

Chapter 13

Logic and Probabilistic Update

Lorenz Demey and Barteld Kooi

Abstract This chapter surveys recent work on probabilistic extensions of epistemic and dynamic-epistemic logics (the latter include the basic system of public announcement logic as well as the full product update logic). It emphasizes the importance of higher-order information as a distinguishing feature of these logics. This becomes particularly clear in the dynamic setting: although there exists a clear relationship between usual Bayesian conditionalization and public announcement, the probabilistic effects of the latter are in general more difficult to describe, because of the subtleties involved in higher-order information. Finally, the chapter discusses some applications of probabilistic dynamic epistemic logic, such as the Lockean thesis in formal epistemology and Aumann's agreement theorem in game theory.

13.1 Introduction

Epistemic logic and probability theory both provide formal accounts of information. Epistemic logic takes a *qualitative* perspective on information, and works with a modal operator K . Formulas such as $K\varphi$ can be interpreted as 'the agent knows that φ ', 'the agent believes that φ ', or, more generally speaking, ' φ follows from the agent's current information'. Probability theory, on the other hand, takes a *quantitative* perspective on information, and works with numerical probability functions P . Formulas such as $P(\varphi) = k$ can be interpreted as 'the probability of φ is k '. In the

L. Demey (✉)

Center for Logic and Analytical Philosophy, KU Leuven—University of Leuven, Kardinaal Mercierplein 2, 3000 Leuven, Belgium
e-mail: Lorenz.Demey@hiw.kuleuven.be

B. Kooi

Faculty of Philosophy, University of Groningen, Oude Boteringestraat 52,
9712 GL Groningen, The Netherlands
e-mail: B.P.Kooi@rug.nl

present context, probabilities will usually be interpreted subjectively, and can thus be taken to represent the agent's degrees of belief or credences.

With respect to one and the same formula φ , epistemic logic is able to distinguish between three epistemic attitudes: knowing its truth ($K\varphi$), knowing its falsity ($K\neg\varphi$), and being ignorant about its truth value ($\neg K\varphi \wedge \neg K\neg\varphi$). Probability theory, however, distinguishes infinitely many epistemic attitudes with respect to φ , viz. assigning it probability k ($P(\varphi) = k$), for every $k \in [0, 1]$. In this sense probability theory can be said to provide a much more *fine-grained* perspective on information.

While epistemic logic thus is a coarser account of information, it certainly has a wider scope. From its very origins in Hintikka's [34], epistemic logic has not only been concerned with knowledge about 'the world', but also with knowledge about knowledge, i.e. with *higher-order information*. Typical discussions focus on principles such as positive introspection ($K\varphi \rightarrow KK\varphi$). On the other hand, probability theory rarely talks about principles involving higher-order probabilities, such as $P(\varphi) = 1 \rightarrow P(P(\varphi) = 1) = 1$.¹ This issue becomes even more pressing in multi-agent scenarios. Natural examples might involve an agent a not having any information about a proposition φ , while being certain that another agent, b , does have this information. In epistemic logic this is naturally formalized as

$$\neg K_a\varphi \wedge \neg K_a\neg\varphi \wedge K_a(K_b\varphi \vee K_b\neg\varphi).$$

A formalization in probability theory might look as follows:

$$P_a(\varphi) = 0.5 \wedge P_a(P_b(\varphi) = 1 \vee P_b(\varphi) = 0) = 1.$$

However, because this statement makes use of 'nested' probabilities, it is rarely used in standard treatments of probability theory.

An additional theme is that of dynamics, i.e. *information change*. The agents' information is not eternally the same; rather, it should be changed in the light of new incoming information. Probability theory typically uses Bayesian updating to represent information change (but other, more complicated update mechanisms are available as well). Dynamic epistemic logic interprets new information as changing the epistemic model, and uses the new, updated model to represent the agents' updated information states. Once again, the main difference is that dynamic epistemic logic takes (changes in) higher-order information into account, whereas probability theory does not.

For all these reasons, the project of *probabilistic epistemic logic* seems very interesting. Such systems inherit the fine-grained perspective on information from probability theory, and the representation of higher-order information from epistemic

¹ A notable exception is 'Miller's principle', which states that $P_1(\varphi | P_2(\varphi) = b) = b$. The probability functions P_1 and P_2 can have various interpretations, such as the probabilities of two agents, subjective probability (credence) and objective probability (chance), or the probabilities of one agent at different moments in time—in the last two cases, the principle is also called the 'principal principle' or the 'principle of reflection', respectively. This principle has been widely discussed in Bayesian epistemology and philosophy of science [29, 32, 38, 40, 41].

logic. Their *dynamic* versions provide a unified perspective on changes in first- and higher-order information. In other words, they can be thought of as incorporating the complementary perspectives of (dynamic) epistemic logic and probability theory, thus yielding richer and more detailed accounts of information and information flow.

The remainder of this chapter is organized as follows. Section 13.2 introduces the static framework of probabilistic epistemic logic, and discusses its intuitive interpretation and technical features. Section 13.3 focuses on a rather straightforward type of dynamics, namely public announcements. It describes a probabilistic version of the well-known system of public announcement logic, and compares public announcement and Bayesian conditionalization. In Sect. 13.4 a more general update mechanism is introduced. This is a probabilistic version of the ‘product update’ mechanism in dynamic epistemic logic. Section 13.5, finally, indicates some applications and potential avenues of further research for the systems discussed in this chapter.

13.2 Probabilistic Epistemic Logic

In this section we introduce the static framework of probabilistic epistemic logic, which will be ‘dynamified’ in Sects. 13.3 and 13.4. Section 13.2.1 discusses the models on which the logic is interpreted. Section 13.2.2 defines the formal language and its semantics. Finally, Sect. 13.2.3 provides a complete axiomatization.

13.2.1 Probabilistic Kripke Models

Consider a finite set I of agents, and a countably infinite set $Prop$ of proposition letters. Throughout this chapter, these sets will be kept fixed, so they will often be left implicit.

Definition 13.1 A *probabilistic Kripke frame* is a tuple $\mathbb{F} = \langle W, R_i, \mu_i \rangle_{i \in I}$, where W is a non-empty finite set of states, $R_i \subseteq W \times W$ is agent i ’s epistemic accessibility relation, and $\mu_i : W \rightarrow (W \rightarrow [0, 1])$ assigns to each state $w \in W$ a partial function $\mu_i(w) : W \rightarrow [0, 1]$, such that

$$\sum_{v \in \text{dom}(\mu_i(w))} \mu_i(w)(v) = 1.$$

Definition 13.2 A *probabilistic Kripke model* is a tuple $\mathbb{M} = \langle \mathbb{F}, V \rangle$, where \mathbb{F} is a probabilistic Kripke frame (with set of states W), and $V : Prop \rightarrow \wp(W)$ is a valuation.

Note that in principle, no conditions are imposed on the agents’ epistemic accessibility relations. However, as is usually done in the literature on (probabilistic) dynamic epistemic logic, we will henceforth assume these relations to be

equivalence relations (so that the corresponding knowledge operators satisfy the principles of the modal logic S5).

The function $\mu_i(w)$ represents agent i 's probabilities (i.e. degrees of belief) at state w . For example, $\mu_i(w)(v) = k$ means that at state w , agent i assigns probability k to state v being the actual state. From a mathematical perspective, this is not the most general approach: one can also define a *probability space* $\mathbb{P}_{i,w}$ for each agent i and state w , and let $\mu_i(w)$ assign probabilities to sets in a σ -algebra on $\mathbb{P}_{i,w}$, rather than to individual states. In this way one can easily drop the requirement that frames and models have finitely many states. This approach is taken in [28] for static probabilistic epistemic logic, and extended to dynamic settings in [47]. However, because all the characteristic features of probabilistic (dynamic) epistemic logic already arise in the simpler approach, in this chapter we will stick to this simpler approach, and take $\mu_i(w)$ to assign probabilities to individual states. These functions are additively extended from individual states to sets of states, by putting (for any set $X \subseteq \text{dom}(\mu_i(w))$):

$$\mu_i(w)(X) := \sum_{x \in X} \mu_i(w)(x).$$

A consequence of our simple approach is that all sets $X \subseteq \text{dom}(\mu_i(w))$ have a definite probability $\mu_i(w)(X)$, whereas in the more general approach, sets X not belonging to the σ -algebra on $\mathbb{P}_{i,w}$ are not assigned any definite probability at all. A similar distinction can be made at the level of individual states. Because $\mu_i(w)$ is a partial function, states $v \in W - \text{dom}(\mu_i(w))$ are not assigned any definite probability at all. An even simpler approach involves putting $\mu_i(w)(v) = 0$, rather than leaving it undefined. In this way, the function $\mu_i(w)$ is total after all. From a mathematical perspective, these two approaches are equivalent. From an informal perspective, however, there is a clear difference: $\mu_i(w)(v) = 0$ means that agent i is certain (at state w) that v is not the actual state, whereas $\mu_i(w)(v)$ being undefined means that agent i has no opinion whatsoever (at state w) about v being the actual state. Again, because all the characteristic features of probabilistic (dynamic) epistemic logic already arise without this intuitive distinction, we will opt for the even simpler approach, and henceforth assume that all probability functions are total.

