Dynamic Logics of Evidence-Based Belief

Johan van Benthem Eric Pacuit

March 3, 2011

1 Introduction and Motivation

It is generally accepted that a *rational* belief must be grounded in the evidence available to an agent. However, there is much less agreement about the precise relationship between an agent's beliefs and her evidence. Understanding this complex mechanism raises many interesting philosophical and technical issues. Modeling evidence requires richer structures than found in standard epistemic semantics where the accessible worlds aggregate all reliable evidence gathered so far. Even recent more finely-grained plausibility models ordering the epistemic ranges identify too much: belief is indistinguishable from aggregated *best* evidence. At the opposite extreme, one might model evidence syntactically as "formulas received", but this seems overly detailed, and we we lose the intuition that evidence can be semantic in nature, zooming in on some actual world.

Our proposal in this paper is to explore an intermediate semantic level, viz. that of *neighborhood semantics*, where evidence is recorded as a family of sets of worlds. Neighborhood models have been proposed as a technical tool for studying weak modal logics in the past. But here, we show how they support a notion of evidence with matching languages for attitudes based on it, as well as an array of natural actions that transform evidence.[1]

Our paper is mainly a pilot study. We develop the basics of neighborhood models for evidence and belief, and the logics that they support. Then we move to the main theme of the paper, showing how these models support natural actions of "evidence management", ranging from external new information to internal rearrangement. Our main results are relative completeness theorems for these actions over the static base logic of evidence and belief, identifying the major 'dynamic equations' that govern the flow of information. Finally, we show via some representation theorems how the neighborhood framework extends the currently popular approach to dynamic logics of belief revision via plausibility pre-orders. All this is just the start of a much larger research project, and we indicate some further directions at the end.

2 Evidence in Neighbourhood Models

The semantics of evidence that we will use in this paper is based on *neighborhood* models (cf. Chellas, 1980, Chapter 7). Segerberg (1971) is an early source, while Hansen (2003), Pacuit (2007)

[1]Precursors to our ideas are found in the treatment of abstract belief revision policies in Segerberg (1995) and further explored in Leitgeb and Segerberg (2007), the PhD thesis of Girard (2008), and in the representation of scientific theories as discussed in van Benthem (2005). Also inspirational was the work by Kaile Su and colleagues (Su et al., 2005) on sensor-based models of knowledge in AI studies of agency. Finally, we mention the congenial research program of 'topologic' (Moss and Parikh, 1992).

and Hansen et al. (2009) give modern motivations and mathematical details. Our main interest is not the standard logic of these models per se, however, but the account of evidence dynamics which they support.[2] Even so, our approach may also shed some new light on neighborhood logics.

2.1 The basic models

Let W be a set of possible worlds or states one of which represents the "actual" situation. An agent gathers evidence about this situation from a variety of sources.[3] To simplify things, we assume these sources provide *binary* evidence, i.e., subsets of W which (may) contain the actual world. We make the following basic assumptions:

1. Sources may or may not be *reliable*: a subset recording a piece of evidence need not contain the actual world. Also, agents need not know which evidence is reliable.

2. The evidence gathered from different sources (or even the same source) may be jointly inconsistent. And so, the intersection of all the gathered evidence may be empty.

3. Despite the fact that sources may not be reliable or jointly inconsistent, they are all the agent has for forming beliefs. [4]

The *evidential state* of the agent is the set of all propositions identified by the agent's sources. In general, this could be any collection of subsets of W; but we impose some constraints:

- No evidence set is empty (evidence per se is never contradictory),

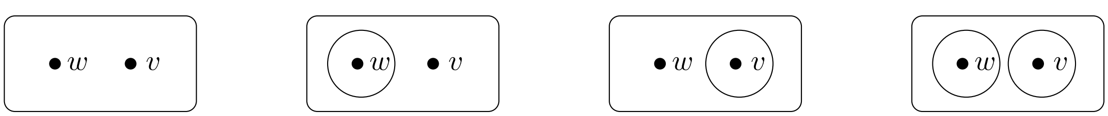
- The whole universe W is an evidence set (agents know their 'space').

In addition, much of the literature would suggest a 'monotonicity' assumption:

If the agent has evidence X and $X \subseteq Y$ then the agent has evidence Y.

To us, however, this is a property of propositions supported by evidence, not of the evidence itself. Therefore, we will model this feature differently, through the semantics of our logic.

Example 2.1 To illustrate these ideas, consider a situation with two worlds $W = \{w, v\}$, where p is true at w and not at v. The following might be evidential states:



There is no evidence for or against p. There is evidence that supports p. There is evidence that rejects p. There is evidence that supports p and also evidence that rejects p.

[2] Other concrete uses of neighborhood models are logics for knowledge: (cf. Vardi, 1986), players' powers in games (cf. Parikh, 1985; Pauly, 2002), and beliefs in the dynamic-epistemic foundations of game theory (cf. Zvesper, 2010).

[3] We basically treat the single-agent case, the many agent case is discussed briefly in Section 5.

[4] Modeling sources and agents' *trust* in these is possible – but we do not develop this theme here.

Definition 2.2 (Evidence Model) An **evidence model** is a tuple $\mathcal{M} = \langle W, E, V \rangle$, where W is a non-empty set of worlds, $E \subseteq W \times \wp(W)$ is an evidence relation, $V : \mathsf{At} \to \wp(W)$ is a valuation function. A **pointed evidence model** is a pair $\mathcal{M}, w$ where w is the "actual world". When E is a constant function, we write $\langle W, \mathcal{E}, V \rangle, w$ where $\mathcal{E}$ is the fixed family of subsets of W related to each state by E. In such a case, we call $\mathcal{M}$ a **uniform evidence model**. $\triangleleft$

We write $E(w)$ for the set $\{X \mid wEX\}$. If $X \in E(w)$, we say "the agent accepts X as evidence at state w". The above two constraints on the evidence function then become:

(Cons) For each state w, $\emptyset \notin E(w)$.
(Triv) For each state w, $W \in E(w)$.

In what follows, we shall mainly work with uniform evidence models. While this may seem very restrictive, the reader will soon see how much relevant structure can be found even at this level.

As stated before, we do not assume that $E(w)$ is closed under supersets. Also, even though evidence pieces are non-empty, their combination through the obvious operations of taking *intersections* need not yield consistent evidence: we allow for disjoint evidence sets, whose combination may lead (and should lead) to trouble. But importantly, even though an agent may not be able to consistently combine *all* of her evidence, there will be maximal collections of admissible evidence that she can safely put together:

Definition 2.3 (Maximal overlapping evidence) A family $\mathcal{X}$ of subsets of W has the **finite intersection property** (f.i.p.) if, for each finite subfamily $\{X_1, \ldots, X_n\} \subseteq \mathcal{X}$, $\bigcap_{1 \leq i \leq n} X_i \neq \emptyset$. The set $\mathcal{X}$ has the **maximal f.i.p.** if $\mathcal{X}$ has the f.i.p. but no proper extension $\mathcal{X}'$ of $\mathcal{X}$ has f.i.p. $\triangleleft$

We will now develop the logic of this framework. Clearly, the above families of sets give us more detail than the usual information states with sets of (accessible) worlds. However, our semantic abstraction level does not provide us with details of code and syntax – and hence it retains a whiff of epistemic closure and logical omniscience. If finer operational aspects of inference and introspection are thought important to one's notion of evidence, then the methods of this paper should be extended to include the dynamic logics of awareness and inference (Fagin and Halpern, 1988; Ågotnes and Alechina, 2007; van Benthem, 2008; Velazquez-Quesada, 2009).

Here is one final restriction: all models in this paper are assumed to be *finite*. Lifting this restriction is possible, but the lure of the infinite must be left for another occasion.

2.2 Reasoning about an agent's beliefs and evidence

In this section, we introduce our basic logic for reasoning about evidence and beliefs.

Language While many formalisms come to mind, we will merely explore a simple modal system.

Definition 2.4 (Evidence and Belief Language) Let At be a fixed set of atomic propositions. Let $\mathcal{L}_0$ be the smallest set of formulas generated by the following grammar

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid B\varphi \mid \Box\varphi \mid A\varphi$$

where $p \in \mathsf{At}$. Additional propositional connectives $(\wedge, \to, \leftrightarrow)$ are defined as usual and the existential modality $E\varphi$ is defined as $\neg A \neg\varphi$. $\triangleleft$

The intended interpretation of $\Box\varphi$ is "the agent has evidence that φ is true" (i.e., "the agent has evidence for φ") and $B\varphi$ says that "the agents believes that φ is true". We also include the universal modality ($A\varphi$: "φ is true in all states") for technical convenience.

Our notion of evidence of having evidence for φ need not imply that the agent *believes* φ. In order to believe a proposition φ, the agent must consider *all* her evidence for or against φ. To model the latter, we will make use of Definition 2.3. Each maximal f.i.p. subset of $E(w)$ represents a maximally consistent theory based on (some of) the evidence collected at w. [5]

Semantics We now interpret formulas of this language on our neighborhood models,

Definition 2.5 (Truth) Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model. Truth of a formula $\varphi \in \mathcal{L}_0$ is defined inductively as follows:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$ $\quad (p \in \mathsf{At})$
- $\mathcal{M}, w \models \neg\varphi$ iff $\mathcal{M}, w \not\models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ iff $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
- $\mathcal{M}, w \models \Box\varphi$ iff there exists X such that wEX and for all $v \in X$, $\mathcal{M}, v \models \varphi$
- $\mathcal{M}, w \models B\varphi$ for each maximal f.i.p. $\mathcal{X} \subseteq E(w)$ and for all $v \in \bigcap\mathcal{X}$, $\mathcal{M}, v \models \varphi$
- $\mathcal{M}, w \models A\varphi$ iff for all $v \in W$, $\mathcal{M}, v \models \varphi$

The truth set of φ is the set $[\![\varphi]\!]_{\mathcal{M}} = \{w \mid \mathcal{M}, w \models \varphi\}$. The standard logical notions of **satisfiability** and **validity** are defined as usual. $\triangleleft$

In uniform evidence models, our main concern, these clauses work out the same. [6]

Logical validities Now we get a set of valid principles for evidence and belief. While the complete static logic of our language is not the main concern of this paper, being rather evidence dynamics, we do make a few simple observations.

Fact 2.6 *(i) A satisfies all laws of modal* **S5**,[7] *B satisfies the modal logic* **KD**,[8] *and $\Box$ satisfies only the principles of the minimal classical modal logic allowing for upward monotonicity (but conjunction under the modality fails)*[9]*. (ii) The operator connections*

[5] Analogous ideas occur in semantic theories of conditionals (cf. Kratzer, 1977; Veltman, 1976) as well as policies for belief revision (cf. Gärdenfors, 1988; Rott, 2001).

[6] One might also be interested in a notion of *knowledge* in our models. One candidate is truth in all worlds occurring in evidence sets, since these are the relevant ones. Given our special assumptions, this is the whole universe, and hence the A-modality serves as knowledge. More refined modelings of knowledge are possible, here and later in this paper, but we omit them here.

[7] By the axioms $A(\varphi \to \psi) \to (A\varphi \to A\psi)$, $A\varphi \to \varphi$, $A\varphi \to AA\varphi$, $\neg A\varphi \to A\neg A\varphi$ and the rule of Necessitation.

[8] By the axioms $B(\varphi \to \psi) \to (B\varphi \to B\psi)$, $B\varphi \to \neg B\neg\varphi$ and the rule of Necessitation.

[9] That is, $(\Box\varphi \wedge \Box\psi) \to \Box(\varphi \wedge \psi)$ is not valid. The logic only contains the axiom scheme dual $\Box\varphi \leftrightarrow \neg\Diamond\neg\varphi$ (given the usual definition of $\Diamond$ as the dual of $\Box$) and the monotonicity rule: "from $\varphi \to \psi$, infer $\Box\varphi \to \Box\psi$".

$$\begin{array}{ccccc} & & B\varphi & & \\ & \nearrow & & \searrow & \\ A\varphi & & & & E\varphi \\ & \searrow & & \nearrow & \\ & & \Box\varphi & & \end{array}$$

are valid, but no other implications hold.

Verifying part (i) of the above fact is straightforward. In particular, the observation about $\Box$ is typical for neighborhood semantics: the basic evidence modality does not allow of automatic "aggregation by conjunction". For aggregation to happen, an agent must do *work* – as we will see later on. The validity in part (ii) follows from our two basic assumptions (Cons) and (Triv): that $A\varphi \rightarrow (\Box\varphi \wedge B\varphi)$ is valid follows from the fact that the evidential states are assumed to be non-empty (by Triv, at least W is in the agent's evidential state); and the fact that $\Box\varphi \rightarrow E\varphi$ is valid follows from the assumption that evidence sets are non-empty.

But over our special class of uniform evidence models, we can say much more about validities. First note that the following are valid:

$$B\varphi \rightarrow AB\varphi \quad \text{and} \quad \Box\varphi \rightarrow A\Box\varphi.$$

It follows easily that belief introspection is trivially true, as reflected in:

$$\Box\varphi \leftrightarrow B\Box\varphi \quad \text{and} \quad \neg\Box\varphi \leftrightarrow B\neg\Box\varphi$$

These observations suggest the following more general proposition:

Proposition 2.7 *On the class of uniform evidence models, each formula of $\mathcal{L}_0$ is equivalent to a formula with modal depth 1.*

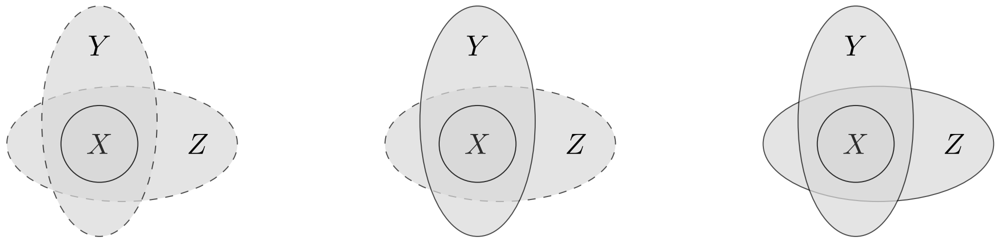
The proof is essentially the same as that for modal **S5** (cf. Meyer and van der Hoek, 2004, Section 1.7.6). Axiomatizing the logic of our models completely seems an application of standard techniques, but we do not pursue this theme here. [10]

Model theory and bisimulation While our basic system seems quite straightforward, our treatment of belief has some special features deviating from modal orthodoxy. These become visible when we consider the fundamental semantic invariance of *bisimulation* appropriate to our evidence models. [11]

For a start, recall that we do not assume that the set of evidence accepted by the agent at a world is closed under supersets, though our semantic definition implies that the set of propositions that the agent has *evidence for* is closed under weakening. This means that our language $\mathcal{L}_0$ is invariant under adding supersets of evidence already contained in a model. As a concrete illustration of this point, consider the three evidential states pictured below:

[10] The semantics for belief may pose some problems, since it does not translate obviously into standard modal or guarded first-order languages.

[11] This section is a brief technical excursion that can be skipped without loss of continuity.



$\mathcal{E}_1 = \{X\}$ $\mathcal{E}_2 = \{X, Y\}$ $\mathcal{E}_3 = \{X, Y, Z\}$

In each model the agent believes any proposition entailed by X. Furthermore, the agent has evidence for X, Y and Z (and all that they entail) in all three models. Of particular interest is evidential state $\mathcal{E}_2$, where the agent not only has evidence for Y (because of the accepted evidence X), but also has accepted Y itself as evidence. However, in general, an agent can have *evidence for* X without *accepting* the set X as evidence.[12] This all suggests the following observation:

Fact 2.8 *Let $\mathcal{M} = \langle W, E, V\rangle$ and $\mathcal{M}' = \langle W, E', V\rangle$ be two models differing only in their evidence functions. Suppose that, for all $w \in W$, (1) $E(w) \subseteq E'(w)$ and (2) if $X' \in E'(w)$, there is a $X \in E(w)$ with $X \subseteq X'$. Then, for all $w \in W$ and all $\varphi \in \mathcal{L}_0$, $\mathcal{M}', w \models \varphi$ iff $\mathcal{M}', w \models \varphi$.*

The proof is an easy induction on the structure of the formulas. A natural generalization here is the "monotonic bisimulation" familiar from the literature on neighbourhood semantics (Hansen et al., 2009; Hansen, 2003) and game logics (Pauly, 2001).

Definition 2.9 (Monotonic bisimulation) Let $\mathcal{M}_1 = \langle W_1, E_1, V_1\rangle$ and $\mathcal{M}_2 = \langle W_2, E_2, V_2\rangle$ be two evidence models. A non-empty relation $Z \subseteq W_1 \times W_2$ is a **bisimulation** if, for all worlds $w_1 \in W_1$ and $w_2 \in W_2$:

Prop If $w_1 Z w_2$, then for all $p \in \mathsf{At}$, $p \in V_1(w_1)$ iff $p \in V_2(w_2)$.

Forth If $w_1 Z w_2$, then for each $X \in E_1^{sup}(w_1)$ there is a $X' \in E_2^{sup}(w_2)$ such that for all $x' \in X'$, there is a $x \in X$ such that xZx'.

Back If $w_1 Z w_2$, then for each $X \in E_2^{sup}(w_2)$ there is a $X' \in E_1^{sup}(w_1)$ such that for all $x' \in X'$, there is a $x \in X$ such that xZx'.

