

A PARTITION-BASED SEMANTICS FOR DEONTIC MODALS

RAMIRO CASO

Instituto de Investigaciones Filosóficas SADAF/CONICET, Universidad de Buenos Aires, ARGENTINA

`caso.ramiro@conicet.gov.ar`

<https://orcid.org/0000-0002-1421-2739>

Abstract. Kolodny and MacFarlane (“Ifs and oughts”, *JPhil* 107(3)) and MacFarlane (*Assessment sensitivity*, Chapter 11) describe the distinctive behavior of deontic modals in deliberation. They highlight two main features that any semantic theory should account for. First, deontic modals are informational in the sense that relativity to a body of information is an essential part of their interpretation. Second, *modus ponens* should be invalid, and it should fail when such modals are involved. And they argue for an account that posits information-neutral semantic values. In this paper, I provide an account of deontic modals that does not involve any deviation from standard possible world semantic values, yet can capture their distinctive semantic behavior. The key element of the proposal is a partition-based semantics that treats available information as placing a constraint on interpretation.

Keywords: contextualism • relativism • information sensitivity • Miner Paradox

RECEIVED: 14/12/2023

REVISED: 15/07/2025

ACCEPTED: 03/11/2025

1. Preliminaries

Deontic modals (DMs) are a class of modals whose working requires that we represent some possibilities as better or worse, or as more or less preferable, than others. In this class, we find, among others, strictly deontic modals, like ‘must’ in (1a), bouletic modals, like ‘should’ in (1b), and teleological modals, like ‘ought’ in (1c):

- (1) a. Johann must return the money (in view of what is morally required).
- b. Susan should take the train (in view of her interest in travelling as safely as possible).
- c. Mary ought to buy a bicycle (in view of her goal to live a healthier life).

Briefly, (1a) says that the possibility of Johann returning the money is preferable (from the viewpoint of morality) to the possibility of his not doing so. (1b) says that Susan’s taking the train is preferable (from the viewpoint of her preference for travelling safely) to her not doing so. And (1c) states that Mary’s buying a bicycle is preferable (from the viewpoint of her goal to live a healthy life) to her doing otherwise (cf. Portner 2009, Chapter 3, on priority modals).



Kolodny and MacFarlane (2010) and MacFarlane (2014, Chapter 11) describe a distinct use of DMs, their *deliberative* use. Deliberative uses of deontic modals rank possibilities as better or worse according to an agent's preferences and goals in the context of deliberation. In describing this class of uses, they highlight two semantic features that an adequate account of DMs should capture. First, available information has two distinct effects on the interpretation of these expressions: it restricts the possibilities under consideration, and it affects the way those possibilities are ordered. Second, as a plausible consequence of this, *modus ponens* for the indicative conditional is an invalid rule of inference, and instances involving DMs provide counterexamples to it. They further argue that only a semantic theory that construes these expressions as information-neutral (that is, expressions whose extension is non-indexically sensitive to an information state) can properly account for their behavior, and consequently push for an assessment-sensitive semantics for sentences containing them according to which those sentences express propositions whose truth value is sensitive to both a possible world and an information state.

In what follows, I propose a contextualist theory that captures both the information relativity of DMs and the invalidity of *modus ponens* without positing non-standard semantic values. In Section 2, I present the practical situation of the Miner Paradox. It will provide all the elements needed for the discussion. I also outline Kolodny and MacFarlane's proposal for comparison purposes. In Section 3, I present a basic contextualist semantics for DMs in general, with special emphasis on the quasi-modal 'ought', based on the work of Kratzer (1981a, 1981b, 1986, 1991). Deliberative uses of DMs are a particular instance of this semantics. As it stands, however, standard Kratzerian semantics fails to capture the informational character of DMs. This is shown in Section 4, which also presents an account of information relativity in a modified version of Kratzer's semantic framework where DMs are interpreted relative to a contextually salient partition of the space of possibilities. This fulfils the first desideratum on an adequate semantics for DMs posited by Kolodny and MacFarlane. Section 5 shows how *modus ponens* is an invalid rule of inference under this semantics and how DMs provide counterexamples to it. This fulfils Kolodny and MacFarlane's second desideratum. Section 6 offers some concluding remarks that highlight, among other things, the precise sense in which the partition-based semantics is to be preferred over an information-sensitive construal.

2. A practical situation

The practical situation depicted in the Miner Paradox is simple (Kolodny & MacFarlane, 2010). Ten miners are trapped in a mine. They are either in shaft A or shaft B. Both locations are equally likely. The mine is about to flood, and you must act to save

as many miners as possible. Your options are blocking shaft A, shaft B, or neither. Blocking both shafts is out of the question because you don't have enough materials. You possess the following information: if the right shaft is blocked, all ten miners survive; if the wrong shaft is blocked, all ten miners die; if neither is blocked, nine miners will survive (for blocking neither shaft will cause water to go evenly into both shafts, thus killing just one miner, no matter where the miners are).

2.1. The information relativity of deontic modals

You deliberate, and you conclude that the best course of action is to block neither shaft, for this will save nine lives with certainty. Consequently, you conclude:

- (2) I ought to block neither shaft.

Now, suppose someone offers you the following advice:

- (3) No, you ought to block shaft A.

Suppose your adviser knows more than you about hydraulics and knows that, if you block neither shaft, water will make shaft A collapse, thereby blocking it. Consequently, water will go into shaft B, drowning anyone in it. If you block shaft A, it will take longer to collapse, and water will go evenly into both shafts. After hearing *this* explanation, you accept the advice and conclude:

- (4) I ought to block shaft A.

You make your decision, and you proceed to block shaft A.

Sentences (2)–(4) exemplify a central feature of DMs: their information relativity. This information relativity manifests itself in two ways. Firstly, available information restricts which possibilities are up for consideration. Secondly, available information affects how those possibilities are ordered in terms of better or worse.

To see that available information determines which possibilities we must consider, consider the Miner Paradox situation more closely. Let *S* be the deliberative agent and *A* the adviser. As far as both *S* and *A* are concerned, the miners may be in shaft A or in shaft B; neither shaft is likelier; the actions available to *S* are either blocking one of the shafts or blocking neither, etc. This means that, for both *S* and *A*, the possibilities under consideration will be ones where all these propositions hold. Different information, however, results in the consideration of different possibilities. For example, given *S*'s information, all possibilities where *S* blocks shaft A are ones where all water goes into shaft B, whereas relative to *A*'s information, all possibilities where *S* blocks shaft A are ones where (shaft A takes longer to collapse and) water goes evenly into both shafts. Thus, different information entails that different possibilities will be considered as live or relevant during deliberation. To see that available

information also determines how possibilities are ordered, we only need to observe that, relative to *S*'s information, the best possibilities are those where *S* blocks neither shaft, whereas relative to *A*'s information, they are those in which *S* blocks shaft *A*.

A final remark before leaving the information relativity of DMs. Kolodny and MacFarlane point out that it is customary to distinguish between subjective and objective readings of 'ought', along the following lines: in a subjective sense, *S* ought to ϕ if and only if, relative to *S*'s information (and *S*'s goals and preferences), ϕ -ing is the best course of action for *S* to take; and in an objective sense, *S* ought to ϕ if and only if, relative to all the relevant information, both known and unknown (and *S*'s goals and preferences), ϕ -ing is the best course of action for *S* to take. That is, under the subjective reading, only the information available to the deliberative agent is relevant, regardless of any other information that may be relevant to the practical situation the agent is in. Under the objective reading, the totality of relevant facts is taken into account, regardless of whether they are available to the deliberative agent (or any other agent, for that matter). The reading of DMs identified by Kolodny and MacFarlane, which will concern us in what follows, coincides with neither of these. This fact may be obscured if we focus only on first-person uses, for in them, deliberative agent and speaker coincide. However, advice situations reveal that the information-relative reading of DMs is a distinct reading, for in uttering (3), *A* is not basing their judgment on *S*'s information, or the totality of relevant information (insofar as *A* also ignores the miners' whereabouts). Rather, the adviser bases their judgment on whatever information they possess, which need not coincide with the deliberative agent's information, nor with the totality of the relevant facts. That is, the adviser's use of 'ought' seems to be governed by something like the following: *S* ought to ϕ iff, relative to *A*'s information (and *S*'s goals and preferences), ϕ -ing is the best course of action for *S* to take. More generally, information-relative uses seem to be anchored to whatever information is relevant in the context of utterance, regardless of whether that information is available to the deliberative agent.

2.2. The invalidity of *modus ponens*

The Miner Paradox is a paradox that arises from the Miner decision situation. As discussed in Sect. 2.1, before any adviser's contribution, the plausible conclusion of the process of deliberation is (2)—or we will take it to be so. Now, a contradiction arises when we consider the following additional (true, or plausibly true) premises:

- (5) If the miners are in shaft *A*, *S* ought to block shaft *A*.
- (6) If the miners are in shaft *B*, *S* ought to block shaft *B*.
- (7) The miners are either in shaft *A* or in shaft *B*.

By constructive dilemma, we seem to be able to conclude:

- (8) Either S ought to block shaft A, or S ought to block shaft B.

This conclusion contradicts the datum that the best course of action, what S ought to do in the situation, is to block neither shaft.

Kolodny & MacFarlane (2010) convincingly argue that we should solve the paradox by rejecting the validity of *modus ponens*. This will be my starting point. A somewhat formal version of the paradox may be of help in seeing why this is their preferred way out:

- | | | |
|-----|------------------------------------|-----------------------|
| 1. | $O(S, \neg bIA) \& O(S, \neg bIB)$ | Prem. |
| 2. | $inA \vee inB$ | Prem. |
| 3. | $inA \Rightarrow O(S, bIA)$ | Prem. |
| 4. | $inB \Rightarrow O(S, bIB)$ | Prem. |
| 5. | inA | Assumption 1 $\vee E$ |
| 6. | $O(S, bIA)$ | $\Rightarrow E$ 3, 5 |
| 7. | $O(S, bIA) \vee O(S, bIB)$ | $\vee I$ 6 |
| 8. | inB | Assumption 2 $\vee E$ |
| 9. | $O(S, bIB)$ | $\Rightarrow E$ 4, 8 |
| 10. | $O(S, bIA) \vee O(S, bIB)$ | $\vee I$ 9 |
| 11. | $O(S, bIA) \vee O(S, bIB)$ | $\vee E$ 2, 5-7, 8-10 |

There are three avenues for dealing with the paradox: (i) rejecting a premise; (ii) attributing a different logical form to the paradoxical reasoning (under which it is no longer a valid form of reasoning); and (iii) providing a semantics that invalidates a rule of inference required by the paradox. The question is, what should we do?

