contradicts (6), there is no contradiction since at no point in the argument are we committed to the conjunction of "Jones ought not murder Smith" and "Jones ought to murder Smith gently." There is a point in the argument at which we hold that Jones ought not murder Smith, gently or otherwise. And under the supposition that Jones murders Smith, we hold that Jones ought to murder Smith gently. But under that supposition we *no longer* hold that Jones ought not murder Smith—additional information in discourse and reasoning may defeat prior deontic truths. To be precise, let Γ be the set of sentences consisting of (6) and (7), i.e. assume that $\Gamma = \{\Box_d \neg m, m \Rightarrow \Box_d g\}$. The hypothesis is that while $\ulcorner \Box_d \neg m \urcorner$ follows trivially from Γ, it no longer follows from Γ if strengthened with the additional premise that Jones murders Smith: $\Gamma \vDash \Box_d \neg m$ yet $\Gamma, m \nvDash \Box_d \neg m$. So even if $\Gamma, m \vDash \Box_d g$ and $\Box_d g \wedge \Box_d \neg m \vDash \bot$, (6)-(8) do not entail a contradiction since the right logic for *ifs* and *oughts* fails to be monotonic.

A nonmonotonic perspective on *ifs* and *oughts* promises an attractive escape route from Forrester's paradox. This is important and deserves further elaboration but it is just as crucial to notice that we already have good reason to expect that a logic for *ifs* and *oughts* fails to be monotonic. Monotonicity requires that additional information in discourse and reasoning preserves what has already been established, but we already know that deontic *ought* fails to be t-persistent and thus that information strengthening fails to preserve deontic truths. Persistence failures alone, of course, are not sufficient to arrive at a nonmonotonic conception of logical

---

[16]For previous discussions of the role of nonmonotonicity in deontic discourse and reasoning, see, e.g., Bonevac (1998), Horty (2003, 2007), as well as the papers collected in Nute (1997). As demonstrated momentarily, the story told here differs from previous nonmonotonic narratives in details, scope, and general aim. To cut a long story short: a nonmonotonic perspective on deontic *ought* is independently motivated and best elaborated by combining an intuitive semantics for deontic modals with a (no less intuitive) dynamic perspective on discourse and reasoning.

consequence: neoclassicism as well as its informational alternative define validity as preservation of truth at a possible world and in light a *fixed* informational parameter. Processing a premise in discourse and reasoning, then, does not strengthen the informational parameter in light of which subsequent claims are evaluated and thus persistence failures are barred from having any nonmonotonic effects in logical reasoning. Such static evaluation procedures for arguments stand in stark contrast with the familiar very *dynamic* evaluation procedure for conditionals: to determine the truth-value of a conditional with respect to some informational parameter i, one first strengthens i with the information carried by the antecedent and then evaluates the consequent in light of the output informational parameter $i + \phi$. On the other hand, the semantics of conditionals also suggests a thoroughly dynamic perspective on discourse and reasoning. Let me explain.

Earlier we exploited persistence failures in our semantic analysis of deontic conditionals, and I have argued that this is attractive because it allows us to diagnose a commonality between the paradoxes under consideration. I now suggest that we exploit persistence failures once more to arrive at a nonmonotonic conception of logical consequence. Doing so requires that processing a premise in an argument does not only rule out certain possibilities but also strengthens the informational parameter in light of which subsequent claims are evaluated. This gives us a dynamic conception of logical consequence:

Dynamic Logical Consequence $\phi_1, \ldots, \phi_n \vDash_D \psi$ iff for all w and i such that $w \in i$: if $w \in \llbracket \phi_1 \rrbracket^i$ and... and $w \in \llbracket \phi_n \rrbracket^{i + \cdots + \phi_{n-1}}$, then $w \in \llbracket \psi \rrbracket^{(i + \cdots + \phi_{n-1}) + \phi_n}$[17]

Dynamic logical consequence preserves the idea of logical consequence as necessary preservation of truth but in addition we keep track of the effects that premises have on the relevant informational parameter.

