Uncertainty IV: Simple Decision-Making

We now examine:

- the concept of a utility function;
- the way in which such functions can be related to reasonable axioms about preferences;
- a generalization of the Bayesian network, known as a decision network;
- how to measure the value of information, and how to use such measurements to design agents that can ask questions.

Reading: Russell and Norvig, chapter 16.

Simple decision-making

We now look at choosing an action by maximising expected utility.

A utility function $U(s)$ measures the desirability of a state.

If we can express a probability distribution for the states resulting from alternative actions, then we can act in order to maximise expected utility.

For an action a, let $\texttt{Result}(a) = \{s_1, \ldots, s_n\}$ be a set of states that might be the result of performing action a. Then the expected utility of a is

$$\text{EU}(a|E) = \sum_{s \in \texttt{Result}(a)} \text{Pr}(s|a, E)U(s)$$

Note that this applies to individual actions. Sequences of actions will not be covered in this course.

Simple decision-making: all of AI?

Much as this looks like a complete and highly attractive method for an agent to decide how to act, it hides a great deal of complexity:

1. it may be hard to compute $U(s)$. You generally don't know how good a state is until you know where it might lead on to: planning *etc*...
2. knowing what state you're currently in involves most of AI!
3. dealing with $\text{Pr}(s|a, E)$ involves Bayesian networks.

Utility in more detail

Overall, we now want to express preferences between different things.

Let's use the following notation:

$X > Y$: X is preferred to Y
$X = Y$: we are indifferent regarding X and Y
$X \geq Y$: X is preferred, or we're indifferent

X, Y and so on are lotteries. A lottery has the form

$$X = [p_1, O_1|p_2, O_2|\cdots|p_n, O_n]$$

where O_i are the outcomes of the lottery and p_i their respective probabilities. Outcomes can be other lotteries or actual states.

Axioms for utility theory

Given we are dealing with preferences it seems that there are some clear properties that such things should exhibit:

Transitivity: if $X > Y$ and $Y > Z$ then $X > Z$.

Orderability: either $X > Y$ or $Y > X$ or $X = Y$.

Continuity: if $X > Y > Z$ then there is a probability p such that

$$[p, X|(1-p), Z] = Y$$

Substitutability: if $X = Y$ then

$$[p, X|(1-p), L] = [p, Y|(1-p), L]$$

Axioms for utility theory

Monotonicity: if $X > Y$ then for probabilities p_1 and p_2, $p_1 \geq p_2$ if and only if

$$[p_1, X|(1-p_1), Y] \geq [p_2, X|(1-p_2), Y]$$

Decomposability:

$$[p_1, X|(1-p_1), [p_2, Y|(1-p_2), Z]] = [p_1, X|(1-p_1)p_2, Y|(1-p_1)(1-p_2), Z]$$

Axioms for utility theory

If an agent's preferences conform to the utility theory axioms—and note that we are only considering preferences, not numbers—then it is possible to define a utility function $U(s)$ for states such that:

1. $U(s_1) > U(s_2) \longleftrightarrow s_1 > s_2$
2. $U(s_1) = U(s_2) \longleftrightarrow s_1 = s_2$
3. $U([p_1, s_1|p_2, s_2|\cdots|p_n, s_n]) = \sum_{i=1}^{n} p_i U(s_i)$.

We therefore have a justification for the suggested approach.

Designing utility functions

There is complete freedom in how a utility function is defined, but clearly it will pay to define them carefully.

Example: the utility of money (for most people) exhibits a monotonic preference. That is, we prefer to have more of it.

But we need to talk about preferences between lotteries.

Say you've won $100,000$ pounds in a quiz and you're offered a coin flip:

- for heads: you win a total of $1,000,000$ pounds;
- for tails: you walk away with nothing!

Designing utility functions

The expected monetary value (EMV) of this lottery is

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

whereas the EMV of the initial amount is $100,000$.

BUT: most of us would probably refuse to take the coin flip.

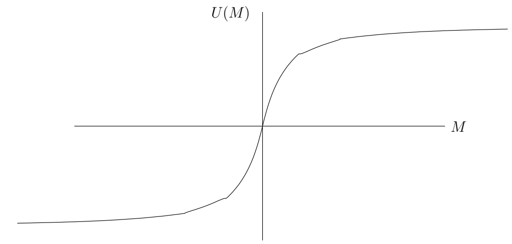
The story is not quite as simple as this though: our attitude probably depends on how much money we have to start with. If I have M pounds to start with then I am in fact choosing between expected utility of $U(M + 100,000)$ and expected utility of

$$(0.5 \times U(M)) + (0.5 \times U(M + 1,000,000))$$

If M is $50,000,000$ my attitude is much different to if it is $10,000$.