Rejecting a premise does not seem to be the right way to go. Premise 2 of the formalized paradox is true by stipulation. Premise 1 can be denied if we adopt an objective reading of DMs, for the best course of action, considering all relevant facts, known or unknown, is to block the shaft in which the miners are. This move would certainly block the paradox. However, the objective reading of ‘ought’ offers no actionable insights in a scenario characterized by uncertainty, such as the Miner practical situation, for uncertainty precludes an assessment of what the objectively correct decision is. As Kolodny and MacFarlane put it, “the objectivist’s ‘ought’ seems useless in deliberation” (in deliberation under uncertainty, I would qualify; Kolodny & MacFarlane 2010, p.117). Thus, even though reading ‘ought’ objectively blocks the paradox, it hardly seems to be a way of dealing with the paradox when deliberation under uncertainty is underway. At most, we have shown that no paradox arises under one possible interpretation of DMs. However, there are other possible interpretations

(arguably more relevant to deliberation under uncertainty) under which the paradox may still arise.

This leaves up for rejection only the deontic conditionals in premises 3 and 4. Can these be rejected? Let's start by noticing that, if we adopt a subjective reading of the DMs occurring in their consequents, these premises are indeed false (Kolodny & MacFarlane 2010, p.118ff). So, adopting a subjective interpretation of 'ought' would seem to be enough to block the paradox. There are good reasons to avoid this route. Firstly, just as before, this move would show, at most, that the argument is not paradoxical under the subjective reading of 'ought', without addressing the issue that it may still arise under the information-relative reading. Secondly, as Kolodny and MacFarlane point out, conditionals such as those in premises 3 and 4 routinely occur in practical deliberation, and declaring them false would force us to deem our everyday forms of practical reasoning defective, which, even if not a knock-down argument, provides some motivation for the endeavor of finding another way out of the paradox. Finally, and perhaps more importantly, the truth of those conditionals, under an information-relative reading of their consequents, can be motivated by an intuitive way of understanding indicative conditionals, which ultimately stems from (Ramsey 1931). In a famous passage, Ramsey discusses what is involved in accepting an indicative conditional. According to him, in assessing a conditional, we reason under the assumption that the antecedent holds. If, under that assumption, the consequent also holds, the conditional holds; and, if the consequent does not hold, the conditional does not hold. Once we assume that something like this operation is involved in the interpretation of the indicative conditional, premises 3 and 4 are plausibly assessed as true: in an information-relative reading of DMs, when we reason under the assumption that the miners are in shaft A (that is, when we take it as true, even if temporarily, that the miners are in shaft A), we conclude that, relative to the information assumed to hold in that scenario, what we should do is to block shaft A; and the same holds, *mutatis mutandis*, for the temporary assumption that the miners are in shaft B. (Notice, in passing, that this also explains the falsity of those same conditionals under the subjective reading: the assumption that the miners are in shaft A does not warrant that *S* ought to block that shaft when what matters is the information available to *S*—rather, what is needed is the stronger assumption that *S* comes to learn, in that situation, that the miners are in shaft A. Only then will the subjective 'ought' statement be evaluated as true.)

Since rejecting a premise does not seem to be a promising way of dealing with the Miner Paradox, we are left with the options of either revising the logical form of the sentences involved in the paradox or invalidating a required rule of inference. For the reasons in Kolodny & MacFarlane (2010) and MacFarlane (2014, Chapter 11), positing a different logical form is of no help.¹ So, we are left with the option of somehow invalidating the paradox under its *prima facie* logical form.² Since there

are versions of the paradox that do not employ disjunction introduction or disjunction elimination to derive a paradoxical conclusion (Kolodny & MacFarlane, 2010), invalidating *modus ponens* for the indicative conditional is the only remaining way out. Thus, we should provide an account of the quasi-modal ‘ought’ that entails (together with a suitable account of the indicative conditional) that *modus ponens* for the indicative conditional is invalid.³ More to the point, *modus ponens* should fail when such modals are involved.

Before moving on to my proposal, let’s briefly review Kolodny & MacFarlane’s (2010) and MacFarlane’s (2014, Chapter 11) account, which posits information-neutral semantic values for DMs (i.e., semantic values whose extensions are determined relative both to a possible world and an information state), and where sentences containing DMs express information-neutral propositions (i.e., propositions whose truth value varies according to an information state, as well as a possible world). They posit the following clause for ‘ought’ (from here on, I follow MacFarlane’s (2014) semantic presentation, since it is the most detailed):

$$(9) \quad \llbracket \text{Ought}(S, \phi) \rrbracket^{c,a} = \{\langle w, i \rangle \mid \forall z \in \text{Optimal}_{w,t_c,i}(S) : z \rightarrow_{S,w,t_c}^x |\phi(x)|\}$$

where c is a context of use, a is an assignment to variables, i is an information state, t_c is the time of c , and the following definitions are in place:

Definition 1. $^x|\phi(x)|$ is the action of making ‘ $\phi(x)$ ’ true of oneself.

Definition 2. $\text{Choice}_{w,t}(\alpha) = \{x \mid \alpha \text{ may choose to perform } x \text{ at } t \text{ at } w\}$

Definition 3. $x \geq_{\alpha,w,t,i} y$ iff, given information i , action x is at least as good as action y for α to perform at w at t .

Definition 4. $x \rightarrow_{\alpha,w,t} y$ iff α cannot perform x at w at t without performing y .

Definition 5. $\text{Optimal}_{w,t,i}(\alpha) = \{x \in \text{Choice}_{w,t}(\alpha) \mid \forall y \in \text{Choice}_{w,t}(\alpha) : y \geq_{\alpha,w,t,i} x \Rightarrow x \geq_{\alpha,w,t,i} y\}$

That is (leaving aside for now the relativization to a world and an information state), a sentence of the form ‘ S ought to ϕ ’ is true just in case each of the optimal choices available to S requires that S performs ϕ . A crucial aspect of these definitions is that the way the actions available to an agent are ordered (Definition 3) is sensitive to whatever information is provided. Consequently, which options are optimal (Definition 5) also depends on available information. It thus follows that which options are quantified over in a sentence of the form ‘ S ought to ϕ ’, and whether ϕ -ing is required by all of them, depends on whatever information is available as well. Consequently, whether a sentence of the form ‘ S ought to ϕ ’ is true also partly depends on available information. So, more precisely, one such sentence will be true at a possible

world w and an information state i just in case ϕ -ing (more properly, the action of making ϕ true of oneself) is required by all the optimal choices available to S at the world and time of utterance, as so determined by information i . This non-indexical sensitivity to a body of information is highlighted by the fact that the semantic value of one such sentence is not a set of worlds, as in standard, possible world semantics, but a set of world-information state pairs.

As for *modus ponens*, MacFarlane offers the following account of conditionals (cf. Yalcin 2007):

$$(10) \quad \llbracket \text{If } \phi, \psi \rrbracket^{c,a} = \{ \langle w, i \rangle : \forall i' \subseteq i (i' \subseteq \llbracket \phi \rrbracket^{c,a} \Rightarrow i' \subseteq \llbracket \psi \rrbracket^{c,a}) \}$$

Here, the function of the *if*-clause is to add the supposition that ϕ to the interpretation of ψ . More precisely, a conditional of the form “If ϕ, ψ ” at a context c expresses a proposition true at (a possible world w and) an information state i just in case every ϕ -substate of i is also a ψ -substate of i . Thus, (10) incorporates the intuition about the evaluation of indicative conditionals contained in Ramsey’s test in the following way: a conditional is true at an information state i just in case every way of consistently adding to i the assumption that the antecedent is true results in an enriched information state that makes the consequent also true. Notice that according to (10) the truth-conditions for an indicative conditional are information-sensitive but not world-sensitive. This can be appreciated from the fact that no free world variable occurs to the right of the set abstractor, whereas the information state variable i does occur free, and is thereby bound by the information state of evaluation.

This account leads to the invalidity of *modus ponens* when coupled with the following definition of validity:

Definition 6. S is a logical consequence of Γ iff, for all c, a , if for all $\gamma \in \Gamma$, $\llbracket \gamma \rrbracket^{c,a}$ is true relative to $\langle w_c, i_c \rangle$, then $\llbracket S \rrbracket^{c,a}$ is true relative to $\langle w_c, i_c \rangle$.

(with w_c and i_c the world and information state of c) plus the additional constraint that all the $\langle w, i \rangle$ pairs we consider are such that $w \in i$. It should be noted that Definition 6 introduces a notion of logical consequence that understands validity as preservation of truth at a context and the circumstance or index of that context. This notion was initially introduced by Kaplan (1989) in his discussion of the logic of demonstratives. There are two useful ways of conceptualizing it: as defining validity as preservation of *truth at a context* for sentences, or preservation of *truth* for occurrences of sentences, or sentence-context pairs, Kaplan’s semantic proxy for the pragmatic notion of *utterance*. In essence, Definition 6 states that an argument from a set of premises Γ to a conclusion S is valid just in case, no matter how the sentences involved in the argument are used, whenever all the premises are true, the conclusion is true as well.

To see how this semantics entails the invalidity of *modus ponens*, let's construct a counterexample. According to Definition 6 such a counterexample should have the following properties: a context c should exist, such that all the premises are true when evaluated with respect to the world of the context w_c and the information state of the context i_c , and the conclusion is false when evaluated with respect to those same parameters.

To have a particular case of *modus ponens* in mind, let's focus on one of the instances involved in the Miner Paradox:

- (11) a. If inA , $Ought(S, blA)$
 b. inA
 c. $Ought(S, blA)$

Let's also make a further stipulation. Let i be an information state and P a proposition. Then, we say that i supports P just in case $i \subseteq P$.

For (11b) to be true at w_c and i_c , w_c must be a world where the miners are in shaft A. Whether the miners are in shaft A does not depend in any way on the information available at c , so the truth of (11b) imposes no constraint on i_c . The truth of (11a), on the other hand, does place a constraint on i_c , and no further constraints on w_c (as discussed, the truth of a conditional under Kolodny and MacFarlane's account is world-insensitive and information-sensitive). Let's stipulate that i_c is the information state of the deliberative agent at the time of deliberation. Consequently, i_c does not contain enough information to decide the whereabouts of the miners. Thus, it contains worlds where the miners are in shaft A, and worlds where they are in shaft B. It follows from clause (10) that (11a) is true just in case every substate of i_c either does not support the antecedent or supports the consequent. Now, every such substate will fall into one of three classes: either it contains only worlds where the miners are in shaft A, only worlds where the miners are in shaft B, or both kinds of worlds. Any substate in the latter two classes satisfies the condition imposed by the conditional, because it does not support the antecedent. So, to secure the truth of (11a), we must ensure that every substate i'_c where the miners are known to be in shaft A, is a substate relative to which blocking shaft A is the optimal action (more properly, is required by all the optimal actions available to the deliberative agent at the time of deliberation). Crucially, at this point, given the information contained in i'_c , blocking shaft A is indeed the best action, for its result, saving all ten miners, is the preferred outcome. Thus, any substate of i_c that supports the antecedent of (11a) is also a substate that supports its consequent. Hence, (11a) is also true, under our construction of c . Finally, it can be shown that (11c) is false relative to c . The key insight here is that (11c) must be evaluated with respect to i_c , and relative to this information state, which is ignorant as to the whereabouts of the miners, the deontic sentence is false.