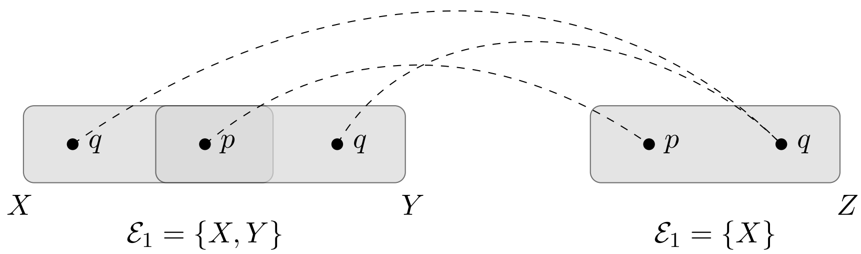
We write $\mathcal{M}_1 \leftrightarroweq \mathcal{M}_2$ there there is a bisimulation Z between $\mathcal{M}_1$ and $\mathcal{M}_2$ and $\mathcal{M}_1, w_1 \leftrightarroweq \mathcal{M}_2, w_2$ if in addition $w_1 Z w_2$. We say Z is a **total bisimulation** if it is a bisimulation and every world in W_1 is related to at least one world in W_2 and vice versa. ◁

It is a standard fact that the sublanguage of $\mathcal{L}_0$ without belief operators is invariant under total bisimulations (totality is needed for the universal modality). Thus, with respect to statements about evidential states two evidence models are the "same" if they are neighborhood bisimilar. Interestingly, beliefs are *not* invariant under this notion of bisimulation.

Fact 2.10 *The belief modality is not definable in terms of evidence modalities only.*

Proof. Consider the following example of two evidence models:

[12]This issue will return when discussing an evidence removal operation: cf. the proof of Proposition 3.12.



It is easy to see that the dashed line is a total bisimulation between the above two models. Still, note that p is believed in the model on the left (Bp is true) but not in the model on the right. QED

Of course, there remains the issue of finding a stronger notion of bisimulation respecting the complete language of evidence and belief. We refer to Section 5 for some thoughts.

2.3 Conditional belief and conditional evidence

While our language so far has a familiar look, it still lacks some basic features that have turned out essential to logics of belief. For this reason, and also in anticipation of the evidence dynamics found in the next section, we introduce *conditional* belief and evidence, replacing absolute belief $B\varphi$ and evidence $\Box\varphi$ with their conditional versions: $B^{\varphi}\psi$ and $\Box^{\varphi}\psi$. Let $\mathcal{L}_1$ be the resulting language. [13]

Conditional evidence The intended interpretation of $\Box^{\varphi}\psi$ is "the agent has evidence that ψ is true conditional on φ being true". This is almost immediate, but not quite. When conditioning on φ the agent may have evidence X inconsistent with φ. Thus, we cannot simply intersect each piece of evidence with the truth set of φ. We say that $X \subseteq W$ is **consistent (compatible) with** φ if $X \cap [\![\varphi]\!]_{\mathcal{M}} \neq \emptyset$. Truth of conditional evidence can then be defined as follows:

- $\mathcal{M}, w \models \Box^{\varphi}\psi$ iff there exists an evidence set $X \in E(w)$ which is consistent with φ such that for all worlds $v \in X \cap [\![\varphi]\!]_{\mathcal{M}}$, $\mathcal{M}, v \models \varphi$.

In particular, if there is no evidence consistent with φ, then $\Box^{\varphi}\psi$ is false. This, in turn means that $\Box^{\varphi}\psi$ is not equivalent to $\Box(\varphi \to \psi)$. [14] Indeed, a simple bisimulation argument shows that no definition exists for conditional evidence in the language with absolute evidence and belief.

Conditional belief Defining conditional belief ($B^{\varphi}\psi$) requires us to understand what it means to "relativize" an evidence model to some formula φ. Some of the agent's current evidence may be inconsistent with φ (i.e., disjoint with $[\![\varphi]\!]_{\mathcal{M}}$). If one is restricting attention to situations where φ is true, then such inconsistent evidence must be "ignored". Here is how we do this:

Definition 2.11 (Relativized maximal overlapping evidence) Suppose that $X \subseteq W$. Given a collection $\mathcal{X}$ of subsets of W (i.e., $\mathcal{X} \subseteq \wp(W)$), the relativization of $\mathcal{X}$ to X is the set $\mathcal{X}^X =$

[13] As usual, we can define absolute belief and evidence as follows: $B\varphi := B^{\top}\varphi$ and $\Box\varphi := \Box^{\top}\varphi$. That these syntactic definitions are semantically correct will become apparent below.

[14] To see this, consider a model where φ is false at all worlds. Then $\Box^{\varphi}\psi$ is also false at all worlds, but $\Box(\varphi \to \psi)$ will be true at all worlds, since $\varphi \to \psi$ is true everywhere.

$\{Y \cap X \mid Y \in \mathcal{X}\}$. We say that a collection $\mathcal{X}$ of subsets of W has the **finite intersection property relative to** X (X-f.i.p.) if, for each $\{X_1, \ldots, X_n\} \subseteq \mathcal{X}^X$, $\bigcap_{1 \le i \le n} X_i \neq \emptyset$. $\mathcal{X}$ has the **maximal** X**-f.i.p.** if $\mathcal{X}$ has X-f.i.p. and no proper extension $\mathcal{X}'$ of X has the X-f.i.p. $\lhd$

To simplify notation, when X is the truth set of formula φ, we write "maximal φ-f.i.p." for "maximal $[\![\varphi]\!]_{\mathcal{M}}$-f.i.p." and "$\mathcal{X}^{\varphi}$" for "$\mathcal{X}^{[\![\varphi]\!]_{\mathcal{M}}}$". Now we define conditional belief:

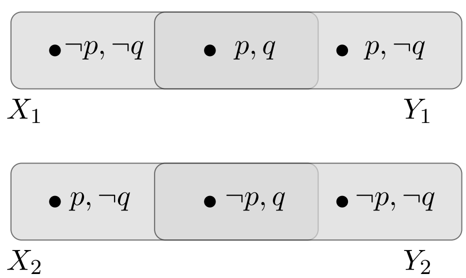
- $\mathcal{M}, w \models B^{\varphi}\psi$ iff for each maximal φ-f.i.p. $\mathcal{X} \subseteq E(w)$, for each $v \in \bigcap \mathcal{X}^{\varphi}$, $\mathcal{M}, v \models \psi$

Of course, $B\varphi \to B^{\psi}\varphi$ is not valid. One can compare this to the failure of monotonicity for antecedents in conditional logic. In our more general setting which allows inconsistencies among accepted evidence, we also see that the following variant is not valid:

$$B\varphi \to B^{\psi}\varphi \vee B^{\neg\psi}\varphi$$

This is interesting as it is valid on *connected* plausibility models.

Example 2.12 Consider the model $\mathcal{M} = \langle W, E, V \rangle$ whose evidence function for world w is pictured below. There are four evidence sets (i.e., $E(w) = \{X_1, X_2, Y_1, Y_2\}$). There are two maximal f.i.p. $\mathcal{X}_1 = \{X_1, Y_1\}$ and $\mathcal{X}_2 = \{X_2, Y_2\}$. Since $X_i \cap Y_i \subseteq [\![q]\!]_{\mathcal{M}}$ for $i = 1, 2$, we have $\mathcal{M}, w \models Bq$.



When the agent conditions on p, there is are two maximal $[\![p]\!]_{\mathcal{M}}$-f.i.p. sets $\mathcal{X}_1 = \{X_1, Y_1\}$ and $\mathcal{X}_2 = \{X_1\}$. Since $\cap\mathcal{X}_2 \not\subseteq [\![p]\!]_{\mathcal{M}}$ we have $\mathcal{M}, w \models \neg B^{p}q$. Furthermore, there are two $[\![\neg p]\!]_{\mathcal{M}}$-f.i.p. sets $\mathcal{X}_1 = \{X_1\}$ and $\mathcal{X}_2 = \{Y_1, Y_2\}$. Since $\cap\mathcal{X}_1 \not\subseteq [\![q]\!]_{\mathcal{M}}$, we also have $\mathcal{M}, w \models \neg B^{\neg p}q$.

The complete logic of the above notions show strong analogies with the basic logic of *conditionals* as found in (Burgess, 1981; Veltman, 1985; Girard, 2008). [15] Conditional belief satisfies

1. $B^{\varphi}\varphi$
2. $B^{\varphi}\psi \to B^{\varphi}\psi \vee \chi$
3. $(B^{\varphi}\psi_1 \wedge B^{\varphi}\psi_2) \to B^{\varphi}(\psi_1 \wedge \psi_2)$
4. $(B^{\varphi}\psi_1 \wedge B^{\varphi}\psi_2) \to B^{\varphi}(\psi_1 \wedge \psi_2)$
5. $(B^{\varphi_1}\psi \wedge B^{\varphi_2}\psi) \to B^{\varphi_1 \vee \varphi_2}\psi$

[15] To make a comparison with conditional semantics easier, our treatment of belief may be viewed as involving a binary relation from any world w to all worlds occurring in the intersection of all maximally f.i.p. families of evidence.

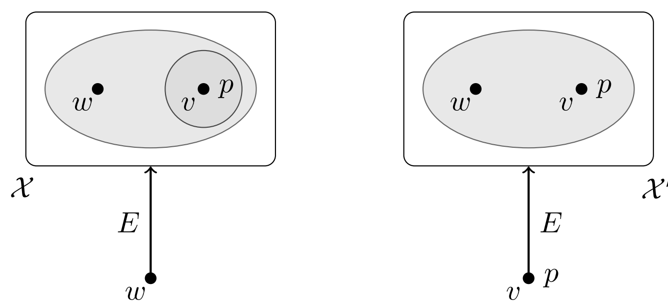
6. $(B^{\varphi}\psi \wedge B^{\psi}\varphi) \rightarrow (B^{\varphi}\chi \leftrightarrow B^{\psi}\chi)$

Compare the fifth validity to the formula discussed in the above example. Further possible validities include "cautious monotonicity": $(B^{\varphi}\psi \wedge B^{\varphi}\chi) \rightarrow B^{\varphi\wedge\psi}\chi$. [16] But of course, the full logic that we need would add evidence and conditional evidence modalities, which amounts to mixing normal and non-normal conditional logics in one system. [17]

2.4 Non-uniform evidence models

For ease of exposition, this paper works with uniform evidence models, the neighborhood counterpart of binary, rather than ternary accessibility relations between worlds. However, it is easy to generalize our analysis to models where different worlds may have different evidence sets.

Example 2.13 Here is how the earlier introspection property for evidence fails. Suppose there are two worlds $W = \{w, v\}$ and an atomic proposition p true only at v. We will construct an evidence relation so that $\mathcal{M}, w \models \Box p \wedge \neg\Box\Box p$. Suppose there are two evidential states, $\mathcal{X} = \{\{v\}, W\}$ and $\mathcal{X}' = \{W\}$. The evidence relation E relates w to $\mathcal{X}$ and v to $\mathcal{X}'$ as pictured below:



In the above model, w is the only state where $\mathcal{M}, w \models \Box p$, so $[\![\Box p]\!]_{\mathcal{M}} = \{w\}$. However, there is no evidence for $\{w\}$ in $\mathcal{X}$, so $\mathcal{M}, w \models \neg\Box\Box p$.

As for the logic, we now get a much weaker set of valid principles, supporting many new distinctions. In particular, no iterations of evidence operators will be collapsed, the same way the basic conditional logic does not collapse nested implications.

Finally, as in standard modal logic, one can study the effects of imposing special conditions on our general models. This was true before already: it is well-known that imposing $(\Box\varphi \wedge \Box\psi) \rightarrow (\Box(\varphi \wedge \psi))$ requires closure of evidence sets under intersections, leading to a collapse into relational semantics. Likewise, for example, the above introspection principle $\Box\varphi \rightarrow \Box\Box\varphi$ is true on an evidence frame (i.e., a model without a valuation) iff the evidence function E has the following property: for all w, if $X \in E(w)$, then $\{v \mid \text{there is an } Y \in E(v) \text{ with } Y \subseteq X\} \in E(w)$. Hansen (2003) develops a systematic modal correspondence theory in this context.

This concludes our discussion of the static base language of evidence models. While this may look like a rich repertoire already, it still follows familiar patterns for absolute and conditional

[16] These depend on evidence counterparts to transitivity in order semantics, but we omit technical details. Our treatment of belief may be viewed as involving a binary relation from any world w to all worlds occurring in the intersection of a maximally f.i.p. families of evidence.

[17] We are not aware of complete axiomatizations for such mixed systems.

belief. However, as we shall see in the following section, there are still further natural evidence modalities, and we will bring them to light through a dynamic analysis of operations that change one's current evidence.

3 Natural Operations on Evidence Models

Evidence is not a static substance that we have once and for all. It is continually affected by new incoming information, and also by processes of internal re-evaluation. Our main point in this paper is to show how this dynamics can be naturally made visible on the neighborhood models that we have introduced so far.

Our methodology in doing so comes from recent dynamic logics of knowledge update (van Ditmarsch et al., 2007; van Benthem, 2011) and belief revision (van Benthem, 2004; Baltag and Smets, 2006a), which model a rich repertoire of informational actions driving agency. Formally, these come as operations that change current models, viewed as snapshots of an agent's information and attitudes in some relevant process over time. Examples range from "hard" information change provided by *public announcements* (Plaza, 2007; Gerbrandy, 1999) to softer announcements encoding different policies of belief revision (cf. Rott, 2006) by more radical or more conservative *upgrade*s of plausibility orderings. Nowadays there are also dynamic logics of acts of inference or introspection that raise 'awareness' (van Benthem, 2008; Velazquez-Quesada, 2009), and of questions that modify the focus of a current process of inquiry (van Benthem and Minica, 2009).

Our neighborhood models of evidence and belief suggest a new scope for all these methods, as they bring out various forms of what may be called *evidence dynamics*. What we expect, and what we will soon find confirmed, in line with our richer structure as compared with standard relational models, is a finer view of possible changes that can take place. [18]

3.1 Public announcements

We start with the situation where the agent considers the (source of) the new evidence infallible. This type of informational action has been extensively studied in the dynamic epistemic logic literature and is usually called *public announcement* (or also "public observation").

Definition 3.1 (Public Announcement) Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model and $\varphi \in \mathcal{L}_1$ a formula. The model $\mathcal{M}^{!\varphi} = \langle W^{!\varphi}, E^{!\varphi}, V^{!\varphi}\rangle$ is defined as follows: $W^{!\varphi} = [\![\varphi]\!]_{\mathcal{M}}$, for each $p \in \mathsf{At}$, $V^{!\varphi}(p) = V(p) \cap W^{!\varphi}$ and for all $w \in W$,

$$E^{!\varphi}(w) = \{X \mid \emptyset \neq X = Y \cap [\![\varphi]\!]_{\mathcal{M}} \text{ for some } Y \in E(w)\}. \qquad \triangleleft$$

There is again a corresponding modal operator $[!\varphi]\psi$, this time, expressing that "ψ is true after the public announcement of φ". Its truth is defined as usual:

$$\text{(PA)} \quad \mathcal{M}, w \models [!\varphi]\psi \text{ iff } \mathcal{M}, w \models \varphi \text{ implies } \mathcal{M}^{!\varphi}, w \models \psi.$$

On evidence models, the standard recursion axioms for public announcement then turn out to be valid, and we get a new set of dynamic equations for evidence under hard information. Again, we get a relative completeness theorem, which we state in slightly less detail than the preceding one:

[18] Dynamic-epistemic views of neighborhood models with different motivations coming from game theory and probability theory are found in (Ågotnes and van Ditmarsch, 2010; Demey, 2010a; Zvesper, 2010).

Theorem 3.2 *The dynamic logic of evidence-based belief under public announcement is axiomatized completely over the chosen static base logic by (a) the minimal modal logic for separate dynamic modalities, (b) the following set of recursion axioms:*

(PA1)	$[!\varphi]p$	$\leftrightarrow$	$(\varphi \to p)$	$(p \in \mathsf{At})$
(PA2)	$[!\varphi](\psi \wedge \chi)$	$\leftrightarrow$	$([!\varphi]\psi \wedge [!\varphi]\chi)$	
(PA3)	$[!\varphi]\neg\psi$	$\leftrightarrow$	$(\varphi \to \neg[!\varphi]\psi)$	
(PA4)	$[!\varphi]\Box\psi$	$\leftrightarrow$	$(\varphi \to \Box^{\varphi}[!\varphi]\psi)$	
(PA5)	$[!\varphi]B\psi$	$\leftrightarrow$	$(\varphi \to B^{\varphi}[!\varphi]\psi)$	
(PA6)	$[!\varphi]\Box^{\alpha}\psi$	$\leftrightarrow$	$(\varphi \to \Box^{\varphi \wedge [!\varphi]\alpha}[!\varphi]\psi)$	
(PA7)	$[!\varphi]B^{\alpha}\psi$	$\leftrightarrow$	$(\varphi \to B^{\varphi \wedge [!\varphi]\alpha}[!\varphi]\psi)$	
(PA8)	$[!\varphi]A\psi$	$\leftrightarrow$	$(\varphi \to A[!\varphi]\psi)$	

Table 2: Public Announcement Recursion Axioms

Proof. The general strategy of the proof is standard. Owing to lack of space, we omit detailed verifications for the key axioms $PA1$ - $PA8$, which may be obtained in analogy with the calculations in Demey (2010b) and (Zvesper, 2010, Chapter 3). QED

From the perspective of our evidence models, a public announcement of φ can be naturally "deconstructed" into a complex combination of three distinct operations:

1. **Evidence addition**: the agent accepts that φ is an "admissible" piece of evidence (perhaps on par with the other available evidence).
2. **Evidence removal**: the agent removes any evidence for $\neg\varphi$.
3. **Evidence modification**: the agent incorporates φ into each piece of evidence gathered so far, making φ the most important piece of evidence.