To summarize: the approach adopted in this chapter is the simplest one possible, in the sense that definite probabilities are assigned to 'everything': (i) to all *sets* (there is no σ -algebra to rule out some sets from having a definite probability), and (ii) to all *states* (the probability functions $\mu_i(w)$ are total on their domain W , so no states are ruled out from having a definite probability).

We finish this subsection by mentioning two typical properties of probabilistic Kripke frames.² In the next subsection we will show that these properties correspond to natural principles about the interaction between knowledge and probability.

Definition 13.3 Consider a probabilistic Kripke frame \mathbb{F} and an agent $i \in I$. Then \mathbb{F} is said to be *i -consistent* iff for all states w, v : if $(w, v) \notin R_i$ then $\mu_i(w)(v) = 0$.

² See [33] for a further discussion of these and other properties, and their correspondence to knowledge/probability interaction principles.

Furthermore, \mathbb{F} is said to be *i-uniform* iff for all states w, v : if $(w, v) \in R_i$ then $\mu_i(w) = \mu_i(v)$.

13.2.2 Language and Semantics

The language \mathcal{L} of (static) probabilistic epistemic logic is defined by means of the following Backus-Naur form:

$$\varphi ::= p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid K_i\varphi \mid a_1 P_i(\varphi) + \cdots + a_n P_i(\varphi) \geq b$$

—where $p \in Prop$, $i \in I$, $1 \leq n < \omega$, and $a_1, \dots, a_n, b \in \mathbb{Q}$. We only allow rational numbers as values for a_1, \dots, a_n, b in order to keep the language countable. As usual, $K_i\varphi$ means that agent i knows that φ , or, more generally, that φ follows from agent i 's information. Its dual is defined as $\hat{K}_i\varphi := \neg K_i\neg\varphi$, and means that φ is consistent with agent i 's information.

Formulas of the form $a_1 P_i(\varphi_1) + \cdots + a_n P_i(\varphi_n) \geq b$ are called *i-probability formulas*.³ Note that mixed agent indices are not allowed; for example, $P_a(p) + P_b(q) \geq b$ is *not* a well-formed formula. Intuitively, $P_i(\varphi) \geq b$ means that agent i assigns probability at least b to φ . We allow for linear combinations in *i-probability* formulas, because this additional expressivity is useful when looking for a complete axiomatization [28], and because it allows us to express comparative judgments such as ‘agent i considers φ to be at least twice as probable as ψ ’: $P_i(\varphi) \geq 2P_i(\psi)$. This last formula is actually an abbreviation for $P_i(\varphi) - 2P_i(\psi) \geq 0$. In general, we introduce the following abbreviations:

$$\begin{aligned} \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b & \quad \text{for} \quad a_1 P_i(\varphi_1) + \cdots + a_n P_i(\varphi_n) \geq b, \\ a_1 P_i(\varphi_1) \geq a_2 P_i(\varphi_2) & \quad \text{for} \quad a_1 P_i(\varphi_1) + (-a_2) P_i(\varphi_2) \geq 0, \\ \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \leq b & \quad \text{for} \quad \sum_{\ell=1}^n (-a_\ell) P_i(\varphi_\ell) \geq -b, \\ \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) < b & \quad \text{for} \quad \neg(\sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b), \\ \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) > b & \quad \text{for} \quad \neg(\sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \leq b), \\ \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) = b & \quad \text{for} \quad \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b \wedge \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \leq b. \end{aligned}$$

Note that because of its recursive definition, the language \mathcal{L} can express the agents’ higher-order information of any sort: higher-order knowledge (for example $K_a K_b \varphi$), but also higher-order probabilities (for example $P_a(P_b(\varphi) \geq 0.5) = 1$), and higher-order information that mixes knowledge and probabilities (for example, $K_a(P_b(\varphi) \geq 0.5)$ and $P_a(K_b \varphi) = 1$).

³ The agents’ probabilities are thus explicitly represented in the logic’s object language \mathcal{L} . Other proposals provide a probabilistic semantics for an object language that is itself fully classical (i.e. that does not explicitly represent probabilities). See [26] for a recent overview of the various ways of combining logic and probability.

The formal semantics for \mathcal{L} is defined as follows. Consider an arbitrary probabilistic Kripke model \mathbb{M} (with set of states W) and a state $w \in W$. We will often abbreviate $\llbracket \varphi \rrbracket^{\mathbb{M}} := \{v \in W \mid \mathbb{M}, v \models \varphi\}$. Then:

$$\begin{array}{ll}
 \mathbb{M}, w \models p & \text{iff } w \in V(p), \\
 \mathbb{M}, w \models \neg\varphi & \text{iff } \mathbb{M}, w \not\models \varphi, \\
 \mathbb{M}, w \models \varphi \wedge \psi & \text{iff } \mathbb{M}, w \models \varphi \text{ and } \mathbb{M}, w \models \psi, \\
 \mathbb{M}, w \models K_i\varphi & \text{iff for all } v \in W: \text{if } (w, v) \in R_i \text{ then } \mathbb{M}, v \models \varphi, \\
 \mathbb{M}, w \models \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b & \text{iff } \sum_{\ell=1}^n a_\ell \mu_i(w)(\llbracket \varphi_\ell \rrbracket^{\mathbb{M}}) \geq b.
 \end{array}$$

Furthermore, we also define:

- $\mathbb{M} \models \varphi$ iff $\mathbb{M}, w \models \varphi$ for all $w \in W$,
- $\mathbb{F} \models \varphi$ iff $\langle \mathbb{F}, V \rangle \models \varphi$ for all valuations V on the frame \mathbb{F} ,
- $\models \varphi$ iff $\mathbb{F} \models \varphi$ for all frames \mathbb{F} .

As promised, we will now provide correspondence results for the frame properties defined at the end of the previous subsection:

Lemma 13.1 *Consider a probabilistic Kripke frame \mathbb{F} . Then:*

1. \mathbb{F} is *i-consistent* iff $\mathbb{F} \models K_i p \rightarrow P_i(p) = 1$,
2. \mathbb{F} is *i-uniform* iff $\mathbb{F} \models (\varphi \rightarrow K_i\varphi) \wedge (\neg\varphi \rightarrow K_i\neg\varphi)$ for all *i-prob. formulas* φ .

From a technical perspective, this lemma indicates how the notion of *frame correspondence* from modal logic [8, 9, 20] can be extended into the probabilistic realm. From an intuitive perspective, this lemma sheds some new light on the various interactions between epistemic and probabilistic information. Probabilistic epistemic logic distinguishes between epistemic impossibility ($(w, v) \notin R_i$) and probabilistic impossibility ($\mu_i(w)(v) = 0$). For example, when a fair coin is tossed, an infinite series of tails is probabilistically impossible, but epistemically possible [37, p. 384]. Item 1 of Lemma 13.1 establishes a connection between the principle that knowledge implies certainty, and the property of consistency (epistemic impossibility entails probabilistic impossibility). Similarly, item 2 establishes a connection between the principle that agents know their own probabilistic setup, and the property of uniformity (the impossibility of epistemic uncertainty about probabilities).

13.2.3 Proof System

Probabilistic epistemic logic can be axiomatized in a highly modular fashion. An overview is given in Fig. 13.1. The propositional and epistemic components shouldn't need any further comments. The probabilistic component is a straightforward translation into the formal language \mathcal{L} of the well-known Kolmogorov axioms of probability; it ensures that the formal symbol $P_i(\cdot)$ behaves like a real probability function. Finally, the linear inequalities component is mainly a technical tool to

1. propositional component
 - all propositional tautologies and the modus ponens rule
2. epistemic component
 - the S5 axioms and rules for the K_i -operators
3. probabilistic component
 - $P_i(\varphi) \geq 0$
 - $P_i(\top) = 1$
 - $P_i(\varphi \wedge \psi) + P_i(\varphi \wedge \neg\psi) = P_i(\varphi)$
 - if $\vdash \varphi \leftrightarrow \psi$ then $\vdash P_i(\varphi) = P_i(\psi)$
4. linear inequalities component
 - $\sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b \leftrightarrow \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) + 0P_i(\varphi_{n+1}) \geq b$
 - $\sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b \leftrightarrow \sum_{\ell=1}^n a_{p(\ell)} P_i(\varphi_{p(\ell)}) \geq b$
for any permutation p of $1, \dots, n$
 - $\sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b \wedge \sum_{\ell=1}^n a'_\ell P_i(\varphi_\ell) \geq b' \rightarrow$
 $\sum_{\ell=1}^n (a_\ell + a'_\ell) P_i(\varphi_\ell) \geq b + b'$
 - $\sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b \leftrightarrow \sum_{\ell=1}^n da_\ell P_i(\varphi_\ell) \geq db$ (for any $d > 0$)
 - $\sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b \vee \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \leq b$
 - $\sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) \geq b \rightarrow \sum_{\ell=1}^n a_\ell P_i(\varphi_\ell) > b'$ (for any $b' < b$)

Fig. 13.1 Componentwise axiomatization of probabilistic epistemic logic

ensure that the logic is strong enough to capture the behavior of linear inequalities of probabilities.

