# Extending Probabilistic Dynamic Epistemic Logic 

Joshua Sack<br>joshua.sack@gmail.com


#### Abstract

This paper aims to extend in two directions the probabilistic dynamic epistemic logic provided in Kooi's paper [8]. Kooi's probabilistic dynamic epistemic logic adds to probabilistic epistemic logic sentences that express consequences of public announcements. The first extension offered in this paper is to add a previous time operator to a probabilistic dynamic epistemic logic similar to Kooi's. The other is to involve action models and update products in a probabilistic dynamic epistemic logic setting. This would allow for more kinds of actions, such as private announcements.


## 1 Introduction

Probabilistic epistemic logic has been developed to express interaction between both qualitative and quantitative beliefs. This logic lets us formally express statements such as "Bob believes the probability of $\varphi$ to be at least $1 / 2$ " or "Ann considers the probability of $\psi$ to be $1 / 4$ ". As we are often concerned about how beliefs and probabilities change over time, there have been papers written that mix probability, belief, and time. Examples include, [7] and [3], which use probabilistic systems of runs, and [8], which combines probability with public announcement logic. The probabilistic systems of runs provides a natural way to view time, both past and future, but conditions need to be imposed in order to ensure that agents' probability measures change in a realistic way. Public announcement logic, and more generally dynamic epistemic logic (DEL), provides a mechanical procedure for changes in belief upon receipt of public information, and [8] extends this mechanical procedure to show how a probability measure may change given public information. But DEL has limitations in its ability to express features of the past and future. By adding temporal logic to DEL in a non-probabilistic setting, the paper [13] captures both some of the temporal flexibility of the system of runs as well as the mechanical method offered by DEL of going from one stage in time to the next. One goal of this paper is to involve probability in the combination of tempo-
ral logic and DEL, focusing on the inclusion of a previoustime operator and exploring the possibility of completeness.

Another goal is to go beyond public announcements. There are other forms of information exchange that are of interest, such as semi-private announcements, where the fact that a message was sent to someone is not a secret, and completely-private announcements, where nonrecipients of the message are completely unaware of the fact that there is a message at all. An illuminating example involving semi-private announcements is given in [3]. The sequence of events provides a context motivating why there are stages in time in which an agent's sample space should differ from the set of states the agent considers possible. A non-probabilistic mechanical method for changing beliefs according to semi-private announcements was given in [2] and [1]. There, semi-private announcements are encoded in action models, and a product is defined between a model and an action model to produce an updated model that encodes the updated beliefs. This goal then is to involve action models in probabilistic dynamic epistemic logic.

## 2 Probabilistic Public Announcement Logic with a Previous Time Operator

As the underlying structure for public announcement logic is the epistemic model, the underlying structure for probabilistic public announcement logic is the probabilistic epistemic model, which adds probability spaces to an epistemic model.

Definition 2.1 [Probabilistic Epistemic Model] Let $\Phi$ be a set of proposition letters, and I be a set of agents. A probabilistic epistemic model is a tuple $\mathbf{M}=(X,\{\xrightarrow{i}$ $\left.\}_{i \in \mathbf{I}},\|\cdot\|,\left\{\mathbf{P}_{i, x}\right\}\right)$, where

- $X$ is a set of "states" or "possible worlds"
- $\xrightarrow{i} \subseteq X^{2}$ is an epistemic relation for each agent $i \in \mathbf{I}$, that is $x \xrightarrow{i} y$ if $i$ considers $y$ possible from $x$
- || $\|$ is a function assigning to each proposition letter $p$ the set of states where it is true.
- $\mathbf{P}_{i, x}$ is a probability space for each agent $i$ and state $x$, that is $P_{i, x}=\left(S_{i, x}, \mathcal{A}_{i, x}, \mu_{i, x}\right)$, where
- $S_{i, x} \subseteq X$ is set called the sample space
- $\mathcal{A}_{i, x}$ is a $\sigma$-algebra over $S_{i, x}$ (that is, a collection of subsets of $S_{i, x}$ that is closed under complements and countable unions). We the sets in the $\sigma$-algebra "measurable sets".
- $\mu_{i, x}: \mathcal{A}_{i, x} \rightarrow[0,1]$ is a probability measure over $S_{i, x}$ (that is, $\mu_{i, x}\left(S_{i, x}\right)=1$ and for each countable collection $A_{1}, A_{2}, \ldots$ of pairwise disjoint sets in $\left.\mathcal{A}, \mu\left(\bigcup_{k=1}^{\infty} A_{k}\right)=\sum_{k=1}^{\infty} \mu\left(A_{k}\right)\right)$.

For the rest of this section, we restrict the set $X$ to be finite and the set $\mathcal{A}_{i, x}$ to be the power set $\mathcal{P}\left(S_{i, x}\right)$. Thus we need not specify the $\sigma$-algebra $\mathcal{A}_{i, x}$ until the next section. It is recommended that $S_{i, x} \subseteq\{z: x \xrightarrow{i} z\}$, as every outcome in the sample space is a state the agent considers possible. For technical convenience in definition 2.2 , we will not impose such a restriction. One might assume that the converse of the recommendation should hold too, thus making $i$ 's sample space $S_{i, x}$ equal to the set $\{z: x \xrightarrow{i} z\}$ of states $i$ considers possible, but the example in the beginning of the next section motivates why we prefer not to make this restriction either. The example presents a situation in which an agent does not know enough to assign a probability to everything she considers possible, and although there are different ways of handling the uncertainty about the probability, omitting some states from the sample space is an attractive solution. As sample spaces are defined for each state, the agent may still be uncertain about which sample space is correct.

