CS761 Artificial Intelligence

19. Planning with Uncertainty: Utilities and Decisions

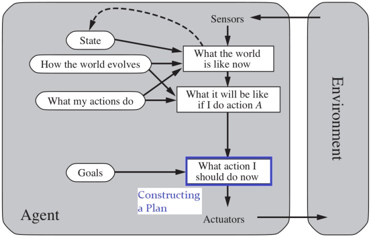
Jiamou Liu
The University of Auckland

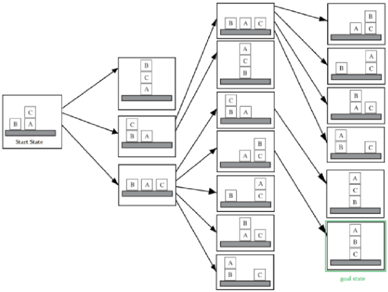
Planning with Uncertainty

Classical planning is a planning task with the following assumptions:

- Finite set of states as the search space.
- State changes defined only by agents' actions.
- Deterministic actions: Each action in a state has one outcome, which can be foreseen by the agent.
- Perfect information.
- Goals must be achieved.

Goal-based agents





Planning with uncertainty differs from classical planning:

- Infinite set of probabilistic outcomes as search space.
- State changes stochastically.
- Non-deterministic actions.
- Imperfect information.
- Goals may not be achieved. Need to evaluate desirability of outcomes.

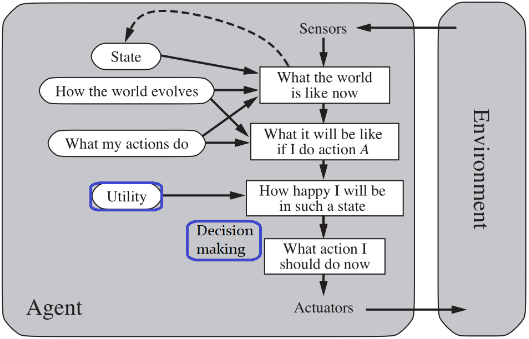
Thus the need to reason about uncertainty using tools like:

- Probability theory: quantifying uncertainty and calculating expectations.
- Bayesian networks: reason about probabilistic outcomes with imperfect information.
- Markov model: reason about discrete time and sequential actions.
- Utility theory: to be covered in this lecture.

Evaluating the desirability of (probabilistic) outcomes:

- The agent hopes to maximise the chance of meeting its goal, i.e., arrive at outcomes that are more desirable.
- **Preferences:** How desirable an outcome is.
- **Utility function:** A numerical value that expresses the agent's preference of a state.

Utility-based agent.

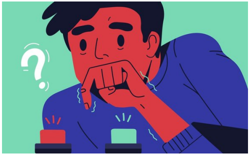


Utility Theory

Question. Can preferences be represented by numbers?

Example 1. Suppose you have \$1000 in savings. Would you risk \$1000 for a 10% chance odds to win \$9000?

Example 2. Would you risk \$1 to enter a lottery with a 0.1% chance odds to win \$999?



Utility Theory

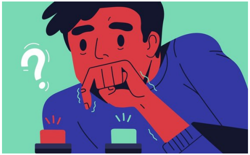
Question. Can preferences be represented by numbers?

Example 1. Suppose you have \$1000 in savings. Would you risk \$1000 for a 10% chance odds to win \$9000?

- Option 1: Keep \$1000. Expected win: 0.
- Option 2: 10% chance of winning \$9000, but 90% chance of losing \$1000. Expected win: $0.1 \times 9000 + 0.9 \times (-1000) = 0$.

Example 2. Would you risk \$1 to enter a lottery with a 0.1% chance odds to win \$999?

- Option 1: Do not enter the lottery. Expected win: 0.
- Option 2: 0.1% chance of winning \$999, but 99.9% chance of losing \$1. Expected win: $0.001 \times 999 + 0.999 \times (-1) = 0$.



Utility theory is the study of utilities and their relations to preferences over uncertain outcomes.

Definition.

- A lottery is a finite distribution over a set $\{a_1, \ldots, a_k\}$, written as

$$[p_1 : a_1, p_2 : a_2, \ldots, p_k : a_k]$$

  where each $s_i \in S$ and $p_i \in \mathbb{R}$ is non-negative with $\sum_i p_i = 1$.