Using standard techniques the following theorem can be proved [28]:

Theorem 13.1 *Probabilistic epistemic logic, as axiomatized in Fig. 13.1, is sound and complete with respect to the class of probabilistic Kripke frames.*

The notion of completeness used in this theorem is *weak* completeness ($\vdash \varphi$ iff $\models \varphi$), rather than *strong* completeness ($\Gamma \vdash \varphi$ iff $\Gamma \models \varphi$). These two notions do not coincide in probabilistic epistemic logic, because this logic is not *compact*; for example, every finite subset of the set $\{P_i(p) > 0\} \cup \{P_i(p) \leq k \mid k > 0\}$ is satisfiable, but the entire set is not.

13.3 Probabilistic Public Announcement Logic

In this section we discuss a first ‘dynamification’ of probabilistic epistemic logic, by introducing public announcements into the logic. Section 13.3.1 discusses updated probabilistic Kripke models, and introduces a public announcement operator into

the formal language to talk about these models. Section 13.3.2 provides a complete axiomatization, and Sect. 13.3.3 focuses on the role of higher-order information in public announcement dynamics.

13.3.1 Semantics

Public announcements form one of the simplest types of epistemic dynamics. They concern the truthful and public announcement of some piece of information φ by an external source. That the announcement is *truthful* means that the announced information φ has to be true; that it is *public* means that all agents $i \in I$ learn about it simultaneously and commonly. Finally, the announcement's source is called 'external' because it is not one of the agents $i \in I$ (and will thus not be explicitly represented in the formal language).

Public announcement logic [27, 31, 44] represents these announcements as updates that change Kripke models, and introduces a dynamic public announcement operator into the formal language to describe these updated models. This strategy can straightforwardly be extended into the probabilistic realm.

Syntactically, we add a dynamic operator $[\cdot]$ to the static language \mathcal{L} , thus obtaining the new language $\mathcal{L}^!$. The formula $[\varphi]\psi$ means that after any truthful public announcement of φ , it will be the case that ψ . Its dual is defined as $\langle \varphi \rangle \psi := \neg[\varphi]\neg\psi$, and means that φ can truthfully and publicly be announced, and afterwards ψ will be the case. These formulas thus allow us to express 'now' (i.e. *before* any dynamics has taken place) what will be the case 'later' (*after* the dynamics has taken place). These formulas are interpreted on a probabilistic Kripke model \mathbb{M} and state w as follows:

$$\begin{aligned} \mathbb{M}, w \models [\varphi]\psi & \quad \text{iff} \quad \text{if } \mathbb{M}, w \models \varphi \text{ then } \mathbb{M}|\varphi, w \models \psi, \\ \mathbb{M}, w \models \langle \varphi \rangle \psi & \quad \text{iff} \quad \mathbb{M}, w \models \varphi \text{ and } \mathbb{M}|\varphi, w \models \psi. \end{aligned}$$

Note that these clauses not only use the model \mathbb{M} , but also the updated model $\mathbb{M}|\varphi$. The model \mathbb{M} represents the agents' information *before* the public announcement of φ ; the model $\mathbb{M}|\varphi$ represents their information *after* the public announcement of φ ; hence the public announcement of φ *itself* is represented by the update mechanism $\mathbb{M} \mapsto \mathbb{M}|\varphi$, which is formally defined as follows:

Definition 13.4 Consider a probabilistic Kripke model $\mathbb{M} = \langle W, R_i, \mu_i, V \rangle_{i \in I}$, a state $w \in W$, and a formula $\varphi \in \mathcal{L}^!$ such that $\mathbb{M}, w \models \varphi$. Then the *updated probabilistic Kripke model* $\mathbb{M}|\varphi := \langle W^\varphi, R_i^\varphi, \mu_i^\varphi, V^\varphi \rangle_{i \in I}$ is defined as follows:

- $W^\varphi := W$,
- $R_i^\varphi := R_i \cap (W \times \llbracket \varphi \rrbracket^{\mathbb{M}})$ (for every agent $i \in I$),
- $\mu_i^\varphi : W^\varphi \rightarrow (W^\varphi \rightarrow [0, 1])$ is defined (for every agent $i \in I$) by

$$\mu_i^\varphi(v)(u) := \begin{cases} \frac{\mu_i(v)(\{u\} \cap \llbracket \varphi \rrbracket^{\mathbb{M}})}{\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}})} & \text{if } \mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}}) > 0 \\ \mu_i(v)(u) & \text{if } \mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}}) = 0, \end{cases}$$

- $V^\varphi := V$.

The main effect of the public announcement of φ in a model \mathbb{M} is that all links to $\neg\varphi$ -states are deleted; hence these states are no longer accessible for any of the agents. This procedure is standard; we will therefore focus on the probabilistic components μ_i^φ .

First of all, it should be noted that the case distinction in the definition of $\mu_i^\varphi(v)(u)$ is made for strictly technical reasons, viz. to ensure that there are no ‘dangerous’ divisions by 0. In all examples and applications, we will be using the ‘interesting’ case $\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}}) > 0$. Still, for general theoretical reasons, *something* has to be said about the case $\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}}) = 0$. Leaving $\mu_i^\varphi(v)(u)$ undefined would lead to truth value gaps in the logic, and thus greatly increase the difficulty of finding a complete axiomatization. The approach taken in this chapter is to define $\mu_i^\varphi(v)(u)$ simply as $\mu_i(v)(u)$ in case $\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}}) = 0$ —so the public announcement of φ has *no effect* whatsoever on $\mu_i(v)$. The intuitive idea behind this definition is that an agent i simply *ignores* new information if she previously assigned probability 0 to it. Technically speaking, this definition will yield a relatively simple axiomatization.

One can easily check that if \mathbb{M} is a probabilistic Kripke model, then $\mathbb{M}|\varphi$ is a probabilistic Kripke model as well. We focus on $\mu^\varphi(v)$ (for some arbitrary state $v \in W^\varphi$). If $\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}}) = 0$, then $\mu_i^\varphi(v)$ is $\mu_i(v)$, which is a probability function on $W = W^\varphi$. If $\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}}) > 0$, then for any $u \in W^\varphi$,

$$\mu_i^\varphi(v)(u) = \frac{\mu_i(v)(\{u\} \cap \llbracket \varphi \rrbracket^{\mathbb{M}})}{\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}})},$$

which is positive because $\mu_i(v)(\{u\} \cap \llbracket \varphi \rrbracket^{\mathbb{M}})$ is positive, and at most 1, because $\mu_i(v)(\{u\} \cap \llbracket \varphi \rrbracket^{\mathbb{M}}) \leq \mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}})$ —and hence $\mu_i^\varphi(v)(u) \in [0, 1]$. Furthermore,

$$\sum_{u \in W^\varphi} \mu_i^\varphi(v)(u) = \sum_{u \in W} \frac{\mu_i(v)(\{u\} \cap \llbracket \varphi \rrbracket^{\mathbb{M}})}{\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}})} = \sum_{\mathbb{M}, u \models \varphi} \frac{\mu_i(v)(u)}{\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}})} = 1.$$

It should be noted that the definition of $\mu_i^\varphi(v)$ —in the interesting case when $\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}}) > 0$ —can also be expressed in terms of conditional probabilities:

$$\mu_i^\varphi(v)(u) = \frac{\mu_i(v)(\{u\} \cap \llbracket \varphi \rrbracket^{\mathbb{M}})}{\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}})} = \mu_i(v)(u \mid \llbracket \varphi \rrbracket^{\mathbb{M}}).$$

In general, for any $X \subseteq W^\varphi$ we have:

$$\mu_i^\varphi(v)(X) = \frac{\mu_i(v)(X \cap \llbracket \varphi \rrbracket^{\mathbb{M}})}{\mu_i(v)(\llbracket \varphi \rrbracket^{\mathbb{M}})} = \mu_i(v)(X \mid \llbracket \varphi \rrbracket^{\mathbb{M}}).$$

In other words, after the public announcement of a formula φ , the agents calculate their new, updated probabilities by means of *Bayesian conditionalization* on the information provided by the announced formula φ . This connection between public announcements and Bayesian conditionalization will be explored more thoroughly in Sect. 13.3.3.

Example 13.1 We finish this subsection by discussing a simple example. Consider the following scenario. An agent does not know whether p is the case, i.e. she cannot distinguish between p -states and $\neg p$ -states. (In fact, p happens to be true.) Furthermore, the agent has no specific reason to think that one state is more probable than any other; therefore it is reasonable for her to assign equal probabilities to all states. This example can be formalized by the following model: $\mathbb{M} = \langle W, R, \mu, V \rangle$, $W = \{w, v\}$, $R = W \times W$, $\mu(w)(w) = \mu(w)(v) = \mu(v)(w) = \mu(v)(v) = 0.5$, and $V(p) = \{w\}$. (We work with only one agent in this example, so agent indices can be dropped.) This model is a faithful representation of the scenario described above; for example:

$$\mathbb{M}, w \models \neg Kp \wedge \neg K\neg p \wedge P(p) = 0.5 \wedge P(\neg p) = 0.5.$$

Now suppose that p is publicly announced (this is indeed possible, since p was assumed to be actually true). Applying Definition 13.4 we obtain the updated model $\mathbb{M}|p$, with $W^p = W$, $R = \{(w, w)\}$, and

$$\mu^p(w)(\llbracket p \rrbracket^{\mathbb{M}|p}) = \mu^p(w)(w) = \frac{\mu(w)(\{w\} \cap \llbracket p \rrbracket^{\mathbb{M}})}{\mu(w)(\llbracket p \rrbracket^{\mathbb{M}})} = \frac{\mu(w)(w)}{\mu(w)(w)} = 1.$$

Using this updated model $\mathbb{M}|p$, we find that

$$\mathbb{M}, w \models [\!|p|\!] (Kp \wedge P(p) = 1 \wedge P(\neg p) = 0).$$

So after the public announcement of p , the agent has come to know that p is in fact the case. She has also adjusted her probabilities: she now assigns probability 1 to p being true, and probability 0 to p being false. These are the results that one would intuitively expect, so Definition 13.4 seems to yield an adequate representation of the epistemic and probabilistic effects of public announcements.