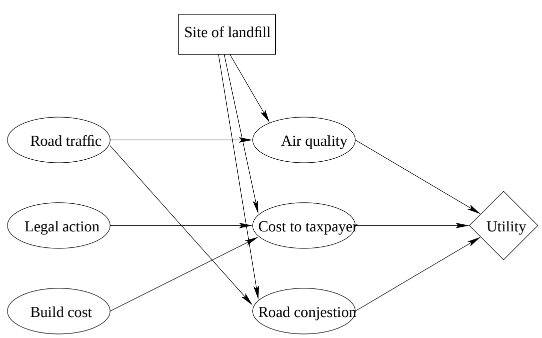
Designing utility functions

In fact, research shows that the utility of M pounds is for most people almost exactly proportional to $\log M$, for $M > 0$.



Decision networks

Decision networks, also known as influence diagrams, allow us to work actions and utilities into the formalism of Bayesian networks.



Decision networks

A decision network has three types of node:

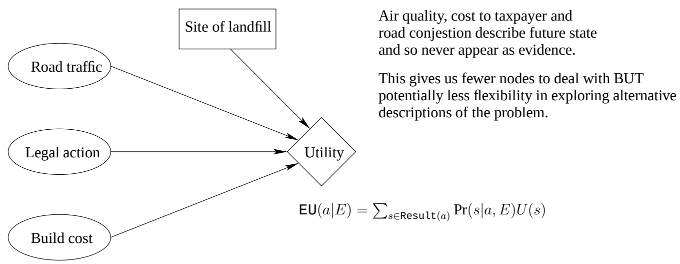
Chance nodes: are denoted by ovals. These are random variables (RVs) represented by a distribution conditional on their parents, as in Bayesian networks. Parents can be other chance nodes or a decision node.

Decision nodes: are denoted by squares. They describe possible outcomes of the decision of interest. Here we deal only with single decisions: multiple decisions require alternative techniques.

Utility nodes: are denoted by diamonds. They describe the utility function relevant to the problem, as a function of the values of the node's parents.

Decision networks

Sometimes such diagrams are simplified by leaving out the RVs describing the new state and converting current state and decision directly to utility:



This is an action-utility table. The utility no longer depends on a state but is the expected utility for a given action.

Evaluation of decision networks

Once a specific action is selected for a decision node it acts like a chance node for which a specific value is being used as evidence.

1. set the current state chance nodes to their evidence values
2. for each potential action
   - fix the decision node
   - compute the probabilities for the utility node's parents
   - compute the expected utility
3. return the action that maximised $\mathtt{EU}(a|E)$

The value of information

We have been assuming that a decision is to be made with all evidence available beforehand. This is unlikely to be the case.

Knowing what questions one should ask is a central, and important part of making decisions. Example:

- doctors do not diagnose by first obtaining results for all possible tests on their patients;
- they ask questions to decide what tests to do;
- they are informed in formulating which tests to perform by probabilities of test outcomes, and by the manner in which knowing an outcome might improve treatment;
- tests can have associated costs.

The value of perfect information

Information value theory provides a formal way in which we can reason about what further information to gather using sensing actions.

Say we have evidence E, so

$$\mathtt{EU}(\mathtt{action}|E) = \max_a \sum_{s \in \mathtt{Result}(a)} \Pr(s|a, E)U(s)$$

denotes how valuable the best action, that is, the best action based on E, must be.

How valuable would it be to learn about a further piece of evidence?

If we examined another RV E' and found that $E' = e'$ then the best action might be altered as we'd be computing

$$\mathtt{EU}(\mathtt{action}'|E, E') = \max_a \sum_{s \in \mathtt{Result}(a)} \Pr(s|a, E, E')U(s)$$

The value of perfect information

BUT: because E' is a RV, and in advance of testing we don't know its value, we need to average over its possible values using our current knowledge.

This leads to the definition of the value of perfect information (VPI)

$$\texttt{VPI}_E(E') = \left\{ \sum_{e'} \text{Pr}(E' = e'|E)\text{EU}(\texttt{action}'|E, E' = e') \right\} - \text{EU}(\texttt{action}|E)$$

The value of perfect information

VPI has the following properties:

- $\texttt{VPI}_E(E') \geq 0$
- It is not necessarily additive, that is, it is possible that
$$\texttt{VPI}_E(E', E'') \neq \texttt{VPI}_E(E') + \texttt{VPI}_E(E'')$$
- It is independent of ordering
$$\begin{aligned}\texttt{VPI}_E(E', E'') &= \texttt{VPI}_E(E') + \texttt{VPI}_{E,E'}(E'') \\ &= \texttt{VPI}_E(E'') + \texttt{VPI}_{E,E''}(E')\end{aligned}$$

Agents that can gather information

In constructing an agent with the ability to ask questions, we would hope that it would:

- use a good order in which to ask the questions;
- avoid asking irrelevant questions;
- trade off the cost of obtaining information against the value of that information;
- choose a good time to stop asking questions.

We now have the means with which to approach such a design.

Agents that can gather information

Assuming we can associate a cost $C(E')$ with obtaining the knowledge that $E' = e'$, an agent can act as follows:

- given a decision network and current percept
- find the piece of evidence E' maximising $\texttt{VPI}_E(E') - C(E')$
- if $\texttt{VPI}_E(E') - C(E')$ is positive then find the value of E', else take the action indicated by the decision network.

This is known as a myopic agent as it requests a single piece of evidence at once.