

Informatics 2D – Reasoning and Agents

Semester 2, 2011-12

Alex Lascarides
alex@inf.ed.ac.uk

 School of
informatics



Lecture 29 – Decision Making Under Uncertainty
29th March 2012
adapted from slides by Michael Rovatsos

Where are we?

Last time ...

- ▶ Looked at Dynamic Bayesian Networks
- ▶ General, powerful method for describing temporal probabilistic problems
- ▶ Unfortunately exact inference computationally too hard
- ▶ Methods for approximate inference (particle filtering)

Today ...

- ▶ **Decision Making under Uncertainty**

Combining beliefs and desires

- ▶ Rational agents do things that are an optimal tradeoff between:
 - ▶ the likelihood of reaching a particular resultant state (given one's actions) and
 - ▶ The desirability of that state
- ▶ So far we have done the 'likelihood' bit: we know how to evaluate the probability of being in a particular state at a particular time.
- ▶ But we've not looked at an agent's preferences or desires
- ▶ Now we will discuss **utility theory** in more detail to obtain a full picture of decision-theoretic agent design

Utility theory & utility functions

- ▶ Agent's preferences between world states are described using a **utility function**
- ▶ UF assigns some numerical value $U(S)$ to each state S to express its desirability for the agent
- ▶ Nondeterministic action a has results $Result(a)$ and probabilities $P(Result(a) = s' | a, \mathbf{e})$ summarise agent's knowledge about its effects given evidence observations \mathbf{e} .
- ▶ Can be combined with probabilities for outcomes to obtain **expected utility** of action:

$$EU(A|E) = \sum_i P(Result(a) = s' | a, \mathbf{e}) U(s')$$

Utility theory & utility functions

- ▶ Principle of **maximum expected utility** (MEU) says agent should use action that maximises expected utility
- ▶ In a sense, this summarises the whole endeavour of AI:
If agent maximises utility function that correctly reflects the performance measure applied to it, then optimal performance will be achieved by averaging over all environments in which agent could be placed
- ▶ Of course, this doesn't tell us how to define utility function or how to determine probabilities for any sequence of actions in a complex environment
- ▶ For now we will only look at **one-shot decisions**, not **sequential decisions** (next lecture)

Constraints on rational preferences

- ▶ MEU sounds reasonable, but why should this be the best quantity to maximise? Why are numerical utilities sensible? Why single number?
- ▶ Questions can be answered by looking at constraints on preferences
- ▶ Notation:

$A \succ B$ A is preferred to B

$A \sim B$ the agent is indifferent between A and B

$A \succsim B$ the agent prefers A to B or is indifferent between them

- ▶ But what are A and B ? Introduce **lotteries** with outcomes $C_1 \dots C_n$ and accompanying probabilities

$$L = [p_1, C_1; p_2, C_2; \dots; p_n, C_n]$$

Constraints on rational preferences

- ▶ Outcome of a lottery can be state or another lottery
- ▶ Can be used to understand how preferences between complex lotteries are defined in terms of preferences among their (outcome) states
- ▶ The following are considered reasonable **axioms of utility theory**
- ▶ **Orderability:** $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
- ▶ **Transitivity:** If agent prefers A over B and B over C then he must prefer A over C : $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
- ▶ Example: Assume $A \succ B \succ C \succ A$ and A, B, C are goods
 - ▶ Agent might trade C for A and some money if he has A
 - ▶ We then offer B for C and some cash and then trade A for B
 - ▶ Agent would lose all his money over time

Constraints on rational preferences

- ▶ **Continuity:** If B is between A and C in preference, then with some probability agent will be indifferent between getting B for sure and a lottery over A and C

$$A \succ B \succ C \Rightarrow \exists p [p, A; 1 - p, C] \sim B$$

- ▶ **Substitutability:** Indifference between lotteries leads to indifference between complex lotteries built from them

$$A \sim B \Rightarrow [p, A; 1 - p, C] \sim [p, B; 1 - p, C]$$

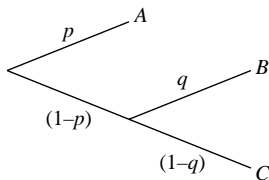
- ▶ **Monotonicity:** Preferring A to B implies preference for any lottery that assigns higher probability to A

$$A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; 1 - p, B] \succeq [q, A; 1 - q, B])$$

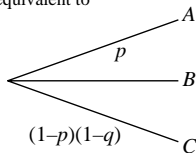
Decomposability example

- **Decomposability:** Compound lotteries can be reduced to simpler one

$$[p, A; 1 - p, [q, B; 1 - q, C]] \sim [p, A; (1 - p)q, B; (1 - p)(1 - q), C]$$



is equivalent to



From preferences to utility

- ▶ The following **axioms of utility** ensure that utility functions follow the above axioms on preference:

- ▶ Utility principle: there exists a function such that

$$U(A) > U(B) \Leftrightarrow A \succ B \quad U(A) = U(B) \Leftrightarrow A \sim B$$

- ▶ MEU principle: utility of lottery is sum of probability of outcomes times their utilities

$$U([p_1, S_1; \dots; p_n, S_n]) = \sum_i p_i U(S_i)$$

- ▶ But an agent might not know even his own utilities!
- ▶ But you can work out his (or even your own!) utilities by observing his (your) behaviour and assuming that he (you) chooses to MEU.