13.3.2 Proof System

Public announcement logic can be axiomatized by adding a set of *reduction axioms* to the static base logic [27]. These axioms allow us to recursively rewrite formulas containing dynamic public announcement operators as formulas without such operators; hence the dynamic language $\mathcal{L}^!$ is equally expressive as the static \mathcal{L} . Alternatively, reduction axioms can be seen as ‘predicting’ what will be the case after the public announcement has taken place in terms of what is the case before the public announcement has taken place.

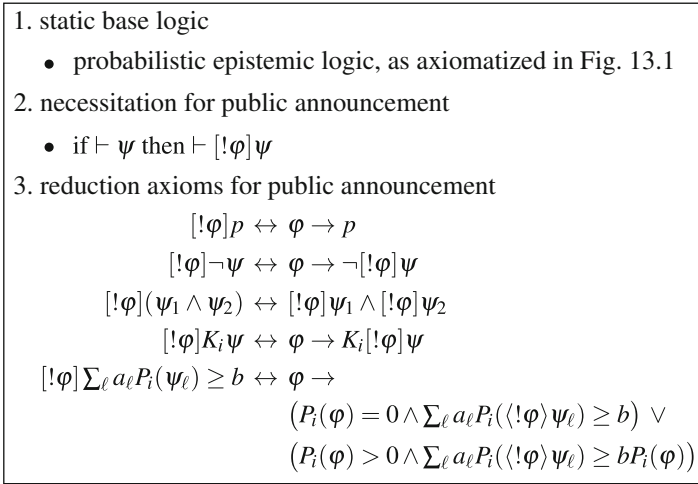


Fig. 13.2 Axiomatization of probabilistic public announcement logic

This strategy can be extended into the probabilistic realm. For the static base logic, we do not simply take some system of epistemic logic (usually S5), but rather the system of probabilistic epistemic logic described in Sect. 13.2.3 (Fig. 13.1), and add the reduction axioms shown in Fig. 13.2. The first four reduction axioms are familiar from classical (non-probabilistic) public announcement logic. Note that the reduction axiom for i -probability formulas makes, just like Definition 13.4, a case distinction based on whether the agent assigns probability 0 to the announced formula φ . The significance of this reduction axiom, and its connection with Bayesian conditionalization, will be further explored in the next subsection.

Once again, standard techniques suffice to prove the following theorem [37]:

Theorem 13.2 *Probabilistic public announcement logic, as axiomatized in Fig. 13.2, is sound and complete with respect to the class of probabilistic Kripke frames.*

13.3.3 Higher-Order Information in Public Announcements

In this subsection we will discuss the role of higher-order information in probabilistic public announcement logic. This will further clarify the connection, but also the distinction, between (dynamic versions of) probabilistic epistemic logic and probability theory proper.

In the previous subsection we introduced a reduction axiom for i -probability formulas. This axiom allows us to derive the following principle as a special case:

$$(\varphi \wedge P(\varphi) > 0) \longrightarrow ([!\varphi]P_i(\psi) \geq b \leftrightarrow P(!\varphi)\psi \geq bP_i(\varphi)). \quad (13.1)$$

The antecedent states that φ is true (because of the truthfulness of public announcements) and that agent i assigns it a strictly positive probability (so that we are in the ‘interesting’ case of the reduction axiom). To see the meaning of the consequent more clearly, note that $\vdash \langle !\varphi \rangle \psi \leftrightarrow (\varphi \wedge [!\varphi]\psi)$, and introduce the following abbreviation of conditional probability into the formal language:

$$P_i(\beta \mid \alpha) \geq b \quad := \quad P_i(\alpha \wedge \beta) \geq bP_i(\alpha).$$

Principle (13.1) can now be rewritten as follows:

$$(\varphi \wedge P(\varphi) > 0) \longrightarrow ([!\varphi]P_i(\psi) \geq b \leftrightarrow P([!\varphi]\psi \mid \varphi) \geq b). \quad (13.2)$$

A similar version can be proved for \leq instead of \geq ; combining these two we get:

$$(\varphi \wedge P(\varphi) > 0) \longrightarrow ([!\varphi]P_i(\psi) = b \leftrightarrow P([!\varphi]\psi \mid \varphi) = b). \quad (13.3)$$

The consequent thus states a connection between the agent’s probability of ψ after the public announcement of φ , and her conditional probability of $[!\varphi]\psi$, given the truth of φ . In other words, after a public announcement of φ , the agent updates her probabilities by Bayesian conditionalization on φ . The subtlety of principle (13.3), however, is that the agent does not take the conditional probability (conditional on φ) of ψ *itself*, but rather of the *updated* formula $[!\varphi]\psi$.

The reason for this is that $[!\varphi]P_i(\psi) = b$ talks about the probability that the agent assigns to ψ after the public announcement of φ has *actually* happened. If we want to describe this probability as a conditional probability, we cannot simply make use of the conditional probability $P_i(\psi \mid \varphi)$, because this represents the probability that the agent *would* assign to ψ if a public announcement of φ *would* happen—hypothetically, not actually! Borrowing a slogan from van Benthem: “The former takes place once arrived at one’s vacation destination, the latter is like reading a travel folder and musing about tropical islands.” [11, p. 417]. Hence, if we want to describe the agent’s probability of ψ after an actual public announcement of φ in terms of conditional probabilities, we need to represent the effects of the public announcement of φ on ψ explicitly, and thus take the conditional probability (conditional on φ) of $[!\varphi]\psi$, rather than ψ .

One might wonder about the relevance of this subtle distinction between actual and hypothetical public announcements. The point is that the public announcement of φ can have effects on the truth value of ψ . For large classes of formulas ψ , this will not occur: their truth value is not affected by the public announcement of φ . Formally, this means that $\vdash \psi \leftrightarrow [!\varphi]\psi$, and thus (the consequent of) principle (13.3) becomes:

$$[!\varphi]P_i(\psi) = b \leftrightarrow P_i(\psi \mid \varphi) = b$$

—thus wiping away all differences between the agent’s probability of ψ after a public announcement of φ , and her conditional probability of ψ , given φ . A typical class of such formulas (whose truth value is unaffected by the public announcement of φ)

is formed by the Boolean combinations of proposition letters, i.e. those formulas which express *ontic* or *first-order information*. Since probability theory proper is usually only concerned with first-order information (‘no nested probabilities’), the distinction between actual and hypothetical announcements—or in general, between actual and hypothetical learning of new information—thus vanishes completely, and Bayesian conditionalization can be used as a universal update rule to compute new probabilities after (actually) learning a new piece of information.

However, in probabilistic epistemic logic (and its dynamic versions, such as probabilistic PAL), higher-order information *is* taken into account, and hence the distinction between actual and hypothetical public announcements has to be taken seriously. Therefore, the consequent of principle (13.3) should really use the conditional probability $P_i([\!\varphi\!]\psi \mid \varphi)$, rather than just $P_i(\psi \mid \varphi)$.⁴

Example 13.2 To illustrate this, consider again the model defined in Example 13.1, and put $\varphi := p \wedge P(\neg p) = 0.5$. It is easy to show that

$$\mathbb{M}, w \models P(\varphi \mid \varphi) = 1 \wedge P([\!\varphi\!]\varphi \mid \varphi) = 0 \wedge [\!\varphi\!]P(\varphi) = 0.$$

Hence the probability assigned to φ after the public announcement is the conditional probability $P([\!\varphi\!]\varphi \mid \varphi)$, rather than just $P(\varphi \mid \varphi)$. Note that this example indeed involves higher-order information, since we are talking about the probability of φ , which itself contains the probability statement $P(\neg p) = 0.5$ as a conjunct. Finally, this example also shows that learning a new piece of information φ (via public announcement) does *not* automatically lead to the agents being certain about (i.e. assigning probability 1 to) that formula. This is to be contrasted with probability theory, where a new piece of information φ is processed via Bayesian conditionalization, and thus always leads to certainty: $P(\varphi \mid \varphi) = 1$. The explanation is, once again, that probability theory is only concerned with first-order information, whereas the phenomena described above can only occur at the level of higher-order information.^{5,6}

⁴ Romeijn [45] provides an analysis that stays closer in spirit to probability theory proper. He argues that the public announcement of φ induces a shift in the interpretation of ψ (in our terminology: from ψ to $[\!\varphi\!]\psi$, i.e. from $\llbracket \psi \rrbracket^{\mathbb{M}}$ to $\llbracket \psi \rrbracket^{\mathbb{M}[\!\varphi\!]}$), and shows that such meaning shifts can be modeled using Dempster-Shafer belief functions. Crucially, however, this proposal is able to deal with the case of ψ expressing *second-order* information (e.g. when it is of the form $P_i(p) = b$), but not with the case of higher-order information *in general* (e.g. when ψ is of the form $P_j(P_i(p) = b) = a$, or involves even more deeply nested probabilities) [45, p. 603].