Our richer evidence models allows us to study each of these operations individually.

3.2 Evidence addition

The first action we investigate is accepting an input from a trusted source. Let $\mathcal{M}$ be an evidence model and φ a new piece of evidence which the agent decides to *accept*. Here "acceptance" does not necessarily mean that the agent believes that φ is true, but rather that she agrees that φ should be considered when "weighing" her evidence. The formal definition of this action is straightforward:

Definition 3.3 (Evidence Addition) Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model, and φ a formula in $\mathcal{L}_1$. The model $\mathcal{M}^{+\varphi} = \langle W^{+\varphi}, E^{+\varphi}, V^{+\varphi}\rangle$ has $W^{+\varphi} = W$, $V^{+\varphi} = V$ and for all $w \in W$,

$$E^{+\varphi}(w) = E(w) \cup \{[\![\varphi]\!]_{\mathcal{M}}\}. \qquad \triangleleft$$

This operation can be described explicitly in our modal language with a dynamic modality $[+\varphi]\psi$ which is intended to mean "ψ is true after φ is accepted as an admissible piece of evidence". The truth condition for this formula is straightforward:

$$\text{(EA)} \qquad \mathcal{M}, w \models [+\varphi]\psi \text{ iff } \mathcal{M}, w \models E\varphi \text{ implies } \mathcal{M}^{+\varphi}, w \models \psi.$$

Here, since evidence sets cannot be empty, the precondition for evidence addition is that φ is true at some state. Compare this with the better-known precondition for public announcement which requires the announced formula to be true.

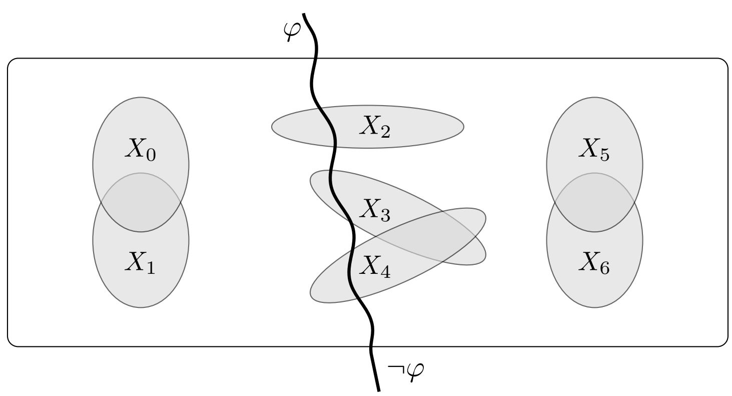
Now we can capture the essence of evidence addition in logical terms. The goal is to find recursive axioms that precisely describe the effect of its action on models, and may be viewed as the 'dynamic equations' of this evidence change.

Definition 3.4 (Compatible/Incompatible) Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model and $\mathcal{X} \subseteq E(w)$ a collection of evidence sets. Let φ be a formula, we say

1. $\mathcal{X}$ is maximally φ-**compatible** provided $\cap\mathcal{X} \cap [\![\varphi]\!]_{\mathcal{M}} \neq \emptyset$ and no proper extension $\mathcal{X}'$ of $\mathcal{X}$ has this property; and
2. $\mathcal{X}$ is **incompatible** with φ provided there are $X_1, \ldots, X_n \in \mathcal{X}$ such that $X_1 \cap \cdots \cap X_n \subseteq [\![\neg\varphi]\!]_{\mathcal{M}}$. ◁

The following example illustrates this definition.

Example 3.5 Consider the evidential state pictured below. There are four maximal f.i.p. sets $\mathcal{X}_1 = \{X_0, X_1\}, \mathcal{X}_2 = \{X_2\}, \mathcal{X}_3 = \{X_3, X_4\}$ and $\mathcal{X}_4 = \{X_5, X_6\}$. Given the a formula φ, we partition the the evidential state into those states satisfying φ and those satisfying $\neg\varphi$:



- Maximally φ-compatible sets: $\mathcal{X}_1 = \{X_0, X_1\}$, $\mathcal{X}_2 = \{X_2\}$, $\mathcal{X}_3 = \{X_3\}$, $\mathcal{X}_4 = \{X_4\}$
- Maximal f.i.p. sets incompatible with φ: $\mathcal{X}_1 = \{X_3, X_4\}$, $\mathcal{X}_2 = \{X_5, X_6\}$
- Maximally $\neg\varphi$-compatible sets: $\mathcal{X}_1 = \{X_2\}$, $\mathcal{X}_2 = \{X_5, X_6\}$
- Maximal f.i.p. sets incompatible with $\neg\varphi$: $\mathcal{X}_1 = \{X_0, X_1\}$, $\mathcal{X}_2 = \{X_2\}$

In particular, note that the maximal f.i.p. sets that are incompatible with φ are not the same as those that are maximally $\neg\varphi$-compatible: For example, $\{X_2\}$ is maximally $\neg\varphi$-compatible but not incompatible with φ, and $\{X_3, X_4\}$ is a maximal f.i.p. set incompatible with φ that is not maximally $\neg\varphi$-compatible.

We can rephrase the definition of conditional belief (which we now denote $B^{+\varphi}\psi$) as follows:

$$\mathcal{M}, w \models B^{+\varphi}\psi \text{ iff for each maximally } \varphi\text{-compatible } \mathcal{X} \subseteq E(w), \bigcap \mathcal{X} \cap [\![\varphi]\!]_{\mathcal{M}} \subseteq [\![\psi]\!]_{\mathcal{M}}$$

But, more interesting, is a new conditional belief operator, now based on *incompatibility*:

Conditional Beliefs (Incompatibility Version) $\mathcal{M}, w \models B^{-\varphi}\psi$ iff for all maximal f.i.p., if $\mathcal{X}$ is incompatible with φ then $\bigcap \mathcal{X} \subseteq [\![\psi]\!]_{\mathcal{M}}$.

Lemma 3.6 *The formula* $[+\varphi]B\psi \leftrightarrow E\varphi \rightarrow (B^{+\varphi}[+\varphi]\psi \wedge B^{-\varphi}[+\varphi]\psi)$ *is valid*

Proof. Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model and φ a formula with $[\![\varphi]\!]_{\mathcal{M}} \neq \emptyset$. We first note the following facts:

1. $\mathcal{X} \subseteq E(w)$ is maximally φ-compatible iff $\mathcal{X} \cup \{[\![\varphi]\!]_{\mathcal{M}}\} \subseteq E^{+\varphi}(w)$ is a maximal f.i.p.
2. $\mathcal{X} \subseteq E(w)$ is a maximal f.i.p. that is incompatible with φ iff $\mathcal{X} \subseteq E^{+\varphi}(w)$ is a maximal f.i.p. that does not contain $[\![\varphi]\!]_{\mathcal{M}}$.

The proof of both of these facts follows from the observation that $E(w) \subseteq E^{+\varphi}(w)$ and any $\mathcal{X}$ that is a maximal f.i.p. in $E^{+\varphi}(w)$ but not in $E(w)$ *must* contain $[\![\varphi]\!]_{\mathcal{M}}$.

Suppose $\mathcal{M}, w \models [+\varphi]B\psi$. Then,

$$(*) \qquad \text{for all maximal f.i.p. } \mathcal{X} \subseteq E^{+\varphi}(w), \text{ we have } \bigcap \mathcal{X} \subseteq [\![\psi]\!]_{\mathcal{M}^{+\varphi}}$$

We must show $\mathcal{M}, w \models B^{+\varphi}[+\varphi]\psi \wedge B^{-\varphi}[+\varphi]\psi$. To see that the left conjunct is true, let $\mathcal{X} \subseteq E(w)$ be any maximally φ-compatible collection of evidence. By (1), $\mathcal{X} \cup \{[\![\varphi]\!]_{\mathcal{M}}\} \subseteq E^{+\varphi}(w)$ is a maximal f.i.p. set. Then, we have

$$\bigcap \mathcal{X} \cap [\![\varphi]\!]_{\mathcal{M}} = \bigcap (\mathcal{X} \cup \{[\![\varphi]\!]_{\mathcal{M}}\}) \subseteq [\![\psi]\!]_{\mathcal{M}^{+\varphi}} = [\![[+\varphi]\psi]\!]_{\mathcal{M}}$$

where the containment follows from $(*)$. Since $\mathcal{X}$ was an arbitrary maximally φ-compatible set, we have $\mathcal{M}, w \models B^{+\varphi}[+\varphi]\psi$. For the right conjunct, let $\mathcal{X} \subseteq E(w)$ be any maximal f.i.p. that is incompatible with φ. By (2), $\mathcal{X} \subseteq E^{+\varphi}(w)$ is a maximal f.i.p. (that does not contain $[\![\varphi]\!]_{\mathcal{M}}$). Again by $(*)$ we have

$$\bigcap \mathcal{X} \subseteq [\![\psi]\!]_{\mathcal{M}^{+\varphi}} = [\![[+\varphi]\psi]\!]_{\mathcal{M}}$$

Hence, since $\mathcal{X}$ was an arbitrary maximal f.i.p. subset of $E(w)$ incompatible with φ, we have $\mathcal{M}, w \models B^{-\varphi}[+\varphi]\psi$. This shows that $[+\varphi]B\psi \rightarrow B^{+\varphi}[+\varphi]\psi \wedge B^{+\varphi}[+\varphi]\psi$ is valid.

Suppose now that $\mathcal{M}, w \models B^{+\varphi}[+\varphi]\psi \wedge B^{+\varphi}[+\varphi]\psi$. Then

A. For all maximally φ-compatible $\mathcal{X} \subseteq E(w)$, we have $\bigcap \mathcal{X} \cap [\![\varphi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\psi]\!]_{\mathcal{M}}$; and

B. For all maximally f.i.p. $\mathcal{X} \subseteq E(w)$ incompatible with φ, we have $\bigcap \mathcal{X} \subseteq [\![[+\varphi]\psi]\!]_{\mathcal{M}}$.

We must show $\mathcal{M}^{+\varphi}, w \models B\psi$. Let $\mathcal{X} \subseteq E^{+\varphi}(w)$ be a maximal f.i.p. set. There are two cases to consider. First, $[\![\varphi]\!]_{\mathcal{M}} \in \mathcal{X}$. Then, by (1), $\mathcal{X} - \{[\![\varphi]\!]_{\mathcal{M}}\} \subseteq E(w)$ is maximally φ-compatible. Furthermore, by (A) we have

$$\bigcap \mathcal{X} = \bigcap (\mathcal{X} - \{[\![\varphi]\!]_{\mathcal{M}}\}) \cap [\![\varphi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\psi]\!]_{\mathcal{M}} = [\![\psi]\!]_{\mathcal{M}^{+\varphi}}$$

The second case is $[\![\varphi]\!]_{\mathcal{M}} \notin \mathcal{X}$. Then by (2), $\mathcal{X} \subseteq E(w)$ is a maximal f.i.p. that is incompatible with φ. By (B), we have

$$\bigcap \mathcal{X} \subseteq [\![[+\varphi]\psi]\!]_{\mathcal{M}} = [\![\psi]\!]_{\mathcal{M}^{+\varphi}}$$

In either case, $\bigcap \mathcal{X} \subseteq [\![\psi]\!]_{\mathcal{M}^{+\varphi}}$; hence, $\mathcal{M}^{+\varphi}, w \models B\psi$, as desired. QED

A complete logical analysis of recursion axioms has to mix compatibility-based conditional belief and the new incompatibility-based notion, generalizing them to one notion of belief that is conditional on one proposition, but only at stages where some other proposition has already been settled.[19] Thus, we find a new conditional belief operator which generalizes both types of conditional belief operators we have encountered thus far.

Language Extension Let $\mathcal{L}_2$ be the smallest set of formulas generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box\varphi \mid B^{\varphi,\psi}\chi \mid A\varphi$$

where $p \in \mathsf{At}$ and $\overline{\varphi}$ is any finite sequence of formulas from the language.[20]

Definition 3.7 (Truth of $\mathcal{L}_2$) We only define the new modal operator:

- $\mathcal{M}, w \models B^{\varphi,\psi}\chi$ iff for all maximally φ-compatible sets $\mathcal{X} \subseteq E(w)$, if $\bigcap \mathcal{X} \cap [\![\varphi]\!]_{\mathcal{M}} \subseteq [\![\psi]\!]_{\mathcal{M}}$, then $\bigcap \mathcal{X} \cap [\![\varphi]\!]_{\mathcal{M}} \subseteq [\![\chi]\!]_{\mathcal{M}}$.

Note that we defined $B^{+\varphi}$ as $B^{\varphi,\top}$ and $B^{-\varphi}$ as $B^{\top,\neg\varphi}$ $\triangleleft$

As usual in dynamic-epistemic logic, we concentrate on the principles that need to be added to the static base logic. A precise choice of models and valid principles for the latter is largely a 'free parameter' in our analysis.

Theorem 3.8 *The dynamic logic of evidence addition is axiomatized completely by (a) the static base logic of evidence models, (b) the minimal modal logic for each separate dynamic modality, and (c) the following set of recursion axioms:*

[19] Excluding φ as in the above incompatibility notion amounts to having settled that $\neg\varphi$ holds. The idea of conditionality relative to some proposition having been settled is reminiscent of the semantics of intuitionistic implication, and analogies with relational models with be pursued in Section 4 below.

[20] Absolute belief and evidence versions again arise by setting some parameters to $\top$.

(EA1)	$[+\varphi]p$	$\leftrightarrow$	$(E\varphi \to p)$ *($p \in \mathsf{At}$)*
(EA2)	$[+\varphi](\psi \wedge \chi)$	$\leftrightarrow$	$([+\varphi]\psi \wedge [+\varphi]\chi)$
(EA3)	$[+\varphi]\neg\psi$	$\leftrightarrow$	$(E\varphi \to \neg[+\varphi]\psi)$
(EA4)	$[+\varphi]\Box\psi$	$\leftrightarrow$	$(E\varphi \to (\Box[+\varphi]\psi \vee A(\varphi \to [+\varphi]\psi)))$
(EA5)	$[+\varphi]B\psi$	$\leftrightarrow$	$(E\varphi \to (B^{+\varphi}[+\varphi]\psi \wedge B^{-\varphi}[+\varphi]\psi))$
(EA6)	$[+\varphi]\Box^{\alpha}\psi$	$\leftrightarrow$	$(E\varphi \to (\Box^{[+\varphi]\alpha}[+\varphi]\psi \vee (E(\varphi \wedge [+\varphi]\alpha) \wedge$ $A((\varphi \wedge [+\varphi]\alpha) \to [+\varphi]\psi))))$
(EA7)	$[+\varphi]B^{\psi,\alpha}\chi$	$\leftrightarrow$	$(E\varphi \to (B^{\varphi\wedge[+\varphi]\psi,[+\varphi]\alpha}[+\varphi]\chi \wedge B^{[+\varphi]\psi,\neg\varphi\wedge[+\varphi]\alpha}[+\varphi]\chi))$
(EA8)	$[+\varphi]A\psi$	$\leftrightarrow$	$(E\varphi \to A[+\varphi]\psi)$

Table 1: Evidence Addition Recursion Axioms

Proof. As for soundness, we mainly explain what the recursion axioms say, which then easily justifies their validity. The first three axioms express the usual relationship between a dynamic-epistemic modality and boolean connectives. For example, axiom $EA3$ says that evidence addition is *functional*: it maps each evidence model to the unique model representing the situation after the evidence is accepted. Axioms $EA4$ - $EA8$ then describe the precise effect of evidence addition on the agent's (conditional) beliefs and accepted evidence. Axiom $EA4$ says that after accepting φ as evidence, the agent has evidence that ψ just in case either the agent had evidence for ψ before adding φ, or ψ was implied by φ in the model. Axiom $EA5$ shows the effect of accepting the evidence φ on the agent's beliefs. An agent comes to believe ψ after accepting φ as evidence just in case the agent believed ψ conditional on φ being true *and* believes ψ conditional on the incompatibility of φ. This should be contrasted with the much stronger law for *public announcement* (cf. axiom $PA5$ in Section 3.1 above) which only considers whether the agent believed ψ conditional on φ being true.

Here are a few more detailed verifications of some key recursion axioms. To simplify the presentation, assume that $[\![\varphi]\!]_{\mathcal{M}} \neq \emptyset$ (so $\mathcal{M}, w \models E\varphi$).

(Axiom $EA4$.) $\mathcal{M}, w \models [+\varphi]\Box\psi$ iff $\mathcal{M}^{+\varphi}, w \models \Box\psi$ iff there is an $X \in E^{+\varphi}(w)$ with $X \subseteq [\![\psi]\!]_{\mathcal{M}^{+\varphi}}$. By definition, we have $[\![\psi]\!]_{\mathcal{M}^{+\varphi}} = [\![[+\varphi]\psi]\!]_{\mathcal{M}}$. There are two cases to consider for the axiom:

1. $X \in E(w)$. Then, $\mathcal{M}, w \models \Box[+\varphi]\psi$.

2. $X = [\![\varphi]\!]_{\mathcal{M}}$. This means that $[\![\varphi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\psi]\!]_{\mathcal{M}}$ and so $\mathcal{M}, w \models A(\varphi \to [+\varphi]\psi)$.

(Axiom $EA5$.) The validity of this axiom is proven in Lemma 3.6.

