MAKING SIMPLE DECISIONS

Outline

- Combining Beliefs and Desires Under Uncertainty
- The Basis of Utility Theory
- (Multiattribute) Utility Functions
- Decision Networks
- The Value of Information
- Decision-Theoretic Expert Systems

Based on the textbook by Stuart Russell & Peter Norvig:

Artificial Intelligence, A Modern Approach (2nd Edition)

Chapter 16

1. COMBINING BELIEFS AND DESIRES

- A **state** S is a complete snapshot of the world.
- An agent's preferences are captured by a **utility function** U which maps a state S to a number $U(S)$ describing the desirability of S.
- Specifying a utility function U for each state S may be tedious.
- The problem can be relieved under some circumstances by decomposing states for the purpose of utility assignment.
- A *nondeterministic action* A may have several outcome states $Result_i(A)$ indexed by the different outcomes of A.
- Prior to executing an action A, the agent assigns a probability $P(Result_i(A) \mid Do(A), E)$ to each outcome (here E summarizes the agent's evidence about the world).

Maximum Expected Utility (MEU)

- The **expected utility** of an action A is $\mathrm{EU}(A \mid E) =$
$$\textstyle\sum_i P(Result_i(A) \mid Do(A), E) \times U(Result_i(A)).$$
- The principle of **maximum expected utility**: a rational agent should choose an action that maximizes its expected utility.
- The MEU principle is closely related to performance measures:
"If an agent maximizes a utility function that correctly reflects the performance measure by which its behavior is being judged, then it will achieve the highest possible performance score if averaged over the environments in which the agent could be placed."
- In this lecture, we concentrate on **one-shot decisions**. The case of making **sequential decisions** will be considered later.

2. THE BASIS OF UTILITY THEORY

- As a justification for the MEU principle, some constraints are imposed on the preferences that a rational agent may possess.
- In utility theory, different attainable outcomes (*prizes*) and the respective probabilities (*chances*) are formalized as **lotteries**:
  - A lottery L having outcomes $A_1, \ldots, A_n$ with probabilities $p_1 + \ldots + p_n = 1$ is written as $[p_1, A_1; \ldots; p_n, A_n]$.
  - A lottery $[1, A]$ with a single outcome is abbreviated as A.
- Preference relations for lotteries (or states) A and B:

$A \succ B \iff A$ is preferred to B,

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

$A \succsim B \iff A \succ B$ or $A \sim B$.

Axioms of Utility Theory

For any lotteries A, B, and C:

1. Orderability: $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
2. Transitivity: $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
3. Continuity: $A \succ B \succ C \Rightarrow \exists p[p,A;1-p,C] \sim B$
4. Substitutability: $A \sim B \Rightarrow [p,A;1-p,C] \sim [p,B;1-p,C]$
5. Monotonicity:
$$A \succ B \Rightarrow (p \geq q \Leftrightarrow [p,A;1-p,B] \succsim [q,A;1-q,B])$$
6. Decomposability (the "no fun in gambling" rule):
$$[p,A;1-p,[q,B;1-q,C]] \sim [p,A;(1-p)q,B;(1-p)(1-q),C]$$

- The existence of a utility function *is guaranteed* by the axioms:
  1. **Utility principle:** if the axioms of utility theory are obeyed, then there is a real-valued function U such that
  $$U(A) > U(B) \iff A \succ B \text{ and}$$
  $$U(A) = U(B) \iff A \sim B.$$
  2. **Maximum Expected Utility principle:** the utility of a lottery
  $$U([p_1,A_1;\ldots;p_n,A_n]) = \sum_i p_i U(A_i).$$
- However, the existence of a utility function U need not imply the agent is *explicitly* maximizing U in its own deliberations.
- By observing an agent's preferences, it is possible to construct a utility function representing what the agent is trying to achieve.

3. UTILITY FUNCTIONS

- Beyond the axioms, an agent can have any preferences it likes.

  Example. An agent prefers to have a prime number of euros in its bank account (having 16€ it would give away 3€).
- Preferences can also interact in complex ways.

  Example. Having a digital TV (in contrast to a conventional one) affects the preferences on soap operas one wishes to watch.
- We are interested in systematic ways of designing utility functions that generate the kinds of behavior we want.

The Utility of Money

- Utility theory has its roots in economy where the utility measure is **money** (an agent's total net assets).
- Money plays a central role in human utility functions because of its almost universal exchangeability for all kinds of goods and services.
- Typically, there is a **monotonic preference** for money.
- Money behaves as a **value function** or **ordinal utility function**: more money is preferred to less when considering *definite amounts*.
- To understand monetary decision making under uncertainty we need to analyze the agent's preferences between lotteries involving money.

