A Remark on Iffy Oughts*

Malte Willer

Abstract

Every adequate semantics for conditionals and deontic *ought* must offer a solution to the miners paradox about conditional obligations. Kolodny and MacFarlane have recently argued that such a semantics must reject the validity of modus ponens. I demonstrate that rejecting the validity of modus ponens is inessential for an adequate solution to the paradox.

1 The Miners Paradox

Every adequate semantics for conditionals and deontic *ought* must offer a solution to the miners paradox about conditional obligations. Kolodny and MacFarlane (2010) have recently argued that such a semantics must reject the validity of modus ponens. My goal in this paper is to demonstrate that rejecting modus ponens is inessential for solving the puzzle. I begin with a brief outline of Kolodny's and MacFarlane's analysis of the paradox and their reasons for rejecting modus ponens. Then I develop and defend a semantics for deontic conditionals that avoids the miners paradox while preserving the validity of modus ponens. The key observation of this paper is that Kolodny's and MacFarlane's case against modus ponens trades heavily on assumptions about logical consequence that their very own semantics shows are dubious. The validity of modus ponens may very well have its limits, but we have no compelling reason to think that it is invalid for deontic conditionals and certainly would need more than the miners paradox to show that it is.[1]

Here is the miners paradox. 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.

*Published in the *Journal of Philosophy* 109(7): 449–461, 2012

[1] I observe elsewhere that modus ponens is in fact a problematic rule of inference for so-called "Thomason conditionals" such as "If Mary is deceiving me, I'll never believe it" (see Willer (2010) for discussion). My goal here is then not to offer an unqualified defense of modus ponens but rather to contribute to the (no less interesting) project of determining the exact scope of its validity.

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). Thus we have a paradox.

After a detailed discussion of various escape routes, Kolodny and MacFarlane conclude that we should take the argument to be invalid with its obvious logical form. Pursuing this strategy requires that we reject at least one of the following rules of inference, which are jointly sufficient to derive the paradoxical conclusion (5) from (2)-(4): modus ponens (MP), disjunction introduction ($\vee$I), or disjunction elimination ($\vee$E). Throughout this paper, I will follow a familiar path in formally representing *ought* as a deontic necessity operator ($\Box_d$) and *if* as a binary connective ($\Rightarrow$).[2]

[2] Kolodny and MacFarlane treat conditional antecedents as operators to account for their role as modifiers of modals, but the difference between their notation and mine is harmless for current purposes.

1	$inA \vee inB$	
2	$inA \Rightarrow \Box_d blA$	
3	$inB \Rightarrow \Box_d blB$	
4	$\quad inA$	
5	$\quad \Box_d blA$	MP, 2, 4
6	$\quad \Box_d blA \vee \Box_d blB$	$\vee$I, 5
7	$\quad inB$	
8	$\quad \Box_d blB$	MP, 3, 7
9	$\quad \Box_d blA \vee \Box_d blB$	$\vee$I, 8
10	$\Box_d blA \vee \Box_d blB$	$\vee$E, 1–9

Kolodny and MacFarlane join McGee (1985) in rejecting modus ponens. To motivate this decision, they present a simpler paradoxical argument that uses only one conditional premise. Notice that (1) entails:

(6) It is not the case that we ought to block shaft A.

But (2) and (6) seem to entail that

(7) The miners are not in shaft A,

which is an unjustified conclusion. Our moral predicaments do not allow us deduce the location of the miners. This paradoxical argument does not appeal to any proof rules for disjunction. But it relies on modus tollens, which can be proved using modus ponens and reductio:

1	$\phi \Rightarrow \psi$	
2	$\neg\psi$	
3	$\quad \phi$	
4	$\quad \psi$	MP, 1, 3
5	$\quad \bot$	$\bot$I, 2, 4
6	$\neg\phi$	Reductio, 3–5

Since modus ponens appears to be the only common factor between this paradoxical argument and the original one, Kolodny and MacFarlane conclude that an adequate solution to the miners paradox must reject the validity of modus ponens. Such radical measures, however, are not necessary to avoid all the trouble. Let me explain.

