Informational Dynamics of 'Might' Assertions

Peter Hawke[1] and Shane Steinert-Threlkeld[1]

1. Department of Philosophy, Stanford University
phawke@stanford.edu and shanest@stanford.edu

Abstract. We investigate, in a logical setting, the proposal that assertion primarily functions to express and coordinate doxastic states and that 'might' fundamentally expresses *lack* of belief. We provide a formal model of an agent's doxastic state and precise assertability conditions for an associated formal language. We thereby prove that an arbitrary assertion (including a complex of 'might' and 'factual' claims) always succeeds in expressing a well-defined doxastic state. We then propose a fully general and intuitive doxastic update operation as a model of an agent coming to accept an arbitrary assertion. We provide reduction axioms for some novel update operations related to this proposal.[1]

1 Introduction

Consider the following conversation in ordinary language:

(1) Context: Mark hasn't been able to find his house keys in his pocket, his bag, or on his nightstand. While searching, Mark looks out the window at his partner Sue's car, but he sees no reason to think the keys could be there: it is extremely rare that Sue uses his house keys without checking with him.
  a) M: I'm so annoyed. I must have accidentally left my keys on the bus.
  b) S: Actually, they might be in my car.
  c) M: Ah, OK. I'll go look.

Intuitively, Sue has raised for Mark the possibility that the keys are in her car, which he had previously decided not to take 'seriously' even though he was aware of it. He acknowledges that possibility and so goes to check her car.

Surprisingly, offering an explanation of this information flow has proven difficult. Most theorists have tried to provide a semantics for the word 'might' which, when combined with a picture of assertion, will generate the right results.[2] The simplest explanation would identify a particular piece of information that Sue puts forward in (1b) and which Mark subsequently adopts. Note, however, that the information that it's compatible with Sue's information that the keys are in her car does not do the trick. If Mark accepts that information, he acquires a belief about Sue, not about the keys. More sophisticated views posit *many* pieces of information[3] or information whose truth depends on who assesses it.[4]

[1] The final publication is available at Springer Link through the following URL: http://link.springer.com/chapter/10.1007/978-3-662-48561-3_12.
[2] The orthodox semantics belongs to Kratzer [1981, 2012]. See also Papafragou [2006].
[3] See von Fintel and Gillies [2011].
[4] See MacFarlane [2011a, 2014].

Although we will not explicitly argue against these views here, their baroqueness merits initial hesitation.

Against this backdrop, expressivists argue that epistemic modals generally serve to express features of agents' doxastic states and assertion helps to coordinate on those features.[5] In this paper, we develop a formal model of the above information flow that captures these expressivist thoughts. In the next section, we motivate the view that 'might' expresses *lack* of belief (which we call abelief).[6] Following that, we provide a formal model of assertability and doxastic states that allows us to precisely identify the set of beliefs and abeliefs expressed by an arbitrary (and possibly complex) assertion. Then, we identify a *simultaneous update operation* for a given set of beliefs and abeliefs which is fully general, gives the right results in many cases, and reduces to natural updates in the particular case of bare indicative or bare 'might' assertions. After presenting this model, we demonstrate how our view handles epistemic contradictions, disagreement, and interactions with conjunction and disjunction.

2 Proposal: 'Might' as Abelief Coordinator

To make precise our view that assertion functions primarily to coordinate doxastic states, we need to say what doxastic states are. For us, such a state is a set of worlds W with a plausibility order $\succeq$. An agent believes that p if and only if p is true in all of the most plausible worlds. This natural model generalizes the standard modal semantics of belief in a way that allows conditional beliefs and various revision policies to be modeled.[7]

To warm up to our analysis of (1), consider a 'factual' version:

(2) Context: as in (1)
  a) M: I must have accidentally left my keys on the bus.
  b) S: They are in my car.
  c) M: Oh, OK. Thanks!

On our picture, Sue, in (2b), does two things: she *expresses* that she believes that the keys are on the table and *invites* Mark to modify his doxastic state so as to acquire that belief. When Mark accepts the assertion, he does so modify his state.

Let c, b, n, and p be the propositions that the keys are in Sue's car, in Mark's bag, on his nightstand, or in his pocket, respectively. We can model Mark's doxastic state with 5 worlds: $W = \{c, b, n, p, L\}$. In our abused notation, the worlds c, b, etc. are worlds in which only the corresponding proposition is true. L is a world in which the keys are 'lost' (i.e. left on the bus). Initially, Mark's doxastic state looks like this:[8]

[5] See Yalcin [2007, 2011] for expressivism about epistemic modals.
[6] An agent abelieves that p iff she does not believe that p. This is different from *disbeleiving* that p, which means believing that not-p.
[7] See, for instance, van Benthem [2011].
[8] $w \succ v$ means w is strictly more plausible than v. In our notation, b, n, p are all equally plausible.

(3) $L \succ c \succ b, n, p$

Because the unique most plausible world is an L world, Mark believes L, that the keys are lost. Upon accepting Sue's assertion, Mark's doxastic state becomes:

(4) $c \succ L \succ b, n, p$

The c world has been upgraded to be the unique most plausible world, and so Mark believes that the keys are in Sue's car.