(Axiom $EA6$.) $\mathcal{M}, w \models [+\varphi]\Box^{\alpha}\psi$ iff $\mathcal{M}^{+\varphi}, w \models \Box^{\alpha}\psi$ iff there exists $X \in E^{+\varphi}(w)$ consistent with α and having $X \cap [\![\alpha]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\psi]\!]_{\mathcal{M}^{+\varphi}}$. Again there are two cases:

1. $X \in E(w)$. Then we have $X \cap [\![[+\varphi]\alpha]\!]_{\mathcal{M}} = X \cap [\![\alpha]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\psi]\!]_{\mathcal{M}^{+\varphi}} = [\![[+\varphi]\psi]\!]_{\mathcal{M}}$. Hence $\mathcal{M}, w \models \Box^{[+\varphi]\alpha}[+\varphi]\psi$.

2. $X = [\![\varphi]\!]_{\mathcal{M}}$. Then we have $[\![\varphi]\!]_{\mathcal{M}} \cap [\![\alpha]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\psi]\!]_{\mathcal{M}^{+\varphi}}$. Therefore, $\mathcal{M}, w \models A((\varphi \wedge [+\varphi]\alpha) \to [+\varphi]\psi)$. Furthermore, since X is consistent with α in $\mathcal{M}^{+\varphi}$, we have $[\![\varphi]\!]_{\mathcal{M}} \cap [\![\alpha]\!]_{\mathcal{M}^{+\varphi}} \neq \emptyset$, and hence $\mathcal{M}, w \models E(\varphi \wedge [+\varphi]\alpha)$.

(Axiom $EA7$.) The proof here is similar to the proof of Lemma 3.6. We first note the following facts:

1. $\mathcal{X} \subseteq E(w)$ is maximally $\varphi \wedge [+\varphi]\psi$-compatible with $\bigcap \mathcal{X} \cap [\![\varphi \wedge [+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\alpha]\!]_{\mathcal{M}}$ iff $\mathcal{X} \cup \{[\![\varphi]\!]_{\mathcal{M}}\} \subseteq E^{+\varphi}(w)$ is maximally ψ-compatible with $\bigcap(\mathcal{X} \cup \{[\![\varphi]\!]_{\mathcal{M}}\}) \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\alpha]\!]_{\mathcal{M}^{+\varphi}}$

2. $\mathcal{X} \subseteq E(w)$ is maximally $[+\varphi]\psi$-compatible with $\bigcap \mathcal{X} \cap [\![[+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![\neg\varphi \wedge [+\varphi]\alpha]\!]_{\mathcal{M}}$ iff $\mathcal{X} \subseteq E^{+\varphi}(w)$ is maximally ψ-compatible such that $[\![\varphi]\!]_{\mathcal{M}} \notin \mathcal{X}$ and $\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\alpha]\!]_{\mathcal{M}^{+\varphi}}$.

The proof of these facts is straightforward. For the proof of (2), note that if $\mathcal{X} \subseteq E^{+\varphi}(w)$ is maximally ψ-compatible and $[\![\varphi]\!]_{\mathcal{M}} \notin \mathcal{X}$ then we must have $\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\neg\varphi]\!]_{\mathcal{M}}$. Suppose that $\mathcal{M}, w \models [+\varphi]B^{\psi,\alpha}\chi$. Then,

(∗) for all maximally ψ-compatible $\mathcal{X} \subseteq E^{+\varphi}(w)$ with $\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\alpha]\!]_{\mathcal{M}^{+\varphi}}$, we have $\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\chi]\!]_{\mathcal{M}^{+\varphi}}$

We must show $\mathcal{M}, w \models B^{\varphi \wedge [+\varphi]\psi, [+\varphi]\alpha}[+\varphi]\chi \wedge B^{[+\varphi]\psi, \neg\varphi \wedge [+\varphi]\alpha}[+\varphi]\chi$. To see that the first conjunct is true, let $\mathcal{X} \subseteq E(w)$ be maximally $\varphi \wedge [+\varphi]\psi$-compatible with $\bigcap \mathcal{X} \cap ([\![\varphi \wedge [+\varphi]\psi]\!]_{\mathcal{M}}) \subseteq [\![[+\varphi]\alpha]\!]_{\mathcal{M}}$. By (1), $\mathcal{X} \cup \{[\![\varphi]\!]_{\mathcal{M}}\} \subseteq E^{+\varphi}(w)$ is maximally ψ-compatible with $\bigcap(\mathcal{X} \cup \{[\![\varphi]\!]_{\mathcal{M}}\}) \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\alpha]\!]_{\mathcal{M}^{+\varphi}}$. By (∗),

$$\bigcap \mathcal{X} \cap [\![\varphi \wedge [+\varphi]\psi]\!]_{\mathcal{M}} = \bigcap \mathcal{X} \cap [\![\varphi]\!]_{\mathcal{M}} \cap [\![[+\varphi]\psi]\!]_{\mathcal{M}} = \bigcap(\mathcal{X} \cup \{[\![\varphi]\!]_{\mathcal{M}}\}) \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\chi]\!]_{\mathcal{M}^{+\varphi}} = [\![[+\varphi]\chi]\!]_{\mathcal{M}}$$

Hence, $\mathcal{M}, w \models B^{\varphi \wedge [+\varphi]\psi, [+\varphi]\alpha}[+\varphi]\chi$. For the second conjunct, let $\mathcal{X} \subseteq E(w)$ be a maximally $[+\varphi]\psi$-compatible set with $\bigcap \mathcal{X} \cap [\![[+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![\neg\varphi \wedge [+\varphi]\alpha]\!]_{\mathcal{M}}$. Then by (2), $\mathcal{X} \subseteq E^{+\varphi}(w)$ is maximally ψ-compatible with $\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\alpha]\!]_{\mathcal{M}^{+\varphi}}$. By (∗), we have

$$\bigcap \mathcal{X} \cap [\![[+\varphi]\psi]\!]_{\mathcal{M}} = \bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\chi]\!]_{\mathcal{M}^{+\varphi}} = [\![[+\varphi]\chi]\!]_{\mathcal{M}}$$

Hence, $\mathcal{M}, w \models B^{[+\varphi]\psi, \neg\varphi \wedge [+\varphi]\alpha}[+\varphi]\chi$

For the converse, suppose that $\mathcal{M}, w \models B^{\varphi \wedge [+\varphi]\psi, [+\varphi]\alpha}[+\varphi]\chi \wedge B^{[+\varphi]\psi, \neg\varphi \wedge [+\varphi]\alpha}[+\varphi]\chi$. Then

A. for all maximally $\varphi \wedge [+\varphi]\psi$-compatible sets $\mathcal{X} \subseteq E(w)$ with $\bigcap \mathcal{X} \cap [\![\varphi \wedge [+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\alpha]\!]_{\mathcal{M}}$, we have $\bigcap \mathcal{X} \cap [\![\varphi \wedge [+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\chi]\!]_{\mathcal{M}}$; and

B. for all maximally $[+\varphi]\psi$-compatible sets $\mathcal{X} \subseteq E(w)$ with $\bigcap \mathcal{X} \cap [\![[+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![\neg\varphi \wedge [+\varphi]\alpha]\!]_{\mathcal{M}}$, we have $\bigcap \mathcal{X} \cap [\![[+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\chi]\!]_{\mathcal{M}}$.

We must show $\mathcal{M}^{+\varphi}, w \models B^{\psi,\alpha}\chi$. Let $\mathcal{X} \subseteq E^{+\varphi}(w)$ be a maximally ψ-compatible set with $\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\alpha]\!]_{\mathcal{M}^{+\varphi}}$. There are two cases. First, $[\![\varphi]\!]_{\mathcal{M}} \in \mathcal{X}$. Then by (1), $\mathcal{X} - \{[\![\varphi]\!]_{\mathcal{M}}\}$ is maximally $\varphi \wedge [+\varphi]\psi$-compatible with $\bigcap(\mathcal{X} - \{[\![\varphi]\!]_{\mathcal{M}}\}) \cap [\![\varphi \wedge [+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\alpha]\!]_{\mathcal{M}}$. By (A), we have

$$\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} = \bigcap(\mathcal{X} - \{[\![\varphi]\!]_{\mathcal{M}}\}) \cap [\![\varphi \wedge [+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\chi]\!]_{\mathcal{M}} = [\![\chi]\!]_{\mathcal{M}^{+\varphi}}$$

The second case is $[\![\varphi]\!]_{\mathcal{M}} \notin \mathcal{X}$. Then, by (2), $\mathcal{X} \subseteq E(w)$ is maximally $[+\varphi]\psi$-compatible with $\bigcap \mathcal{X} \cap [\![[+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![\neg\varphi \wedge [+\varphi]\alpha]\!]_{\mathcal{M}}$. By (B), we have

$$\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} = \bigcap \mathcal{X} \cap [\![[+\varphi]\psi]\!]_{\mathcal{M}} \subseteq [\![[+\varphi]\chi]\!]_{\mathcal{M}} = [\![\chi]\!]_{\mathcal{M}^{+\varphi}}$$

In either case, $\bigcap \mathcal{X} \cap [\![\psi]\!]_{\mathcal{M}^{+\varphi}} \subseteq [\![\chi]\!]_{\mathcal{M}^{+\varphi}}$, so $\mathcal{M}^{+\varphi}, w \models B^{\psi,\alpha}\chi$, as desired.

The remainder of the completeness proof follows what is by now a standard pattern in dynamic epistemic logic. Working inside out, the stated recursion axioms suffice for successively removing all dynamic modalities from a given formula, leading to a provably equivalent formula in the base language, whose logic was assumed to be complete. QED

We have now found a complete system of evidence addition with its natural associated static base modalities of conditional belief. But we do not just want to work in a piece-meal fashion. How does this language extension fit with our earlier analysis of public announcement? Things turn out to be in harmony.

Fact 3.9 *The following principle is valid principle for our extended notion of conditional belief under public announcement:*

$$[!\varphi]B^{\psi,\alpha}\chi \leftrightarrow B^{\varphi\wedge[!\varphi]\psi,\varphi\to[!\varphi]\alpha}[!\varphi]\chi$$

Further algebraic laws of evidence change? As usual, this completeness result raises further questions. The preceding completeness proof does not require a recursion axiom for combinations of the form $[+\varphi][+\psi]\alpha$, but it is of interest to ask whether one exists. Can two acts of evidence addition be compressed into one? Maybe the above combination is equivalent to $[+(\varphi \wedge \psi)]\alpha$? As it happens, in general, it is not, though it can be in special cases – say, when adding some evidence and then adding weaker evidence. Thus, there is more to the general logic of operator combinations than meets the eye.

For instance, it is well-known that public announcement satisfies valid combination principles such as

$$[!\varphi][!\psi]\alpha \leftrightarrow [!(\varphi \wedge [!\varphi]\psi)]\alpha.$$

Next, combining our two operations so far, what about valid recursion axioms for mixed combinations of evidence actions like $[+\varphi][!\psi]\alpha$? It is easy to see that no general commutation law holds, but there may be valid special cases. Here is one case where evidence addition and public announcement can be interchanged.

Fact 3.10 *The following equivalence is valid for arbitrary formulas ψ and* factual *formulas* [21] φ*,*

$$[!\varphi][+\psi]\alpha \leftrightarrow [+[!\varphi]\psi][!\varphi]\alpha$$

This is easy to show [22], as is the fact that this equivalence may fail when φ is not factual, containing modal operators sensitive to the dynamic effects of the evidence operation $+\varphi$. There are many further valid principles in this spirit, but we leave their study to some other occasion.

[21] *Factual* formulas are purely propositional without modal operators. Their truth values at worlds remain unchanged by any of the model transformations considered in this paper.

[22] The formula $[!\varphi]\psi$ describes the set of worlds that will be ψ after φ has been announced, and factual formulas do not undergo side effects of the two different orders of reaching the same model.

3.3 Evidence removal

Evidence addition and public announcement represent two different ways in which an agent can incorporate a proposition φ into her evidence. Adding φ to her evidence set means that the agent agrees to consider φ when figuring out what she believes. A public announcement of φ is stronger in that the agent also agrees to *ignore* states inconsistent with φ. But this latter attitude is interesting by itself, suggesting an act of *evidence removal*, something that is also well-known in studies of belief revision, as a natural converse to addition. Here is how such an act can be defined in our setting:

Definition 3.11 (Evidence Removal) Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model, and $\varphi \in \mathcal{L}_1$. The model $\mathcal{M}^{-\varphi} = \langle W^{-\varphi}, E^{-\varphi}, V^{-\varphi}\rangle$ has $W^{-\varphi} = W$, $V^{-\varphi} = V$ and for all $w \in W$,

$$E^{-\varphi}(w) = E(w) - \{X \mid X \subseteq [\![\varphi]\!]_{\mathcal{M}}\}. \qquad \triangleleft$$

This time, the corresponding dynamic modal operator is $[-\varphi]\psi$ ("after removing the evidence that φ, ψ is true"), whose truth is defined as follows:

$$\text{(ER)} \qquad \mathcal{M}, w \models [-\varphi]\psi \text{ iff } \mathcal{M}, w \models \neg A\varphi \text{ implies } \mathcal{M}^{-\varphi}, w \models \psi$$

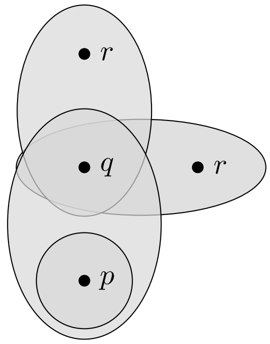
Removal or contraction[23] has been a challenge to existing systems of dynamic-epistemic logic, since it is not lear how to modify a current epistemic or doxastic relation to "retreat" into some weaker position. But our neighborhood models record more relevant history, and hence we have more structure to work with.

Again, we will do a complete reduction-axiom analysis in the earlier style. But as before, this dynamic analysis is not purely a passive imposition of an action superstructure. It affects the very choice of the base language itself, and hence it is an instrument for discovering new static logical structure concerning evidence structure. Once more, we are going to find a need for several extensions of what we had earlier on, before we find a total dynamic language that is in expressive harmony between its dynamic and its static parts. Let $\mathcal{L}_1^-$ be the extension of the language $\mathcal{L}_1$ with the dynamic operator $[-\varphi]$.

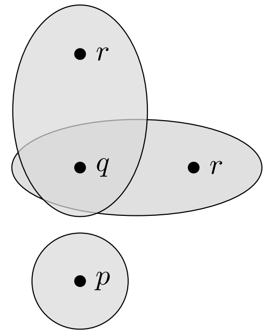
Proposition 3.12 *$\mathcal{L}_1^-$ is strictly more expressive than $\mathcal{L}_1$ over evidence models.*

Proof. Let $\mathcal{M}_1 = \langle W, \mathcal{E}_1, V\rangle$ and $\mathcal{M}_2 = \langle W, \mathcal{E}_2, V\rangle$ be two evidence models pictured below:

[23] Note that, in our setting, removing the evidence for φ is much weaker the usual notion of *contracting* by φ found in the belief revision literature (cf. Rott, 2001). In particular, it is possible to remove the evidence for φ and yet the agent maintains her belief in φ! Formally, $[-\varphi]\neg B\varphi$ is not valid. To see this, suppose that $W = \{w_1, w_2, w_3\}$ where p is true only at w_3. Consider an evidential state with two pieces of evidence: $\mathcal{E} = \{\{w_1, w_3\}, \{w_2, w_3\}\}$. The agent believes p and, since the model does not change when removing the evidence for p, we have that $[-p]Bp$ is true. Note that the same is true for the evidential state that has explicit evidence for p, i.e., $\mathcal{E}' = \{\{w_1, w_3\}, \{w_2, w_3\}, \{w_3\}\}$.



$\mathcal{E}_1$



$\mathcal{E}_2$

First of all, note that the formula $[-p]\Box(p \vee q)$ of $\mathcal{L}_1^-$ is true in $\mathcal{M}_1$ but not in $\mathcal{M}_2$. However, there is no formula of $\mathcal{L}_1$ that can distinguish $\mathcal{M}_1$ from $\mathcal{M}_2$. To see this, just note that $\mathcal{E}_1^{sup} = \mathcal{E}_2^{sup}$, while the agent has the same beliefs in both models. QED

Adding compatibility At this stage, we find the need for another key notion in our language of evidence-based belief. So far, we have looked at conditional evidence and beliefs where we added evidence that implies some given proposition. But this time, we also need to look at evidence that is merely compatible with some relevant proposition.

An agent had evidence that ψ conditional on φ if there is some evidence consistent with φ such that restriction of the evidence set to the worlds where φ is true entails ψ. Our new weaker conditional operator, denoted $\Box_\varphi\psi$, drops this latter condition. So, $\Box_\varphi\psi$ is true if the agent has evidence compatible with φ that entails ψ. [24] In general, we include operators $\Box_{\overline{\varphi}}\psi$ where $\overline{\varphi}$ is a sequence of formulas. The intended interpretation is that "ψ is entailed by some admissible evidence *compatible* with each of $\overline{\varphi}$".