3. A Kratzerian semantics for deontic

Kolodny & MacFarlane's (2010) and MacFarlane's (2014, Chapter 11) account of the information relativity of DMs in general, and the Miner Paradox in particular, departs from standard semantic theory. For the relativization of extension (and thus propositional truth) to an information state calls for the introduction of information-neutral semantic values, witnessed by the modeling of propositions as sets of world-information state pairs (with the accompanying risk of disconnecting propositions from their intuitive—and thus, content-giving—anchoring as representational devices (Einheuser 2008; Wright 2008).) Even though relativism is a fairly well-established semantics by now, it is important to ask whether we may deal with the information sensitivity of DMs, and the invalidity of *modus ponens*, without deviating from standard possible world semantic values.

As it turns out, we can. In Sect. 4, I put forward a semantic analysis of DMs that captures the information relativity of 'ought' without positing information-neutral semantic values. And in Sect. 5, I show that this account, together with a Kratzerian take on conditionals, invalidates *modus ponens* for the indicative conditional in a principled way. In the remainder of this section, I introduce a standard, Kratzerian framework for natural language modals, which will form the basis for the semantic proposal of Sect. 4 and 5.

In Kratzer's framework (Kratzer 1981a, 1981b, 1986, 1991), modals are interpreted as restricted quantifiers over possible worlds. Modals have two quantificational forces: universal (necessity modals like 'must' and 'ought') and existential (possibility modals like 'might' and 'may'). The worlds over which a modal quantifies are determined by two contextual parameters, a modal base f and an ordering source g . Modal bases and ordering sources are implemented in terms of *conversational backgrounds*, i.e., functions from worlds to sets of propositions, or from agents and worlds to sets of propositions. Let $PROP$ be the set of all propositions, defined here as $\mathcal{P}(W)$, with W the set of all possible worlds. A modal base is a function $f : W \rightarrow \mathcal{P}(PROP)$ that, given a possible world w , determines a set of propositions $f(w)$, whose nature may vary. For example, a doxastic modal base determines, for every w , the set of propositions that are believed by a particular agent (or group of agents) at w , and a circumstantial modal base determines, for every world w , a set of propositions that are true at w . (Since propositions are being modeled as sets of worlds, a proposition P is true at a possible world w just in case $w \in P$.) For any w , $\cap f(w)$ yields the set of worlds at which all the propositions in $f(w)$ are true. An ordering source is, in the simplest case, a function $g : W \rightarrow \mathcal{P}(PROP)$ that orders the worlds in $\cap f(w)$ according to some criterion. For example, a normalcy-based ordering source orders the worlds in $\cap f(w)$ according to how much they respect what is the normal course of events in w , and a similarity-based ordering source orders the worlds in $\cap f(w)$

according to how much they resemble w in certain respects.

In Kratzer's theory, DMs may have doxastic or circumstantial modal bases, and a variety of ordering sources, which rank worlds according to how much they respect moral norms (strictly deontic modals), how good they are from the viewpoint of the agent's interests or desires (bouletic modals), etc. A characteristic of DMs in general is that the ordering source is a function $g : D_a \times W \rightarrow \mathcal{P}(PROP)$ from agents and possible worlds to sets of propositions (with D_a the set of agents). For the case of deliberative uses of DMs, the modal base orders worlds according to how well they satisfy the preferences and goals of the deliberative agent.

As for the syntax of DMs, I make the syntactic assumption that, with an important amount of simplification, 'ought' enters the structural configuration depicted in Figure 1:

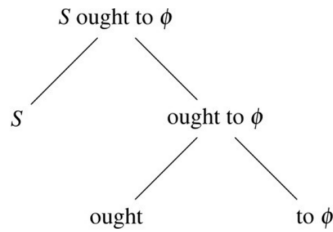


Figure 1: General syntactic parsing of bare 'ought' sentences

This structural assumption coheres with an analysis of 'ought' as akin to a control predicate (see endnote 2 for details). From the semantic standpoint, this translates into the following clause (adopting Portner's (2009) semantic style):

$$(12) \quad \llbracket \text{Ought} \rrbracket^{c,a,f,g} = \lambda P \lambda x \lambda w. \text{Best}(f(w), g(x, w)) \subseteq P(x)$$

where c , a are a context of use and an assignment to free variables, f is a modal base, and g is an ordering source. According to (12), 'ought' takes a proposition P and an individual x as arguments and returns a proposition (a property of worlds) according to which the best worlds, relative to a given selection of modal base and ordering source, are all worlds where P is true. As for Best , it is defined as follows:

Definition 7. $\text{Best}(f(w), g(x, w)) = \{u \in \cap f(w) \mid \neg \exists v (v \in \cap f(w) \ \& \ u <_{g(x,w)} v)\}^4$

with $<_{g(x,w)}$ the ordering relation derived from g . According to Definition 7, the best worlds within the modal base are those such that no world in the modal base is strictly better, according to the ordering source. The ordering relation $<_g$ is obtained through the following definition:

Definition 8. For all $w', w'' \in \cap f(w)$, $w' <_{g(x,w)} w''$ iff for all $P \in g(x, w)$, if $w'' \in P$, then $w' \in P$, and there is a $P \in g(x, w)$ such that $w' \in P$ and $w'' \notin P$.

That is, a world w' is strictly better than a world w'' , from the viewpoint of g (anchored in the world of evaluation w and a deliberative agent x), just in case w' satisfies every preference of x that w'' satisfies, and it also satisfies at least one preference not satisfied by w'' . (Notice that, following usual practice in formal semantics, the better world is the one to the left of $<$, not the one to the right.) To get the gist of this definition, assume that the ordering source contains three propositions, $P = \{w_1, w_2\}$, $Q = \{w_2\}$, and $R = \{w_3\}$. It follows from Definition 8 that $w_2 < w_1$ (that is, w_2 is strictly better than w_1), because w_2 satisfies preferences P and Q , and w_1 only satisfies preference P . On the other hand, w_3 is incomparable to both w_1 and w_2 , that is, neither $w_3 < w_1$ nor $w_1 < w_3$, and neither $w_3 < w_2$ nor $w_2 < w_3$. Consequently, x is predicted to prefer w_2 over w_1 , and to be undecided when considering w_3 as an alternative to w_1 and w_2 .

If we move from clause to (12) complete ‘ought’ sentences, we obtain:

$$(13) \quad \llbracket \text{Ought}(S, \phi) \rrbracket^{c,a,f,g} = \lambda w. \text{Best}(f(w), g(S, w)) \subseteq \llbracket \phi \rrbracket^{c,a,f,g}(S)$$

Here, ϕ has taken the place of P in (12), and S that of x . According to (13), a sentence of the form ‘ S ought to ϕ ’ at a context c , relative to a modal base f and an ordering source g , expresses a proposition that contains a world w just in case the best worlds in $\cap f(w)$ from the viewpoint of S at w , are worlds where S ϕ -ies.

The difference between the various flavors of deontic modality is provided by what counts as best from the viewpoint of S at w —that is, by the nature of the ordering source. (As we already remarked, for deliberative uses, the ordering source tracks the goals and preferences of the deliberative agent.) As for the syntactic realization of the modal base and the ordering source, I assume, for concreteness, that they are the values of covert variables introduced by the modal, and that when these variables occur free, context supplies their values.

Before moving on, notice that, in line with information-relativity, (13) allows for the relativization of the truth of an *ought*-claim to different bodies of information, as it does not anchor the relevant information to the deliberative agent. This is in contradistinction to the treatment of the ordering source, which is anchored to the goals and preferences of the deliberative agent through that same clause.

4. The information relativity of DMs

To get a sense of how this semantics works, let’s consider the predictions it makes in the simplest case, namely, what the deliberative agent S should do in the Miner situation, before any advice by A , relative to S ’s own information state. Of course, the semantics also yields predictions about other cases, like what S ought to do in that same practical situation, relative to A ’s information. However, we will see that, even in the simplest case, the prediction is wrong.

To draw the prediction for the simplest case, let's consider a doxastic modal base for 'ought' that contains all the information available to S in the Miner decision situation, before any intervention by A . Since this modal base characterizes the belief state of the deliberative agent S at the time of deliberation (at least as far as the relevant information is concerned), let's call it f_S . Relative to the world of deliberation w , f_S yields the propositions that the miners are either in shaft A or in shaft B, that the mine is about to flood, that we do not have enough materials to block both shafts, and propositions about the consequences of the different actions (*if we block shaft A, all water will go into shaft B; if all water goes into shaft B, everyone in it will drown; if we block neither shaft, water will go evenly into both shafts; if water goes evenly into both shafts, nine miners will survive; etc.*) It also contains propositions about likelihoods (thus, it includes propositions like *it is as likely that the miners are in shaft A as it is that they are in shaft B*). As for the ordering source, let g rank the worlds in $\cap f_S(w)$ according to how much they satisfy the preferences and goals of S during deliberation. Some care must be given to how the contents of the ordering source are spelled out, since not every way of doing so yields the required ordering (this is indeed a limitation of Kratzer's semantics with which I deal in Sect. 4.3). The following will do:

$$(14) \quad g(S, w) = \{\text{at least 0 miners survive, at least 9 miners survive, at least 10 miners survive}\}$$

In this case, worlds where ten miners survive are strictly better than worlds where nine miners survive, because the former satisfy three preferences in $g(S, w)$, namely that at least zero, at least nine, and at least ten miners survive, whereas the latter only two, namely that at least zero and at least nine miners survive. And worlds where nine miners survive are strictly better than those where no miners survive, for the same reason. To see how sensitive ordering is to the way the contents of the ordering source are conveyed, consider instead the following way of spelling out S 's preferences:

$$(15) \quad g(S, w) = \{\text{exactly 0 miners survive, exactly 9 miners survive, exactly 10 miners survive}\}$$

This way of giving the contents of the ordering source coincides with the known possible outcomes of S 's actions, the very information S uses during deliberation. However, if we had adopted (15) instead of (14), it would have followed from Definition 8, *inter alia*, that worlds where ten miners survive are incomparable to worlds where zero miners survive, so S should not have a preference for one outcome over the other, which is clearly wrong. So, we adopt (14) as the official way of spelling out the contents of the ordering source in the context of Kratzer's account.

