# CS 561: Artificial Intelligence

Instructor: Sofus A. Macskassy, macskass@usc.edu
TAs: Nadeesha Ranashinghe (nadeeshr@usc.edu)
William Yeoh (wyeoh@usc.edu)
Harris Chiu (chiciu@usc.edu)

Lectures: MW 5:00-6:20pm, OHE 122 / DEN
Office hours: By appointment
Class page: http://www-rcf.usc.edu/~macskass/CS561-Spring2010/

This class will use http://www.uscden.net/ and class webpage
- Up to date information
- Lecture notes
- Relevant dates, links, etc.

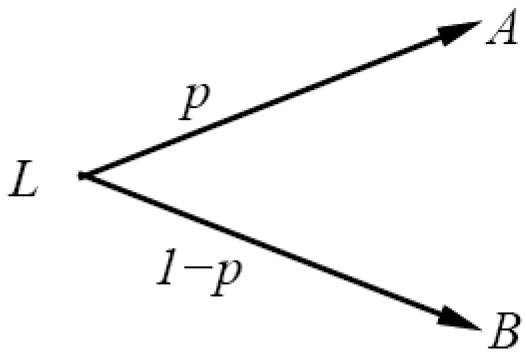
Course material:
[AIMA] Artificial Intelligence: A Modern Approach,
by Stuart Russell and Peter Norvig. (2nd ed)

# Rational Decisions [AIMA Ch. 16]

- Rational preferences
- Utilities
- Money
- Multi-attribute 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:

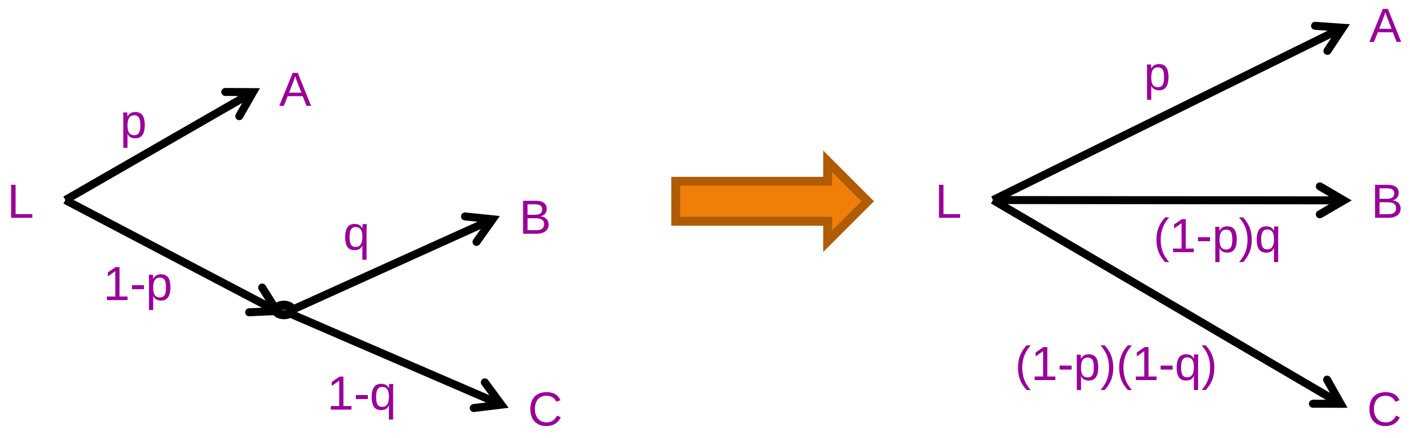
| | |
|---|---|
| $A \succ B$ | A preferred to B |
| $A \sim B$ | indifference between A and B |
| $A \succsim B$ | B not preferred to A |

