Lecture Notes for AI (CS 382)
Spring 2012, Utility Theory

Instructor: Kostas Bekris

April 17, 2012

1 Utility Theory

In artificial intelligence, we design agents that make intelligent decisions. This section focuses on decision-making under uncertainty. In this context, agents pick states based on their utilities.

A utility function assigns a single number to express the desirability of a state. $U : S \Rightarrow R$ is used to denote the utility of a state, where S is the state space of a problem and R is the set of real numbers. These utilities are used in combination with probabilities of outcomes to get expected utilities for every action. Agents choose the actions that maximize their expected utility.

For example, consider a non-deterministic action A, which has possible outcome states $Result_i(A)$. Index i spans over the different outcomes. Each outcome has a probability assigned to it by the agent before the action is performed:

$$P(Result_i(A)|Do(A), E)$$

E corresponds to the evidence variables. Now we want to maximize the expected utility. If $U(Result_i(A))$ is the utility of state $(Result_i(A))$, then the expected utility is:

$$EU(A|E) = \sum_i P(Result_i(A)|Do(A), E)U(Result_i(A))$$

Maximum Expected Utility (MEU) principle: A rational agent will choose the action that maximizes the expected utility. To find this maximum, we would have to enumerate all actions, which is not practical for long sequences. Instead we must find a different way to handle these problems, starting with simple decisions.

1.1 Constraints on Rational Preferences

First we will define some notation for the constraints on preferences that an agent should have.

$A \succ B$ (A is preferred to B)
$A \sim B$ (A is indifferent to B)
$A \succeq B$ (The agent prefers A to B or is indifferent to them)

A and B here are lotteries. A Lottery is a probability distribution over a set of outcomes. These outcomes can be called prizes of the lottery. A lottery can be defined according to the equation below, where $C_1, ..., C_n$ are possible outcomes and $p_1, ..., p_n$ are the probabilities of these outcomes.

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

The outcome of a lottery is either an atomic state or another lottery.

There are also constraints on lotteries, using the notation introduced above:

- **Orderability:**

  $$\forall A, B : (A \succ B) \vee (B \succ A) \vee (A \sim B)$$

  A rational agent must prefer one lottery or the other, or else the agent considers the two lotteries to be equivalent.

- **Transitivity:**

  $$(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$$

  If an agent prefers A over B, and prefers B over C, then the agent prefers A over C.

- **Continuity:**

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

  If B is between A and C as a preference, then there is a probability that makes the selection of B equivalent to the selection of the lottery that yields A with a probability p and C with probability 1-p.

  For example, consider a case where there are three lotteries A, B, and C. Winning the A lottery will yield $1,000,000; winning the B lottery will yield $10,000; and winning the C lottery will yield $0. The agent has the following choice: Either select B or gamble between A and C with probability p. For a certain probability p of winning A, these two choices will be the same for the agent.

  $$p * 1,000,000 = 10,000 \quad \Rightarrow \quad p = 0.01$$

  When the probability is 1% that the agent can win lottery A, the two choices are considered equivalent.

- **Substitutability:**

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

  If an agent is indifferent between two lotteries, A and B, then the agent is indifferent between two more complex lotteries where the only difference is that B is substituted for A. This is true no matter what the probability p is and no matter what the other outcomes are in the lotteries.

- **Monotonicity:**

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

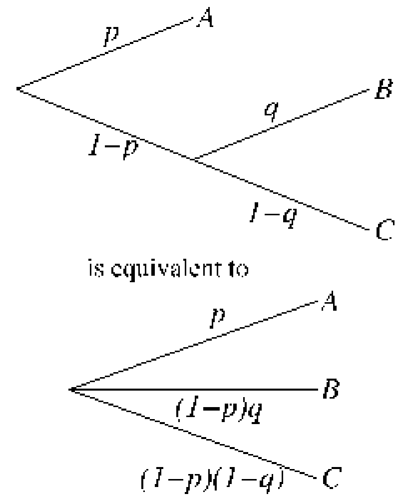
  If A is a preferred outcome over B, then a lottery that selects with higher probability A over B is preferred over one with a lower probability for A.

- **Decomposability:**

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

  Compound lotteries can be simplified using the laws of probability. This also means that two consecutive lotteries can be combined into one equivalent lottery. This is also known as the “no fun in gambling” rule. Figure 1.1 provides an illustration of this constraint.

