Artificial Intelligence
4 – Utility Theory

Arthur Bit-Monnot

INSA 4IR

Context: Making Simple Decisions

- **probability theory**: what an agent believes
- **utility theory**: what an agent desires

Decision theory: making rational decision based on the beliefs and desires of the agent

Section 1

Agents Beliefs

Uncertainty in the current state

An agent is necessarily in a state, denoted s^*

Under partial observability, an agent

- may not know precisely what the state s^* is,
- only capable of isolating a set S_{belief} of possible states it believes it may be in.

$$s^* \in S_{belief} \subseteq S$$

Agents will typically have an explicit or implicit characterization of the likelihood of each state: We note

- $P(s)$ the probability of being in state s, as believed by the agent (probability that $s^* = s$)

Uncertainty in the current state of a coin

- S contains
  - S_{HH} = (on=hand, side=head)
  - S_{HT} = (on=hand, side=tails)
  - S_{FH} = (on=floor, side=head)
  - S_{FT} = (on=floor, side=tails)

After throw, I know the coin is on my hand but not on which side is landed:

- S_{belief} contains
  - S_{HH} = (on=hand, side=head)
  - S_{HT} = (on=hand, side=tails)

Probability distribution:

- $P(S_{HH}) = 0.5$
- $P(S_{HT}) = 0.5$
- $P(S_{FH}) = P(S_{FT}) = 0$

Uncertainty in actions' outcome

There may also be uncertainty in the outcome come of an action

- $P(s'|s,a)$: (transition model)
  - **probability** of ending in s'
  - knowing that I am in s
  - given that I do action a

- $P(\textsc{Result}(a) = s') = \sum_s P(s) \times P(s'|s,a)$
  - **probability** of ending in state s'
  - after doing action a
  - integrating uncertainty on the current state

Uncertainty in actions' outcome

s	a	s'	$P(s'\|s,a)$	Comment
s_{HH}	flip	s_{HH}	0.45	90% chances to land on hand, evenly split head/tails
s_{HH}	flip	s_{HT}	0.45	
s_{HH}	flip	s_{FH}	0.05	10% chances to land on floor, evenly split H/T
s_{HH}	flip	s_{FT}	0.05	
–	–	–	–	
s_{HT}	flip	s_{HT}	0.45	Symmetric when starting from tail position
s_{HH}	flip	s_{HT}	0.45	
s_{HH}	flip	s_{FH}	0.05	
s_{HH}	flip	s_{FT}	0.05	
–	–	–	–	
s_{FT}	flip	s_{FT}	1	No-op action (i.e. does not affect the state)
s_{FT}	flip	s_{HT}	0	

Expected utility

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

The **expected utility** of an action a is just the average utility of the outcome, weighted by the probability that the outcome occurs:

$$EU(a) = \sum_{s'} P(\textsc{Result}(a) = s')U(s')$$

The principle of **maximum expected utility** (MEU) states that a rational agent should choose the action that maximizes the agent's expected utility:

$$action = \arg\max_{a} EU(a)$$

Challenges of decision theory

Maximum expected utility formalizes the general notion of rational agents.

Exploiting it however comes with challenges, notably:

- estimating $P(s)$ (probability distribution of the current state)
  - requires perception, learning, knowledge representation, inference
- computing $P(s'|s,a)$ (action consequences)
  - requires causal model of the environment
- computing $U(s')$ (utility of any state)
  - requires planning/search (what's the utility of an intermediate step over long term objectives)

Section 2

Utility Theory

Lottery

A situation is a given state of the environment

- e.g., \$1: I have one dollar[1]

A **lottery** is a set of **outcomes** $(S_1, S_2, \dots, S_n)$ each with a **probability** $(p_1, p_2, \dots, p_n)$ of occurring.

$$L = [p_1, S_1 \parallel p_2, S_2 \parallel \dots \parallel p_n, S_n]$$

where an outcome can be either:

- a situation, e.g., a lottery where I win either \$10 or \$20 depending on a coin flip
  - $[0.5, \$10 \parallel 0.5, \$20]$
- another lottery, e.g., when I need to win two coin flips to get \$20
  - $\left[0.5, \$0 \parallel [0.5, \$0 \parallel 0.5, \$20]\right]$

[1] Simplified: this is assuming that I have initially \$0

Preferences

Over situations and lotteries, an agent will have preferences:

- $A \succ B$: the agent prefers A over B
- $A \sim B$: the agent is indifferent A and B
- $A \succsim B$: the agent prefers A over B or is indifferent

Axioms of utility theory: the preferences of any rational agent must follow some rules (orderability, transitivity, ...)

