Artificial Intelligence

Simple Decisions (AIMA 16)

Christopher Simpkins

Kennesaw State University

Making Simple Decisions

A decision-theoretic agent—an agent can make rational decisions based on what it believes and what it wants.

- A goal-based agent has a binary distinction between good (goal) and bad (non-goal) states.
- A decision-theoretic agent assigns a continuous range of values to states, enabling decision-making even when no best option is available.

This lesson deals with simple decisions, that is, decisions based on immediate outcomes in episodic environments.

Combining Beliefs and Desires under Uncertainty

- An agent has uncertainty about the current state, so each state has an associated probability $Pr(s)$.
- Action outcomes are also uncertain, so the transition model is $Pr(s' \mid s, a)$.

$$Pr(RESULT(a) = s') = \sum_{s} Pr(s)Pr(s' \mid s, a)$$

Decision Theory

Expected utility:

$$EU(a) = \sum_{s'} Pr(RESULT(a) = s')U(s') \tag{16.1}$$

Principle of **maximum expected utility** (MEU):

$$action = \operatorname*{argmax}_a EU(a)$$

A few points:

- The MEU principle *formalizes* rational decisions but does not *operationalize* them.
- If an agent acts so as to maximize a utility function *that correctly reflects the performance measure*, then the agent will achieve the highest possible performance score (averaged over all the possible environments).

Basis of Utility Theory

Notation:

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

Lottery L with outcomes $S_1, \ldots, S_n$ that occur with probabilities $p_1, \ldots, p_n$:

$$L = [p_1, S_1; p_2, S_2; \ldots p_n, S_n]$$

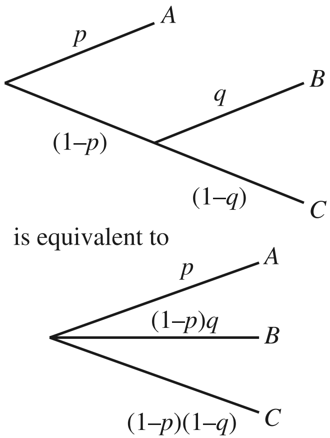
Axioms of Utility Theory (1/2)

Six constraints that we require any reasonable preference relation to obey:

- **Orderability**: Given any two lotteries, a rational agent must either prefer one or else rate them as equally preferable. That is, the agent cannot avoid deciding. Refusing to bet is like refusing to allow time to pass.
  - Exactly one of $(A \succ B)$, $(B \succ A)$, or $(A \sim B)$ holds.
- **Transitivity**: Given any three lotteries, if an agent prefers A to B and prefers B to C, then the agent must prefer A to C.
  - $(A \succ B) \land (B \succ C) \implies (A \succ C)$.
- **Continuity**: If some lottery B is between A and C in preference, then there is some probability p for which the rational agent will be indifferent between getting B for sure and the lottery that yields A with probability p and C with probability $1 - p$.
  - $A \succ B \succ C \implies \exists p[p, A; 1-p, C] \sim B$.
- **Substitutability**: If an agent is indifferent between two lotteries A and B, then the agent is indifferent between two more complex lotteries that are the same except that B is substituted for A in one of them. This holds regardless of the probabilities and the other outcome(s) in the lotteries.
  - $A \sim B \implies [p, A; 1-p, C] \sim [p, B; 1-p, C]$.
  - This also holds if we substitute $\succ$ for $\sim$ in this axiom.

Axioms of Utility Theory (2/2)

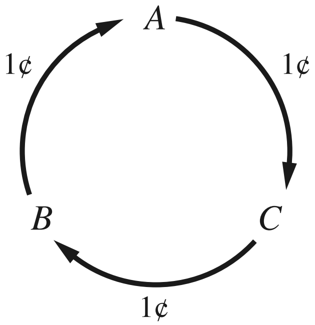
- **Monotonicity**: Suppose two lotteries have the same two possible outcomes, A and B. If an agent prefers A to B, then the agent must prefer the lottery that has a higher probability for A (and vice versa).
  - $A \succ B \implies (p > q \iff [p, A; 1-p, B] \succ [q, A; 1-q, B])$.
- **Decomposability**: Compound lotteries can be reduced to simpler ones using the laws of probability. This has been called the "no fun in gambling" rule: as Figure 15.1(b) shows, it compresses two consecutive lotteries into a single equivalent lottery.
  - $[p, A; 1-p, [q, B; 1-q, C]] \sim [p, A; (1-p)q, B; (1-p)(1-q), C]$.



Nontransitive Preferences

Suppose an agent held the following preferences for freely exchangeable goods: $A \succ B \succ C \succ A$.

