Reasoning with Probabilities

Eric Pacuit Joshua Sack

July 28, 2009

Plan for the Course

Day 1: Introduction and Background

Day 2: Probabilistic Epistemic Logics

Day 3: Dynamic Probabilistic Epistemic Logics

Day 4: Reasoning with Probabilities

Day 5: Conclusions and General Issues

Probability language

Let Φ be a set of proposition letters.
Propositional Formulas:

$$F ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi$$

Terms:

$$t ::= aP(F) \mid t + t$$

where $a \in \mathbb{Q}$ and F is a propositional formula.
Formulas:

$$\varphi ::= t \geq r \mid \neg\varphi \mid \varphi \wedge \varphi$$

where $r \in \mathbb{Q}$ and t is a term.
This language is from:

- R. Fagin, J. Halpern, N. Megiddo (1990) Reasoning about Probabilities. *Information and Computation* 87:1, pp. 76–128.

Probability models and semantics

Let Φ be a set of proposition letters.
$M = (X, \mathcal{A}, \mu, \| \cdot \|)$, where

- $(X, \mathcal{A}, \mu)$ is a probability space
- $\| \cdot \| : \Phi \to \mathcal{P}(X)$

The semantics of propositional formulas is defined by a function $[\![\cdot]\!]$ from propositional formulas to subsets of X.

$$\begin{array}{lcl} [\![\top]\!] & = & X \\ [\![p]\!] & = & \|p\| \\ [\![\neg\varphi]\!] & = & X - [\![\varphi]\!] \\ [\![\varphi \wedge \psi]\!] & = & [\![\varphi]\!] \cap [\![\psi]\!] \end{array}$$

The semantics of probability formulas are defined by

$\models a_1 P(\varphi_1) + \cdots + a_n P(\varphi_n) \geq r$ iff
$a_1 \mu([\![\varphi_1]\!]) + \cdots + a_n \mu([\![\varphi_n]\!]) \geq r$.

A note about σ-algebras

- Probability spaces may be finite or infinite.
- Possible probability distributions may include the *uniform probability distribution* over an interval, in which case $\mathcal{A}$ cannot be $\mathcal{P}(X)$ (Recall Vitali sets).

Abbreviations

Let

$$\sum_{k=1}^{n} a_k P(\varphi_k) \equiv a_1 P(\varphi_1) + \cdots + a_n P(\varphi_n)$$

Then if $t = \sum_{k=1}^{n} a_k P(\varphi_k)$, let $bt = \sum_{k=1}^{n} b a_k P(\varphi_k)$

$$\begin{aligned}
t \leq r &\equiv -t \geq -r \\
t = r &\equiv (t \leq r) \wedge (t \geq r) \\
t > r &\equiv \neg(t \leq r) \\
t_1 \geq t_2 &\equiv t_1 - t_2 \geq 0 \\
t_1 \leq t_2 &\equiv t_1 - t_2 \leq 0 \\
t_1 = t_2 &\equiv t_1 - t_2 = 0
\end{aligned}$$

Proof system

- All propositional tautologies
- $P(\varphi) \geq 0$
- $P(\top) = 1$
- $P(\varphi \wedge \psi) + P(\varphi \wedge \neg\psi) = P(\varphi)$
- $P(\varphi) = P(\psi)$ whenever $\varphi \leftrightarrow \psi$ is a propositional tautology
- If $\vdash \varphi$ and $\vdash \varphi \rightarrow \psi$, then $\vdash \psi$.

In addition there are inequality axioms.

Inequality axioms

- (permutation)
  $a_1 P(\varphi_1) + \cdots + a_n P(\varphi_n) \geq r \rightarrow$
  $a_{j_1} P(\varphi_{j_1}) + \cdots + a_{j_n} P(\varphi_{j_n}) \geq r$
- (adding coefficients)
  $(\sum_{k=1}^{n} a_k P(\varphi_k) \geq r) \wedge (\sum_{k=1}^{n} b_k P(\varphi_k) \geq s) \rightarrow$
  $(\sum_{k=1}^{n} (a_k + b_k) P(\varphi_k) \geq (r + s))$
- (adding and deleting 0 terms)
  $(t \geq r) \leftrightarrow (t + 0P(\varphi) \geq r)$
- (multiplying by non-zero coefficient)
  $t \geq r \leftrightarrow at \geq ar$ whenever $a > 0$.
- (dichotomy)
  $t \geq r \vee t \leq r$
- (monotonicity)
  $t \geq r \rightarrow t > s$, whenever $r > s$.