- otherwise, one can show that the agent would display some pathologically irrational behavior

Rational preferences: Orderability

Given any two lotteries, the agent either prefer one or is indifferent

I.e., Exactly one on $(A \succ B)$, $(B \succ A)$ *or* $(A \sim B)$ *holds*

- the agent cannot avoid deciding

Rational preferences: Transitivity

Given any three lotteries, if an agent prefers A to B and B to C, then the agent must prefer A to C

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

Rational preferences: Continuity

If some lottery B is between A and C in preferences, then there is a probability p for which the rational agent will be indifferent between 1) getting B for sure and 2) the lottery than gives A with probability p and C with probability $1-p$

$$A \succ B \succ C \Rightarrow \exists p \text{ such that } [p, A \parallel 1-p, C] \sim B$$

- If $\$100 \succ \$40 \succ \$0$, there exists a lottery where the agent would be indifferent in
  - getting \$40 for sure
  - getting \$100 with probability p (and \$0 otherwise)

Rational preferences: Substitutability

If an agent is indifferent between (resp. prefers) A and B, then it must be indifferent between (resp. prefers) two more complex lotteries that only differ by A being substituted by B

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

Example (with preference): in a gambling game, you would prefer a new lottery, that replaces a price of \$10 by a price of \$100

Rational preferences: Monotonicity

If two lotteries have the same outcomes A and B and the agent prefers A over B then the agent must prefer the lottery with the highest probability for A

$$A \succ B \Rightarrow (p > q \Leftrightarrow [p, A \parallel 1-p, B] \succ [q, A \parallel 1-q, C])$$

All other things identical, I would prefer the lottery where the high-price is more probable

Section 3

From rational preferences to utility

Existence of a utility function[2]

If an agent's preferences **obey the axioms of utility**, then, there **exists a function** U such that:

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

I.e., there is a function that captures the preferences of an agent by assigning a single numeric value to each situation.

[2] Theory of Games and Economic Behavior (1944), von Neumman & Morgenstern

Expected utility of a lottery

The utility of a lottery is the sum of the utilities of the outcomes, weighted by there probability:

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

Corollary: if the agent knows

- the probabilities of the lotteries
- the utility of each outcome

it can compute the utility of any lottery.

Acting with utilities

A non-deterministic action is a lottery:

- several possible outcomes,
- each with probability of occurring

$$EU(a) = \sum_{s'} P(\textsc{Result}(a) = s')U(s')$$

An agent can act rationally by:

- computing the expected utility of each action,
- selecting the action with the maximum expected utility

$$action = \arg\max_a EU(a)$$

Exercise

- I have two available actions: a_1 and a_2 which can result in one of three outcomes A, B, C
- I know:
  - the utility of each outcome S: $U(S)$
  - the probability of having outcome S for each action a: $P(S|a)$

S	U(S)
A	9
B	4
C	6

a	S	$P(S\|a)$
a_1	A	0.5
a_1	B	0.5
a_1	C	0
a_2	A	0
a_2	B	0
a_2	C	1

Which action should I choose?

Section 4

Utility functions

(Non-)uniqueness of the utility function

We have a function $U(S)$ consistent with the agent's preferences, is this function unique?

(Non-)uniqueness of the utility function

We have a function $U(S)$ consistent with the agent's preferences, is this function unique?

$\Rightarrow$ No, a function

$$U'(S) = 283 \times U(S) + 74.6$$

would also be consistent with the agent's preferences (preserves order, equality and distribution of probabilities).

In fact, any function affine function $U'(S) = a \times U(S) + b$ would work (assuming $a > 0$)

Normalized utility

A utility function can be normalized by:

- setting a utility $u_\perp = 0$ to the worst possible outcome
- setting a utility $u_\top = 1$ to the best possible outcome

All utilities would be in the $[0, 1]$ range.

Preferences elicitation

Preference elicitation: presenting choices to the human and using the observed preferences to determine the underlying utility function.

Scenario:

- worst outcome: win \$0 (with utility $u_\perp = 0$)
- best outcome: win \$100k (with utility $u_\top = 1$)

What's the utility of winning \$50k ?

Preferences elicitation

Preference elicitation: presenting choices to the human and using the observed preferences to determine the underlying utility function.