- The set of outcomes Ω contains all lotteries over states S, and all lotteries over outcomes (inductively defined).

Utility theory is the study of utilities and their relations to preferences over uncertain outcomes.

Definition.

- A lottery is a finite distribution over a set $\{a_1, \ldots, a_k\}$, written as

$$[p_1 : a_1, p_2 : a_2, \ldots, p_k : a_k]$$

  where each $s_i \in S$ and $p_i \in \mathbb{R}$ is non-negative with $\sum_i p_i = 1$.
- The set of outcomes Ω contains all lotteries over states S, and all lotteries over outcomes (inductively defined).

- Define the weak preference relation $\succeq$:

$$o_1 \succeq o_2 \text{ if outcome } o_1 \text{ is at least as desirable as outcome } o_2$$

- Define the indifference relation $\sim$:

$$o_1 \sim o_2 \text{ if } o_1 \succeq o_2 \text{ and } o_2 \succeq o_1 \text{, i.e., } o_1 \text{ and } o_2 \text{ are equally preferred.}$$

- Define the strict preference relation $\succ$:

$$o_1 \succ o_2 \text{ if } o_1 \succeq o_2 \text{ and } o_2 \not\succeq o_1.$$

Axioms for Rationality

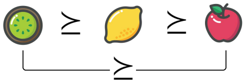
Axioms defining the preferences of a rational agent:

- Axiom 1 Completeness. An agent has preference between all pairs of outcomes:

$$o_1 \succeq o_2 \text{ or } o_2 \succeq o_1$$

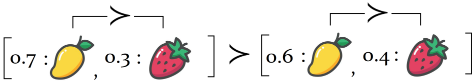
- Axiom 2 Transitivity. Preference must be transitive:

$$(o_1 \succeq o_2 \,\&\, o_2 \succeq o_3) \Rightarrow o_1 \succeq o_3$$



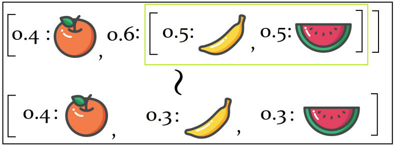
- Axiom 3 Monotonicity. An agent prefers a larger chance of getting a better outcome than a smaller chance of getting the better outcome.

$$\text{if } o_1 \succ o_2 \,\&\, p > q \text{ then } [p : o_1, (1-p) : o_2] \succ [q : o_1, (1-q) : o_2]$$



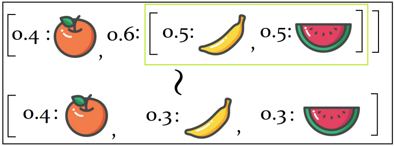
- Axiom 4 Decomposability. An agent is indifferent between lotteries that have the same probabilities over the same outcomes. **E.g.,**

$$[p: o_1, (1-p): [q: o_2, (1-q): o_3]] \sim [p: o_1, (1-p)q: o_2, (1-p)(1-q): o_3]$$



- Axiom 4 Decomposability. An agent is indifferent between lotteries that have the same probabilities over the same outcomes. **E.g.,**

$$[p\colon o_1, (1-p)\colon [q\colon o_2, (1-q)\colon o_3]] \sim [p\colon o_1, (1-p)q\colon o_2, (1-p)(1-q)\colon o_3]$$

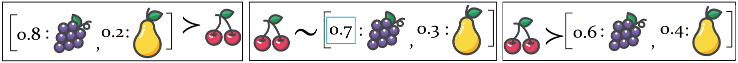


- Axiom 5 Continuity. Suppose $o_1 \succ o_2 \succ o_3$. Then there exists $p \in [0,1]$ such that

$$o_2 \sim [p\colon o_1, (1-p)\colon o_3]$$

Note. For p above, by monotonicity and transitivity, we have

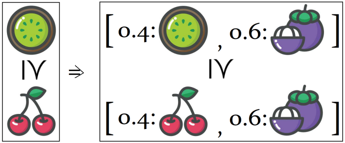
$$\forall p' > p\colon o_2 \succ [p'\colon o_1, (1-p')\colon o_3] \text{ and } \forall p'' < p\colon [p''\colon o_1, (1-p'')\colon o_3] \succ o_2$$