Public announcement logic is concerned with how an agent revises his/her beliefs given new information, knowing that this information is received by all other agents. Of greatest interest is new information that is consistent with the agent's beliefs, and PAL provides an enlightening mechanical procedure called an update for producing a new epistemic model from an old one given the new information. But although the updates provide a reasonable method of revising beliefs upon consistent information, they do not upon inconsistent information. The goal of probabilistic public announcement logic is to provide an update procedure that shows how to produce a new probabilistic epistemic model from an old one given information that is not only consistent with the agents' beliefs, but is also given positive probability. The case where the probability of the new information is 0 poses difficulties, and the goal of the definition of such a case is more to provide technical convenience than to offer a realistic result.

Definition 2.2 [updates] Given a probabilistic epistemic model $\mathbf{M}=\left(X,\{\stackrel{i}{\rightarrow}\},\|\cdot\|,\left\{\mathbf{P}_{i, x}\right\}\right)$ and a subset $Y$ of $X$, the update of $\mathbf{M}$ given $Y$ is written $\mathbf{M} \otimes Y$ and is the model $\left(X^{\prime},\{\xrightarrow{i}\}^{\prime},\|\cdot\|^{\prime},\left\{\mathbf{P}_{i, x}^{\prime}\right\}\right)$, where

- $X^{\prime}=X \cap Y$
- $x \xrightarrow{i}{ }^{\prime} y$ iff $x, y \in Y$ and $x \xrightarrow{i} y$
- $\|p\|^{\prime}=\|p\| \cap Y$
- If $\mu_{i, x}(Y)=0$, then let $\mathbf{P}_{i, x}^{\prime}$ be the only probability space definable on the singleton $x$. Otherwise, let $\mathbf{P}_{i, x}^{\prime}$ be defined by

$$
\begin{aligned}
& \text { - } S_{i, x}^{\prime}=S_{i, x} \cap Y \\
& \text { - For each subset } Z \subseteq Y, \mu_{i, x}^{\prime}(Z)= \\
& \mu_{i, x}(Z) / \mu_{i, x}(Y)
\end{aligned}
$$

This definition differs from the one in [8] in that here updating removes states while in [8] it removes relational connections but not states. Probability is updated in the same way as long as the set $Y$ has positive probability. The reason for these differences is to aid in proving completeness for the language that adds a previous time operator to probabilistic public announcement logic. Although completeness is still under construction (and the semantics in [8] may later demonstrate itself to be easier), the following discussions about the previous time operator may provide some intuition for why this new semantics may help, particularly for the conditions in definition 2.4. ${ }^{1}$

A natural choice for semantics that includes a previous time operator for this language is to involve structures that consist of a list of all past and present models. This is what was done in [13]

Definition 2.3 [History] A history $H$ is a list of models $\left(\mathbf{M}_{0}, \mathbf{M}_{1}, \ldots, \mathbf{M}_{n}\right)$, where for each $k, \mathbf{M}_{k}=\left(X_{k}, \stackrel{i}{\rightarrow}{ }_{k}\right.$ $\left.\},\|\cdot\|_{k},\left\{\mathbf{P}_{k i, x}\right\}\right)$, and $M_{k+1}=M_{k} \otimes X_{k+1}$.

Given a history $H=\left(\mathbf{M}_{0}, \mathbf{M}_{1}, \ldots, \mathbf{M}_{n}\right)$, let $\widehat{P}(H)=$ $\left(\mathbf{M}_{0}, \mathbf{M}_{1}, \ldots \mathbf{M}_{n-1}\right)$ be the previous history, $\widehat{\mathbf{M}}(H)=$ $\mathbf{M}_{n}$ be the last (most recent) model in the list, and let $\widehat{X}(H)=X_{n}$. We may write $x \in H$ for $x \in \widehat{X}(H)$.

[^0]
## Language

Let $\Phi$ be a set of proposition letters and $\mathbf{I}$ be a set of agents. We define by mutual recursion a multi-sorted language $\mathcal{L}$ with sentences and terms for each agent. The sentences (also called formulas) are given by

$$
\varphi::=\operatorname{true}|p| \neg \varphi\left|\varphi_{1} \wedge \varphi_{2}\right| \square_{i} \varphi\left|\left[\varphi_{1}\right] \varphi_{2}\right| t_{i} \geq q \mid \bar{Y} \varphi
$$

where $t$ is a term, $p \in \Phi$, and $i \in \mathbf{I}$.
The terms for agent $i$ are given by

$$
t_{i}::=q P_{i}(\varphi) \mid t_{i}+u_{i}
$$

where $q \in \mathbb{Q}$ is a rational number, $t_{i}$ and $u_{i}$ are terms for agent $i$, and $\varphi$ is a sentence.

The semantics is defined by a function $\llbracket \cdot \rrbracket$ from formulas to functions $f$ that map each history $H$ to a subset of $\widehat{X}(H)$, the carrier set of the most recent model in $H$. Then

- 【true】 is the function that maps each $H$ to the whole set $\widehat{X}(H)$,
- $\llbracket \neg \varphi \rrbracket(H)=\widehat{X}(H)-\llbracket \varphi \rrbracket(H)$
- $\llbracket \varphi \wedge \psi \rrbracket(H)=\llbracket \varphi \rrbracket(H) \cap \llbracket \psi \rrbracket(H)$.
- $x \in \llbracket \square_{i} \varphi \rrbracket(H)$ if and only $y \in \llbracket \square_{i} \varphi \rrbracket(H)$ for every $y$ in which $x \xrightarrow{i} y$, where $\xrightarrow{i}$ is $i$ 's epistemic relation in $\widehat{\mathbf{M}}(H)$ (the most recent model in $H$ ).
- $x \in \llbracket\left[\varphi_{1}\right] \varphi_{2} \rrbracket(H)$ if and only if either $x \notin \llbracket \varphi_{1} \rrbracket(H)$ or $x \in \llbracket \varphi_{2} \rrbracket\left(H \otimes \llbracket \varphi_{1} \rrbracket(H)\right)$.
- $x \in \llbracket q_{1} P_{i}\left(\varphi_{1}\right)+\cdots+q_{n} P_{i}\left(\varphi_{n}\right) \geq q \rrbracket(H)$ if and only if $\left.\left.q_{1} \mu_{i, x}\left(\llbracket \varphi_{1} \rrbracket(H)\right)+\cdots+q_{n} \mu_{i, x} \llbracket \varphi_{n} \rrbracket(H)\right) \geq q\right)$.
- $x \in \llbracket \bar{Y} \varphi \rrbracket(H)$ if and only if $H=(\mathbf{M})$ has just one model or $x \in \llbracket \varphi \rrbracket(\widehat{P}(H))$.