Scenario:

- worst outcome: win \$0 (with utility $u_\perp = 0$)
- best outcome: win \$100k (with utility $u_\top = 1$)

What's the utility of winning \$50k ?

Propose a series of lotteries between two outcomes with known utilities (e.g. best/worst) and ask the human to pick its preferred one

- $L(p) = [p, \$100k \parallel (1-p), \$0]$ vs $\$50k$

Preferences elicitation

Preference elicitation: presenting choices to the human and using the observed preferences to determine the underlying utility function.

Scenario:

- worst outcome: win \$0 (with utility $u_\perp = 0$)
- best outcome: win \$100k (with utility $u_\top = 1$)

What's the utility of winning \$50k ?

Propose a series of lotteries between two outcomes with known utilities (e.g. best/worst) and ask the human to pick its preferred one

- $L(p) = [p, \$100k \parallel (1-p), \$0]$ vs $\$50k$

probability (p)	0.2	0.4	0.6	0.8
preference	L(p)	L(p)	$\sim$	\$50k

With these preferences, we can infer that the human has the same utility of 0.6 for \$50k and the lottery with p=0.6.

Utility of money

- Utility theory is rooted in economics
- money is almost universally exchangeable[3]

What is the utility of a $100 bill?

[3] There are some things money can't buy; for everything else, there's... money

Utility of money

The value of money comes primarily from what it can buy us.

with $100, I can buy:

- food for the week (critical)
- a cheap smartphone (important)
- a super nice bottle champagne (nice to have?)

Utility of money

The value of money comes primarily from what it can buy us.

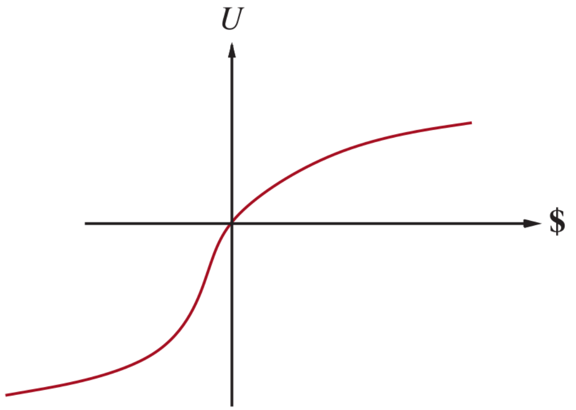
with \$100, I can buy:

- food for the week (critical)
- a cheap smartphone (important)
- a super nice bottle champagne (nice to have?)

Surely, having food for the week is critical, but if I am wealthy I probably have that covered already.

$\Rightarrow$ the utility of what I can buy with \$100 **more** decreases with the amount I already have.

Utility of money



Typical shape of the utility of money

(the actual function will change between individuals)

Risk-aversion

The *decreasing utility/\$* makes the agent **risk-averse**

- prefer a sure thing with a pay off

A typical person would accept \$400 would in place of gamble that would give \$1000 with probability 0.5

$\Rightarrow$ willing to “loose” a \$100 expected dollar.

Decreasing risk-aversion

The risk-aversion for a given gamble decreases with the amount of money you have (as the function is closer to linear)

E.g., a very wealthy person would require $490 to turn off the same bet.

Insurance premium

The decreasing risk-aversion is the basis of the insurance market:

- I am indifferent between
  - paying $100
  - taking the tiny risk of losing the monetary value of my car ($15k)
- The (very wealthy) insurance company is indifferent between:
  - loosing $10
  - taking the same tiny risk of losing $15k

Both would be strictly happier if the insurance company would take the risk for a price in $[\$11, \$99]$[4]

[4] The price proposed/accepted would then depend on the market competition.

Section 5

The Value of information

The Value of information

In practice, an agent rarely has all the relevant information before making a decision.

One of the most important part of decision-making is knowing what question to ask.

A simple example

An oil company is considering buying one of three ocean-drilling rights.

- exactly one of the areas contains oil that would generate a profit of $300m.
- the cost of a drilling right is $100m
- the oil company is risk-neutral (utility proportional to profit)

- What is the expected profit of buying a block?

A simple example

An oil company is considering buying one of three ocean-drilling rights.

- exactly one of the areas contains oil that would generate a profit of $300m.
- the cost of a drilling right is $100m
- the oil company is risk-neutral (utility proportional to profit)

- What is the expected profit of buying a block?

$$0 = \frac{1}{3} \times 300m + \frac{2}{3} \times 0 - 100m$$

A simple example

An oil company is considering buying one of three ocean-drilling rights.