- Axiom 6 Independence. If $o_1 \succeq o_2$, then for any number $p \in [0,1]$ and outcome o_3:

$$[p : o_1, (1-p) : o_3] \succeq [p : o_2, (1-p) : o_3].$$

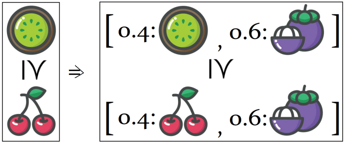
Note: By definition of $\sim$, if $o_1 \sim o_2$, then for any $p \in [0,1]$ and o_3, $[p : o_1, (1-p) : o_3] \sim [p : o_2, (1-p) : o_3]$.



- Axiom 6 Independence. If $o_1 \succeq o_2$, then for any number $p \in [0, 1]$ and outcome o_3:

$$[p : o_1, (1-p) : o_3] \succeq [p : o_2, (1-p) : o_3].$$

Note: By definition of $\sim$, if $o_1 \sim o_2$, then for any $p \in [0, 1]$ and o_3, $[p : o_1, (1-p) : o_3] \sim [p : o_2, (1-p) : o_3]$.



Definition

An agent is defined to be (VNM-)rational if it obeys the completeness, transitivity, monotonicity, decomposability, continuity, and independence axioms.

Theorem (von Neumann and Morgenstern 1947)

If an agent is rational, then for every outcome o_i, there is a real number $u(o_i)$ such that

- **(a)** $o_i \succ o_j$ if and only if $u(o_i) > u(o_j)$ and
- **(b)** utilities are linear with probabilities:

$$u([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) = p_1 u(o_1) + p_2 u(o_2) + \cdots + p_k u(o_k).$$

Theorem (von Neumann and Morgenstern 1947)

If an agent is rational, then for every outcome o_i, there is a real number $u(o_i)$ such that

- **(a)** $o_i \succ o_j$ if and only if $u(o_i) > u(o_j)$ and
- **(b)** utilities are linear with probabilities:

$$u([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) = p_1 u(o_1) + p_2 u(o_2) + \cdots + p_k u(o_k).$$

Proof (sketch). If the agent is indifferent between *all* outcomes, then define $u(o) = 0$ for all outcomes o. Otherwise,

- Choose the best outcome $\overline{o}$, and the worst outcome $\underline{o}$.
- For any outcome o, define $u(o)$ as the value p such that

$$o \sim [p : \overline{o}, (1-p) : \underline{o}] \qquad \text{(by continuity)}$$

Theorem (von Neumann and Morgenstern 1947)

If an agent is rational, then for every outcome o_i, there is a real number $u(o_i)$ such that

- **(a)** $o_i \succ o_j$ if and only if $u(o_i) > u(o_j)$ and
- **(b)** utilities are linear with probabilities:

$$u([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) = p_1 u(o_1) + p_2 u(o_2) + \cdots + p_k u(o_k).$$

Proof (sketch). If the agent is indifferent between *all* outcomes, then define $u(o) = 0$ for all outcomes o. Otherwise,

- Choose the best outcome $\overline{o}$, and the worst outcome $\underline{o}$.
- For any outcome o, define $u(o)$ as the value p such that

$$o \sim [p : \overline{o}, (1-p) : \underline{o}] \qquad \text{(by continuity)}$$

We prove that the function u satisfies the following:

- **(a).** $o_i \succ o_j$, if and only if
  $[u(o_i) : \overline{o}, (1-u(o_i)) : \underline{o}] \succ [u(o_j) : \overline{o}, (1-u(o_j)) : \underline{o}]$ (by independence),
  if and only if $u(o_i) > u(o_j)$ (by monotonicity).
- **(b).** By decomposability, any outcome $[p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]$ can be reduced to a lottery of the form $[p : \overline{o}, (1-p) : \underline{o}]$ where

$$p = p_1 u(o_1) + p_2 u(o_2) + \cdots + p_k u(o_k).$$

□

Consequences of (VNM) rationality.