Dynamic validity entails informational validity since whenever an argument is dynamically valid, it will be so in the special case in which all its premises are accepted in the relevant informational operator and thus strengthening idles:

Fact 4   If $\phi_1, \ldots, \phi_n \vDash_D \psi$ then $\phi_1, \ldots, \phi_n \vDash_I \psi$

This is good news since it immediately shows that dynamic logical consequence is alike to the notions of logical consequence that we have considered so far in that it blocks the miners paradox on its obvious logical form. However, it also differs from its static competitors in several important ways. Unlike its neoclassical alternative, dynamic logical consequence preserves the validity of modus ponens:

Fact 5 $\phi \Rightarrow \psi, \phi \vDash_D \psi$

Modus ponens is dynamically valid just in case for all w and i such that $w \in i$: if $w \in \llbracket \phi \Rightarrow \psi \rrbracket^i$ and $w \in \llbracket \phi \rrbracket^{i + (\phi \Rightarrow \psi)}$, then $w \in \llbracket \psi \rrbracket^{(i + (\phi \Rightarrow \psi)) + \phi}$. Since conditionals

---

[17]See also Gillies (2009); for some for some alternative dynamic conceptions of logical consequence, see Veltman (1985, 1996).

are locally invariant, this condition is equivalent to the requirement that whenever $w \in [\![\phi \Rightarrow \psi]\!]^i$ and $w \in [\![\phi]\!]^i$, then $w \in [\![\psi]\!]^{i+\phi}$. But our semantics for conditionals guarantees that whenever $w \in [\![\phi]\!]^i$, then $w \in [\![\psi]\!]^{i+\phi}$, which is just to say that modus ponens is dynamically valid. And this is not a purely formal result but captures something important about the connection between conditionals and information in discourse and reasoning: if a conditional is true in light of some body of information i, its consequent is true in light of the result of strengthening i with its antecedent. But of course, processing a premise ϕ in discourse and reasoning crucially involves strengthening the contextual information with ϕ. The validity of modus ponens, then, is a natural outcome of the concord between the evaluation procedure for conditionals and the way information is processed in discourse and reasoning. It is only when we postulate a mismatch between these two—as the neoclassicists do—that modus ponens becomes something mysterious.

We have seen that dynamic validity is alike to its informational alternative in that it licenses factual detachment and thus accounts for the intuition that (9) follows from (7) and (8). At the same time, it avoids the problems of informational consequence by preserving the intuition that (6)-(8) are consistent. The first crucial observation is that dynamic logical consequence avoids a purely eliminative conception of the role of premises in discourse and reasoning and thus a commitment to monotonicity. To get this into clearer view, it is helpful to define an update function [.] on sets of proper world-information state pairs, as follows:

Update Consider any ϕ and $\sigma \subseteq \Sigma$. An update on a set of proper indices is a function $[.] : \mathcal{P}(\Sigma) \to \mathcal{P}(\Sigma)$ is defined as follows:

$$\sigma[\phi] = \{\langle w, i + \phi \rangle : \langle w, i \rangle \in \sigma \wedge w \in [\![\phi]\!]^i\}$$

Updating a set of proper indices σ with ϕ proceeds by first eliminating all elements of σ in light of which ϕ is false, and then strengthening the informational parameters of the remaining elements with ϕ. Notice that [.] is well-defined since whenever $w \in i$ and $w \in [\![\phi]\!]^i$, then $w \in i + \phi$.

Dynamic logical consequence can then be captured as follows:

Fact 6 $\phi_1, \ldots, \phi_n \vDash_D \psi$ iff $\Sigma[\phi_1] \ldots [\phi_n] \subseteq [\![\psi]\!]$

The crucial observation is that updating a set of proper indices is not guaranteed to be a purely eliminative affair since it is not guaranteed that whenever $\langle w, i \rangle$ is an element of σ, so is $\langle w, i + \phi \rangle$. Dynamic logical consequence, then, is not wedded to a purely eliminative conception of premises in discourse and reasoning. And noneliminative effects are a real possibility whenever some formulas of our language fail to be t-persistent: whenever ϕ fails to be t-persistent, we can expect that there are $w, i, i' \subseteq i$ such that $\langle w, i \rangle \in \Sigma[\phi]$ but $\langle w, i' \rangle \notin \Sigma[\phi]$, and thus updating $\Sigma[\phi]$ with another formula ψ may very well re-introduce indices eliminated by the update with ϕ. For instance, let $i = W$ and consider $w \in [\![p]\!]^i$: then $\langle w, i + p \rangle \notin \Sigma[\Diamond_e \neg p]$. But since $\langle w, i \rangle \in \Sigma[\Diamond_e \neg p]$ and $w \in [\![p]\!]^{i + \Diamond_e \neg p}$, $\langle w, i + p \rangle \in \Sigma[\Diamond_e \neg p][p]$ and so updating $\Sigma[\Diamond_e \neg p]$ with p has a noneliminative effect.