Utility functions

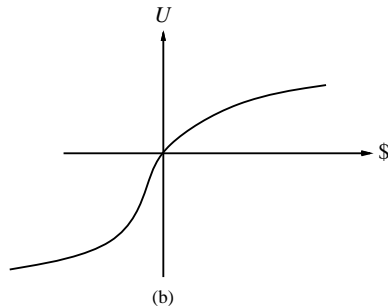
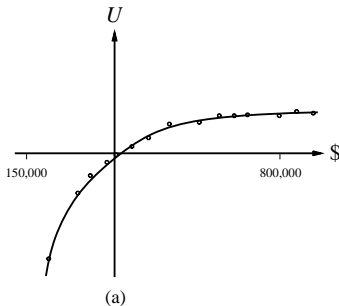
- ▶ According to the above axioms, arbitrary preferences can be expressed by utility functions
 - ▶ I prefer to have a prime number of £ in my bank account; when I have £10 I will give away £3.
- ▶ But usually preferences are more systematic, a typical example being money (roughly, we like to maximise our money)
- ▶ Agents exhibit **monotonic preference** toward money, but how about lotteries involving money?
- ▶ “Who wants to be a millionaire” -type problem, is pocketing a smaller amount irrational?
- ▶ **Expected monetary value** (EMV) is actual expectation of outcome

Utility of money

- ▶ Assume you can keep 1 million or risk it with the prospect of getting three millions at the toss of a (fair) coin
- ▶ EMV of accepting gamble is $0.5 \times 0 + 0.5 \times 3,000,000$ which is greater than 1,000,000
- ▶ Use S_n to denote state of possessing wealth “ n dollars”, current wealth S_k
- ▶ Expected utilities become:
 - ▶ $EU(\text{Accept}) = \frac{1}{2}U(S_k) + \frac{1}{2}U(S_{k+3,000,000})$
 - ▶ $EU(\text{Decline}) = U(S_{k+1,000,000})$
- ▶ But it all depends on utility values you assign to levels of monetary wealth (is first million more valuable than second?)

Utility of money (empirical study)

- ▶ It turns out that for most people this is usually concave (curve (a)), showing that going into debt is considered disastrous relative to small gains in money—**risk averse**.



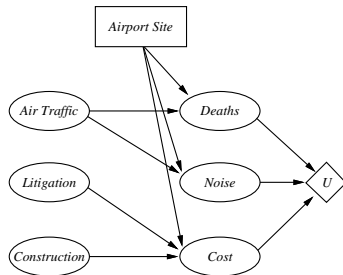
- ▶ But if you're already \$10M in debt, your utility curve is more like (b)—**risk seeking** when desperate!

Utility scales

- ▶ Axioms don't say anything about scales
- ▶ For example transformation of $U(S)$ into $U'(S) = k_1 + k_2 U(S)$ (k_2 positive) doesn't affect behaviour
- ▶ In deterministic contexts behaviour is unchanged by any monotonic transformation (utility function is **value function/ordinal function**)
- ▶ One procedure for assessing utilities is to use **normalised utility** between “best possible prize” ($u^\top = 1$) and “worst possible catastrophe” ($u^\perp = 0$)
- ▶ Ask agent to indicate preference between S and the standard lottery $[p, u^\top : (1 - p), u^\perp]$, adjust p until agent is indifferent between S and standard lottery, set $U(S) = p$

Decision networks

- ▶ What we now need is a way of integrating utilities into our view of probabilistic reasoning
- ▶ **Decision networks (influence diagrams)** combine BNs with additional node types for actions and utilities
- ▶ Illustrate with airport siting problem:

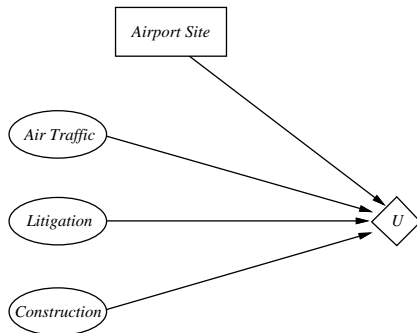


Representing decision problems with DNs

- ▶ **Chance nodes** (ovals) represent random variables with CPTs, parents can be decision nodes
- ▶ **Decision nodes** represent decision-making points at which actions are available
- ▶ **Utility nodes** represent utility function connected to all nodes that affect utility directly
- ▶ Often nodes describing outcome states are omitted and expected utility associated with actions is expressed (rather than states) – **action-utility tables**

Representing decision problems with DNs

- ▶ Simplified version with action-utility tables
- ▶ Less flexible but simpler (like pre-compiled version of general case)



Evaluating decision networks

- ▶ Evaluation of a DN works by setting decision node to every possible value
- ▶ “Algorithm”:
 1. Set evidence variables for current state
 2. For each value of decision node:
 - 2.1 Set decision node to that value
 - 2.2 Calculate posterior probabilities for parents of utility node
 - 2.3 Calculate resulting utility for action
 3. Return action with highest utility
- ▶ Using any algorithm for BN inference, this yields a simple framework for building agents that make single-shot decisions

Summary

- ▶ Foundations for rational decision making under uncertainty
- ▶ Utility theory and its axioms, utility functions
- ▶ Possible points of criticism?
- ▶ Decision networks nicely blend with our BN framework
- ▶ Only looked at one-shot decisions so far
- ▶ Next time: **Markov Decision Processes**