Decision Theory Intro: Preferences and Utility

CPSC 322 Lecture 29

March 22, 2006
Textbook §9.5

Lecture Overview

Markov chain

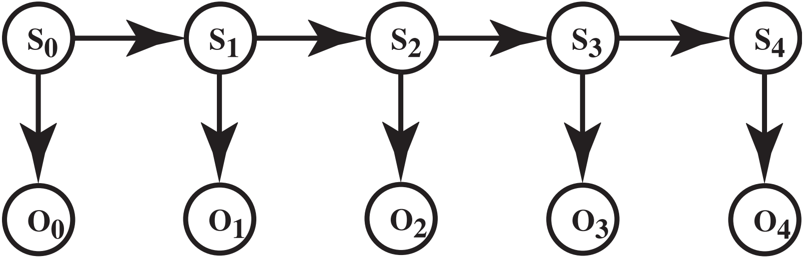
- A Markov chain is a special sort of belief network:

- Thus $P(S_{t+1}|S_0, \ldots, S_t) = P(S_{t+1}|S_t)$.
- "The past is independent of the future given the present."

- A stationary Markov chain is when for all $t > 0$, $t' > 0$, $P(S_{t+1}|S_t) = P(S_{t'+1}|S_{t'})$.
  - We specify $P(S_0)$ and $P(S_{t+1}|S_t)$.

Hidden Markov Model

- A Hidden Markov Model (HMM) starts with a Markov chain, and adds a noisy observation about the state at each time step:



- $P(S_0)$ specifies initial conditions
- $P(S_{t+1}|S_t)$ specifies the dynamics
- $P(O_t|S_t)$ specifies the sensor model

Lecture Overview

Recap

Decision Theory Intro

Preferences

Utility

Decisions Under Uncertainty

- In the first part of the course we focused on decision making in domains where the environment was understood with certainty
  - Search/CSPs: single decisions
  - Planning: sequential decisions
- In uncertain domains, we've so far only considered how to represent and update beliefs
- What if an agent has to make decisions in a domain that involves uncertainty?
  - this is likely: one of the main reasons to represent the world probabilistically is to be able to use these beliefs as the basis for making decisions

Decisions Under Uncertainty

- An agent's decision will depend on:
  1. what actions are available
  2. what beliefs the agent has
     - note: this replaces "state" from the deterministic setting
  3. the agent's goals

- We've spoken quite a lot about (1) and (2).
  - today let's consider (3)
  - we'll move from all-or-nothing goals to a richer notion: rating how happy the agent is in different situations

Lecture Overview

Preferences

- Actions result in outcomes
- Agents have preferences over outcomes
- A rational agent will take the action that leads to the outcome which he most prefers
- Sometimes agents don't know the outcomes of the actions, but they still need to compare actions
- Agents have to act (doing nothing is (often) an action).

Preferences Over Outcomes

If o_1 and o_2 are outcomes

- $o_1 \succeq o_2$ means o_1 is at least as desirable as o_2.
  - read this as "the agent weakly prefers o_1 to o_2"
- $o_1 \sim o_2$ means $o_1 \succeq o_2$ and $o_2 \succeq o_1$.
  - read this as "the agent is indifferent between o_1 and o_2."
- $o_1 \succ o_2$ means $o_1 \succeq o_2$ and $o_2 \not\succeq o_1$
  - read this as "the agent strictly prefers o_1 to o_2"

Lotteries

- An agent may not know the outcomes of his actions, but may instead only have a probability distribution over the outcomes.
- A lottery is a probability distribution over outcomes. It is written

$$[p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]$$

where the o_i are outcomes and $p_i > 0$ such that

$$\sum_i p_i = 1$$

- The lottery specifies that outcome o_i occurs with probability p_i.
- We will consider lotteries to be outcomes.

Our Goal

- We want to reason about preferences mathematically
- To do this, we must give some rules that allow us to allow us to relate and transform expressions involving the symbols $\succeq$, $\succ$ and $\sim$, as well as lotteries.
- Just as we did with probabilities, we will axiomatize preferences.
  - These rules will allow us to derive consequences of preference statements
  - In the end, one has to either accept these consequences or reject one of the axioms

Preference Axioms

- Completeness: A preference relationship must be defined between every pair of outcomes:

$$\forall o_1 \forall o_2 \; o_1 \succeq o_2 \text{ or } o_2 \succeq o_1$$

Preference Axioms

- Transitivity: Preferences must be transitive:

$$\text{if } o_1 \succeq o_2 \text{ and } o_2 \succeq o_3 \text{ then } o_1 \succeq o_3$$

  - This makes good sense: otherwise $o_1 \succeq o_2$ and $o_2 \succeq o_3$ and $o_3 \succ o_1$.
  - An agent should be prepared to pay some amount to swap between an outcome they prefer less and an outcome they prefer more
  - Intransitive preferences mean we can construct a "money pump"!

