Rational decisions

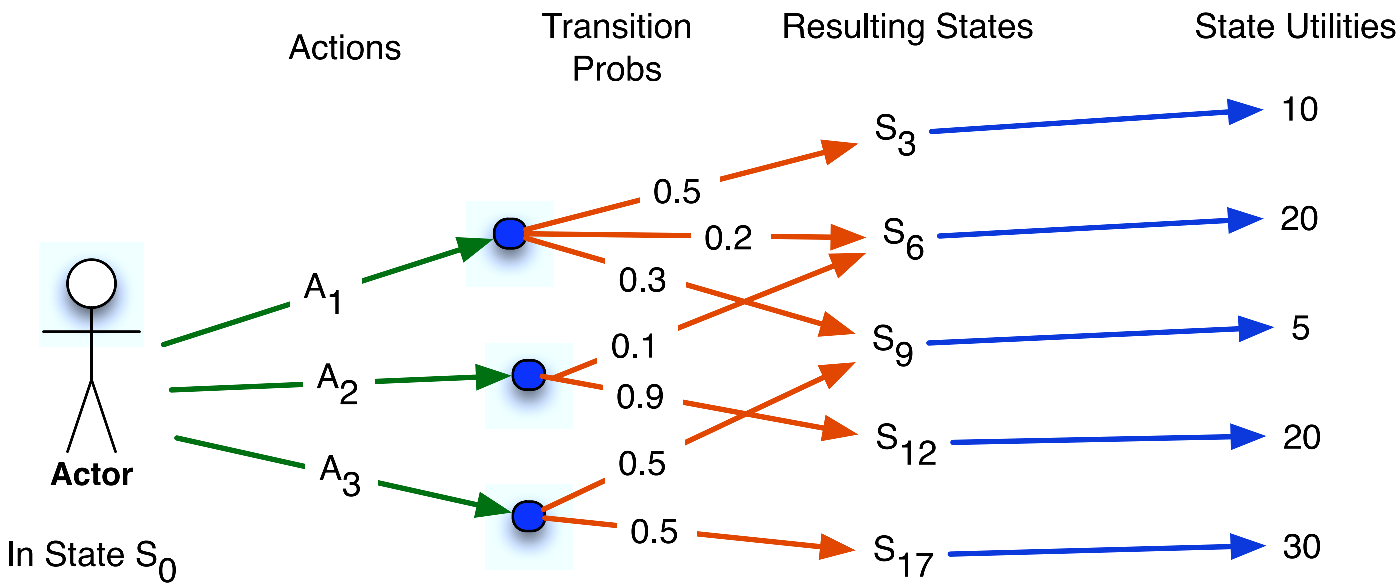
Chapter 16

Outline

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

Rational Decision Making

Do the Right Thing = Choose Action that Maximizes Expected Utility



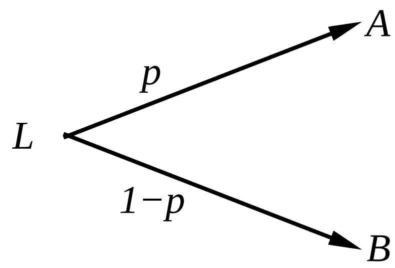
$$
\begin{aligned}
EU(A_1) = (0.5)(10) + (0.2)(20) + (0.3)(5) &= 10.5 \\
EU(A_2) = (0.1)(20) + (0.9)(20) &= 20 \\
EU(A_3) = (0.5)(5) + (0.5)(30) &= 17.5
\end{aligned}
$$

Do A_2!

Preferences

An agent chooses among prizes (A, B, etc.) and lotteries, i.e., situations with uncertain prizes/outcomes.

Lottery $L = [p, A;\ (1-p), B]$
Complete, mutually-exclusive set of probabilistic outcomes.



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.

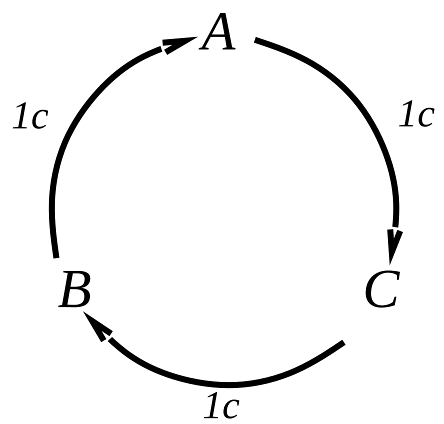
Violating the constraints leads to self-evident irrationality

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

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

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

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

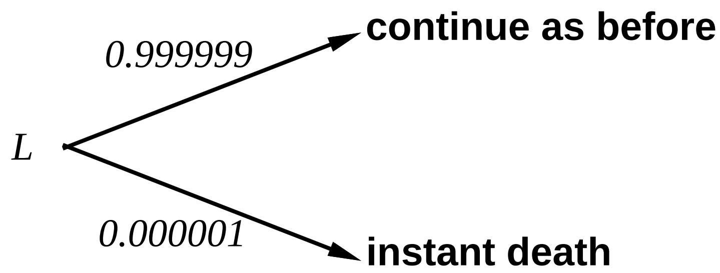
E.g., a lookup table for perfect tictactoe

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" $u_\top$ with probability p
"worst possible catastrophe" $u_\bot$ with probability $(1-p)$
adjust lottery probability p until $A \sim L_p$
I.e., indifference to receiving A or doing the lottery, or, $U(A) = U(L_p)$.

pay \$30 ~



Utility scales

Normalized utilities: $u_\top = 1.0$, $u_\bot = 0.0$

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

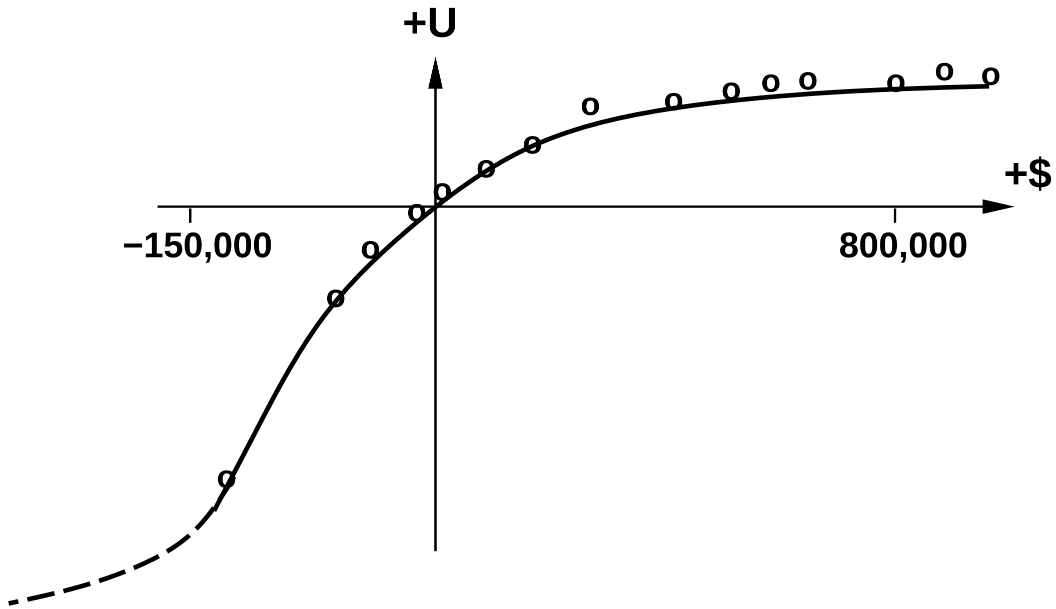
Money

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:



Importance of Money

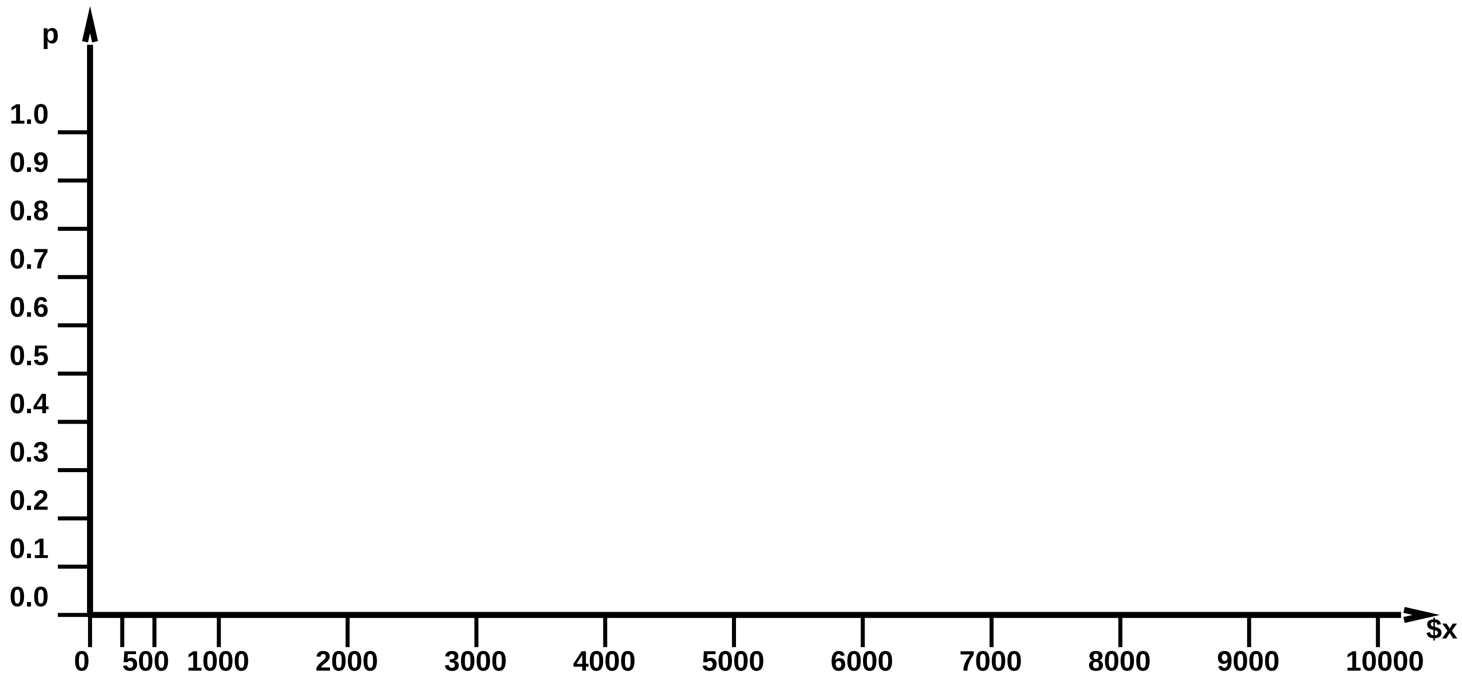
- The value (utility) of money is relative to the amount that you have.
- $\$1000$ has a much higher utility to a poor than a rich person.
- That is, if S is the amount that a person has, then $U(S+\$1000)-U(S)$ is much larger for a poor than for a rich person.
- I.e. U(Money) is a non-linear function: the slope is not constant for all values of S.
- Bernoulli (1738): $U(Money) \approx log(Money)$

Risk Aversion and Insurance

- Given a lottery L $= [p, \$M;\ (1-p), \$0]$, what amount, X, are you willing to receive instead of doing the lottery?
- If $X < EMV(L)$, then you are **risk averse**: You will accept a value LOWER than the expected lottery result just to avoid risky lottery.
- X is called the **certainty equivalent** of the lottery.
- **insurance premium** $= EMV(L) - X$.
  - Assume you have a house worth $M = \$1,000,000$
  - Assume that the probability of losing the house to fire is .001
  - Then lottery involves possible amounts that need to be paid out.
  - $L_h = [p = .001, \$M;\ (1-p) = .999, \$0]$ and $EMV(L_h) = \$1,000$
  - Hypothetically, how much, X, are you willing to receive (from the insurance company) to avoid that lottery?
  - The remainder, $EMV(L_h) - X$ is amount you pay for insurance.
  - If you are very risk averse, then X is low, and you are willing to pay a high insurance premium to avoid a personal economic catastrophe.

Student group utility

For each x, adjust p until half the class votes for lottery (M=10,000)



A (Real) Lottery Example

- L $= [p = .02, \$10; q = .0000005, \$1,000,000; (1-p-q) = .9799995, \$0]$
- A lottery ticket costs \$1.
- What is EMV(L)?
- When is it rational to buy a lottery ticket?

Answer:

- First, include the cost of the ticket in the lottery outcomes.
- L $= [p = .02, \$9; q = .0000005, \$999,999; (1-p-q) = .9799995, \$-1]$
- EMV(L) $=$ (.02)(9) $+$ (.0000005)(999,999) $+$ (.9799995)(-1) $=$ \$-0.3
- If we ignore the ticket cost, EMV(L) $=$ \$+0.7

Rationality and Lottery Utility

- Let S_k be the state involving the person's current assets, k.
- Then the utility of the lottery is based on $U(S_{k+n})$ for different amounts, n.
- Ignoring the cost of the ticket:
- $U(L) = (.02)U(S_{k+10}) + (.0000005)U(S_{k+1,000,000})$
- Now we can ask whether $U(L) > U(S_{k+1})$ (the utility of saving the ticket cost)
- And this depends upon U and k.
- If $U(L) > U(S_{k+1})$, then the ticket purchase is a rational one.

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 behaviour?