Example. A competitor in a TV game show is offered two prizes: either

A: 1000000€ for sure, or

B: after flipping a fair coin, either 3000000€ (heads) or 0€ (tails).

Is it irrational to choose the prize A?

1. The **expected monetary values** (EMV) of the choices are:

   $\mathrm{EMV}(A) = 1 \times 1000000€ = 1000000€$ and

   $\mathrm{EMV}(B) = 0.5 \times 3000000€ + 0.5 \times 0€ = 1500000€$.

2. If S_k denotes the current wealth of k€, **expected utilities** are:

   $\mathrm{EU}(A) = U(S_{k+1000000})$ and

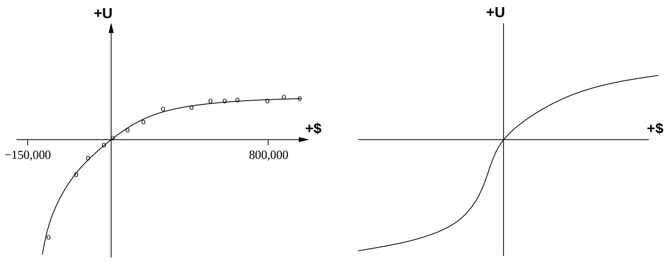
   $\mathrm{EU}(B) = 0.5U(S_k) + 0.5U(S_{k+3000000})$.

☞ The choice depends on the respective utilities and k especially!

➤ Grayson [1960] found an almost perfect fit to the logarithmic form.

➤ Mr. Beard's preferences (a) turned out to be consistent with

$$U(S_{k+n}) = (22.09 \times \mathit{log}(n + 150000) - 263.91)\ \$.$$



(a) (b)

➤ Going into debt is usually considered disastrous.

➤ Preferences between different levels of debt (b) may be analogous (but reverse) to those of positive wealth.

Example. St. Petersburg paradox [Bernoulli, 1738]: a fair coin is tossed n times until it comes up heads and the prize is 2^n€.

How much would you pay for a chance to play this game?

➤ The expected monetary value for this game is

$$\mathrm{EMV} = \sum_i P(Heads_i) \times 2^i = \sum_{i=1}^{\infty} \frac{1}{2^i} 2^i = \infty.$$

☞ A player should be willing to pay any finite sum!

➤ Bernoulli solved the paradox by setting $U(S_{k+n}) = \log_2 n$:

$$\mathrm{EU} = \sum_i P(Heads_i) \times U(Heads_i) = \sum_{i=1}^{\infty} \frac{i}{2^i} = 2.$$

➤ A rational agent (with the given utility scale) should be willing to pay 4€ for playing the game, because $U(S_{k+4}) = \log_2 4 = 2$.

☞ The utility of money is measured on a logarithmic scale (at least for positive amounts).

Insurance Premium

➤ Typically, for any lottery L, the utility of being faced with L is less than the utility of being handed $\mathrm{EMV}(L)$ for sure.

➤ A **risk-averse** agent prefers a sure thing with a payoff that is less than the expected monetary value of a gamble.

➤ A desperately debted agent may behave in a **risk-seeking** way.

➤ A **certainty equivalent** of a lottery L is the sum that an agent is ready to accept as a substitute for participating L.

Example. The certainty equivalent is 400€ for a lottery L that gives 1000€ half the time and 0€ otherwise ($\mathrm{EMV}(L) = 500€$).

➤ An insurance is based on a positive **insurance premium**, i.e., the difference between $\mathrm{EMV}(L)$ and the certainty equivalent for L.

Utility Scales and Assessment

- The axioms of utility do not specify a unique utility function.

  Example. For instance, two agents based on $U(S)$ and $U'(S) = k_1 + k_2 \times U(S)$ with $k_2 > 0$ behave identically.

- A way to assess utilities is to establish a scale with a "*best possible prize*" $u_{\max}$ and a "*worst possible catastrophe*" $u_{\min}$.

- **Normalized utilities** use a scale with $u_{\min} = 0$ and $u_{\max} = 1$.

- An intermediate utility $U(S) = p$ is determined by *indifference* between S and a **standard lottery** $L = [p, u_{\max}; (1-p), u_{\min}]$.

- Trade-offs in decision making let us assess the value of human life.