In the discussion that follows, we simplify this entire procedure by thinking of g as simply entailing value assignments to the outcomes of the different possible actions in a way that reflects Definition 8 and the preferences in (14). Value assignments may be captured in a variety of ways. We choose the most straightforward approach, i.e., assigning an arbitrary utility value to each of them: the utility value of saving all ten miners is 10, the utility value of saving nine miners is 9, and the utility value of saving zero miners is 0.

4.1. Shortcomings of the standard framework

Let's see how the modal base and ordering source determine what S ought to do in the Miner situation, relative to the information S has at that point in time. Given that f_S contains the proposition that the miners are either in shaft A or in shaft B, $\cap f_S(w)$ contains both *inA*-worlds and *inB*-worlds (that is, worlds at which the proposition that the miners are in shaft A is true, and worlds where the proposition that the miners are in shaft B is true—the same reading applies to '*blA*-worlds', '*inA&blA*-worlds', and the like). Now, the best worlds in $\cap f_S(w)$ from the viewpoint of S 's preferences are the *inA&blA*-worlds and the *inB&blB*-worlds. That is, the best worlds are worlds where we block “the right shaft.” The $\neg blA \& \neg blB$ -worlds follow suit, and the *inA&blB*-worlds and the *inB&blA*-worlds are the worst. Thus, the standard Kratzerian semantics for DMs developed in Sect. 3 yields the ordering of worlds represented in Figure 2:

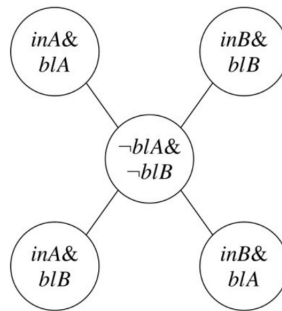


Figure 2: Information-insensitive ordering of possibilities in the Miner decision situation

That is, under standard semantics for DMs, in the best worlds, the agent either performs the action of *blocking-shaft-A-with-the-miners-in-A*, or that of *blocking-shaft-B-with-the-miners-in-B*. This is the prediction made concerning what agent S ought to do, in the sense of performing the action that best satisfies their preferences, relative to S 's own information state.

This, however, is the wrong result, for the course of action of “blocking the right shaft” is something the agent cannot decide to do, given their ignorance of the miners’

whereabouts. This incorrect prediction stems from an informationally unconstrained, or information-insensitive, ordering of possibilities. That is, it is made possible by an ordering source that can classify the blA -worlds into those in which the miners are in shaft A and those in which they are in shaft B, thus ordering them differently (and the same for the blB -worlds).

The ordering that results from this informationally unconstrained operation is at odds with the assumption that, given the limited information, the agent ought to block neither shaft. What we need to represent the choice situation of the Miner Paradox is, rather, an order like the one depicted in Figure 3:

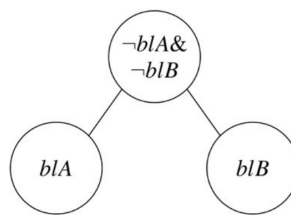


Figure 3: Ordering of possibilities relative to S's prior information

That is, we need the semantics to rank the $\neg blA \& \neg blB$ -worlds in general higher than the blA -worlds and blB -worlds, when taking into account the information S possesses at that point in the deliberation process. To achieve this ordering, we need to place an informational constraint on the ordering source.

4.2. Information relativity from a conceptual point of view

What is missing from the account is a way of properly incorporating information relativity into the semantics of DMs. Figure 2 roughly corresponds to the idea that, were the agent to gain further knowledge about the location of the miners, they could make an ideal choice that would bring about the best possible outcome. That is, the ordering of Figure 2 corresponds to what we may call an ideally discriminating ordering source: an ordering source that has access to perfect information about the worlds it is ordering.

Another way of addressing the issue is to consider the structuring of the logical space induced by the modal base and the ordering source. According to the standard semantics for DMs, modal base and ordering source operate along the lines depicted in Figure 4.

That is, the modal base f_S selects, from the common ground (or more precisely, from the context set (Stalnaker 1978)), those worlds compatible with the information it contains relative to the world of deliberation w (in the upper-left corner of Figure 4). Then, the ordering source gorders those worlds by considering how well each

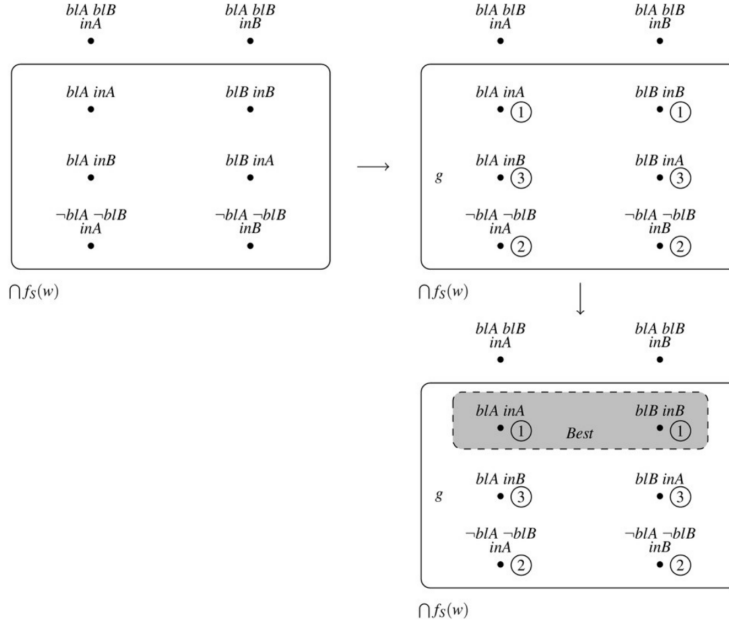


Figure 4: Information-insensitive carving of the logical space

satisfies the preferences of the agent S (upper-right corner), and the highest-ranking worlds are selected as best (lower-right corner). This process results in a structuring of the logical space into the ordered equivalence classes depicted in Figure 5:

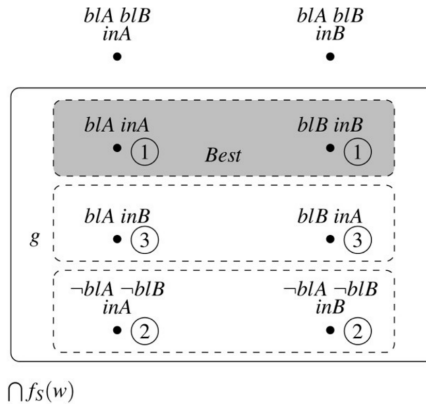


Figure 5: Information-insensitive structuring of the logical space

(Strictly speaking, there is a fourth equivalence class that comprises the worlds outside of $\cap f(w)$, i.e., the worlds that are not live options for the purposes of deliberation.)

eration.) Conceptually, this partitioning of logical space structures contextually live possibilities in a way that reflects the workings of an ordering source that can access complete information about the worlds it is ranking, thus distinguishing *blA*&*inA*-worlds from *blA*&*inB*-worlds, and *blB*&*inA*-worlds from *blB*&*inB*-worlds. In this way of structuring logical space, conceptually, the individual ranking of worlds comes first, and their grouping into equivalence classes comes as a consequence of that process.

However, making such distinctions between worlds requires access to precise information as to the whereabouts of the miners, information the agent does not have. This is the root of the incorrect prediction made by standard modal semantics for our test case: given the limited information that *S* possesses, *S* cannot distinguish *inA*-worlds from *inB*-worlds. Consequently, in ranking the worlds in f_S relative to *S*'s own, limited information, the ordering source should not be able to "look inside" the block of *blA*-worlds and distinguish those that are *inA*-worlds from those that are *inB*-worlds. And it should not be able to "look inside" the block of *blB*-worlds and distinguish those that are *inB*-worlds from those that are *inA*-worlds. That is, the ordering source should rank all *blA*-worlds as a whole and all *blB*-worlds as a whole. If the ordering source is constrained in this way, it will not rank *inA*&*blA*-worlds and *inB*&*blB*-worlds as best, and *inA*&*blB*-worlds and *inB*&*blA*-worlds as worst.

The issue at this point then seems to be how to constrain the power of discrimination of the ordering source, so that *g* can only rank *blA*-worlds, *blB*-worlds, and $\neg blA \& \neg blB$ -worlds as groups. Any internal distinctions we may impose on these sets should be beyond the fineness of grain with which *g* can look at the worlds in them. That is, what we need is a way to induce the structure depicted in Figure 6.

In this way of structuring logical space, we have an element, π , that carves the logical space determined by the modal base ($\cap f_S(w)$) into equivalence classes (upper right corner of Figure 6) *before* the ordering source takes charge and orders the worlds in terms of how well they satisfy the preferences of the agent (lower right corner). This carving of $\cap f_S(w)$ into equivalence classes should capture the way the information available to the agent allows them to distinguish between possibilities, thereby introducing an informational constraint on the ordering source. Conceptually, the partition of worlds into equivalence classes no longer comes as a consequence of the ranking of individual worlds by an informationally unconstrained ordering source. Rather, the partition of the logical space comes *before* the ordering source does its work, and restricts the way the ordering source ranks the worlds in $\cap f_S(w)$ by setting an informational constraint on which distinctions between worlds are accessible to it. (I say *conceptually* because the actual formal implementation will only capture the end-product of the structuring process described in Figure 6, i.e., the resulting ordering of the logical space, not its internal workings, i.e., the ordering source will not order equivalence classes directly, but only as a by-product of its informational constraint.)

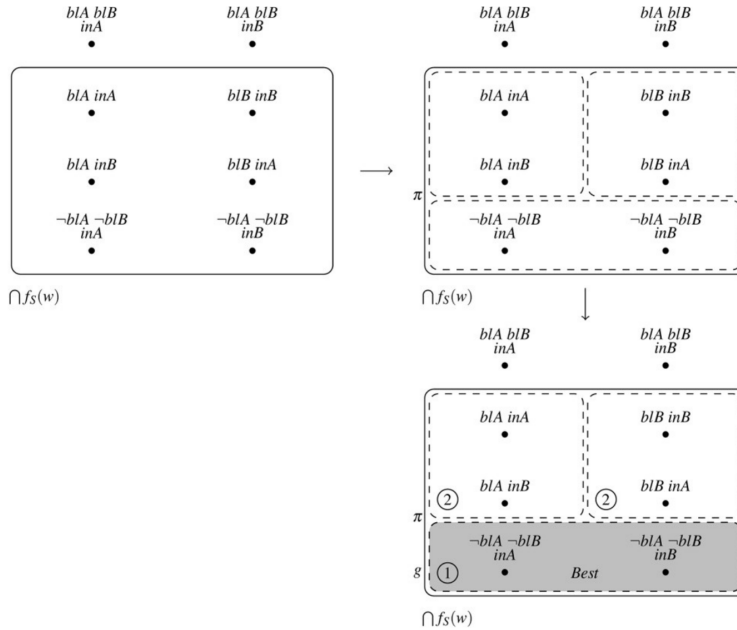


Figure 6: Informationally constrained carving of the logical space

4.3. Introducing information relativity into the semantics of DMs

In the previous subsections, we saw that standard Kratzerian semantics for DMs makes the wrong predictions concerning what S ought to do, in the Miner practical situation, relative to S 's own (incomplete) information. We also saw that, to overcome this, the semantics must incorporate information relativity into the way the ordering source orders worlds, so that, in ordering the worlds compatible with S 's information, it cannot make distinctions S cannot make. In this subsection, we show how information relativity can be incorporated into the ordering source, so as to yield correct predictions, both in this case and in the more complex adviser scenario, where the ordering of worlds is relative, not to S 's information, but to the adviser's.