⁵ Similarly, the *success postulate* for belief expansion in the (traditional) AGM framework [1, 30] states that after expanding one’s belief set with a new piece of information φ , the updated (expanded) belief set should always contain this new information. Also here the explanation is that AGM is only concerned with first-order information. (Note that we talk about the success postulate for belief *expansion*, rather than belief *revision*, because the former seems to be the best analogue of public announcement in the AGM framework.)

⁶ The occurrence of higher-order information is a *necessary* condition for this phenomenon, but not a *sufficient* one: there exist formulas φ that involve higher-order information, but still $\models [\!\varphi\!]P_i(\varphi) = 1$ (or epistemically: $\models [\!\varphi\!]K_i\varphi$).

13.4 Probabilistic Dynamic Epistemic Logic

In this section we will move from a probabilistic version of public announcement logic to a probabilistic version of ‘full’ dynamic epistemic logic. Section 13.4.1 introduces a probabilistic version of the *product update* mechanism that is behind dynamic epistemic logic. Section 13.4.2 introduces dynamic operators into the formal language to talk about these product updates, and discusses a detailed example. Section 13.4.3, finally, shows how to obtain a complete axiomatization in a fully standard (though non-trivial) fashion.

13.4.1 Probabilistic Product Update

Classical (non-probabilistic) dynamic epistemic logic models epistemic dynamics by means of a product update mechanism [4, 5]. The agents’ *static* information (what is the current state?) is represented in a Kripke model \mathbb{M} , and their *dynamic* information (what type of event is currently taking place?) is represented in an update model \mathbb{E} . The agents’ new information (after the dynamics has taken place) is represented by means of a product construction $\mathbb{M} \otimes \mathbb{E}$. We will now show how to define a probabilistic version of this construction.

Before stating the formal definitions, we show how they naturally arise as probabilistic generalizations of the classical (non-probabilistic) notions. The probabilistic Kripke models introduced in Definition 13.2 represent the agents’ static information, in both its epistemic and its probabilistic aspects. This static probabilistic information is called the *prior probabilities of the states* in [17]. We can thus say that when w is the actual state, agent i considers it epistemically possible that v is the actual state ($(w, v) \in R_i$), and, more specifically, that she assigns probability b to v being the actual state ($\mu_i(w)(v) = b$).

Update models are essentially like Kripke models: they represent the agents’ information about events, rather than states. Since probabilistic Kripke models represent both epistemic and probabilistic information about *states*, by analogy probabilistic update models should represent both epistemic and probabilistic information about *events*. Hence, they should not only have epistemic accessibility relations R_i over their set of events E , but also probability functions $\mu_i : E \rightarrow (E \rightarrow [0, 1])$. (Formal details will be given in Definition 13.5.) We can then say that when e is the actually occurring event, agent i considers it epistemically possible that f is the actually occurring event ($(e, f) \in R_i$), and, more specifically, that she assigns probability b to f being the actually occurring event ($\mu_i(e)(f) = b$). This dynamic probabilistic information is called the *observation probabilities* in van Benthem et al. [17].

Finally, how probable it is that an event e will occur, might vary from state to state. We assume that this variation can be captured by means of a set Φ of (pairwise inconsistent) sentences in the object language (so that the probability that an event e will occur can only vary between states that satisfy *different* sentences of Φ). This will

be formalized by adding to the probabilistic update models a set of preconditions Φ , and probability functions $\text{pre}: \Phi \rightarrow (E \rightarrow [0, 1])$. The meaning of $\text{pre}(\varphi)(e) = b$ is that if φ holds, then event e occurs with probability b . In van Benthem et al. [17] these are called *occurrence probabilities*.⁷

We are now ready to formally introduce probabilistic update models:

Definition 13.5 A *probabilistic update model* is a tuple $\mathbb{E} = \langle E, R_i, \Phi, \text{pre}, \mu_i \rangle_{i \in I}$, where E is a non-empty finite set of events, $R_i \subseteq E \times E$ is agent i 's epistemic accessibility relation, $\Phi \subseteq \mathcal{L}^\otimes$ is a finite set of pairwise inconsistent sentences called *preconditions*, $\mu_i: E \rightarrow (E \rightarrow [0, 1])$ assigns to each event $e \in E$ a probability function $\mu_i(e)$ over E , and $\text{pre}: \Phi \rightarrow (E \rightarrow [0, 1])$ assigns to each precondition $\varphi \in \Phi$ a probability function $\text{pre}(\varphi)$ over E .

All components of a probabilistic update model have already been commented upon. Note that we use the same symbols R_i and μ_i to indicate agent i 's epistemic and probabilistic information in a probabilistic Kripke model \mathbb{M} and in a probabilistic update model \mathbb{E} —from the context it will always be clear which of the two is meant. The language \mathcal{L}^\otimes that the preconditions are taken from will be formally defined in the next subsection. (As is usual in this area, there is a non-vicious simultaneous recursion going on here.)

We now introduce occurrence probabilities for events at states:

Definition 13.6 Consider a probabilistic Kripke model \mathbb{M} , a state w , a probabilistic update model \mathbb{E} , and an event e . Then the *occurrence probability of e at w* is defined as

$$\text{pre}(w)(e) = \begin{cases} \text{pre}(\varphi)(e) & \text{if } \varphi \in \Phi \text{ and } \mathbb{M}, w \models \varphi \\ 0 & \text{if there is no } \varphi \in \Phi \text{ such that } \mathbb{M}, w \models \varphi. \end{cases}$$

Since the preconditions are pairwise inconsistent, $\text{pre}(w)(e)$ is always well-defined. The meaning of $\text{pre}(w)(e) = b$ is that in state w , event e occurs with probability b . Note that if two states w and v satisfy the same precondition, then always $\text{pre}(w)(e) = \text{pre}(v)(e)$; in other words, the occurrence probabilities of an event e can only vary 'up to a precondition' (cf. supra).

The probabilistic product update mechanism can now be defined as follows:

Definition 13.7 Consider a probabilistic Kripke model $\mathbb{M} = \langle W, R_i, \mu_i, V \rangle_{i \in I}$ and a probabilistic update model $\mathbb{E} = \langle E, R_i, \Phi, \text{pre}, \mu_i \rangle_{i \in I}$. Then the *updated model* $\mathbb{M} \otimes \mathbb{E} := \langle W', R'_i, \mu'_i, V' \rangle_{i \in I}$ is defined as follows:

- $W' := \{(w, e) \mid w \in W, e \in E, \text{pre}(w)(e) > 0\}$,
- $R'_i := \{((w, e), (w', e')) \in W' \times W' \mid (w, w') \in R_i \text{ and } (e, e') \in R_i\}$ (for every agent $i \in I$),

⁷ Occurrence probabilities are often assumed to be *objective frequencies*. This is reflected in the formal setup: the function pre is not agent-dependent.

- $\mu'_i: W' \rightarrow (W' \rightarrow [0, 1])$ is defined (for every agent $i \in I$) by

$$\mu'_i(w, e)(w', e') := \frac{\mu_i(w)(w') \cdot \text{pre}(w')(e') \cdot \mu_i(e)(e')}{\sum_{\substack{w'' \in W \\ e'' \in E}} \mu_i(w)(w'') \cdot \text{pre}(w'')(e'') \cdot \mu_i(e)(e')}$$

if the denominator is strictly positive, and $\mu'_i(w, e)(w', e') := 0$ otherwise,

- $V'(p) := \{(w, e) \in W' \mid w \in V(p)\}$ (for every $p \in \text{Prop}$).

We will only comment on the probabilistic component of this definition (all other components are fully classical). After the dynamics has taken place, agent i calculates at state (w, e) her new probability for (w', e') by taking the arithmetical product of (i) her *prior probability* for w' at w , (ii) the *occurrence probability* of e' in w' , and (iii) her *observation probability* for e' at e , and then normalizing this product. The factors in this product are not *weighted* (or equivalently, they all have weight 1)—van Benthem et al. [17] also discusses weighted versions of this update mechanism, and shows how one of these weighted versions corresponds to the rule of *Jeffrey conditioning* from probability theory [36]. Finally, note that $\mathbb{M} \otimes \mathbb{E}$ might fail to be a probabilistic Kripke model: if the denominator in the definition of $\mu'_i(w, e)$ is 0, then $\mu'_i(w, e)$ assigns 0 to all states in W' . We will not care here about the interpretation of this feature, but only remark that technically speaking it is harmless and, perhaps most importantly, still allows for a reduction axiom for i -probability formulas (cf. Sect. 13.4.3).