We have the usual modal abbreviations, such as $\diamond_{A} \varphi \equiv$ $\neg \square_{A} \neg \varphi$ and $\langle\psi\rangle \varphi \equiv \neg[\psi] \neg \varphi$, and we let $\widehat{Y} \varphi \equiv \neg \bar{Y} \neg \varphi$, which asserts that there is a previous time and $\varphi$ is true then. Here are some abbreviations in the language that express a variety of inequalities and equality.

- $t \leq q \equiv-t \geq-q$
- $t<q \equiv \neg(t \geq q)$
- $t>q \equiv \neg(t \leq q)$
- $t=q \equiv t \leq q \wedge t \geq q$.
- $t \geq s \equiv t-s \geq 0$
- $t=s \equiv t-s \geq 0 \wedge s-t \geq 0$


## Proof system

Include axioms of proposition logic together with the following:


Include axioms of proposition logic together with the following:

| $\square_{i}$-necessitation | From $\vdash \varphi$ infer $\vdash \square_{i} \varphi$ |
| :--- | :--- |
| $[\varphi]$-necessitation | From $\vdash \varphi$ infer $\vdash[\varphi] \varphi$ |
| $\bar{Y}$-necessitation | From $\vdash \varphi$ infer $\vdash \bar{Y} \varphi$ |
| Equivalence |  |
| $\left.\quad \begin{array}{l}\text { From } \vdash \varphi \leftrightarrow \psi, \\ \end{array}\right)$ |  |

## Approach to completeness

A general strategy for proving weak completeness is to start with a consistent formula and then prove that it is satisfiable. Modal logic and probabilistic epistemic logic provide techniques for finding filtrations that may satisfy either the formula or a formula provably equivalent to the first. But such filtrations are single models, not lists of models. There is a similar difficulty for DEL to use filtrations, because the semantics of DEL involves the construction of a new model. But it turns out that in DEL, a semantics for a subset of formulas, each called normal form formulas, can be defined in which no new model needs to be constructed in order to determine whether a formula is true. In addition, each formula is provably equivalent to a normal form formula, and the two semantics relate to each other in a natural and convenient way: a formula is true given one semantics if and only if it is true given the other.

To employ a strategy similar to this, we need an alternative set of models and an alternative semantics. Let us define a non-standard model as a probabilistic epistemic model together with a binary relation $Y$. Ideally $x Y z$ can be read as " $x$ is one stage later than $z$ ", but this interpretation may be difficult to achieve unless there are some restrictions placed on this new non-standard model. We thus provide the following definition:

Definition 2.4 [non-standard history] Let

$$
\mathcal{M}=\left(X,\{\xrightarrow{i}\}_{i \in \mathbf{I}},\|\cdot\|,\left\{\mathbf{P}_{i, x}\right\}, Y\right)
$$

be a non-standard model. Define

$$
A_{0}=\{x: \text { there is no } z \text { such that } x Y z\}
$$

and for each $n>0$,

$$
\begin{aligned}
& A_{n}=\left\{x \text { :there is a } z \text { such that } x Y^{n} z\right. \\
& \left.\quad \text { and there is no } z \text { such that } x Y^{n+1} z\right\}
\end{aligned}
$$

Define for each set $A$ and binary relation $R$,

$$
R(A)=\{z: \text { there is a } x \in A \text { such that } x R z\}
$$

Then $\mathcal{M}$ is a non-standard history if the following conditions hold:

1. Partial functionality of $Y$ : if $x Y z$ and $x Y z^{\prime}$, then $z=$ $z^{\prime}$.
2. Bounded age: There exists $N$ such that for all $x$ there is no $z$ for which $x Y^{N} z$.
3. Epistemic synchronicity: if $x \xrightarrow{i} z$, then for each $n$, $x Y^{n} x^{\prime}$ for some $x^{\prime}$ iff $z Y^{n} z^{\prime}$ for some $z^{\prime}$.
4. Probabilistic synchronicity: if $x, z, w \in X$ and $x, z \in$ $S_{i, w}$, then for each $n, x Y^{n} x^{\prime}$ for some $x^{\prime}$ iff $z Y^{n} z^{\prime}$ for some $z^{\prime}$.
5. Update product relation condition a: if $x \xrightarrow{i} z$ and $z Y z^{\prime}$ then there exists $x^{\prime}$ such that $x Y x^{\prime}$ and $x^{\prime} \xrightarrow{i} z^{\prime}$.
6. Update product relation condition b: if $x Y x^{\prime}, x^{\prime} \xrightarrow{i} z^{\prime}$, and $z Y z^{\prime}$, then $x \xrightarrow{i} z$.
7. Update product sample space condition a: for each $n \geq 1, i \in \mathbf{I}, x \in A_{n}$, and $z$ such that $x Y z$, if $\mu_{i, z}\left(Y\left(A_{n}\right)\right)>0$, then $Y\left(S_{i, x}\right)=Y\left(A_{n}\right) \cap S_{i, z}$.
8. Update product sample space condition b : for each $n \geq 1, i \in \mathbf{I}, x \in A_{n}$, and $z$ such that $x Y z$, if $\mu_{i, z}\left(Y\left(A_{n}\right)\right)=0$, then $S_{i, x}=\{x\}$.
9. Update product probability condition a: for each $n \geq 1, i \in \mathbf{I}, x \in A_{n}$, and $z$ such that $x Y z$, if $\mu_{i, z}\left(Y\left(A_{n}\right)\right)>0$, then for each $A \subseteq S_{i, x}$,

$$
\mu_{i, x}(A)=\frac{\mu_{i, z}(Y(A))}{\mu_{i, z}\left(Y\left(A_{n}\right)\right)}
$$