In the current framework, the information relativity of the ordering source may be captured by adding an extra argument for a *partition* of the modal base. A partition Π of a set Σ is a set of non-empty subsets of Σ such that every element of Σ belongs to exactly one element of Π . The elements of Π are called its *cells*. Two elements of Σ are called cellmates, given a partition Π , just in case they belong to the same cell of Π . If two elements of Σ belong to the same cell of Π , then we say that Π cannot distinguish them. That is, a partition can make distinctions between elements just in case they belong to different cells. We say that Π is a partition of a modal base f , relative to a world w , just in case Π is a partition of $\cap f(w)$.

We now use partitions to account for information relativity in the following way: we posit that a deliberative ordering source g has an extra argument that selects a partition Π of the modal base f , and then we force g to consider any two worlds to be equally good just in case they are cellmates relative to Π . In this way, which possibilities the ordering source can discriminate (i.e., rank differently), and which should be treated as equal, is governed by the way in which Π partitions the modal base.

Not any partition of the modal base captures the way available information determines the order of worlds. What we need is a partition that comes from the choices available to the agent at the world and time of deliberation. Consequently, we let g receive an additional parameter, a partition Π or more properly, a function π from a world w , a modal base f , and the set of actions available to the agent at the world and time of deliberation, to partitions of $\cap f(w)$. Thus, g will now be a function $g(x, w, \pi)$, where π is a partition selection function.

With this in place, the informational constraint on the ordering source may be stated more precisely. To this end, let's introduce some notation. The basic function of the partition selection function π is to map modal bases to partitions thereof, which group the worlds in the modal base by considering the actions performed in them by the deliberative agent. So, officially, π is a function of the form $\pi(x, f, w, t)$. To avoid notational clutter, we assume in what follows that a particular selection of values for π is provided and denote by Π the partition that is the value of π given those arguments. We also slightly abuse the notation for g , and let Π occur in it, rather than π together with all its parameters. A final piece of notation: given a partition Π of a set Σ and an element σ of Σ , we let $[\sigma]_{\Pi}$ denote the cell of Π to which σ belongs. If two elements $\sigma, \sigma' \in \Sigma$ belong to the same cell of Π , that is, if $[\sigma]_{\Pi} = [\sigma']_{\Pi}$, then σ and σ' are indistinguishable as far as Π can discriminate.

Keeping this in mind, the informational constraint on g , more formally expressed, is:

$$(16) \quad u =_{g(x, w, \Pi)} v \Leftrightarrow [u]_{\Pi} = [v]_{\Pi}$$

Here, $=_{g(x, w, \Pi)}$ is the equivalence relation derived from the ordering imposed by the ordering source, and Π is a partition of the modal base determined by the actions available to the deliberative subject at the time and world of deliberation.⁵ Clause (16) states that the ordering source cannot make distinctions between worlds that the partition treats as equals. Rather, it must treat them as equals with respect to the fulfilment of the deliberative agent's goals and preferences. Since the partition distinguishes worlds based exclusively on the actions performed in them, the result is that the ordering source can distinguish worlds only on that same basis. Any information beyond that (particularly, information about the outcome of an action at a given world) is, in principle, beyond the reach of the ordering source.

To implement this informational restriction, the way in which the ordering relation $<_g$ is derived from an ordering source g must be modified. The most straightforward way is:

Definition 9 (Preliminary version). For all $w', w'' \in \cap f(w)$, $w' <_{g(x, w, \Pi)} w''$ iff for all $P \in g(x, w, \Pi)$, if $[w'']_{\Pi} \subseteq P$, then $[w']_{\Pi} \subseteq P$, and there is a $P \in g(x, w)$ such that $[w']_{\Pi} \subseteq P$ and $[w'']_{\Pi} \not\subseteq P$.

That is, a world w' is strictly better than another world w'' relative to a partition Π of the modal base (and with respect to a world of evaluation w and an agent x) just in case (i) every preference in $g(x, w, \Pi)$ whose satisfaction is guaranteed by performing action $[w'']_{\Pi}$ is also guaranteed to be satisfied by performing action $[w']_{\Pi}$, and (ii) there is a preference in $g(x, w, \Pi)$ whose satisfaction is guaranteed by performing $[w']_{\Pi}$, but not by performing $[w'']_{\Pi}$. Notice that this encodes the idea that the ordering source cannot access the full information about what particular outcome obtains in a world w , but only what outcome is known to obtain relative to the possibly less specific information contained in the class of worlds to which w belongs.

After adding the new parameter for partitions, the semantic clause for a sentence of the form ‘ S ought to ϕ ’ becomes:

$$(17) \quad \llbracket \text{Ought}(S, \phi) \rrbracket^{c, a, f, g} = \lambda w. \text{Best}(f(w), g(S, w, \Pi)) \subseteq \llbracket \phi \rrbracket^{c, a, f, g}(S)$$

According to this modified clause, whether S ought to ϕ , at a context c , relative to a modal base f and an ordering source g , ultimately depends on the goals and preferences of the deliberative agent (as it did before), as well as on whatever information is available at the context of utterance, both in the form of which possibilities are considered as live (i.e., those in $\cap f(w)$) and which actions are available to the deliberative agent at the time of utterance (encoded in Π). Notice that the relevant partition Π (officially, the partition selection function π) does not occur as a parameter of the interpretation function $\llbracket \cdot \rrbracket$. In the current setting, such an inclusion is not required, for the partition (i.e., the selection function) is not subject to contextual determination but always operates by drawing on the set of available actions. However, it could be added as a further parameter of $\llbracket \cdot \rrbracket$, should there exist other interpretations of modals that call for different informational constraints. I leave this as an open empirical question.

Let’s apply this reformulated semantics to our working example. In the Miner Paradox, the available options are blocking shaft A, blocking shaft B, and blocking neither shaft. Consequently, the relevant partition will carve the modal base f_S (relative to the world of deliberation w) into three cells: the *bIA*-cell (which contains exactly those worlds in $\cap f_S(w)$ in which the agent blocks shaft A), the *bIB*-cell (which

	<i>inA</i>	<i>inB</i>	Expected Value
<i>blA</i>	10	0	$\text{ev}(\text{blA}) = 10\text{p}(\text{inA}) + 0\text{p}(\text{inB}) = 5$
<i>blB</i>	0	10	$\text{ev}(\text{blB}) = 0\text{p}(\text{inA}) + 10\text{p}(\text{inB}) = 5$
$\neg\text{blA} \& \neg\text{blB}$	9	9	$\text{ev}(\neg\text{blA} \& \neg\text{blB}) = 9\text{p}(\text{inA}) + 9\text{p}(\text{inB}) = 9$

Table 1: Expected values relative to *S*'s prior information state (f_S)

contains exactly those worlds in $\cap f_S(w)$ in which the agent blocks shaft B), and the $\neg\text{blA} \& \neg\text{blB}$ -cell (which contains exactly those worlds in $\cap f_S(w)$ in which the agent blocks neither shaft). As for the contents of the ordering source, unlike Definition 8, Definition 9 allows the adoption of (15) as a way of spelling out *S*'s preferences. Indeed, relative to *S*'s information before any advice, all cells of the partition contain both *inA* and *inB*-worlds. Given this, the *zero.survive* and *ten.survive* propositions cut across cell boundaries, and only the *nine.survive* proposition is contained within a single cell:

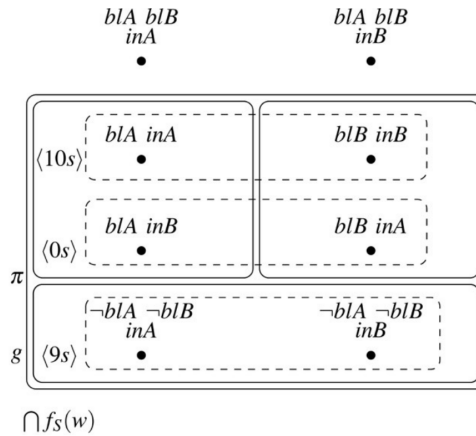


Figure 7: Satisfaction of preferences by available actions under uncertainty

Hence there is no possible action available to *S* that guarantees the *zero.survive* and *ten.survive* outcomes. The *nine.survive* proposition, on the other hand, is guaranteed to obtain by blocking neither shaft. An application of Definition 9 shows that the worlds in the $\neg\text{blA} \& \neg\text{blB}$ -cell are deemed strictly better than those in the other two cells, hence they come on top, and blocking neither shaft is deemed the best available choice. The result is the ranking of partitions depicted in Figure 8.

In the same way as before, we may represent this ordering in terms of expected utility, and we will do so for the ensuing discussion. The result is that the actions of blocking shaft A and of blocking shaft B have lesser expected value than the action of blocking neither, as Table 1 shows.