What, then, about (1)? Sue's 'might' assertion in (1b) *expresses* that she abelieves that $\neg c$ and *invites* Mark to modify his doxastic state so as to acquire that abelief. After Mark accepts the assertion, his doxastic state looks like:

(5) $c, L \succ b, n, p$

Because there is a c world *among* the most plausible worlds, Mark no longer believes $\neg c$, i.e. he now abelieves $\neg c$.

In general, assertion functions to coordinate doxastic states by expressing a state and inviting one's interlocutors to adopt the same state. The two most fundamental such states are belief and abelief. Let us call assertions which primarily express beliefs B-assertions and those which primarily express abeliefs A-assertions. In general, then, we can say the following about the informational effect of accepting assertions of each type:

- B-assertion triggers *conservative revision*: $\uparrow p\,(\succeq)$ is just like $\succeq$ with the most plausible p-worlds made more plausible than all others
- A-assertion triggers *conservative contraction*: $\upharpoonleft p\,(\succeq)$ is just like $\succeq$ with the most plausible p-worlds merged with the previous most plausible worlds

These notions of update – and our terminology – are not new: conservative revision is closely related to standard notions of *revision* in the AGM belief revision literature; while conservative contraction is closely related to standard notions of belief *contraction*.[9]

We note two points. First, given this picture of assertion, it would be very surprising if we had no means of expressing abelief. Secondly, viewing 'might' as expressing abelief provides a plausible model of its role in the dynamics of conversation. We substantiate this claim more below.

3 Two Problems for Mixed Assertions

While our previous story gave a precise and intuitive account of 'bare' B-assertions and A-assertions, it must be generalized to handle assertions of higher complexity, potentially mixing expressions of belief and abelief. A simple example: $p \wedge \Diamond q$.

[9] Conservative revision corresponds in a precise sense to *transitively relational partial meet revision* and conservative contraction corresponds to *transitively relational partial meet contraction*. See Hansson [2014], especially sect. 4, for an overview of these results. See Rott [2009] for a comprehensive list of belief update procedures, including those that appear in this paper.

Intuitively, an assertion thereof expresses belief in p and abelief in $\neg q$. But now consider an assertion of $p \vee \Diamond q$. What doxastic state is thereby expressed? Or, even worse: $(p \wedge \Diamond q) \vee \Diamond (s \wedge (\Diamond t \wedge \neg q))$? In the current section, we provide two logical frameworks (3.1-3.3) which together give precise answers (3.4-3.5) to the following two questions about an *arbitrary* assertion: (i) can it be understood to express a doxastic state (and, if so, what state is expressed)? (ii) what update operation is performed on acceptance?

3.1 Language

We work with a standard logical language containing: atomic proposition letters $(p, q, r, \dots)$; boolean operators $\neg, \vee, \wedge$; $\Diamond\varphi$ ("φ might be the case"); and $B\varphi$ ("the agent believes that φ").

3.2 Assertability Logic

Let $\mathbf{s}$ be an information set (a set of possible worlds). We will define what it means for a formula to be *assertable*[10] relative to an information set. In what follows, read $\mathbf{s} \Vdash \varphi$ as "φ is assertable relative to information $\mathbf{s}$". For the sake of simplicity, we save the case of $B\varphi$ for an extended version of this paper. Call the fragment of our language that ignores the B operator the *assertability language.*

Definition 1 (General Assertability Conditions). *Given a set of worlds W, an information state $\boldsymbol{s} \subseteq W$, and a valuation V:*

- $\boldsymbol{s} \Vdash p$ *iff:* $\forall w \in \boldsymbol{s}$*:* $w \in V(p)$
- $\boldsymbol{s} \Vdash \neg\varphi$ *iff:* $\forall w \in \boldsymbol{s}$*:* $\{w\} \nVdash \varphi$
- $\boldsymbol{s} \Vdash \varphi \wedge \psi$ *iff:* $\boldsymbol{s} \Vdash \varphi$ *and* $\boldsymbol{s} \Vdash \psi$
- $\boldsymbol{s} \Vdash \varphi \vee \psi$ *iff:* $\exists \boldsymbol{s}_1, \boldsymbol{s}_2$*:* $\boldsymbol{s} = \boldsymbol{s}_1 \cup \boldsymbol{s}_2$ *and* $\boldsymbol{s}_1 \Vdash \varphi$ *and* $\boldsymbol{s}_2 \Vdash \psi$
- $\boldsymbol{s} \Vdash \Diamond\varphi$ *iff:* $\boldsymbol{s} \nVdash \neg\varphi$

We intend these conditions to reflect compelling pre-theoretic intuitions. The final clause is, in particular, worth remarking on: this clause is inspired by the strongly felt illegitimacy of asserting both "it might be that φ" and "φ is not the case" in a single context (an intuition emphasized by Yalcin [2007, 2011]). Certain important consequences of these conditions are immediate:[11]

- $\mathbf{s} \Vdash \Diamond\varphi$ iff $\exists w \in \mathbf{s} : \{w\} \Vdash \varphi$
- Relative to the $\Diamond$-free fragment of our language, this logic is classical
- Relative to singletons, $\Diamond\varphi$ is assertable just in case φ is assertable

[10] It is not our goal to here offer an account that does full justice to our ordinary conception of assertion, nor the many facets of the theoretical role that assertion is intended to play in linguistic theorizing. For a more thoroughgoing discussion of assertion, see MacFarlane [2011b]. Our immediate goal is to offer a simple and natural account of when a sentence is assertable relative to a particular body of information, predominantly thought of as the belief worlds of a relevant agent.

