Artificial Intelligence 2

Roman Barták
Department of Theoretical Computer Science and Mathematical Logic

Rational decisions

We are designing **rational agents** that maximize expected utility.

Probability theory is a tool for dealing with degrees of belief (about world states, action effects etc.).

Now, we explore **utility theory** to represent and reason with preferences.

Finally, we combine preferences (as expressed by utilities) with probabilities in the general theory of rational decisions – **decision theory**.

Utility theory

The agent‘s preferences are captured by a **utility function,** U(s), which assigns a single number to express desirability of a state.

The **expected utility** of an action given the evidence is just the average value of outcomes, weighted by their probabilities

$$EU(a|\mathbf{e}) = \sum_s P(Result(a)=s|a,\mathbf{e})\ U(s)$$

A rational agent should choose the action that **maximizes** the agent’s **expected utility** (MEU)

$$action = argmax_a\ EU(a|\mathbf{e})$$

The MEU principle formalizes the general notion that the agent should “do the right thing”, but we need make it operational.

Rational preferences

Frequently, it is easier for an agent to express **preferences** between states:

- A > B: the agent prefers A over B
- A < B: the agent prefers B over A
- A ~ B: the agent is indifferent between A and B

What sort of things are A and B?

- They could be states of of the world, but more often tan not there is uncertainty about what is really being offered.
- We can think of the set of outcomes for each action as a **lottery** (possible outcomes $S_1, \ldots, S_n$ that occur with probabilities $p_1, \ldots, p_n$)
  - $[p_1,S_1; \ldots; p_n,S_n]$

An example of lottery (food in airplanes)
Chicken or pasta?

- [0.8, juicy chicken; 0.2, overcooked chicken]
- [0.7, delicious pasta; 0.3, congealed pasta]

Properties of rational preferences

Rational preferences should lead to maximizing expected utility (if the agent violates them it will exhibit patently irrational behavior in some situations.

We require several constraints (the **axioms of utility theory**) that rational preferences should obey.

- **orderability**:
  exactly one of $(A > B)$ or $(A < B)$ or $(A \sim B)$ holds
- **transitivity**:
  $(A < B) \land (B < C) \Rightarrow (A < C)$
- **continuity**:
  $(A > B > C) \Rightarrow \exists p\ [p,A;\ 1\text{-}p,C] \sim B$
- **substitutability**:
  $A \sim B \Rightarrow [p,A;\ 1\text{-}p,C] \sim [p,B;\ 1\text{-}p,C]$
- **monotonicity**:
  $A > B \Rightarrow (p > q \Leftrightarrow [p,A;\ 1\text{-}p,B] > [q,A;\ 1\text{-}q,B]$
- **decomposability**:
  $[p,A;\ 1\text{-}p,\ [q,B;\ 1\text{-}q,C]] \sim [p,A;\ (1\text{-}p)q,B;\ (1\text{-}p)(1\text{-}q),C]$

Preferences lead to utility

Teh axioms of utility theory are axioms about preferences but we can derive the following consequences form them..

Existence of utility function such that :

$U(A) < U(B) \Leftrightarrow A < B$

$U(A) = U(B) \Leftrightarrow A \sim B$

Expected utility of a lottery:

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

A utility function exists for any rational agent but it is not unique:

$U'(S) = a\ U(S) + b$

Existence of a utility function does not necessarily mean that the agent is explicitly maximizing that utility function. By observing its preferences an observer can construct the utility function (even if the agent does not know it).

Utility functions

Utility is a function that maps from lotteries to real numbers.

We must first work out what the agent‘s utility function is (**preference elicitation**).

- We will be looking for a **normalized utility function**.
- We fix the utility of a “best possible prize” S_{max} to 1, $U(S_{max}) = 1$.
- Similarly, a “worst possible catastrophe” S_{min} is mapped to 0, $U(S_{min}) = 0$.
- Now, to assess the utility of any particular prize S we ask the agent to choose between S and a **standard lottery** $[p, S_{max}; 1\text{-}p, S_{min}]$
- The probability p is adjusted until the agent is indifferent between S and the standard lottery.
- Then the utility of S is given by, $U(S) = p$.

The utility of money

Universal exchangeability of money for all kinds of goods and services suggests that money plays a significant role in human utility functions.

- An agent prefers more money to less, all other things being equal.

But this does not mean that money behaves as a utility function (because it says nothing about preferences between lotteries involving money).

Assume that you won a competition and the host offers you a choice: either you can take the 1 mil. USD price or you can gamble it on the flip of coin. If the coin comes up heads, you end up with nothing, but if it comes up tails, you get 2.5 mil. USD.
What is your choice?

- Expected monetary value of the gamble is 1.250.000 USD.
- Most people decline the gamble and pocket the million. Are they being irrational?