# Rational preferences

- Idea: preferences of a rational agent must obey constraints.
- Rational preferences $\Rightarrow$
  behavior describable as maximization of expected utility
- Constraints:
  Orderability
  $(A \succ B) \lor (B \succ A) \lor (A \sim B)$
  Transitivity
  $(A \succ B) \land (B \succ C) \Rightarrow (A \succ C)$
  Continuity
  $A \succ B \succ C \Rightarrow \exists\ p\ [p, A; 1\text{-}p, C] \sim B$
  Substitutability
  $A \sim B \Rightarrow [p, A; 1\text{-}p, C] \sim [p, B; 1\text{-}p, C]$
  Monotonicity
  $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; 1\text{-}p, B] \succsim [q, A; 1\text{-}q, B])$
  Decomposability
  $[p, A; 1\text{-}p, [q, B, (1\text{-}q), C]] \sim [p, A; (1\text{-}p)q, B; (1\text{-}p)(1\text{-}q), C]$

# Decomposability

- A "complex" hierarchical lottery can be collapsed into a single multi-choice lottery:



L p A 1-p q B 1-q C

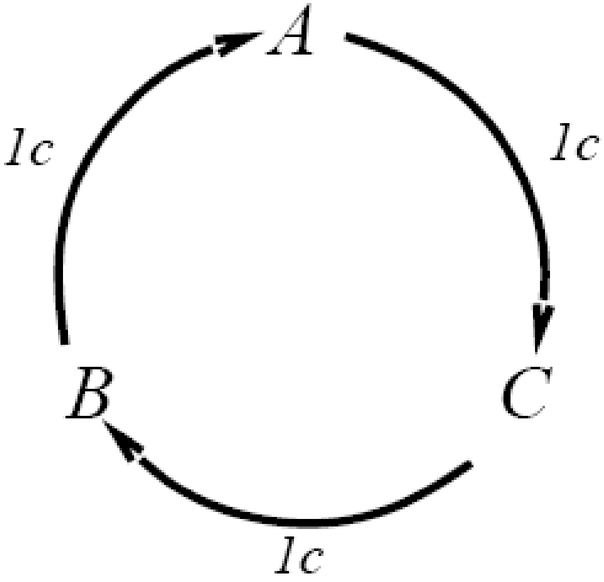
L p A (1-p)q B (1-p)(1-q) C

[p, A; 1-p, [q, B, (1-q), C]]

[p, A; (1-p)q, B; (1-p)(1-q), C]

# Rational preferences cont'd

- 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

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

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

  $U(A) \geq U(B) \Leftrightarrow A \succsim B$

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

- MEU principle:

  Choose the action that maximizes expected utility

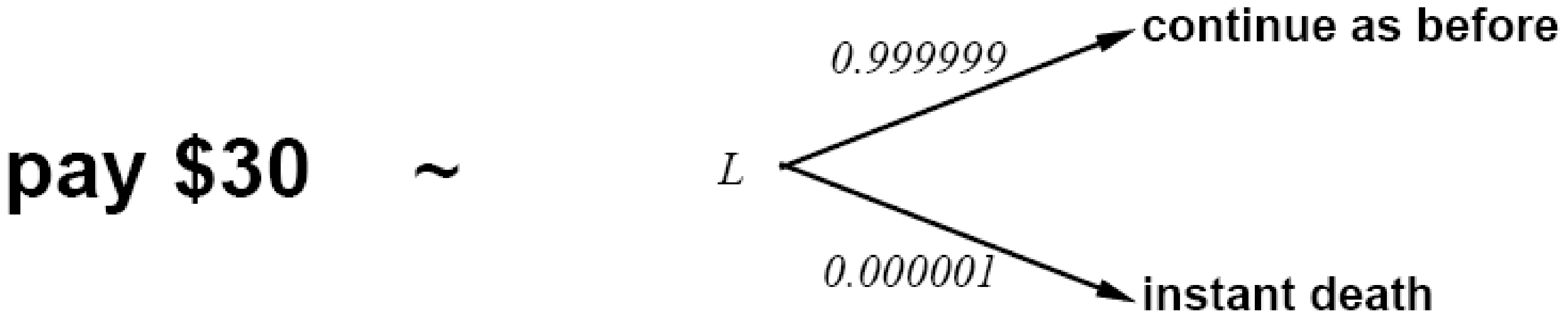
- Note: an agent can be entirely rational (consistent with MEU) without ever representing or manipulating utilities and probabilities
- E.g., a lookup table for perfect tic-tac-toe