[11] Our framework of assertability conditions is similar in technical spirit to the expressivist semantics of Lin [2013], a connection we do not detail here. At any rate, the formulation, conceptual underpinnings, dialectical role and technical consequences of the current framework diverge from that of Lin [2013] in significant ways.

3.3 Doxastic Logic

We now present a *truth-conditional semantics* for our language, with the intended purpose of making precise the manner of thinking about an agent's doxastic states that we have so far utilized in this paper. Our semantics is in the tradition of *dynamic doxastic logic*,[12] though we only add a dynamic component in the next section. Our goal is to have a precise language for describing the beliefs and abeliefs of an agent. In particular, it suits our purpose to exclude $\Diamond\varphi$ sentences from our semantics.

Definition 2 (Doxastic model). *A* doxastic model *is a tuple* $\mathcal{M} = \langle W, \{\succeq_w\}, V \rangle$ *where:*

- W *is a set of* worlds
- $\succeq_w$*, the* plausibility order *on* W *at* w*, is a* total pre-order *on* W*: a reflexive, transitive, total relation.*
- V *is a* valuation function *assigning a proposition (i.e. a set of worlds) to each atom* p*.*

Moreover, we require the orderings $\succeq_w$ *to be* reverse well-founded*: every non-empty* $X \subseteq W$ *has a maximal element.*[13]

For a given plausibility order $\succeq$, $\succ$ denotes its strict counterpart: $v \succ w$ iff $v \succeq w$ and $w \not\succeq v$. For $X \subseteq W$, we define $\textsc{Best}_{\succeq}(X) := \{w \in X \mid \forall v \in X, w \succeq v\}$. This is the set of maximal, or 'best', worlds among X. We will denote by $\mathbf{b}_w$ the set of 'belief worlds' at w, that is the set of worlds maximal in $\succeq_w$, i.e. $\textsc{Best}_{\succeq_w}(W)$. By the assumption of reverse well-foundedness, these sets are always non-empty.

Definition 3 (Static Semantics).

- $\mathcal{M}, w \vDash p$ *iff:* $w \in V(p)$
- $\mathcal{M}, w \vDash \neg\varphi$ *iff:* $\mathcal{M}, w \nvDash \varphi$
- $\mathcal{M}, w \vDash \varphi \wedge \psi$ *iff:* $\mathcal{M}, w \vDash \varphi$ *and* $\mathcal{M}, w \vDash \psi$
- $\mathcal{M}, w \vDash B\varphi$ *iff: for every* $v \in \boldsymbol{b}_w$*,* $\mathcal{M}, v \vDash \varphi$

Two more definitions will be useful in what follows. We will write $[\![\varphi]\!]^{\mathcal{M}} := \{w \in W \mid \mathcal{M}, w \vDash \varphi\}$ and omit the superscript when context allows. For brevity, we will also write $\textsc{Best}_w(\varphi) := \textsc{Best}_{\succeq_w}\left([\![\varphi]\!]^{\mathcal{M}}\right)$.

3.4 From Assertion to Doxastic State Expression

In this section we work towards a theorem that addresses our first problem: what doxastic state is expressed by an arbitrary assertion? The theorem will state that for every assertable sentence there exists a well-defined doxastic state expressed by that sentence. What's more, the proof for this result supplies the ingredients for a method for constructing such a doxastic state, though we will not state such an algorithm explicitly.

[12] See van Benthem [2011] for an overview of this tradition.

[13] In terms of frame correspondence, we can impose this requirement via the Löb axiom $\Box(\Box p \to p) \to p$.

Definition 4 (Assertoric Equivalence). *We say that two sentences φ and ψ in the assertability language are* assertorically equivalent *just in case*

$$\boldsymbol{s} \Vdash \varphi \textit{ iff: } \boldsymbol{s} \Vdash \psi$$

for every information state $\boldsymbol{s}$ *(and every doxastic model $\mathcal{M}$).*

Definition 5 ($\Diamond$-Free Formulae). *A sentence in our assertability language is* $\Diamond$-free *just in case it contains no occurrence of $\Diamond$.*

Lemma 1. *If φ is $\Diamond$-free, then $\boldsymbol{s} \Vdash \varphi$ iff $\boldsymbol{s} \subseteq \overline{V}(\varphi)$, where $\overline{V}$ is the unique extension of V interpreting $\neg$ as complement, $\wedge$ as intersection, and $\vee$ as union.*

Proof. By induction on $\Diamond$-free formulas (exercise).

Proposition 1 (Assertability Facts).

(1) $\boldsymbol{s} \Vdash \Diamond(\psi \wedge \Diamond\varphi_1 \wedge \ldots \wedge \Diamond\varphi_n)$ *iff:* $\boldsymbol{s} \Vdash \Diamond(\psi \wedge \varphi_1 \wedge \ldots \wedge \varphi_n)$.
In particular: $\boldsymbol{s} \Vdash \Diamond\Diamond\varphi$ *iff:* $\boldsymbol{s} \Vdash \Diamond\varphi$.
(2) $\boldsymbol{s} \Vdash \neg(\psi \wedge \Diamond\varphi_1 \wedge \ldots \wedge \Diamond\varphi_n)$ *iff:* $\boldsymbol{s} \Vdash \neg(\psi \wedge \varphi_1 \wedge \ldots \wedge \varphi_n)$.
In particular: $\boldsymbol{s} \Vdash \neg\Diamond\varphi$ *iff:* $\boldsymbol{s} \Vdash \neg\varphi$.
(3) If ψ^1 and ψ^2 are both $\Diamond$-free, then:

