Informatics 2D – Reasoning and Agents

Semester 2, 2011-12

Alex Lascarides
alex@inf.ed.ac.uk

Lecture 20 – Acting under Uncertainty
8th March 2012
adapted from slides by Michael Rovatsos

Where are we?

Last time ...

- Previous part of course discussed planning as an efficient way of determining actions that will achieve goals
- Used more elaborate representations than in search, but avoided full complexity of logical reasoning
- Allowed uncertainty to some extent (e.g. conditional planning, replanning)
- However the approaches seen so far don't allow for a *quantification* of uncertainty

Today ...

- **Acting under uncertainty**

Handling uncertain knowledge

- So far we have always assumed that propositions are assumed to be true, false, or unknown
- But in reality, we have hunches rather than complete ignorance or absolute knowledge
- Approaches like conditional planning and replanning handle things that might go wrong
- But they don't tell us how likely it is that something might go wrong...
- And **rational decisions** (i.e. 'the right thing to do') depend on the relative importance of various goals and the **likelihood** that (and degree to which) they will be achieved

Handling uncertain knowledge

- To develop theories of uncertain reasoning we must look at the nature of uncertain knowledge
- Example: rules for dental diagnosis
  - A rule like $\forall p\ Symptom(p, Toothache) \Rightarrow Disease(p, Cavity)$ is clearly wrong
  - Disjunctive conclusions require long lists of potential diagnoses:

$$\forall p\ Symptom(p, Toothache) \Rightarrow$$
$$Disease(p, Cavity) \vee Disease(p, GumDisease) \vee Disease(p, Abscess) \ldots$$

  - Causal rules like $\forall p\ Disease(p, Cavity) \Rightarrow Symptom(p, Toothache)$ can also cause problems
  - Even if we know all possible causes, what if the cavity and the toothache are not connected?

Uncertain knowledge, logic, and probabilities

- Clearly, using (classical) logic is not very useful to capture uncertainty, because of . . .
  - complexity (can be impractical to include all antecedents and consequents in rules, and/or too hard to use them)
  - theoretical ignorance (don't know a rule completely)
  - practical ignorance (don't know the current state)
  - How *likely* an unknown factor is influences how we reason and act
- One possible approach: express **degrees of belief** in propositions using **probability theory**

  Probability can **summarise** *the uncertainty that comes from our 'laziness' and ignorance*

- Probabilities between 0 and 1 express the degree to which we believe a proposition to be true

Degrees of belief and probabilities

- In probability theory, propositions themselves are actually true or false!
- **Degrees of truth** are the subject of other methods (like **fuzzy logic**) not dealt with here
- Degrees of belief depend on **evidence** and should change with new evidence
- Don't confuse this with change in the world that might make the proposition itself true or false!
- Before evidence is obtained we speak of **prior/unconditional probability**, after evidence of **posterior probability**

Uncertainty and rational decisions

- Logical agent has a goal and executes any plan guaranteed to achieve it
- Different with degrees of belief: If plan P has a 90% chance of success, how about another P' with a higher probability? Or how about P'' with higher cost but same probability?
- Agent must have **preferences** over **outcomes** of plans
- **Utility theory** can be used to reason about those preferences
- Based on idea that every state has a degree of usefulness and agents prefer states with higher utility
- Utilities vary from one agent to another.

Decision theory

- A general theory of rational decision making
- **Decision theory** = probability theory + utility theory
- Foundation of decision theory:

  An agent is rational if and only if it chooses the action that yields the highest expected utility, averaged over all possible outcomes of the action

- Principle of **Maximum Expected Utility**
- Although we follow it here, some points of criticism:
  - Knowledge of preferences?
  - Consistency of preferences?
  - Risk-taking attitude?

Are We Rational?

A: 80% chance of £4000	C: 20% chance of £4000
B: 100% chance of £3000	D: 25% chance of £3000

- Only 19% of you chose lottery A over lottery B.
- But 100% of you chose lottery C over lottery D.
- So lot's of you chose B and C.
- If $U(3000) > 0.8 * U(4000)$, then $0.25 * U(3000) > 0.2 * U(4000)$!!
- It's not rational to choose B and C!
- Our ability to MEU also higly effectly by emotion, social relationships, relationships among our choices. . .
- In fact, we're predictably irrational.
- If we were always rational, we wouldn't have self-help, life coaches etc.

Design for a decision-theoretic agent

- For the time being, we will focus on probability and not utility.
- But still useful to have an idea of general abstract design for a decision-theoretic (utility-based) agent
- Characterised by basic perception-action loop as follows:
  1. Update belief state based on previous action and percept
  2. Calculate outcome probabilities for actions given action descriptions and belief states
  3. Select action with highest expected utility given probabilities of outcomes and utility information
- Very simple but broadly accepted as a general principle for building agents able to cope with real-world environments

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** $P(a \wedge b) = P(a|b)P(b)$
- Intuitively, for $a \wedge b$ to be true, we need b to be true, and a to be true if b is true
- 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 (!):

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

- 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

Summary

- 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
- Atomic events, propositions, random variables
- Probability distributions, conditional probabilities
- Axioms of probability
- Next time: **Introduction to Coursework 2**