Week 2: Decision_Theory

1. Background

1.1. Preference

preference: an agent chooses among prizes (A,B, etc.) and lotteries (for uncertain prizes)

1.1.1.Notation

$$A \succ B \Leftrightarrow A \quad preferred \quad B$$

$$A \succcurlyeq B \Leftrightarrow B \quad not \quad preferred \quad to \quad A$$

$$A \sim B \Leftrightarrow indifference \quad between \quad A \quad and \quad B$$

1.1.2. Rational preferences

behavior describable as **maximization of expected utility** (MEU principles).

However, an agent can be entirely rational (consistent with MEU) without ever representing or manipulating utilities and probabilities (E.g., a lookup table for perfect tictactoe)

1.1.3. Constraints

- Orderability: $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
- Transitivity: $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
- Continuity: $A \succ B \succ C \Rightarrow \exists p\ [p, A; 1\text{-}p, C] \sim B$
- Substitutability: $A \sim B \Rightarrow [p, A; 1\text{-}p, C] \sim [p, B; 1\text{-}p, C]$
- Monotonicity: $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; 1\text{-}p, B] \succsim [q, A; 1\text{-}q, B])$

violating constraints $\Rightarrow$ **self-evident irrationality**

1.2. Lottery:

use for decision under uncertainty:

A lottery L with possible outcomes $S_1, ..., S_n$ that occur with probabilities $p_1, ..., p_n$ is written

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

each outcome S_i of a lottery can be either an atomic state or another lottery

1.2.1. Preference and Utility

1. Existence of Utility Function:

$$\exists U(A) > U(B) \Leftrightarrow A \succ B$$

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

2. Expected Utility of a Lottery

$$U([p1, S1; ...; pn, Sn]) = \sum_i p_i U(Si)$$

Understanding:

This part means, we can **always** transfter **preference** to a **mathematical value,** and the transfer may not unique

1.2.2. Preference Elicitation

1. How to compare a given state A to a standard lottery L_p

    "best possible prize" $u_\top$ with probability p
    "worst possible catastrophe" $u_\perp$ with probability $(1-p)$
    adjust lottery probability p until $A \sim L_p$

2. With deterministic prizes only (no lottery choices), only ordinal utility can be determined

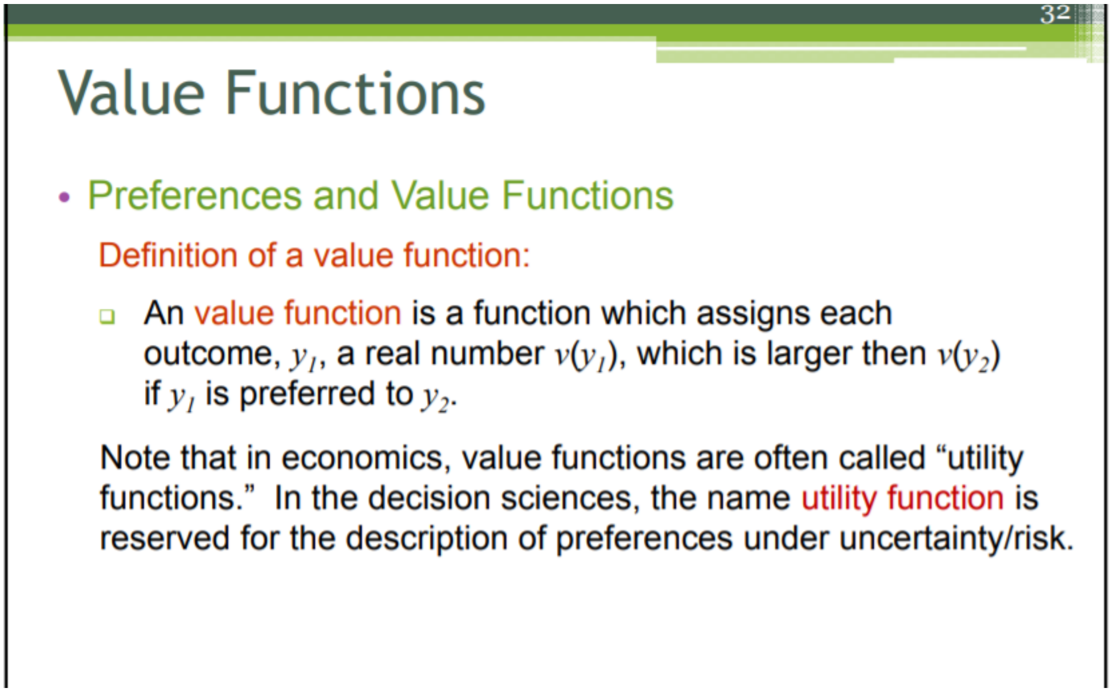
$$U'(x) = k_1 U(x) + k_2 \quad where \quad k_1 > 0$$