10. Update product probability condition b : for each $n \geq 1, i \in \mathbf{I}, x \in A_{n}$, and $z$ such that $x Y z$, if $\mu_{i, z}\left(Y\left(A_{n}\right)\right)=0$, then $\mu_{i, x}(\{x\})=1$.
11. Update product valuation condition: if $x Y z$ then $x \in$ $\|p\|$ iff $z \in\|p\|$

Semantics can be defined for formulas that do not include public announcement operators $[\varphi]$. These formulas constitute a language which we call normal form. The operator $\bar{Y}$ is treated as the box modality for the relation $Y$. The semantics for the other operators remain the same.

To show that every formula is provably equivalent to one in normal form, we employ a term rewriting system similar to one used in [1]. Our term rewriting system will make use of the following algebraic semantics.

Definition 2.5 [Signature] We define $\Delta$ to be the following signature. It is multi-sorted, with one sort for sentence terms $s$ and another for weight terms $t_{i}$ for each agent $i \in \mathbf{I}$. Here are the symbols in the signature:

1. Each $p \in \Phi$ and true is a constant symbol of sort $s$.
2. $\neg, \square_{A}, \square_{\mathcal{B}}^{*}, \bar{Y}$ are function symbols of type $s \rightarrow s$
3. $\wedge, \rightarrow$, and [ ] are binary function symbols of type $s \times$ $s \rightarrow s$
4. $P_{i}$ is a function of type $\mathbb{Q} \times s \rightarrow t_{i}$
5. $+_{i}$ is a function of type $t_{i} \times t_{i} \rightarrow t_{i}$
6. $\geq_{i}$ is a function of type $t_{i} \times \mathbb{Q} \rightarrow s$
7. $\operatorname{triv}_{i}$ is a function of type $t_{i} \rightarrow t_{i}$
8. bay $_{i}$ is a function of type $s \times t_{i} \rightarrow t_{i}$

We will in general write $[s] s$ for []$(s, s)$, and we choose to write $+_{i}$ and $\geq_{i}$ in infix notation too. In addition, we often drop the subscripts when it is understood from context. The functions bay and triv are just tools for reducing formulas of the form $[x]\left(t \geq_{i} q\right)$ in the next definition. The choice of symbol bay is supposed to indicate a relationship to Bayesian updating, and the choice of the symbol triv is to indicate that the probabilities are trivialized (that is, we will take the probability of true).

Let $\mathcal{L}^{+}$be the algebraic language defined by this signature, and let $\mathcal{L}^{+}(X)$ be the language $L$ augmented with a set of variables $X$. Syntactically, occurrences of a variable must agree on the sort, that is, $x+{ }_{i} x$ implies that $x$ is a weight term for $i$, and $x \geq_{i} x$ is not allowed, since the first occurrence of $x$ would have to be a weight term and the second a sentence term.

A term rewriting system is a collection of rewrite rules, written $\varphi \rightsquigarrow \psi$, where $\varphi, \psi \in \mathcal{L}^{+}(X)$. Executing a rewrite rule on a formula $\chi$ would identify a substitution instance of $\varphi$ in $\chi$ and replace it with a substitution instance of $\psi$ (using the same assignment of variables to terms). For our purposes, we use the following rewrite system.

Definition 2.6 [Rewriting system $\mathcal{R}$ ] Here is a rewriting system of use to us

| (r1) | $x \rightarrow y$ | $\rightsquigarrow$ | $\neg(x \wedge \neg y)$ |
| :---: | :---: | :---: | :---: |
| (r2) | $[x]$ true | $\rightsquigarrow$ | true |
| (r3) | $[x] p$ | $\rightsquigarrow$ | $x \rightarrow p$ |
| (r4) | $[x] \neg y$ | $\rightsquigarrow$ | $x \rightarrow \neg[x] y$ |
| (r5) | $[x](y \wedge z)$ | $\rightsquigarrow$ | $[x] y \wedge[x] z$ |
| (r6) | $[x] \square_{A} y$ | $\rightsquigarrow$ | $x \rightarrow \square_{A}[x$ |
| (r7) | $[x] \bar{Y} z$ | $\rightsquigarrow$ | $x \rightarrow$ |
| (r8) | $\operatorname{triv}\left(P_{i}(q, x)\right)$ | $\rightsquigarrow$ | $P_{i}(q$, true $)$ |
| (r9) | $\operatorname{triv}\left(t_{1}+t_{2}\right)$ | $\rightsquigarrow$ | $\operatorname{triv}\left(t_{1}\right)+$ |
| (r10) | $\operatorname{bay}\left(x, P_{i}(q, z)\right)$ | $\rightsquigarrow$ | $P_{i}(q, x \wedge[x]$ |
| (r11) | $\operatorname{bay}\left(x, t_{1}+t_{2}\right)$ |  | $\left.x, t_{1}\right)+ \text { bay }$ |
| (r12) | $\begin{array}{r} {[x]\left(t \geq_{i} q\right)} \\ \left(P_{i}(-1, x\right. \\ \quad(\neg) \\ (x \rightarrow \end{array}$ | $\begin{aligned} & \rightsquigarrow \\ & 0 \wedge \\ & -1, x \\ & (x, t) \end{aligned}$ | $\begin{aligned} & \rightarrow(\operatorname{triv}(t) \geq \\ & 0) \wedge \\ & P_{i}(-q, x) \end{aligned}$ |

These rules correspond to either biconditional axioms schema or provable biconditiionals, which is the core reason why a rewritten formula is provably equivalent to the first. Note that there is a natural translation between our original language $\mathcal{L}$ and the algebraic language $\mathcal{L}^{+}$. The rewriting is done in $\mathcal{L}^{+}$and the provable equivalence is determined between corresponding formulas in $\mathcal{L}$.