13.4.2 Language and Semantics

To talk about these updated models, we add dynamic operators $[E, e]$ to the static language \mathcal{L} , thus obtaining the new language \mathcal{L}^\otimes . Here, E, e are formal names for the probabilistic update model $\mathbb{E} = \langle E, R_i, \Phi, \text{pre}, \mu_i \rangle_{i \in I}$ and event $e \in E$ (recall our remark about the mutual recursion of the dynamic language and the updated models). The formula $[E, e]\varphi$ means that after the event e has occurred, it will be the case that φ . It has the following semantics:

$$\mathbb{M}, w \models [E, e]\psi \quad \text{iff} \quad \text{if } \text{pre}(w)(e) > 0, \text{ then } \mathbb{M} \otimes \mathbb{E}, (w, e) \models \psi.$$

Example 13.3 Consider the following scenario. While strolling through a flea market, you see a painting that you think might be a real Picasso. Of course, the chance that the painting is actually a real Picasso is very slim, say 1 in 100,000. You know from an art encyclopedia that Picasso signed almost all his paintings with a very characteristic signature. If the painting is a real Picasso, the chance that it bears the characteristic signature is 97 %, while if the painting is not a real Picasso, the chance that it bears the characteristic signature is 0 % (nobody is capable of imitating Picasso's signature). You immediately look at the painting's signature, but determining whether it is Picasso's characteristic signature is very hard, and—not being an

expert art historian—you remain uncertain and think that the chance is 50 % that the painting’s signature is Picasso’s characteristic one.

Your initial information (before having looked at the painting’s signature) can be represented as the following probabilistic Kripke model: $\mathbb{M} = \langle W, R, \mu, V \rangle$, where $W = \{w, v\}$, $R = W \times W$, $\mu(w)(w) = \mu(v)(w) = 0.00001$, $\mu(w)(v) = \mu(v)(v) = 0.99999$, and $V(\text{real}) = \{w\}$. (We work with only one agent in this example, so agent indices can be dropped.) Hence, initially you do not rule out the possibility that the painting in front of you is a real Picasso, but you consider it highly unlikely:

$$\mathbb{M}, w \models \hat{K} \text{real} \wedge P(\text{real}) = 0.00001.$$

The event of looking at the signature can be represented with the following update model: $\mathbb{E} = \langle E, R, \Phi, \text{pre} \rangle$, where $E = \{e, f\}$, $R = E \times E$, $\Phi = \{\text{real}, \neg\text{real}\}$, $\text{pre}(\text{real})(e) = 0.97$, $\text{pre}(\text{real})(f) = 0.03$, $\text{pre}(\neg\text{real})(e) = 0$, $\text{pre}(\neg\text{real})(f) = 1$, and $\mu(e)(e) = \mu(f)(e) = \mu(e)(f) = \mu(f)(f) = 0.5$. The event e represents ‘looking at Picasso’s characteristic signature’; the event f represents ‘looking at a signature that is not Picasso’s characteristic one’.

We now construct the updated model $\mathbb{M} \otimes \mathbb{E}$. Since $\mathbb{M}, v \not\models \text{real}$, it holds that $\text{pre}(v)(e) = \text{pre}(v)(f) = 0$, and hence (v, e) does not belong to the updated model. It is easy to see that the other states (w, e) , (w, f) and (v, f) do belong to the updated model. Furthermore, one can easily calculate that $\mu'(w, e)(w, e) = 0.0000003$ and $\mu'(w, e)(w, f) = 0.0000097$, so $\mu'(w, e)([\text{real}])^{\mathbb{M} \otimes \mathbb{E}} = 0.0000003 + 0.0000097 = 0.00001$, and thus

$$\mathbb{M}, w \models [\mathbb{E}, e]P(\text{real}) = 0.00001.$$

Hence, even though the painting in front of you is a real Picasso (in state w), after looking at the signature (which is indeed Picasso’s characteristic signature!—the event that actually happened was event e) you still assign a probability of 1 in 100,000 to it being a real Picasso.

Note that if you had been an expert art historian, with the same prior probabilities, but with the reliable capability of recognizing Picasso’s characteristic signature—let’s formalize this as $\mu(e)(e) = 0.99$ and $\mu(e)(f) = 0.01$ —, then the same update mechanism would have implied that

$$\mathbb{M}, w \models [\mathbb{E}, e]P(\text{real}) = 0.00096.$$

In other words, if you had been an expert art historian, then looking at the painting’s signature would have been highly informative: it would have led to a significant change in your probabilities.

13.4.3 Proof System

A complete axiomatization for probabilistic dynamic epistemic logic can be found using the standard strategy, viz. by adding a set of reduction axioms to static

probabilistic epistemic logic. Implementing this strategy, however, is not entirely trivial. The reduction axioms for non-probabilistic formulas are familiar from classical (non-probabilistic) dynamic epistemic logic, but the reduction axiom for i -probability formulas is more complicated.

First of all, this reduction axiom makes a case distinction on whether a certain sum of probabilities is strictly positive or not. We will show that this corresponds to the case distinction made in the definition of the updated probability functions (Definition 13.7). In the definition of $\mu'_i(w, e)$, a case distinction is made on the value of the denominator of a fraction, i.e. on the value of the following expression:

$$\sum_{\substack{v \in W \\ f \in E}} \mu_i(w)(v) \cdot \text{pre}(v)(f) \cdot \mu_i(e)(f). \quad (13.4)$$

But this expression can be rewritten as

$$\sum_{\substack{v \in W \\ f \in E \\ \varphi \in \Phi \\ \mathbb{M}, v \models \varphi}} \mu_i(w)(v) \cdot \text{pre}(\varphi)(f) \cdot \mu_i(e)(f).$$

Using the definition of $k_{i,e,\varphi,f}$ (cf. Fig. 13.3), this can be rewritten as

$$\sum_{\substack{\varphi \in \Phi \\ f \in E}} \mu_i(w)(\llbracket \varphi \rrbracket^{\mathbb{M}}) \cdot k_{i,e,\varphi,f}.$$

Since E and Φ are finite, this sum is finite and corresponds to an expression in the formal language \mathcal{L}^\otimes , which we will abbreviate as σ :

$$\sigma := \sum_{\substack{\varphi \in \Phi \\ f \in E}} k_{i,e,\varphi,f} P_i(\varphi).$$

This expression can be turned into an i -probability formula by ‘comparing’ it with a rational number b ; for example $\sigma \geq b$. Particularly important are the formulas $\sigma = 0$ and $\sigma > 0$: exactly these formulas are used to make the case distinction in the reduction axiom for i -probability formulas.⁸

Next, the reduction axiom for i -probability formulas provides a statement in each case of the case distinction: $0 \geq b$ in the case $\sigma = 0$, and χ (as defined in Fig. 13.3) in the case $\sigma > 0$. We will only explain the meaning of χ in the ‘interesting’ case

⁸ Note that E and Φ are components of the probabilistic update model \mathbb{E} named by \mathbf{E} ; furthermore, the values $k_{i,e,\varphi,f}$ are fully determined by the model \mathbb{E} and event e named by \mathbf{E} and \mathbf{e} , respectively (consider their definition in Fig. 13.3). Hence any i -probability formula involving σ is fully determined by \mathbb{E} , e , and can be interpreted at any probabilistic Kripke model \mathbb{M} and state w .

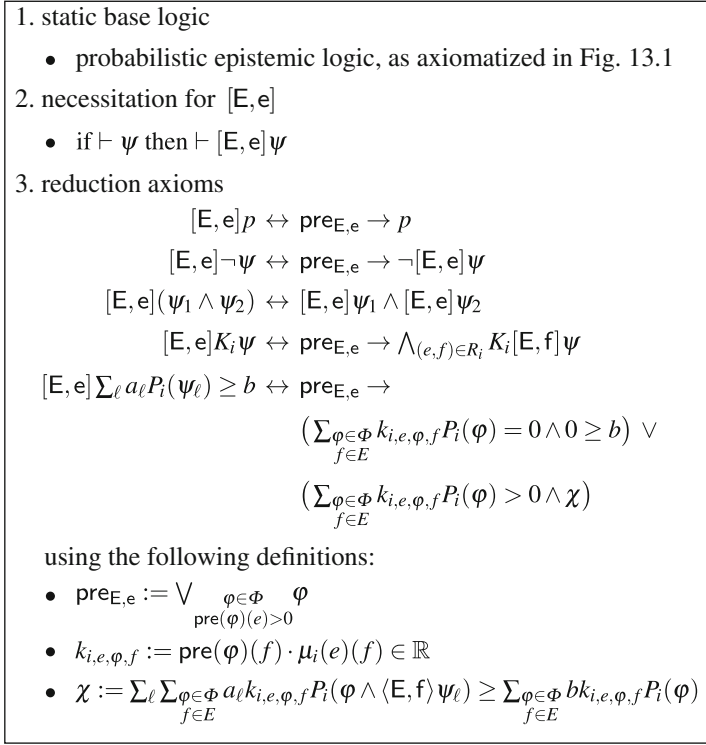


Fig. 13.3 Axiomatization of probabilistic dynamic epistemic logic

$\sigma > 0$. If $\mathbb{M}, w \models \sigma > 0$, then the value of (13.4) is strictly positive (cf. supra), and we can calculate:

$$\begin{aligned}
 \mu'_i(w, e)(\llbracket \psi \rrbracket^{\mathbb{M} \otimes \mathbb{E}}) &= \sum_{\mathbb{M} \otimes \mathbb{E}, (w', e') \models \psi} \mu'_i(w, e)(w', e') \\
 &= \sum_{\substack{w' \in W, e' \in E \\ \mathbb{M}, w' \models \langle E, e' \rangle \psi}} \frac{\mu_i(w)(w') \cdot \text{pre}(w')(e') \cdot \mu_i(e)(e')}{\sum_{v \in W} \sum_{f \in E} \mu_i(w)(v) \cdot \text{pre}(v)(f) \cdot \mu_i(e)(f)} \\
 &= \frac{\sum_{\varphi \in \Phi} \sum_{f \in E} \mu_i(w)(\llbracket \varphi \wedge \langle E, f \rangle \psi \rrbracket^{\mathbb{M}}) \cdot k_{i, e, \varphi, f}}{\sum_{\varphi \in \Phi} \sum_{f \in E} \mu_i(w)(\llbracket \varphi \rrbracket^{\mathbb{M}}) \cdot k_{i, e, \varphi, f}}.
 \end{aligned}$$