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

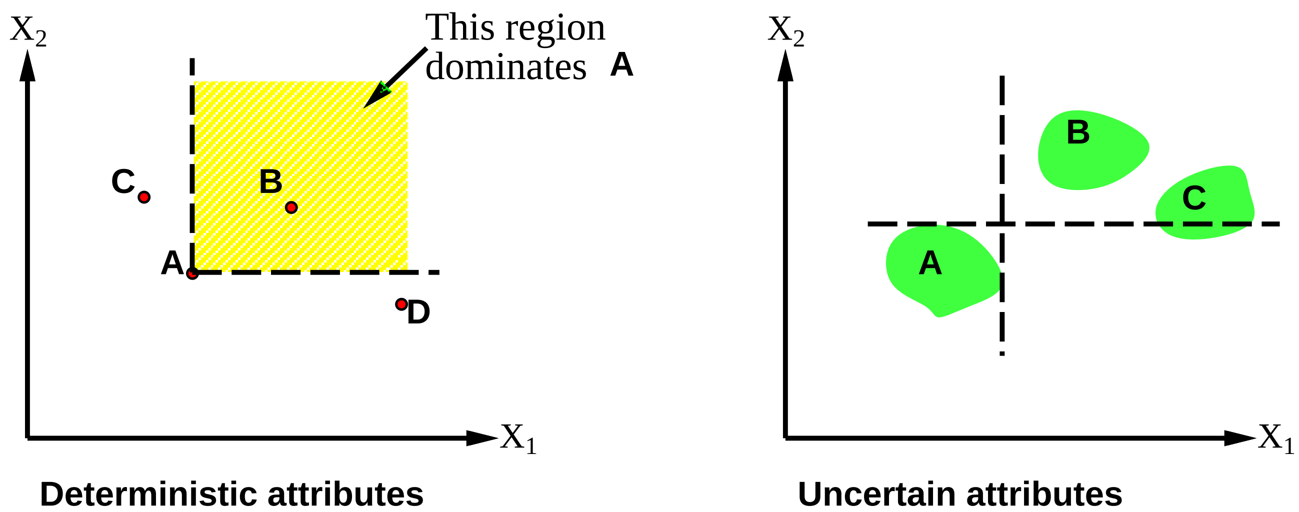
Idea 2: identify various types of **independence** in preferences and derive consequent canonical forms for $U(x_1, \ldots, x_n)$

Strict dominance

Typically define attributes such that U is monotonic in each

Strict dominance: choice B strictly dominates choice A iff

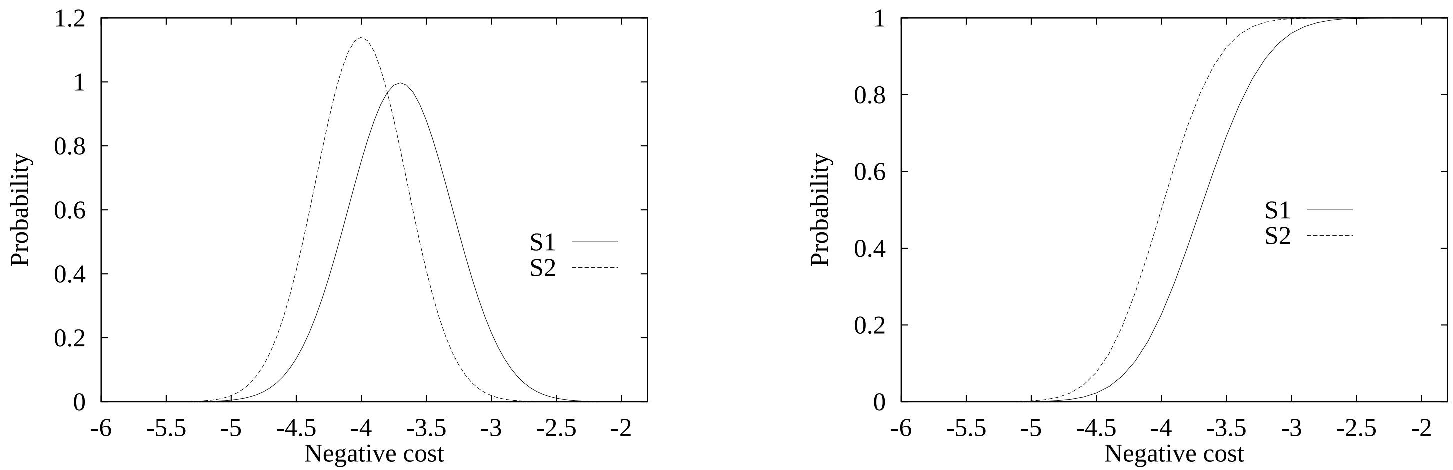
$$\forall i \;\; X_i(B) \geq X_i(A) \qquad \text{(and hence } U(B) \geq U(A)\text{)}$$



Strict dominance seldom holds in practice

Stochastic dominance

Table 1: default



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.

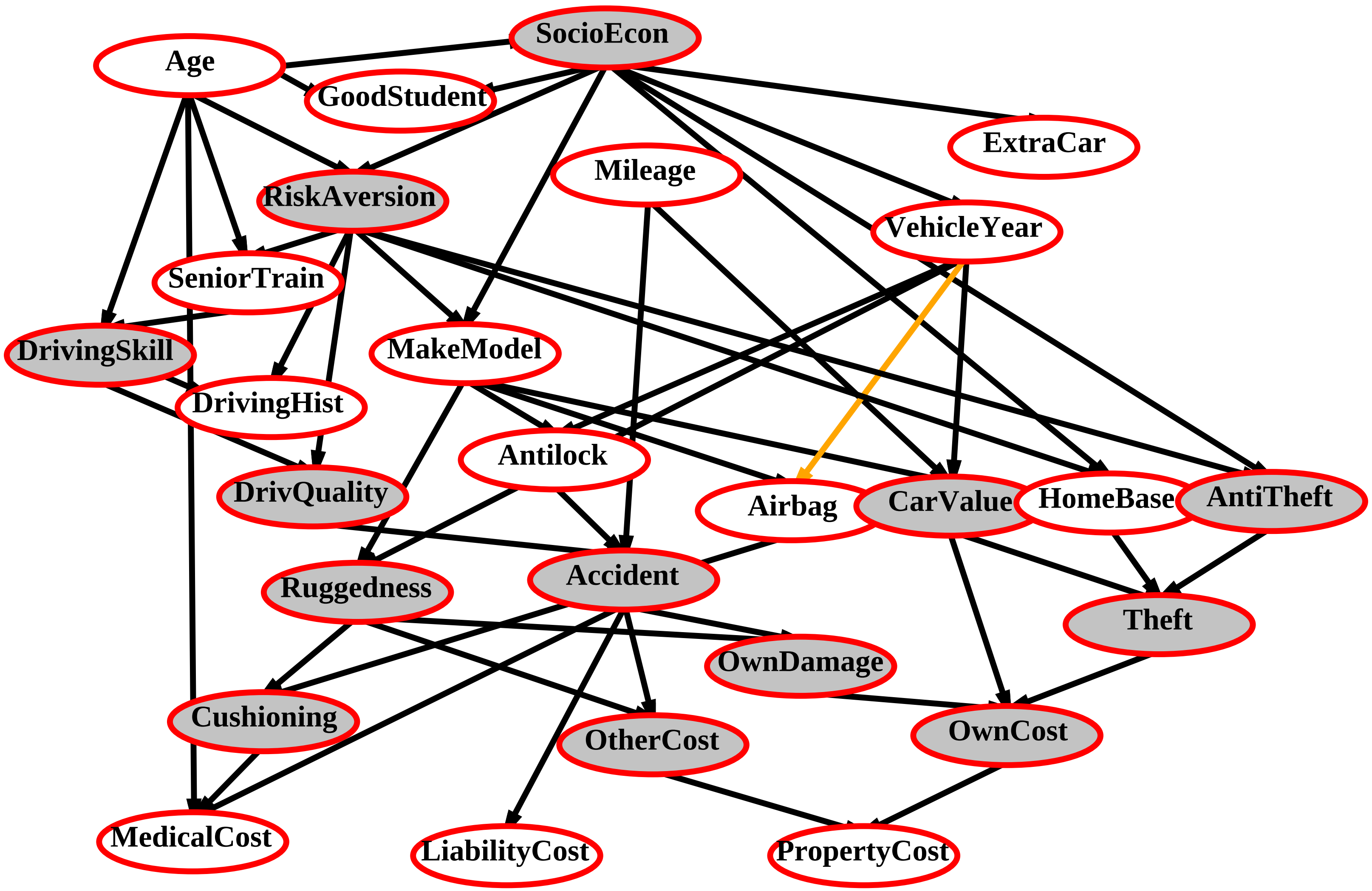
Stochastic dominance can often be determined without exact distributions using **qualitative** reasoning