Since f_S captures the belief state of the deliberative agent at the time of deliber-

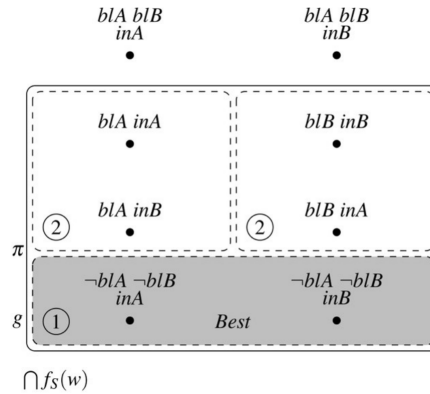


Figure 8: Structuring of logical space relative to S 's prior information (f_s)

ation, Table 1 captures the expected values of the actions available to S at that time, relative to S 's prior information. Relative to this information, blocking neither shaft is the dominant action. Hence, the $\neg\text{blA} \& \neg\text{blB}$ class of worlds is ranked higher than the blA and blB classes (which have the same expected value relative to that information state), and we obtain the desired order of Figure 3.⁶

Let's take stock. Standard Kratzerian semantics cannot appropriately capture the order of possibilities in deliberation under uncertainty. Consequently, partitions were added to the semantics of DMs as a way of formally representing their information relativity. With this modification, the semantics now yields the correct prediction for the simplest case, namely, what S ought to do relative to their own information: what S ought to do, relative to S 's information before any advice, is to block neither shaft. However, as the semantics currently stands, correct predictions for the other cases cannot be derived. This is due to a limitation of Kratzer's account of the ordering source we have already pointed out, and that now surfaces (given the current, preliminary definition of order) when we move back to deliberation under certainty about outcomes. As we already noticed, Definition 8 entails that there are worlds that are incorrectly deemed incomparable. The same problem affects Definition 9 when we consider deliberation under certainty, since in that case, each cell of the relevant partition entails a different outcome (Figure 9). So, by Definition 9, they are all incomparable. Hence, the agent is incorrectly predicted to have no preference for any action over any other. The issue is that Kratzer's account of order treats every preference as having the same weight, and this is incorrect. And this issue is inherited by Definition 9. So, we either come up with a clever way of spelling out the contents of the ordering source, as in (14), or we add weights to the propositions in (15). Since adding weights is the most general solution, we do so in:

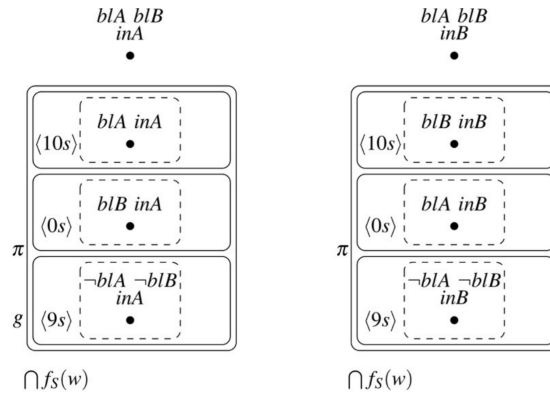


Figure 9: Satisfaction of preferences by available actions under certainty

Definition 10 (Final version). For all $w', w'' \in \cap f(w), w' <_{g(x,w,\Pi)} w''$ iff $\sum_{P \in \text{sat}(w', \Pi)} u_x(P) > \sum_{P \in \text{sat}(w'', \Pi)} u_x(P)$,

where u_x is the utility function of agent x and $\text{sat}(v, \Pi)$ is the set of all preferences $P \in g(x, w, \Pi)$ such that $[v]_{\Pi} \subseteq P$, that is, the set of all preferences guaranteed to be satisfied by performing action $[v]_{\Pi}$. According to Definition 10, then, a world w' is strictly better than another world w'' relative to a partition Π of the modal base (and with respect to a world of evaluation w and an agent x) just in case the aggregated utility of all the preferences in $g(x, w, \Pi)$ whose satisfaction is guaranteed by performing action $[w']_{\Pi}$ is greater than the aggregated utility of all the preferences in $g(x, w, \Pi)$ whose satisfaction is guaranteed by performing action $[w'']_{\Pi}$.

Going back to Figure 9, we can now derive the correct predictions for deliberation under certainty as well. Let's concentrate on the case where the miners are known to be in shaft A (to the left of the figure). In this case, the blA -cell entails the *ten.survive* ($\langle 10s \rangle$) outcome, the $\neg blA \& \neg blB$ -cell entails the *nine.survive* ($\langle 9s \rangle$) outcome, and the blB -cell entails the *zero.survive* ($\langle 0s \rangle$) outcome. Under Definition 9, all cells turned out to be incomparable. Now, after factoring in the corresponding value assignments for each outcome according to Definition 10, blocking shaft A is correctly predicted to be the preferred action: the blA -cell has the highest aggregated utility value (with a value of $u_S(\langle 10s \rangle) = 10$), the $\neg blA \& \neg blB$ -cell follows suit (with $u_S(\langle 9s \rangle) = 9$), and the blB -cell is ranked lowest (with $u_S(\langle 0s \rangle) = 0$).

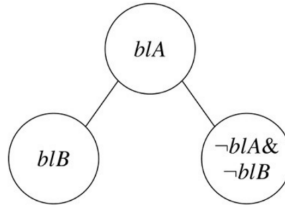
We can now derive correct predictions for the more complex advice scenario as well. As we did before, we use expected values as a way of abstracting away from the particulars of how the ordering source does its work. Indeed, with information-relativity incorporated into the semantics through partitions, we can also explain the truth of the adviser's injunction to block shaft A in (3), relative to the adviser's

	<i>inA</i>	<i>inB</i>	Expected Value
<i>blA</i>	9	9	$ev(blA) = 9p(inA) + 9p(inB) = 9$
<i>blB</i>	0	10	$ev(blB) = 0p(inA) + 10p(inB) = 5$
$\neg blA \& \neg blB$	10	0	$ev(\neg blA \& \neg blB) = 10p(inA) + 0p(inB) = 5$

Table 2: Expected values relative to A's information state

own information. Let's briefly recall that *S*'s adviser *A* possesses better information than *S* and knows that blocking shaft A is the action that *S* ought to perform (not because this will save more lives, but because it will have the desired outcome of saving nine miners with certainty). We may describe the situation as follows. *A* is starting from a different modal base f_A that contains all the information *A* has concerning hydraulics and the consequences of the different actions predictable given that knowledge (plus the mutually shared, incomplete information about the miners' whereabouts, etc.) Given all this information, *A* works with a different decision matrix that makes blocking shaft A the dominant action, as captured by Table 2:

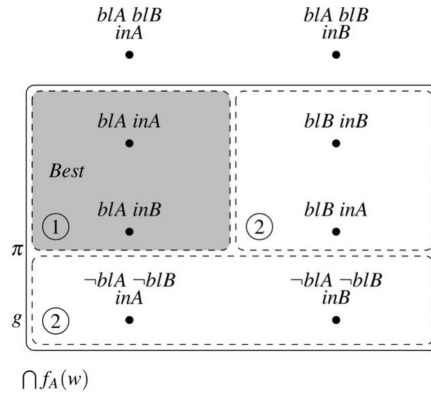
Provided that the ordering source $g(S, w, \Pi)$ makes the value assignments captured by Table 2, *A* gets the ordering of the worlds in $R \cap f_A(w)$ depicted in Figure 10.

Figure 10: Ordering of possibilities relative to A's information (f_A)

That is, given *A*'s information, the best course of action for *S* is to block shaft A (that is, the best worlds in $\cap f_A(w)$ according to $g(S, w, \Pi)$ are *blA*-worlds). This results from a different way of carving the logical space, produced by a different information state f_A (Figure 11).

Finally, *S*'s change of mind in (4) after hearing *A*'s explanation can be accounted for by assuming that *S*'s posterior information state is, for all practical purposes, equivalent to *A*'s, in the sense that it yields the expected values in Table 2, and thus the ordering of worlds in Figure 10, via the structuring of logical space depicted in Figure 11.

These considerations explain the intuitive truth of sentences (2)–(4): relative to *S*'s prior information state f_S , blocking neither shaft is what *S* ought to do (Figure 3 and Figure 8); relative to *A*'s information state f_A , blocking shaft A is what *S* ought

Figure 11: Structuring of logical space relative to A's information (f_A)

to do (Figure 10 and Figure 11); and relative to S 's posterior information state $f_{S'}$, blocking shaft A is what S ought to do (once again, Figure 10 and Figure 11).

It should be noted that the semantics developed in this section resembles Cariani's resolution semantics (Cariani 2013) in an important respect: both make use of partitions to refine the interpretation of DMs. The two approaches, however, differ crucially. The central difference is that Cariani coerces DMs into quantifying over the cells of the partition instead of possible worlds. Thus, he offers what he calls an anti-boxing account of DMs, i.e., an account that forgoes the interpretation of DMs as restricted quantifiers over possible worlds and treats them instead as quantifiers over sets of possible worlds. In the semantics developed here, partitions are an extra parameter of the ordering source, and they impose a constraint on the way an ordering source orders the worlds compatible with the modal base (i.e., by forcing the ordering source to treat equally worlds where the same action is performed). In contradistinction to Cariani's account, the DM itself still quantifies over possible worlds, as in standard, Kratzerian semantics. In Cariani's terms, I am proposing a boxing account of DMs. An important consequence of this way of incorporating partitions into the semantics of DMs, abandoned in Cariani's, is that DMs have the same basic semantic behavior as other natural language modals, thus belonging to the same natural semantic class: a unified theory of natural language modals is still possible under the account proposed in Sect. 3 and 4, whereas such unification is impossible under Cariani's resolution semantics.

A final note on the role of partitions in the present account, which may help better understand the difference between the theories. The informational constraint on the ordering source has been described as an ordering of partition cells. This, however, was a conceptual description of how the ordering source works (a heuristic device to grasp the informational constraint more straightforwardly). At the level

of formal implementation, the ordering source still works as in Kratzer's account: it ranks worlds individually, not classes of worlds. The ordering of the classes themselves is simply a by-product of the fact that the ordering source is informationally constrained to equally rank cellmate worlds (i.e., of the fact that it is not allowed to make distinctions between worlds that are cellmates). The semantics proposed in this section mimics the conceptual description (it captures the resulting structure of the logical space) without implementing the logic behind that description (without making partition cells part of the domain of the ordering source). Thus, even though conceptually we may think of the informational constraint of the ordering source as introducing a partition whose cells are then ordered by the ordering source, formally, the ordering source still ranks worlds individually, thereby maintaining the Kratzerian view of modals as restricted quantifiers over possible worlds.

5. The invalidity of *modus ponens*

In the previous section, I showed how information relativity can be incorporated into the semantics of DMs within a standard, Kratzerian framework for natural language modals. The main contention of this semantics is that there is no single answer across all contexts, concerning what any given agent ought to do in a particular practical situation (in the deliberative, information-relative sense of 'ought'). On the contrary, different contexts characterized by different bodies of relevant information yield different verdicts as to what an agent ought to do. Through partitions of the contextually relevant information based on the actions available to the deliberative agent at the time of deliberation, the relativization of DMs to contextually salient information states yields the correct predictions for sentences (2)–(4). This meets the first semantic desideratum highlighted by Kolodny and MacFarlane. In this section, I address the second desideratum, the principled invalidity of *modus ponens* for the indicative conditional, by showing how the semantics developed in Sect. 4, together with a natural account of the indicative conditional, invalidates that rule.