The observed noneliminative effects also illustrate that dynamic logical consequence fails to be monotonic, as desired. It follows from the local invariance of *might* that $\Diamond_e\neg p \vDash_D \Diamond_e\neg p$, but our earlier observation that $\langle w, i{+}p\rangle \in \Sigma[\Diamond_e\neg p][p]$ shows that $\Diamond_e\neg p, p \nvDash_D \Diamond_e\neg p$, which just delivers the following result:

Fact 7 Dynamic logical consequence is nonmonotonic.

And this, bear in mind, not by fiat but by the interaction between a dynamic evaluation procedure for arguments with independently motivated persistence failures for informational modals.

All of this allows us to endorse factual detachment without regrets. Specifically, a dynamic supplement to our semantics for *ifs* and *oughts* resolves the stubborn paradox of gentle murder. Start again with (7) and (8), i.e. assume that $\Gamma = \{\Box_d\neg m, m \Rightarrow \Box_d g\}$. Then trivially $\Gamma \vDash_D \Box_d\neg m$ and $\Gamma, m \vDash_D \Box_d g$ by modus ponens. Also, $\Box_d\neg m \wedge \Box_d g \vDash_D \bot$ since gentle murders are murders and given the non-emptiness constraint on our deontic selection function. Nonetheless, there is no paradox—(6)-(8) are dynamically consistent—for even though "Jones ought not murder Smith" is true in light of our initial state of information, it is *no longer* true in light of the result of strengthening that state with the information that Jones murders Smith: if $w \in [\![\Box_d\neg m]\!]^i$, $w \in [\![m \Rightarrow \Box_d g]\!]^i$, and $w \in [\![m]\!]^i$, then $w \in [\![\Box_d g]\!]^{i+m}$ but $w \notin [\![\Box_d\neg m]\!]^{i+m}$, which is just to say that $\Gamma, m \nvDash_D \Box_d\neg m$ and thus $\Gamma, m \nvDash_D \bot$.[18]

Dynamic logical consequence preserves factual detachment for contrary-to-duty obligations without negative side effects since it pays proper attention to how information in discourse and reasoning affects contextual parameters that are relevant for evaluating *ifs* and *oughts*. Information strengthening may license factual detachment but at the same time defeat prior commitments whenever we have lack of t-persistence. These two features conspire in the dynamic solution to the gentle murder paradox: the assumption that Jones murders Smith licenses detachment of a contrary-duty-obligation while at the same time preserving consistency through defeating a conflicting primary obligation.

The upshot of the discussion so far is that a dynamic supplement to our semantics for *ifs* and *oughts* has a lot going for it. It matches the evaluation procedure for arguments against what is commonly recognized to be our best semantic take on the evaluation procedure for conditionals. It licenses factual detachment for deontic conditionals and thus preserves the intuitive role of conditional obligations in discourse and reasoning. It avoids the gentle murder paradox by a taking a nonmonotonic perspective on deontic discourse and reasoning. And we can achieve all this with a very simple move: all that is required is that our notion of validity is sensitive to the very same semantic feature that we already exploited in our semantic analysis of deontic conditionals, that is, the fact that deontic *ought* fails

---

[18]It the proposal committed to "Jones ought not murder Smith" being *false* in light of $i + m$? Not if we make the reasonable assumption that deontic *ought* carries the presupposition that its prejacent is a possibility in the modal base. In other words, "Jones ought not murder Smith" is neither true nor false in light of $i + m$, at least on its deliberative interpretation: the question whether Jones ought or ought not murder Smith is moot if it is taken for granted that Jones murders Smith.

to be persistent. This, I submit, is sufficient reason to see how a dynamic story for *ifs* and *oughts* can be further elaborated to take a few more crucial data into account. This is what I will do in the next section.