E.g., construction cost increases with distance from city
S_1 is closer to the city than S_2
$\Rightarrow$ S_1 stochastically dominates S_2 on cost

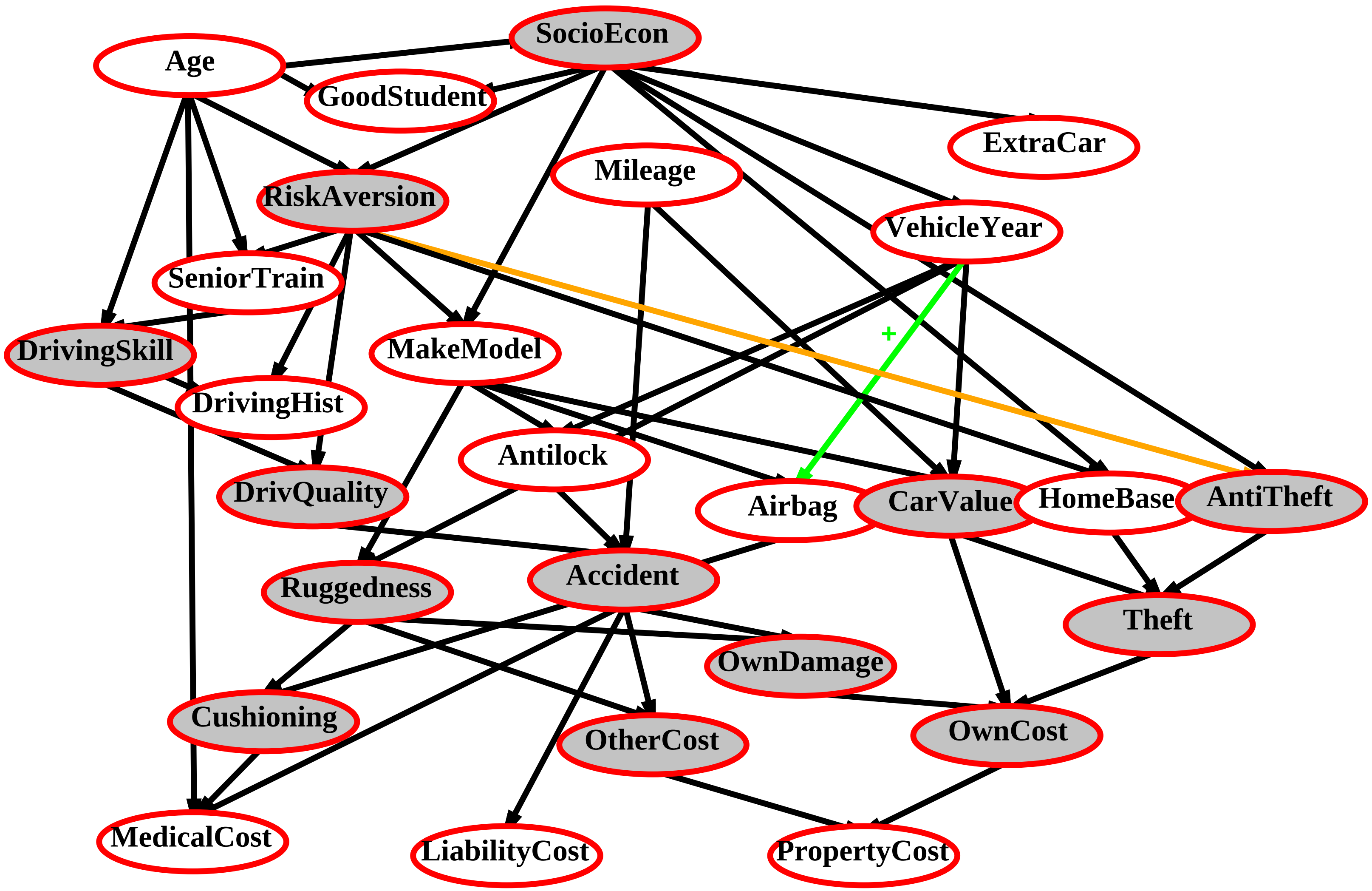
E.g., injury increases with collision speed

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 $\mathbf{P}(Y|x_2, \mathbf{z})$