$\boldsymbol{s} \Vdash (\psi^1 \wedge \Diamond\varphi_1^1 \wedge \ldots \wedge \Diamond\varphi_m^1) \vee (\psi^2 \wedge \Diamond\varphi_1^2 \wedge \ldots \wedge \Diamond\varphi_n^2)$ *iff:*
$\boldsymbol{s} \Vdash (\psi^1 \vee \psi^2) \wedge \Diamond(\psi^1 \wedge \varphi_1^1) \wedge \ldots \wedge \Diamond(\psi^1 \wedge \varphi_m^1) \wedge \Diamond(\psi^2 \wedge \varphi_1^2) \wedge \ldots \wedge \Diamond(\psi^2 \wedge \varphi_n^2)$

Proof.
(1) $\mathbf{s} \Vdash \Diamond\varphi$ is equivalent to $\exists w \in \mathbf{s} : \{w\} \Vdash \varphi$. Further, $\{w\} \Vdash \Diamond\psi$ holds just in case $\{w\} \Vdash \psi$ holds. Hence: $\exists w \in \mathbf{s} : \{w\} \Vdash \psi \wedge \Diamond\varphi_1 \wedge \ldots \wedge \Diamond\varphi_n$ is equivalent to $\exists w \in \mathbf{s} : \{w\} \Vdash \psi \wedge \varphi_1 \wedge \ldots \wedge \varphi_n$.
(2) $\mathbf{s} \Vdash \neg\varphi$ is equivalent to $\forall w \in \mathbf{s} : \{w\} \nVdash \varphi$. Further, $\{w\} \Vdash \Diamond\psi$ holds just in case $\{w\} \Vdash \psi$ holds. Hence: $\forall w \in \mathbf{s} : \{w\} \nVdash \psi \wedge \Diamond\varphi_1 \wedge \ldots \wedge \Diamond\varphi_n$ is equivalent to $\forall w \in \mathbf{s} : \{w\} \nVdash \psi \wedge \varphi_1 \wedge \ldots \wedge \varphi_n$.
(3) We illustrate the proof with a particular instance. The general case uses Lemma 1. We show that

$$\mathbf{s} \Vdash (p \wedge \Diamond q) \vee (r \wedge \Diamond s) \text{ iff: } \mathbf{s} \Vdash (p \vee r) \wedge \Diamond(p \wedge q) \wedge \Diamond(r \wedge s)$$

$\mathbf{s} \Vdash (p \wedge \Diamond q) \vee (r \wedge \Diamond s)$
iff: $\exists \mathbf{s}_1, \mathbf{s}_2 : \mathbf{s}_1 \cup \mathbf{s}_2 = \mathbf{s}$, and $\mathbf{s}_1 \Vdash p$, and $\mathbf{s}_1 \Vdash \Diamond q$, and $\mathbf{s}_2 \Vdash r$, and $\mathbf{s}_2 \Vdash \Diamond s$
iff: $\mathbf{s} \Vdash p \vee r$ and $\exists v_1 \in \mathbf{s} : \{v_1\} \Vdash p \wedge q$ and $\exists v_2 \in \mathbf{s} : \{v_2\} \Vdash r \wedge s$
iff: $\mathbf{s} \Vdash (p \vee r) \wedge \Diamond(p \wedge q) \wedge \Diamond(r \wedge s)$ □

Lemma 2. *Let φ be a sentence in the assertability language. Then there exist sentences β, α_1, ..., α_n (for some $n \geq 0$) such that:*

- *β, α_1, ..., α_n contain no occurrences of $\Diamond$,*
- *$\boldsymbol{s} \Vdash \varphi$ iff $\boldsymbol{s} \Vdash \beta \wedge \Diamond\alpha_1 \wedge \ldots \wedge \Diamond\alpha_n$*

Proof. By induction on the complexity of formulae. The non-trivial cases (taking the assumption that φ is assertorically equivalent to $\beta \wedge \Diamond\alpha_1 \wedge \ldots \wedge \Diamond\alpha_n$ as the induction hypothesis):

– $\neg\varphi$: using fact 2 of proposition 1, we conclude that $\neg\varphi$ is assertorically equivalent to $\neg(\beta \wedge \alpha_1 \wedge \ldots \wedge \alpha_n)$.
– $\varphi_1 \vee \varphi_2$: assume that φ_1 is assertorically equivalent to $\beta^1 \wedge \Diamond\alpha_1^1 \wedge \ldots \wedge \Diamond\alpha_m^1$, and that φ_2 is assertorically equivalent to $\beta^2 \wedge \Diamond\alpha_1^2 \wedge \ldots \wedge \Diamond\alpha_n^2$. Now use fact 3 of proposition 1.
– $\Diamond\varphi$: fact 1 of proposition 1 shows that this is assertorically equivalent to $\Diamond(\beta \wedge \alpha_1 \wedge \ldots \wedge \alpha_n)$ □

Lemma 3. *Let φ be a $\Diamond$-free sentence in the assertability language and $\mathcal{M}$ and w be an arbitrary doxastic model and world. Then:*

$$\boldsymbol{b}_w \Vdash \varphi \text{ iff: } \mathcal{M}, w \vDash B\varphi$$

Proof. By induction on the complexity of formulae (exercise). □

Lemma 4. *Let $\mathcal{M}$ and w be an arbitrary doxastic model and world. Then, for any φ in the assertability language:*

$$\boldsymbol{b}_w \Vdash \Diamond\varphi \text{ iff: } \mathcal{M}, w \vDash \neg B(\neg\varphi)$$

Proof. By induction on the complexity of formulae (exercise). □

Definition 6 (Doxastic State Description). *A* doxastic state description *is a sentence of the form*

$$B\varphi \wedge \neg B(\neg\psi_1) \wedge \ldots \wedge \neg B(\neg\psi_n)$$

where φ and ψ_i, for $i \leq n$, are all $\Diamond$-free.