4 Bells and Whistles

4.1 Chisholm's Paradox

Chisholm's paradox starts with the observation that (12)-(15) are consistent and none of these statements logically implies any other one. This is especially puzzling since (12) and (13) seem to imply (16) while (14) and (15) seem to imply (17) (repeated):

(12) Jones ought to go to the aid of his neighbors,

(13) If Jones goes to the aid of his neighbors, then he ought to tell them he is coming,

(14) If Jones does not go to the aid of his neighbors, then he ought not tell them he is coming,

(15) Jones does not go to the aid of his neighbors,

(16) Jones ought to tell his neighbors that he is coming,

(17) Jones ought not tell his neighbors that he is coming.

We already saw that dynamic logical consequence preserves modus ponens and thus predicts the inference of (17) from (14) and (15). The inference of (16) from (12) and (13)—deontic detachment—is not licensed so far but easy to enforce by imposing the following constraint on our deontic selection function:

Weak Stability   For all $i, i' \subseteq i$, if $d(i) \subseteq i'$, then $d(i) \subseteq d(i')$

Weak Stability requires that strengthening with what is considered to be deontically ideal preserves deontically ideal worlds: doing what one ought to do at best absolves us from some our duties but does not create any new ones. This requirement lives happily with what I said before in response to the miners and the gentle murder paradox, and it is just what we need to license the inference of (16) from (12) and (13):

Fact 8   Given Weak Stability, $\Box_d go, go \Rightarrow \Box_d tell \vDash_D \Box_d tell$

This holds just in case for all w and i such that $w \in i$: if $w \in [\![\Box_d go]\!]^i$ and $w \in [\![go \Rightarrow \Box_d tell]\!]^{i+\Box_d go}$, then $w \in [\![\Box_d tell]\!]^{(i+\Box_d go)+(go \Rightarrow \Box_d tell)}$. Since conditionals and deontic *ought* are locally invariant, this condition reduces to the requirement that whenever $w \in [\![\Box_d go]\!]^i$ and $w \in [\![go \Rightarrow \Box_d tell]\!]^i$, then $w \in [\![\Box_d tell]\!]^i$. Observe that $d(i) \subseteq i$ because of Realism and suppose that $w \in [\![\Box_d go]\!]^i$ and $w \in [\![go \Rightarrow \Box_d tell]\!]^i$:

then $d(i) \subseteq [\![go]\!]^i$ and thus $d(i) \subseteq i + go$. Weak Stability then guarantees that $d(i) \subseteq d(i + go)$. But $d(i + go) \subseteq [\![tell]\!]^{i+go}$ and so $d(i) \subseteq [\![tell]\!]^{i+go}$. Since $tell$ is f-persistent, $d(i) \subseteq [\![tell]\!]^i$ and thus $w \in [\![\Box_d tell]\!]^i$, as required.

The fact that we can find constraints on our deontic selection that license deontic detachment is perhaps not terribly surprising, and in any case those who are committed to a neoclassical or informational approach to logical consequence can do the same. What *is* worth mentioning, however, is that we can endorse both factual and deontic detachment while avoiding Chisholm's paradox. And once again it is the already familiar nonmonotonicity of dynamic logical consequence that does the trick. Deontic detachment licenses the inference of (16) from (12) and (13), but this inference is defeated by the premise (15) that triggers detachment of the contrary-to-duty obligation articulated by (17). So to be precise, assume that $\Sigma = \{\Box_d go, go \Rightarrow \Box_d tell\}$ and that $\Sigma' = \{\neg go, \neg go \Rightarrow \Box_d \neg tell\}$: then $\Sigma \vDash_D \Box_d tell$ and $\Sigma' \vDash_D \Box_d \neg tell$ as required by deontic and factual detachment, respectively. Yet there is no contradiction, for even though $\Sigma \vDash_D \Box_d tell$, we also have $\Sigma, \neg go \nvDash_D \Box_d tell$, which is compatible with deontic detachment since our deontic selection function is required to be realistic and thus $\Sigma, \neg go \nvDash_D \Box_d go$. So under the assumption that Jones does not go, he ought not tell his neighbors that he is coming, but that is consistent with the intuition that he ought to tell his neighbors that he is coming in case he ought to go and also ought to tell them if he goes. What makes all this possible is that additional information in discourse and reasoning may defeat prior deontic truths and thus, in turn, inferences such as those licensed by deontic detachment.