My account shares Kolodny's and MacFarlane's goal of explaining how our problematic arguments can be invalid with their obvious logical forms. There is then not much to be said in defense of modus tollens. I insist, however, that rejecting modus ponens is not necessary to explain how our paradoxical arguments can be invalid with their obvious logical forms. So my immediate goal is to disentangle modus ponens from modus tollens. It is, of course, tempting to think that given reductio, the only way to block the canonical proof of modus tollens is to give modus ponens the boot and deny the step from lines 1 and 3 to 4. But this is simply not so, as the proof also relies on the possibility of appealing to line 2 from within the subproof. Importing this line (or any other line derived from it) into the subproof is invalid in case the truth of $\ulcorner \neg\psi \urcorner$ fails to be preserved under the additional assumption of ϕ. So the proof presupposes that the logical consequence relation is monotonic in the following sense:

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

In defense of modus ponens, one may thus challenge the assumption that the right logic for *ifs* and *oughts* is monotonic, i.e. deny the principle that whenever $\phi_1, \ldots, \phi_n \vDash \psi$, then $\phi_1, \ldots, \phi_n, \phi_{n+1} \vDash \psi$.[3] Denying the monotonicity principle allows us to maintain that in some cases additional assumptions fail to preserve logical entailments, and my specific suggestion is that we exploit this option to account for the invalidity of modus tollens. Going back to the second paradoxical argument, let Γ be the set of sentences consisting of (2) and (6), i.e. assume that $\Gamma = \{inA \Rightarrow \Box_d blA, \neg\Box_d blA\}$. The hypothesis is that while $\ulcorner \neg\Box_d blA \urcorner$ follows trivially from Γ, it no longer follows from Γ if strengthened with the additional premise that the miners are in shaft A: $\Gamma \vDash \neg\Box_d blA$ yet $\Gamma, inA \nvDash \neg\Box_d blA$. So even if $\Gamma, inA \vDash \Box_d blA$ due to modus ponens, nonetheless $\Gamma, inA \nvDash \bot$ and thus $\Gamma \nvDash \neg inA$. In other words, modus tollens is invalid and (7) does not follow from (2) and (6). But this is *not* because modus ponens is invalid, but rather because the right logic for iffy oughts is nonmonotonic.

What we have seen so far is that there is a potential escape route from Kolodny's and MacFarlane's attempt at tying modus ponens to modus tollens. To make this escape route real, one has to make sense of the idea that a logic for deontic conditionals should be nonmonotonic. The good news of the following sections is that we can indeed make sense of this idea. The even better news is that the resulting semantics for deontic conditionals also offers a comprehensive

[3] For recent discussions of the issue of monotonicity in deontic logic, see also Horty (2003, 2007). If one sets aside the possibility that logical consequence is sensitive to the order of premises, Monotonicity is equivalent to the familiar requirement that whenever $\Gamma \vDash \phi$ and $\Gamma \subseteq \Delta$, $\Delta \vDash \phi$. In principle, however, the proof of modus tollens only appeals to a sequence rule that allows for thinning by adding premises on the right. Thus its minimal requirement is that logical consequence is "right monotonic" in the way articulated by Monotonicity. This requirement is satisfied in case logical consequence is closed under the superset relation, but the reverse need not hold in the absence of further assumptions about permutability.

solution to Kolodny's and MacFarlane's miners paradox. Let us go through the details.

2 Monotonicity Failures

There is some evidence that truth-conditional semantics should assign to at least some modalized sentences truth-conditions relative to a possible world and some body of information (modeled as a set of possible worlds). Yalcin (2007) puts this idea to good use in his discussion of epistemic modals—modals that express what, in view of some body of information, *might* or *must* be the case—and so do Kolodny and MacFarlane in their treatment of deontic modals. The basic idea is that such modals are specifications of informational modal operators: they are quantifiers over a set of possible worlds that is determined by a separate informational parameter. So if w is a possible world and i is an information state, the semantics for generic informational modals looks as follows:

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

Here f is a selection function mapping information states to modal quantifier domains. Different modals require different selection functions. Epistemic *must*, for instance, is a necessity operator whose domain is selected by an epistemic selection function e. Deontic *ought*, in turn, is treated as a necessity operator whose domain is selected by a deontic selection function d.

For the purposes of this paper, I will assume that a treatment of deontic modals as specifications of informational modal operators is on the right track.[4] Doing so, however, leaves a lot of room for substantial disagreement. Specifically, it does not decide how one wants to think about logical consequence in a semantics that assigns truth-conditions relative to possible worlds and information states. And this is anything but a trivial decision.

