Last update: April 15, 2010

Rational decisions

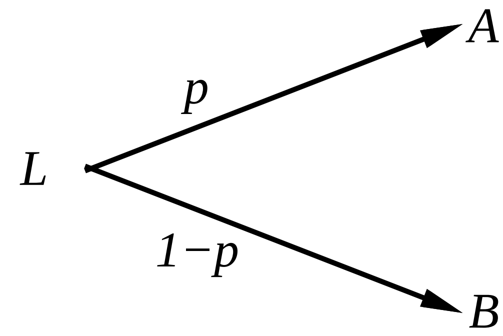
CMSC 421: Chapter 16

Outline

- ♢ Rational preferences
- ♢ Utilities
- ♢ Money
- ♢ Multiattribute utilities
- ♢ Decision networks
- ♢ Value of information

Preferences

An agent chooses among *prizes* (A, B, 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
$A \succsim B$	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:
$(A \succ B) \vee (B \succ A) \vee (A \sim B)$

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

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

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

Monotonicity:
$A \succ B \ \Rightarrow\ (p \geq q \ \Leftrightarrow\ [p, A;\ 1-p, B] \succsim [q, A;\ 1-q, B])$

Rational preferences contd.

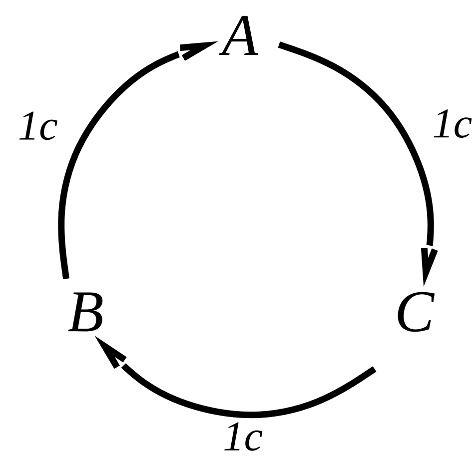
What happens if an agent's preferences violate the constraints?

Example: intransitive preferences

If $B \succ C$, then an agent who has C would trade C plus some money to get B

If $A \succ B$, then an agent who has B would trade B plus some money to get A

If $C \succ A$, then an agent who has A would trade A plus some money to get C



Rational preferences contd.

What happens if an agent's preferences violate the constraints?

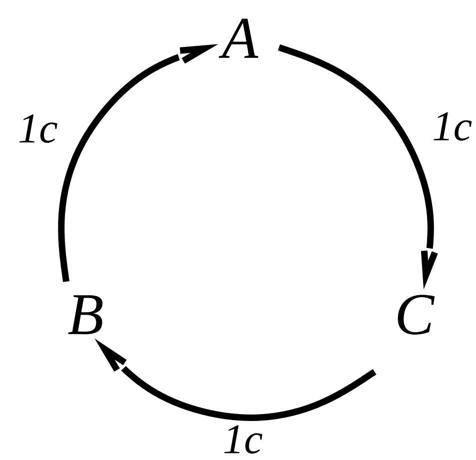
It leads to self-evident irrationality

Example: intransitive preferences

If $B \succ C$, then an agent who has C would trade C plus some money to get B

If $A \succ B$, then an agent who has B would trade B plus some money to get A

If $C \succ A$, then an agent who has A would trade A plus some money to get C



An agent with intransitive preferences can be induced to give away all its money

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

$$U(A) \geq U(B) \quad \Leftrightarrow \quad A \succsim B$$
$$U([p_1, S_1; \; \ldots \; ; \; p_n, S_n]) = \sum_i \, p_i U(S_i)$$

MEU principle:
Choose the action that maximizes the expected utility

Note: an agent can maximize the expected utility without ever representing or manipulating utilities and probabilities
E.g., a lookup table to play tic-tac-toe perfectly

Human utilities

Utilities map states to real numbers. Which numbers?