3. Normalized utilities: $U_{max} = 1.0; \quad U_{min} = 0.0$
4. Given a utility scale between u⊤ and u⊥, we can **assess the utility of any particular prize S** by asking the agent to choose between S and a standard lottery [p, u⊤; (1−p), u⊥]
5. Once this is done for each prize, the utilities for all lotteries involving those prizes are determined

Understanding

This part propose a way to map lotteries to utility and how to compare them

1.3. Value function and Utility function



2. Multi-attribute Utility Theory

2.1. Notation

1. **multi-attribute**: $[X_1,\ X_2,\ ...X_n]$⇔ dimensions to compare, in this chpater, this is same to the outcome of an anction a
2. **Action** $a\ \in\ Actions$ ↔ value(for determinisitc)/lottery(for uncertain situation)
3. A, B in lottery may be a vector like $[x_1,\ x_2,\ ...,\ x_n]$, now the x_i means the different outcomes after an action a.
4. $V(x_1, x_2, ...x_n) or U(x_1, x_2, ...x_n)$

2.2. Dominance (How to compare)

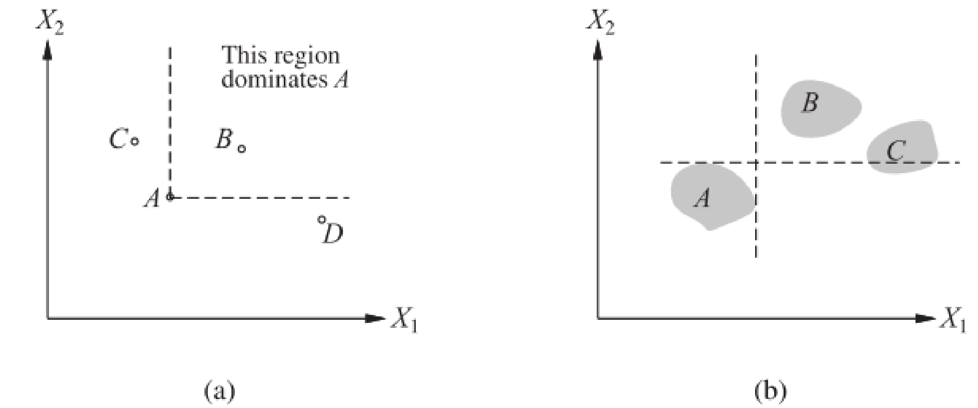


Figure 16.4 Strict dominance. (a) Deterministic: Option A is strictly dominated by B but not by C or D. (b) Uncertain: A is strictly dominated by B but not by C.

2.2.1. Strict dominance (monotonic)

like 16(A)

2.2.2.Uncertain

like 16(B)

2.2.3. Stochastically dominates

1. two distributions p_1, p_2 are stochastically dominates $\iff \forall x \int_{-\infty}^{x} p_1(t)dt \leq \int_{-\infty}^{x} p_2(t)d(t)$
2. If A_1 stochastically dominates A_2, and U(x) is monotonically nondecreasing utility function in x $\longrightarrow$ $\int_{-\infty}^{+\infty} p_1(x)u(x)dx \geq \int_{-\infty}^{+\infty} p_2(x)u(x)dx$

2.2. Preference structure

2.2.1. Determinisitic Situation(Preferentially independent):

Preferentially Independent

An attribute is **preferentially independent** from all other attributes when changes in the rank ordering of preferences of other attributes does not change the preference order of the attribute.

that is:

$$\forall x_3 \in X_3 \quad if \quad (x_1, x_2, x_3) \succ (x_1', x_2', x_3) \Leftrightarrow X1 \quad and \quad X2 \quad P.I. \quad X3$$

Mutual Preferentially Independent

Attributes A are **mutually perferentially independent**, if any subset of attributes $X \subseteq A$ is preferentially independent of the other attributes $Y = A \setminus X$. That is, for any $X \subseteq A, Y = A \setminus X$:

$$(x, y') \geqslant (x', y') \Rightarrow (x, y) \geqslant (x', y) \text{ for all } y \in Y$$

- if every pair of attributes is P.I. of its complement, then every subset of attributes is P.I of its complement
-