Definition 7 (Doxastic State Expression). φ *(in the assertability language)* expresses *doxastic state description δ just in case: for every doxastic model $\mathcal{M}$ and world w,*

$$\boldsymbol{b}_w \Vdash \varphi \text{ iff: } \mathcal{M}, w \vDash \delta$$

Theorem 1 (From Assertion to Doxastic State Expression). *For every sentence φ in the assertability language, there exists a doxastic state description δ_φ that is expressed by φ.*

Proof. $\mathbf{b}_w \Vdash \varphi$
iff: $\mathbf{b}_w \Vdash \beta \wedge \Diamond\alpha_1 \wedge \ldots \wedge \Diamond\alpha_n$ [Lemma 2]
iff: $\mathbf{b}_w \Vdash \beta$ and $\mathbf{b}_w \Vdash \Diamond\alpha_1$ and ... and $\mathbf{b}_w \Vdash \Diamond\alpha_n$
iff: $\mathcal{M}, w \vDash B\beta$ and $\mathcal{M}, w \vDash \neg B(\neg\alpha_1)$ and ... and $\mathcal{M}, w \vDash \neg B(\neg\alpha_n)$ [Lemmas 3 and 4]
iff: $\mathcal{M}, w \vDash B(\beta) \wedge \neg B(\neg\alpha_1) \wedge \ldots \wedge \neg B(\neg\alpha_n)$ □

3.5 Dynamics

Recall our second problem: for an arbitrary assertion, what update does an agent's doxastic state undergo upon acceptance of that assertion? For example, consider an assertion $p \wedge \Diamond q$. This expresses the doxastic state $Bp \wedge \neg B \neg q$ in the sense of Definition 7. How can an agent update her doxastic state so that this is an adequate description thereof? First, note that *sequentially* applying conservative revision and contraction will not work. Consider $W = \{w_1, w_2, w_3\}$ where w_1 satisfies p but not q, w_2 satisfies q but not p, and w_3 satisfies both p and q. Suppose the agent's doxastic state has the form

(6) $w_1 \succ w_2 \succ w_3$

Then $\uparrow p \, (\upharpoonleft \neg q \, (\succeq))$ will be the same order (6). But at a world with that order, $B \neg q$ is true. On the other hand, $\upharpoonleft \neg q \, (\uparrow p \, (\succeq))$ is

(7) $w_1, w_2 \succ w_3$

But at a world with this order, Bp is not true. Thus, $Bp \wedge \neg B \neg q$ is not an accurate description of either doxastic state. This example shows that the update to perform upon accepting a mixed assertion cannot simply be an iteration of our earlier updates. To address this issue, we enrich the language with expressions of the form $[\upuparrows \varphi] \, \psi$ with intended reading: "after conservative expansion by φ, ψ holds" and use this to define a simultaneous update operation.

Definition 8 (Conservative Expansion). *Given an order $\succeq$ and $X \subseteq W$, we denote by $\upuparrows X \, (\succeq)$ the* conservative expansion *of $\succeq$ by X, where: $\upuparrows X \, (\succeq)$ is the order that is just like $\succeq$ except with all of X made most plausible and all worlds in X made equally plausible to each other.*

We extend this to a model-changing operation as follows: $\mathcal{M} \upuparrows X$ is just like $\mathcal{M}$, except with each $\succeq_w$ replaced with $\upuparrows X \, (\succeq_w)$. We focus on the case where $X = [\![\varphi]\!]^{\mathcal{M}}$ for some formula in our language, in which case we will write $\mathcal{M} \upuparrows \varphi$, calling this the conservative expansion of $\mathcal{M}$ by φ.

Definition 9 (Dynamic Semantics). *The static semantics can be extended:*

– $\mathcal{M}, w \vDash [\upuparrows \varphi] \, \psi$ *iff:* $\mathcal{M} \upuparrows \varphi, w \vDash \psi$

Using this framework, we have the resources to define *conservative revision* and *conservative contraction* operations, respectively as follows:

i. $\uparrow \varphi \, (\succeq_w) \; := \; \upuparrows \textsc{Best}_w(\varphi) \, (\succeq_w)$
ii. $\upharpoonleft \varphi \, (\succeq_w) \; := \; \upuparrows (\textsc{Best}_w(\varphi) \cup \textsc{Best}_w(\top)) \, (\succeq_w)$

Now, we can define an operation that tells us how to update on an arbitrary doxastic state description. Intuitively, it is the operation of *simultaneously performing* the conservative revisions and conversative contractions suggested by the set of beliefs and abeliefs expressed by that description.