Examples. **Micromort** (1/1000000 chance of death) and **QALY** (quality-adjusted life year) are measures for the value of human life.

4. MULTIATTRIBUTE UTILITY FUNCTIONS

- **Multiattribute utility theory** deals with utility functions $U(X_1, \ldots, X_n)$ that depend on several attributes $X_1, \ldots, X_n$.

- Each attribute X_i ranges over discrete/continuous scalar values.

- For simplicity, it is assumed that (*all other things being equal*) greater values of an attribute X_i correspond to higher utilities.

- We would like to identify regularities in the preference behavior as **representation theorems** for the corresponding utility functions:

$$U(x_1, \ldots, x_n) = f[f_1(x_1), \ldots, f_n(x_n)]$$

  where f is a simple function such as addition.

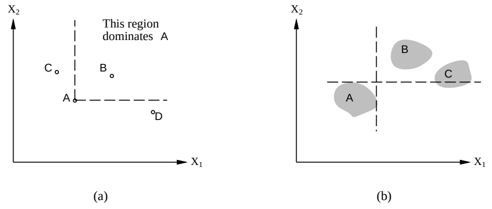
Dominance

- There is **strict dominance** of an option S_1 over other option S_2 if S_1 is better than S_2 with respect to all attributes.

  Example. An airport site S_1 costs less, generates less noise pollution, and is safer than another site S_2.

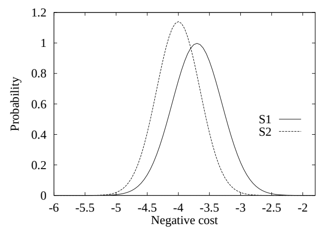
- Uncertain attribute values can be handled analogously.

- Strict dominance is useful in narrowing down the choices.



Stochastic Dominance

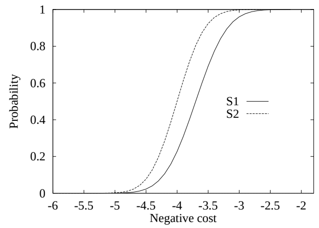
Example. The costs of siting the airport at S_1 and S_2 are 3.7×10^9€ and 4.0×10^9€ with standard deviations 0.4×10^9€ and 0.35×10^9€.



- Knowing that the cost of S_1 is *exactly* 3.7×10^9€ does not enable decision making, because S_2 could be cheaper.

- But S_1 **stochastically dominates** S_2 $\implies$ S_2 can be discarded.

- Stochastic dominance is best detected from the respective *cumulative* probability distributions for the costs of S_1 and S_2:



- If actions A_1 and A_2 lead to probability distributions $p_1(x)$ and $p_2(x)$ on attribute X, then A_1 stochastically dominates A_2 on X if and only if for all x, $\int_{-\infty}^{x} p_1(y)dy \leq \int_{-\infty}^{x} p_2(y)dy$.
- In many cases, stochastic dominance is easily detected. E.g., construction costs depend on the distance to the city center.

Preferences without Uncertainty

- Attributes X_1 and X_2 are **preferentially independent** of a third attribute X_3 if the preference between outcomes $\langle x_1, x_2, x_3 \rangle$ and $\langle x_1', x_2', x_3 \rangle$ is independent of the particular value x_3 of X_3.
- **Mutual preferential independence** (MPI) of $X_1, \ldots, X_n$: each *pair of variables* is preferentially independent from others.
- If attributes $X_1, \ldots, X_n$ are mutually preferentially independent, then the agent's behavior can be described as maximizing

  $$V(S) = \sum_{i=1}^{n} V_i(X_i(S))$$

  where each V_i is a value function referring only to X_i.
- A value function like $V(S)$ is called an **additive value function**.

Preferences with Uncertainty

- **Utility independence** extends preferential independence to cover lotteries: a set of attributes $\mathbf{X}$ is utility-independent of $\mathbf{Y}$ if lotteries involving $\mathbf{X}$ are independent of the particular values of $\mathbf{Y}$.
- A set of attributes $\mathbf{X}$ is **mutually utility-independent** (MUI) if each subset $\mathbf{Y} \subseteq \mathbf{X}$ is utility-independent of $\mathbf{X} - \mathbf{Y}$.
- If MUI holds, the agent's behavior can be described in terms of a **multiplicative utility function**. For three attributes, $U_i =$