Lemma for Completeness

- $\Phi = \{p_1, \ldots, p_n\}$ is set of proposition letters,
- $At(\Phi) = \{\bigwedge_{i=1}^n q_i \mid q_i \in \{p_i, \neg p_i\}\}$ is set of atoms.

Lemma

Let $t \geq r$ be a probability formula, and Φ a set of proposition letters containing all letters occurring in t. Let $At(\Phi) = \{\alpha_1, \ldots, \alpha_{2^n}\}$. Then there are rationals $a_1, \ldots, a_{2^n}$ such that $t \geq r$ is equivalent to $a_1 P(\alpha_1) + \cdots + a_{2^n} P(\alpha_{2^n}) \geq r$.

Let $At(\Phi, \varphi) = \{\alpha \in At(\Phi) \mid \vdash \alpha \to \varphi\}$. Then

$$P(\varphi) \equiv \sum_{\alpha \in At(\Phi,\varphi)} P(\varphi \wedge \alpha) \equiv \sum_{\alpha \in At(\Phi,\varphi)} P(\alpha).$$

The first equivalence comes from multiple applications of additivity proposition letter by proposition letter.

Completeness of Halpern's Probability Logic

Let φ be a formula. It is a Boolean combination of probability terms.

- Transform φ into disjunctive normal form: a disjunction of conjunctions of probability formulas.
- Consider a disjunct

$$\psi = (t_1 \geq r_1) \wedge \cdots \wedge (t_k \geq r_k) \\ \wedge \neg(t_{k+1} \geq r_{k+1}) \wedge \cdots \wedge \neg(t_m \geq r_m).$$

- Let $\Phi = \{p_1, \ldots, p_n\}$ be the set of proposition letters occurring in ψ
- Let $At = \{\delta_1, \ldots, \delta_{2^n}\}$ be the set of all atoms: conjunctions of n literals from Φ
- Each conjunct $t_i \geq r_i$ of ψ is equivalent to $a_{i,1}P(\delta_1) + \cdots + a_{i,2^n}P(\delta_{2^n}) \geq r_i$

System of inequalities

The disjunct ψ is equivalent to the following system of inequalities:

$$
\begin{array}{|rcrcl|}
\hline
a_{1,1}P(\delta_1)+ & \cdots & +a_{1,2^n}P(\delta_{2^n}) & \geq & r_1 \\
 & & & \vdots & \\
a_{k,1}P(\delta_1)+ & \cdots & +a_{k,2^n}P(\delta_{2^n}) & \geq & r_k \\
\hline
a_{k+1,1}P(\delta_1)+ & \cdots & +a_{k+1,2^n}P(\delta_{2^n}) & < & r_{k+1} \\
 & & & \vdots & \\
a_{m,1}P(\delta_1)+ & \cdots & +a_{m,2^n}P(\delta_{2^n}) & < & r_m \\
\hline
P(\delta_1)+ & \cdots & +P(\delta_{2^n}) & \geq & 1 \\
-P(\delta_1)- & \cdots & -P(\delta_{2^n}) & \geq & -1 \\
\hline
 & & P(\delta_1) & \geq & 0 \\
 & & & \vdots & \\
 & & P(\delta_{2^n}) & \geq & 0 \\
\hline
\end{array}
$$

Final step

Completeness follows from the fact that the logic can follow the along with the steps of a mathematical algorithm that checks whether a solution to the system of inequalities exists. If there were no solution, then the logic would prove false.