Hence, in this case ($\sigma > 0$) we can express that $\mu'_i(w, e)(\llbracket \psi \rrbracket^{\mathbb{M} \otimes \mathbb{E}}) \geq b$ in the formal language, by means of the following i -probability formula:

$$\sum_{\substack{\varphi \in \Phi \\ f \in E}} k_{i, e, \varphi, f} P_i(\varphi \wedge \langle E, f \rangle \psi) \geq \sum_{\substack{\varphi \in \Phi \\ f \in E}} b k_{i, e, \varphi, f} P_i(\varphi).$$

Moving to linear combinations, we can express that $\sum_{\ell} a_{\ell} \mu'_i(w, e)(\llbracket \psi_{\ell} \rrbracket^{\mathbb{M} \otimes \mathbb{E}}) \geq b$ in the formal language using an analogous i -probability formula, namely χ (cf. the definition of this formula in Fig. 13.3).

We thus obtain the following theorem [17]:

Theorem 13.3 *Probabilistic dynamic epistemic logic, as axiomatized in Fig. 13.3, is sound and complete with respect to the class of probabilistic Kripke frames.*

13.5 Further Developments and Applications

Probabilistic extensions of dynamic epistemic logic are a recent development, and there are various open questions and potential applications to be explored. In this section we discuss a selection of such topics for further research; more suggestions can be found in [17] and [15, ch. 8].

We distinguish between *technical* and *conceptual* open problems.⁹ A typical technical problem that needs further research is the issue of *surprising information*. In the update mechanisms described in this chapter, the agents' new probabilities are calculated by means of a fraction whose denominator might take on the value 0. The focus has been on the 'interesting' (non-0) cases, and the 0-case has been treated as mere 'noise': a technical artefact that cannot be handled convincingly by the system. However, sometimes such 0-cases *do* represent very intuitive scenarios; for example, one can easily think of an agent being absolutely certain that a certain proposition φ is false ($P(\varphi) = 0$), while that proposition is actually true, and can thus be announced! In such cases, the system of probabilistic public announcement logic described in Sect. 13.3 predicts that the agent will simply *ignore* the announced information (rather than performing some sensible form of *belief revision*). More can, and should be said about such cases [2, 6, 46].

Another technical question is whether other representations of *soft information* can learn something from the probabilistic approach to dynamic epistemic logic. Probabilistic Kripke models represent the agents' soft information via the probability functions μ_i , and interpret formulas of the form $P_i(\varphi) \geq b$. Plausibility models, on the other hand, represent the agents' soft information via a (non-numerical) plausibility ordering \leq_i , and interpret more qualitative notions of belief [7, 12, 15, 22]. In particular, if we use $Min_{\leq_i}(X)$ to denote the set of \leq_i -minimal states in the set X , then the formula $B_i\varphi$ is interpreted in a plausibility model \mathbb{M} as follows:¹⁰

$$\mathbb{M}, w \models B_i\varphi \text{ iff for all } v \in Min_{\leq_i}(R_i[w]): \mathbb{M}, v \models \varphi.$$

The product update for probabilistic Kripke models described in Definition 13.7 takes into account prior probabilities ($\mu_i(w)(v)$ for states w and v), observation probabilities ($\mu_i(e)(f)$ for events e and f), and occurrence probabilities ($\mathbf{pre}(w)(e)$ for a state w and event e). One can also define a product update for plausibility models; a widely used rule to define the updated plausibility ordering is the so-called 'priority update' [7, 15]:

⁹ In practice, this distinction will not always be clear-cut, of course.

¹⁰ As usual, $R_i[w]$ denotes the set $\{v \in W \mid (w, v) \in R_i\}$.

$$(w, e) \leq_i (v, f) \text{ iff } e <_i f \text{ or } (e \cong_i f \text{ and } w \leq_i v).$$

The updated plausibility ordering thus gives priority to the plausibility ordering on events, and otherwise preserves the original plausibility ordering on states as much as possible. In analogy with the probabilistic setting, the plausibility orderings on states and events can be called the ‘prior plausibility’ and ‘observation plausibility’, respectively. However, the notion of occurrence probability does *not* seem to have a clear analogue in the framework of plausibility models. van Benthem [16] defines a notion of ‘occurrence plausibility’, which can be expressed as $e \leq_w f$: at state w , event e is at least as plausible as f to occur (this ordering is not agent-dependent; recall Footnote 7). New product update rules thus have to merge *three* plausibility orderings: prior plausibility, observation plausibility, and occurrence plausibility. van Benthem [16] makes some proposals for such rules, but finding a fully satisfactory definition remains a major open problem in this area.

An important conceptual issue that is currently actively being investigated, is the exact relation between the quantitative (probabilistic) and qualitative perspectives on soft information. A widespread proposal is to connect *belief* with *high probability*, where ‘high’ means ‘above some threshold $\tau \in (0.5, 1]$ ’; formally: $B_i \varphi \Leftrightarrow P_i(\varphi) \geq \tau$. An immediate problem of this proposal is that belief is standardly taken to be closed under conjunction, while ‘high probability’ is not closed under conjunction (unless $\tau = 1$). Despite this initial problem, there’s also a lot in favor of this proposal. Plausibility models not only interpret a notion of belief, but also a notion of *conditional belief*: $B_i^\alpha \varphi$ means that agent i believes that φ , conditional on α . The connection between belief and high probability can perfectly be extended to conditional belief and high conditional probability:

$$B_i^\alpha \varphi \Leftrightarrow P_i(\varphi | \alpha) \geq \tau. \quad (13.5)$$

Furthermore, (conditional) belief and high (conditional) probability seem to display highly similar dynamic behavior under public announcements. We saw in Sect. 13.3.3 that $[\!|\varphi]P_i(\psi) \geq \tau$ can sometimes diverge in truth value from $P_i(\psi | \varphi) \geq \tau$, because of the presence of higher-order information. In exactly the same way (and for the same reason), $[\!|\varphi]B_i \psi$ and $B_i^\varphi \psi$ can diverge in truth value on plausibility models. Furthermore, (conditional) belief and high (conditional) probability have exactly the ‘same’ reduction axiom. This means that (13.6) (which is interpreted on probabilistic Kripke models) and (13.7) (which is interpreted on plausibility models) are intertranslatable, using principle (13.5) above:

$$[\!|\varphi]P_i(\psi | \alpha) \geq \tau \Leftrightarrow \left(\varphi \rightarrow P_i((\!|\varphi)\psi | (\!|\varphi)\alpha) \geq \tau \right), \quad (13.6)$$

$$[\!|\varphi]B_i^\alpha \psi \Leftrightarrow \left(\varphi \rightarrow B_i^{(\!|\varphi)\alpha} (\!|\varphi)\psi \right). \quad (13.7)$$

The significance of these observations is further discussed in [24].

Several fruitful *applications* of probabilistic dynamic epistemic logic can be expected in the fields of *game theory* and *cognitive science*. In recent years, dynamic

epistemic logic has been widely applied to explore the epistemic foundations of game theory [10, 13, 18]. However, given the importance of probability in game theory (for example, in the notion of mixed strategy), it is surprising that very few of these logical analyses have a probabilistic component.¹¹ Probabilistic dynamic epistemic logic provides the required tools to explore the epistemic *and* the probabilistic aspects of game theory.

For example, [21, 23] uses a version of probabilistic public announcement logic to analyze the role of common knowledge and communication in Aumann's well-known agreeing to disagree theorem. Classically, this theorem is stated as follows: "If two people have the same prior, and their posteriors for an event $[\varphi]$ are common knowledge, then these posteriors are equal" [3, p. 1236]. If we represent the experiments (with respect to which the agents' probabilities are called 'prior' and 'posterior') with a dynamic operator [EXP], then this version can be formalized as (13.8), which is derivable in the underlying logical system:

$$[\text{EXP}]C(P_1(\varphi) = a \wedge P_2(\varphi) = b) \rightarrow a = b. \quad (13.8)$$

However, this version does not say *how* the agents are to obtain this common knowledge; it just assumes that they have been able to obtain it one way or another. The way to obtain common knowledge is via a certain communication protocol, which is described explicitly in the intuitive scenario that is used to motivate and explain this theorem. Once this communication dynamics is made explicitly part of the story, common knowledge of the posteriors need no longer be *assumed* in the formulation of the agreement theorem, since it will now simply *follow* from the communication protocol. If we represent the communication protocol with a dynamic operator [DIAL(φ)], this new version of the theorem can be formalized as (13.9):

$$[\text{EXP}][\text{DIAL}(\varphi)](P_1(\varphi) = a \wedge P_2(\varphi) = b) \rightarrow a = b. \quad (13.9)$$

The notion of common knowledge is thus less central to the agreement theorem than is usually thought: if we compare (13.8) and (13.9), it is clear that common knowledge and communication are two sides of the same coin; the former is only needed to formulate the agreement theorem if the latter is not represented explicitly.