- Although preferences may seem to be complex and multifaceted, a rational agent's value for an outcome can be measure by a (one-dimensional) number $u(o_i)$, i.e., the utility of the outcome o_i.
- The utility of a (probabilistic) outcome $[p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]$ can be described as the linear sum, i.e., the expected utility:

$$p_1 u(o_1) + p_2 u(o_2) + \cdots + p_k u(o_k)$$

- Linear scalability: Suppose $u : \Omega \to \mathbb{R}$ is a utility function. Then for any constant $c > 0$, $u' : \Omega \to \mathbb{R}$ defined by $u'(o) = cu(o)$ is also a utility function.
- Thus the utility functions of two agents cannot be added.

Expected utility hypothesis: The utility function correctly reflects the performance measure of an agent, i.e., if the agent acts so as to maximize the expected utility, then the agent will achieve the highest performance.

[1]

[1] von Neumann, Morgenstern, *Theory of Games and Economics Behavior*. 1953.

Example. Suppose you have \$1000 in savings. Would you risk \$1000 for a 10% chance odds to win \$9000?

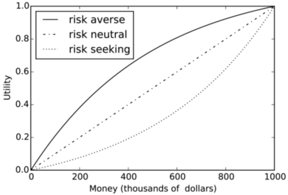
- Outcome s_0: Not participating, monetary value \$0
- Outcome s_{9000}: Participating, monetary value \$9000
- Outcome s_{-1000}: Participating, monetary -\$1000

Then the agent has lotteries:

- Not participating: $[1: s_0, 0: s_{9000}, 0: s_{-1000}]$
  Utility: $u(s_0)$
- Participating: $[0: s_0, , 0.1: s_{9000}, 0.9: s_{-1000}]$
  Utility: $u(0.1s_{9000} + 0.9s_{-1000}) = 0.1s(o_{9000}) + 0.9u(s_{-1000})$

Monetary Monotonicity Assumption: $s_{9000} \succ s_0 \succ s_{-1000}$.

- $u(s_0) > 0.1u(s_{9000}) + 0.9u(s_{-1000})$: the agent is risk averse.
- $u(s_0) = 0.1u(s_{9000}) + 0.9u(s_{-1000})$: the agent is risk neutral.
- $u(s_0) < 0.1u(s_{9000}) + 0.9u(s_{-1000})$: the agent is risk seeking.



Multi-dimensional Utility

Question.

- Atomic representation: A utility function is defined on each state $s \in S$.
- Factored representation: How to extend the definition of utility functions to factored representations (over variables)?

Example [warrior]. What items should the warrior choose?

- Weapon: Sword, Wooden club
- Armor: Chain vest, Cloth armor

Additive independence property: A state is described by k variables $V_1, \ldots, V_k$:

- a utility function $u_i\colon \mathrm{dom}(V_i) \to [0,1]$ is defined for any variable V_i, $1 \leq i \leq k$.
- For state $s = (V_1 = v_1, \ldots, V_k = v_k)$, $u(s) = w_1 u_1(v_1) + \cdots + w_k u_k(v_k)$ for weights $w_1, \ldots, w_k \in \mathbb{R}$.

E.g. $u(Weapon, Armor) = w_1 u_1(Weapon) + w_2 u_2(Armor)$.

Additive independence property usually does *not* hold:

Example. [delivery robot] Consider a delivery robot aiming to move to a target location.

- Variable 1 $Acc \in \{0,1\}$: The robot may have an accident.
- Variable 2 $Pad \in \{0,1\}$: The robot can choose to put on a pad, which avoids severe damage, but increases weight.
- Variable 3 $Way \in \{0,1\}$: The robot can choose to use either a short or a long route. The short one has higher chance of having an accident.

The variables satisfy the following:

- The utility depends on all three variables.
- The value of Acc depends on Way, i.e. $\mathbf{P}(Acc, Way) \neq \mathbf{P}(Acc)\mathbf{P}(Way)$.
- The effect of Acc on the utility is different for different values of Pad:

$$u(Acc=0, Pad=0, Way) - u(Acc=1, Pad=0, Way)$$
$$> u(Acc=0, Pad=1, Way) - u(Acc=1, Pad=1, Way)$$

One-off Decisions

Definition

The one-off decision problem is defined as follows:

- A set of decision variables $V_1, \ldots, V_k$
- A set of random variables $R_1, \ldots, R_\ell$
- State space $S = \prod_{i=1}^{k} \mathrm{dom}(V_i) \times \prod_{j=1}^{\ell} \mathrm{dom}(R_j)$.
- Conditional probability $\mathbf{P}(R_1, \ldots, R_\ell \mid V_1, \ldots, V_k)$.
- Preferences over S which satisfy the rationality axioms.