- exactly one of the areas contains oil that would generate a profit of $300m.
- the cost of a drilling right is $100m
- the oil company is risk-neutral (utility proportional to profit)

- What is the expected profit of buying a block?

$$0 = \frac{1}{3} \times 300m + \frac{2}{3} \times 0 - 100m$$

- A survey indicates with certainty whether there is oil in the first area. How much should the company be willing to pay for this?

The value of perfect information

X is a random variable that influences the agent's beliefs.

Without knowing the value of X, the agent would select the action α that maximizes the expected utility:

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

If I know that $X = x_i$ is would select the action α_{x_i}

$$EU(\alpha_{x_i}) = \max_a \sum_{s'} P(Result(a) = s'|x_i)U(s')$$

Value of perfect information

The expected utility after the knowing the value of X should take into account all possible outcomes. Averaging the utility of the best action in all situation:

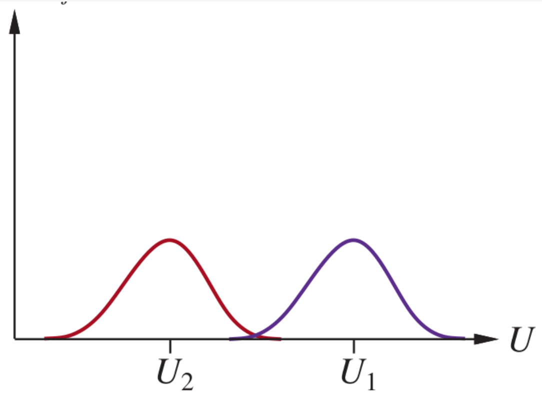
$$\sum_{x_i} P(X = x_i) EU(x_{x_i} | x_i)$$

The **value of perfect information** (VPI) is the difference with the expected utility of the action I would have selected without knowing the value of X:

$$VPI(X) = \Big(\sum_{x_i} P(X = x_i) EU(x_{x_i} | x_i) \Big) - EU(\alpha)$$

Intuitively, the value of information lies in that it may enable us to select another action.

Information Value: Example 1



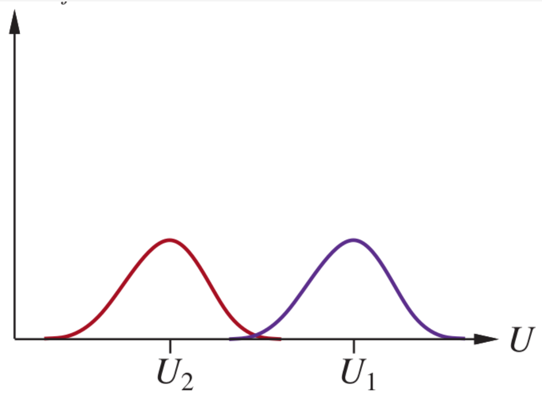
Two actions a_1 and a_2 whose utility has a given probability distribution.

Example 1:

- a_1 take the highway (with uncertainty about the traffic)
- a_2 take the dirt road (uncertainty about the road state)

I can observe the precise status of each road (for a small cost). Should I do it?

Information Value: Example 1



Two actions a_1 and a_2 whose utility has a given probability distribution.

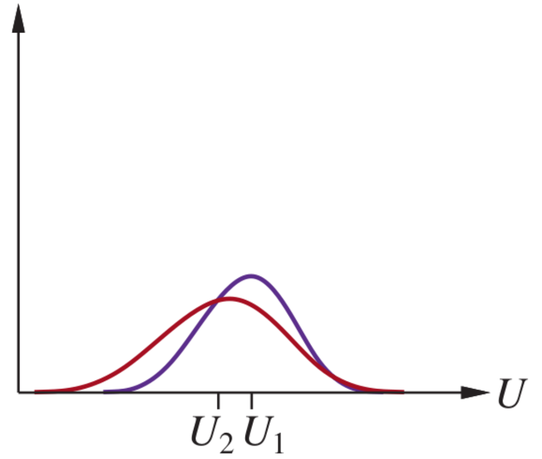
Example 1:

- a_1 take the highway (with uncertainty about the traffic)
- a_2 take the dirt road (uncertainty about the road state)

I can observe the precise status of each road (for a small cost). Should I do it?