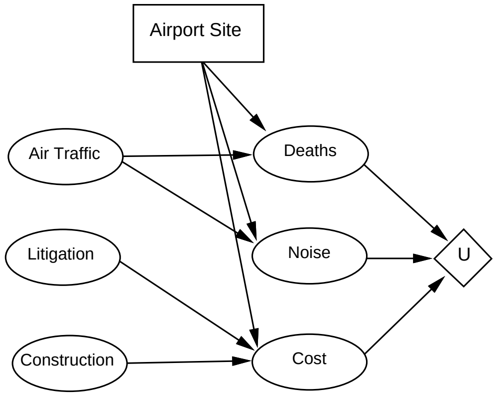
  $k_1U_1 + k_2U_2 + k_3U_3 + k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_1k_3U_1U_3 + k_1k_2k_3U_1U_2U_3$

  where U_i denotes $U_i(X_i(S))$ for $i \in \{1,2,3\}$.
- In general, an n-attribute problem exhibiting MUI can be modeled using n single-attribute utilities and n constants.

5. DECISION NETWORKS

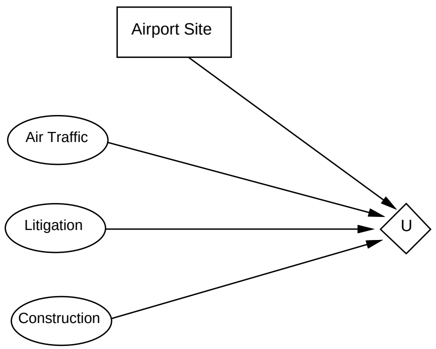
- **Decision networks** (or **influence diagrams**) extend Bayesian networks with additional nodes for actions and utilities:
  1. **Chance nodes** (ovals) represent random variables with CPTs.
  2. **Decision nodes** (rectangles) represent points where the decision-maker has a choice of actions to perform.
  3. **Utility nodes** (diamonds) represent the agent's utility function (a tabulation of the agent's utility as a function of attributes).
- Chance nodes (as well as utility nodes) may have both chance nodes and decision nodes as parents.
- We concentrate on decision networks with a single decision node.

Example. Consider the airport siting problem. In addition to the choice being made, factors including *AirTraffic*, *Litigation*, and *Construction* affect utility indirectly via *Deaths*, *Noise*, and *Cost*.



➤ A way to simplify a decision network is to represent the *expected* utility of actions using **action-utility tables**.

Example. The decision network for the airport siting problem can be simplified by factoring out chance nodes describing outcome states:



☞ Less flexible to update!

Evaluating Decision Networks

The algorithm for evaluating a decision network in the following:

1. Set the evidence variables for the current state.
2. For each possible value of the decision node:
   (a) Set the decision node to that value (like any evidence variable).
   (b) Calculate the posterior probabilities for the parent nodes of the utility node using standard probabilistic inference algorithms.
   (c) Calculate the resulting utility for the action.
3. Return the action with the highest utility.

➤ We will later consider the possibility of executing several actions in sequence which makes the problem much more interesting.

6. THE VALUE OF INFORMATION

➤ One of the most important parts of decision making is knowing what questions to ask to obtain all relevant information.

Example. A doctor cannot expect to be provided with the results of all possible diagnostic tests when meeting a patient.

➤ The **value of information** is the difference between the expected utilities of the best actions before and after obtaining information.

➤ The acquisition of information is achieved by **sensing actions**.

➤ **Information value theory** is a form of sequential decision making.

Example. An oil company is willing to buy one of n indistinguishable blocks of ocean drilling rights. The setting is as follows:

1. There are n blocks for sale.
2. Exactly one block contains oil worth C€.
3. The price of a single block is $\frac{C}{n}$€.

A seismologist offers the company the results of a survey of block 3.

- How much is the company willing to pay for knowing the results?
- The expected value of this piece of information is

$$\frac{1}{n}(C-\frac{C}{n})+\frac{n-1}{n}(\frac{C}{n-1}-\frac{C}{n})=\frac{C}{n} \qquad (€).$$

- The information is is worth as much as the block itself!