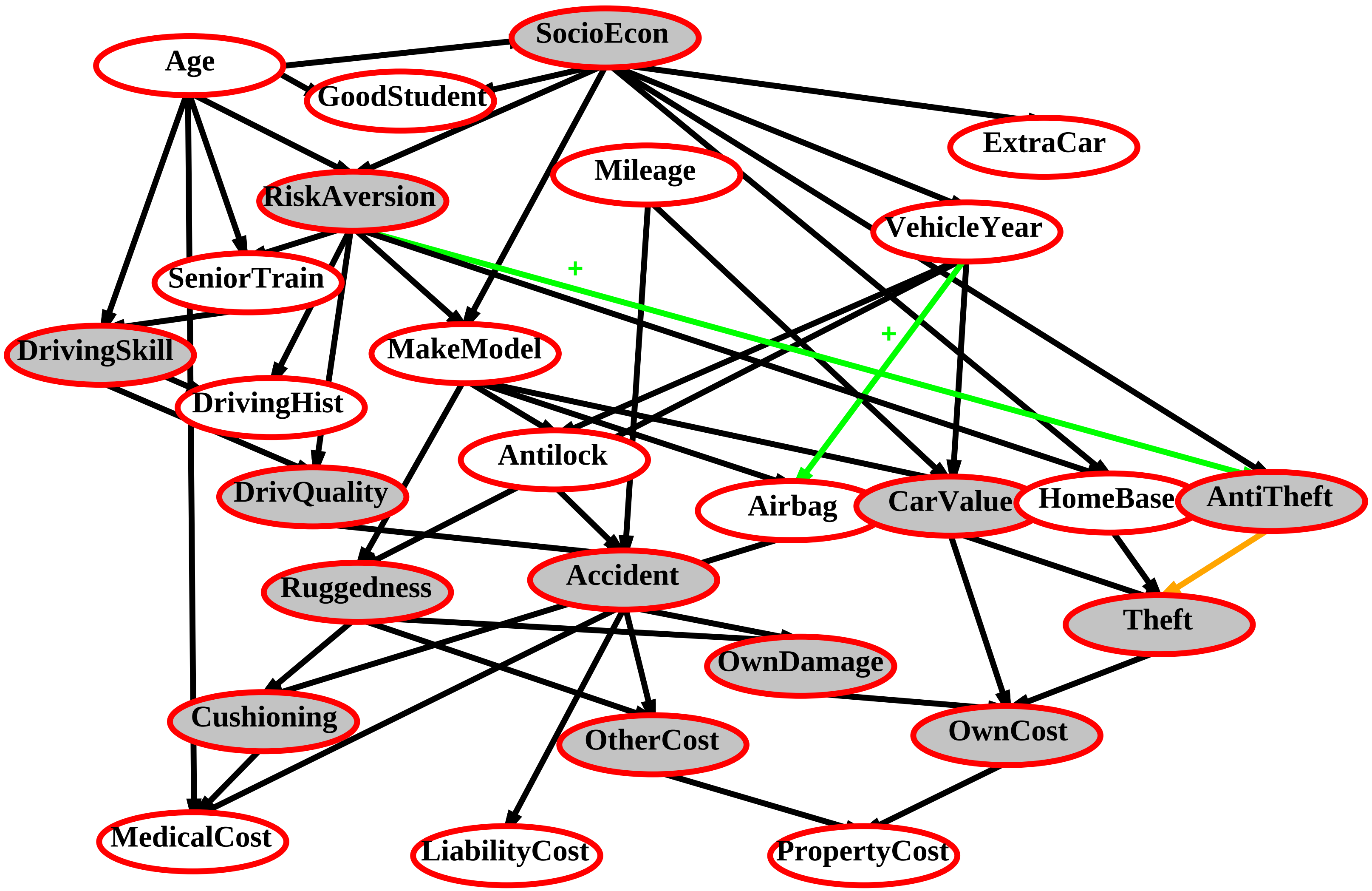
Label the arcs + or −



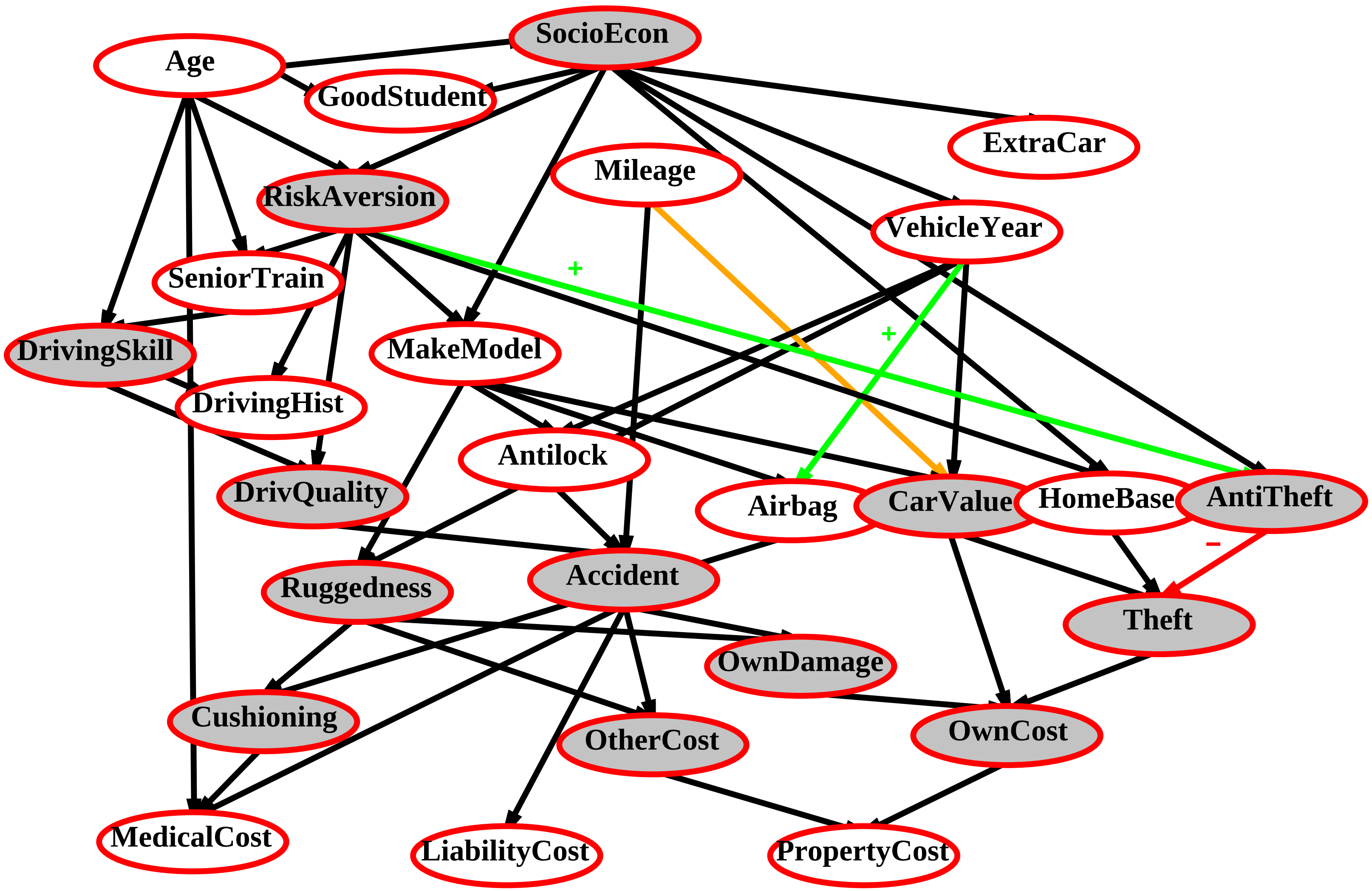
Label the arcs + or −



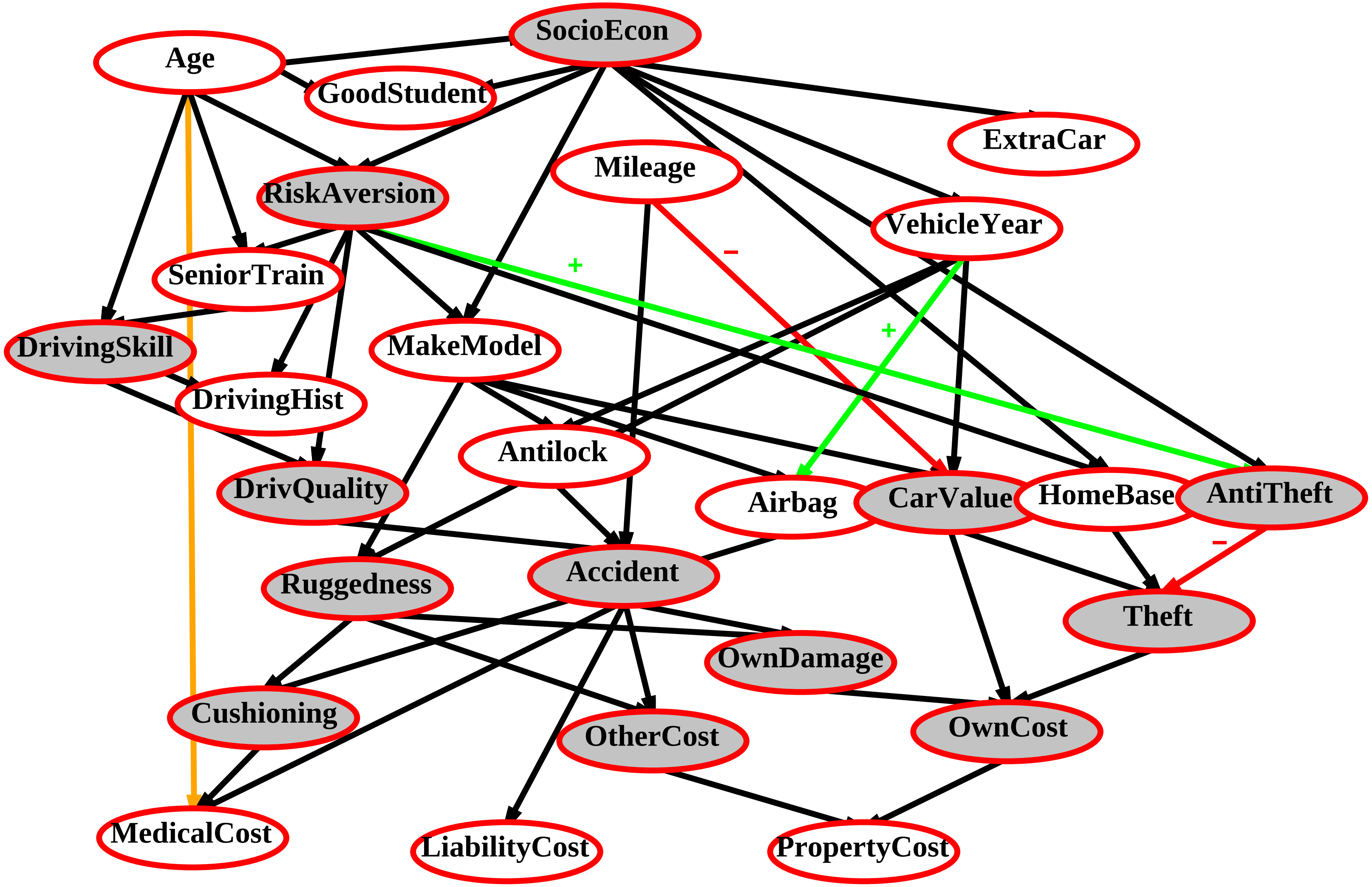
Label the arcs + or −