In a classical intensional setting, logical consequence amounts to necessary preservation of truth at a point of evaluation, understood as a possible world. It is then natural to think that our new semantic framework only adds a new wrinkle to an already familiar setup. The suggestion is that logical consequence remains nothing but necessary preservation of truth at a point of evaluation. However, a point of evaluation is no longer a plain possible world, but rather an ordered pair consisting of a possible world w and an informational parameter i. This is the path Kolodny and MacFarlane take, with the proviso that we only consider "proper" points of evaluation at which $w \in i$. Precisely:

[4] My general view is that a non-truth-conditional semantics for epistemic and deontic modals is more promising than a truth-conditional approach. It is also evident that a very rich conception of an informational parameter (distinguishing, for example, genuine beliefs from mere assumptions) is required to account for the full variety of informational modals in natural language, with corresponding complexities in how such parameters get modified in light of new information. All such reservations are irrelevant for the present issue.

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

This setup is conservative enough to preserve the structural features of classical logical consequence. In particular, neoclassical logical consequence is monotonic. To see this, notice that we may associate, with each set of sentences Γ, a set of proper indices $V(\Gamma) \subseteq W \times \mathcal{P}(W)$ that make all members of Γ true. $\Gamma \vDash \phi$ just in case ϕ is true at every member of $V(\Gamma)$. But whenever $\Gamma \subseteq \Delta$, $V(\Delta) \subseteq V(\Gamma)$, which just means that adding premises to an argument amounts to nothing but restricting the set of indices at which we evaluate the conclusion. So if $\phi_1, \ldots, \phi_n \vDash \psi$, then $\phi_1, \ldots, \phi_n, \phi_{n+1} \vDash \psi$.

On the neoclassical view, logical consequence amounts to necessary preservation of truth at a point of evaluation, which is now understood as a pair consisting of a possible world and an informational parameter. This setup is monotonic by design, but it is not the only available option. Another way of thinking about logical consequence assigns to informational parameters the role that Stalnaker (1978) assigns to contexts. Like Stalnaker's contexts, informational parameters determine what proposition a sentence expresses in discourse or reasoning, where a proposition is understood as a set of possible worlds. On this view, a point of evaluation is a plain possible world. But sentential truth is nonetheless information-sensitive, since the proposition that a sentence expresses may vary across different informational parameters.

Bringing Stalnaker into the picture adds more than a traditional gloss to the idea that truth depends on some informational parameter. This is because Stalnaker's theory about context-content interaction suggests a quite non-classical way of thinking about information aggregation in rational deliberation. In a well-run conversation, the information a sentence carries strengthens the context by eliminating all possibilities that are incompatible with it. In logical reasoning, additional information may then just do what sentences achieve in everyday discourse: they *affect* the informational parameter in light of which subsequent claims are evaluated. To make this idea more precise, we will require that ϕ affects i by strengthening it with the proposition that ϕ expresses relative to i, as follows:

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

An informational parameter i affects the proposition expressed by ϕ, and ϕ in turn affects i by ruling out all possibilities that are incompatible with the proposition expressed by ϕ in i.

Logical consequence can then be thought of as necessary preservation of truth at a possible world, but in addition we keep track of the changes that the premises of an argument induce on the proposition determining informational parameter. This leads to a dynamic notion of logical consequence:

Dynamic Logical Consequence $\phi_1, \ldots, \phi_n \vDash \psi$ iff for all w and i such that $w \in i$: if $w \in [\![\phi_1]\!]^i$ and... and $w \in [\![\phi_n]\!]^{i+\ldots+\phi_{n-1}}$, then $w \in [\![\psi]\!]^{(i+\ldots+\phi_{n-1})+\phi_n}$[5]

The current proposal predicts that information strengthening plays two distinct roles. First, adding another premise to an argument has the classical effect of restricting the set of possible worlds at which its conclusion is to be evaluated. Second, and since we now require that truth at a world is information sensitive, adding another premise to the argument strengthens the informational parameter with respect to which subsequent claims are evaluated.[6]