Probabilistic Epistemic Logic

Let Φ be a set of proposition letters and *Agt* a set of agents.
Formulas:

$$F ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid [i]\varphi \mid t_i \geq r$$

where $p \in \Phi$, $r \in \mathbb{Q}$, and t_i is a term for agent i
Terms for $i \in Agt$:

$$t_i ::= aP_i(F) \mid t_i + t_i$$

where $a \in \mathbb{Q}$ and F is a propositional formula.
This language is from This example is from

- R. Fagin & J. Halpern (1994) Reasoning about Knowledge and Probability. *Journal of the ACM* 41:2, pp. 340–367.

Probabilistic epistemic models and semantics

Let Φ be a set of proposition letters and Agt a set of agents. $M = (X, R, \|\cdot\|, \mathbf{P})$, where

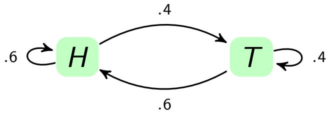
- $(X, R, \|\cdot\|)$ is an epistemic model
- $\mathbf{P}$ is a collection of probability spaces $(S_{i,x}, \mathcal{A}_{i,x}, \mu_{i,x})$ for each $i \in Agt$ and $x \in X$, such that $S_{i,x} \subseteq X$.

The semantics of formulas is defined by a function $[\![\cdot]\!]$ from formulas to subsets of X.

$$
\begin{array}{lcl}
[\![\top]\!] & = & X \\
[\![p]\!] & = & \|p\| \\
[\![\neg\varphi]\!] & = & X - [\![\varphi]\!] \\
[\![\varphi \wedge \psi]\!] & = & [\![\varphi]\!] \cap [\![\psi]\!] \\
[\![[i]\varphi]\!] & = & l_i([\![\varphi]\!]) \\
[\![\sum_{k=1}^{n} a_j P_i(\varphi_k) \geq r]\!] & = & \{x \mid \sum_{j=1}^{n} a_k \mu_{i,x}([\![\varphi_k]\!] \cap S_{i,x}) \geq r\}
\end{array}
$$

Intuition about semantics

When X is finite and all $\mathcal{A}_{i,x} = \mathcal{P}(X)$, then depict a probability function as a directed graph labelled with probabilities:
For example, we represent the uncertainty of an agent about the result of flipping a weighted coin:



Notice that the sum of the numbers on arrows leaving a state is 1.

Fagin, Halpern, and Tuttle example

Suppose there are two agents i and k.

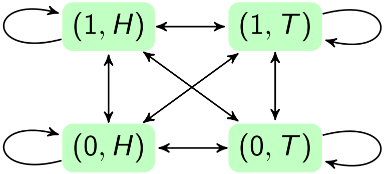
1. k is first given a bit 0 or 1. k learns he has this bit, i is aware that k received a bit, but i does not know what bit k received.
2. k flips a fair coin and looks at the result. i sees k look at the result, but does not what the result is.
3. k performs action s if the coin agrees with the bit (given that heads agrees with 1 and tails agrees with 0), and performs action d otherwise.

This example is from

- R. Fagin & J. Halpern (1994) Reasoning about Knowledge and Probability. *Journal of the ACM* 41:2, pp. 340–367.

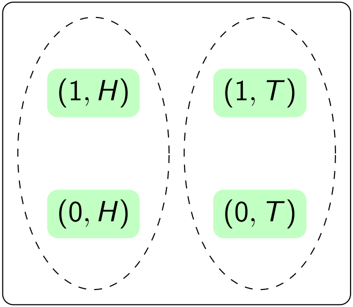
Discussion

There are four possible sequences of events: $(1, H), (1, T), (0, H), (0, T)$ (note that the action s or d is determined from the first two steps). Until k performs the action s or d, agent i considers any of these four states possible.



We indicate i's uncertainty between two states using a bidirectional arrow between the two states. In particular, an arrow from state x to state y indicates that i considers y possible if x is the actual state. (Before the bit is given, k's epistemic relation will be the same).

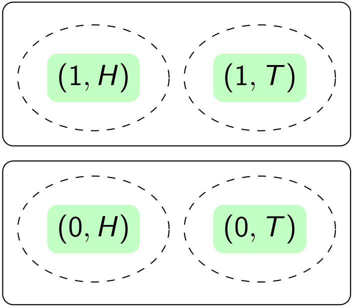
Here is a possibility for i's probability spaces. The sample space enclosed in a box, and the σ-algebra equivalence classes are enclosed in the dotted ovals.