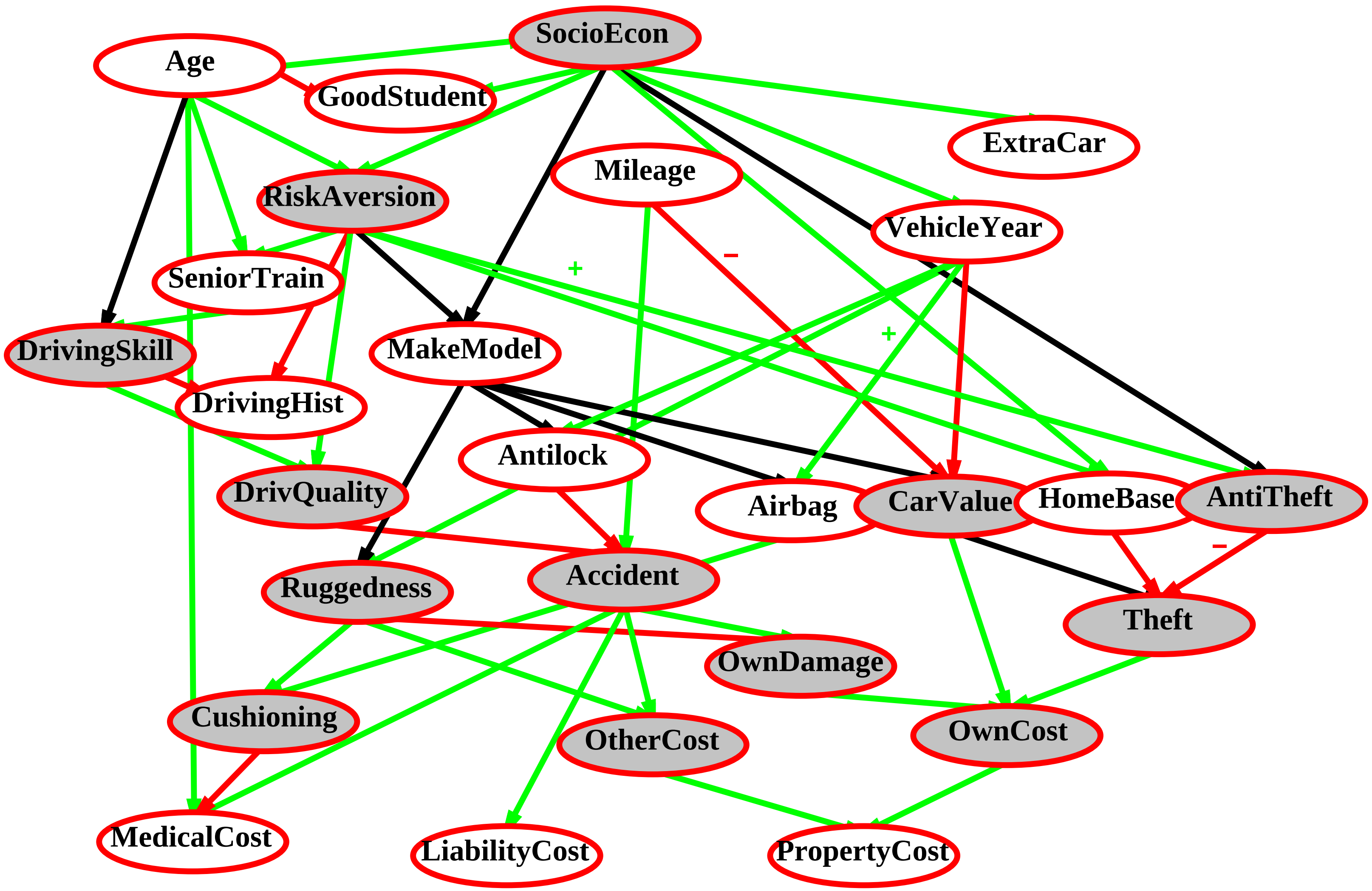
Label the arcs + or −



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., $\langle Noise, Cost, Safety \rangle$:
⟨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