But it is also important that we can apply rules finitely many times in order to obtain a term corresponding to a formula in normal form. We first observe that no rule can be applied to terms if and only if the terms correspond to formulas in in normal form. But we must also show that only finitely many applications of the rules can be applied, something that we may doubt, given that rule (r12) appears to produce a much more complicated term. But the following interpretation of symbols in the signature can help us show that the rewriting system terminates.

Definition 2.7 [Interpretation of Signature] Let us overload the symbol $\llbracket \cdot \rrbracket$ to indicate interpretation. Our signature has a carrier $\mathbb{N}_{\geq 3}$ for sentences, and a carrier $\mathbb{N}_{\geq 3}$ for actions. The function symbols are then interpreted as the following arithmetic functions on these numbers:

We recursively extend this interpretation to all terms and sentences.

It turns out that every application of a term rewriting rule results in a term with a strictly smaller interpretation. Thus
the interpretation of the term is an upper bound to the number of times rules can be applied before terminating.

To show completeness, we start with a consistent formula in $\mathcal{L}$, translate it into $\mathcal{L}^{+}$, and then apply rewrite rules until we obtain a term in which no more rules can be applied. The result of translating this term back into $\mathcal{L}$ is a normal form formula provably equivalent to the original formula. We then show for every non-standard history $\mathcal{H}$ and state $x$ in the non-standard history, there is an actual history $H$ and a state in that history that agrees with $x$ on the truth of every normal form formula. We then form a filtration for the non-standard semantics. A number of model transformations will likely be needed to turn the filtration into a history. One transformation that might be useful is an unravelling (or partial unravelling) about a point in the canonical model that satisfies the consistent formula. This was done in [12]. But producing a discrete probability after unravelling is not always straightforward or possible. Thus it would be helpful to get a better grasp of issues regarding updating probabilistic epistemic models that have unmeasurable sets (sets not in the $\sigma$-algebra). This is the greatest challenge discussed in the next section concerning a dynamic probabilistic epistemic logic.

## 3 Involving Action Models

The example in [3] is as follows. There are two agents: $i$ and $k$. Agent $k$ receives a bit: 0 or 1 . Agent $i$ is aware that $k$ learns what the bit is, but $i$ does not know what the bit is. Then agent $k$ flips a fair coin, and observes the result. Again $i$ is aware that $k$ learns the result of the flip, but does not learn the result. Viewing heads as 1 and tails as 0 , agent $k$ performs action $s$ if the coin agrees with the bit, and $d$ if it does not.

Fagin and Halpern viewed this experiment through a system of runs. There are 4 possible runs of this example based on the outcome of the bit together with the outcome of the coin. The action $d$ or $s$ is determined from the first two outcomes. Let us consider 4 states, one for each run: $(1, H),(1, T),(0, H),(0, T)$. Before agent $k$ performs action $s$ or $d$, agent $i$ considers all four possible. But what should agent $i$ 's probability space be at each state? Three possibilities are discussed in [3] and are depicted below. The solid box depicts what $i$ would consider to be the sample spaces from each state within its borders, and the dotted lines depict the smallest non-empty sets in the $\sigma$-algebra for $i$ from each element in the sample space. Note that $i$ 's probability spaces from one state to another need not always differ.


The diagram on the left depicts the situation where the sample space should consist of all 4 states, thus making the sample space the same as the set of states considered possible. Notice that the set $\{(1, H),(1, T)\}$ is not measurable (that is, is not in the $\sigma$-algebra). Unlike a fair coin, we do not know enough about how the bit is to be assigned to give it a probability. But even without knowing the probability of the bits 1 and 0 , we can determine that $s$ (represented by $\{(1, H),(0, T)\})$ has probability $1 / 2$, and similarly for $d$. So the second diagram provides model where $i$ can give a probability to $s$, but still does not give a probability to bits 1 or 0 . The set of states $i$ considers possible is still all four states, but now the sample spaces are different from that. Thus $i$ can also be uncertain about whether the bit turned out to be 0 or 1 , but can still determine that the probability of $s$ is $1 / 2$ in each case, thus concluding that the probability is $1 / 2$. In the third diagram, agent $i$ believes the probability of any outcome is either 1 or 0 , but does not know which.

It is suggested that these three diagrams may be viewed as three different stages of the example. The first diagram would correspond to the time before the bit is given. The second diagram would correspond to the time after the bit is given but before the coin is flipped. The third diagram would correspond to the time after the coin is flipped. The transition from one stage to the next might make more sense if we view the probabilities as objective. Indeed, these diagrams would just as appropriately represent $k$ 's probabilities at the three stages. But one consequence of using these objective probabilities is that $i$ 's degree of certainty about the probabilities changes as a result of learning that $k$ was informed of something. From the first diagram to the second, $i$ becomes more certain about the probability of events $s$ and $d$, but in both steps, $i$ becomes less certain about the probability space. Being more sensitive to probabilities as subjective, we may prefer that there is more information revealed to $i$ between stages, to help the change in the degree of certainty. Suppose that at first $i$ is not aware of any plan for $k$ to perform either action $s$ or $d$, and hence does not wonder about the probabilities of these two events. The sequence of actions may be as follows:

1. $k$ receives the bit.
2. $k$ informs $i$ of his plan to perform action $s$ if the result of the coin matches the bit and $d$ if it does not match.
3. $k$ observes the outcome of the coin.
4. $k$ offers $i$ a bet that the result was heads.