If the above axioms hold, then $\exists$ a real-valued function U that operates on states so that:
    $U(A) > U(B)$ iff $A \succ B$ and
    $U(A) = U(B)$ iff $A \sim B$
then the utility of a lottery is $U([p_1, S_1; ...; p_n, S_n]) = \sum_i p_i * U(S_i)$ **[Maximum Expected Utility Principle]**



2 Utility Functions

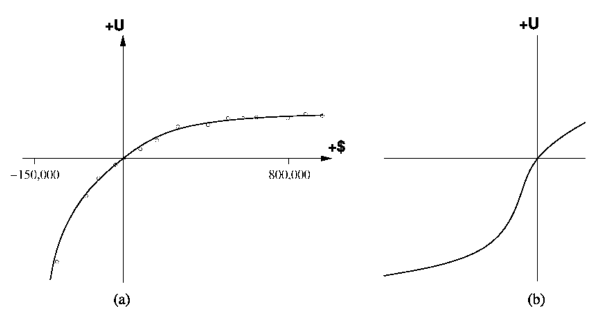
Defining utilities is useful in that we can design utility functions that can change the behavior of the agent to a desired behavior.

2.1 Utility of Money

One example that everyone can relate to in some form is the utility of money. An agent would exhibit monotonic preference which would mean that the agent would always prefer more money to less. But this is not always exaclty the case. Consider the case where you can either take \$1,000,000 or gamble between winning \$0 and \$3,000,000 by flipping a coin. So our expected monetary value of the gamble looks like this:

$$0.5 \cdot (\$0) + 0.5 \cdot (\$3,000,000) = \$1,500,000$$

And the expected monetary value for the first case is \$1,000,000. This does not necessarily mean that gambling is the better deal. Some people would rather take \$1,000,000 if it is worth a lot to them, while people with billions of dollars may gamble because the \$1,000,000 will probably not make a huge difference to them. Studies have shown that the utility of money is proportional to the logarithm of the amount, shown in the first graph below. After a certain amount of money, risk-averse agents prefer a sure payoff that is less than the expected monetary value of the gamble. However, if someone is \$10,000,000 in debt, they may consider gambling with a 50/50 chance of winning \$10,000,000 or losing \$20,000,000 because of desperation. This type of agent is risk seeking and the curve of this behavior can be seen in the second graph below.

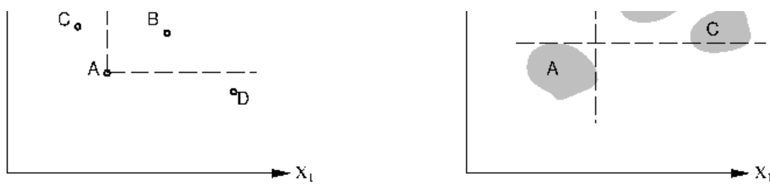


3 Multi-attribute Utility Functions

Multi-attribute problems are problems in which there is more than one attribute. If an airport needs to be built, then cost of land, distance from centers of population, noise levels, and safety issues are all attributes associated with this problem. There are cases, however, where a multi-attribute problem can be equivalent to a single-attribute problem.

3.1 Dominance

Suppose that an airport site S_1 costs less, generates less noise, and is safer than an airport site S_2. Since S_2 is worse on all attributes, it does not even need to be considered. S_1 has a strict dominance over S_2 in this case. Strict dominance can be useful in narrowing choices.



(a) (b)

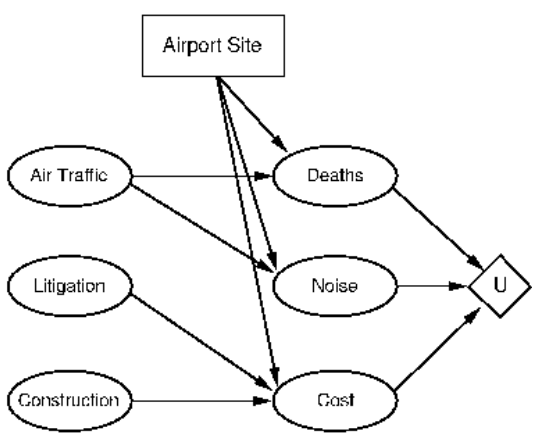
In the figure above, section (a) shows a deterministic case and section (b) shows an uncertain case. In section (a), Option A is strictly dominated by B, but not by C or D. In section (b) of the figure, A is strictly dominated by B, but not by C. Here Option B would be S_2 in our airport example.