Definition 3.13 (Compatible evidence) Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model and $\overline{\varphi} = (\varphi_1, \ldots, \varphi_n)$ a finite sequence of formulas. We say that a subset $X \subseteq W$ is **compatible with** $\overline{\varphi}$ provided that, for each formula φ_i, $X \cap [\![\varphi_i]\!]_\mathcal{M} \neq \emptyset$. $\triangleleft$

Truth of the new formula $\Box_{\overline{\varphi}}\psi$ is then defined as follows:

$$\mathcal{M}, w \models \Box_{\overline{\varphi}}\psi \text{ iff there is some } X \in E(w) \text{ compatible with } \overline{\varphi} \text{ where } X \subseteq [\![\psi]\!]_\mathcal{M}$$

This new operator gives us a very natural reduction axiom for $\Box$: [25]

Fact 3.14 *The formula* $[-\varphi]\Box\psi \leftrightarrow (\neg A\varphi \rightarrow \Box_{\neg\varphi}[-\varphi]\psi)$ *is valid on evidence models.*

Proof. Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model with $[\![\varphi]\!]_\mathcal{M} \neq W$ (otherwise, we have that, for all w, $E^{-\varphi}(w) = \emptyset$). We will show that $[-\varphi]\Box\psi \leftrightarrow \Box_{\neg\varphi}[-\varphi]\psi$ is valid on $\mathcal{M}$. Suppose $w \in W$. The key observation is that for all $X \subseteq W$, $X \in E^{-\varphi}(w)$ iff $X \in E(w)$ and X is compatible with $\neg\varphi$. Then we get

[24] Note that a set X may be consistent with both φ_1 and φ_2, yet not consistent with $\varphi_1 \wedge \varphi_2$.
[25] The precondition is because each evidence set must contain at least the set of all worlds W.

$$
\begin{array}{lll}
\mathcal{M}, w \models [-\varphi]\Box\varphi & \text{iff} & \mathcal{M}^{-\varphi}, w \models \Box\varphi \\
& \text{iff} & \text{there is a } X \in E^{-\varphi}(w) \text{ such that } X \subseteq [\![\psi]\!]_{\mathcal{M}^{-\varphi}} \\
& & (\text{note that } [\![\psi]\!]_{\mathcal{M}^{-\varphi}} = [\![[-\varphi]\psi]\!]_{\mathcal{M}}) \\
& \text{iff} & \text{there is a } X \in E(w) \text{ compatible with } \neg\varphi \\
& & \text{such that } X \subseteq [\![[-\varphi]\psi]\!]_{\mathcal{M}} \\
& \text{iff} & \mathcal{M}, w \models \Box_{\neg\varphi}[-\varphi]\psi.
\end{array}
$$

QED

Note how this principle captures the logical essence of evidence removal.

But we are not done yet. In this new language we also need a reduction axiom for our new operator $\Box_{\overline{\varphi}}$. This can be stated in the same style. But we are not done even then. In general, with conditional evidence present as well, we need an operator $\Box^{\alpha}_{\overline{\varphi}}\psi$ expressing there is some evidence compatible with $\overline{\varphi}$ and α such that the restriction of that evidence to α entails ψ. Moreover, we must make some adjustments to our earlier notion of (conditional) beliefs:

Definition 3.15 (Compatibility evidence - set version) A maximal f.i.p. set $\mathcal{X}$ is **compatible with** a sequence of formulas $\overline{\varphi}$ provided for each $X \in \mathcal{X}$, X is compatible with $\overline{\varphi}$. $\triangleleft$

Language and dynamic logic We are now ready to formally define our new language. Let $\mathcal{L}_3$ be the smallest set of formulas generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid B^{\alpha}_{\overline{\varphi}}\psi \mid \Box^{\alpha}_{\overline{\varphi}}\psi \mid A\varphi$$

where $p \in \mathsf{At}$ and $\overline{\varphi}$ is any finite sequence of formulas from the language.[26]

Definition 3.16 (Truth of $\mathcal{L}_3$) We only define the new modal operators:

- $\mathcal{M}, w \models \Box^{\alpha}_{\overline{\varphi}}\psi$ iff there is $X \in E(w)$ compatible with $\overline{\varphi}$, α such that $X \cap [\![\alpha]\!]_{\mathcal{M}} \subseteq [\![\psi]\!]_{\mathcal{M}}$.
- $\mathcal{M}, w \models B^{\alpha}_{\overline{\varphi}}\psi$ iff for each maximal α-f.i.p. $\mathcal{X}$ compatible with $\overline{\varphi}$, $\bigcap \mathcal{X}^{\alpha} \subseteq [\![\psi]\!]_{\mathcal{M}}$.

We write $\Box^{\alpha}_{\varphi_1,\ldots,\varphi_n}$ for $\Box^{\alpha}_{(\varphi_1,\ldots,\varphi_n)}$ and $\overline{\varphi}, \alpha$ for $(\varphi_1, \ldots, \varphi_n, \alpha)$. $\triangleleft$

We are now ready to give the reduction axioms for evidence removal over the language $\mathcal{L}_3$. Some final notation: if $\overline{\varphi} = (\varphi_1, \ldots, \varphi_n)$, then we write $[-\varphi]\overline{\varphi}$ for $([-\varphi]\varphi_1, \ldots, [-\varphi]\varphi_n)$.

Theorem 3.17 *The complete dynamic logic of evidence removal is axiomatized, over the complete logic of the static base language as enriched above, by the following reduction axioms:*

[26] Absolute belief and evidence versions again arise by setting some parameters to $\top$.

(ER1)	$[-\varphi]p$	$\leftrightarrow$	$(\neg A\varphi \to p)$	$(p \in \mathsf{At})$
(ER2)	$[-\varphi](\psi \wedge \chi)$	$\leftrightarrow$	$([-\varphi]\psi \wedge [-\varphi]\chi)$	
(ER3)	$[-\varphi]\neg\psi$	$\leftrightarrow$	$(\neg A\varphi \to \neg[-\varphi]\psi)$	
(ER4)	$[-\varphi]\Box^{\alpha}_{\overline{\psi}}\chi$	$\leftrightarrow$	$(\neg A\varphi \to \Box^{[-\varphi]\alpha}_{[-\varphi]\overline{\psi},\neg\varphi}[-\varphi]\chi)$	
(ER5)	$[-\varphi]B^{\alpha}_{\overline{\psi}}\chi$	$\leftrightarrow$	$(\neg A\varphi \to B^{[-\varphi]\alpha}_{[-\varphi]\overline{\psi},\neg\varphi}[-\varphi]\chi)$	
(ER6)	$[-\varphi]A\psi$	$\leftrightarrow$	$(\neg A\varphi \to A[-\varphi]\psi)$	

Table 3: Evidence Removal Recursion Axioms

Proof. We only do axiom $ER5$: Suppose that $\mathcal{M} = \langle W, E, V\rangle$ is an evidence model, $w \in W$ and $[\![\varphi]\!]_{\mathcal{M}} \neq W$. First of all, note that the key observation in the proof of Fact 3.14 extends to sets of evidence sets (cf. Definition 3.15). That is, for all worlds w, $\mathcal{X} \subseteq E^{-\varphi}(w)$ is compatible with $\overline{\psi}$ iff $\mathcal{X} \subseteq E(w)$ is compatible with $[-\varphi]\overline{\psi}, \neg\varphi$. Furthermore, for all states w, $\mathcal{X} \subseteq E^{-\varphi}(w)$ is a maximal α-f.i.p. iff $\mathcal{X} \subseteq E(w)$ is a maximal $[-\varphi]\alpha$-f.i.p. *compatible with* $\neg\varphi$.[27] Then,

$\mathcal{M}, w \models [-\varphi]B^{\alpha}_{\overline{\psi}}\chi$	iff	$\mathcal{M}^{-\varphi}, w \models B^{\alpha}_{\overline{\psi}}\chi$
	iff	for each maximal α-f.i.p. $\mathcal{X} \subseteq E^{-\varphi}(w)$ compatible with $\overline{\varphi}$, $\bigcap \mathcal{X}^{\alpha} \subseteq [\![\chi]\!]_{\mathcal{M}^{-\varphi}} = [\![[-\varphi]\chi]\!]_{\mathcal{M}}$
	iff	for each maximal $[-\varphi]\alpha$-f.i.p. $\mathcal{X} \subseteq E(w)$ compatible with $[-\varphi]\overline{\varphi}$ and $\neg\varphi$, $\bigcap \mathcal{X}^{[-\varphi]\alpha} \subseteq [\![[-\varphi]\chi]\!]_{\mathcal{M}}$
	iff	$\mathcal{M}, w \models B^{[-\varphi]\alpha}_{\overline{\psi},\neg\varphi}[-\varphi]\chi$.

QED

The above principles state the essence of evidence removal, as well as the beliefs one can still have after such an event. The additional insight is that removal essentially involves compatibility as well as implication between propositions – something that seems of independent interest.

Languages once more Again, this is a beginning rather than an end. For a start, extending the base language in this manner will have repercussions for our earlier analyses. Using the style of analysis presented here, it is possible to also find reduction axioms for the new evidence and belief operators under actions of evidence addition and public announcement. For example, for the compatible evidence operator $\Box_{\overline{\psi}}$ with $\overline{\psi} = (\psi_1, \ldots, \psi_n)$, we have the following validities:

- $[+\varphi]\Box_{\overline{\psi}}\chi \leftrightarrow [E\varphi \to (\Box_{[+\varphi]\overline{\psi}}[+\varphi]\chi \vee (\bigwedge_{i=1,\ldots,n} E(\varphi \wedge \psi_i) \wedge A(\varphi \to [+\varphi]\psi)))]$
- $[!\varphi]\Box_{\overline{\psi}}\chi \leftrightarrow (\varphi \to \Box^{\varphi}_{[!\varphi]\overline{\psi}}[!\varphi]\chi)$

In the interest of space, we do not include all combinations here. The important point is that the combinations do not lead to further extensions of the base language.

[27] Note that this last clause about be compatible with $\neg\varphi$ is crucial: it is not true that every $\mathcal{X} \subseteq E^{-\varphi}(w)$ that is a maximal α-f.i.p. corresponds to a maximal $[-\varphi]\alpha$-f.i.p. subset of $E(w)$.

Further, and more challenging open problems have to do with the 'action algebra' of combined modalities for our three basic actions on evidence so far. What happens when we compose them? We must leave this to another occasion.

3.4 Evidence modification

Our examples so far represent the major operations on evidence that we can see. Nevertheless, the space of potential operations on neighborhood models for evidence is much larger, even if we impose suitable conditions of bisimulation invariance similar to those found in process algebra (cf. Nicola (1987) and also the discussion in Hansen et al. (2009)). Instead of exploring this wide mathematical realm, we conclude with a different type of definition that might make sense. What we have done so far is adding or removing evidence, or perhaps combining it. But one could also *modify* the existing pieces of evidence. To see, the difference, here is a straightforward new operation, making some proposition φ more prominent:

Definition 3.18 (Evidence Upgrade) [28] Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model and $\varphi \in \mathcal{L}_1$. The model $\mathcal{M}^{\Uparrow\varphi} = \langle W^{\Uparrow\varphi}, E^{\Uparrow\varphi}, V^{\Uparrow\varphi}\rangle$ has $W^{\Uparrow\varphi} = W$, $V^{\Uparrow\varphi} = V$, and for all $w \in W$,

$$E^{\Uparrow\varphi}(w) = \{X \cup [\![\varphi]\!]_{\mathcal{M}} \mid X \in E(w)\} \cup [\![\varphi]\!]_{\mathcal{M}}. \qquad \triangleleft$$

This operation is stronger than simply adding $[\![\varphi]\!]_{\mathcal{M}}$ as evidence, since it also modifies each admissible evidence set. But it is still weaker than publicly announcing φ, since the agent retains the ability to consistently condition on $\neg\varphi$. The result of this operation is to make φ the most "important" piece of evidence, as exemplified in the following principles:

Fact 3.19 *The following recursion principles are valid:*

1. $[\Uparrow\varphi]\Box\psi \leftrightarrow (E\varphi \to A(\varphi \to [\Uparrow\varphi]\psi))$
2. $[\Uparrow\varphi]B\psi \leftrightarrow (E\varphi \to A(\varphi \to [\Uparrow\varphi]\psi))$

Proof. For the second principle, note that in $E^{\Uparrow\varphi}(w)$, there is only one maximal f.i.p. whose intersection must be $[\![\varphi]\!]_{\mathcal{M}}$. The proof of the first principle is left to the reader – to see how it goes, cf. the proof below of Fact 3.20. QED

As these laws show, $\Uparrow\varphi$ is a strong operation which gives a very special status to the incoming information φ and consequently blurs the distinction between evidence and belief. This suggests a weaker operation that modifies the evidence sets in favor of φ, but does not add explicit support for φ. Define $\mathcal{M}^{\Uparrow_w\varphi}$ as in Definition 3.18 except for setting $E^{\Uparrow_w\varphi}(w) = \{X \cup [\![\varphi]\!]_{\mathcal{M}} \mid X \in E(w)\}$. A simple modification to Principle 2 in the above fact gives us a valid principle for our evidence operator. However, the case of belief poses some problems. [29]

Fact 3.20 *The formula* $[\Uparrow_w\varphi]\Box\psi \leftrightarrow (\Box[\Uparrow_w\varphi]\psi \wedge A(\varphi \to [\Uparrow_w\varphi]\psi))$ *is valid.*

[28] This operation is somewhat similar to the "radical upgrade" in dynamic logics of belief change.

[29] The new complication is that, without adding φ to the evidence sets, intersections of maximal f.i.p. sets in the upgraded model may contain more than just φ states.

Proof. Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model with $w \in W$. Then,

$$\begin{array}{lll} \mathcal{M}, w \models [\Uparrow_w\varphi]\Box\psi & \text{iff} & \mathcal{M}^{\Uparrow\varphi}, w \models \Box\psi \\ & \text{iff} & \text{there is a } X \in E^{\Uparrow\varphi}(w) \text{ such that } X \subseteq [\![\psi]\!]_{\mathcal{M}^{-\varphi}} \\ & & (\text{note that } [\![\psi]\!]_{\mathcal{M}^{\Uparrow\varphi}} = [\![[\Uparrow\varphi]\psi]\!]_{\mathcal{M}}) \\ & \text{iff} & \text{there is } X' \in E(w) \text{ with } X' \cup [\![\varphi]\!]_{\mathcal{M}} = X \subseteq [\![[\Uparrow\varphi]\psi]\!]_{\mathcal{M}} \\ & \text{iff} & \text{there is } X' \in E(w) \text{ with } X' \subseteq [\![[\Uparrow\varphi]\psi]\!]_{\mathcal{M}} \text{ and } [\![\varphi]\!]_{\mathcal{M}} \subseteq [\![[\Uparrow\varphi]\psi]\!]_{\mathcal{M}} \\ & \text{iff} & \mathcal{M}, w \models \Box[\Uparrow\varphi]\psi \wedge A(\varphi \to [\Uparrow\varphi]\psi) \end{array}$$

QED

Other operations in the same vein would intersect the current evidence sets with $[\![\varphi]\!]_{\mathcal{M}}$, or rather subtracting the latter set from all evidence sets – as long as this can be done consistently. However, the latter operations seem close to the public update that we have already discussed before.

3.5 Evidence combination

We have now brought to light a rich repertoire of evidence-modifying actions. Still, the operations discussed above all exemplify "external evidence dynamics" responding to some outside source, where the agent reacts appropriately, either by incorporating φ or removing φ from consideration. But our neighborhood models also suggest internal operations that arise from pondering the evidence, without external triggers. We will discuss only one such internal operation in this paper, be it a basic one.

One natural operation available to an agent is to *combine* her evidence. Of course, as we have noted, an agent's evidence may be contradictory, so she can only combine evidence that is not inconsistent. For example, one round of combination in an evidence model $\mathcal{M} = \langle W, E, V\rangle$ results in the new evidence set:

$$E^{\#_1}(w) = E(w) \cup \{X \mid \text{there are } Y_1, Y_2 \in E(w) \text{ such that } \emptyset \neq X = Y_1 \cap Y_2\}$$

The corresponding operator $[\#_1]\varphi$ says that "φ is true after one round of evidence combination". This validates, for instance, the formula

$$(E(\varphi \wedge \psi) \wedge \Box\varphi \wedge \Box\psi) \to [\#_1]\Box(\varphi \wedge \psi).$$

Of course, further rounds of evidence combination may generate even more evidence sets. The limit of this internal process of introspection leads to the following notion:

Definition 3.21 (Evidence combination) Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model. The model $\mathcal{M}^{\#} = \langle W^{\#}, E^{\#}, V^{\#}\rangle$ has $W^{\#} = W$, $V^{\#} = V$ and for all $w \in W$, $E^{\#}(w)$ is the smallest set closed under (non-empty) intersection and containing $E(w)$. The corresponding modal operator is defined as $\mathcal{M}, w \models [\#]\varphi$ iff $\mathcal{M}^{\#}, w \models \varphi$. ◁

A complete analysis of this operation will be left for future work, since it poses some challenges to our recursive style so far. [30] Nevertheless, we can observe the following interesting facts:

Fact 3.22 *The following formulas are valid on the class of evidence models:*

[30]The problem may be that standard modal languages are too poor – but we suspend judgment here.