The first diagram would correspond to $i$ 's probabilities before any of these actions have taken place. The second diagram would correspond to $i$ 's probabilities after the second action, but before the third. The third diagram would correspond to $i$ 's probabilities after all four actions have taken place. Perhaps action two, where $k$ informs $i$ of the plan to perform either $s$ or $d$ prompts $i$ to rework his probabilities so that $s$ and $d$ are assigned probabilities. The contribution of the fourth action, where $k$ offers $i$ a bet, may be more convincing. Now $i$ considers it highly unlikely that $k$ would have offered the bet knowing that he would lose, yet $i$ could not have a quantitive grasp of this likelihood, and thus could not assign a probability.

Now in terms of probabilistic epistemic logic, the first model poses a difficulty; if there is any formula for the bit 1 , the bit 0 , the action $s$, or the action $d$, then the set of states making that formula true is not measurable. A temporary fix is proposed in [3], which is to use the inner measure function. If $(S, \mathcal{A}, \mu)$ is a probability space, then the inner measure of $\mu$ is $\mu_{*}$ defined by

$$
\mu_{*}(T)=\sup \{\mu(A): A \in \mathcal{A}, A \subseteq T\}
$$

for each $T \subseteq S$. We could alternatively use the outer measure:

$$
\mu^{*}(T)=\inf \{\mu(A): A \in \mathcal{A}, T \subseteq A\}
$$

for each $T \subseteq S$. Inner and outer measures are related according to $\mu^{*}(T)=1-\mu_{*}(\bar{T})$, where $\bar{T}$ is the complement of $T$ in $S$. Thus either the inner or outer measure can be taken as primitive in the language, and the other can be defined according to the other. Either the inner or outer measure lets us define the semantics of formulas such as $P_{i}(\varphi) \geq 1 / 2$ even when the set of states making $\varphi$ true is unmeasurable. But as the inner measure and the outer measure need not be a measures themselves, the probability axioms would fail. One suggestion given in [3] is that we explicitly require that the $\sigma$-algebra be large enough to contain all sets corresponding to each formula. But nonmeasurable cases have been considered with a more relaxed set axioms

In defining an update product between a probabilistic epistemic model and an action model, we shall first consider the general case, where sets of states that make the formulas true need not be measurable. In particular, the formulas inducing the update might not correspond to measurable sets (the elements of the $\sigma$-algebra). We thus define an "outer (or inner) probability dynamic epistemic logic". To make this task manageable, we restrict the models to be finite.

Definition 3.1 [action model] An action model $(\Sigma,\{\xrightarrow{i}$ $\},\left\{\mathcal{P}_{i, \sigma}\right\}$, pre) is a probabilistic epistemic model with the valuation function $\|\cdot\|$ replaced by a function pre which assigns to each $\sigma \in \Sigma$ a function that assigns to each probabilistic epistemic model a subset of the carrier set of that model. Each element $\sigma \in \Sigma$ is called an action type. $\triangleleft$

We define the update product between a probabilistic epistemic model and an action model in two stages. We first define the product between the original probabilistic epistemic model and an action signature (which is just a probabilistic epistemic frame (no valuation, and without the pre function), and then relativize the result according to the pre function. The first product is called the unrestricted product. The second is called the relativization.

Definition 3.2 [unrestricted product] The unrestricted product between a probabilistic epistemic model $\mathbf{M}$ and an action model $\boldsymbol{\Sigma}$ is $\mathbf{M} \otimes_{U} \boldsymbol{\Sigma}$ with the following components:

1. $X_{\otimes}=X \times \Sigma$
2. $(x, \sigma) \xrightarrow{i}(z, \tau)$ iff $x \xrightarrow{i} z$ and $\sigma \xrightarrow{i} t$
3. $\|p\|_{\otimes}=\|p\| \times \Sigma$
4. We define $\mathcal{P}_{i,(x, \sigma)}$ as follows:
(a) The sample space is the Cartesian product $S_{i,(x, \sigma)}=S_{i, x} \times S_{i, \sigma}$
(b) The $\sigma$-algebra $\mathcal{A}_{i,(x, \sigma)}$ is the smallest $\sigma$-algebra containing

$$
\left\{A \times B: A \in \mathcal{A}_{i, x}, B \in \mathcal{A}_{i, \sigma}\right\}
$$

(c) The probability measure is defined as

$$
\mu_{i,(x, \sigma)}(A)=\sum_{k=1}^{n} \mu_{i, x}\left(B_{k}\right) \mu_{i, \sigma}\left(C_{k}\right)
$$

where $B_{k} \in \mathcal{A}_{i, x}, C_{k} \in \mathcal{A}_{i, \sigma}$, and $A=$ $\biguplus_{i=1}^{n} B_{k} \times C_{k}$

This product is a probabilistic epistemic model. The usual definition for product measures (for finite spaces) is given to our new probability measure. Product measures need not be restricted to finite spaces, and hence this unrestricted product can be defined between any probabilistic epistemic model and action model of infinite size.

But the relativization of our probabilistic epistemic model requires some restriction be placed on the probabilistic epistemic model. Requiring the carrier set of the probabilistic epistemic model to be finite is sufficient and still allows us to explore a wealth of examples.