4.2 Order-Sensitivity

Dynamic logical consequence, as I have shown, offers an attractive perspective on deontic discourse and reasoning since it allows us to preserve intuitive rules of inference for deontic conditionals as well as the consistency of the premises in Forrester's and Chisholm's paradox. But, as so often in dynamic approaches to discourse and reasoning, order matters. It is easy to verify that reversing the order of premises in Forrester's paradox results in a contradiction, i.e. $m, m \Rightarrow \Box_d g, \Box_d \neg m \vDash_D \bot$ (and similarly for Chisholm's paradox). The question whether this is a damaging result is quite subtle and must be left to another day, but at a minimum it is fair to worry that we have identified an unwelcome limitation on the explanatory power of the dynamic approach outlined so far. The good news is that one can arrive at a slightly modified conception of logical consequence that preserves the key ideas of my dynamic proposal and sidesteps the issue of order-sensitivity.

Dynamic Logical Consequence, v.2 $\phi_1, \ldots, \phi_n \vDash_{D*} \psi$ iff for all w and i such that $w \in i$: if $w \in [\![\phi_1]\!]^i$ and... and $w \in [\![\phi_n]\!]^i$, then $w \in [\![\psi]\!]^{(i+\phi_1) \cap \ldots \cap (i+\phi_n)}$

On this conception, premises of an argument play the classical role of restricting the set of possible worlds at which its conclusion is evaluated. But in addition, each premise updates the relevant informational parameter, and the conclusion is

then evaluated with respect to the intersection of those restrictions. This may look like an ad hoc move, but in fact it preserves the key idea of the previous proposal—that logical consequence should be sensitive to the natural language phenomenon of persistence failures—while avoiding a contingent commitment endorsed before. Exploiting persistence failures in logical consequence requires that premises of an argument update the informational parameter in light of which the conclusion is evaluated, but it does not require that the update is sequential.

The resulting version of dynamic logical consequence shares the virtues of its predecessor. Above all, it resolves Forrester's and Chisholm's paradox while preserving factual as well as deontic detachment, and for familiar reasons. Factual information continues to play its usual dynamic role of strengthening the informational parameter in light of which the conclusion of an argument is evaluated. In combination with the independently plausible failure of t-persistence for deontic *ought*, this yields the desired nonmonotonic effects. But in addition, the setup guarantees that the premises of Forrester's and Chisholm's paradox are dynamically consistent regardless of the order of premises.

5   Concluding Remarks

The goal of this paper has been to look for a semantic approach to conditionals and deontic modals that does not only avoid the miners paradox but also Forrester's and Chisholm paradox while preserving intuitive rules of inference for deontic conditionals. A treatment of deontic *ought* as an informational modal is on the right track but requires that we re-think the notion of logical consequence. Neoclassical and informational consequence resolve the miners paradox but fail to deliver fully adequate responses to the more classical paradoxes from Forrester and Chisholm, and this is not a coincidence since both notions of validity are committed to a monotonic conception of logical consequence. A dynamic conception of logical consequence, in contrast, naturally leads to a nonmonotonic logic for deontic discourse and reasoning and this, I have argued, offers an attractive escape route from the puzzles surrounding conditionals articulating contrary-to-duty obligations.

In arguing for a dynamic approach to deontic discourse and reasoning, I have put a lof of weight on the observation that dynamic logical consequence avoids our deontic paradoxes while preserving factual and deontic detachment for iffy oughts. It strikes me as uncontroversial that this is the best result one might hope for, and neither neoclassical nor informational consequence have something similar to offer. But one might still worry that I have overstated my case. The complaint about neoclassical consequence was that it fails to deliver the validity of factual detachment and thus does not account for the role that conditionals articulating contrary-to-duty obligations play in everyday discourse and reasoning. However, and as I mentioned earlier, modus ponens is quasi-valid, which is just to say that it is neoclassically valid whenever the antecedent is epistemically necessary: $\phi \Rightarrow \psi, \Box_e \phi \vDash_N \psi$. It follows that contrary-to-duty obligations detach on the neoclassical picture whenever it is settled that the relevant primary obligation is violated, and one might insist that this is all we need to account for the role that

such conditional obligations play in discourse and reasoning.[19]

We are well advised not to be satisfied with what the neoclassicist has to offer. Above all, it is nonnegotiable that plain factual information triggers detachment, and this still remains a mystery on the neoclassical picture. To get the problem into clearer view, consider again the miners scenario. Alex and Mary are discussing what to do; Mary is a bit better informed about the miners' whereabouts than Alex:

(18)   Alex: We ought to block neither shaft.
        Mary: But the miners are in shaft A.
        Alex: Oh, so we ought to block shaft A, then.