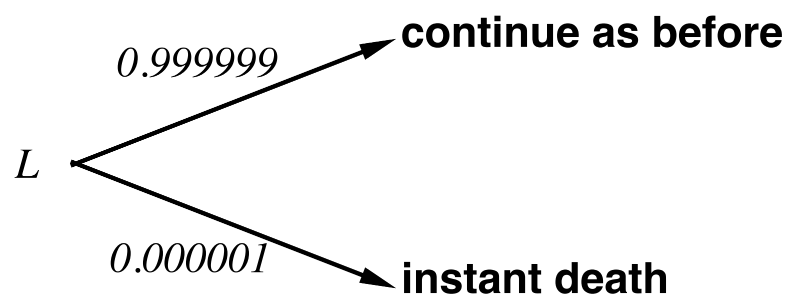
Standard approach to assessing human utilities:

Compare a given state A to a *standard lottery* L_p that has

- "best possible prize" $u_{\max}$ with probability p
- "worst possible catastrophe" $u_{\min}$ with probability $(1-p)$

Adjust lottery probability p until $A \sim L_p$

How much would you pay to avoid a $1/1{,}000{,}000$ chance of death?



Human utilities

Utilities map states to real numbers. Which numbers?

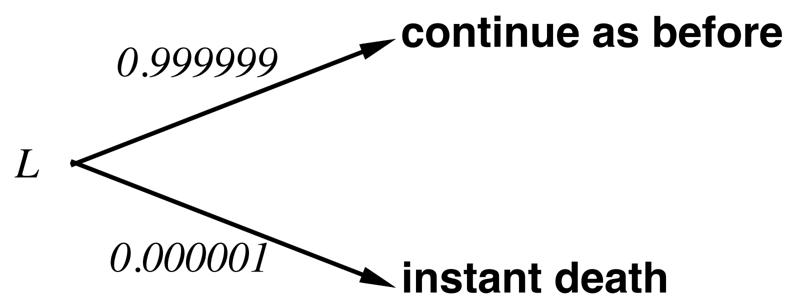
Standard approach to assessing human utilities:

Compare a given state A to a *standard lottery* L_p that has

- "best possible prize" $u_{\max}$ with probability p
- "worst possible catastrophe" $u_{\min}$ with probability $(1-p)$

Adjust lottery probability p until $A \sim L_p$

Judging from people's **actions**, they will pay about **$20** to avoid a 1/1,000,000 chance of death



One *micromort*

$\approx P$(accidental death in 370 km of car travel)

$\approx P$(accidental death in 9700 km of train travel)

Utility scales

Note: behavior is **invariant** w.r.t. positive linear transformation

Let

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

Then U' models the same preferences that U does.

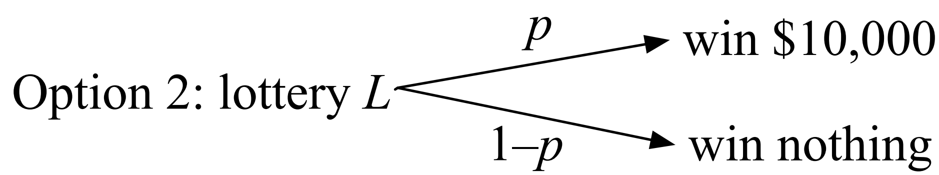
Normalized utilities:

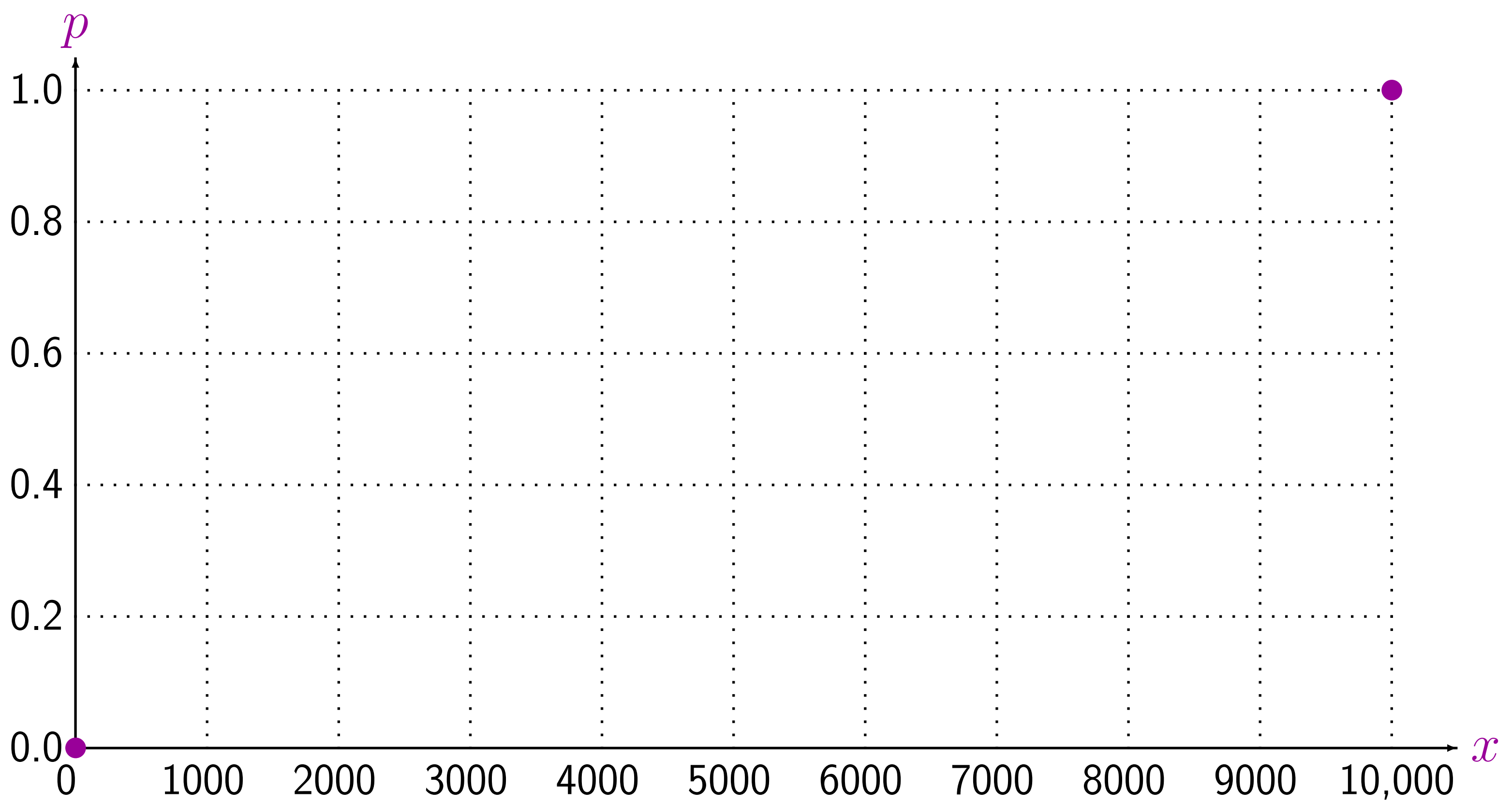
define U' such that $0 \le U'(x) \le 1$ for all x

The utility of money

For each amount x, adjust p until half the class votes for each option:

Option 1: you win \$$x$.