What makes the dynamic conception of logical consequence interesting for our discussion is that it fails to be monotonic by design. The reason is that adding premises to an argument now affects the informational parameter in light of which its conclusion is evaluated. At least in principle, this effect may *change* the truth-value of the conclusion at a possible world from true to false. More precisely, suppose that $\phi_1, \ldots, \phi_n \vDash \psi$ and consider some $w \in i$ such that $w \in [\![\phi_1]\!]^i$ and... and $w \in [\![\phi_n]\!]^{i+\ldots+\phi_{n-1}}$. Then $w \in [\![\psi]\!]^{(i+\ldots+\phi_{n-1})+\phi_n}$, i.e. the proposition that ψ expresses relative to $(i + \ldots + \phi_{n-1}) + \phi_n$ is true at w. But suppose we add a new premise ϕ_{n+1} to the argument. Even if $w \in [\![\phi_{n+1}]\!]^{(i+\ldots+\phi_{n-1})+\phi_n}$, it may very well be that $w \notin [\![\psi]\!]^{((i+\ldots+\phi_{n-1})+\phi_n)+\phi_{n+1}}$, since the proposition that ψ expresses can change as the result of strengthening $(i+\ldots+\phi_{n-1})+\phi_n$ with the additional information that is carried by ϕ_{n+1}. So it may very well be that $\phi_1, \ldots, \phi_n, \phi_{n+1} \nvDash \psi$, which is just to say that dynamic logical consequence is not guaranteed to be monotonic.

Dynamic logical consequence, then, leaves room for monotonicity failures. Whether or not this is a *real* possibility depends, of course, on the details of our information-sensitive semantics. Specifically, monotonicity is guaranteed in case truth at a point is persistent:

Persistence If $i' \subseteq i$, then $[\![\phi]\!]^i \subseteq [\![\phi]\!]^{i'}$

Persistence requires that truth at a possible world is preserved under information strengthening: it is excluded that additional information switches a sentence's truth-value at a world from true to false. The role of extra premises in an

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

[6] I should say here that the dynamic conception of logical consequence also makes sense if points of evaluation are pairs consisting of a possible world and an information state:

Dynamic Logical Consequence* $\phi_1, \ldots, \phi_n \vDash \psi$ iff for all w and i such that $w \in i$: if $\langle w, i\rangle \in [\![\phi_1]\!]$ and...and $\langle w, i + \ldots + \phi_{n-1}\rangle \in [\![\phi_n]\!]$, then $\langle w, (i + \ldots + \phi_{n-1}) + \phi_n\rangle \in [\![\psi]\!]$

That nothing technical hinges on how one wants to divide the labor between points of evaluation and features of the context should not be surprising (see Lewis (1980) for discussion). However, the alternative proposal misses a clear conception of logical consequence as preservation of truth at a point of evaluation. It is also hard to see how it can be motivated using Stalnaker's seminal work on assertion. All of this, I think, favors the dynamic conception of logical consequence that treats its points of evaluation as plain possible worlds.

argument is then limited to reducing the set of possible worlds at which its conclusion is evaluated. This guarantees that logical consequence is monotonic, even on its dynamic conception. In other words, monotonicity is guaranteed once we combine a dynamic conception of logical consequence with a suitably static conception of truth at a point.

The last observation notwithstanding, there is every reason to think that logical consequence as it is currently understood fails to be monotonic. In fact, one of Kolodny's and MacFarlane's very own insights into the semantics of *ought* shows that it is not. As they observe, our deontic selection function d must be "seriously information dependent:"

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. To see that this is plausible, assume that i represents our information about the miners' whereabouts, and let i' be the result of strengthening i with the information that the miners are in shaft A. Then $d(i)$ selects those possible worlds as deontically ideal at which we block neither shaft, but $d(i')$ selects those worlds as deontically ideal at which we block shaft A. Such information dependence immediately entails persistence failures: there are ϕ and $i' \subseteq i$ such that $[\![\phi]\!]^i \nsubseteq [\![\phi]\!]^{i'}$. For example, "It is not the case that we ought to block shaft A" is true at the actual world given i but fails to be so given i'.

