Informatics 2D: Reasoning and Agents

Alex Lascarides

Lecture 20c: Probability: Notation and Axioms

Where are we?

- Explained why logic in itself is insufficient to model uncertainty
- Discussed principles of decision making under uncertainty
  - Decision theory, MEU principle
- Probability theory provides useful tools for quantifying degree of belief/uncertainty in propositions
- Now: **basic probablity notation and axioms of probability theory**

Propositions & atomic events

- Degrees of belief concern propositions
- Basic notion: **random variable**, a part of the world whose status is unknown, with a **domain** (e.g. *Cavity* with domain $\langle true, false \rangle$)
- Can be boolean, discrete or continuous
- Can compose complex propositions from statements about random variables (e.g. $Cavity = true \wedge Toothache = false$)
- **Atomic event** = complete specification of the state of the world
  - Atomic events are mutually exclusive
  - Their set is exhaustive
  - Every event entails truth or falsehood of any proposition (like models in logic)
  - Every proposition logically equivalent to the disjunction of all atomic events that entail it

Propositions & atomic events

- **Unconditional/prior probability** = degree of belief in a proposition a in the absence of any other information
- Can be between 0 and 1, write as $P(Cavity = true) = 0.1$ or $P(cavity) = 0.1$
- **Probability distribution** = the probabilities of all values of a random variable
- Write $\mathbf{P}(Weather) = \langle 0.7, 0.2, 0.1 \rangle$ for

$$P(Weather = sunny) = 0.7$$
$$P(Weather = rain) = 0.2$$
$$P(Weather = cloudy) = 0.1$$

Probability distributions/conditional probabilities

- For a mixture of several variables, we obtain a **joint probability distribution** (JPD) – cross-product of individual distributions
- A JPD ("joint") describes one's uncertainty about the world as it specifies the probability of every atomic event
- For continuous variables we use **probability density function** (we cannot enumerate values)
- Will talk about these in detail later
- **Conditional probability** $P(a|b)$ = the probability of a given that all we know is b
- Example: $P(cavity|toothache) = 0.8$ means that if patient is observed to have toothache, then there is an 80% chance that he has a cavity

Conditional probabilities

- Can be defined using unconditional probabilities: $P(a|b) = \frac{P(a \wedge b)}{P(b)}$
- Often written as **product rule** $P(a \wedge b) = P(a|b)P(b)$
- Good for describing JPDs (which then become "CPDs") as $\mathbf{P}(X, Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$
- Set of equations, not matrix multiplication (!):

$$
\begin{gathered}
P(X = x_1 \wedge Y = y_1) = P(X = x_1 | Y = y_1)P(Y = y_1) \\
P(X = x_1 \wedge Y = y_2) = P(X = x_1 | Y = y_2)P(Y = y_2) \\
\vdots \\
P(X = x_n \wedge Y = y_m) = P(X = x_n | Y = y_m)P(Y = y_m)
\end{gathered}
$$

- Conditional probability does **not** mean logical implication!

The axioms of probability

- **Kolmogorov's axioms** define basic semantics for probabilities:

  1. $0 \leq P(a) \leq 1$ for any proposition a
  2. $P(true) = 1$ and $P(false) = 0$
  3. $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

- From this, a number of useful facts can be derived, e.g:
  - $P(\neg a) = 1 - P(a)$
  - For variable D with domain $\langle d_1, \ldots, d_n \rangle$, $\sum_{i=1}^{n} P(D = d_i) = 1$
  - And so any JPD over finite variables sums to 1
  - If $\mathbf{e}(a)$ is the set of atomic events that entail a, then (because they are mutually exclusive) it holds that

$$P(a) = \sum_{e_i \in \mathbf{e}(a)} P(e_i)$$

  - With this, we can calculate the probability of any proposition from a JPD

Example derivation: $P(\neg a) = 1 - P(a)$

$$
\begin{aligned}
P(\neg a \vee a) &= P(true) && \textit{logic} \\
&= 1 && \textit{axiom1} \\
&= P(\neg a) + P(a) - P(\neg a \wedge a) && \textit{axiom3} \\
&= P(\neg a) + P(a) - P(false) && \textit{logic} \\
&= P(\neg a) + P(a) - 0 && \textit{axiom1} \\
&= P(\neg a) + P(a) && \textit{arithmetic} \\
P(\neg a) &= 1 - P(a) && \textit{arithmetic}
\end{aligned}
$$

Summary

- Probability theory provides useful tools for quantifying degree of belief/uncertainty in propositions
- Atomic events, propositions, random variables
- Probability distributions, conditional probabilities
- Axioms of probability
- Next time: **Introduction to Coursework 2**