1. *$\Box[\#]\varphi \to [\#]\Box\varphi$ (combining evidence does not remove any of the original evidence[31])*

2. *$B[\#]\varphi \leftrightarrow [\#]B\varphi$ (beliefs are immune to evidence combination)*

3. *$B\varphi \to [\#]\Box\varphi$ (beliefs are explicitly supported after consistently combining evidence)*

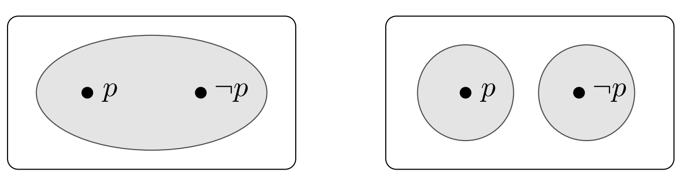
4. *For factual φ, $B\varphi \to \neg[\#]\Box\neg\varphi$ (if an agent believes φ then the agent cannot combine her evidence so that there is evidence for $\neg\varphi$)*

Proof. The proofs for the stated principles are easy but instructive. The proof that the first three are valid is left to the reader. For the fourth principle, note that $\Box\neg\varphi \to \neg B\varphi$ is valid. The proof is as follows: First of all, in any evidence model $\mathcal{M} = \langle W, E, V \rangle$, every piece of evidence in $X \in E(w)$ is contained in a maximal f.i.p. $\mathcal{X} \subseteq E(w)$ (models are finite, so simply find the maximal f.i.p. containing X which may be $\{X, W\}$). Suppose that $\Box\neg\varphi$ is true at a state w, then there is an $X \in E(w)$ such that $X \subseteq [\![\neg\varphi]\!]_{\mathcal{M}}$. Let $\mathcal{X}$ be the maximal f.i.p. containing X. Hence, $\bigcap \mathcal{X} \subseteq X \subseteq [\![\neg\varphi]\!]_{\mathcal{M}}$. Therefore, $B\varphi$ is not true at w. This shows that $\Box\neg\varphi \to \neg B\varphi$ is valid, as desired. We can then derive Principle 3 by noting the following series of implications:

$$B\varphi \to [\#]B\varphi \to [\#]\neg\Box\neg\varphi \to \neg[\#]\Box\neg\varphi$$

Here the first implication follows from the second principle applied to factual formulas φ (for which $\varphi \leftrightarrow [\#]\varphi$ is valid), the second implication follows from the fact that $B\varphi \to \neg\Box\neg\varphi$ is valid (as argued above) while $[\#]$ is a normal modal operator and the third implication follows from the fact that the evidence combination operation is functional. QED

We conclude with a few comments. First, recall that $B\varphi \to \Box\varphi$ is, in general, not valid (believing that φ is true need not imply that the agent has evidence for φ). The agent must do some work (i.e., consistently combine her evidence) to find explicit support for her beliefs. Our second observation is that the converses of principles 3 and 4 are *not* valid. This is surprising as one may expect that, in our setting, believing that φ is true *corresponds* to not having evidence for $\neg\varphi$ after consistent combining all available evidence (or perhaps having evidence for φ after consistently combining all available evidence). But the agent's evidence may not be able to make the "relevant distinctions". Consider the following evidence sets:



$\mathcal{E}_1$ $\mathcal{E}_2$

Combining evidence does not add any new pieces of evidence to either evidence set and the agent does not believe p in either model. However, in $\mathcal{E}_1$ the agent does not have evidence for $\neg p$ and in $\mathcal{E}_2$ the agent does have evidence for p. Thus, the principles $[\#]\Box\varphi \to B\varphi$ and $\neg[\#]\Box\neg\varphi \to B\varphi$ are not valid. The problem is that the evidence in $\mathcal{E}_1$ does not make enough distinctions while the

[31] That is, we assume in Definition 3.21 that for all states w, $E(w) \subseteq E^{\#}(w)$. This means that during the process of combination, the agent does not notice any inconsistencies that may be present in her evidential state. A more fine-grained analysis would include policies for removing such inconsistencies.

evidence in $\mathcal{E}_2$ makes too many distinctions. There may be interesting "separation properties" we can impose on the evidence sets that would find a balance between these two extremes, validating both principles, but we leave a full discussion for another time.

Internal evidence combination has no counterpart in the standard modal semantics of belief, as we will see in Section 4. There may be other actions of this internal kind, too, such as versions of *introspection* on one's evidence. In particular, our distinction between external and internal evidential acts is not hard and fast. Think of an agent noting that she has two contradictory pieces of evidence: this is an act of "seeing an inconsistency". This act does not license combination in the above sense, but it might well lead to *removal* of one of the two disjoint sets. In that sense, removal fits with a purely internal process of restoring consistency in one's evidence.

We conclude our dynamic analysis of operations on evidence by juxtaposing a few recursion axioms found above. Note the subtle differences in varieties of evidence management and resulting belief changes:

Public announcement:	$[!\varphi]B\psi$	$\leftrightarrow$	$(\varphi \to B^{\varphi}[!\varphi]\psi)$
Evidence addition:	$[+\varphi]B\psi$	$\leftrightarrow$	$(E\varphi \to (B^{+\varphi}[+\varphi]\psi \wedge B^{-\varphi}[+\varphi]\psi))$
Evidence removal:	$[-\varphi]B\psi$	$\leftrightarrow$	$(\neg A\varphi \to B_{\neg\varphi}[-\varphi]\psi)$
Evidence upgrade:	$[\Uparrow\varphi]B\psi$	$\leftrightarrow$	$(E\varphi \to A(\varphi \to [\Uparrow\varphi]\psi))$

Table 4: Recursion axioms describing belief change under different forms of evidence management

Summarizing, we have shown how neighborhood models support a rich dynamics of evidence transformations, and there are bound to be more than those we have studied. Along the way, we found that it was necessary to extend our basic static language of evidence and belief $\mathcal{L}_0$ with several new types of conditional evidence and beliefs. We conclude this section by taking stock of these various extensions:

$\Box\psi$:	"there is evidence for ψ"
$\Box^{\varphi}\psi$:	"there is evidence compatible with φ for ψ"
$\Box_{\overline{\gamma}}\psi$:	"there is evidence compatible with each of the γ_i for ψ"
$B\psi$:	"the agent believe χ"
$B^{\varphi}\psi$:	"the agent believe χ conditional on φ"
$B^{\varphi}_{\overline{\gamma}}\psi$:	"the agent believe χ conditional on φ assuming compatibility with each of the γ_i"
$B^{\varphi,\alpha}\psi$:	"the agent believe ψ, after having settled on α and conditional on φ"

Table 5: Different notions of conditional evidence and belief

We leave open a complete logical analysis in the full static language including all of these operators. Validities include interesting connections between varieties of conditional belief such as:

$$B^{\varphi}\psi \to B(\varphi \to \psi) \quad \text{and} \quad B(\varphi \to \psi) \to B^{\top,\varphi}\psi$$

4 Plausibility Models

In this section, we will contrast our neighborhood models with another general framework for belief change in the modal tradition. This will put our findings so far in a broader perspective, while connecting up to more standard approaches.

4.1 Plausibility orders

Originally used as a semantics for conditionals (cf. Lewis, 1973), *plausibility models* are wide-spread in modal logics of belief (van Benthem, 2004, 2011; Baltag and Smets, 2006a; Girard, 2008). The main idea is to endow epistemic ranges with further structure, viz. an ordering $w \preceq v$ of relative *plausibility* on worlds (usually uniform across epistemic equivalence classes): "(according to the agent) world v is at least as plausible as w". [32] Plausibility orders are typically assumed to be reflexive and transitive, and often also *connected*, making every two worlds comparable. We will look at both cases, though our general set-up will allow for incomparable worlds.

Example 4.1 Recall the simple illustration from the introduction (Example 2.1: $W = \{w, v\}$ with p true at w). Plausibility models over W can represent the following states of belief, with belief read as usual as *truth in all most plausible worlds*:

- $w \preceq v$ and $v \not\preceq w$: the agent believes that p is true
- $v \preceq w$ and $w \not\preceq v$: the agent believes that p is false
- $w \preceq v$ and $v \preceq w$: the agent suspends judgement on p
- $w \not\preceq v$ and $v \not\preceq w$: the agent cannot compare w and v (perhaps for lack of evidence)

We will see later on how to relate these belief states to evidential states.

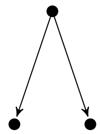
Here is the formal definition of our semantic structures:

Definition 4.2 (Plausibility model) A **plausibility model** is a tuple $\mathcal{M} = \langle W, \preceq, V \rangle$ where W is a finite nonempty set, $\preceq \subseteq W \times W$ is a reflexive and transitive ordering on W, and $V : \mathsf{At} \to \wp(W)$ is a valuation function. If $\preceq$ is also *connected* (for each $w, v \in W$, either $w \preceq v$ or $v \preceq w$) then we say $\mathcal{M}$ is a **connected plausibility model**. A pair $\mathcal{M}, w$ where w is a state is called a **pointed (connected) plausibility model**. ◁

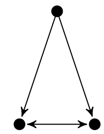
We conclude by discussing the interpretation of plausibility orders that are not connected. [33] When two worlds w and v cannot be compared in the ordering, the agent has either accepted contradictory evidence or lacks enough evidence to compare the two states. For example, consider the following two plausibility models:

[32] In conditional semantics, plausibility or 'similarity' orders are typically world-dependent.

[33] The recent article Swanson (2011) has an extensive discussion of incomparability in modeling conditionals.



Evidence coming from different sources



Evidence coming from the same source

We can interpret the model on the left concretely as representing a situation where the agent has two sources of evidence which are not fully compatible.

4.2 Language and logic

Plausibility models interpret a standard doxastic language. Let $\mathcal{L}_{\preceq}$ be the smallest set of formulas generated by the following language

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid B^{\varphi}\psi \mid [\preceq]\varphi \mid A\varphi$$

As usual, $B\varphi$ is defined as $B^{\top}\varphi$. To define truth, we need some notation. For $X \subseteq W$, let

$$Min_{\preceq}(X) = \{v \in W \mid v \preceq w \text{ for all } w \in X\ \}$$

Given a set X, $Min_{\preceq}(X)$ is the set of most plausible worlds in X (i.e., minimal elements of X according to the plausibility order)[34]. Truth of the boolean connectives and atomic propositions is explained as usual. We only consider the modal operators:

- $\mathcal{M}, w \models B^{\varphi}\psi$ iff $Min_{\preceq}([\![\varphi]\!]_{\mathcal{M}}) \subseteq [\![\psi]\!]_{\mathcal{M}}$
- $\mathcal{M}, w \models [\preceq]\varphi$ iff for all $v \in W$, if $v \preceq w$ then $\mathcal{M}, v \models \varphi$
- $\mathcal{M}, w \models A\varphi$ iff for all $v \in W$, $\mathcal{M}, v \models \varphi$.

With $B\varphi$ defined as $B^{\top}\varphi$, we get the usual notion of belief as truth in all minimal worlds. We can think of this as follows. As is well-known, any pre-order forms a partial order of 'clusters', maximal subsets where the relation is universal. A finite pre-order has one or more final clusters, not having any proper successors. (If the order is connected, there is only one final cluster.) Belief means truth in all final clusters.

The logic of this system is basically the minimal conditional logic over pre-orders that we have encountered before. Instead of pursuing it, we make some further comments on definability. Plausibility orders are binary relations supporting a standard modal language. Indeed, as has been noted by Boutilier (1992), on finite models, belief and conditional belief are already definable in the language with A and $[\preceq]$ only:

Fact 4.3 *Belief and conditional belief can be explicitly defined as follows:*

- $B\varphi \ := \ A\langle\preceq\rangle[\preceq]\varphi$

[34] As before in this paper, we assume that the set of worlds is finite, so this minimal set always exist. One needs a well-foundedness condition to guarantee this when the set of state is infinite.

- $B^{\varphi}\psi \ := \ A(\varphi \to \langle\preceq\rangle(\varphi \wedge [\preceq](\varphi \to \psi)))$

The proof is straightforward since the combination $\langle\preceq\rangle[\preceq]$ refers to final clusters, while the second equivalence is a standard relativization of the former to the antecedent predicate.

While the plausibility modality may look like a technical device, Baltag and Smets (2006a) interpret $[\preceq]\varphi$ as "a *safe belief* in φ". Following Stalnaker (1996), they show that this amounts to the beliefs the agent retains under all new true information about the actual world. [35] Indeed, we see this simple modal language over plausibility models as a natural limit of expressive power. In Section 4.4, we will show how it passes the additional test of being able to define all new evidence modalities that were brought to light by our dynamic analysis of Section 3.

4.3 Dynamics on plausibility models

Plausibility models support a dynamics of informational action through *model change*. In this subsection, we quickly survey some relevant themes from the literature, and then make a few new points relevant to linking up with our evidence setting.

Belief change under hard information One paradigmatic action has been discussed before, in Section 3.1. "Hard information" reduces current models to definable submodels:

Definition 4.4 (Public announcement - plausibility models) Let $\mathcal{M} = \langle W, \preceq, V\rangle$ be a plausibility model. The model $\mathcal{M}^{!\varphi} = \langle W^{!\varphi}, \preceq^{!\varphi}, V^{!\varphi}\rangle$ is defined as follows: $W = [\![\varphi]\!]_{\mathcal{M}}$, for all $p \in \mathsf{At}$, $V^{!\varphi}(p) = V(p) \cap W^{!\varphi}$ and $\preceq^{!\varphi} = \preceq \cap (W^{!\varphi} \times W^{!\varphi})$. ◁

Dynamic logics exist that describe belief change under such events of new hard information, i.e., the logical laws governing $[!\varphi]B\psi$. [36] Indeed, the crucial recursion axioms for belief are the same as those for evidence models in Section 3.1:

$$[!\varphi]B\psi \leftrightarrow (\varphi \to B^{\varphi}[!\varphi]\psi)$$

$$[!\varphi]B_i^{\psi}\chi \leftrightarrow (\varphi \to B_i^{\varphi \wedge [!\varphi]\psi}[!\varphi]\chi)$$

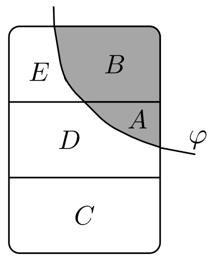
Public announcement assumes that agents treat the source of the new information as *infallible*. But in many scenarios, agents *trust* the source of the information only up to a point. This calls for *softer* announcements, that can also be brought under our framework. In this section, we will only make some basic remarks on how this works: (van Benthem, 2011, Chapter 7) and (Baltag and Smets, 2009) have much more extensive discussion.

Belief change under soft information How to incorporate evidence that φ is true into a doxastic or epistemic-doxastic model $\mathcal{M}$? Soft announcements of a formula φ do not eliminate worlds, but rather *modify the plausibility ordering* that represents an agent's current information state. The goal is to *rearrange* all states in such a way that φ is believed, and perhaps other desiderata are met. There are many "policies" for doing this (cf. Rott, 2006), but here, we only mention two, that have been widely discussed in the literature on belief revision.

[35] For the same notion in the computational literature on agency (cf. Shoham and Leyton-Brown, 2009).

[36] One might think this is taken care of by conditional belief $B_i^{\varphi}\psi$, and indeed it is, when ψ is a *factual formula* not containing any modal operators. But in general, the laws are different (van Benthem, 2011, Chapter 7).

Example 4.5 The following picture illustrates soft update as plausibility change:



One policy that has been extensively studied is *radical upgrade* where *all* φ worlds are moved ahead of all other worlds, while keeping the order inside these two zones the same. In the above example, the radical upgrade by φ would result in the ordering $A \prec B \prec C \prec D \prec E$.

More formally, the model transformation here is *relation change*:

Definition 4.6 (Radical Upgrade.) Given an epistemic-doxastic model $\mathcal{M} = \langle W, \preceq, V\rangle$ and a formula φ, the *radical* upgrade of $\mathcal{M}$ with φ is the model $\mathcal{M}^{\Uparrow\varphi} = \langle W^{\Uparrow\varphi}, \preceq^{\Uparrow\varphi}, V^{\Uparrow\varphi}\rangle$ with $W^{\Uparrow\varphi} = W$, $V^{\Uparrow\varphi} = V$, where $\preceq^{\Uparrow\varphi}$ is defined as follows:

1. for all $x \in [\![\varphi]\!]_{\mathcal{M}}$ and $y \in [\![\neg\varphi]\!]_{\mathcal{M}}$, set $x \prec^{\Uparrow\varphi} y$,
2. for all $x, y \in [\![\varphi]\!]_{\mathcal{M}}$, set $x \preceq^{\Uparrow\varphi} y$ iff $x \preceq y$, and
3. for all $x, y \in [\![\neg\varphi]\!]_{\mathcal{M}}$, set $x \preceq^{\Uparrow\varphi} y$ iff $x \preceq y$. $\triangleleft$

A logical analysis of this type of information change uses modalities $[\Uparrow\varphi]\psi$ meaning "after radical upgrade of φ, ψ is true", interpreted as follows:

$$\mathcal{M}, w \models [\Uparrow\varphi]\psi \text{ iff } \mathcal{M}^{\Uparrow\varphi}, w \models \psi.$$

Here is the crucial recursion axiom for belief change under soft information: [37]

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

This is the key principle driving the complete logic of belief revision from (van Benthem, 2004), and its proof shows quite concretely how revision policies as plausibility transformations give agents not just new beliefs, but also new conditional beliefs.

But radical upgrade is not the only way for an agent to accept incoming information. Perhaps the most ubiquitous policy is *conservative upgrade*, which lets the agent only tentatively accept the incoming information φ by making the best φ the new minimal set and keeping the old plausibility ordering the same on all other worlds. In the above picture a conservative upgrade with φ results in the new ordering $A \prec C \prec D \prec B \cup E$. The general idea for dealing with such policies is this:

"Information update is model transformation"[38]

[37] There are some technical issues here with the treatment of possible epistemic structure, and also, with possible non-connectedness of the plausibility order. We forego these in this survey section.

[38] Baltag and Smets (2006b); van Benthem and Liu (2004) provide general logical formats for belief revision policies.

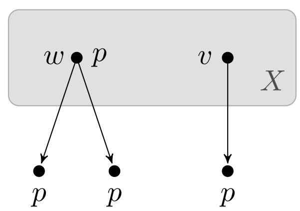
Digression: pre-orders and safety Many intuitions underlying dynamic logics of belief change were formed with connected plausibility orders in mind. However, as we shall soon see, the plausibility orders that are most congenial to our evidence models need to allow for arbitrary *pre-orders* where worlds can also be incomparable. In such cases, not every insight from the study of relational belief revision can be adopted as it stands. Here is one example.