Definition 10 (Simultaneous Update). *The simultaneous update to believe* β *and abelieve* $\alpha_1, \ldots, \alpha_n$ *is the following operation:*

$$[\Uparrow \beta, \alpha_1, \ldots, \alpha_n](\succeq_w) := \left[\Uparrow \textsc{Best}_w(\beta) \cup \bigcup_{1 \leq i \leq n} \textsc{Best}_w(\neg\alpha_i \wedge \beta)\right](\succeq_w)$$

For all of conservative revision, conservative contraction, and simultaneous update, we can define the appropriate model-changing operations and extend the syntax with dynamic operators in exactly the same way as was done for conservative expansion above. Our definiton of simultaneous update has many attractive consequences. First, note that this update handles our earlier counter-example. $[\Uparrow p, \neg q]$ applied to the order in (6) yields

(8) $w_1, w_3 \succ w_2$

In a world with this order, however, $Bp \wedge \neg B \neg q$ is true. Moreover, this definition handles our motivating cases (1) and (2) with aplomb. If φ^* is a doxastic state description, we will abbreviate the above by $[\Uparrow \varphi^*]$. In the case when φ^* has no conjunct $B(\beta)$, replace β with $\top$. In the case when φ^* has no conjunct $\neg B\psi_i$, set $n = 1$ and $\alpha_1 = \bot$.

Proposition 2. *Let* φ *be a sentence in the assertability language. Then, for every model and order:*

i. If φ *expresses no abeliefs, then* $[\Uparrow \varphi^*](\succeq) = [\uparrow \beta](\succeq)$
ii. If φ *expresses a single abelief, then* $[\Uparrow \varphi^*](\succeq) = [\upharpoonleft \alpha_1](\succeq)$

When working with dynamic operators like this, a natural question to ask is: is every sentence in the language with dynamic operators equivalent to some sentence in the static 'base' language? One usually provides a 'yes' answer to this question by giving *reduction axioms* which show how to push the dynamic operators to simpler subformulas. We can provide such axioms for many of our operators. Conservative revision is already well understood,[14] so we focus on conservative expansion and conservative contraction. We start with conservative expansion and the doxastic language. We must augment the language with an existential modality E and its dual universal modality U.

Proposition 3. *The following reduction axioms are valid for the class of doxastic models:*

$$\begin{array}{lcl}
[\Uparrow \varphi]\, p & \leftrightarrow & p \\
[\Uparrow \varphi]\, \neg\psi & \leftrightarrow & \neg [\Uparrow \varphi]\, \psi \\
[\Uparrow \varphi]\, \psi \wedge \chi & \leftrightarrow & [\Uparrow \varphi]\, \psi \wedge [\Uparrow \varphi]\, \chi \\
[\Uparrow \varphi]\, B\psi & \leftrightarrow & (E\varphi \wedge U(\varphi \rightarrow [\Uparrow \varphi]\, \psi)) \vee (\neg E\varphi \wedge B\, [\Uparrow \varphi]\, \psi) \\
[\Uparrow \varphi]\, E\psi & \leftrightarrow & E\, [\Uparrow \varphi]\, \psi
\end{array}$$

[14] See chapter 8 of van Benthem [2011], where it goes by 'conservative upgrade'.

For the case of conservative contraction, we extend the language with a *conditional belief* operator $B^{\varphi}\psi$ with the following semantics:

Definition 11 (Conditional Belief). *For a doxastic model $\mathcal{M}$ and world w:*

$$\mathcal{M}, w \vDash B^{\varphi}\psi \text{ iff: for every } v \in \text{Best}_w(\varphi), \mathcal{M}, v \vDash \psi$$

Note that $B\varphi$ is the special case $B^{\top}\varphi$.

Proposition 4. *The following reduction axioms are valid for the class of doxastic models: those above for atoms, $\neg$, $\wedge$, and E but with $\upharpoonleft\varphi$ and*

$$[\upharpoonleft\varphi] B\psi \quad \leftrightarrow \quad B[\upharpoonleft\varphi]\psi \wedge B^{\varphi}[\upharpoonleft\varphi]\psi$$

This axiom makes good intuitive sense. After the update, the agent believes ψ iff the new best worlds are all ψ. The new best worlds are: the old best worlds merged with the old best φ worlds. The first conjunct handles the former and the second conjunct the latter. Of course, if we have conditional belief in our language, one would like a reduction axiom for that operator.

Theorem 2. *The following reduction axioms are valid for doxastic models:*

$$\begin{aligned}
[\upuparrows\varphi] B^{\chi}\psi \quad &\leftrightarrow \quad \left(\neg E(\varphi \wedge [\upuparrows\varphi]\chi) \wedge B^{[\upuparrows\varphi]\chi}[\upuparrows\varphi]\psi\right) \vee \\
&\qquad (E(\varphi \wedge [\upuparrows\varphi]\chi) \wedge U(\varphi \wedge [\upuparrows\varphi]\chi \to [\upuparrows\varphi]\psi)) \\
[\upharpoonleft\varphi] B^{\chi}\psi \quad &\leftrightarrow \quad \left(B^{\varphi}\neg[\upharpoonleft\varphi]\chi \wedge B^{[\upharpoonleft\varphi]\chi}[\upharpoonleft\varphi]\psi\right) \vee \\
&\qquad \left(\neg B^{\varphi}\neg[\upharpoonleft\varphi]\chi \wedge B^{\varphi\wedge[\upharpoonleft\varphi]\chi}[\upharpoonleft\varphi]\psi \wedge \left(\neg B\neg[\upharpoonleft\varphi]\chi \to B^{[\upharpoonleft\varphi]\chi}[\upharpoonleft\varphi]\psi\right)\right)
\end{aligned}$$