The utility of money

The decision in the previous game does not depend on the prize only but also on the wealth of the player!

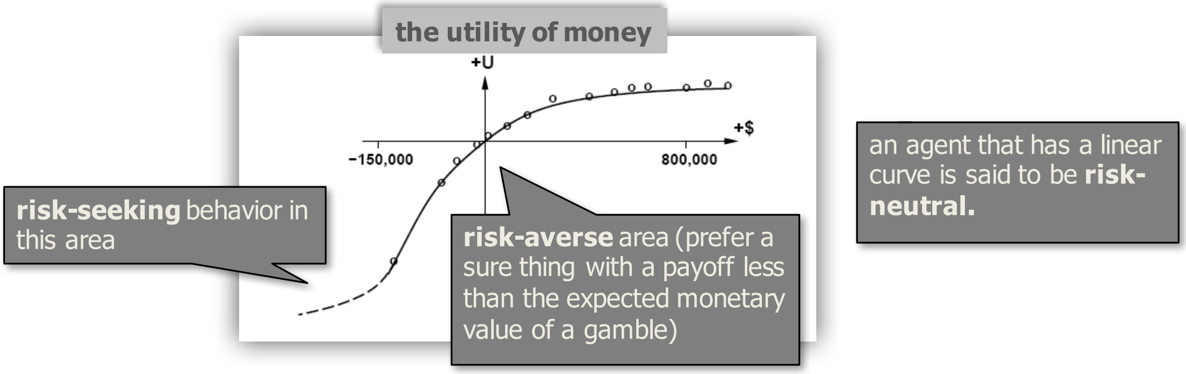
Let S_n denote a state of possessing total wealth n USD, and the current wealth is k USD.

The the expected utilities of two actions are:

- EU(Accept) = ½ $U(S_k)$ + ½ $U(S_{k+2.500.000})$
- EU(Decline) = $U(S_{k+1.000.000})$

Suppose we assign $U(S_k) = 5$, $U(S_{k+1.000.000}) = 8$, $U(S_{k+2.500.000}) = 9$.

Then the rational decision would be to decline!



Human judgment (certainty effect)

The evidence suggests that humans are “predictable irrational “.

Allais paradox

- A: 80% chance of 4000 USD
- B: 100% chance of 3000 USD

  What is your choice?
  - Most people consistently prefer B over A (taking the sure thing!)
- C: 20% chance of 4000 USD
- D: 25% chance of 3000 USD

  What is your choice?
  - Most people prefer C over D (higher expected monetary value)

Certainty effect – people are strongly attracted to gains that are certain

Ellsberg paradox

The urn contains 1/3 red balls, and 2/3 either black or yellow balls.

- A: 100 USD for a red ball
- B: 100 USD for a black ball

  What is your choice?

  - Most people prefer A over B (A gives a 1/3 chance of winning, while B could be anywhere between 0 and 2/3)
- C: 100 USD for a red or yellow ball
- D: 100 USD for a black or yellow ball

  What is your choice?

  - Most people prefer D over C (D gives you a 2/3 chance, while C could anywhere between 1/3 and 3/3)

However, if you think there are more red than black balls then your should prefer A over B and C over D.

Ambiguity aversion – most people elect the known probability rather than the unknown unknown.

Framing effect – the exact wording of a decision problem can have a big impact on the agent‘s choices

- medical procedure A has 90% survival rate
- medical procedure B has 10% death rate

  What is your choice?

  - Most people prefer A over B though both choices are identical

Anchoring effect – people feel more comfortable making relative utility judgments rather than absolute ones

- The restaurant takes advantage of this by offering a $200 bottle that it knows nobody will buy, but which serves to skew upward the customer’s estimate of the value of all wines and make the $55 bottle seem like a bargain.

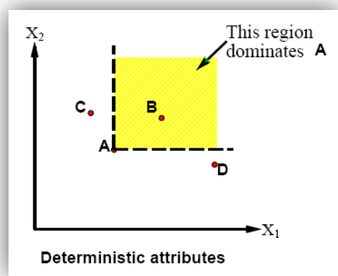
In real-life the outcomes are characterized by two or more attributes such as cost and safety issues – **multi-attribute utility theory**.

We will assume that higher values of an attribute correspond to higher utilities.

The question is how to get preferences for more attributes

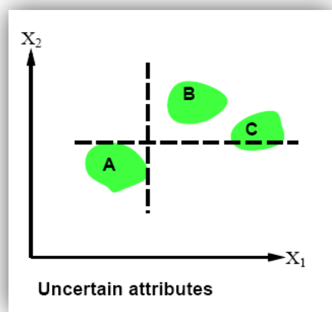
- without combining the attribute values into a single utility value – **dominance**
- **combining the attribute values** into a single utility value

If an option is of lower value on all attributes that some other option, it need not be considered further – **strict dominance**.