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



# Utility scales

- Normalized utilities: $u_\top = 1{:}0$, $u_\perp = 0{:}0$
- 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

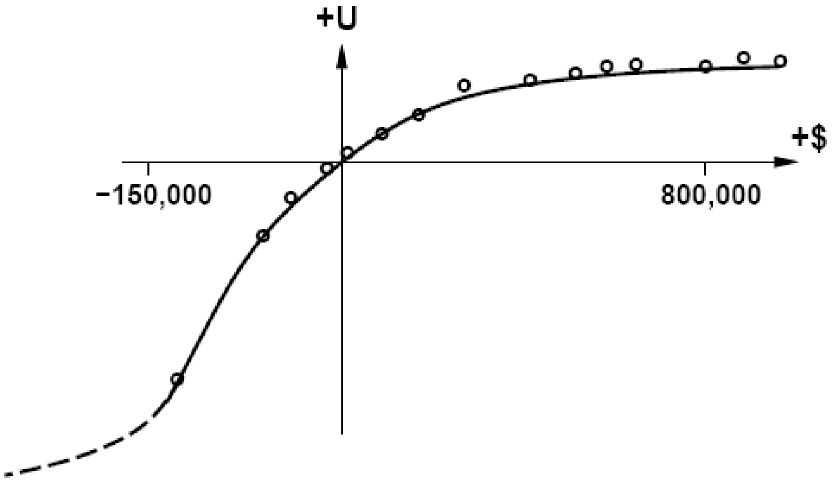
- Note: behavior is **invariant** w.r.t. +ve linear transformation

