Slide 1

COMP3152/9152
Lecture 12
Knowledge and Probability
Ron van der Meyden

Slide 2

The Monty Hall Puzzle

Monty Hall is a television quiz show host. You are the contestant, and have the opportunity to win a prize.

There are three doors. Behind one door is a car. Behind each of the other doors is a goat. You will win whatever is behind the door you pick.

You pick a door, and then Monty opens another door. There is a goat there.

Monty asks: "Would you like to switch doors, or stick with the door that you have picked?"

What should you do?

Slide 3

Probability Spaces

Recall that a *probability space* is a tuple $(W, \mathcal{F}, \mu)$, where

1. W is a nonempty set,
2. $\mathcal{F}$ is an algebra over W, i.e., a set of subsets of $\mathcal{F}$ that contains W and is closed under union and complementation: if $U, V \in \mathcal{F}$ then $U \cup V \in \mathcal{F}$, $U \cap V \in \mathcal{F}$ and $W \setminus U \in \mathcal{F}$.
3. $\mu : \mathcal{F} \to [0, 1]$ satisfies
   (a) $\mu(W) = 1$
   (b) $\mu(U \cup V) = \mu(U) + \mu(V)$ if $U \cap V = \emptyset$.

Slide 4

If W is infinite we also require that

1. $\mathcal{F}$ is a σ-algebra, i.e., if $U_1, U_2, \ldots \in \mathcal{F}$ then $\bigcup_{i \in \mathbf{N}} U_i \in \mathcal{F}$,
2. μ is countably additive, i.e., if $U_i \cap U_j = \emptyset$ for $i \neq j$, then

$$\mu(\bigcup_{i \in \mathbf{N}} U_i) = \sum_{i \in \mathbf{N}} \mu(U_i)$$

Slide 5

Example

Tossing two coins:

$W = \{HH, HT, TH, TT\}$

$\mathcal{F} = \mathcal{P}(W)$

$\mu(\{HH\}) = 1/4, \quad \mu(\{HT\}) = 1/4,$
$\mu(\{TH\}) = 1/4, \quad \mu(\{TT\}) = 1/4$

Note that it follows that, e.g.,

$\mu(\{HH, TT\}) = \mu(\{HH\}) + \mu(\{TT\})) = 1/4/ + 1/4 = 1/2$

In general, if W is finite, μ is defined by its values on singletons, and we write $\mu(w)$ for $\mu(\{w\})$ for $w \in W$.

Slide 6

Conditioning

Conditional Probability captures revision of uncertainty when given new information.

$$\mu(U|V) = \mu(U \cap V)/\mu(V)$$

Example: Two coins are tossed. You are told that the outcome of the two coin tosses was not HH. What is the probabilty

1. the first coin was H?
2. the coins tosses were the same?

Slide 7

Multi-agent Probability

A *probability structure* for n agents is a tuple $(W, \mathcal{PR}_1, \ldots, \mathcal{PR}_n, \pi)$ where

1. W is a set of worlds and
2. each $\mathcal{PR}_i$ is a *probability assignment*, mapping each world $w \in W$ to probability space $\mathcal{PR}_i(w) = (W_{w,i}, \mathcal{F}_{w,i}, \mu_{w,i})$
3. $\pi : W \times \Phi \rightarrow \{0, 1\}$ is an intrerpretation of atomic propositions Φ.

Slide 8

Example - Two agent Coin Tossing

Agent 1 and 2 each toss a coin, and see only their own coin toss.

$W = \{HH, HT, TH, TT\}$

$\mathcal{PR}_1(HT) = (\{HH, HT\}, \mathcal{P}(\{HH, HT\}), \mu_{1,HT})$

$\mathcal{PR}_2(HT) = (\{TT, HT\}, \mathcal{P}(\{TT, HT\}), \mu_{2,HT})$

where

$$\mu_{1,HT}(HH) = 1/2, \quad \mu_{1,HT}(HT) = 1/2$$

$$\mu_{2,HT}(TT) = 1/2, \quad \mu_{2,HT}(HT) = 1/2$$

Slide 9

Conditions on Probability Strctures

Uniformity. For all i, v, w, if $\mathcal{PR}_i(w) = (W_{w,i}, \mathcal{F}_{w,i}, \mu_{w,i})$ and $v \in W_{w,i}$, then $\mathcal{PR}_i(v) = \mathcal{PR}_i(w)$.