$\mathbf{M}_1$

The sample space is the same as the set of states i considers possible. Individual states cannot be measurable (otherwise 0 or 1 must be assigned a probability).

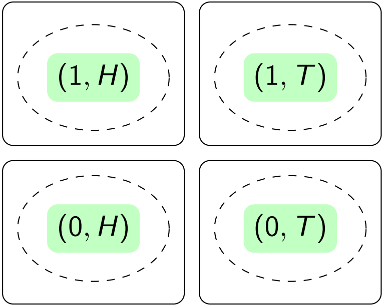
Another possibility has a sample space containing only the states with the correct bit (but recall that i considers all states possible and both sample spaces possible).



$\mathbf{M}_2$

Without assigning probability to the bit, i can now assign a probability to the actions s and d.

Here i is uncertain among 4 probability spaces.



$\mathbf{M}_3$

Mixing qualitative and quantitative

When mixing probability and epistemics, each represents beliefs about different aspects of a situation. In the previous example, there may be

1. quantitative (probability) beliefs about the coin toss
2. qualitative beliefs about the bit or about the probabilities themselves

Representing uncertainty about probabilities

- unmeasurable sets:
  - advantage of allowing us to clearly represent an agent's complete uncertainty about the probability of an situation.
  - disadvantage of excluding potentially reasonable sets from having a probability (such as the probability of $\{(H,1),(T,0)\}$, that is agent k doing action s).
- uncertainty about probabilities
  - advantage of allowing us to divide an unmeasurable set into subsets each in different probability spaces.
  - advantage of allowing us to reflect uncertainty between/among specific probability spaces.
  - disadvantage of requiring all probability measures considered possible be explicit; complete uncertainty requires all infinitely many possible probability measures.

Proof System for PEL

- All propositional tautologies
- $[i](\varphi \to \psi) \to ([i]\varphi \to [i]\psi)$
- $[i]\varphi \to \varphi$
- $[i]\varphi \to [i][i]\varphi$
- $\neg[i]\varphi \to [i]\neg[i]\varphi$
- $P_i(\varphi) \geq 0$
- $P_i(\top) = 1$
- $P_i(\varphi \wedge \psi) \wedge P_i(\varphi \wedge \neg\psi) = P_i(\varphi)$
- Inequality axioms
- If $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$.
- If $\vdash \varphi$, then $\vdash [i]\varphi$.
- If $\vdash \varphi \leftrightarrow \psi$, then $\vdash P_i(\varphi) = P_i(\psi)$.

Inequality axioms

- (permutation)
  $a_1 P_i(\varphi_1) + \cdots + a_n P_i(\varphi_n) \geq r \rightarrow a_{j_1} P_i(\varphi_{j_1}) + \cdots + a_{j_n} P_i(\varphi_{j_n}) \geq r$
- (adding coefficients)
  $(\sum_{k=1}^{n} a_k P_i(\varphi_k) \geq r) \wedge (\sum_{k=1}^{n} b_k P_i(\varphi_k) \geq s) \rightarrow (\sum_{k=1}^{n} (a_k + b_k) P_i(\varphi_k) \geq (r + s))$
- (adding and deleting 0 terms)
  $(t \geq r) \leftrightarrow (t + 0 P_i(\varphi) \geq r)$
- (multiplying by non-zero coefficient)
  $t \geq r \leftrightarrow at \geq ar$ whenever $a > 0$.
- (dichotomy)
  $t \geq r \vee t \leq r$
- (monotonicity)
  $t \geq r \rightarrow t > s$, whenever $r > s$.

Completeness

- Fix a consistent formula θ
- Let Δ be the set of subformulas and negations of subformulas of θ. (Δ is finite.)