This agent would be willing, e.g., to exchange B plus \$.01 for A, and so on. But the nontransitivity of the agent's preferences could lead to a cycle that ends in complete financial depletion:



The axioms of utility theory are rational because violating them leads to bad outcomes.

From Rational Preferences to Untilities

- **Existence of Utility Function**: If an agent's preferences obey the axioms of utility, then there exists a function U such that $U(A) > U(B)$ if and only if A is preferred to B, and $U(A) = U(B)$ if and only if the agent is indifferent between A and B. That is,
  - $U(A) > U(B) \iff A \succ B$ and $U(A) = U(B) \iff A \sim B$.
- **Expected Utility of a Lottery**: The utility of a lottery is the sum of the probability of each outcome times the utility of that outcome.
  - $U([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i U(S_i)$.

Utility functions create relative scales, not absolute scales. For example, if we apply a positive affine transformation:

$$U'(S) = aU(S) + b \tag{16.2}$$

Then U' and U are effectively equivalent because they lead to the same decisions.

So $U(S)$ is a **value function** or **ordinal utility function**, in which an agent needs only a preference ranking on states – the numbers don't matter.

Utility Functions

A utility functions

- map from lotteries to real numbers, and
- obey the axioms of utility theory.

Otherwise, they are arbitrary.

- I might prefer pepperoni pizza to pineapple, another might prefer the reverse.
- Decisions based on either preference ordering, as long as it follows the two properties above, are rational.

To build a decision support system for humans we must try to infer the human's utility function, a process called **preference elicitation**.

Utility Scales and Preference Elicitation

There is no absolute scale for utilities, but we can establish some scale. Let

- $U(S) = u_\top$ be the best possible prize,
- $U(S) = u_\perp$ be the worst possible catastrophe, and
- use a **normalized utility scale** in which $u_\perp = 0$ and $u_\top = 1$.

Then preference elicitation can proceed by

- asking the agent to choose between a particular prize S and a **standard lottery** $[p, u_\top; (1-p), u_{bot}]$, and
- adjusting the probability p until the agent is indifferent between S and the standard lottery.

Assuming normalized probabilities, the utility of S is then given by p. We repeat this process for every S to get the full utility function.

Example:

The "Value" of Human Life

- Asbestos in schools paradox.
- US govt agencies like FDA and EPA use the **value of a statistical life** to determine the costs and benefits of regulations. In 2019 this was $10M.
- A **micromort** is a one in a million chance of death.
  - Say driving a car for 230 miles incurs a risk of one micromort.
  - If you drive 92,000 miles that's 400 micromorts.
  - If you're willing to pay $12,000 more for a car that halves your risk of death, then a micromort has a value of $60 to you.
- **QALY** is quality-adjusted life year. Patients are willing to accept a shorter life expectancy to avoid a disability.

A new car built by my company leaves somewhere traveling at 60 mph.

The rear differential locks up.

The car crashes and burns with everyone trapped inside.

Now, should we initiate a recall?

Take the number of vehicles in the field, A, multiply by the probable rate of failure, B, multiply by the average out-of-court settlement, C.

A times B times C equals X.

If X is less than the cost of a recall, we don't do one.

The Utility of Money

Money is an obvious candidate for a utility measure.

- Most agents exhibit a **monotonic preference** for more money.
- However, money is not necessarily a utility function because it says nothing about preferences between *lotteries* involving money.
  - Example: Choose between \$1,000,000 and a 50% chance for \$2,500,000.
  - The **expected monetary value** (EMV) of the gamble is $.5(\$0) + .5(\$2,500,000) = \$1,250,000$

Key idea: value of money not directly proportional to utility.

- Let k be current wealth and S_n be the state of posessing total wealth of n dollars.

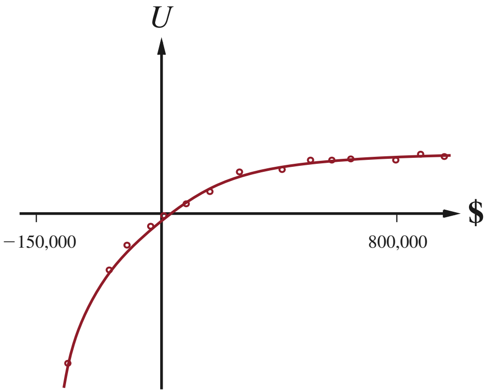
$$EU(Accept) = \frac{1}{2}U(S_k) + \frac{1}{2}U(S_{k+2,500,000})$$
$$eu(Decline) = U(S_{k+1,000,000})$$