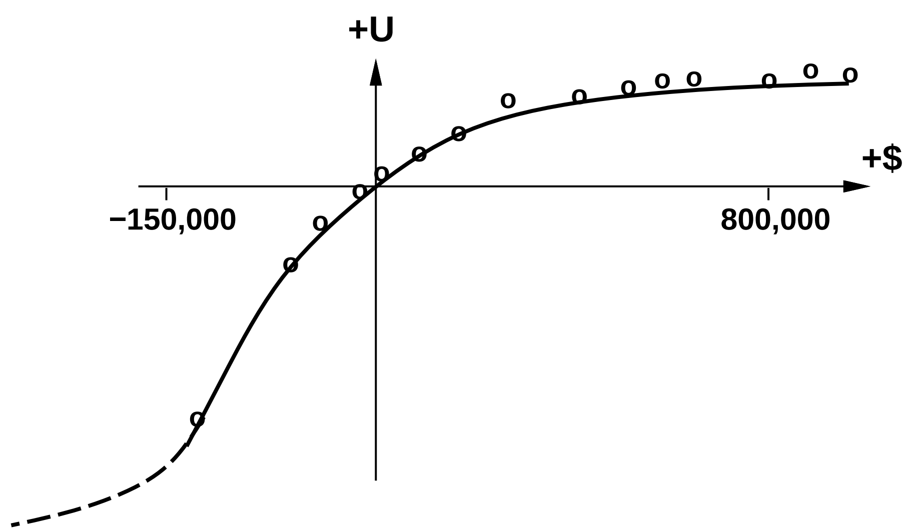
What the book says

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*

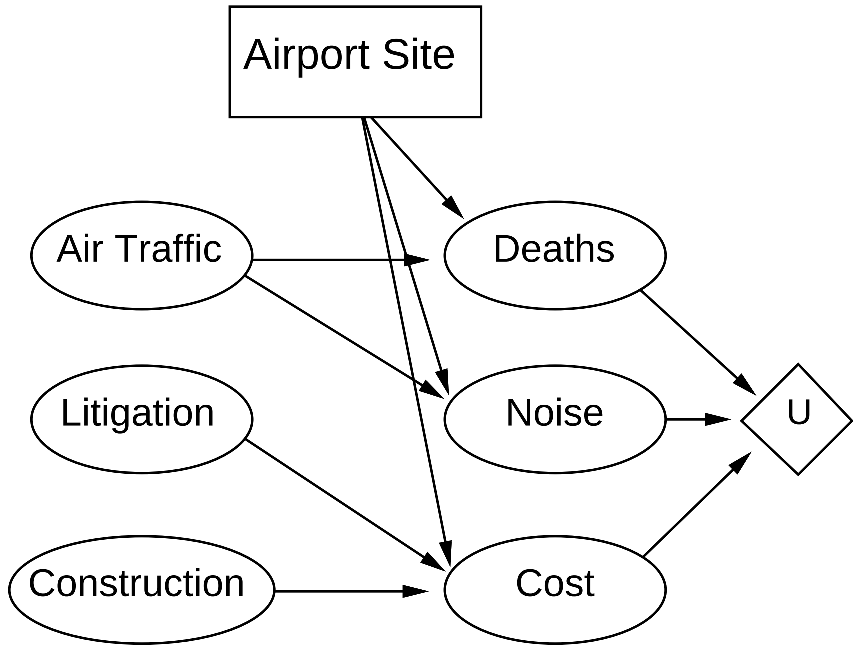
Utility curve: for what probability p am I indifferent between a prize x and a lottery $[p, \$M; \ (1-p), \$0]$ for large M?

Typical empirical data, extrapolated with *risk-prone* behavior:



Decision networks

Add *action nodes* and *utility nodes* to causal networks
to enable rational decision making



Algorithm:

For every possible value of the action node
compute E(utility node | action, evidence)
Return MEU action

Multiattribute utility

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 behavior?

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

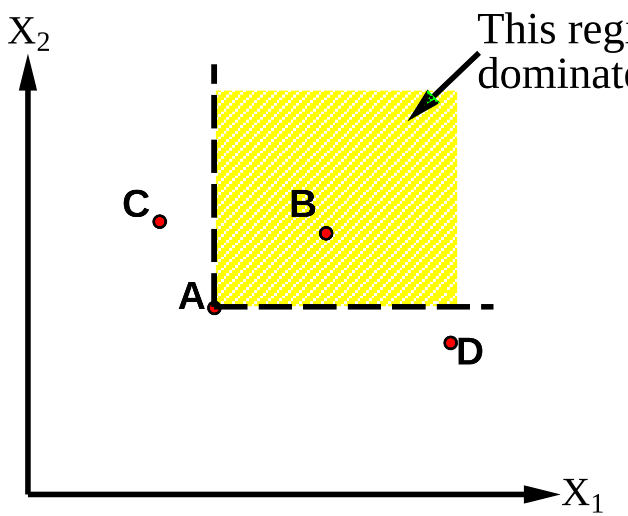
Idea 2: identify various types of **independence** in preferences and derive consequent canonical forms for $U(x_1, \ldots, x_n)$