Another potential field of application is cognitive science. The usefulness of (epistemic) logic for cognitive science has been widely recognized [14, 35, 43]. Of course, as in any other empirical discipline, one quickly finds out that real-life human cognition is rarely a matter of all-or-nothing, but often involves degrees (probabilities). Furthermore, a recent development in cognitive science is toward probabilistic (Bayesian) models of cognition [42]. If epistemic logic is to remain a valuable tool here, it will thus have to be a thoroughly 'probabilized' version. For example, probabilistic dynamic epistemic logic has been used to model the cognitive phenomenon of surprise and its epistemic aspects [25, 39].

¹¹ The logic in [19] does have a probabilistic component, but this logic is fully *static*.

13.6 Conclusion

In this chapter we have introduced probabilistic epistemic logic, and several of its dynamic versions. These logics provide a standard epistemic (possible-worlds) analysis of the agents' hard information, and supplement it with a fine-grained probabilistic analysis of their soft information. Higher-order information of any kind (knowledge about probabilities, probabilities about knowledge, etc.) is represented explicitly. The importance of higher-order information in dynamics is clear from our discussion of the connection between public announcements and Bayesian conditionalization. The probabilistic versions of both public announcement logic and dynamic epistemic logic with product updates can be completely axiomatized in a standard way (via reduction axioms). The fertility of the research program of probabilistic dynamic epistemic logic is illustrated by the variety of technical and conceptual issues that are still open for further research, and its (potential) use in analyzing theorems and phenomena from game theory and cognitive science.

Acknowledgments The authors wish to thank Alexandru Baltag, Johan van Benthem and Sonja Smets for their valuable feedback on earlier versions of this chapter. Lorenz Demey is financially supported by a PhD fellowship of the Research Foundation–Flanders (FWO).

References

1. Alchourrón C, Gärdenfors P, Makinson D (1985) On the logic of theory change: partial meet contraction and revision functions. *J Symb Logic* 50:510–530
2. Aucher G (2003) A combined system for update logic and belief revision. Master's thesis, Institute for Logic, Language and Computation, Universiteit van Amsterdam
3. Aumann R (1976) Agreeing to disagree. *Ann Stat* 4:1236–1239
4. Baltag A, Moss LS (2004) Logics for epistemic programs. *Synthese* 139:1–60
5. Baltag A, Moss LS, Solecki S (1998) The logic of common knowledge, public announcements, and private suspicions. In: Gilboa I (ed) *Proceedings of the 7th conference on theoretical aspects of rationality and knowledge (TARK '98)*, pp 43–56
6. Baltag A, Smets S (2008) Probabilistic dynamic belief revision. *Synthese* 165:179–202
7. Baltag A, Smets S (2008) A qualitative theory of dynamic interactive belief revision. In: Bonanno G, van der Hoek W, Woolridge M (eds) *Texts in logic and games*, vol 3. Amsterdam University Press, pp 9–58
8. van Benthem J (1983) *Modal logic and classical logic*. Bibliopolis, Napoli
9. van Benthem J (2001) Correspondence theory. In: Gabbay DM, Guenther F (eds) *Handbook of philosophical logic*, revised 2nd edn, vol 3. Kluwer, Dordrecht, pp 325–408
10. van Benthem J (2001) Games in dynamic epistemic logic. *Bull Econ Res* 53:219–248
11. van Benthem J (2003) Conditional probability meets update logic. *J Logic Lang Inf* 12:409–421
12. van Benthem J (2007) Dynamic logic for belief revision. *J Appl Nonclass Logics* 17:129–155
13. van Benthem J (2007) Rational dynamics and epistemic logic in games. *Int Game Theory Rev* 9:13–45
14. van Benthem J (2008) Logic and reasoning: do the facts matter? *Stud Logica* 88:67–84
15. van Benthem J (2011) *Logical dynamics of information and interaction*. Cambridge University Press, Cambridge
16. van Benthem J (2012) A problem concerning qualitative probabilistic update. p 10. (Unpublished manuscript)

17. van Benthem J, Gerbrandy J, Kooi BP (2009) Dynamic update with probabilities. *Stud Logica* 93:67–96
18. Bonanno G, Dégrémont C (2014) Logic and game theory. In: Baltag A, Smets S (eds) *Johan van Benthem on logic and information dynamics*. Springer, Dordrecht, pp 421–449 (Chapter 15 in this volume)
19. de Bruin B (2010) *Explaining games: the epistemic programme in game theory*. Springer, Dordrecht
20. Conradie W, Ghilardi S, Palmigiano A (2014) Unified correspondence. In: Baltag A, Smets S (eds) *Johan van Benthem on logic and information dynamics*. Springer, Dordrecht, pp 933–975 (Chapter 36 in this volume)
21. Demey L (2010) *Agreeing to disagree in probabilistic dynamic epistemic logic*. Master's thesis, Institute for Logic, Language and Computation, Universiteit van Amsterdam
22. Demey L (2011) Some remarks on the model theory of epistemic plausibility models. *J Appl Nonclass Logics* 21:375–395
23. Demey, L. (2013) *Contemporary epistemic logic and the Lockean thesis*. *Found Sci* 18:599–610
24. Demey, L. (2014) *Agreeing to disagree in probabilistic dynamic epistemic logic*. *Synthese* 191:409–438
25. Demey L (forthcoming) *The dynamics of surprise*. *Logique Anal*
26. Demey L, Kooi B, Sack J (2013) Logic and probability. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy*, Stanford
27. van Ditmarsch H, van der Hoek W, Kooi B (2007) *Dynamic epistemic logic*. Springer, Dordrecht
28. Fagin R, Halpern J (1994) Reasoning about knowledge and probability. *J ACM* 41:340–367
29. van Fraassen B (1984) Belief and the will. *J Philos* 81:235–256
30. Gärdenfors P (1988) *Knowledge in flux*. MIT Press, Cambridge
31. Gerbrandy J, Groeneveld W (1997) Reasoning about information change. *J Logic Lang Inf* 6:147–169
32. Halpern JY (1991) The relationship between knowledge, belief, and certainty. *Ann Math Artif Intell* 4:301–322 (errata in the same journal, 26:59–61, 1999)
33. Halpern JY (2003) *Reasoning about uncertainty*. MIT Press, Cambridge
34. Hintikka J (1962) *Knowledge and belief: an introduction to the logic of the two notions*. Cornell University Press, Ithaca
35. Isaac A, Szymanik J, Verbrugge R (2014) Logic and complexity in cognitive science. In: Baltag A, Smets S (eds) *Johan van Benthem on logic and information dynamics*. Springer, Dordrecht, pp 787–824 (Chapter 30 in this volume)
36. Jeffrey R (1983) *The logic of decision*, 2nd edn. University of Chicago Press, Chicago
37. Kooi BP (2003) Probabilistic dynamic epistemic logic. *J Logic Lang Inf* 12:381–408
38. Lewis D (1980) A subjectivist's guide to objective chance. In: Jeffrey RC (ed) *Studies in inductive logic and probability*, vol 2. University of California Press, Berkeley, pp 263–293
39. Lorini E, Castelfranchi C (2007) The cognitive structure of surprise: looking for basic principles. *Topoi* 26:133–149
40. Meacham CJG (2010) Two mistakes regarding the principal principle. *Br J Philos Sci* 61:407–431
41. Miller D (1966) A paradox of information. *Br J Philos Sci* 17:59–61
42. Oaksford M, Chater N (2008) *The probabilistic mind: prospects for Bayesian cognitive science*. Oxford University Press, Oxford
43. Pietarinen AV (2003) What do epistemic logic and cognitive science have to do with each other? *Cogn Syst Res* 4:169–190
44. Plaza J (1989) Logics of public communications. In: Emrich ML, Pfeifer MS, Hadzikadic M, Ras ZW (eds) *Proceedings of the 4th international symposium on methodologies for intelligent systems: poster session program*, Oak Ridge National Laboratory, Oak Ridge, pp 201–216 (Reprinted in: *Synthese* 158:165–179, 2007)
45. Romeijn JW (2012) Conditioning and interpretation shifts. *Stud Logica* 100:583–606
46. Rott H, Girard P (2014) Belief revision and dynamic logic. In: Baltag A, Smets S (eds) *Johan van Benthem on logic and information dynamics*. Springer, Dordrecht, pp 203–233 (Chapter 8 in this volume)
47. Sack J (2009) Extending probabilistic dynamic epistemic logic. *Synthese* 169:241–257