I follow Kratzer (1981a, 1981b, 1986) in considering that the antecedent of an indicative conditional works as a restrictor of an operator occurring in the consequent. The general clause for a conditional statement is:

$$(18) \quad \llbracket \text{If } \phi, \psi \rrbracket^{c,a,f,g} = \llbracket \psi \rrbracket^{c,a,f^{+\phi},g}$$

where $f^{+\phi}$ is the modal base that results from adding the proposition expressed by ϕ at c to f , i.e., the modal base such that, for any w , $f^{+\phi}(w) = f(w) \cup \{\llbracket \phi \rrbracket^{c,a,f,g}\}$. According to this clause, an indicative conditional instructs us to evaluate its consequent relative to a modal base enlarged by the addition of the antecedent. If the consequent holds relative to this informationally enriched modal base, the conditional

holds relative to the original, unenriched modal base; otherwise, it is false. This is conceptually equivalent to the operation of assuming that the antecedent holds and assessing whether the consequent holds under that assumption as well. This understanding of the indicative conditional is in line with the intuition highlighted in Sect. 2 through Ramsey's test for conditionals, and it also agrees with MacFarlane's understanding of the indicative conditional in (10). Crucially, though, in the Kratzerian story for indicative conditionals, the consequent is always governed by an operator, which can be either explicit or implicit. That is, in an indicative conditional of the form 'If ϕ , ψ ', under this analysis, ψ is always of the form $\text{Op}(\chi)$. Thus, the enriched modal base $f^{+\phi}$ forms the information background against which to evaluate the operator occurring in the consequent. If there is no overt operator in the consequent, a covert epistemic necessity modal is posited.⁷ In the case we are interested in, though, there is an overt modal ('ought'). Thus, for the current discussion, a deontic conditional works by adding the proposition that ϕ to the modal base relative to which the 'ought' occurring in the consequent is to be interpreted. The procedure for evaluating a deontic conditional amounts, then, to the following: assume that the antecedent of the conditional holds (i.e., temporarily enrich the information contained in the modal base with the proposition expressed by the antecedent) and determined whether the obligation expressed by the consequent holds under that assumption (i.e., whether the bare deontic sentence that occurs in the consequent is true relative to the informationally enriched modal base). If it does, the deontic conditional is true relative to the original, unenriched modal base. Otherwise, it is false.

For (5) and (6), the conditionals involved in the Miner Paradox, (18) yields the following interpretations:

$$(19) \quad \llbracket \text{If } \text{inA}, \text{Ought}(S, \text{bLA}) \rrbracket^{c,a,f,g} = \llbracket \text{Ought}(S, \text{bLA}) \rrbracket^{c,a,f^{+\text{inA}},g}$$

$$(20) \quad \llbracket \text{If } \text{inB}, \text{Ought}(S, \text{bLB}) \rrbracket^{c,a,f,g} = \llbracket \text{Ought}(S, \text{bLB}) \rrbracket^{c,a,f^{+\text{inB}},g}$$

That is, the conditional in (5) is true just in case its consequent holds when evaluated with respect to $f^{+\text{inA}}$, i.e., the modal base f enriched by the assumption that the miners are in shaft A. The same applies, *mutatis mutandis*, to the conditional in (6). The propositions expressed by those conditionals relative to particular values for c, a, f, g are:

$$(21) \quad \lambda w. \text{Best}(f^{+\text{inA}}(w), g(S, w, \Pi)) \subseteq \llbracket \text{bLA}(S) \rrbracket^{c,a,f^{+\text{inA}},g}$$

$$(22) \quad \lambda w. \text{Best}(f^{+\text{inB}}(w), g(S, w, \Pi)) \subseteq \llbracket \text{bLB}(S) \rrbracket^{c,a,f^{+\text{inB}},g}$$

In other words, the first conditional says that the best worlds from the viewpoint of S , relative to a modal base that includes the information that the miners are in shaft A (and the partition provided by the actions available to S at the time of the

context), are worlds where S blocks shaft A . The same goes, *mutatis mutandis*, for the second conditional.

It is now straightforward to show that *modus ponens* for the indicative conditional is invalid under this semantics. First, we introduce a definition of logical consequence parallel to Definition 6 of Sect. 2:

Definition 11. S is a logical consequence of Γ iff, for all c, a , if all $\gamma \in \Gamma$ are true at c, a , then S is true at c, a .

where a sentence is true at a context c under an assignment a just in case it is true relative to c, a, w_c, f_c, g_c . In this case, a formula ϕ is true relative to c, a, w, f, g just in case $w \in \llbracket \phi \rrbracket^{c,a,f,g}$. Note that Definition 11 also understands logical consequence as preservation of truth-at-a-context for sentences, or truth preservation for occurrences of sentences. In a nutshell, Definition 11 states that, for an argument from Γ to S to be valid, there can be no context c such that every sentence in Γ is true at c , and S is false at c . To provide a counterexample to *modus ponens* for the indicative conditional, we need to construct one such context.

To have a concrete instance of *modus ponens* to discuss, let's focus again on the Miner Paradox, which crucially appeals to instances of *modus ponens* like the following:

- (23) a. If inA , $Ought(S, blA)$
 b. inA
 c. $Ought(S, blA)$

If *modus ponens* is to be invalid, and DMs are to provide a counterexample to it, we should be able to construct a context that makes both (23a) and (23b) true, and (23c) false. Constructing such a context is fairly easy.

Since the modal base and the ordering source are determined by context, we focus on contexts with doxastic modal bases that contain all the information available to S at the time of deliberation and ordering sources that capture S 's preferences regarding the practical situation of the Miner Paradox. Also, we need to focus only on contexts whose worlds are inA -worlds, for only such contexts will make (23b) true.

Now recall that, under the semantics for conditionals proposed in this section, a conditional is true at a context c just in case its consequent is true when evaluated relative to a temporary modal base that has been informationally enriched with the assumption that the antecedent of the conditional holds. For the case we are concerned with, the conditional in (23a), 'If inA , $Ought(S, blA)$ ', is true at any context c that complies with the restrictions we are now imposing, since ' $Ought(S, blA)$ ' is true at c relative to an enriched modal base f_c^{+inA} . To see how this is so, we must show that the ordering source induces on $\cap f_c^{+inA}(w_c)$ the ordering depicted in Figure 12.



Figure 12: Ordering of possibilities relative to the assumption that the miners are in shaft A

This is the crux of the invalidity result for *modus ponens*. By the restrictions that were placed on the context we are constructing, (23b) is true by stipulation. If we can show that temporarily adding the information that the miners are in shaft A to a modal base that contains *S*'s prior information results in the order of worlds depicted in Figure 12, we will have shown that (23a) is true in that same context as well. After that, all we need to do is to show that (23c) is false in that context, which yields a counterexample to *modus ponens*.

The order in Figure 12 results from the carving of logical space depicted in Figure 13.

Recall the explanation target here: we need to show that the ordering source induces on $\cap f_c^{+inA}(w_c)$ an order where the *blA*-worlds come on top. Since f_c^{+inA} contains the assumption that the miners are in shaft A, the worlds selected by this temporarily enriched modal base relative to the world of the context w_c are all *inA*-worlds (upper left corner). Once again, the actions available to the agent induce a three-cell partition Π of $\cap f_c^{+inA}(w_c)$ into the *blA*-worlds, the *blB*-worlds, and the $\neg blA \& \neg blB$ -worlds (upper right corner). Since $\cap f_c^{+inA}(w_c)$ contains only *inA*-worlds, the ordering source knows (that is, relative to the assumption that the miners are in shaft A) that all the *blA*-worlds in $\cap f_c^{+inA}(w_c)$ are *inA*-worlds, and that all the *blB*-worlds in $\cap f_c^{+inA}(w_c)$ are *inA*-worlds. So, now the expected values for the different actions are as depicted in Table 3.

The ordering source follows these expected values in ranking the available classes of worlds (lower left corner of Figure 13), and the order depicted in Figure 12 is obtained. Thus, '*Ought*(*S*, *blA*)' comes out true at *c*, *a* relative to w_c, f_c^{+inA}, g_c , which entails that '*If inA*, *Ought*(*S*, *blA*)' is true at *c*, *a* relative to w_c, f_c, g_c . Notice that this ordering is obtained, not because the ordering source can make internal distinctions

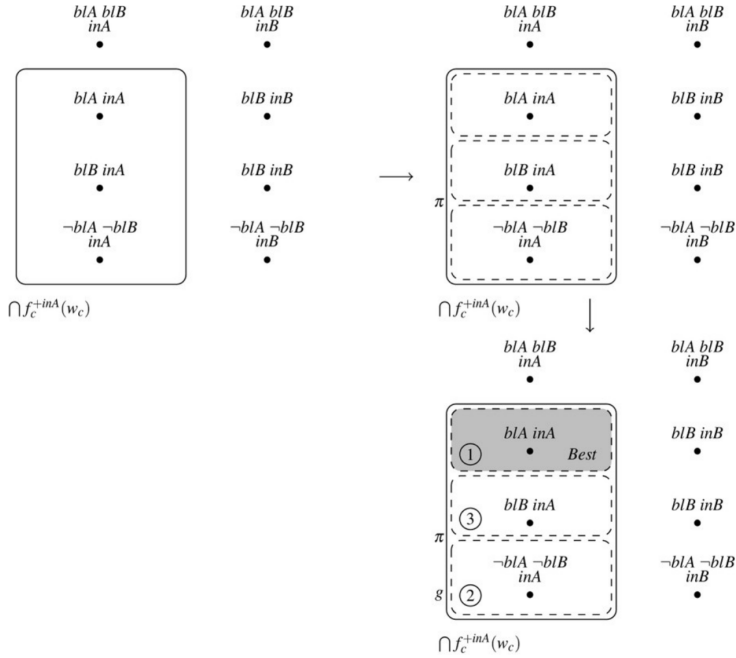


Figure 13: Structuring of the logical space relative to the assumption that the miners are in shaft A

	<i>inA</i>	<i>inB</i>	Expected Value
<i>blA</i>	10	0	$ev(blA) = 10p(inA) + 0p(inB) = 10$
<i>blB</i>	0	10	$ev(blB) = 0p(inA) + 10p(inB) = 0$
$\neg blA \ \& \ \neg blB$	9	9	$ev(\neg blA \ \& \ \neg blB) = 9p(inA) + 9p(inB) = 9$

Table 3: Expected values relative to the assumption of proposition *inA*

within the cells of the partition, but because it has external warrant (in the form of determinate, even if only suppositionally added, information concerning the whereabouts of the miners) to the effect that all the worlds in each cell are worlds where the miners are in shaft A.