$\mathcal{M} = (X, R, \|\cdot\|, \mathbf{P})$, where

- X is the set of maximally consistent subsets of Δ
- xR_iy iff for all $[i]\varphi \in \Delta$, $[i]\varphi \in x$ iff $[i]\varphi \in y$.
- $\|p\| = \{x \in X \mid p \in x\}$
- $\mathbf{P} = \{(S_{i,x}, \mathcal{A}_{i,x}, \mu_{i,x})\}$
  - $S_{i,x} = X$
  - $\mathcal{A}_{i,x} = \mathcal{P}(X)$
  - $\mu_{i,x}$ is any function satisfying conditions of next slides.

Lemma for Completeness

- $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$ be the set of subsets of θ,
- $At(\Sigma) = \{\bigwedge_{i=1}^{n} \delta_i \mid \delta_i \in \{\sigma_i, \neg\sigma_i\}\}$

Lemma

Let $t \geq r$ be a probability formula. Let $At(\Sigma) = \{\alpha_1, \ldots, \alpha_{2^n}\}$. Then there are rationals $a_1, \ldots, a_{2^n}$ such that $t \geq r$ is equivalent to $a_1 P_i(\alpha_1) + \cdots + a_{2^n} P_i(\alpha_{2^n}) \geq r$.

Let $At(\Sigma, \varphi) = \{\alpha \in At(\Sigma) \mid \vdash \alpha \to \varphi\}$. Then

$$P(\varphi) \equiv \sum_{\alpha \in At(\Sigma,\varphi)} P(\varphi \wedge \alpha) \equiv \sum_{\alpha \in At(\Sigma,\varphi)} P(\alpha).$$

The first equivalence comes from multiple applications of additivity for each subformula σ_i.

For each $x \in X$, let $\widehat{x} = \bigwedge_{\{\delta \in x\}} \delta$.
Note: $\{\widehat{x} \mid x \in X\} \subseteq At(\Sigma)$, and

$$\{\widehat{x} \mid \psi \in x\} = At(\Sigma, \psi) := \{\alpha \in At(\Sigma) \mid \vdash \alpha \to \varphi\}.$$

- Fix i and x.
- Let $\{t_1 \geq r_1, \ldots, t_k \geq r_k\}$ be the i inequality formulas in x.
- Let $\{t_{k+1} \geq r_{k+1}, \ldots, t_m \geq r_m\}$ be the i inequality formulas in $\Delta - x$.
- Each formula $t_j \geq r_j$ is equivalent to $a_{j,1} P_i(\alpha_1) + \cdots + a_{j,2^n} P_i(\alpha_{2^n}) \geq r_j$
- Each formula $t_j \geq r_j$ is equivalent to $\sum_{y \in X} a_{j,x} P_i(\widehat{y}) \geq r_j$

System of inequalities

Let $X = \{y_1, \ldots, y_\ell\}$. Let $\mu_{i,x}$ be defined on X as a solution to:

$$
\begin{array}{|rcl|}
\hline
\sum_{y \in X} a_{1,y}\mu_{i,x}(y) & \geq & r_1 \\
 & \vdots & \\
\sum_{y \in X} a_{k,y}\mu_{i,x}(y) & \geq & r_k \\
\hline
\sum_{y \in X} a_{k+1,y}\mu_{i,x}(y) & < & r_{k+1} \\
 & \vdots & \\
\sum_{y \in X} a_{m,y}\mu_{i,x}(y) & < & r_m \\
\hline
\sum_{y \in X} \mu_{i,x}(y) & \geq & 1 \\
-\sum_{y \in X} \mu_{i,x}(y) & \geq & -1 \\
\hline
\mu_{i,x}(y_1) & \geq & 0 \\
 & \vdots & \\
\mu_{i,x}(y_\ell) & \geq & 0 \\
\hline
\end{array}
$$

Completeness follows from a truth lemma:

Lemma

For every formula $\varphi \in \Delta$ *and state* $x \in X$,

$$\varphi \in x \textit{ iff } (M, x) \in [\![\varphi]\!]$$

- This is proved by induction on the structure of the formula, and is similar to the proof of the truth lemma for basic epistemic logic.
- Note that the case for probability formulas $t \geq r$ does not make use of the induction hypothesis, but follows directly from the choice of the probability measure.