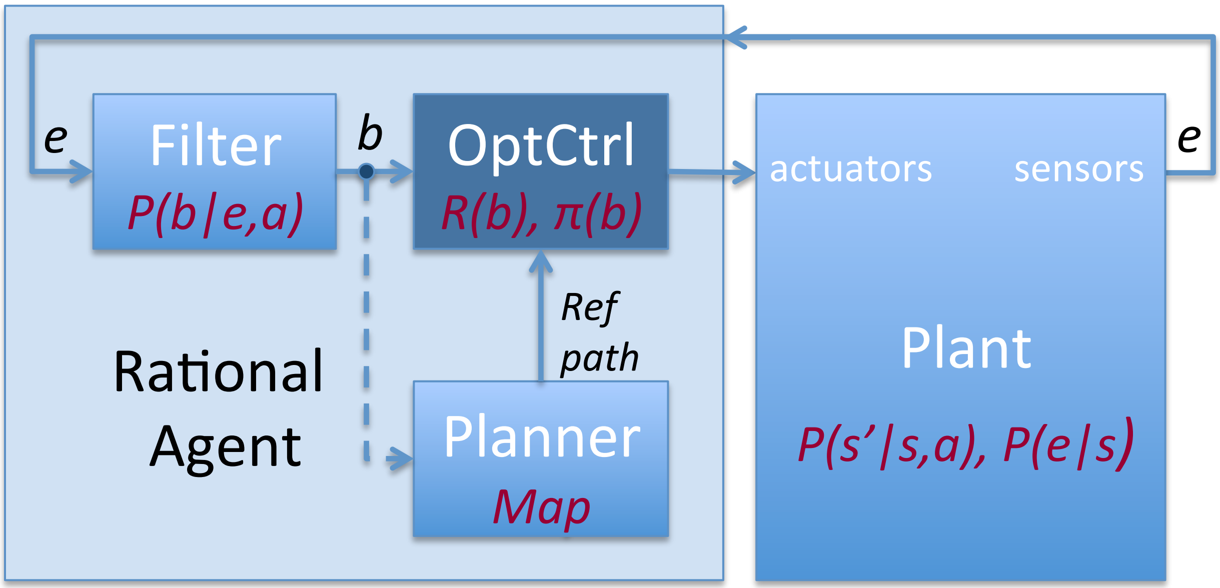
Making Simple Decisions

Chapter 16



Outline

- Beliefs and desires under uncertainty
- Rational preferences
- Utilities
- Money
- Multi-attribute utilities
- Decision networks
- Value of information

Beliefs/Desires under Uncertainty

Decision-theoretic agent:

- Rational decisions: Based on what it believes and what it wants
- Decides: In contexts with uncertainty and conflicting goals

Decision theory:

- Choose action based on desirability of immediate outcomes
- Environment: Stochastic, partially obs. $P(\mathrm{Result}(a) = s' \mid a, \mathbf{e})$

Utility function U(s): No. expressing desirability of s

Expected utility of an action given evidence EU(a|e):

$$EU(a \mid \mathbf{e}) = \sum_{s'} P(\mathrm{Result}(a) = s' \mid a, \mathbf{e})\, U(s')$$

Principle of maximum expected utility MEU(a|e):

$$\mathrm{action} = \arg\max_a EU(a \mid \mathbf{e})$$

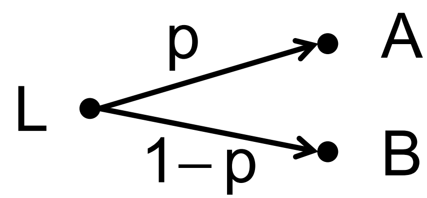
- MEU priciple can be seen as defining all of AI
- Operationalization: perception, learning, inference, planing

Preferences

An agent chooses among:

- Prizes (states): A, B, C, etc.
- Lotteries: Set of outcomes for each action

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



Notation:

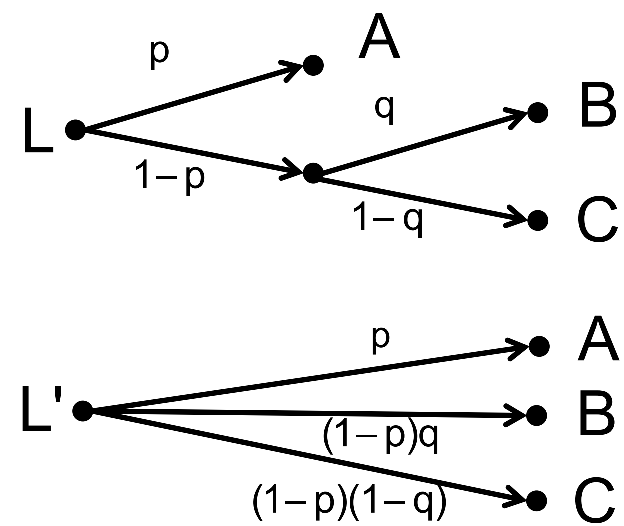
- $A \succ B$ Lottery A is preferred to lottery B
- $A \sim B$ indiference between A and B
- $A \succsim B$ B is not preferred to A

Utility-Theory Axioms

Preferences of a rational agent must obey constraints:

- Rational preferences: Maximization of expected utility

Constraints on lotteries:



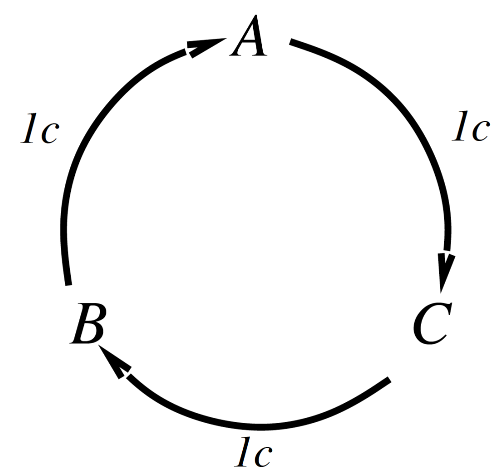
- Orderability: $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
- Transitivity: $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
- Continuity: $A \succ B \succ C \Rightarrow$
  $\exists p.\ [p,A;\ (1-p),C] \sim B$
- Substitutivity: $A \sim B \Rightarrow$
  $[p,A;\ (1-p),C] \sim [p,B;\ (1-p),C]$
- Monotonicity: $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p,A;\ (1-p),B] \succ [q,A;\ (1-q),B]$
- Decomposability: $[p,A;\ (1-p),[q,B;\ (1-q)C]] \sim$
  $[p,A;\ (1-p)q,B;\ (1-p)(1-q)C]$

Rational Preferences Continued

Violating the constraints leads to self evident irrationality

Example: An agent with intransitive preferences

- Can be induced to: give away all its money
- If $B \succ C$ then: An agent that has C would pay, say 1 cent, to get B
- If $A \succ B$ then: An agent that has B would pay, say 1 cent, to get A
- If $C \succ A$ then: An agent that has A would pay, say 1 cent, to get C



Maximizing Expected Utility

Theorem (Ramsey'31, von Neumann/Morgenstern'44):

Given preferences satisfying the constraints
there exists a real-valued function U such that:

- $U(A) > U(B) \;\Leftrightarrow\; A \succ B, \quad U(A) = U(B) \;\Leftrightarrow\; A \sim B$
- $U([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i U(S_i)$

MEU principle:

- Choose the action that maximizes the expected utility
- Note: An agent can be entirely rational (consistent with MEU)
- Without representing/manipulating utilities and probabilities
- Example: A lookup table for perfect tic-tac-toe

Utilities and Preference Elicitation

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 probabilty p
- Worst possible catastrophe $u_\perp$ with probabilty $(1-p)$
- Adjust lotery probability p until $A \sim L_p = [p, u_\top;\ (1\text{-}p), u_\perp]$

$$\text{pay } \$30 \quad \sim \quad L \begin{cases} \xrightarrow{0.999999} u_\top : \text{continue as before} \\ \xrightarrow{0.000001} u_\perp : \text{instant death} \end{cases}$$

Utility Scales

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

Micromorts: one-millionth chance of death

- Useful for Russian roulette
- Driving a car for 230 miles: a risk of one micromort

QALY: quality-adjusted life years

- Useful for medical decisions involving substantial risk

Behavior is invariant wrt. positive linear transformations:

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

With deterministic prizes only (no lottery choices), only:

- Ordinal utility can be determined, i.e., total order on prizes

Money: Obvious Utility Candidate

Money does not behave as a utility function

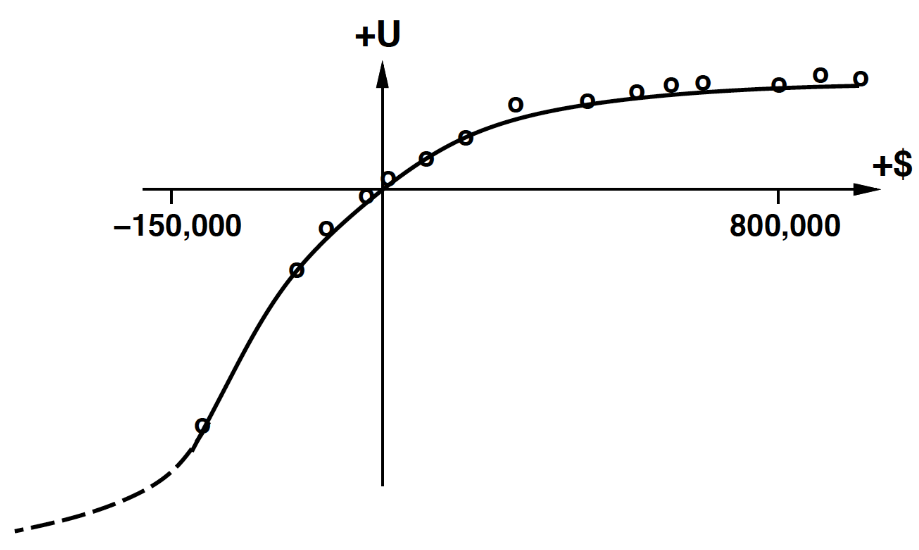
Given a lottery L with expected monetary value EMV(L)

- Usually for $U(L) < U(EMV(L))$ i.e., people are risk averse

Utility curve: For what probability p am I indifferent

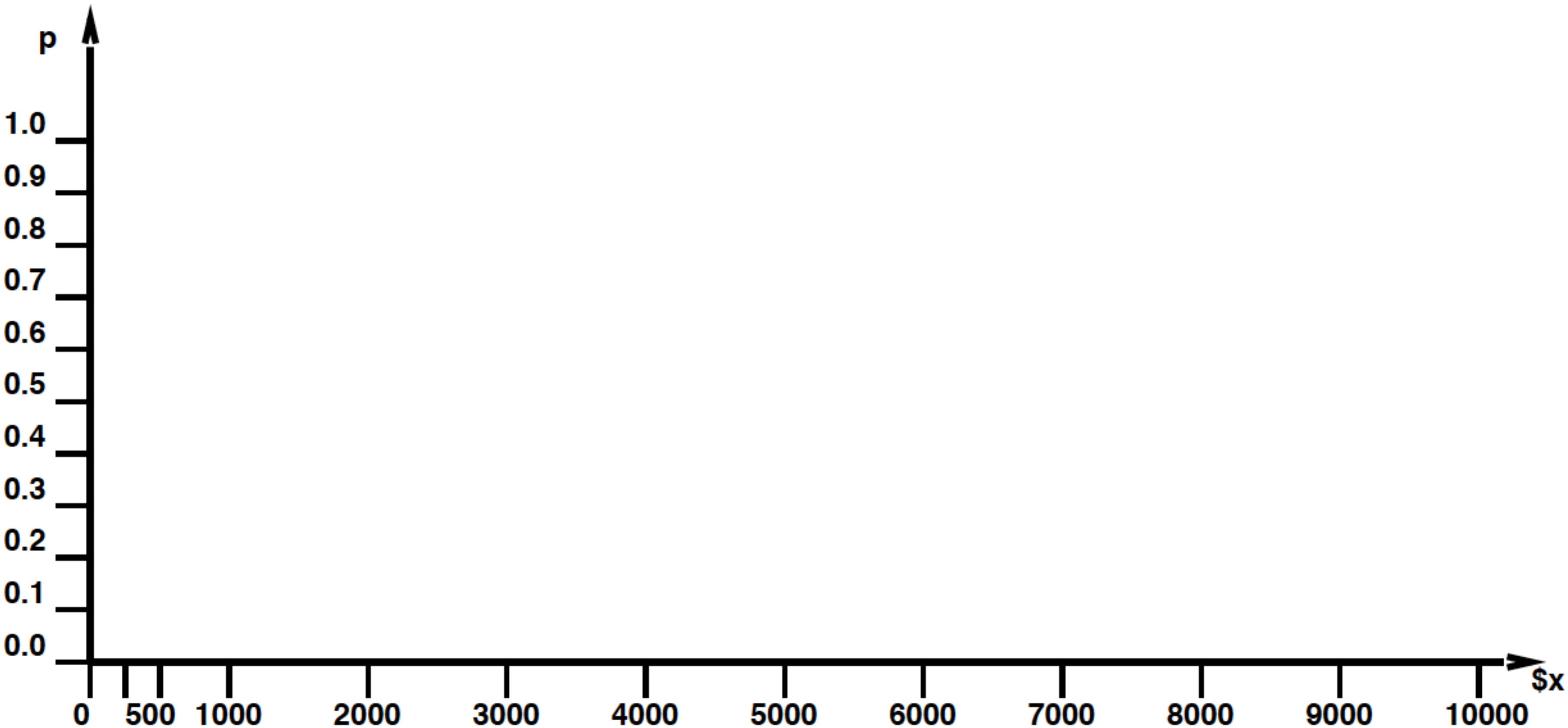
- Between prize x and lottery [p,$M; (1-p),$0] for large M?