One conspicuous feature of current logics over plausibility models is their *extending* the repertoire of basic doxastic attitudes with new ones suggested by the dynamic perspective. A famous example is the above notion of *safe belief*. This has been characterized dynamically as those beliefs,[39] which an agent keeps having when given true information about the actual world:

$$\mathcal{M}, w \models B^s\varphi \text{ iff for each } X \subseteq W \text{ with } w \in X, \text{ we have } Min_{\preceq}(X) \subseteq [\![\varphi]\!]_{\mathcal{M}}$$

But this neat equivalence fails on pre-orders:

Example 4.7 The following model $\mathcal{M}$ shows that $B^s\varphi$ is not equivalent to $[\preceq]\varphi$ – and also that $B^s\varphi$ is not equivalent to $B\varphi$:



First, note that both $[\preceq]p$ and Bp are true at w. However, since X is a set with $w \in X$ and $Min_{\preceq}(X) = \{w, v\}$, we have $\mathcal{M}, w \not\models B^s\varphi$.

The correct result for general plausibility models needs a wider-ranging modality:

Fact 4.8 *On arbitrary pre-orders $\mathcal{M}$, safe belief in its dynamic sense $B^s\varphi$ holds at a world w iff φ is true at all worlds $v \not\prec w$ that are not strictly less plausible than w.*

We omit the simple proof, but the reader can easily see the point of the new broader modality by looking at a model consisting of two disjoint worlds.

This concludes our brief tour of relational plausibility models. For further notions and results (including the role of knowledge entangled with belief), we refer to the cited literature.

4.4 From plausibility models to evidence models

Plausibility pre-orders are a very general format for describing beliefs. Let us know make a connection with our evidence models.

[39]One needs to restrict attention to *factual propositions* for this result.

Transformation Let $\mathcal{M} = \langle W, \preceq, V \rangle$ be a plausibility model. Our main intuition is very simple:

the appropriate evidence sets are the downward $\preceq$-closed sets of worlds.

To be more precise, we fix some notation:

- Given a $X \subseteq W$, let $X{\downarrow_{\preceq}} = \{v \in W \mid \exists x \in X \text{ and } v \preceq x\}$
  (we write $X{\downarrow}$ when it is clear which plausibility ordering is being used).

- A set $X \subseteq W$ is $\preceq$-closed if $X{\downarrow_{\preceq}} \subseteq X$.

Here is the formal definition for the above simple idea:

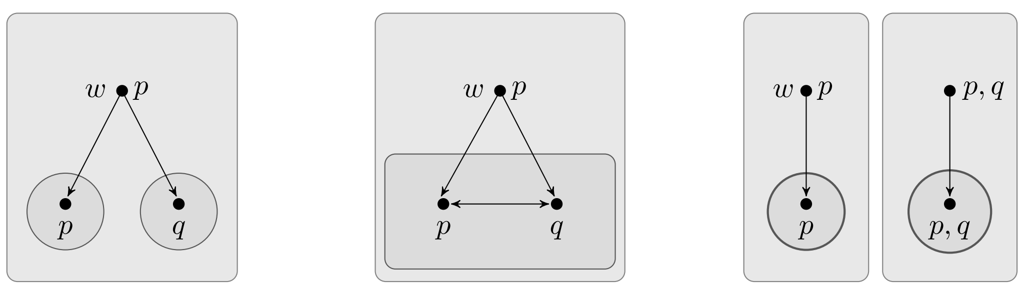
Definition 4.9 (Plausibility-Based Evidence Model) Let $\mathcal{M} = \langle W, \preceq, V \rangle$ be a plausibility model. The **evidence model generated from** $\mathcal{M}$ is[40] $EV(\mathcal{M}) = \langle W, \mathcal{E}_{\preceq}, V \rangle$ with $\mathcal{E}_{\preceq}$ as follows:

$$\mathcal{E}_{\preceq} = \{X \mid \emptyset \neq X \text{ is } \preceq\text{-closed } \} \qquad \triangleleft$$

Comparison with general evidence models Given a plausibility model $\mathcal{M}$, the evidence model generated by the plausibility order of $\mathcal{M}$ clearly satisfies the basic properties that we have required in Section 2: the sets are non-empty, and the whole universe is among them. [41] But more can be said. Here is an extremely simple, but important observation:

Fact 4.10 *The evidence sets of generated models $EV(\mathcal{M})$ are closed under intersections.*

Example 4.11 The following three plausibility models – with their induced evidence sets drawn in gray – highlight three key situations that can occur:



$EV(\mathcal{M}_1), w \models \neg B(p \wedge q)$ $\qquad$ $EV(\mathcal{M}_2), w \models \neg B(p \wedge q)$ $\qquad$ $EV(\mathcal{M}_3), w \models Bp \wedge \neg Bq$

$EV(\mathcal{M}_1), w \models \Box p \wedge \Box q$ $\qquad$ $EV(\mathcal{M}_2), w \models \neg\Box p \wedge \neg\Box q$ $\qquad$ $EV(\mathcal{M}_3), w \models \Box(p \wedge q)$

But not every evidence model, not even a uniform one, comes from a plausibility model.

Example 4.12 Let $\mathcal{M}$ be an evidence model with $W = \{w, v, x\}$, a constant evidence function with range $\mathcal{E} = \{\{w, v\}, \{v, x\}\}$ and the valuation function defined by $V(p) = \{w, v\}$ and $V(q) = \{v, x\}$. Note that we have $\mathcal{M}, w \models B(p \wedge q)$ (the agent believes p and q) but $\mathcal{M}, w \models \Box p \wedge \Box q \wedge \neg\Box(p \wedge q)$ (even though there is evidence for p and evidence for q, there is no evidence for $p \wedge q$).

[40] Here the set of worlds and valuation function remains the same as in the plausibility model $\mathcal{M}$.
[41] But our evidence sets need not be upward closed: we need upward relational closures of extensions.

Here is the intuitive explanation. Plausibility models represent a situation where the agent has "combined" all of her evidence, as reflected in this technical property:

$$\text{If } X, Y \in \mathcal{E}_{\preceq} \text{ and } X \cap Y \neq \emptyset \text{ then } X \cap Y \in \mathcal{E}_{\preceq}.$$

In other words, the combining operation # of Section 3.4 has already happened.

Translation Having connected the two kinds of structure in one direction, we now want to compare their languages. Let $\mathcal{L}$ be the full static language of evidence models (so $\mathcal{L} = \mathcal{L}_0 \cup \mathcal{L}_1 \cup \mathcal{L}_2 \cup \mathcal{L}_3$) and $\mathcal{L}_{\preceq}$ be the language of plausibility models with modalities $[\preceq]$ and A. The following key shows how this basic modal inclusion language can deal with all operators introduced in the course of Section 3.

Definition 4.13 (P-translation) The translation $(\cdot)^P : \mathcal{L} \to \mathcal{L}_{\preceq}$ is defined as follows:

- $p^P = p$, $(\neg\varphi)^P = \neg\varphi^P$, $(\varphi \wedge \psi)^P = \varphi^P \wedge \psi^P$, $(A\varphi)^P = A\varphi^P$,
- $(\Box\varphi)^P = E[\preceq]\varphi^P$,
- $(\Box^{\varphi}\psi)^P = E\langle\preceq\rangle(\varphi^P \wedge [\preceq](\varphi^P \to \psi^P))$,
- $(\Box^{\varphi}_{\vec{\gamma}}\psi)^P = E(\bigwedge_i \langle\preceq\rangle\gamma_i^P \wedge \langle\preceq\rangle(\varphi^P \wedge [\preceq](\varphi^P \to \psi^P)))$,
- $(B^{\varphi}\psi)^P = A(\varphi^P \to (\langle\preceq\rangle\varphi^P \wedge [\preceq](\varphi^P \to \psi^P)))$,
- $(B^{\varphi,\alpha}\psi)^P = A(([\preceq]\alpha^P \wedge [\preceq]\langle\preceq\rangle\varphi^P) \to \langle\preceq\rangle(\varphi^P \wedge [\preceq](\varphi^P \to \psi^P)))$, and
- $(B^{\varphi}_{\vec{\gamma}}\psi)^P = A((\varphi^P \wedge \bigwedge_i\langle\preceq\rangle\gamma_i^P) \to (\langle\preceq\rangle(\varphi^P \wedge \bigwedge_i\langle\preceq\rangle\gamma_i^P) \wedge [\preceq]((\varphi^P \wedge \bigwedge_i\langle\preceq\rangle\gamma_i^P) \to \psi^P)))$. ◁

Here is how the connection works:

Lemma 4.14 *Let $\mathcal{M} = \langle W, \preceq, V\rangle$ be a plausibility model. For any $\varphi \in \mathcal{L}_1$ and world $w \in W$,*

$$\mathcal{M}, w \models \varphi^P \text{ iff } EV(\mathcal{M}), w \models \varphi$$

The proof follows by noting that the intersection of maximally overlapping families of evidence sets are exactly the final clusters in an plausibility models. This shows how our earlier basic modal language of evidence can be reduced to that of plausibility on these special evidence models. [42]

4.5 From evidence models to plausibility models

The transformation We have seen how to turn plausibility models into evidence models. Going in the opposite direction, we start with a family of evidence sets, and need to induce a natural ordering. Here we borrow a ubiquitous idea, occurring in point-set topology, but also in theories of relation merge (cf. Andreka et al., 2002; Liu, 2008): the so-called *specialization (pre)-order*"

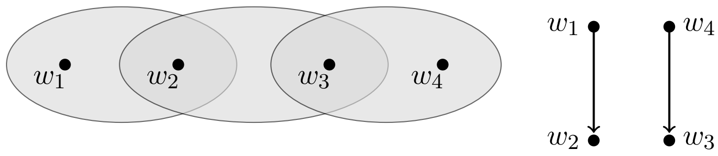
[42]Similar issues make sense for the next subsection, but we will not repeat them there.

Definition 4.15 (Plausibility based evidence model) Suppose that $\mathcal{M} = \langle W, \mathcal{E}, V\rangle$ is an evidence model (with constant evidence function E whose range is $\mathcal{E}$). The plausibility model generated by $\mathcal{M}$ is the structure $ORD(\mathcal{M}) = \langle W, \preceq_E, V\rangle$ where $\preceq_E$ is an ordering on W defined as follows:

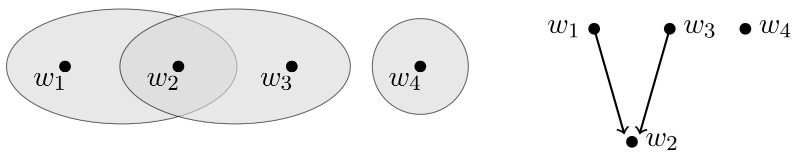
$$w \preceq_{\mathcal{E}} v \text{ iff } \forall X \in \mathcal{E},\ v \in X \text{ implies } w \in X$$

It is easy to see that $\preceq_E$ is reflexive and transitive, so $ORD(\mathcal{M})$ is indeed a plausibility model.[43] ◁

To make this definition a bit more concrete, here is a simple illustration.



$\mathcal{M}$ $ORD(\mathcal{M})$



$\mathcal{M}$ $ORD(\mathcal{M})$

Now one obvious question arises: how are our two representations related?

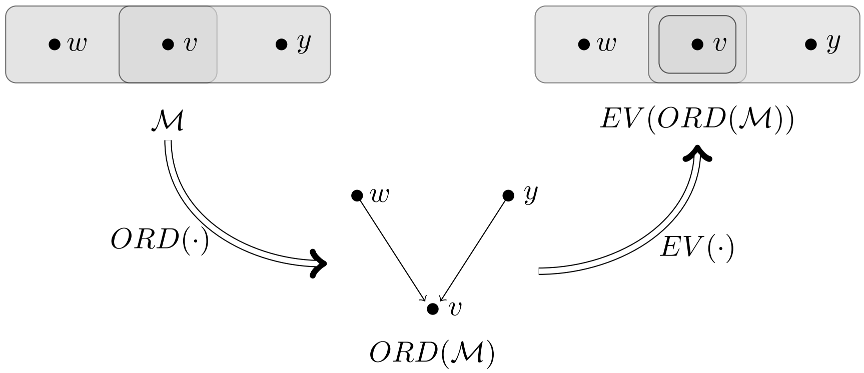
Fact 4.16 *(i) For all models plausibility models $\mathcal{M}$, $ORD(EV((\mathcal{M})) = \mathcal{M}$, (ii) The identity $EV(ORD(\mathcal{M})) = \mathcal{M}$ does not hold for all evidence models $\mathcal{M}$. (iii) For all evidence models $\mathcal{M}$, $EV(ORD(\mathcal{M})) = \mathcal{M}^{\#}$, where # is the combination operation of Section 3.21.*

Proof. (i) The first identity follows by a simple calculation. The main underlying point is the observation that, on any pre-order,

$$w \preceq v \text{ iff for all } \preceq\text{-closed sets } U\text{: } v \in U \text{ implies } w \in U.$$

(ii) Here is a counter-example to the second identity. Consider an evidence model $\mathcal{M}$ with two evidence sets as pictured below on the left. Notice that representing this model as a plausibility model and then turning that model back into an evidence model adds an evidence set:

[43]The intuition is that if there is no evidence that distinguishes (or "separates", in a topological sense) w and v, then w and v should be considered equally plausible. Furthermore, w should be considered "more plausible" if there is more evidence in support of w. At one extreme here, all of the agent's sources may provide just the trivial evidence set W, and all worlds will be equally plausible. At the other extreme is an overabundance of information with all singletons as evidence sets. In this case, all states are incomparable. Interestingly, according to our definition of belief, in both cases the agent will believe a proposition φ iff it is true at *all* worlds.



(iii) The preceding example explains what is going on: composing the two transformations adds all intersections of evidence sets, adding which does not change the induced world order. QED

Translations We now ask again, given the above transformation turning evidence models into plausibility models, can we also show that this transformation is correct by reducing the modal language of plausibility models to our basic modal language of evidence and belief? Our first observation is that the interpretation of (conditional) beliefs is indeed invariant under the above transformation. Let $\mathcal{L}_B$ be the language with only (conditional) belief modalities. Note that $\mathcal{L}_B$ is interpreted over both evidence models and plausibility models. The following is easy to show:

Fact 4.17 *Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model. For each formula $\varphi \in \mathcal{L}_B$ and world $w \in W$,*

$$\mathcal{M}, w \models \varphi \textit{ iff } ORD(\mathcal{M}), w \models \varphi$$

This remains true if we add the universal modality. But, recall that over plausibility models (conditional) beliefs were *defined* in a more basic language $\mathcal{L}_{\preceq}$ including only A and the base modality $[\preceq]$ (safe belief). One can show that this modality does not have a counterpart in our basic modal language for evidence modals. Instead, it suggests a natural *extension* to our language beyond what we have seem earlier:

Reliable evidence and belief Suppose that $\mathcal{M} = \langle W, E, V\rangle$ is an evidence model. Let $E^C(w) = \{X \in E(w) \mid w \in X\}$ denote the set of evidence that is "correct" or reliable at state w. Unlike the agent's full evidential state at w, the set of reliable evidence at w can always be consistently combined, suggesting a new modality $\Box^C\varphi$ meaning "the agent's reliable evidence entails φ" or, for lack of a better term, "φ is reliably believed". The formal definition is:

$$\mathcal{M}, w \models \Box^C\varphi \text{ iff for all } v \in \bigcap E^C(w),\ \mathcal{M}, v \models \varphi$$

This is indeed the counterpart to the basic plausibility modality. Let $\mathcal{L}^C$ be the extension of $\mathcal{L}_1$ with the operator $\Box^C$ and define the translation $(\cdot)^N : \mathcal{L}_{\preceq} \to \mathcal{L}^C$ as follows: $(\cdot)^N$ commutes with boolean connectives and the universal modality and $([\preceq]\varphi)^N = \Box^C\varphi^N$. Then,

Fact 4.18 *Let $\mathcal{M}$ be a uniform evidence model. Then, for all formulas $\varphi \in \mathcal{L}_{\preceq}$ and worlds w,*

$$\mathcal{M}, w \models \varphi^N \textit{ iff } ORD(\mathcal{M}), w \models \varphi$$

The easy proof is left to the reader. The key observation is that $\bigcap E^C(w) = \{v \mid v \preceq_{\mathcal{E}} w\}$.[44]

There is one technical issue we need to address here, which has interesting consequences. Recall that we transform *uniform* evidence models into plausibility models. But note that the set of reliable evidence differs from state to state. That is, even if the evidence function E is assumed to be constant, the *reliable* evidence function E^C will not be constant. This means that, unlike plain belief, even in uniform evidence models, reliable belief does not satisfy introspection properties. So, what type of doxastic attitude is $\Box^C$? We do not have the space here for an extensive discussion, but here are a few comments.

First, notice that $\Box^C$ validates the truth axiom ($\Box^C\varphi \to \varphi$), so believing something based on reliable evidence implies truth. Of course, it is only the modeler, from a third person perspective, who can actually determine what the agent believes based only on reliable evidence. Since the agent does not have access to the actual world, she herself cannot determine which evidence is reliable and which is not. Second, the restriction to "truthful" evidence suggests that reliable belief might be *safe belief* on evidence models. However, plausibility models of the form $ORD(\mathcal{M})$ are typically not connected, and we showed that on such models safe belief quantifies over all worlds not strictly less plausible than the current world (cf. Fact 4.8). In particular, if φ is safely believed at a world w in a non-connected plausibility model, then φ must be true at all worlds that are incomparable with w. Now, in a plausibility model $ORD(\mathcal{M})$, there are two reasons why a state v may be incomparable to the actual state w: either there is reliable evidence not containing v, or there is evidence containing v but not w. This suggests yet another extension to the language $\mathcal{L}_1$.

