RATIONAL DECISIONS

CHAPTER 16
Outline

◇ Rational preferences
◇ Utilities
◇ Money
◇ Multiattribute utilities
◇ Decision networks
◇ Value of information
Preferences

An agent chooses among prizes \((A, B, \text{etc.})\) and lotteries, i.e., situations with uncertain prizes

Lottery \(L = [p, A; (1-p), B]\)

Notation:
- \(A \succ B\) \(A\) preferred to \(B\)
- \(A \sim B\) indifference between \(A\) and \(B\)
- \(A \simeq B\) \(B\) not preferred to \(A\)
Rational preferences

Idea: preferences of a rational agent must obey constraints. Rational preferences \( \Rightarrow \)
behavior describable as maximization of expected utility

Constraints:

- **Orderability**
  \[ (A \succ B) \lor (B \succ A) \lor (A \sim B) \]

- **Transitivity**
  \[ (A \succ B) \land (B \succ C) \Rightarrow (A \succ C) \]

- **Continuity**
  \[ A \succ B \succ C \Rightarrow \exists p \ [p, A; 1 - p, C] \sim B \]

- **Substitutability**
  \[ A \sim B \Rightarrow [p, A; 1 - p, C] \sim [p, B; 1 - p, C] \]

- **Monotonicity**
  \[ A \succ B \Rightarrow (p \geq q \iff [p, A; 1 - p, B] \succeq [q, A; 1 - q, B]) \]
Rational preferences contd.

Violating the constraints leads to self-evident irrationality

For example: an agent with intransitive preferences can be induced to give away all its money

If $B \succ C$, then an agent who has $C$ would pay (say) 1 cent to get $B$

If $A \succ B$, then an agent who has $B$ would pay (say) 1 cent to get $A$

If $C \succ A$, then an agent who has $A$ would pay (say) 1 cent to get $C$
Maximizing expected utility

**Theorem** (Ramsey, 1931; von Neumann and Morgenstern, 1944): Given preferences satisfying the constraints there exists a real-valued function $U$ such that

\[
U(A) \geq U(B) \iff A \simeq B
\]

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

**MEU principle:**
Choose the action that maximizes expected utility

Note: 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
Utilities

Utilities map states to real numbers. Which numbers?

Standard approach to assessment of human utilities:
- compare a given state $A$ to a standard lottery $L_p$ that has
  - “best possible prize” $u_T$ with probability $p$
  - “worst possible catastrophe” $u_\perp$ with probability $(1 - p)$
- adjust lottery probability $p$ until $A \sim L_p$

pay $30$ $\sim$

```
0.999999
L
0.000001
```

continue as before

instant death
Utility scales

Normalized utilities: $u_T = 1.0, u_\perp = 0.0$

Micromorts: one-millionth chance of death
   useful for Russian roulette, paying to reduce product risks, etc.

QALYs: quality-adjusted life years
   useful for medical decisions involving substantial risk

Note: behavior is invariant w.r.t. +ve linear transformation

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

With deterministic prizes only (no lottery choices), only ordinal utility can be determined, i.e., total order on prizes
Money

Money does not behave as a utility function

Given a lottery \( L \) with expected monetary value \( EMV(L) \), usually \( U(L) < U(EMV(L)) \), i.e., people are risk-averse

Utility curve: for what probability \( p \) am I indifferent between a prize \( x \) and a lottery \([p, M; (1 - p), 0]\) for large \( M \)?

Typical empirical data, extrapolated with risk-prone behavior:
For each $x$, adjust $p$ until half the class votes for lottery ($M=10,000$)
Decision networks

Add action nodes and utility nodes to belief networks to enable rational decision making.

Algorithm:
For each value of action node
    compute expected value of utility node given action, evidence
Return MEU action
How can we handle utility functions of many variables $X_1 \ldots X_n$? E.g., what is $U(Deaths, Noise, Cost)$?

How can complex utility functions be assessed from preference behaviour?

Idea 1: identify conditions under which decisions can be made without complete identification of $U(x_1, \ldots, x_n)$

Idea 2: identify various types of independence in preferences and derive consequent canonical forms for $U(x_1, \ldots, x_n)$
Strict dominance

Typically define attributes such that $U$ is \textbf{monotonic} in each

\textbf{Strict dominance}: choice $B$ strictly dominates choice $A$ iff

$\forall i \; X_i(B) \geq X_i(A)$ \quad (and \ hence $U(B) \geq U(A)$)

Strict dominance seldom holds in practice
Stochastic dominance

Distribution \( p_1 \) stochastically dominates distribution \( p_2 \) iff
\[
\forall t \quad \int_{-\infty}^{t} p_1(x) \, dx \leq \int_{-\infty}^{t} p_2(t) \, dt
\]

If \( U \) is monotonic in \( x \), then \( A_1 \) with outcome distribution \( p_1 \) stochastically dominates \( A_2 \) with outcome distribution \( p_2 \):
\[
\int_{-\infty}^{\infty} p_1(x) U(x) \, dx \geq \int_{-\infty}^{\infty} p_2(x) U(x) \, dx
\]

Multiattribute case: stochastic dominance on all attributes \( \Rightarrow \) optimal
Stochastic dominance contd.