Typical empirical data, extrapolated with risk-prone behavior:



Student-Group Utility

$\forall x$, adjust p, until half the class votes for lottery ($M = 10K$):



Multi-Attribute Utility

How can we handle utility functions of vars $X_1,\ldots,X_n$?

- Example: What is U(Deaths, Noise, Cost)?

How can complex utility functions:

- Be assessed from preference behaviour?

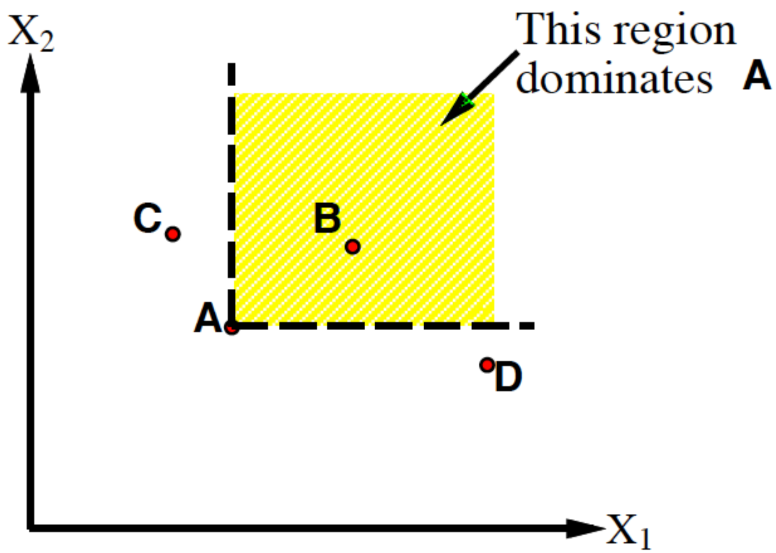
- Identify conditions under which decisions can be made:
  - Without complete identification of $U(x_1,\ldots,x_n)$
- Identify various types of independence in preferences, and:
  - Derive consequent canonical forms for $U(x_1,\ldots,x_n)$