Unreliable evidence Let $\mathcal{M} = \langle W, E, V\rangle$ be an evidence model and define $E^U(w) = \{X \in E(w) \mid w \notin X\}$. This is the set of *unreliable* (or incorrect) evidence at state w. The corresponding modality is $\Box^U$: "φ follows from the unreliable evidence at w". Of course, the agent cannot necessarily consistently combine this evidence, but in the formal definition we can take the *union* of these sets:

$$\mathcal{M}, w \models \Box^U\varphi \text{ iff for all } v \in \bigcup E^U(w),\ \mathcal{M}, v \models \varphi$$

It is the conjunction of these two operators that corresponds to safe belief on evidence models:

Fact 4.19 *Let $\mathcal{M}$ be a uniform evidence model, then for all factual formulas*[45] *φ:*

$$\mathcal{M}, w \models \Box^C\varphi \wedge \Box^U\varphi \textit{ iff } ORD(\mathcal{M}), w \models B^s\varphi$$

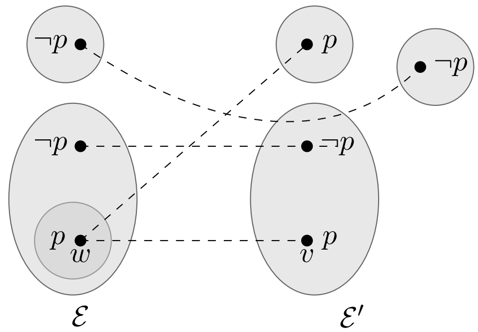
Finally, we show that these new operators are not definable in our basic language $\mathcal{L}_1$.

Fact 4.20 *The operators $\Box^C$ and $\Box^U$ are not definable in evidence belief language $\mathcal{L}_1$.*

Proof. Consider two uniform evidence models $\mathcal{M}$ and $\mathcal{M}'$ whose evidence sets are pictured below:

[44] For a proof, suppose that $v \in \bigcap \mathcal{E}^C(w) = \bigcap\{X \in E(w) \mid w \in X\}$. Then for all $X \in E(w)$, if $v \in X$ then $w \in X$. But this means $v \preceq_E w$. The converse is straightforward.

[45] This fact extends to more complex formulas, but we leave the details for a later analysis.



The dashed line is a monotonic bisimulation, so $\mathcal{M}, w \leftrightarroweq \mathcal{M}', v$. Furthermore, it is easy to see that the agent has the same beliefs in both models: recall that in uniform models all higher-order beliefs are reducible to first-order beliefs (Proposition 2.7), and the agent only has trivial first-order beliefs in both models. So, $\mathcal{M}, w$ and $\mathcal{M}', v$ satisfy the same formulas of $\mathcal{L}_0$. However, we have[46] $\mathcal{M}, w \models \Box^C p \wedge \Box^U \neg p$ while $\mathcal{M}', v \models \neg\Box^C p \wedge \neg\Box^U \neg p$. QED

This concludes our comparison of relational and neighborhood semantics for belief and evidence. We have clarified their relationship as one of generalization, and we have also shown how some interesting new questions arise at the interface.

5 Further Directions

Having explored our basic framework, we conclude by pointing out a number of interesting avenues for future research. Some are more technical, some increase coverage: [47]

Complete static logic We have found quite a few new evidence-based modalities of conditional belief. What is the complete logic of this system? This is a new question of standard axiomatization, that can be appreciated also outside of our dynamic perspective. One reason for its complexity may be that we are mixing a language of neighbourhood-based modalities with normal operators of belief that can also be interpreted in a matching relational semantics.

What is the right notion of bisimulation? Here is a more systematic perspective on our study of evidence models. Designing logical languages invites matching up with notions of structural invariance for evidence models. We have already seen that standard bisimulation for neighborhood models is appropriate for modal logics with only evidence operators. But Fact 2.10 showed that this does not extend to the modal language with *belief* referring to intersections of maximally consistent families of evidence sets. And we introduced even stronger modal evidence languages in our discussion of dynamics in Section 3 as well as the transformation of evidence models into plausibility models in Section 4.5. What would be well-motivated stronger notions of bisimulation, respecting more evidence structure?

[46] In general, it need not be the case that the agent reliably believes φ iff the agent unreliably believes $\neg\varphi$. We do not discuss the logic of these operators here, but the reader might note that the set of reliable evidence at a state is a *filter* while the set of unreliable evidence is an *ideal*.

[47] Of course, there are some obvious technical generalizations that need to be made, say to *infinite models* and to *DEL*-style *product update* mechanisms for rich input.

Reliable evidence. In our discussion of plausibility models versus evidence models, we found notions of "reliable" evidence, based only on sets containing the actual world. What is the complete logic of this operator? This observation also suggests a broader study of types of belief based on reliable evidence, in line with current trends in epistemology. Eventually, we also want explicit modeling of sources of evidence and what agents know or believe about their reliability.

Social notions. We have seen that interesting evidence structure arises in the single agent case, the focus of this paper. But, *multi-agent situations* are also natural: e.g., argumentation is social confrontation of evidence, which may result in new group attitudes among participants. This raises a number of interesting issues of its own. The most pressing is to find natural notions of *group evidence* and belief. Here one soon finds that the additional evidence structure takes us beyond the usual notions of group beliefs or group knowledge that are found in the epistemic logic literature based on relational models.

To illustrate this, suppose there are two agents i and j and a multi-agent uniform evidence model $\mathcal{M} = \langle W, \mathcal{E}_i, \mathcal{E}_j, V\rangle$. We can then ask what evidence the group i, j has. One option here is mere throwing together into one "unprocessed" new evidence set:

$$\mathcal{M}, w \models \Box^{\{i,j\}}\varphi \text{ iff there is a } X \in \mathcal{E}_i \cup \mathcal{E}_j \text{ such that } X \subseteq [\![\varphi]\!]_{\mathcal{M}}$$

At another extreme, the group might only take into account only evidence that is *shared* between the two agents:

$$\mathcal{M}, w \models \Box_{\{i,j\}}\varphi \text{ iff there is a } X \in \mathcal{E}_i \cap \mathcal{E}_j \text{ such that } X \subseteq [\![\varphi]\!]_{\mathcal{M}}$$

For a rough analogy on relational models, the latter form of group evidence is more like common knowledge, while the former is more related to distributed knowledge. But our richer neighborhood models also allow for further distinctions. In particular, the agents can also "pool" their evidence creating a *new* evidential states by *combining* their evidence:

$$\mathcal{E}_i \sqcap \mathcal{E}_j = \{Y \mid \emptyset \neq Y = X \cap X' \text{ with } X \in \mathcal{E}_i \text{ and } X' \in \mathcal{E}_j\ \}$$

We can then define group evidence and belief modalities for this evidential state as we did in Section 2.2. For instance,

$$\mathcal{M}, w \models [i \sqcap j]\varphi \text{ iff there exists } X \in \mathcal{E}_i \sqcap \mathcal{E}_j \text{ with } X \subseteq [\![\varphi]\!]_{\mathcal{M}}$$

How, precisely, are these notions related to the familiar common/distributed belief found in the epistemic logic literature? What other natural notions of group evidence can we find?

Priority structures. The evidence dynamics in this paper treats evidence sets on a par. As a consequence, removal may seem arbitrary and non-deterministic, since there is nothing in the structure of the evidence itself which directs the process. A next reasonable step would be to model *levels of reliability* of evidence. One natural format for this are the "priority graphs" of Andreka et al. (2002), which have already been used extensively in dynamic-epistemic logic (Liu, 2008; Girard, 2008). These graphs provide much richer input to evidence management, and can break stalemates between conflicting pieces of evidence. It should be possible to extend the above framework to one with ordered evidence sets – and conversely, then, our logics may help provide something that has been missing so far: modal logics working directly on priority graphs.

Evidence structure once more. As noted at the start of this paper, "evidence" is a notion that has many different aspects. Our proposal has been set-theoretic and semantic, while there are many other treatments of evidence for a proposition φ, in particular, in terms of proofs for φ, or using the balance of probability for φ versus $\neg\varphi$. What we find particularly pressing is a junction with more syntactic approaches making evidence something coded that can be operated on in terms of inference and computation. As we have noted before, in seeking this balance between syntax and semantics (cf. Velazquez-Quesada, 2011), we may have to merge our approach with current extensions of dynamic-epistemic logic with actions on syntactic code.

Related work. Our final desideratum is of a different kind. As we said right in the introduction, there are several related lines in the literature, such as the seminal work by Dempster (1967) and Shafer (1976) on evidence, the probabilistic logics of evidence (cf. Halpern and Pucella, 2006), or the "topologic" of Moss and Parikh (1992). And one can add more, such as the "priority graphs" inducing preference orders in (Liu, 2011), or the "belief base" account of belief revision in (see Hansson, 1999, and references therein). We intend to clarify these connections in future work.

6 Conclusion

We have made a pilot proposal for using neighborhood models as fine-grained evidence structures that allow for richer representation of information than current relational models of belief. We have shown how, in particular, these structures support a rich dynamics of evidence change that goes beyond current dynamic logics of belief revision. A number of completeness theorems identified the key dynamic equations governing this process, while also suggesting new static languages of evidence and belief. We also clarified the connections between plausibility models and evidence models, leading to further questions across frameworks. Finally, we discussed some of the interesting new issues that lie ahead now, such as finding logics with priority structure and group evidence that exploit the more finely-grained neighborhood setting.

References

Ågotnes, T. and N. Alechina (2007). The dynamics of syntactic knowledge. *Journal of Logic and Computation 17*(1), 83–116.

Ågotnes, T. and H. van Ditmarsch (2010). What will they say?—public announcement games. *Synthese*, 1–29.

Andreka, H., M. Ryan, and P. Y. Schobbens (2002). Operators and laws for combining preference relations. *Journal of Logic and Computation 12*(1), 13 – 53.

Baltag, A. and S. Smets (2006a). Conditional doxastic models: A qualitative approach to dynamic belief revision. In G. Mints and R. de Queiroz (Eds.), *Proceedings of WOLLIC 2006, Electronic Notes in Theoretical Computer Science*, Volume 165, pp. 5 – 21.

Baltag, A. and S. Smets (2006b). Dynamic belief revision over multi-agent plausiblity models. In G. Bonanno, W. van der Hoek, and M. Wooldridge (Eds.), *Proceedings of the 7th Conference on Logic and the Foundations of Game and Decision (LOFT 2006)*, pp. 11 – 24.

Baltag, A. and S. Smets (2009). ESSLLI 2009 course: Dynamic logics for interactive belief revision. Slides available at `http://alexandru.tiddlyspot.com/#%5B%5BESSLLI09%20COURSE%5D%5D`.

Boutilier, C. (1992). *Conditional Logics for Default Reasoning and Belief Revision.* Ph. D. thesis, University of Toronto.

Burgess, J. (1981). Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic 22*, 76 – 84.

Chellas, B. (1980). *Modal Logic: An Introduction.* Cambridge University Press.

Demey, L. (2010a). Agreeing to disagree in probabilistic dynamic epistemic logic. Master's thesis, ILLC University of Amsterdam (LDC 2010-14).

Demey, L. (2010b). Neighborhood semantics for public announcement logic. Manuscript.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics 38*(2), 325 – 339.

Fagin, R. and J. Halpern (1988). Belief, awareness and limited reasoning. *Artificial Intelligence 34*, 39 – 76.

Gärdenfors, P. (1988). *Knowledge in Flux: Modeling the Dynamics of Epistemic States.* Bradford Books, MIT Press.

Gerbrandy, J. (1999). *Bisimulations on Planet Kripke.* Ph. D. thesis, Institute for Logic, Language and Computation (DS-1999-01).

Girard, P. (2008). *Modal Logic for Belief and Preference Change.* Ph. D. thesis, ILLC University of Amsterdam Dissertation Series DS-2008-04.

Halpern, J. and R. Pucella (2006). A logic for reasoning about evidence. *Journal of AI Research 26*, 1 – 34.

Hansen, H. H. (2003). Monotonic modal logic. Master's thesis, Universiteit van Amsterdam (ILLC technical report: PP-2003-24).

Hansen, H. H., C. Kupke, and E. Pacuit (2009). Neighbourhood structures: Bisimilarity and basic model theory. *Logical Methods in Computer Science 5*(2), 1 – 38.

Hansson, S. O. (1999). *A Textbook of Belief Dynamics. Theory Change and Database Updating.* Kluwer.

Kratzer, A. (1977). What *must* and *can* must and can mean. *Linguistics and Philosophy 1*, 337 – 355.

Leitgeb, H. and K. Segerberg (2007). Dynamic doxastic logic: why, how and where to? *Synthese 155*(2), pgs. 167 – 190.

Lewis, D. (1973). *Counterfactuals.* Oxford: Blackwell Publishers.

Liu, F. (2008). *Changing for the Better: Preference Dynamics and Agent Diversity.* Ph. D. thesis, Institute for Logic, Language and Computation (DS-2008-02).

Liu, F. (2011). A two-level perspective on preference. *Journal of Philosophical Logic*. (To Appear).

Meyer, J.-J. and W. van der Hoek (2004). *Epistemic Logic for AI and Computer Science.* Cambridge University Presse.

Moss, L. and R. Parikh (1992). Topological reasoning and the logic of knowledge. In Y. Moses (Ed.), *Proceedings of TARK IV*. Morgan Kaufmann.

Nicola, R. D. (1987). Extensional equivalences for transition systems. *Acta Informatica 24*, 211–237.

Pacuit, E. (2007). Neighborhood semantics for modal logic: An introduction. ESSLLI 2007 course notes (`ai.stanford.edu/~epacuit/classes/esslli/nbhdesslli.pdf`).

Parikh, R. (1985). The logic of games and its applications. In M. Karpinski and J. v. Leeuwen (Eds.), *Topics in the Theory of Computation*, Annals of Discrete Mathematics 24. Elsevier.

Pauly, M. (2001). *Logic for Social Software.* Ph. D. thesis, ILLC University of Amsterdam Dissertation Series DS 2001-10.

Pauly, M. (2002). A modal logic for coalitional power in games. *Journal of Logic and Computation 12*(1), 149–166.

Plaza, J. (2007). Logics of public communications. *Synthese: Knowledge, Rationality, and Action 158*(2), 165 – 179.

Rott, H. (2001). *Change, Choice and Inference: A Study in Belief Revision and Nonmonotonic Reasoning.* Oxford University Press.

Rott, H. (2006). Shifting priorities: Simple representations for 27 itereated theory change operators. In H. Lagerlund, S. Lindström, and R. Sliwinski (Eds.), *Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg*, Volume 53 of *Uppsala Philosophical Studies*, pp. 359 – 384.

Segerberg, K. (1971). *An Essay in Classical Modal Logic.* Filosofisska Stuier, Uppsala Universitet.

Segerberg, K. (1995). Belief revision fromt he point of view of doxastic logic. *Journal of the IGPL 3*(4), 535–553.

Shafer, G. (1976). *A Mathematical Theory of Evidence.* Princeton University Press.

Shoham, Y. and K. Leyton-Brown (2009). *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations.* Cambridge University Press.

Stalnaker, R. (1996). Knowledge, belief and counterfactual reasoning in games. *Economics and Philosophy 12*(02), 133 – 163.

Su, K., A. Sattar, G. Governatori, and Q. Chen (2005). A computationally grounded logic of knowledge, belief and certainty. In *Proceedings of the fourth international joint conference on autonomous agents and multiagent systems*, AAMAS '05, pp. 149 – 156.

Swanson, E. (2011). On the treatment of incomparability in ordering semantics and premise semantics. *Journal of Philosophical Logic*. (To Appear, DOI: 10.1007/s10992-010-9157-z).

van Benthem, J. (2004). Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics 14*(2), 129 – 155.

van Benthem, J. (2005). A note on modeling theories. In R. Festa, A. Aliseda, and J. Peijnenburg (Eds.), *Poznan Studies in the Philosophy of the Sciences and Humanities: Confirmation, Empiral Progress and Truth Approximaiton. Essays in Debate with Theo Kuipers*, Volume 17, pp. 403 – 419.

van Benthem, J. (2008). Merging observation and access in dynamic logic. *Studies in Logic 1*(1), 1 – 17.

van Benthem, J. (2011). *Logical Dynamics of Information Flow.* Cambridge University Press.

van Benthem, J. and F. Liu (2004). Diversity of agents in games. *Phiosophia Scientiae 8*(2).

van Benthem, J. and S. Minica (2009). Toward a dynamic logic of questions. In X. He, J. F. Horty, and E. Pacuit (Eds.), *Logic, Rationality, and Interaction*, Volume 5834 of *Lecture Notes in Computer Science*, pp. 27–41. Springer.

van Ditmarsch, H., W. van der Hoek, and B. Kooi (2007). *Dynamic Epistemic Logic.* Synthese Library. Springer.

Vardi, M. (1986). On epistemic logic and logical omniscience. In J. Halpern (Ed.), *Theoretical Aspects of Reasoning about Knowledge: Proceedings of the 1986 Conference*, pp. 293–305. Morgan Kaufmann.

Velazquez-Quesada, F. R. (2009). Inference and update. *Synthese (Knowledge, Rationality & Action) 169*(2), 283 – 300.

Velazquez-Quesada, F. R. (2011). *Small steps in dynamics of information.* Ph. D. thesis, ILLC University of Amsterdam Dissertation Series DS 2011-02.

Veltman, F. (1976). Prejudices, presuppositions and the theory of conditionals. In J. Groenendijk and M. Stokhof (Eds.), *Amsterdam Papers in Formal Grammar*, Volume 1, pp. 248 – 281.

Veltman, F. (1985). *Logics for Conditionals.* Ph. D. thesis, Universiteit van Amsterdam.

Zvesper, J. (2010). *Playing with Information.* Ph. D. thesis, ILLC University of Amsterdam Dissertation Series DS-2010-02.