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

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

# Money

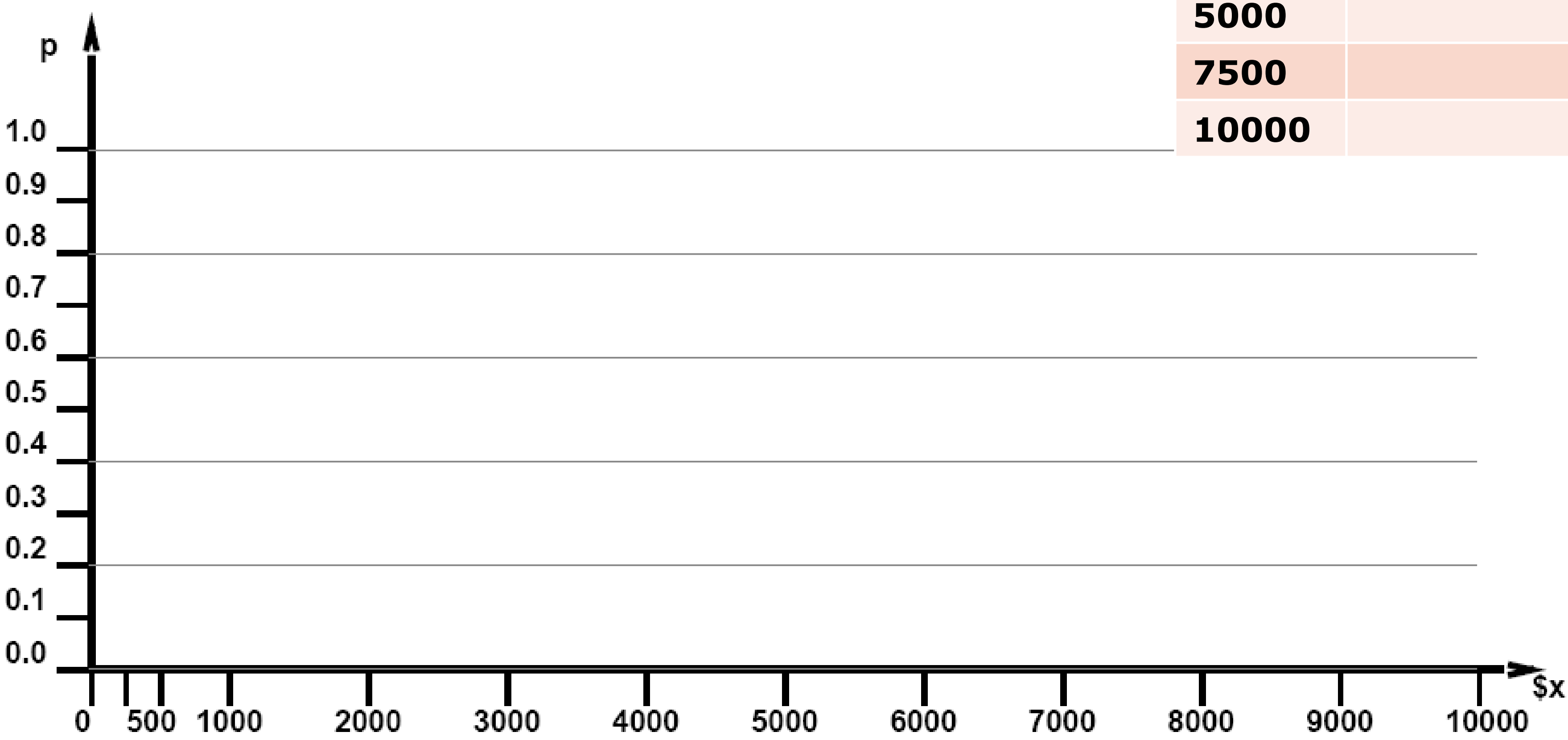
- Money does **not** behave as a utility function
- Given a lottery L with expected monetary value EMV (L), usually U(L) < U(EMV (L)), i.e., people are risk-averse
- Utility curve: for what probability p am I indifferent between a prize x and a lottery [p, \$M; (1-p), \$0] for large M?
- Typical empirical data, extrapolated with risk-prone behavior:



# Student group utility

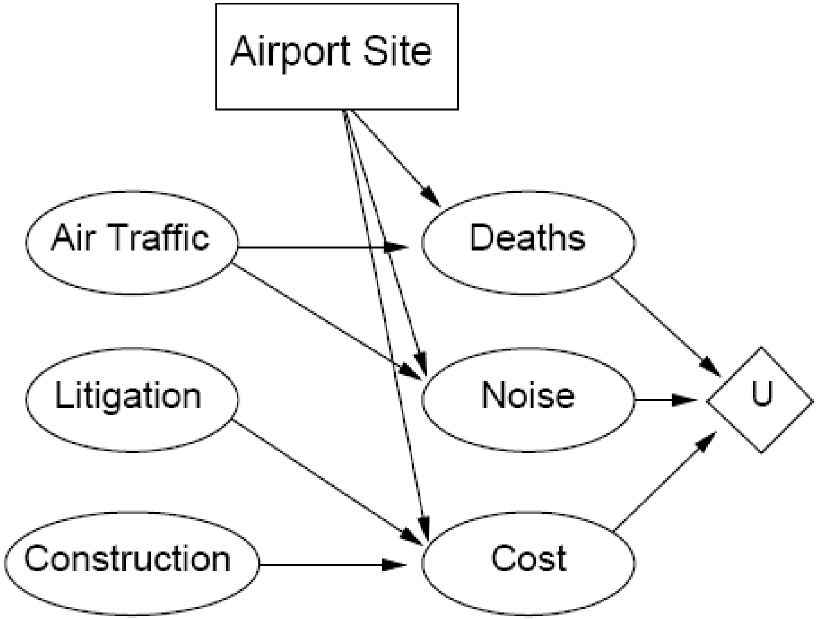
| Pay | Prob |
|---|---|
| 500 | |
| 1000 | |
| 2000 | |
| 3000 | |
| 4000 | |
| 5000 | |
| 7500 | |
| 10000 | |