Slide 10

Logic of Probability - Syntax

Let Φ be a set of atomic propositions. The following are formulas:

1. p, where $p \in \Phi$
2. $\phi \wedge \psi$, $\neg\phi$, where ϕ, ψ are formulas
3. $a_1 l_{i_1}(\phi_1) + \ldots + a_k l_{i_k}(\phi_k) > a_{k+1}$, where the a_i are real numbers, $i_1, \ldots, i_k$ are agents, and $\phi_1, \ldots, \phi_k$ are formulas.

Slide 11

Other sorts of inequalities can be defined, e.g.,

$l_i(\phi) > l_i(\psi)$ by $l_i(\phi) - l_i(\psi) > 0$

$al_i(\phi) + bl_j(\psi) \geq c$ by $\neg(al_i(\phi) + bl_j(\psi) < c)$

Examples:

$$l_1(p) + l_1(\neg p) = 1$$

$$l_1(p \wedge q) + 2l_2(p \vee q) > 1$$

Slide 12

Logic of Probability, semantics

Let $M = (W, \mathcal{PR}_1, \ldots, \mathcal{PR}_n, \pi)$ be a probability structure

Define the semantics using a function $t \mapsto [t]_{M,w}$ (where $w \in W$) whose value is a number if t is a term:

$[l_i(\phi)]_{M,w} = \mu_{w,i}(\{u \in W |\ M, u \models \phi\} \cap W_{i,w})$

$[a \cdot t]_{M,w} = a \cdot [t]_{M,w}$

$[t_1 + t_2]_{M,w} = [t_1]_{M,w} + [t_2]_{M,w}$

Define satisfaction using:

$M, w \models t > c$ if $[t]_{M,w} > c$

Slide 13

Example

In the two agent coin tossing example, suppose

$p_{i,x}$ means "agent i tossed x"

Then

$M, HH \models l_1(p_{1,H}) = 1$

$M, HH \models l_2(p_{1,H}) = 1/2$

$M, HT \models l_2(p_{1,H}) = 1/2$

$M, HH \models l_1(l_2(p_{1,H}) = 1/2) = 1$

Slide 14

Axiomatizing the Logic of Probability

Prop. All substitution instances of propositional logic

MP. From ϕ and $\phi \Rightarrow \psi$ deduce ψ

QU1. $l_i(\phi) \geq 0$

QU2. $l_i(\texttt{true}) = 1$

QU3. $l_i(\phi \wedge \psi) + l_i(\phi \wedge \neg\psi) = l_i(\phi)$

QUGen. From $\phi \iff \psi$ infer $l_i(\phi) = l_i(\psi)$.

Ineq. All substitution instances of valid linear inequality formulas.

Slide 15

Examples of valid linear inequality formulas:

$$x + 3y \geq 0 \Rightarrow 2y + x + y \geq 0$$

$$x + y \geq 1 \wedge 2x + y \geq 0 \Rightarrow 3x + 2y \geq 1$$

Slide 16

Let $\mathcal{M}_n^{meas}$ be the set of probability structures for n agents in which all sets are *measurable*, i.e., for which for each w, i, there exists a set X such that $\mathcal{F}_{w,i} = \mathcal{P}(X)$.