One can now also demonstrate that dynamic logical consequence fails to be monotonic for a language involving deontic *ought*. To streamline the upcoming discussion, notice that some sentences are "locally invariant:"

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

It follows that whenever ϕ is locally invariant and $w \in [\![\phi]\!]^i$, then $i + \phi = i$ and thus $w \in [\![\phi]\!]^{i+\phi}$. Deontically modalized sentences are locally invariant because their truth-value does not vary across possible worlds given some fixed informational parameter. It follows that whenever $w \in [\![\Box_d\phi]\!]^i$, then $w \in [\![\Box_d\phi]\!]^{i+\Box_d\phi}$, i.e that $\Box_d\phi \vDash \Box_d\phi$. A similar line of reasoning establishes that $\neg\Box_d\phi \vDash \neg\Box_d\phi$ for all ϕ. This is all we need to prove the point.

Consider the miners scenario again. Given our information i, we ought to block neither shaft, and thus it is false that we ought to block shaft A. Assume that $w \in [\![inA]\!]^i$, i.e. that the miners are in shaft A at w. Then $w \in [\![\neg\Box_d blA]\!]^i$. But since blocking shaft A is deontically ideal with respect to the result of strengthening i with the information that the miners are in shaft A, $w \notin [\![\neg\Box_d blA]\!]^{i+inA}$. The simple observation then is that $\neg\Box_d blA \vDash \neg\Box_d blA$ but $\neg\Box_d blA, inA \nvDash \neg\Box_d blA$. Thus dynamic logical entailment fails to be monotonic.

The resulting nonmonotonic logic for deontic *ought* does not only cut the classical connection between modus tollens and modus ponens that proves to be so damaging for the latter. It also makes perfect sense of the observation that (2) and (6) fail to entail (7) (repeated):

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

(6) It is not the case that we ought to block shaft A,

(7) The miners are not in shaft A.

Given that it is an open question where exactly the miners are, we ought to block neither shaft, and thus it is not the case that we ought to block shaft A. But now strengthen the information from which we reason by assuming that the miners are in shaft A. Then in light of *that* information, blocking neither shaft is *not* the right thing to do. Quite to the contrary, given that the miners are in shaft A, we ought to block shaft A, just as modus ponens together with (2) would predict. No contradiction follows since adding another premise has changed the truth-value of "It is not the case that we ought to block shaft A" from true to false, which is why importing the initial premise (6) into the subproof is an invalid step. Since the assumption that the miners are in shaft A does not lead to a contradiction, it simply does not follow that the miners are not in shaft A. And of course, all of this is perfectly compatible with the unrestricted validity of modus ponens. So *pace* Kolodny and MacFarlane, the fact that (7) does not follow from (2) and (6) gives us no reason to think that modus ponens is invalid. In fact, it is in part due to their very own insights into the semantics of deontic *ought* that we can see how to disentangle modus ponens from modus tollens.

A sober analysis of the transition from classical to information-sensitive semantics thus reveals that modus ponens is compatible with the failure of modus tollens. This, of course, does not establish that my preferred dynamic conception of logical consequence can avoid the first formulation of the miners paradox while preserving the validity of modus ponens. Another potential worry concerns the substance of the preceding discussion. After all, one may think that what has been said so far amounts to nothing but a verbal dispute about "logical consequence" in a semantics that assigns truth-conditions relative to possible worlds and informational parameters. I address these remaining issues in the next section.

3 Conditionals and Logical Consequence

One remaining challenge is to explain how conditionals can support modus ponens in a semantics for *ifs* and *oughts* that assigns truth-values relative to possible worlds and informational parameters. There is no need to be very creative here: Kolodny and MacFarlane give us everything we need to establish the validity of modus ponens. The key assumption is that conditional antecedents function as modifiers of informational modals. By default, this modal is an (perhaps implicit) epistemic necessity operator. But in a conditional such as "If the miners are in shaft A, we ought to block shaft A," the modal is deontic. The simple suggestion then is that a conditional is true with respect to some information state i just in case its consequent is true at the result of strengthening i with the information carried by the antecedent. Thus

$$[\![\phi \Rightarrow \psi]\!]^{w,i} \text{ is true iff } [\![\psi]\!]^{w,i+\phi} \text{ is true}$$[7]

The neoclassical conception of logical consequence predicts that modus ponens fails to be reliable for conditionals. For suppose that $\ulcorner\phi \Rightarrow \psi\urcorner$ and ϕ are true at some proper index $\langle w, i\rangle$. Then we know that ψ is true at $\langle w, i + \phi\rangle$. This, however, is compatible with ψ being false at the original index $\langle w, i\rangle$. So there are points of evaluation at which a conditional and its antecedent are true yet the consequent fails to be true.