4 Decision Networks

Decision networks combine Bayesian networks with additional nodes for types of actions and utilities. In particular, there are:

- **Chance Nodes**(ovals): as in a typical Bayesian network they represent the possible attributes that effect the problem's state.
- **Decision Nodes**(rectangles): They correspond to the different actions that the agent can take to solve a problem.
- **Utility nodes**(diamonds): They are used to show which attributes affect the utility.

Consider the airport figure below. The airport site is a decision node and U is a utility node.



4.1 Evaluating decision networks

The airport site action can take different values (actions of the agent) , which influences cost, noise, and death (safety). The actions should be selected by evaluating the network for each setting of the airport site. The algorithm for evaluating decisions networks is the following:

- Set evidence variables for the current state
- For each possible value of the decision node
  - Set decision node to that value
  - Calculate posterior probabilities for the state nodes that are parents to the utility node
  - Compute expected utility for this action
- Return action that maximizes the expected utility

5 Value of Information

A problem in real life is that often not all information is available and we must collect data to build a proper decision network for our problem. But then the following question arises: what kind of data should we collect? And most importantly: how much should we invest in acquiring data? The example below studies the value of information provided to assist an agent in taking a decision.
Consider the following example: An oil company has n indistinguishable blocks of ocean drilling rights.

- Only one of them will contain oil worth \$C
- The price of each block is $\$\frac{C}{n}$

A seismologist offers to survey block 3 and will definitely indicate whether the block has oil or not. How much should the company pay the seismologist? In other words, what is the "value of information" that the seismologist offers?

There are two outcomes when using the seismologist:

1. $\exists$ oil in block 3, which has a probability of $\frac{1}{n}$
   Then the best action is to buy block 3
   Because we want profit (C) - cost ($\frac{C}{n}$), the total profit is $C - \frac{C}{n} = (n-1)\frac{C}{n}$
2. Does not $\exists$ oil in block 3, which has a probability of $\frac{n-1}{n}$
   Then the best action is to buy a block other than 3
   Because expected profit is $\frac{C}{n-1}$ and since we choose among n-1 blocks, cost is $\frac{C}{n}$ of buying a block.
   The total profit is therefore $\frac{C}{n-1} - \frac{C}{n} = \frac{C}{n(n-1)}$

The expected utility in the case that we use the seismologist is:

$$EU = \frac{1}{n} * \frac{(n-1)C}{n} + \frac{n-1}{n} * \frac{C}{n(n-1)} = \frac{C}{n}$$

If we do not use the seismologist:

$$EU = \text{ expected profit - cost } = \frac{C}{n} - \frac{C}{n} = 0$$

Consequently, if the company can pay less than $\frac{C}{n}$ to acquire the seismologist's information, then it pays off to hire him.

5.1 Generalization

The above problem, however, generalizes to many problems. Lets say that:
E: the current evidence variables
E_j: potentially new evidence (new information) we can acquire
In order to compute the value of acquiring the new information, we must compute the expected utility in the two cases:

1. If we do not acquire the evidence E_j

$$E(U)(a|E) = \max_A \sum_i U(Result_i(A)) * P(Result_i(A)|D_0(A), E)$$

   where A is the action space for the agent.

2. If we do acquire the evidence E_j

$$E(U)(a_{E_j}|E, E_j) = \max_A \sum_i U(Result_i(A)) * P(Result_i(A)|(A), E, E_j)$$

   In this case we have to consider the fact that E_j can take many values. So the expected utility of acquiring E_j is the weighted sum of the expected utilities we get for each possible value of E_j (e.g. whether the seismologist finds or does not find oil in block 3). The weights in the sum are the probabilities of E_j acquiring these values. Then we have:

$$E(U)_{E_j} = \sum_k P(E_j = e_{jk}|E) * E(U)(a_{E_{jk}}|E, E_j = e_{jk})$$

Overall, the value of perfect information is given by:

$$VPI_E(E_j) = (\sum_k P(E_j = e_{jk}|E) * E(U)(a_{E_{jk}}|E, E_j = e_{jk})) - EU(a|E)$$

The above equation shows that information has value:

- if it is likely to cause change of plan
- and the new plan is significantly better than the old