Theorem: The above axiomatization is sound and complete with respect to $\mathcal{M}_n^{meas}$.

lide 17

Combining Multi-agent Knowledge and Probability

An *epistemic probability structure* for n agents is a tuple $(W, \mathcal{K}_1, \ldots, \mathcal{K}_n, \mathcal{PR}_1, \ldots, \mathcal{PR}_n, \pi)$ where

1. W is a set of worlds and
2. each $\mathcal{K}_i$ is an equivalence relation on W,
3. each $\mathcal{PR}_i$ is a *probability assignment*, mapping each world $w \in W$ to probability space $\mathcal{PR}_i(w) = (W_{w,i}, \mathcal{F}_{w,i}, \mu_{w,i})$
4. $\pi : W \times \Phi \to \{0, 1\}$ is an interpretation of atomic propositions Φ.

lide 18

Conditions on Probability Structures

Uniformity. For all i, v, w, if $\mathcal{PR}_i(w) = (W_{w,i}, \mathcal{F}_{w,i}, \mu_{w,i})$ and $v \in W_{w,i}$, then $\mathcal{PR}_i(v) = \mathcal{PR}_i(w)$.

State Dependent Probability. For all i, v, w, if $v \in \mathcal{K}_i(w)$, then $\mathcal{PR}_i(v) = \mathcal{PR}_i(w)$.

Consistency. For all i and w, if $\mathcal{PR}_i(w) = (W_{w,i}, \mathcal{F}_{w,i}, \mu_{w,i})$ then $W_{w,i} \subseteq \mathcal{K}_i(w)$.

If M satisfies State Dependent Probability and Consistency then it satisfies Uniformity.

Slide 19

A language for knowledge and probability

Extend the language for probability, by adding formulas $K_i\phi$, with the usual semantics...

$M, w \models K_i\phi$ if $M, w' \models \phi$ for all $w'\mathcal{K}_i w$.

Slide 20

Example- Two agent Coin tossing

Extending the previous example...

$\mathcal{K}_1(HH) = \{HT, HH\}$ and

$M, HH \models K_1(l_2(p_{1,H}) = 1/2)$

lide 21

Example - differing intuitions

Alice has a coin, which she tosses. Bob does not see the outcome. *After* the toss, what is Bob's probability of Heads?

Intuition 1: $K_B(l_B(p_{A,H}) = 0 \vee l_B(p_{A,H}) = 1$
$\neg K_B(l_B(p_{A,H}) = 0) \wedge \neg K_B(l_B(p_{A,H}) = 1)$

Intuition 2: $K_B(l_B(p_{A,H}) = 1/2)$

We can model both intuitions.

lide 22

Axiomatizing Knowledge and Probability

Theorem: The axiomatic system that combines all the axioms and inference rules for probability with the axioms and inference rules of $S5_n$ is sound and complete for measurable epistemic probability structures.

Slide 23

Adding probability to interpreted systems

A *probability system* is a tuple $(\mathcal{R}, \mathcal{PR}_1, \ldots \mathcal{PR}_n)$ where

1. $\mathcal{R}$ is a set of runs
2. each $\mathcal{PR}_i$ is a probability assignment that associates each point (r, m) of $\mathcal{R}$ with a probability space $\mathcal{PR}_i(r, m) = (W_{r,m,i}, \mathcal{F}_{r,m,i}, \mu_{r,m,i})$.

Generally, $W_{r,m,i}$ will be a set of points of $\mathcal{R}$.

Slide 24

Assigning probability to runs

Suppose that

1. runs have finite length $k + 1$,
2. there is a probability distribution P_0 on the initial states of the system, and
3. for each run r and time $j < k + 1$, we can assign a probability to each next possible state, i.e., there is a probability distribution $P_{r,j}$ over the set of states

$$\{r'(j+1) \mid r[0 \ldots j] = r'[0 \ldots j]\}$$

at time $j + 1$ in runs that extend $r[0 \ldots j]$.

lide 25

Then we can assign a run r the probability

$$\mu_{\mathcal{R}}(r) = P_0(r(0)) \cdot P_{r,0}(r(1)) \cdot P_{r,1}(r(2)) \cdot \ldots \cdot P_{r,k}(r(k+1))$$

Fact: $\mu_{\mathcal{R}}$ is a probability distribution on the set of runs $\mathcal{R}$ of length $k+1$.

lide 26

Example- Tossing a coin

Suppose that Alice has a fair (1/2 H, 1/2 T) coin F and baised (1/3 H, 2/3 T) coin B in her pocket. She randomly (1/2, 1/2) picks a coin and tosses it. What is the probability of heads?