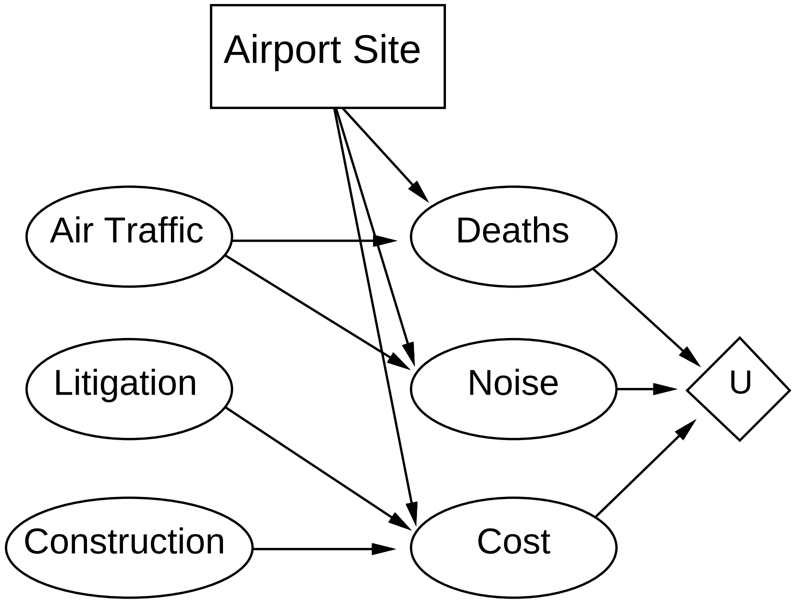
$\Rightarrow$ $\exists$ multiplicative utility function:

$$U = k_1U_1 + k_2U_2 + k_3U_3 + k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_3k_1U_3U_1 + k_1k_2k_3U_1U_2U_3$$

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

Decision networks

Add action nodes (rectangles) and utility nodes (diamonds) to belief networks to enable rational decision making



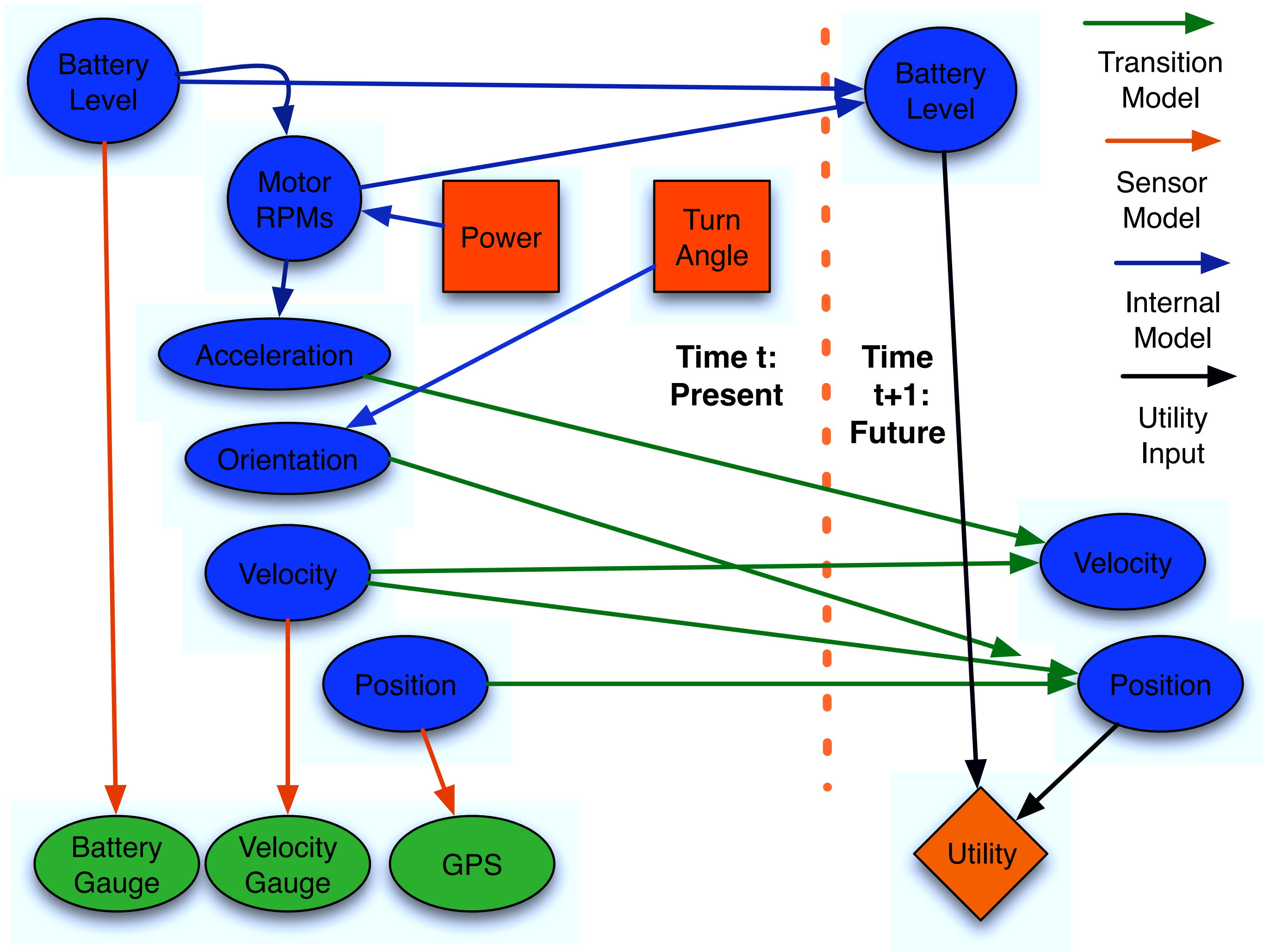
Algorithm:

For each value of action node

compute expected value of utility node given action, evidence

Return MEU action

Robot Decision Network



Decision Network Evaluation

1. Set evidence variables for the current state of the system.
2. For each vector of value assignments for the DECISION variables (e.g. Steering and Power):
   (a) Set the decision nodes to those values.
   (b) Use Bayesian Inference (either exact or approximate) to compute the posterior probabilities of all parents of the utility node.
   (c) Calculate the utility based on the probability distribution over the parent values.
3. Return the decision vector with the highest utility.

Robot Decision Making

1. Assign known values to evidence vars:
   Battery Gauge, Velocity Gauge and GPS (all at time t).
2. For each possible combination of power and turn-angle:
   (a) Use Bayesian Inference to calculate the probability distributions for Battery Level and Position at time t+1.
   (b) Then use those distributions to calculate utility.
3. Perform the power-turn combo with the highest utility

Note: Using the utility of the PROJECTED FUTURE values of Battery Level and Position as the basis for our current action choices.

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?