Here the factual information that the miners are in shaft A triggers detachment, and it is just as easy to imagine discourses in which plain factual information allows us to infer contrary-to-duty obligations. Restricting modus ponens to cases in which the antecedent is epistemically necessary, then, does not give us everything we need to make sense of deontic conditionals.

Of course, there is some room for pragmatic maneuvers here. The obvious response on behalf of the neoclassicist takes some inspiration from Stalnaker's (1978) seminal work on assertion and maintains that Mary's utterance modifies the conversational context: it strengthens the common ground between Alex and Mary by eliminating all possible worlds at which the miners are not in shaft A. With respect to the modified context, it is epistemically necessary that the miners are in shaft A, and so once again we have a case in which even the neoclassicist can predict detachment.

The outlined pragmatic rejoinder is not unreasonable but what does the trick is the assumption that epistemic *must* interacts dynamically with factual information in discourse: the assertion that the miners are in shaft A updates the common ground, triggering a transition from a state of information according to which the miners might be in shaft A but also might be in shaft B (and thus we ought to block neither shaft) to one according to which the miners must be in shaft A (and thus we ought to block shaft A). The problem for the neoclassicist is that there is no reason at all to think that the dynamics of *must* is a merely pragmatic affair—it is just as robust in reasoning, as the following example demonstrates:

(19)   a.   The miners are in shaft A or they are in shaft B.
        b.   They are not in shaft B.
        c.   *So*, they must be in shaft A.

This is as an intuitive entailment as we are likely to find even if (19a) and (19b) occur in a merely hypothetical context, and to make sense of it one needs to admit that the information that the miners are not in shaft B eliminates the epistemic possibility that they are in shaft B. In short, ϕ defeats $\ulcorner \diamondsuit_e \neg \phi \urcorner$ not only in everyday discourse but also in everyday reasoning. What started out as a pragmatic response

---

[19]This response is very much in the spirit of Kolodny's and MacFarlane's view on how life goes on without modus ponens. The result that modus ponens is generally invalid but holds whenever the antecedent is settled can also be found in the frameworks developed by Belnap et al. (2001), Horty (2001) and Loewer and Belzer (1983, 1986), who interpret settledness in a temporal framework.

to the shortcomings of neoclassicism, then, collapses into an argument for the general validity of factual detachment: ϕ entails $\ulcorner\Box_e\phi\urcorner$ and since $\ulcorner\Box_e\phi\urcorner$ together with $\ulcorner\phi \Rightarrow \Box_d\psi\urcorner$ entail $\ulcorner\Box_d\psi\urcorner$, we get factual detachment. There is, however, a bit more to learn from this exercise.

The preceding considerations matter since we are now in a position to diagnose the troubles with neoclassicism. There is nothing wrong with the neoclassical contention that $\ulcorner\phi \Rightarrow \psi\urcorner$ together with $\ulcorner\Box_e\phi\urcorner$ entail ψ, but everything wrong with its contention that $\ulcorner\phi \Rightarrow \psi\urcorner$ together with ϕ do *not* entail ψ. These facts are related because a commitment to ϕ brings in its wake a commitment to $\ulcorner\Box_e\phi\urcorner$, and not only in discourse but also in reasoning. This is what we have on the dynamic conception of logical consequence, i.e. $\phi \vDash_D \Box_e\phi$, but not on the neoclassical view since $\phi \nvDash_N \Box_e\phi$. Neoclassicists are thus right to say that modus ponens is valid whenever the antecedent is epistemically necessary but at the same time overlook the interaction between factual information and epistemic necessity: the neoclassical failure to get the basic facts about deontic *ought* straight (factual detachment) is grounded in its failure to capture the dynamics of epistemic *might* and *must*. Once again, we arrive at the dynamic perspective on discourse and reasoning as the right supplement to information-sensitive semantics.