Now say $U((S_k) = 5$, $U(S_{k+2,500,000}) =$ 9\$, and $U(S_{k+1,000,000} = 8$. Then

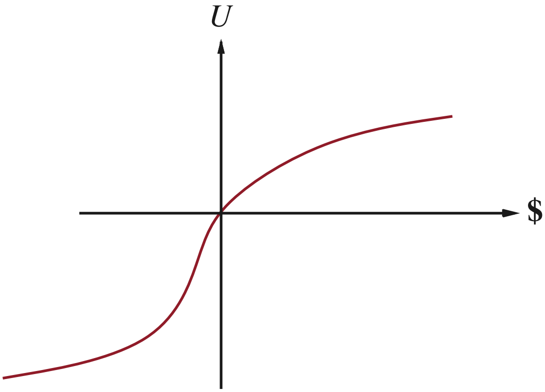
- $EU(Accept) = 7$
- $EU(Decline) = 8$

"Utility of your first million is higher than your second." A billionaire would likely have a locally linear utility function over such small amounts, so would accept the gamble.

Utility of Money Example

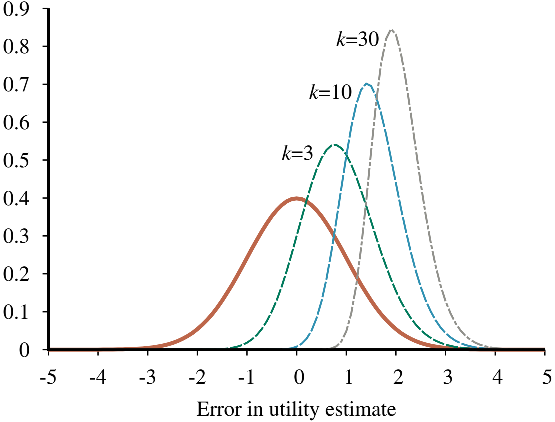


Utility of Money in General



The Optimzer's Curse

The greater the number of choices, the more biased your estimates are because you always pick "optimal" action.



Human Judgement and Irrationality

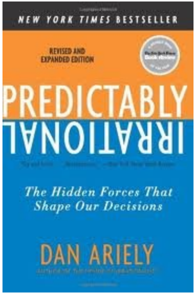
- Decision theory is a **normative theory** – how people *should* act.
- Actualy behavior is described by a **descriptive theory**.

Humans aren't "rational." (Ariely, 2009)

Allais Paradox: people don't act in accordance with utility theory.

Common errors in thinking:

- Ambiguity aversion
- Framing effect
- Anchoring effect




[1]

[1] https://web.mit.edu/ariely/www/

Multiattribute Utility Functions

Let the attributes be $\boldsymbol{X} = X_1, \ldots, X_n$ and the vector $\boldsymbol{x} = \langle x_1, \ldots, x_n \rangle$ be a complete set of variable values.

- Each x_i, whether continuous or discrete, must be from a totally ordered set.
- Not required, but better if higher values mean higher utilities, i.e., each x_i monotonically increases.
  - $-d$ deaths better than d, assuming you prefer fewer deaths.
  - Measures with a peak that falls off in either direction could be split. E.g., instead of temperature t with peak at 70, have:
    - Warmth, $w = \min(0, t - 70)$ with max value of 0.
    - Coolness, $c = \min(0, 70 - t)$ with max value of 0.

Two ways to deal with multiattribute utility funcitons:

- Keep attributes separate.
- Combine attributes into a single utility value.

Airport Siting Problem

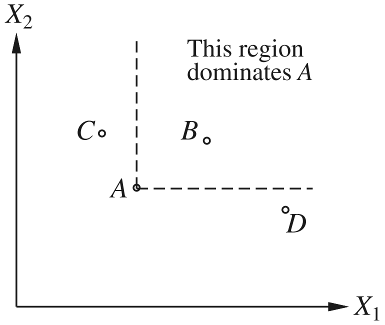
Attributes of airport siting problem (where to locate an aiport):

- X_1: Throughput, measured by the number of flights per day;
- X_2: Safety, measured by minus the expected number of deaths per year;
- X_3: Quietness, measured by minus the number of people living under the flight paths;
- X_4: Frugality, measured by the negative cost of construction.

Strict Dominance

Given two airport sites $S_1 = \langle x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_4^{(1)} \rangle$ and $S_2 = \langle x_1^{(2)}, x_2^{(2)}, x_3^{(2)}, x_4^{(2)} \rangle$:

- If $x_i^2 < x_i^1, \forall i$, then S_1 **strictly dominates** S_2.
- Helpful for eliminating choices, rarely sufficient to find a unique best choice.