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
```

# 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 under which decisions can be made without complete identification of $U(x_1, \ldots, x_n)$

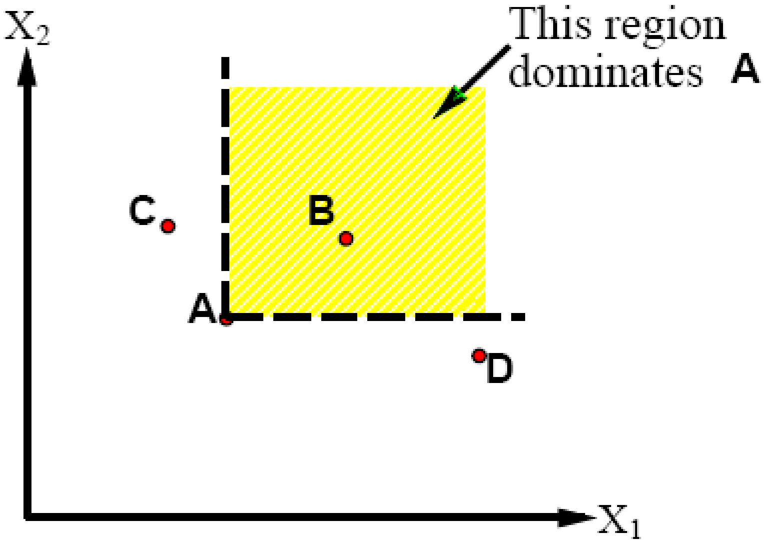
- Idea 2: identify various types of **independence** in preferences and derive consequent canonical forms for $U(x_1, \ldots, x_n)$

# Strict dominance

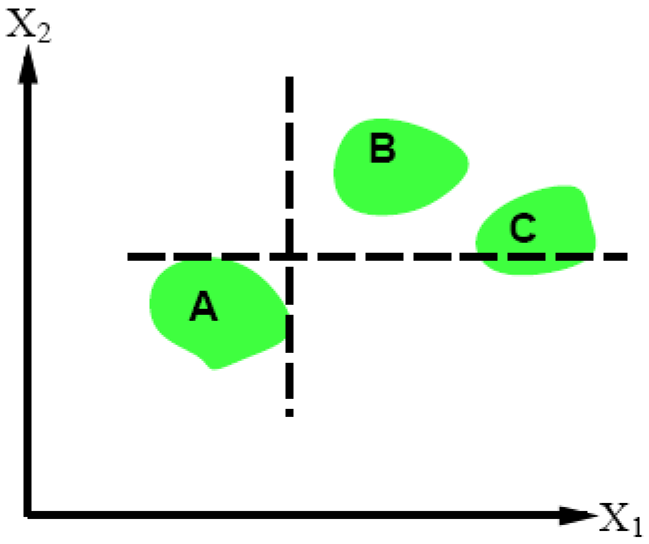
Typically define attributes s.t. U is monotonic in each

Strict dominance: choice B strictly dominates choice A iff

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



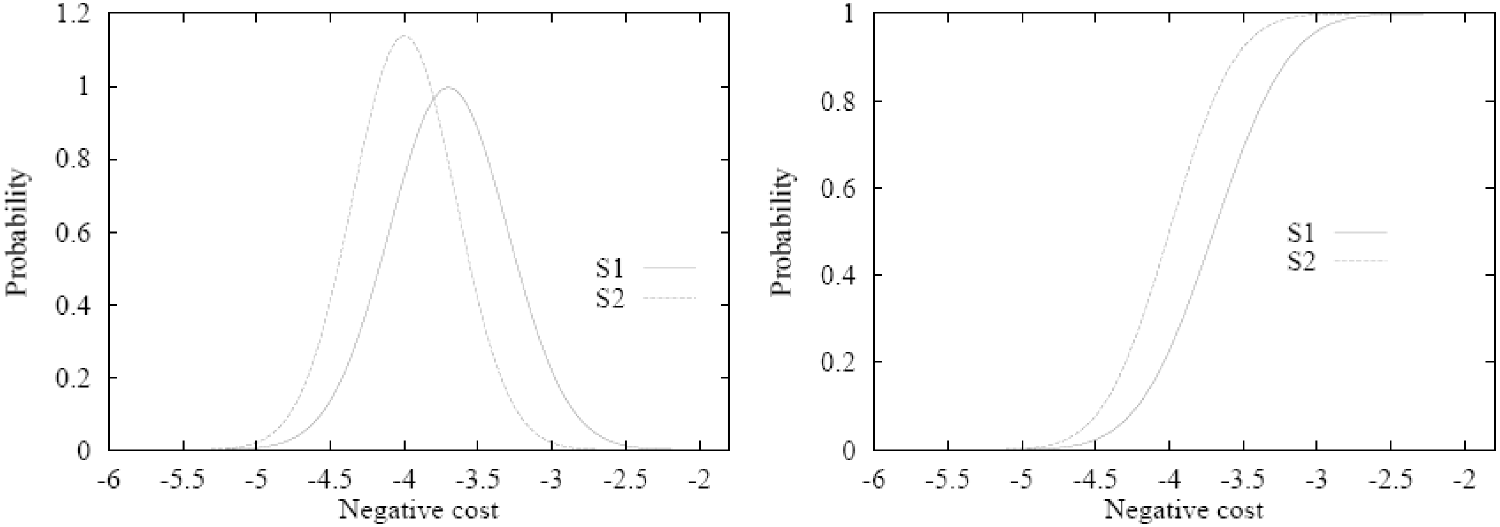
**Deterministic attributes**



**Uncertain attributes**

Strict dominance seldom holds in practice

# Stochastic dominance



- Distribution $p_1$ stochastically dominates distribution $p_2$ iff

$$\forall x \int_{-\infty}^{x} p_1(x')dx' \le \int_{-\infty}^{x} p_2(x')dx'$$

- If $U$ is monotonic in $x$, then $A_1$ with outcome distribution $p_1$ stochastically dominates $A_2$ with outcome distribution $p_2$:

$$\int_{-\infty}^{\infty} p_1(x)U(x)dx \ge \int_{-\infty}^{\infty} p_2(x)U(x)dx$$

- Multiattribute case: stochastic dominance on all attributes $\Rightarrow$ optimal