Let's relate this result to the informal discussion of conditionals in Sect. 2.2. There, the truth of the deontic conditionals under an information-relative reading of DMs was motivated by citing that, under the assumption that the antecedent is true, the bare deontic statement in the consequent also holds, because the assumption incorporated into the informational background relative to which the consequent is evaluated need not be known by the deliberative agent (neither in the actual situation, nor in the temporarily enriched scenario with respect to which the consequent is being evaluated). This is the distinctive way in which information-relative read-

ings of DMs are relativized to a body of information. The current semantics captures this nicely. Firstly, the evaluation of an indicative conditional involves adding the antecedent to the modal base relative to which the consequent is evaluated. This captures the intuition that, in evaluating a conditional, we add the supposition that the antecedent holds to our current information, and evaluate whether the consequent holds as well, relative to that enriched informational background. Secondly, a DM statement of the form ‘ S ought to ϕ ’ is evaluated as true, relative to a modal base f , just in case the best worlds in f are ϕ -worlds. It is not required by the semantics of DMs that the information contained in f be available to the deliberative agent for the corresponding ‘ought’ claim to be true. This captures the information-relative reading of DMs, as opposed to their purely subjective reading. With these two features in place, the evaluation of the deontic conditionals in the Miner Paradox as true parallels the informal argument in favor of their truth. This also explains why the conditional premise in (23a) is true.

Thus far, we have constructed a context in which both (23a) and (23b) are true. To provide a counterexample to *modus ponens*, we now need to show that the conclusion of the argument, (23c), does not hold in that context. This is straightforward, as the action of blocking shaft A, given the unmodified modal base f_c that contains all the information S possesses before the advice, is ranked lower than the action of blocking neither shaft, as shown in Figure 3 (and as entailed by Table 1). Thus, ‘Ought(S , bla)’ comes out false at c .

We should note that the invalidity of *modus ponens* for the indicative conditional, far from being an *ad hoc* move to block a deontic paradox, follows from a straightforward account of the indicative conditional in line with the standard Kratzerian take, and from the informational character of DMs (which accounts for the truth of the consequent relative to the assumption of the antecedent, regardless of whether the deliberative agent comes to learn that information in the hypothetical evaluation scenario). Notice that the invalidity result cannot be reproduced if we adopt a subjective interpretation of DMs instead of an information-relative one: under a subjective interpretation, the conditionals in the Miner Paradox, thus the conditional premise in (23), are false, and a counterexample to *modus ponens* cannot be constructed on that basis. It is only when information-relativity is brought into the picture that *modus ponens* for the indicative conditional becomes an invalid form of reasoning. Thus, Kolodny & MacFarlane’s second desideratum for an adequate account of DMs is also met in a principled way.

6. Conclusion

Throughout this paper, I attempted to defend the claim that we can capture the distinctive semantic behavior of DMs within a standard Kratzerian view of modality without positing non-standard, information-neutral semantic values (contra Kolodny & MacFarlane (2010) and MacFarlane (2014, Chapter 11)). I proposed a Kratzerian semantics of DMs that accounts for their information-relative interpretation by relativizing their ordering source to a partition provided by the actions available to the deliberative agent at the time of deliberation and showed how this allows us to explain the linguistic data concerning information-relativity, and thus the details of the Miner decision situation. I then showed how this account, together with a standard Kratzerian account of the indicative conditional, invalidates *modus ponens*, thereby also resolving the Miner Paradox, without rejecting as false the conditional premises involved in the paradoxical argument. With these results, Kolodny & MacFarlane (2010)'s semantic *desiderata* are met: not only is the informational character of DMs captured by the semantic proposal in Sect. 4, but also that character plays a central role in the invalidity of *modus ponens*. Hence, we do not need to deviate from standard possible-world semantic values to account for the distinctive semantic features of DMs: a standard semantic account in terms of propositions as sets of possible worlds is enough both to capture the information relativity of DMs and to invalidate *modus ponens* for the indicative conditional in a principled way.⁸

References

- Barbiers, S. 2001. Current issues in modality. In: S. Barbiers; F. Beukema; W. van der Wurff (eds.), *Modality and its Interaction with the Verbal System*, p.11–29. Amsterdam: John Benjamins.
- Bhatt, R. 1998. Obligation and possession. In: H. Harley (ed.), *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, p.21–40. Cambridge: MITWPL.
- Brennan, V. 1993. *Root and epistemic modal auxiliary verbs*. PhD thesis, University of Massachusetts, Amherst.
- Cariani, F. 2013. 'Ought' and resolution semantics. *Nous* 47(3): 534–558.
- Chisholm, R. 1963. Contrary-to-duty imperatives and deontic logic. *Analysis* 24(2): 33–36.
- Einheuser, I. 2008. Three forms of truth relativism. In: M. García-Carpintero; M. Kölbel (eds.), *Relative Truth*, p.187–203. Oxford: Oxford University Press.
- Gillies, A. 2004. Epistemic conditionals and conditional epistemics. *Nous* 38(4): 585–616.
- Kaplan, D. 1989. Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In: J. Almog; J. Perry; H. Wettstein (eds.), *Themes From Kaplan*, p.481–563. New York: Oxford University Press.
- Kolodny, N.; MacFarlane, J. 2010. Ifs and oughts. *The Journal of Philosophy* 107(3): 115–143.
- Kratzer, A. 1981a. The notional category of modality. In: H. Eikmeyer; H. Rieser (eds.), *Words, Worlds, and Contexts*, p.38–74. Berlin: de Gruyter.

- Kratzer, A. 1981b. Blurred conditionals. In: W. Klein; W. Levelt (eds.), *Crossing the Boundaries in Linguistics*, p.201–209. Dordrecht: Reidel.
- Kratzer, A. 1986. Conditionals. In: A. Farley; P. Farley; K. McCollough (eds.), *Papers from the Parasession on Pragmatics and Grammatical Theory*, p.115–135. Chicago: Chicago Linguistics Society.
- Kratzer, A. 1991. Modality. In: A.v. Stechow; D. Wunderlich (eds.), *Semantics: An International Handbook of Contemporary Research*, p.639–650. Berlin: de Gruyter.
- Lewis, D. 1973. *Counterfactuals*. Cambridge, MA: Harvard University Press.
- MacFarlane, J. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Clarendon Press.
- Portner, P. 2009. *Modality*. Oxford: Oxford University Press.
- Ramsey, F. 1931. General propositions and causality. In: J.B. Braithwaite (ed.), *The Foundations of Mathematics and Other Logical Essays*, p.237–255. London: Kegan Paul, 237–255.
- Roberts, I. 1985. Agreement parameters and the development of English auxiliaries. *Natural Language and Linguistic Theory* 3: 21–58.
- Stalnaker, R. 1978. Assertion. In: P. Cole (ed.), *Syntax and Semantics*, 9: Pragmatics, p.315–332. New York: Academic Press.
- Wright, C. 2008. Relativism about truth itself: Haphazard thoughts about the very idea. In: M. García-Carpintero; M. Kölbel (eds.), *Relative Truth*, p.157–186. Oxford: Oxford University Press.
- Wurmbrand, S. 1999. Modal verbs must be raising verbs. In: S. Bird; A. Carnie; J. Haugen; P. Norquest (eds.), *Proceedings of the 18th West Coast Conference on Formal Linguistics*, p.599–612. Somerville, MA: Cascadia.
- Yalcin, S. 2007. Epistemic modals. *Mind* 116(464): 983–1026.
- Zubizarreta, M. 1982. *On the relationship of the lexicon to syntax*. PhD thesis, MIT.

Notes

¹There are two ways of revising the logical form of the paradox, both related to the logical form of the deontic conditionals. The first proposal considers that the DM scopes over the conditional, so that instead of parsing these premises as ‘If ϕ , *Ought*(ψ)’, we should parse them as ‘*Ought*(If ϕ , ψ)’. The second proposal is to replace the deontic conditional with a dyadic deontic operator, thus rendering a deontic conditional as ‘*Ought*(ϕ | ψ)’. The reader is referred to Kolodny and MacFarlane’s discussion of these options for arguments against them.

²This logical form is based on two ideas. First, that premises 3 and 4 involve conditionals (as opposed to, say, dyadic operators). Second, that ‘ought’ behaves like a control predicate in the sense that it takes two arguments, an individual and a proposition (Zubizarreta 1982, Roberts 1985, Brennan 1993). As almost everything, this is disputed (Bhatt 1998, Wurmbrand 1999, Babiers 2001). It is, however, a safe assumption to make in this context, and is in line with the syntactic assumptions made by MacFarlane (2014, Chapter 11).

³This is not the only deontic paradox that puts pressure on *modus ponens*. Indeed, Chisholm’s paradox (Chisholm 1963) makes a claim for the incompatibility between factual detachment (i.e., instances of *modus ponens* of the form ‘If ϕ , *Ought*(ψ); $\phi \models \text{Ought}(\psi)$ ’) and

deontic detachment (inferences of the form ‘If ϕ , *Ought*(ψ); *Ought*(ϕ) \models *Ought*(ψ)’). Some, like Lewis (1973) (and Kratzer (1991) following suit), have taken it that we should abandon *modus ponens* (factual detachment) as a solution to this problem. Noteworthy about the Miner Paradox is that deontic detachment is not involved in the derivation of the problematic result.

⁴I make the limit assumption, to simplify things (cf. Kratzer (1981)). Kolodny & MacFarlane (2010) and MacFarlane (2014, Chapter 11) make this assumption as well.

⁵More properly, where $\geq_{g(x,w,\Pi)}$ is the partial order derived from the ordering source g , we let $u =_{g(x,w,\Pi)} v$ just in case both $u \geq_{g(x,w,\Pi)} v$ and $v \geq_{g(x,w,\Pi)} u$. Recall that $g(x,w,\Pi)$ is actually an unofficial abbreviation of $g(x,w,\pi(x,f,w,t))$, so that we should officially write that $u =_{g(x,w,\pi(x,f,w,t))} v$ just in case both $u \geq_{g(x,w,\pi(x,f,w,t))} v$ and $v \geq_{g(x,w,\pi(x,f,w,t))} u$. Similarly, the official form of the right-hand side of (13) is $[u]_{\pi(x,t,w,f)} = [v]_{\pi(x,t,w,f)}$.

⁶Here I am making a few relatively safe assumptions. First, that certainty and uncertainty as to the whereabouts of the miners can be captured in terms of probability assignments. Second, that in the choice situation of the Miner Paradox, the actions available to the agent are mutually exclusive, in the sense that performing one action requires the agent not to perform the others. These assumptions are required if choice situations are to be representable at all in decision-theoretic terms.

⁷In Kratzer’s account, indicative conditionals without explicit operators are *epistemic* conditionals, that is, conditionals with an (implicitly) epistemically modalized consequent. A full discussion of why indicative conditionals are sensibly epistemic conditionals would take us too far away. The upshot is that this understanding of the indicative conditional is in line with the insight derived from Ramsey’s test. See Gillies (2004) for a sustained argument in favor of indicative conditionals as epistemic conditionals (though in a dynamic semantic setting).

⁸It should be noted that the relativist’s case against contextualism about DMs does not depend solely on the semantic argument in favor of information-neutral semantic values, but also on a pragmatic argument based on data concerning assertion, rejection, and retraction conditions on utterances involving DMs. The discussion of this pragmatic side of the debate lies outside the scope of the present paper and will be addressed on a separate occasion.