Preference Structure

Specifying complete utility function $U(x_1, \ldots, x_n)$ requires d^n values in the worst case. Avoid this complexity by encoding some structure of preferences into **representation theorems**:

$$U(x_1, \ldots, x_n) = F\left[f_1(x_1), \ldots, f_n(x_n)\right]$$

Where

- F is a simple function (like addition), and
- each f_i converts utility attributes into a common measure.

Example: Each x_i is an amount of money in an abitrary currency like Euros or Rupees, and each f_i converts the amount into USD.

Deterministic Preferences

Most common regularity in deterministic preferences is **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 for attribute X_3.

Airport siting example: let $X_1 = Quietness$, $X_2 = Frugality$, $X_1 = Safety$, with

- $x_1 =$ 70,000 people living under flight path,
- $X_2 =$ \$3.7 Billion constuction cost, and
- $x_3 =$ 0.006 deaths per billion passenger miles.

If we prefer $\langle 20,000, 4B, 0.006 \rangle$, then this preference would hold for any x_3.

Mutial Preference Independence

Mutual prefence independence (MPI) means that preference for each attribute has no effect on preference for other attributes. When this holds, the agent's preferences can be represnented by a value function of the form:

$$V(x_1,,\ldots,x_n) = \sum_i V_i(x_i)$$

For example, we could assume MPI for the airport siting problem and use the value funciton:

$$V(quietness, frugality, safety) = quietness \times 10^4 + frugality + safety \times 10^{12}$$

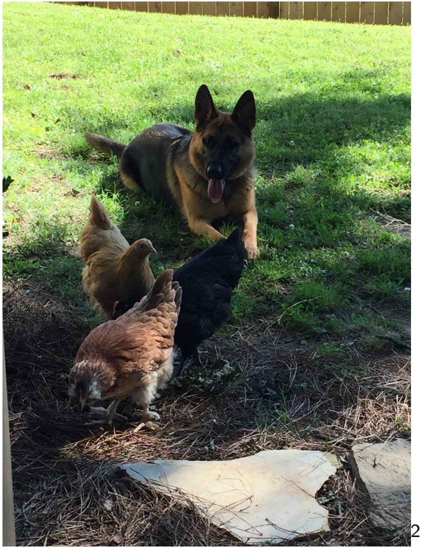
This is an example of an **additive value function**. Even when MPI doesn't hold, can be a useful simplifying assumption.

Non-MPI Example

You're purchasing some hunting dogs, some chickens, and some cages for the chickens.

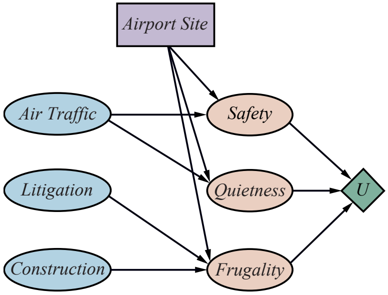
- The hunting dogs are very valuable, but will eat uncaged chickens.
- So you need enough cages for the chickens.

Hence, the tradeoff between dogs and chickens depends strongly on the number of cages, and MPI is violated.


[2]

[2] https://www.reddit.com/r/germanshepherds/comments/p4sou1/our_girl_reyna_watches_her_flock_8_of_backyard/

Syntax of Decision Networks



- Chance nodes (ovals) represent random variables, like in Bayesian networks.
  - Agent could be uncertain about the construction cost, level of air traffic, potential for litigation, and Safety, Quietness, Frugality variables, each of which also depends on the site chosen.
  - Each chance node has a conditional distribution indexed by state of its parent nodes.
  - Parent nodes can include decision nodes and chance nodes.
  - Each current-state chance node could be part of a large Bayesian network for assessing construction costs, air traffic levels, or litigation potentials.
- Decision nodes (rectangles) represent action choice points.
  - $AirportSite$ action can take a different value for each candidate site.
  - The choice influences the safety, quietness, and frugality of the solution.
  - For simple decisions, we have a single decision node.
- Utility nodes, .a.k.a. value nodes, (diamonds) represent the agent's utility function.
  - Utility node's parents are all variables describing outcomes directly affecting utility.
  - Utility node has description of the agent's utility as a function of parent attributes.
  - Description could be a tabulation of the function, or a parameterized additive or linear function of the attribute values.
  - For now, assume function is deterministic – given values of its parent variables, value of utility node is fully determined.

Semantics of Decision Networks

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.
   - (b) Calculate the posterior probabilities for the parent nodes of the utility node, using a standard probabilistic inference algorithm.
   - (c) Calculate the resulting utility for the action.
3. Return the action with the highest utility.