Proof. First, consider conservative expansion. $[\upuparrows\varphi] B^{\chi}\psi$ says: after making all of the φ worlds most plausible (and equally plausible), the best χ worlds are ψ worlds. We make a case distinction: (i) no φ worlds become χ-worlds or (ii) some φ worlds become χ-worlds. If (i), then the best χ-worlds after the update are the best worlds pre-update that *become* χ. We then need to check that those worlds become ψ worlds. That is what the first disjunct in the reduction axiom states. If (ii), the best χ-worlds post-update are exactly the current φ worlds that *become* χ worlds since all of the current φ worlds become best overall. We thus need to check that every φ world which becomes χ also becomes a ψ world. That's what the second disjunct in the recursiom axiom states.

Now, consider conservative contraction. $[\upharpoonleft\varphi] B^{\chi}\psi$ says: after merging the best φ worlds with the best-overall worlds, the best χ worlds are ψ worlds. We again make a case distinction: (i) no best φ worlds become χ worlds or (ii) some best φ worlds become χ worlds. In case (i), the best χ worlds post-update are simply the best worlds that *become* χ. We need those to become ψ, which is just what the first disjunct states. In case (ii), the best χ worlds post-update come from two sources: (a) previous best φ worlds that become χ and (b) previous best-overall worlds that become χ. The conjunct

$$\neg B^{\varphi}\neg[\upharpoonleft\varphi]\chi \wedge B^{\varphi\wedge[\upharpoonleft\varphi]\chi}[\upharpoonleft\varphi]\psi$$

handles case (a) by requiring that the best φ worlds that become χ also become ψ. The conjunct

$$\neg B \neg [\Uparrow \varphi] \chi \to B^{[\Uparrow \varphi]\chi} [\Uparrow \varphi] \psi$$

handles case (b). □

We can derive reduction axioms for full belief in Propositions 3 and 4 as special cases of the above.

Corollary 1. *The following reduction axioms are valid for doxastic models:*

$$[\Uparrow\varphi] B\psi \quad \leftrightarrow \quad (E\varphi \wedge U(\varphi \to [\Uparrow\varphi]\psi)) \vee (\neg E\varphi \wedge B[\Uparrow\varphi]\psi)$$
$$[\Uparrow \varphi] B\psi \quad \leftrightarrow \quad B[\Uparrow \varphi]\psi \wedge B^{\varphi}[\Uparrow \varphi]\psi$$

Proof. We do the conservative contraction case and leave conservative expansion as an exercise. Substituting $\top$ for χ in the above reduction axiom yields:

$$[\Uparrow \varphi] B\psi \quad \leftrightarrow \quad \left(B^{\varphi}\neg [\Uparrow \varphi] \top \wedge B^{[\Uparrow \varphi]\top} [\Uparrow \varphi] \psi\right) \vee \left(\neg B^{\varphi}\neg [\Uparrow \varphi] \top \wedge B^{\varphi \wedge [\Uparrow \varphi]\top} [\Uparrow \varphi] \psi \wedge \left(\neg B \neg [\Uparrow \varphi] \top \to B^{[\Uparrow \varphi]\top} [\Uparrow \varphi] \psi\right)\right)$$

Notice that the first disjunction is a contradiction: $B^{\varphi}\neg [\Uparrow \varphi] \top$ is always false since every world satisfies $\top$. The conjunct $\neg B^{\varphi}\neg [\Uparrow \varphi] \top$ is always true since it merely states the existence of a best φ world, which the assumption of reverse well-foundedness ensures. The conjunct $B^{\varphi \wedge [\Uparrow \varphi]\top} [\Uparrow \varphi] \psi$ simplifies to $B^{\varphi} [\Uparrow \varphi] \psi$. Now, the antecedent of the conditional is trivially true since it merely asserts that there are best worlds, which again holds by reverse well-foundedness. The consequent simplifies to $B [\Uparrow \varphi] \psi$. Thus, we are left with the desired equivalence

$$[\Uparrow \varphi] B\psi \quad \leftrightarrow \quad B [\Uparrow \varphi] \psi \wedge B^{\varphi} [\Uparrow \varphi] \psi$$

as desired. □

4 Some Welcome Consequences

Epistemic contradictions. As emphasized by Yalcin [2007, 2011], statements of the following form (so-called 'epistemic contradictions') seem defective: "John is in his office. But it might be that John is not in his office". Fortunately, then, our assertability conditions immediately yield (for any information state $\mathbf{s}$):

$$\mathbf{s} \not\Vdash p \wedge \Diamond \neg p$$

This does not yet entirely deal with the observation that sentences that *embed* epistemic contradictions are notably defective, for a key case for Yalcin is sentences of the form "supposing that $p \wedge \Diamond\neg p$, then ...", and our current setup does not have resources capturing suppositional actions/operators. In the extended version of this paper, we add such operators to our language, and treat their assertability conditions in a fashion inspired by Yalcin's domain semantics, yielding pleasing results.

Disagreement. Consider the following variation of (1), where Mark does not accept Sue's assertion.