Definition 3.3 [relativization] The relativization of a probabilistic epistemic model $\mathbf{M}$ to $Y \subseteq X$ is given by $\mathbf{M} \otimes_{R} Y$ with the following components:

1. $X_{Y}=Y$
2. $x \xrightarrow{i}_{Y} z$ iff $x \xrightarrow{i} z$ and $x, z \in Y$
3. $\|p\|_{Y}=\|p\| \cap Y$
4. For $x \in Y$, if $\mu_{i, x}^{*}(Y)=0$, then define $\mathcal{P}_{i, x}$ to be the trivial probability space on the singleton $x$. Otherwise
(a) $S_{Y i, x}=S_{i, x} \cap Y$
(b) $\mathcal{A}_{Y i, x}$ is the $\sigma$-algebra generated by $\{A \cap Y$ : $\left.A \in \mathcal{A}_{i, x}\right\}$
(c) The probability measure is defined by

$$
\mu_{Y i, x}(A)=\frac{\mu_{i, x}^{*}(B)}{\mu_{i, x}^{*}(Y)}
$$

The choice to update using outer measures rather than inner measures is mostly arbitrary. The outer measure, however is less likely to be zero. When the $\sigma$-algebra $\mathcal{A}$ of a space
$(S, \mathcal{A}, \mu)$ is finite, the outer measure of $\mu$ applied to a set $T \subseteq S$ becomes

$$
\begin{aligned}
\mu^{*}(T) & =\bigcap\{\mu(A): A \in \mathcal{A}, T \subseteq A\} \\
& =\mu(\bigcap\{A: A \in \mathcal{A}, T \subseteq A\})
\end{aligned}
$$

Thus the outer measure of a (not necessarily measurable) set is equal to the measure of an appropriate measurable set. This is not guaranteed in the infinite case. But this property helps us guarantee that the updated function is indeed a measure. The most difficult case is the additivity condition. If $A_{1}, \ldots, A_{n}$ is a set of pairwise disjoint sets measurable in the relativized model, let $\hat{A}_{i}=\bigcap\left\{B: A_{i} \subseteq B, B \in \mathcal{A}_{i, x}\right\}$, where $\mathcal{A}_{i, x}$ is the $\sigma$-algebra for the original model. Unlike the $A_{i}$, the $\hat{A}_{i}$ are necessarily measurable in the first space. Also $\hat{A}_{1}, \ldots, \hat{A}_{n}$ is pairwise disjoint, for if $Y$ is the set with which we relativized, then $B=\hat{A}_{j} \cap \hat{A}_{k} \subseteq \bar{Y}$ (otherwise $A_{j}$ and $A_{k}$ would not be disjoint). But $A_{j} \subseteq \hat{A}_{j}-B$ and $\hat{A}_{j}-B \in \mathcal{A}_{i, x}$, thus $\hat{A}_{j}=\hat{A}_{j}-B$, and so we conclude that $B=\emptyset$. Also observe that $\widehat{\bigcup A_{i}}=\bigcup \hat{A}_{i}$. We can then make use of this and the fact that $\mu^{*}(C)=\mu(\hat{C})$ for any set $C$ in order to establish the additivity property of the new measure.

Definition 3.4 [update product] Let $\boldsymbol{\Sigma}=(\Sigma,\{\xrightarrow{i}$ $\}_{i \in \mathbf{I}},\left\{\mathcal{P}_{i, x}\right\}$, pre) be an action model and $\mathbf{M}=(X,\{\xrightarrow{i}$ $\left.\},\|\cdot\|,\left\{\mathcal{P}_{i, x}\right\}\right)$. Let $Y=\{(x, \sigma): x \in \operatorname{pre}(\sigma)(\mathbf{M})\}$. The update product between $\mathbf{M}$ and $\boldsymbol{\Sigma}$ is written $\mathcal{M} \otimes \boldsymbol{\Sigma}$ and is defined as $\left(\mathbf{M} \otimes_{U} \boldsymbol{\Sigma}\right) \otimes_{R} Y$.

Returning to the example above, the action of revealing the bit to $k$ in such a way that $i$ knows $k$ learned something is a semi-private announcement. Similarly, $k$ 's learning the result of the flip in such a way that $i$ knows $k$ learned something is also a semi-private announcement. The relational part of the action signature for semi-private announcements may be depicted by the following diagram:

$$
\begin{aligned}
& i, j \subset \sigma \longleftrightarrow \quad i \quad(\tau) i, j \\
& R_{1}
\end{aligned}
$$

From each of the two action types, $i$ 's probability space is the only probability space where the sample space is that single action type. This action signature may be used for the action models of both stages. For the first stage, the precondition of $\sigma$ is 1 , and the precondition of $\tau$ is 0 . For the second action signature, the precondition of $\sigma$ could be $H$, while the precondition for $\tau$ could be $T$.

But what should be the probability spaces of the action model? Let us assume that one action model will capture both the semi-private announcement of the bit to $k$ and announcement that $k$ plans to do either $s$ or $d$. Then both $i$ and
$k$ 's probability spaces in the action model could be $i$ and $k$ 's probability spaces:


From each of the action types, the probability space is the only space that can be defined over a sample space with one element. Dotted ovals are therefore not needed. That $i$ and $k$ share the same probability spaces agrees with the view that probabilities should be objective.

But let us consider what happens if we break down the transition from $M_{1}$ to $M_{2}$ into two steps (giving us an intermediate model) and similarly break down the transition from $M_{2}$ to $M_{3}$ into two steps. The action where $k$ is informed of the bit will still be considered a semi-private announcement, and the relational structure will be the same. The only difference shall be $i$ 's probability space, which we change to the following structure:


The action where $i$ is informed that $k$ plans to do either action $s$ or $d$ will serve the purpose of splitting $i$ 's probability space into two. This can be done by using the probability structure $P_{1}$. We will also use $P_{1}$ for $k$ 's probability structure. As $k$ 's probability spaces are already split, using $P_{1}$ for $k$ as well will not affect $k$ 's probability structure in the update model. For the relational structure we use


As the precondition of $\sigma$ is the bit 1 and the precondition of $\tau$ is the bit 0 , the updated model will have the same relational structure as it did right before updating.

We use the same relational and probabilistic structures in the actions model from $M_{2}$ to $M_{3}$, but we use different preconditions. We may let the precondition for $\sigma$ be $H$ rather than 1 , and we may let the precondition for $\tau$ be $T$ rather than 0 .