A General Formula

- It is expected that the exact value of some random variable E_j is obtained: hence the term **value of perfect information** (VPI).
- The utility $\mathrm{EU}(\alpha|E)$ of the **current best action** α is defined by

$$\max_A \sum_i U(Result_i(A))P(Result_i(A) \mid E, Do(A)).$$

- Given a piece of evidence E_j this becomes $\mathrm{EU}(\alpha_{E_j} \mid E, E_j) =$

$$\max_A \sum_i U(Result_i(A))P(Result_i(A) \mid E, Do(A), E_j).$$

- But the value of E_j is currently *unknown*, and we have to average over all possible values e_{jk} of E_j. Thus $\mathrm{VPI}_E(E_j) =$

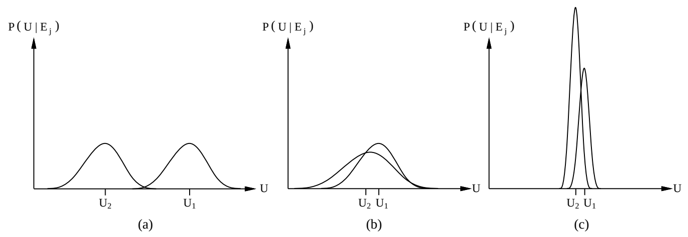
$$(\sum_k P(E_j = e_{jk} \mid E)\mathrm{EU}(\alpha_{e_{jk}} \mid E, E_j = e_{jk})) - \mathrm{EU}(\alpha \mid E).$$

Example. Consider different routes through a mountain range.

(a) A straight highway through a low pass (action A_1) is clearly preferable to a winding dirt road over the top (action A_2)

(b) The choice between two different winding dirt roads of slightly different lengths – each of which may be blocked or not.

(c) The differences are likely to be small in summertime.



☞ Additional information becomes valuable in the case (b).

☞ "*Information has value to the extent that it is likely to cause a change of plan, and to the extent that the new plan will be significantly better than the old plan*".

Properties of the Value of Information

The value of perfect information shares the following properties:

1. *Nonnegativeness*: $\mathrm{VPI}_E(E_j) \geq 0$.
2. *Nonadditivity* (VPI depends on the evidence E obtained so far):
$$\mathrm{VPI}_E(E_j, E_k) \neq \mathrm{VPI}_E(E_j) + \mathrm{VPI}_E(E_k).$$
3. *Order-independence*:
$$\begin{aligned}\mathrm{VPI}_E(E_j, E_k) &= \mathrm{VPI}_E(E_j) + \mathrm{VPI}_{E,E_j}(E_k)\\ &= \mathrm{VPI}_E(E_k) + \mathrm{VPI}_{E,E_k}(E_j).\end{aligned}$$

Implementing an Information-Gathering Agent

- For now, it is assumed that with each observable evidence variable E_j, there is an associated cost $Cost(E_j)$ of obtaining E_j via tests.
- An information gathering agent should request the most valuable piece of information E_j compared to $Cost(E_j)$:

```
function INFORMATION-GATHERING-AGENT(percept) returns an action
   static: D, a decision network

   integrate percept into D
   j ← the value that maximizes VPI(E_j) − Cost(E_j)
   if VPI(E_j) > Cost(E_j)
       then return REQUEST(E_j)
   else return the best action from D
```

- The procedure implements **myopic** information gathering, since VPI is short-sightedly applied to single pieces of evidence.

DECISION-THEORETIC EXPERT SYSTEMS

The knowledge engineering process for a decision-theoretic system:

1. Determine the scope of the problem (decide nodes).
2. Lay out the topology of the network (analyze dependencies).
3. Assign probabilities to chance nodes.
4. Assign utilities to utility nodes.
5. Enter available evidence to the network.
6. Evaluate posterior probabilities and utilities for the nodes.
7. Gather new evidence using value of information as a criterion.
8. Perform sensitivity analysis for the assigned probabilities/utilities.

SUMMARY

- **Decision theory** = **probability theory** + **utility theory**.
- A **rational agent** considers all possible actions and chooses the one that leads to the best expected outcome.
- **Decision networks** – a generalization of Bayesian networks – provide a simple formalism for expressing and solving decision problems.
- The **value of information** is defined as the expected improvement in utility compared to making a decision without the information.
- **Expert systems** that incorporate utility information have additional capabilities compared to pure inference systems.

QUESTIONS

Recall the domain of soccer playing agents and formalize a ball tracking system using a Bayesian network with the following variables:

Variable	Values	Explanation
Tired	*True*, *False*	Is the agent feeling tired?
Angle	*Left*, *Center*, *Right*	Angle with respect to the ball
Distance	*Far*, *Close*, *Touch*	Distance to the ball

- For each variable X of these, introduce an additional variable X_{next} referring to the outcome of actions available to the agent:

  TurnLeft, *TurnRight*, *Run* and *Nop*.

- Add a utility node that depends on $Tired_{\text{next}}$, $Angle_{\text{next}}$, and $Distance_{\text{next}}$. Define a utility function based on these attributes.