# Stochastic dominance cont'd

- 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

# 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., ⟨Noise, Cost, Safety⟩:
  ⟨20,000 suffer, \$4.6 billion, 0.06 deaths/mpm⟩ vs.
  ⟨70,000 suffer, \$4.2 billion, 0.06 deaths/mpm⟩
- **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. $\Rightarrow \exists$ 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:
  $\mathbf{X}$ is utility-independent of $\mathbf{Y}$ iff
  preferences over lotteries in $\mathbf{X}$ do not depend on $\mathbf{y}$

- Mutual U.I.: each subset is U.I of its complement
  $\Rightarrow \exists$ multiplicative utility function:
  $$U = k_1U_1 + k_2U_2 + k_3U_3 + k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_3k_1U_3U_1 + k_1k_2k_3U_1U_2U_3$$
  where $U_i = U_i(x_i)$

- 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.

    What is a fair price?

# Value of information (cont'd)

- Solution: compute expected value of information
  = expected value of an action given the information
  minus expected value of action without information
- Select action with highest value of information
- Survey may say "oil in A" or "no oil in A", **prob. 0.5 each** (given!)

  = [0.5 × value of "buy A" given "oil in A"

  + 0.5 × value of "buy B" given "no oil in A"]

  - 0

  = (0.5 × k/2) + (0.5 × k/2) - 0 = k/2

# General formula

Current evidence $E$, current best action $\alpha$

Possible action outcomes $S_i$, potential new evidence $E_j$

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

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

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

$E_j$ is a random variable whose value is currently unknown

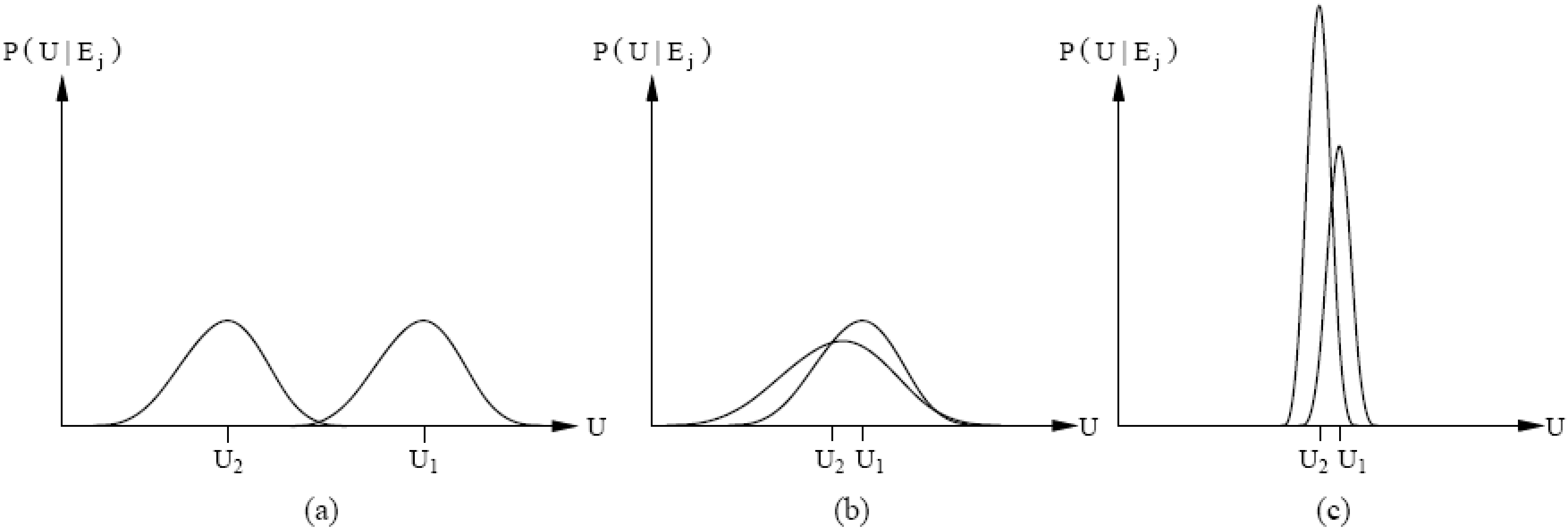
⇒ must compute expected gain over all possible values:

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

(VPI = value of perfect information)

# Qualitative behaviors

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



# Properties of VPI

- **Nonnegative**—in **expectation**, not **post hoc**
  $\forall\, j,\ E\ \mathrm{VPI}_E(E_j) \geq 0$

- **Nonadditive**—consider, e.g., obtaining $E_j$ twice
  $\mathrm{VPI}_E(E_j,E_k) \neq \mathrm{VPI}_E(E_j) + \mathrm{VPI}_E(E_k)$

- **Order-independent**
  $\mathrm{VPI}_E(E_j,E_k) = \mathrm{VPI}_E(E_j) + \mathrm{VPI}_{E,E_j}(E_k) = \mathrm{VPI}_E(E_k) + \mathrm{VPI}_{E,E_k}(E_j)$

- 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