Then

$$\mu_{\mathcal{R}}(FH) = P_{FH,0}(F) \cdot P_{FH,1}(FH) = 1/2 \cdot 1/2 = 3/12$$

$$\mu_{\mathcal{R}}(BH) = P_{BH,0}(B) \cdot P_{BH,1}(BH) = 1/2 \cdot 1/3 = 2/12$$

So the probability of obtaining heads in a run is 5/12

We can draw this as a tree...

Slide 27

Transition Probabilities from Probabilistic Protocols

A *probabilistic protocol* for agent i (or the environment e) is a mapping P that maps each local state s of agent i (or global state s) to a probability space $P_i(s) = (A_i, \mathcal{P}(A_i), \mu_{i,s})$ over the set of actions of agent i (the environment).

Given a transition function τ, and a joint probabilistic protocol $\mathbf{P}$, we define $P_{r,m}$ as follows:

$$P_{r,m}(t) = \sum_{\mathbf{a}:\tau(\mathbf{a})(r(m))=t} \mu_{e,r(m)}(\mathbf{a}_e) \cdot \mu_{1,r_i(m)}(\mathbf{a}_1) \cdot \ldots \mu_{n,r_n(m)}(\mathbf{a}_n)$$

Slide 28

From probability on runs to probability on points

If U is a set of points and $\mathcal{S}$ a set of runs, both from a set of runs $\mathcal{R}$, define

$$\mathcal{S}(U) = \{r \in \mathcal{S} \mid (r,m) \in U \text{ for some } m\}$$

$$U(\mathcal{S}) = \{(r,m) \in U \mid r \in \mathcal{S}\}$$

lide 29

Suppose that agents are synchronous (this is the easiest case to handle).

We can now assign each agent i a probability space $\mathcal{PR}_i(r,m) = (W_{r,m,i}, \mathcal{F}_{r,m,i}, \mu_{r,m,i})$. at each point (r,m) as follows:

$W_{r,m,i} = \mathcal{K}_i(r,m) = \{(r',m) \mid r_i(m) = r'_i(m)\}$

$\mathcal{F}_{r,m,i} = \{\mathcal{K}_i(r,m)(\mathcal{S}) \mid \mathcal{S} \in \mathcal{F}_\mathcal{R}\}$

$\mu_{r,m,i}(U) = \mu_\mathcal{R}(\mathcal{R}(U) \mid \mathcal{R}(\mathcal{K}_i(r,m)))$ for $U \in \mathcal{F}_{r,m,i}$

Fact: This is a probability space.

lide 30

If agents have perfect recall, we can understand the evolution of their probabilities as obtained by conditioning...

For U a set of points, define $U^- = \{(r,m) \mid (r,m+1) \in U\}$

Proposition: If $\mathcal{R}$ is a system with synchronous perfect recall, then for all points (r,m) and agents i, if $U \in \mathcal{F}_{r,m+1,i}$ then $U^- \in \mathcal{F}_{r,m,i}$ and

$$\mu_{r,m+1,i}(U) = \mu_{r,m,i}(U^- \mid \mathcal{K}_i(r,m+1)^-)$$

Slide 31

Analysis of the Monty Hall Problem

Answer: what you should do depends on a number of extra assumptions, in particular on Monty's protocol.

Assume: the location of the car is uniformly distributed.

Case 1: whatever door you pick, Monty randomly opens another door.
(It doesn't help to switch)

Case 2: If you pick the door with a car, Monty randomly chooses another door. If you pick a door with a goat, Monty chooses the door with the other goat.
(It helps to switch).