Stochastic dominance can often be determined without exact distributions using qualitative reasoning.

E.g., construction cost increases with distance from city.

\[ S_1 \text{ is closer to the city than } S_2 \]
\[ \Rightarrow S_1 \text{ stochastically dominates } S_2 \text{ on cost} \]

E.g., injury increases with collision speed.

Can annotate belief networks with stochastic dominance information:

\[ X \rightarrow Y \text{ (} X \text{ positively influences } Y \text{)} \text{ means that} \]

For every value \( z \) of \( Y \)'s other parents \( Z \):

\[ \forall x_1, x_2 \ x_1 \geq x_2 \Rightarrow P(Y|x_1, z) \text{ stochastically dominates } P(Y|x_2, z) \]
Label the arcs + or −
Label the arcs + or −
Label the arcs + or −
Label the arcs + or −
**Preference structure: Deterministic**

$X_1$ and $X_2$ preferentially independent of $X_3$ iff
preference between \( \langle x_1, x_2, x_3 \rangle \) and \( \langle x'_1, x'_2, x_3 \rangle \)
does not depend on $x_3$

E.g., \( \langle Noise, Cost, Safety \rangle \):
\( \langle 20,000 \text{ suffer, } $4.6 \text{ billion, } 0.06 \text{ deaths/mpm} \rangle \) vs.
\( \langle 70,000 \text{ suffer, } $4.2 \text{ billion, } 0.06 \text{ deaths/mpm} \rangle \)

**Theorem** (Leontief, 1947): if every pair of attributes is P.I. of its complement, then every subset of attributes is P.I of its complement: mutual P.I..

**Theorem** (Debreu, 1960): mutual P.I. \( \Rightarrow \exists \text{ additive value function:} \)

\[
V(S) = \Sigma_i V_i(X_i(S))
\]

Hence assess $n$ single-attribute functions; often a good approximation
Preference structure: Stochastic

Need to consider preferences over lotteries:
\( X \) is utility-independent of \( Y \) iff
preferences over lotteries in \( X \) do not depend on \( Y \)

Mutual U.I.: each subset is U.I of its complement
\[ \Rightarrow \exists \text{ multiplicative utility function:} \]
\[ U = k_1U_1 + k_2U_2 + k_3U_3 \]
\[ + k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_3k_1U_3U_1 \]
\[ + k_1k_2k_3U_1U_2U_3 \]

Routine procedures and software packages for generating preference tests to identify various canonical families of utility functions
Value of information

Idea: compute value of acquiring each possible piece of evidence
Can be done directly from decision network

Example: buying oil drilling rights
  Two blocks $A$ and $B$, exactly one has oil, worth $k$
  Prior probabilities 0.5 each, mutually exclusive
  Current price of each block is $k/2$
  “Consultant” offers accurate survey of $A$. Fair price?

Solution: compute expected value of information
  $=$ expected value of best action given the information
    minus expected value of best action without information
Survey may say “oil in $A$” or “no oil in $A$”, prob. 0.5 each (given!)
  $= [0.5 \times \text{value of “buy } A\text{” given “oil in } A\text{”}
      + 0.5 \times \text{value of “buy } B\text{” given “no oil in } A\text{”}]$
  $\quad - 0$
  $= (0.5 \times k/2) + (0.5 \times k/2) - 0 = k/2$
General formula

Current evidence $E$, current best action $\alpha$
Possible action outcomes $S_i$, potential new evidence $E_j$

$$EU(\alpha|E) = \max_a \sum_i U(S_i) \ P(S_i|E, a)$$

Suppose we knew $E_j = e_{jk}$, then we would choose $\alpha_{e_{jk}}$ s.t.

$$EU(\alpha_{e_{jk}}|E, E_j = e_{jk}) = \max_a \sum_i U(S_i) \ P(S_i|E, a, E_j = e_{jk})$$

$E_j$ is a random variable whose value is currently unknown
⇒ must compute expected gain over all possible values:

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

(VPI = value of perfect information)
Properties of VPI

**Nonnegative**—in expectation, not post hoc

\[ \forall j, E \ VPI_E(E_j) \geq 0 \]

**Nonadditive**—consider, e.g., obtaining \( E_j \) twice

\[ VPI_E(E_j, E_k) \neq VPI_E(E_j) + VPI_E(E_k) \]

**Order-independent**

\[ VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,E_j}(E_k) = VPI_E(E_k) + VPI_{E,E_k}(E_j) \]

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

\[ \Rightarrow \] evidence-gathering becomes a **sequential** decision problem
Qualitative behaviors

a) Choice is obvious, information worth little
b) Choice is nonobvious, information worth a lot
c) Choice is nonobvious, information worth little