Strict dominance

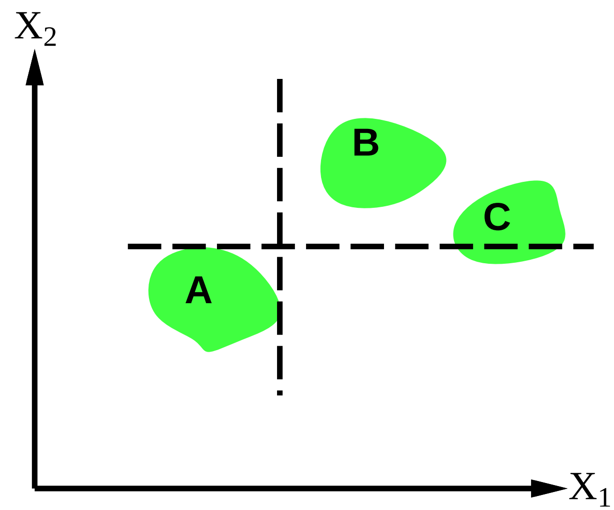
Typically define attributes such that U is *monotonic* in each attribute

Strict dominance: choice B strictly dominates choice A iff

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



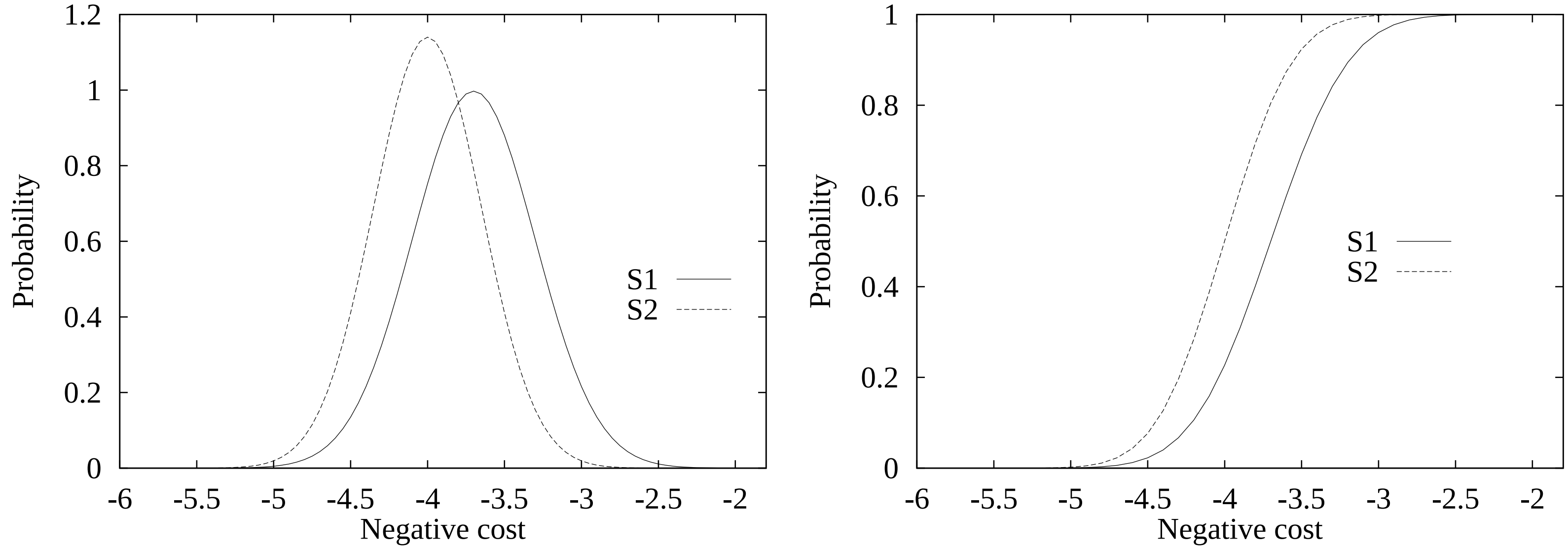
Deterministic attributes



Uncertain attributes

Strict dominance seldom holds in practice

Stochastic dominance



Choices S_1 and S_2 with continuous distributions p_1 and p_2

S_1 *stochastically dominates* S_2 iff $\forall\, t \;\; P(S_1 \leq t) \leq P(S_2 \leq t)$,

i.e., $\forall\, t \;\; \int_{-\infty}^{t} p_1(x)dx \leq \int_{-\infty}^{t} p_2(t)dt$

If S_1 stochastically dominates S_2 and U is monotonic in x, then

$$EU(S_1) = \int_{-\infty}^{\infty} p_1(x)U(x)dx \geq \int_{-\infty}^{\infty} p_2(x)U(x)dx = EU(S_2)$$

If p_1, p_2 are discrete, use sums instead of integrals

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 is closer to the city than S_2

$\Rightarrow$ S_1 stochastically dominates S_2 on cost

E.g., injury increases with collision speed

Can annotate belief networks with stochastic dominance information:

$X \xrightarrow{+} Y$ (X positively influences Y) means that

For every value $\mathbf{z}$ of Y's other parents $\mathbf{Z}$

$\forall\, x_1, x_2 \;\; x_1 \geq x_2 \Rightarrow \mathbf{P}(Y|x_1, \mathbf{z})$ stochastically dominates $\mathbf{P}(Y|x_2, \mathbf{z})$

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$20,000 suffer, \$4.6 billion, 0.06 deaths/mpm$\rangle$ vs.
$\langle$70,000 suffer, \$4.2 billion, 0.06 deaths/mpm$\rangle$

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

Theorem (Debreu, 1960): mutual P.I. $\Rightarrow$ $\exists$ *additive* value function:
If the attributes of S are $X_1, X_2, \ldots, X_n$, then

$$V(S) = \Sigma_i V_i(X_i(S))$$

Hence assess n single-attribute functions; often a good approximation

Preference structure: Stochastic

Need to consider preferences over lotteries:
$\mathbf{X}$ is *utility-independent* of $\mathbf{Y}$ iff
  preferences over lotteries in $\mathbf{X}$ do not depend on $\mathbf{y}$

The lotteries in $\mathbf{X} = \{X_1, \ldots, X_k\}$ are mutually U.I. if every subset of $\mathbf{X}$ is U.I. of its complement
$\Rightarrow \;\; \exists$ *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 $k/2$
How much to pay a consultant for an accurate survey of A?

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** (from above)
$= [0.5 \times$ value of "buy A" given "oil in A"
$+ 0.5 \times$ value of "buy B" given "no oil in A"]
– 0
$= (0.5 \times k/2) + (0.5 \times k/2) - 0 = k/2$

General formula

Current evidence E, current best action α
Possible action outcomes $\{S_1, S_2, \ldots\}$

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

Potential new evidence E_j

If we knew $E_j = e_j$, then we would choose α_{e_j} s.t.

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

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

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

(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, greedy selection (next one to gather = the one of maximum VPI) isn't always optimal

Can have situations where
$VPI(E_1|E) > VPI(E_1'|E)$ and $VPI(E_2|E, E_1) > VPI(E_2'|E, E_1)$

but $VPI(E_1, E_2|E) < VPI(E_1', E_2'|E)$

$\Rightarrow$ 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