(9) Context: as before, except that Mark actually went out and checked Sue's car, where he did not find the keys.
 a) M: I'm so annoyed. I must have accidentally left my keys on the bus.
 b) S: They might be in my car.
 c) M: No, I already checked your car.

In such a case two things must be explained: (i) How is it possible to disagree with an assertion of $\Diamond c$? (ii) Why, when one does so disagree, are the reasons provided about the prejacent c itself? Our story provides natural answers to both questions. To (i): Mark's disagreement consists in *rejecting* Sue's invitation to update his doxastic state to incorporate a c-world. To (ii): Mark rejects this invitation because he thinks he has good reason to have already ruled out the c worlds (for example, having already checked the car). Thus, when explaining his disagreement, he will argue about the prejacent c itself.

Interactions with conjunction and disjunction. It has been observed that conjunction and disjunction display unusual behavior when connecting 'might' claims (Zimmermann [2000], Ciardelli et al. [2009]). Namely, 'or' and 'and' seem *equivalent* in this linguistic context: to say "John might be in his office or he might be at home" seems equivalent to saying "John might be in his office and he might be at home". It may be seen as a virtue then that our assertability conditions yield (for any information state $\mathbf{s}$):

$$\mathbf{s} \Vdash \Diamond p \wedge \Diamond q \text{ iff } \mathbf{s} \Vdash \Diamond p \vee \Diamond q$$

This is a consequence of proposition 1, fact 3. To see this, set $\psi^1 := \top$, $\psi^2 := \top$, $m = 1$ and $n = 1$ in the statement of the fact.

5 Conclusion and Further Work

We have developed the idea that 'might' fundamentally functions to express abelief, providing a formal model which explains the doxastic state expressed by and the update operation performed upon accepting an arbitrary assertion. Our theory can also handle various problematic phenomena involving 'might' that have proven tricky to accommodate in the context of other approaches.

Further work remains to be done. First, more needs to be said to relate our current results to the elaborate debate on the semantics and pragmatics of 'might' and other epistemic modals in the philosophy and linguistics literature. Second, there are various intriguing avenues for further technical results. It would be of interest to identify a complete axiomatization for our assertability logic. We also note that the motivation for our study of conservative contraction – that it captures the idea of "coming to take a possibility seriously" – resembles that for the *suggestion operation* introduced in van Benthem and Liu [2007], Liu [2011]. However, these operations have very different technical consequences. For example, conservative contraction preserves totality, while suggestion does not. It is of interest therefore to thoroughly contrast conservative contraction and suggestion as alternative proposals for 'might' updates.

Bibliography

Ivano Ciardelli, Jeroen Groenendijk, and Floris Roelofsen. Attention! Might in Inquisitive Semantics. In *Proceedings of Semantics and Linguistic Theory (SALT) 19*, number 1, pages 91–108, 2009.

Sven Ove Hansson. Logic of Belief Revision. *Stanford Encyclopedia of Philosophy*, (Winter), 2014. URL `http://plato.stanford.edu/archives/win2014/entries/logic-belief-revision/`.

Angelika Kratzer. The Notional Category of Modality. In Hans-Jürgen Eikmeyer and Hannes Rieser, editors, *Words, Worlds, and Context*, pages 38–74. Walter de Gruyter, 1981.

Angelika Kratzer. *Modals and Conditionals*. Oxford University Press, 2012.

Hanti Lin. Acceptance-Conditional Semantics & Modality-Disjunction Interaction. In *USC Deontic Modality Workshop*, 2013.

Fenrong Liu. *Reasining About Preference Dynamics*. Springer, Dordrecht, 2011.

John MacFarlane. Epistemic Modals are Assessment-Sensitive. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality*, pages 144–179. Oxford University Press, 2011a.

John MacFarlane. What Is Assertion? In Jessica Brown and Herman Cappelen, editors, *Assertion: New Philosophical Essays*, pages 79–97. Oxford University Press, 2011b.

John MacFarlane. *Assessment Sensitivity*. Oxford University Press, 2014. doi: 10.1093/acprof:oso/9780199682751.001.0001.

Anna Papafragou. Epistemic modality and truth conditions. *Lingua*, 116(10): 1688–1702, 2006. doi: 10.1016/j.lingua.2005.05.009.

Hans Rott. Shifting Priorities: Simple Representations for Twenty-seven Iterated Theory Change Operators. In D Makinson, J Malinkowski, and H Wansing, editors, *Towards Mathematical Philosophy*, volume 28 of *Trends in Logic*, pages 269–296. Springer, Dordrecht, 2009.

Johan van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, Cambridge, 2011.

Johan van Benthem and Fenrong Liu. Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, 17(2):157–182, 2007. doi: 10.3166/jancl.17.157-182.

Kai von Fintel and Anthony S Gillies. 'Might' Made Right. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality*, pages 108–130. Oxford University Press, Oxford, 2011.

Seth Yalcin. Epistemic Modals. *Mind*, 116(464):983–1026, 2007. doi: 10.1093/mind/fzm983.

Seth Yalcin. Nonfactualism About Epistemic Modality. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality*, chapter 10, pages 295–332. Oxford University Press, Oxford, 2011.

Thomas Ede Zimmermann. Free Choice Disjunction and Epistemic Possibility. *Natural Language Semantics*, 8:255–290, 2000.