Oily Answers

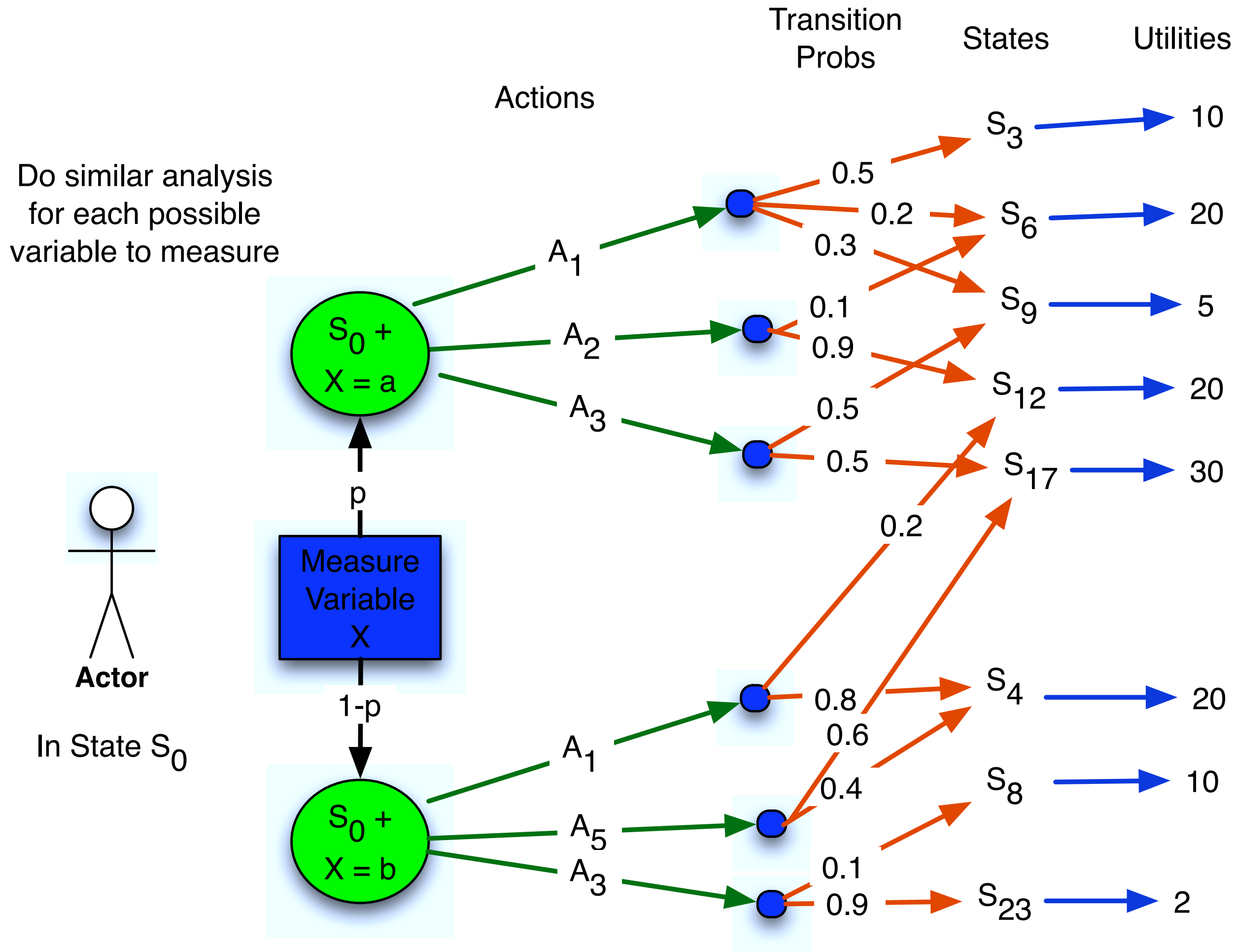
- Solution: compute expected value of information
  = expected value of best action, given the information
  minus expected value of best action without information

- The actions are Buy(A) and Buy(B).
- Without info, the expected value of either action is:
  $-\frac{k}{2} + (0.5)(k) + (0.5)(0) = 0$
- Possible info = (Oil(A), ¬Oil(A)), both with prob = 0.5
- With info, the expected value of the best action is:
  [$0.5 \times$ value of Buy (A) given Oil(A)
  $+ 0.5 \times$ value of Buy(B) given ¬Oil(A)]
  $= (0.5 \times k/2) + (0.5 \times k/2) = k/2$
- So Expected Value of Info $= \frac{k}{2} - 0 = \frac{k}{2}$

Measurement Selection

- Information is not always free.
- Taking measurements can be costly: medical tests, deep-sea chemical analyses, etc.
- Need to prioritize measurements/tests.
- Choose those that will lead to states with the highest expected utility.
- Measurements $\rightarrow$ Information States $\rightarrow$ Action Choices $\rightarrow$ New General States (Info + Physical) $\rightarrow$ Utility Assessment
- Since many of these links are stochastic, we deal with probability distributions across the results of both measurement and regular actions.

Measurement Selection (2)



General formula

Current evidence E, current best action α
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
$\Rightarrow$ 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
- It's the expected improvement in our action-assessment abilities given the new information (the value of E_j).

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

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

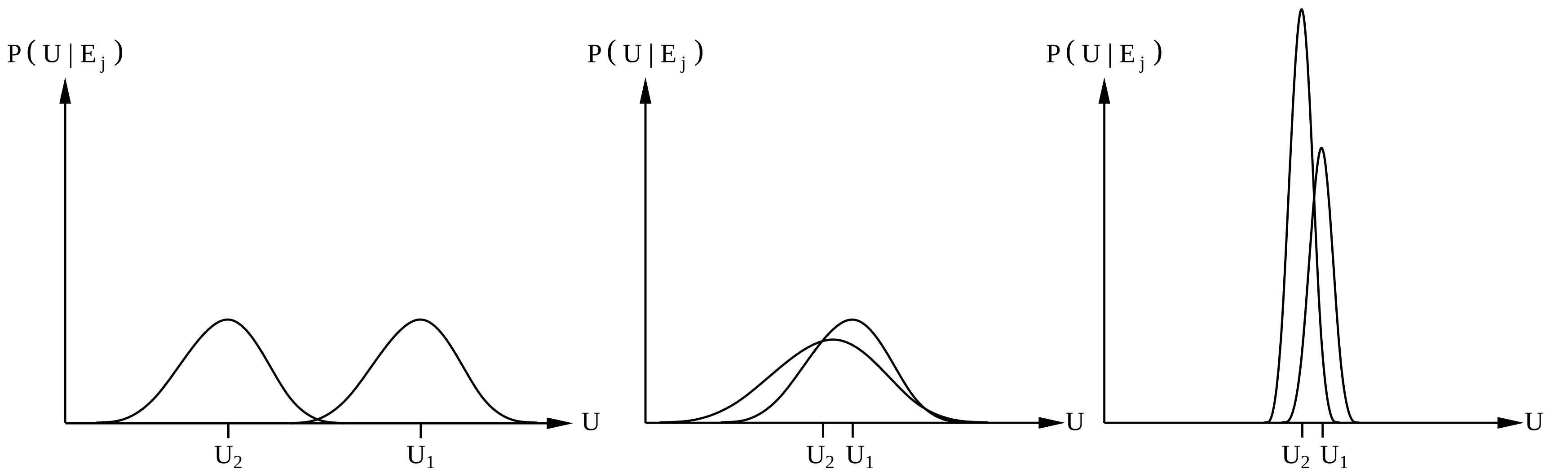
Order-independent

$$VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,E_j}(E_k) = VPI_E(E_k) + VPI_{E,E_k}(E_j)$$

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 (due to high U variance)
c) Choice is nonobvious, information worth little (due to low U variance)



(a) (b) (c)