- No $\rightarrow$ very unlikely to change my choice (and with minimal impact even then)

Information Value: Example 2

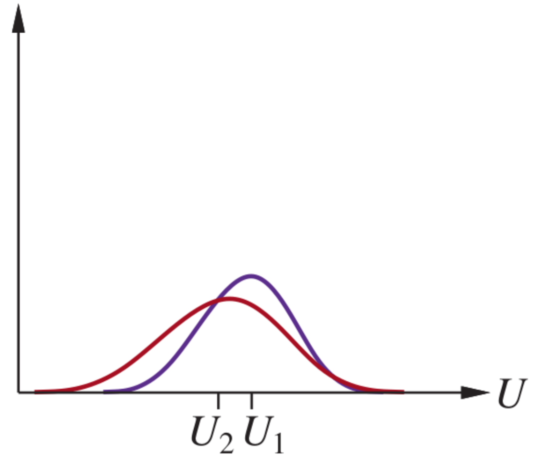


It has been snowing a lot and any road where the snowplow hasn't passed would be blocked

- a_1: the left road
- a_2: the slightly longer right road

I can observe the precise status of each road (for a small cost). Should I do it?

Information Value: Example 2



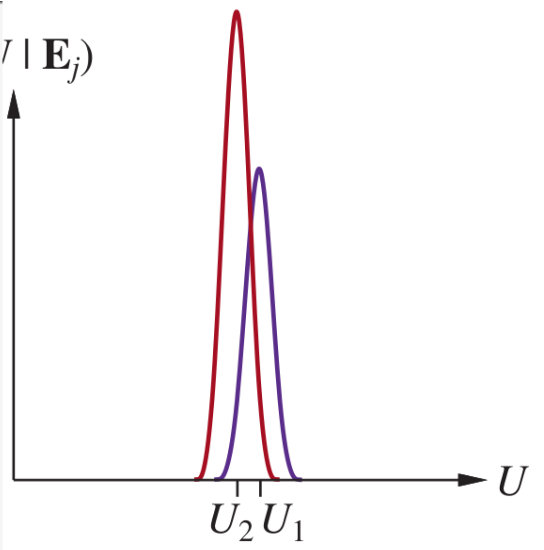
It has been snowing a lot and any road where the snowplow hasn't passed would be blocked

- a_1: the left road
- a_2: the slightly longer right road

I can observe the precise status of each road (for a small cost). Should I do it?

- Yes $\rightarrow$ may change my choice with high impact

Information Value: Example 3

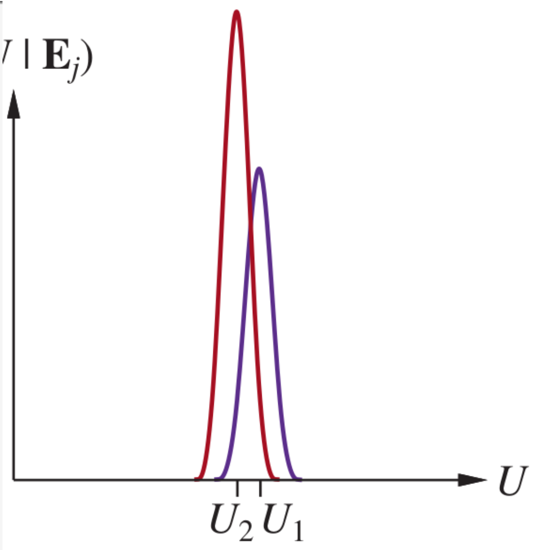


It's sunny and far from the peak of traffic

- a_1: the left road
- a_2: the slightly longer right road

I can observe the precise status of each road (for a small cost). Should I do it?

Information Value: Example 3



It's sunny and far from the peak of traffic

- a_1: the left road
- a_2: the slightly longer right road

I can observe the precise status of each road (for a small cost). Should I do it?

- No $\rightarrow$ may change my choice but with very limited impact.

Section 6

Summary

Summary

- **probability theory** describes what an agent believes
- **utility theory** describes what an agent wants
- **decision theory** combines the two to describe what an agent should do
- an agent that shows **consistent preferences** possesses a **utility function**
- a rational agent can act by selecting the action that **maximizes the expected utility**
- the **value of information** describes the increase of utility gained through information-gathering, *prior to making a decision*

Utility function vs Performance Measure

- performance measure:
    - is only computed once (at the end)
- utility function:
    - is a guide towards good performance
    - means of comparing states
    - often approximated