The dynamic conception of logical consequence preserves the validity of modus ponens. The reason is that on this conception, a premise does not only restrict the points at which the conclusion of an argument is evaluated, but also modifies the informational parameter in light of which subsequent claims are evaluated. Modus ponens is valid just in case for all w and i such that $w \in i$: if $w \in [\![\phi \Rightarrow \psi]\!]^i$ and $w \in [\![\phi]\!]^{i+(\phi\Rightarrow\psi)}$, then $w \in [\![\psi]\!]^{(i+(\phi\Rightarrow\psi))+\phi}$. Due to the local invariance of conditionals, this condition is equivalent to the requirement that if $w \in [\![\phi \Rightarrow \psi]\!]^i$ and $w \in [\![\phi]\!]^i$, then $w \in [\![\psi]\!]^{i+\phi}$. Notice that ψ is to be evaluated with respect to the result of strengthening i with the conditional antecedent. This marks an important distinction between the dynamic and the neoclassical view (which keeps the relevant informational parameter fixed). And in combination with our semantics of conditionals, this feature also delivers the validity of modus ponens. For notice that our semantics guarantees that $w \in [\![\psi]\!]^{i+\phi}$ whenever $w \in [\![\phi \Rightarrow \psi]\!]^i$. So the validity of modus ponens is an immediate result of the interaction between our semantic clauses and the dynamic conception of logical consequence.

We have already seen that the dynamic framework avoids the derivation of (7) from (2) and (6) once it is combined with a reasonable semantics for deontic *ought*. A bit of checking also verifies that it just as easily avoids the derivation of (5) from (2)-(4) once we add conditionals to the picture (repeated):

(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,

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

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

[7] Kolodny and MacFarlane further finesse their analysis to account for problems involving conditionals with modalized antecedents. Such conditionals are irrelevant for our purposes, and thus it is harmless to ignore the complications Kolodny and MacFarlane have in mind. What is said here can be easily modified in case the problems they discuss should indeed require a more subtle evaluation procedure for conditionals.

Suppose that i represents our knowledge about the miners' whereabouts. Then strengthening i with the premises (2)-(4) idles. Thus all we need to demonstrate the invalidity of the argument is to find some $w \in i$ so that (2)-(4) are true yet (5) is false. This is easy to do, and once again the crucial work is done by the flexibility of our deontic selection function and the resulting non-persistence of truth at a point of evaluation.

The simple observation is that our deontic selection function is designed in such a way that disjunction elimination fails for deontic conditionals. Notice that $d(i)$ selects those worlds as ideal at which we block neither shaft. But $d(i + inA)$ selects those worlds at which we block shaft A, and $d(i + inB)$ selects those at which we block B. As a consequence, given arbitrary w, $[\![\Box_d blA \vee \Box_d blB]\!]^{w,i}$ is false yet both $[\![\Box_d blA \vee \Box_d blB]\!]^{w,i+inA}$ and $[\![\Box_d blA \vee \Box_d blB]\!]^{w,i+inB}$ are true. Since the miners are in shaft A or in shaft B at all possible worlds in i, there are $w \in i$ such that $[\![inA \vee inB]\!]^{w,i}$, $[\![inA \Rightarrow (\Box_d blA \vee \Box_d blB)]\!]^{w,i}$, and $[\![inB \Rightarrow (\Box_d blA \vee \Box_d blB)]\!]^{w,i}$ are true yet $[\![\Box_d blA \vee \Box_d blB]\!]^{w,i}$ is false. Thus (5) does not follow from (2)-(4), as required. Conclusion: dynamic logical consequence gives us everything we want while preserving the validity of modus ponens.[8]

All of this should come as good news if one already believes that modus ponens is valid for deontic conditionals. Someone less committed, however, may still wonder why one should prefer the dynamic conception of logical consequence over Kolodny's and MacFarlane's neoclassical view. The decisive argument is that one's conception of logical consequence should accord with what one takes to be the best semantics for conditionals. On the face of it, the evaluation procedure for conditionals is similar to the one for arguments in that both proceed by evaluating a sentence under certain assumptions. This alone, of course, does not fix the details of our semantic theory. But it entails that a coherent semantics should deliver a *uniform* picture about the role of assumptions in conditional and deontic reasoning. The simple observation is that given what we have said about the meaning of conditionals, a coherent semantics must be based on a dynamic rather than a neoclassical conception of logical consequence. Let me explain.