Strict dominance can be defined for uncertain outcomes too.

- if all possible outcomes of B strictly dominate all possible outcomes of A



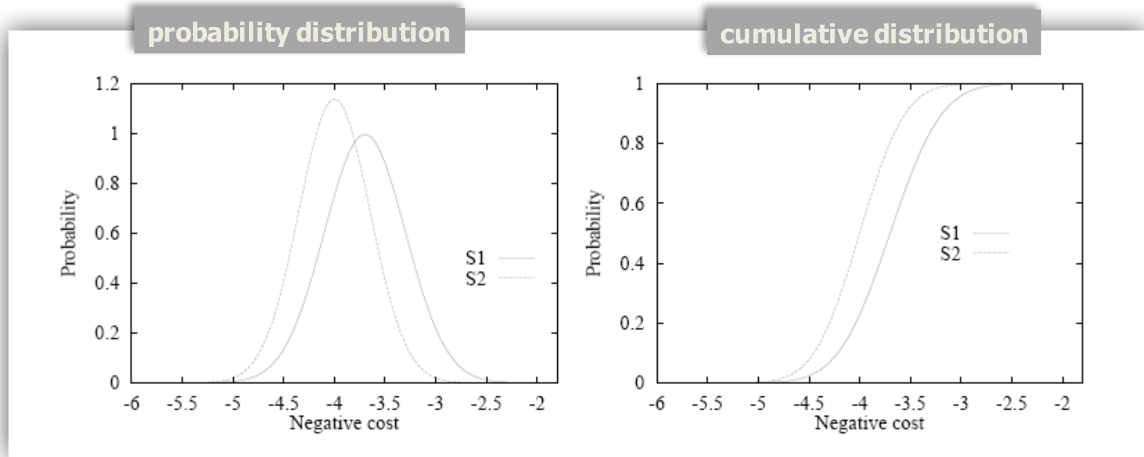
Strict dominance will probably occur less often than in the deterministic case.

Stochastic dominance

Stochastic dominance occurs more frequently in real problems. It is easier to understand in the context of a single variable.

Stochastic dominance is best seen by examining the **cumulative distribution** that measures the probability that that the cost is less than or equal any given amount (it integrates the original distribution).

$$\forall t \quad \int_{-\infty}^{t} p_1(x)dx \leq \int_{-\infty}^{t} p_2(t)dt$$



Preference structure

To specify the **complete utility function** for **n attributes** each having **d values**, we need $\mathbf{d^n}$ values in the worst case.

– This corresponds to a situation in which agent‘s preferences have **no regularity** at all.

Preferences of typical agents have much **more structure** so the the utility function can be expressed as :

$$U(x_1,\dots,x_n) = F[f_1(x_1),\dots,f_n(x_n)]$$

Preference structure (without uncertainty)

The basic regularity is called **preference independence**.

Two 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 $\langle x'_1, x'_2, x_3 \rangle$
does not depend on the particular value x_3.

If each pair of attributes is preferentially independent of any other attribute, we talk about **mutual preferential independence (MPI)**.

If attributes are mutually preferentially independent then the agent's preference behavior can be described as maximizing the function:

$$U(x_1,\ldots,x_n) = \sum_i U_i(x_i)$$

A value function of this type is called an **additive value function**.

Preference structure (with uncertainty)

When uncertainty is present we need to consider the structure of preferences between lotteries.

For **mutually utility independent (MUI)** attributes we can use **multiplicative utility function**:

$$\begin{aligned} U = {} & k_1U_1 + k_2U_2 + k_3U_3 \\ & + k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_1k_3U_1U_3 \\ & + k_1k_2k_3U_1U_2U_3 \end{aligned}$$

For n attributes exhibiting MUI we can represent the utility function using n constants and n single-attribute utilities.

The value of information

So far we have assumed that all relevant information is provided to the agent before it makes its decision.

In practice, this is hardly ever the case. For example, a doctor cannot expect to be provided with the result of all possible diagnostic tests.

One of the most important parts of decision making is knowing what questions to ask.

We will now look at **information value theory**, which enables an agent to choose which information to acquire.

The value of information (example)

Suppose an oil company is hoping to buy one of the n indistinguishable blocks of ocean-drilling rights.

Let us assume further that exactly one of the blocks contain oil worth C dollars, while others are worthless. The asking price of each block is C/n.

Expected monetary value of buying one block is **C/n – C/n = 0**.

Now suppose that a seismologist offers the company the result of a survey of one specific block, which indicates definitely whether the block contains oil.