Combining insights from dynamic semantics with recent work on deontic modality allows us to make some real progress towards a satisfying solution for a number of notorious paradoxes about deontic conditionals. Without doubt, there are many more issues about deontic modality to explore. Above all, I have abstracted from the question how deliberative interpretations relate to non-deliberative ones. I have also not addressed the possibility of genuine moral dilemmas, the topic of weak vs strong deontic necessity, or issues pertaining to degree of value, comparing one option to another, and the number of deontic notions connected to these ideas.[20] But I submit that as soon as things get iffy in the realm of deontic modality, a dynamic perspective is an attractive one to take.

References

Åqvist, Lennart, and Jaap Hoepelman. 1981. "Some Theorems About a 'Tree' System of Deontic Tense Logic." In *New Studies in Deontic Logic*, ed. Risto Hilpinen, 187–221. Dordrecht: Reidel.

---

[20]For a discussion of strong vs weak necessity in deontic modality, see von Fintel and Gillies (2008). For a discussion of supererogation and related notions, see McNamara (1996a,b).

Arregui, Ana. 2010. "Detaching *If*-clauses from *Should*." *Natural Language Semantics* 18(3): 241–293.

Belnap, Nuel, Michael Perloff, and Ming Xu. 2001. *Facing the Future: Agents and Choices in Our Indeterminist World*. New York: Oxford University Press.

Bonevac, Daniel. 1998. "Against Conditional Obligation." *Noûs* 32(1): 37–53.

Cariani, Fabrizio, Magdalena Kaufmann, and Stefan Kaufmann. 2011. "Deliberative Modality under Epistemic Uncertainty." http://semanticsarchive.net/Archive/TM5ZmM2Z/epi_delib_final.pdf. Manuscript, Northwestern University/University of Göttingen.

Castañeda, Hector-Neri. 1985. "Aspectual Actions and Davidson's Theory of Events." In *Aspects and Events*, ed. Ernest LePore and Brian P. McLaughlin, 294–310. Oxford: Blackwell.

———. 1986. "Obligations, Aspectual Actions and Circumstances." *Philosophical Papers* 15: 155–170.

Charlow, Nate. Forthcoming. "What We Know and What To Do." *Synthese.*

Chisholm, Roderick M. 1963. "Contrary-to-Duty Imperatives and Deontic Logic." *Analysis* 24(2): 33–36.

Dowell, Janice L. 2011. "A Flexible Contextualist Account of Epistemic Modals." *Philosophers' Imprint* 11(14).

Dreier, Jamie. 2009. "Practical Conditionals." In *Reasons for Action*, ed. David Sobel and Steven Wall, 116–133. New York: Cambridge University Press.

van Eck, Job A. 1982. "A System of Temporally Relative Modal and Deontic Predicate Logic and Its Philosophical Applications (2)." *Logique et Analyse* 25: 339–381.

Feldman, Fred. 1986. *Doing the Best We Can*. Dordrecht: Reidel.

———. 1990. "A Simpler Solution to the Paradoxes of Deontic Logic." *Philosophical Perspectives* 4, Action Theory and Philosophy of Mind: 309–341.

von Fintel, Kai. 2012. "The Best We Can (Expect to) Get? Challenges to the Classic Semantics for Deontic Modals." http://mit.edu/fintel/fintel-2012-apa-ought.pdf. Manuscript, MIT.

von Fintel, Kai, and Anthony S. Gillies. 2008. "CIA Leaks." *Philosophical Review* 117(1): 77–98.

Forrester, James W. 1984. "Gentle Murder, or the Adverbial Samaritan." *Journal of Philosophy* 81(4): 193–197.

Gillies, Anthony S. 2009. "On Truth-Conditions for *If* (but Not Quite Only *If*)." *Philosophical Review* 118(3): 325–349.

———. 2010. "Iffiness." *Semantics and Pragmatics* 3(4): 1–42.

Goble, Lou. 1991. "Murder Most Gentle: The Paradox Deepens." *Philosophical Studies* 64(2): 217–227.