Property:

$$V(x_1, x_2, ..., x_n) = \sum_i V_i(x_i)$$

if you keep x_j constant, and change others it is easy to understand inuitively

2.2.2. Uncertain situation

Utility Independence

- It is similar to preferential independence, except that the assessments are made with uncertainty present. It is **a stronger assumption**

Attribute 1 is ***utility-independent*** of attribute 2, if the conditional preferences on lotteries on attribute 1 given a constant value of attribute 2, do not depend on that constant value.

that is

$$\forall x_2 \in X_2 \quad if \quad < (x_1, x_2), (x_1', x_2) > \succ < (y_1, x_2), (y_1', x_2) > \Leftrightarrow X1 \quad U.I.X2$$

Mutual Utility Independence

each subset is utility independent of its complement

Property

if an multi-attributes situation is M.U.I, then we can find a **multiplicative utility function**

for example, for 3 attibutes

$$\begin{aligned} U = & k_1U_1 + k_2U_2 + k_3U_3 + k_1k_2U_1U_2 + k_2k_3U_2U_3 \\ & + k_3k_1U_3U_1 + k_1k_2k_3U_1U_2U_3 \end{aligned}$$

3. Value of Information

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

3.1. Value of perfect information (VPI)

To evaluate the information, we computed expected value of information:

expected value of best action **given the additional information -** expected value of best action **without the additional information**

α represents the best choice

$$EU(\alpha|E) = \max_{a \in A} \sum_i U(S_i)P(S_i|E, a)$$

$$EU(\alpha|E, e) = \max_{a \in A} \sum_i U(S_i)P(S_i|E, a, e)$$

$$VPI_E(E_j) = (\sum_k P(E_j = e_{jk}|E)EU(\alpha_{e_{jk}}|E, E_j = e_{jk})) - EU(\alpha|E)$$

3.2. Cases for VPI

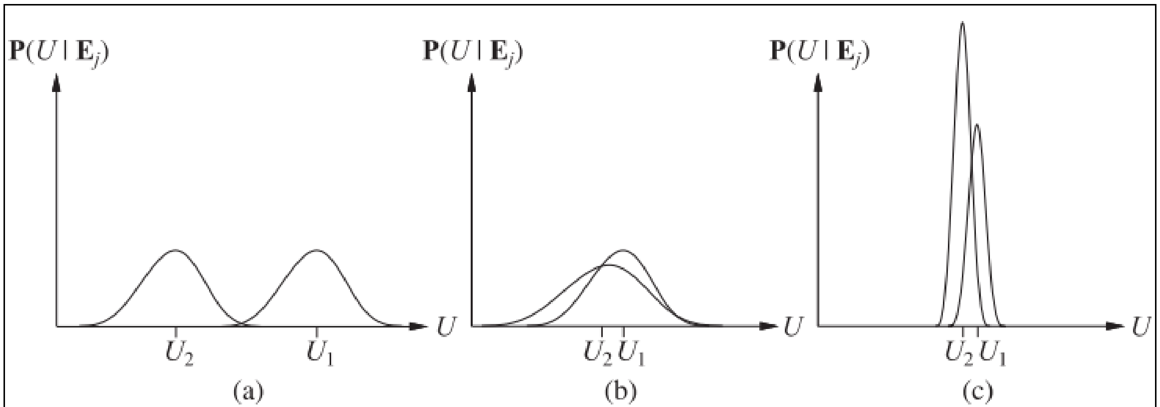


Figure 16.8 Three generic cases for the value of information. In (a), a_1 will almost certainly remain superior to a_2, so the information is not needed. In (b), the choice is unclear and the information is crucial. In (c), the choice is unclear, but because it makes little difference, the information is less valuable. (Note: The fact that U_2 has a high peak in (c) means that its expected value is known with higher certainty than U_1.)

1. Understanding of C: One thing may happen with certainty, for example route2 is very likely to be blocked. Event that time, the difference between values are small, but we will not be bothered when make a decision.

3.3. Properties of VPI

1. **Nonnegative**

   $$\forall e, E_j \quad VPI_e(E_j) \geq 0$$

   innuitively, because VPI is a theorem about expected value, not actual value, one could in the worst case just ignore the information and pretend that one has never received it.