How much should the company to pay for that information?

- With probability 1/n, the survey will indicate oil in a given block and the the company will buy it and make a profit **C – C/n**.
- With probability (n-1)/n, the survey will show that the block contains no oil, in which case the company will buy another block. Now the probability of finding oil in that other block is 1/(n-1), so the expected profit is **C/(n-1) – C/n**.
- Together the expected profit given the survey information is:
  1/n (C – C/n) + (n-1)/n (C/(n-1) – C/n) = C/n

Therefore the company should be willing to pay the seismologist up to C/n dollars for the information: the information is worth as much as the block itself.

The value of information (a general formula)

We assume that exact evidence can be obtained about the value of some random variable E_j – this is called **value of perfect information** (VPI).

The value of the current best action α (with the initial evidence **e**) is defined by:

$$EU(\alpha \mid \mathbf{e}) = \max_a \sum_{s'} P(Result(a)=s' \mid a,\mathbf{e})\ U(s')$$

The value of the best action α_{jk} after the new evidence $E_j = e_{jk}$ is obtained is defined by :

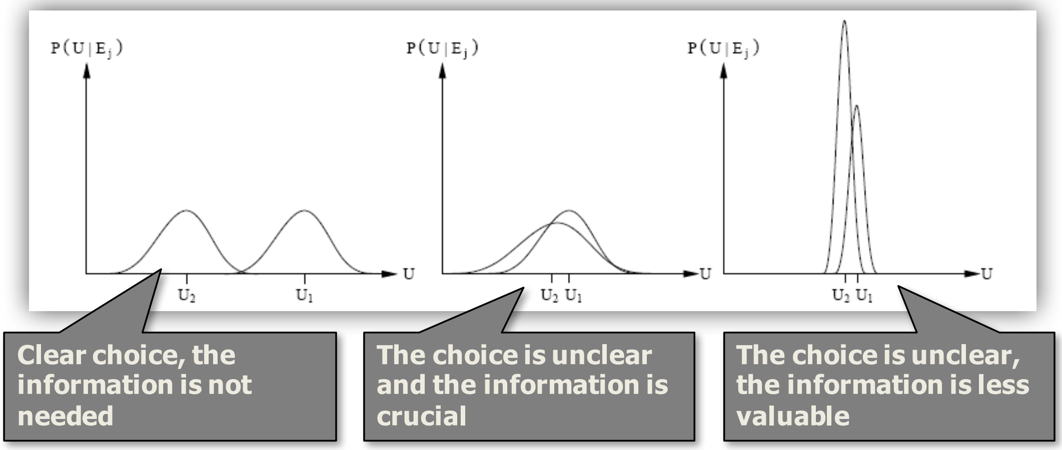
$$EU(\alpha_{jk} \mid \mathbf{e}, E_j=e_{jk}) = \max_a \sum_{s'} P(Result(a)=s' \mid a,\mathbf{e}, E_j=e_{jk})\ U(s')$$

But the value of E_j is currently unknown so we must average over all possible values that we might discover for E_j:

$$VPI_{\mathbf{e}}(E_j) = \left(\sum_k P(E_j = e_{jk} \mid \mathbf{e})\ EU(\alpha_{jk} \mid \mathbf{e}, E_j=e_{jk})\right) - EU(\alpha \mid \mathbf{e})$$

The value of information (qualitatively)

When is it beneficial to obtain new information ?



Information has value to the extend that

– it is likely to cause a change of a plan and
– the new plan will be significantly better that the old plan.

Properties of the value of information

Is it possible for information to be deleterious?

The expected value of information is **nonnegative**.

$\forall \mathbf{e}, E_j \; VPI_{\mathbf{e}}(E_j) \geq 0$

The value of information is **not additive**.

$VPI_{\mathbf{e}}(E_j,E_k) \neq VPI_{\mathbf{e}}(E_j) + VPI_{\mathbf{e}}(E_k)$

The expected value of information is **order independent**.

$VPI_{\mathbf{e}}(E_j,E_k) = VPI_{\mathbf{e}}(E_j) + VPI_{\mathbf{e},ej}(E_k) = VPI_{\mathbf{e}}(E_k) + VPI_{\mathbf{e},ek}(E_j)$

Information gathering

A sensible agent should

- ask questions in a reasonable order
- avoid asking questions that are irrelevant
- take into account the importance of each piece of information in relation its cost
- stop asking questions when that is appropriate

We assume that with each observable evidence variable E_j, there is an associated cost, $Cost(E_j)$.

Information-gathering agent can select (greedily) the most efficient observation until no observation worth its costs (myopic approach)

```
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
```

Department of Theoretical Computer Science and Mathematical Logic
bartak@ktiml.mff.cuni.cz