When we consider the language and semantics, we may wish that every formula correspond to a measurable set. But even this might not guarantee that the set $Y$ in definition 3.3 is itself measurable. Consider an action signature for which $k$ 's probability space is given by diagram $P_{2}$, that is, there are only two measurable sets: the whole set and the empty set. Suppose a probabilistic epistemic model M
has two states: $x$ and $y$, and $k$ 's probability sample space is $\{x, y\}$ and all subsets are measurable. Then in the product measure, the measurable sets are

$$
\begin{aligned}
& \{\emptyset,\{(x, \sigma),(y, \sigma)\},\{(x, \tau),(y, \tau)\} \\
& \{(x, \sigma),(y, \sigma),(x, \tau),(y, \tau)\}\}
\end{aligned}
$$

Suppose there were a formula $\varphi$ for which only $x$ is true, and another formula for which $y$ is true. Then these formulas correspond to measurable sets. Let the function pre reflect these two formulas, by defining $\operatorname{pre}(\sigma)(\mathbf{M})=x$ and $\operatorname{pre}(\tau)(\mathbf{M})=y$. Then when taking the full update product, we would be relativizing with respect to the set $Y=\{(x, \sigma),(y, \tau)\}$, which is not measurable. In general, if an action signature has only discrete probability spaces (probability spaces where the $\sigma$-algebras are power sets of the sample spaces), then the measurability of the sets $\operatorname{pre}(\sigma)(\mathbf{M})$ for each $\sigma \in \Sigma$, does guarantee that the set $Y$ in definition 3.3 is measurable. It remains to be seen that in an updated model, every formula still corresponds to a measurable set.

## 4 Conclusion

This paper is a synthesis of two related projects. One is to add a previous time operator to a probabilistic dynamic epistemic logic similar to the one given in [8], and the other is to involve action models and update products in a probabilistic dynamic epistemic logic. Although it appears that these projects are independent, the second project may help support the first. We have so far approached the first project with the initial goal of maintaining simplicity in hope that technical results will be easier to achieve. But sometimes extra structure makes it easier to prove certain results, and we have yet to see if the involvement of unmeasurable sets will facilitate the completeness proof of a probabilistic dynamic epistemic logic with a previous time operator.

The second project explores the possibility and motivation of non-discrete probability measures and updating based on non-measurable sets. We have seen one way to update finite probabilistic epistemic models that are not necessarily discrete upon finite action models that are not necessarily discrete in order to yield a new finite probabilistic epistemic model. Although this updating can guarantee that the updated model is indeed a probabilistic epistemic model, it does not guarantee that in the updated model, all the formulas correspond to measurable sets; we have yet to see which conditions would ensure the updated model does have that property. This is only a concern if we wish to enforce additivity axioms of probability. Otherwise we may have the machinery for a nice inner (or outer) probability dynamic epistemic logic.

As this update product is quite flexible, with nonmeasurable sets in both probabilistic dynamic epistemic
models and action models, questions open up as to how to interpret particular instances of updating. We have looked at an example in [3] to help us with this. While doing so, we distinguished between subjective and objective probabilities and considered breaking down each action into two. The second action for each only affects the probability spaces being updated, and does not affect the structure of the epistemic relations. There so far is no language for this probabilistic dynamic epistemic logic with action models and update products, and in coming up with a language, we should determine what fundamentally is driving this change in probability spaces. May the source of information play an important role, as suggested by the phrasing of the action " $k$ offers $i$ a bet that the result was heads"? What is the essential component of the action phrased " $k$ informs $i$ of his plan to perform action $s$ if the result of the coin matches the bit and $d$ it does not match"? I suggest that in future work, finding more examples will help reveal underlying patterns that will enable us to adequately answer these questions.

## References

[1] A. Baltag, L. Moss, and S. Solecki. Logics for epistemic actions: Completeness, decidability, expressivity. ms. Indiana University, 2003.
[2] A. Baltag and L. S. Moss. Logics for epistemic programs. Synthese, 139(2, Knowledge, Rationality \& Action):165224, 2004.
[3] R. Fagin and J. Halpern. Reasoning about knowledge and probability. Journal of the ACM, 41(2):340-367, March 1994.
[4] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. Reasoning About Knowledge. The MIT Press, 1995.
[5] T. French, R. van der Meyden, and M. Reynolds. Axioms for logics of knowledge and past time: Synchrony and unique initial states. Proceedings of Advances in Modal Logic, 2005.
[6] D. Gerbrandy. Bisimulations on Planet Kripke. PhD thesis, ILLC, Universiteit van Amstardam, 1998.
[7] J. Halpern and M. Tuttle. Knowledge, probability, and adversaries. Journal of the ACM, 40(4):917-962, 1993.
[8] B. P. Kooi. Probabilistic dynamic epistemic logic. J. Logic Lang. Inform., 12(4):381-408, 2003. Special issue on connecting the different faces of information.
[9] J. Miller and L. Moss. Undecidability of iterated modal relativization. Studia Logica, 79(3):373-407, April 2005.
[10] R. Parikh and R. Ramanujam. A knowledge based semantics of messages. J. Logic Lang. Inform., 12(4):453-467, 2003. Special issue on connecting the different faces of information.
[11] J. Plaza. Logics of public communications. Synthese, 158(2, Knowledge, Rationality \& Action):165-179, 2007.
[12] J. Sack. Logic for update products and steps into the past. ms., 2007.
[13] J. Sack. Temporal languages for epistemic programs. Journal of Logic, Language and Information, 17(2):183-216, April 2008.
[14] A. Yap. Product update and looking backward. ms. www.illc.uva.nl/lgc/papers/bms-temporal.pdf, 2005.


[^0]:    ${ }^{1}$ Definition 2.4 makes use of sets $A_{n}$ consisting of all states corresponding to time $n$. It is helpful that this set is equal to the set $B_{n}$ consisting of all of the more recent versions of states that satisfied the formula that induced the update. Finding an appropriate characterization of $B_{n}$ so far appears more difficult using the semantics of [8].