The goal is to choose values for $(V_1, \ldots, V_k)$.

One-off Decisions

Definition

The one-off decision problem is defined as follows:

- A set of decision variables $V_1, \dots, V_k$
- A set of random variables $R_1, \dots, R_\ell$
- State space $S = \prod_{i=1}^{k} \mathrm{dom}(V_i) \times \prod_{j=1}^{\ell} \mathrm{dom}(R_j)$.
- Conditional probability $\mathbf{P}(R_1, \dots, R_\ell \mid V_1, \dots, V_k)$.
- Preferences over S which satisfy the rationality axioms.

The goal is to choose values for $(V_1, \dots, V_k)$.

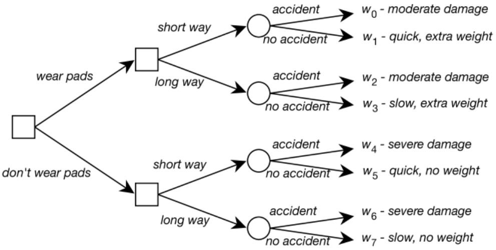
Solution.

- By VNM theorem, there is a utility function $u\colon \Omega \to \mathbb{R}$.
- By the expected utility hypothesis: The agent wants to maximise the expected utility $\arg\max_{(v_1,\dots,v_k)\in\prod_{i=1}^{k}\mathrm{dom}(i)} E(u \mid v_1, \dots, v_k)$ where

$$E(u \mid v_1, \dots, v_k) = \sum_{(r_1,\dots,r_\ell)\in\prod_{j=1}^{\ell}\mathrm{dom}(R_j)} P(r_1, \dots, r_\ell \mid v_1, \dots, v_k)\, u(v_1, \dots, v_k, r_1, \dots, r_\ell)$$

Example. [delivery robot]

- Decision variable 1 $Pad \in \{0, 1\}$
- Decision variable 2 $Way \in \{0, 1\}$
- Effect variable $Acc \in \{0, 1\}$
- States: $S = (Pad, Way, Acc) \in \{0, 1\}^3$
- Conditional probability: $\mathbf{P}(Acc \mid Pad, Way)$.
- Preference: $\overline{o} = w_5, \underline{o} = w_6, w_i \sim [p_i : w_5, (1 - p_i) : w_6]$ for $0 \leq i \leq 7$
- Decision tree:



Task. Decide on *Pad* and *Way* to maximise the expected utility.

Example. [delivery robot]

Conditional probability:

Way	Acc	$\mathbf{P}(Acc \mid Way)$
0	0	0.8
0	1	0.2
1	0	0.99
1	1	0.01

Utility function: $u\colon S \to \mathbb{R}$

Pad	Way	Acc	Outcome	$u(Pad, Way, Acc)$
0	0	0	w_5	100
0	0	1	w_4	3
0	1	0	w_7	80
0	1	1	w_6	0
1	0	0	w_1	95
1	0	1	w_0	35
1	1	0	w_3	75
1	1	1	w_2	30

Expected utility: $E(u \mid Pad, Way)$

Pad	Way	$E(u \mid Pad, Way)$
0	0	$0.2 \times 3 + 0.8 \times 100 = 80.6$
0	1	$0.01 \times 0 + 0.99 \times 80 = 79.2$
1	0	$0.2 \times 35 + 0.8 \times 95 = 83$
1	1	$0.01 \times 30 + 0.99 \times 75 = 74.55$

Summary of The Topic

The following are the main knowledge points covered:

- Planning with uncertainty v.s. Classical planning
- **Utility theory:** Lottery and outcomes, Preferences relations $\succeq$, $\sim$, $\succ$
- **Axioms of rationality:**
  1. Completeness
  2. Transitivity
  3. Monotonicity
  4. Decomposability
  5. Continuity
  6. Independence
- **VNM rationality theorem:** The existence and linearity of utility function, expected utility hypothesis
- **Multi-dimensional utility:** Additive independence property
- **One-off decision problem:**
  - Decision variables
  - random variables
  - Conditional probability
  - Utility/preferences
  - Decision tree