Preference Axioms

Monotonicity: An agent prefers a larger chance of getting a better outcome than a smaller chance:

- If $o_1 \succ o_2$ and $p > q$ then

$$[p : o_1, 1 - p : o_2] \succ [q : o_1, 1 - q : o_2]$$

Consequence of axioms

- Suppose $o_1 \succ o_2$ and $o_2 \succ o_3$. Consider whether the agent would prefer
  - o_2
  - the lottery $[p : o_1, 1 - p : o_3]$

  for different values of $p \in [0, 1]$.
- You can plot which one is preferred as a function of p:

Preference Axioms

Continuity: Suppose $o_1 \succ o_2$ and $o_2 \succ o_3$, then there exists a $p \in [0, 1]$ such that

$$o_2 \sim [p : o_1, 1 - p : o_3]$$

Preference Axioms

Decomposability: ("no fun in gambling"). An agent is indifferent between lotteries that have same probabilities and outcomes. This includes lotteries over lotteries. For example:

$$\begin{aligned}&[p : o_1, 1 - p : [q : o_2, 1 - q : o_3]]\\ \sim\quad &[p : o_1, (1 - p)q : o_2, (1 - p)(1 - q) : o_3]\end{aligned}$$

Preference Axioms

- Substitutivity: if $o_1 \sim o_2$ then the agent is indifferent between lotteries that only differ by o_1 and o_2:

$$[p : o_1, 1 - p : o_3] \sim [p : o_2, 1 - p : o_3]$$

Lecture Overview

What we would like

- We would like a measure of preference that can be combined with probabilities. So that

$$\begin{aligned} & value([p : o_1, 1-p : o_2]) \\ = \quad & p \times value(o_1) + (1-p) \times value(o_2) \end{aligned}$$

- Can we use money as this measure of preference?
  - Would you you prefer

$$\$1,000,000 \text{ or } [0.5 : \$0, 0.5 : \$2,000,000]?$$

What we would like

- We would like a measure of preference that can be combined with probabilities. So that

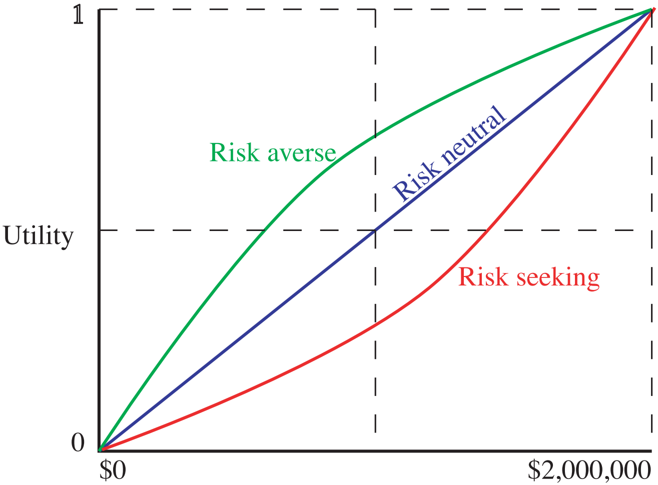
$$\begin{aligned} & value([p : o_1, 1-p : o_2]) \\ = \;& p \times value(o_1) + (1-p) \times value(o_2) \end{aligned}$$

- Can we use money as this measure of preference?
  - Would you you prefer

$$\$1,000,000 \text{ or } [0.5 : \$0, 0.5 : \$2,000,000]?$$

  - Money is not going to work. Let's invent an abstract concept called "utility".

Utility as a function of money



Theorem

Is it possible that preferences are too complex and muti-faceted to be represented by single numbers?
If preferences satisfy the preceding axioms, then preferences can be measured by a function

$$utility : outcomes \rightarrow [0, 1]$$

such that

1. $o_1 \succeq o_2$ if and only if $utility(o_1) \geq utility(o2)$.
2. Utilities are linear with probabilities:

$$utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) = \sum_{i=1}^{k} p_i \times utility(o_i)$$

Proof

Part 1: $o_1 \succeq o_2$ if and only if $utility(o_1) \geq utility(o2)$.

- If all outcomes are equally preferred, set $utility(o_i) = 0$ for all outcomes o_i.
- Otherwise, suppose the best outcome is $best$ and the worst outcome is $worst$.
- For any outcome o_i, define $utility(o_i)$ to be the number u_i such that

$$o_i \sim [u_i : best, 1 - u_i : worst]$$

This exists by the Continuity property.

Proof (cont.)

Part 1: $o_1 \succeq o_2$ if and only if $utility(o_1) \geq utility(o2)$.

- Suppose $o_1 \succeq o_2$ and $utility(o_i) = u_i$, then by Substitutivity,

$$[u_1 : best, 1 - u_1 : worst] \succeq [u_2 : best, 1 - u_2 : worst]$$

Which, by completeness and monotonicity implies $u_1 \geq u_2$.

Proof (cont.)

Part 2: $utility([p_1 : o_1, \ldots, p_k : o_k]) = \sum_{i=1}^{k} p_i \times utility(o_i)$

- Suppose $p = utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k])$.
- Suppose $utility(o_i) = u_i$. We know:

$$o_i \sim [u_i : best, 1 - u_i : worst]$$

- By substitutivity, we can replace each o_i by $[u_i : best, 1 - u_i : worst]$, so

$$\begin{aligned} p = utility(\quad [\quad & p_1 : [u_1 : best, 1 - u_1 : worst] \\ & \ldots \\ & p_k : [u_k : best, 1 - u_k : worst]]) \end{aligned}$$

Proof (cont.)

Part 2: $utility([p_1 : o_1, \ldots, p_k : o_k]) = \sum_{i=1}^{k} p_i \times utility(o_i)$

- By decomposability, this is equivalent to:

$$p = utility(\quad [\quad p_1 u_1 + \cdots + p_k u_k$$
$$: best,$$
$$p_1(1 - u_1) + \cdots + p_k(1 - u_k)$$
$$: worst]])$$

- Thus, by definition of utility,

$$p = p_1 \times u_1 + \cdots + p_k \times u_k$$