Greenspan, Patricia S. 1975. "Conditional Oughts and Hypothetical Imperatives." *The Journal of Philosophy* 72(10): 259–276.

Horty, John F. 2001. *Agency and Deontic Logic.* New York: Oxford University Press.

———. 2003. "Reasoning with Moral Conflicts." *Noûs* 37(4): 557–605.

———. 2007. "Reasons as Defaults." *Philosophers' Imprint* 7(3).

Kolodny, Niko, and John MacFarlane. 2010. "Ifs and Oughts." *Journal of Philosophy* 107(3): 115–143.

Kratzer, Angelika. 1991. "Modality." In *Semantics: An International Handbook of Contemporary Research*, ed. Armin von Stechow and Dieter Wunderlich, 639–650. Berlin: Walter de Gruyter & Co.

Lewis, David K. 1973. *Counterfactuals.* Cambridge, MA: Harvard University Press.

Loewer, Barry, and Marvin Belzer. 1983. "Dyadic Deontic Detachment." *Synthese* 54(2): 295–318.

———. 1986. "Help for the Good Samaritan Paradox." *Philosophical Studies* 50(1): 117–127.

McNamara, Paul. 1996a. "Making Room for Going Beyond the Call." *Mind* 105(419): 415–450.

———. 1996b. "Must I Do What I Ought? (or Will the Least I Can Do Do?)." In *Deontic Logic, Agency and Normative Systems*, ed. Mark A. Brown and Jose Carmo, 154173. New York: Springer.

Nute, Donald E., ed. 1997. *Defeasible Deontic Logic.* Dordrecht: Kluwer.

Parfit, Derek. 1988. "What We Together Do." http://individual.utoronto.ca/stafforini/parfit/parfit_-_what_we_together_do.pdf. Manuscript, Oxford.

———. 2011. *On What Matters.* New York: Oxford University Press.

Prakken, Henry, and Marek Sergot. 1996. "Contrary-to-Duty Obligations." *Studia Logica* 57(1): 73–90.

———. 1997. "Dyadic Deontic Logic and Contrary-to-Duty Obligations." In *Defeasible Deontic Logic*, ed. Donald E. Nute, 223–262. Dordrecht: Kluwer.

Reagan, Donald H. 1980. *Utilitarianism and Cooperation.* New York: Oxford University Press.

Schroeder, Mark A. 2011. "*Ought*, Agents, and Actions." *Philosophical Review* 120(1): 1–41.

Silk, Alex. 2012. "Evidence Sensitivity in Weak Necessity Deontic Modals." http://www-personal.umich.edu/~asilk/Alex_Silk/home_files/silk%20evidence-sensitivity.pdf. Manuscript, University of Michigan.

Sinnott-Armstrong, Walter. 1985. "A Solution to Forrester's Paradox of Gentle Murder." *The Journal of Philosophy* 82(3): 162–168.

Stalnaker, Robert C. 1975. "Indicative Conditionals." *Philosophia* 5(3): 269–286.

———. 1978. "Assertion." In *Syntax and Semantics 9*, ed. Peter Cole, 315–332. New York: New York Academic Press.

Thomason, Richmond H. 1981a. "Deontic Logic as Founded on Tense Logic." In *New Studies in Deontic Logic*, ed. Risto Hilpinen, 165–176. Dordrecht: Reidel.

———. 1981b. "Deontic Logic and the Role of Freedom in Moral Deliberation." In *New Studies in Deontic Logic*, ed. Risto Hilpinen, 177–186. Dordrecht: Reidel.

Veltman, Frank. 1985. *Logics for Conditionals*. Ph.D. thesis, University of Amsterdam.

———. 1996. "Defaults in Update Semantics." *Journal of Philosophical Logic* 25(3): 221–261.

von Wright, George H. 1951. "Deontic Logic." *Mind* 60(237): 1–15.

———. 1956. "A Note on Deontic Logic and Derived Obligation." *Mind* 65(260): 507–509.

Yalcin, Seth. 2007. "Epistemic Modals." *Mind* 116(464): 983–1026.

———. 2011. "Nonfactualism about Epistemic Modals." In *Epistemic Modality*, ed. Andy Egan and Brian Weatherson, 295–332. New York: Oxford University Press.