2. **non-additive**

   $$VPI_e(E_j, E_k) \neq VPI_e(E_j) + VPI_e(E_k)$$

innuitively, new information will change the value of initial information

3. **order-dependent**

$$VPI_e(E_j, E_k) = VPI_e(E_j) + VPI_{e,e_j}(E_k) = VPI_e(E_k) + VPI_{e,e_k}(E_j)$$

Note:

when more than one piece of evidence can be gathered, maximizing VPI for each to select one **is not always optimal:**

At that time: evidence-gathering becomes a **sequential decision problem**

3.4. Implementation of an Information-gathering agent

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

  integrate *percept* into D
  $j \leftarrow$ the value that maximizes $VPI(E_j) \,/\, Cost(E_j)$
  if $VPI(E_j) > Cost(E_j)$
    return REQUEST(E_j)
  else return the best action from D

Figure 16.9 Design of a simple information-gathering agent. The agent works by repeatedly selecting the observation with the highest information value, until the cost of the next observation is greater than its expected benefit.

4. Decision Network

4.1. Resenting a decision problem with a decision network

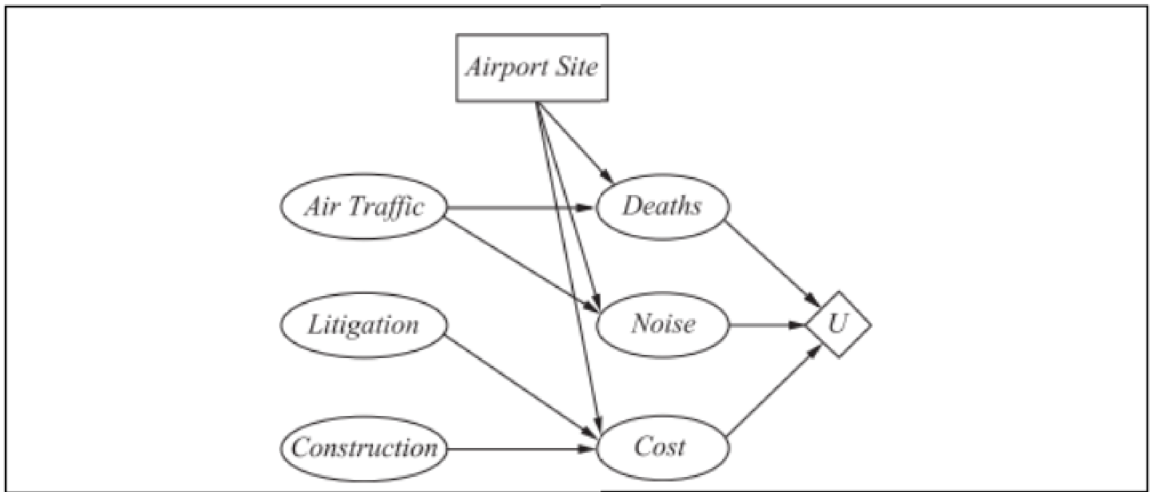


Figure 16.6 A simple decision network for the airport-siting problem.

Chance Node:

- represent **random variables**, just as they do in Bayesian networks
- In decision networks, the parent nodes can include decision nodes as well as chance nodes

Decision Nodes:

- represent points where the decision maker has a choice of actions, like "Airport Site" node

Utility Nodes:

- The utility node has as parents all variables describing the outcome that directly affect utility
- The description could be just a tabulation of the function, or it might be a parameterized additive or linear function of the attribute values.

4.2. Evaluating decision networks

- Once the decision node is set, it behaves exactly like a chance node that has been set as an evidence variable

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.

5. Example of real-life Utility and Decision

5.1. Micromorts

Introduction

A micromort is a **unit** of risk measuring a **one-in-a-million probability of death**

Application

An application of micromorts is measuring **the value that humans place on risk**:

- What is the amount of **money** one would have to pay a person to get him or her to **accept** a one-in-a-million chance of death
- What **amount** is someone willing to pay to **avoid** a one-in-amillion chance of death

Examples

5.2. Quality Adjusted Life Years (QALYs)

Introduction

The quality-adjusted life year (QALY) is a **measure of disease burden**, including both the **quality and the quantity** of life lived

Application

It is used in assessing the value for money of a **medical intervention**

Properties

- The QALY model requires utility independent, risk neutral, and constant proportional tradeoff behaviour
- based on the number of years of life that would be added by the intervention

Debate

- Perfect health is hard, if not impossible, to define
- There are health states worse than death, therefore there should be **negative values possible** on the health spectrum
- Measures place disproportionate importance on **physical pain** or disability over **mental health**
- The **effects** of a patient's health **on the quality of life of others** (e.g. caregivers or family) do not figure into these calculations