Consider again the semantic proposal for conditionals that was outlined at the beginning of this section. This semantics is "internally dynamic," since the truth-value of a conditional is determined by first strengthening one's infor-

[8] Another attractive feature of the dynamic perspective is that it has no need for notions such as Kolodny's and MacFarlane's "quasi-validity" to explain why modus ponens is in general a reliable rule of inference. In this context, it is worth stressing that the notion of dynamic logical consequence is very different from the one of quasi-validity. Unlike dynamic logical consequence, quasi-validity only makes sense for inferences that depart from known premises and thus lacks any implications for hypothetical reasoning. As Levi (1996) observes, reasoning from premises that are accepted merely for the sake of the argument is of outstanding relevance for all kinds of practical and theoretical purposes. Relatedly, the notion of quasi-validity fails to make sense of some crucial forms of inference, including reductio: while $\neg\Box_d blA, inA \vDash \bot$ is quasi-valid, $\neg\Box_d blA \vDash \neg inA$ fails to be quasi-valid. It follows that the notion of quasi-validity—very much unlike the dynamic conception of logical consequence—offers an at best limited perspective on the logic of deontic and conditional reasoning.

mation with the antecedent and then evaluating the consequent on that basis. Dynamic logical consequence is internally dynamic just in the same way. On this view, processing an argument proceeds by first strengthening one's information with its premises and then evaluating its conclusion on that basis. The dynamic conception of logical consequence thus guarantees a nice match between the semantic evaluation procedure for conditionals and the way we semantically evaluate arguments in everyday reasoning.

In contrast, the proposed semantics for conditionals stands in tension with Kolodny's and MacFarlane's neoclassical conception of logical consequence. On the neoclassical view, evaluating an argument does *not* proceed by strengthening one's information with its premises, as the informational parameter in light of which claims are evaluated remains static throughout the evaluation process. But if the evaluation procedure for arguments is internally static, there is no reason to think that the one for conditionals is internally dynamic. And if, as Kolodny and MacFarlane themselves propose, our best semantics for conditionals is internally dynamic, there is no reason to believe in a neoclassical and thus internally static evaluation procedure for arguments. Even the most ardent skeptic about modus ponens should find this problem for neoclassicism worrying.

The upshot is that a coherent semantics that relativizes truth to bodies of information must be based on a dynamic conception of logical consequence. The point is not merely that the dynamic view offers a comprehensive solution to the miners paradox while preserving the validity of modus ponens for deontic conditionals. In addition, it delivers an intuitive match between the semantic evaluation procedure for conditionals and the way we semantically evaluate arguments in everyday reasoning. This does not diminish the importance of Kolodny's and MacFarlane's positive proposal for the semantics of *ifs* and *oughts*. As we have seen, their semantic insights are essential for the solution to the miners paradox that has been suggested in this paper. But at the same time, this solution undermines the main conclusion of Kolodny's and MacFarlane's discussion, i.e. a conclusive refutation of the validity of modus ponens. We have no reason to think that the miners paradox poses a problem for modus ponens, but every reason to believe that deontic logic has a lot to learn from nonmonotonic logic and the dynamic perspective on meaning and communication.

Acknowledgments Special thanks to Josh Dever and Hans Kamp for their detailed comments on earlier versions of this paper.

References

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

Horty, John F. 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.

Levi, Isaac. 1996. *For the Sake of the Argument.* New York: Cambridge University Press.

Lewis, David K. 1980. "Index, Context, and Content." In *Philosophy and Grammar*, edited by Stig Kanger and Sven Öhman, 79–100. Dordrecht: Reidel.

McGee, Vann. 1985. "A Counterexample to Modus Ponens." *The Journal of Philosophy* 82(9): 462–471.

Stalnaker, Robert C. 1978. "Assertion." In *Syntax and Semantics 9*, edited by Peter Cole, 315–332. New York: New York Academic Press.

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.

Willer, Malte. 2010. "New Surprises for the Ramsey Test." *Synthese* 176(2): 291–309.

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