Strict Dominance

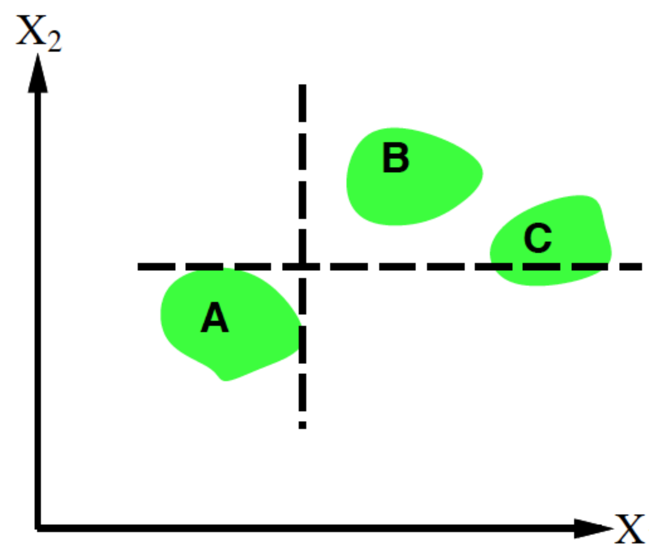
Typically define attributes such that U is monotonic in each

Strict dominance: choice B strictly dominates choice A iff

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



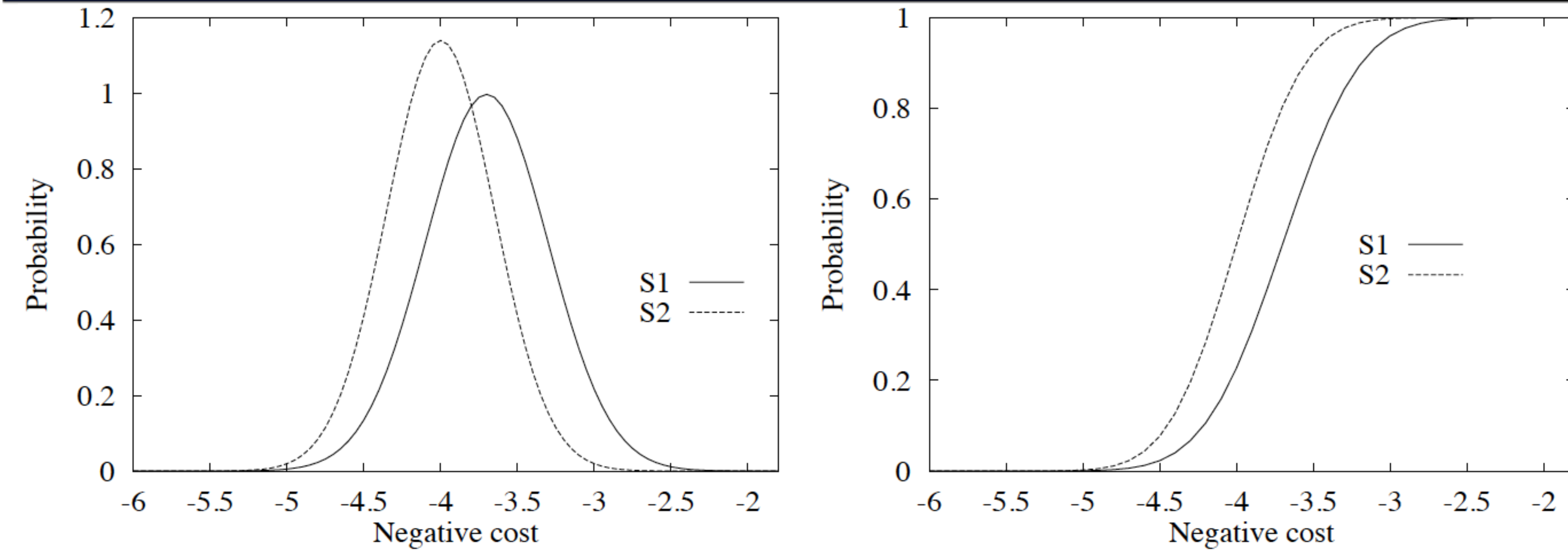
deterministic attributes



uncertain attributes

Strict dominance: seldom in practice

Stochastic Dominance



Distribution p_1 stochastically dominates distribution p_2 iff:

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

If U is monotonic in x: A_1 with pdf p_1 SD A_2 with pdf p_2

- $\int_{-\infty}^{+\infty} p_1(x)\,U(x)\,dx \leq \int_{-\infty}^{+\infty} p_2(x)\,U(x)\,dx$

Multi-attribute case: SD on all attributes $\Rightarrow$ optimality

Stochastic Dominance

Stochastic dominance can often be determined without:

- Exact distributions using qualitative reasoning

Example: construction cost increases with distance from city

- S_1 is closer to the city than S_2 implies
- S_1 stochastically dominates S_2 on cost

Example: injury cost increases with collision speed

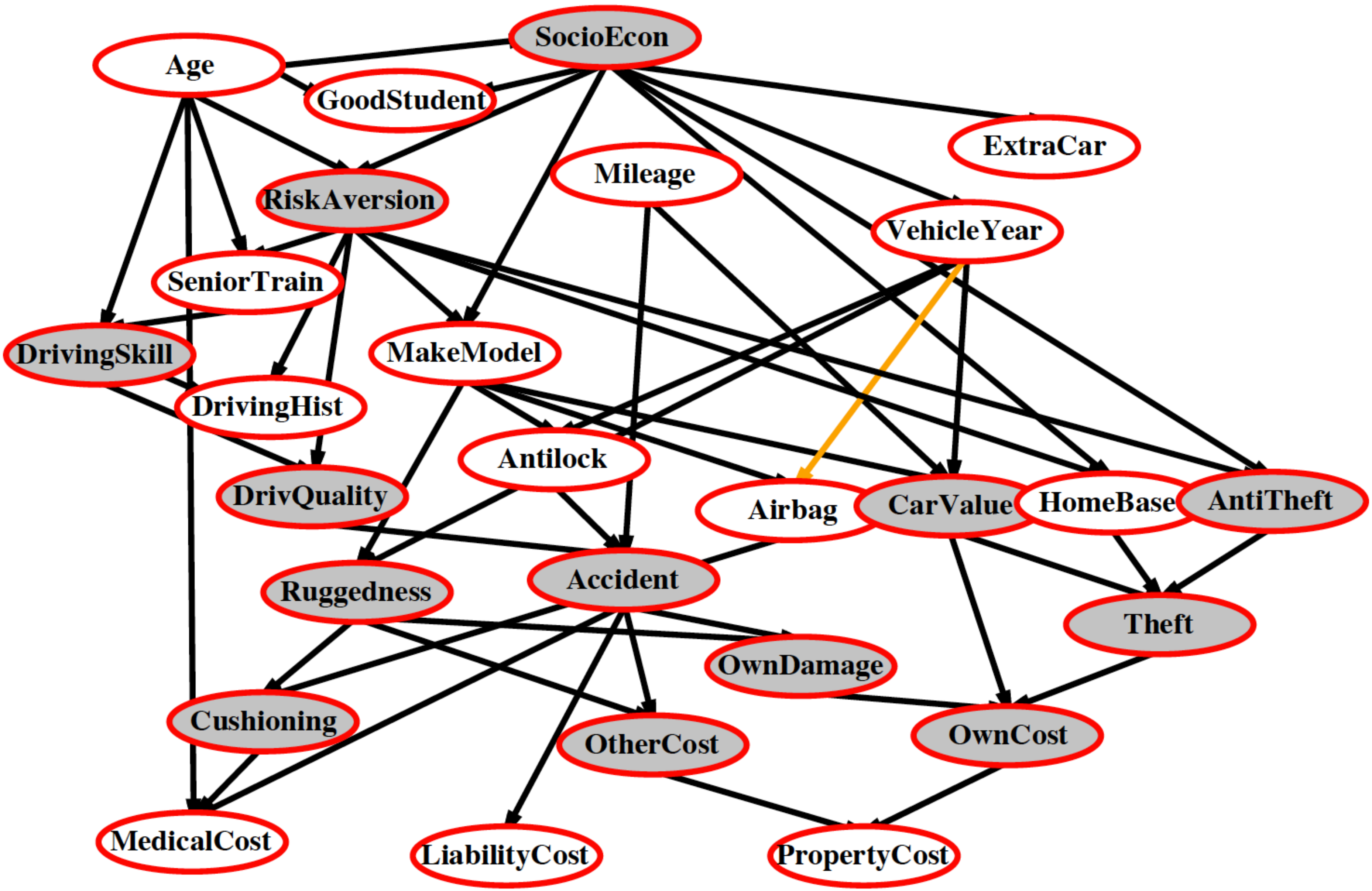
Can annotate belief networks with SD information:

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

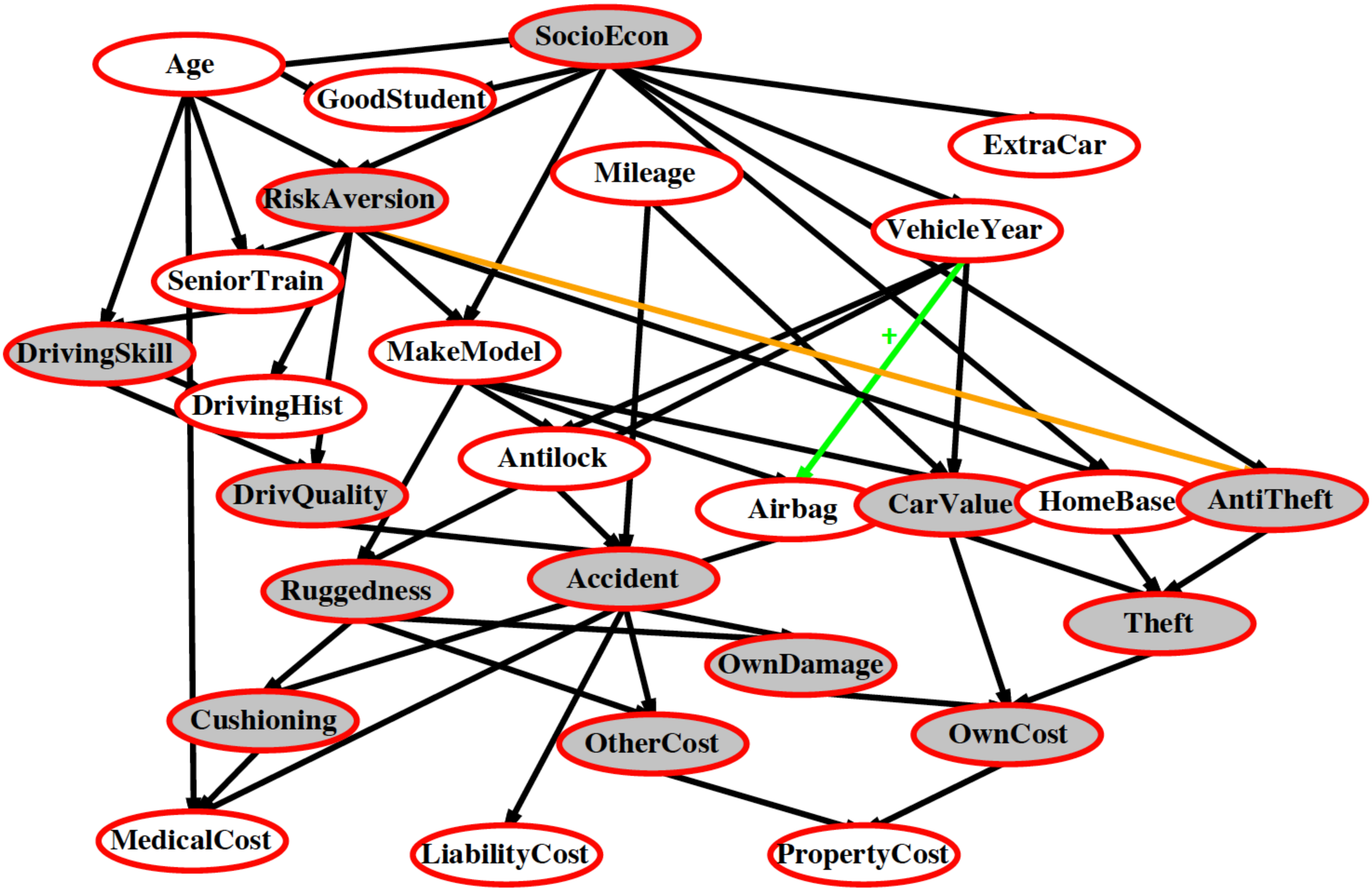
$$\forall z \in (Y\text{'s other parents } \mathbf{Z}).\ \forall x_1, x_2.$$

$$x_1 \geq x_2 \Rightarrow P(Y \mid x_1, \mathbf{z}) \text{ SD } P(Y \mid x_2, \mathbf{z})$$

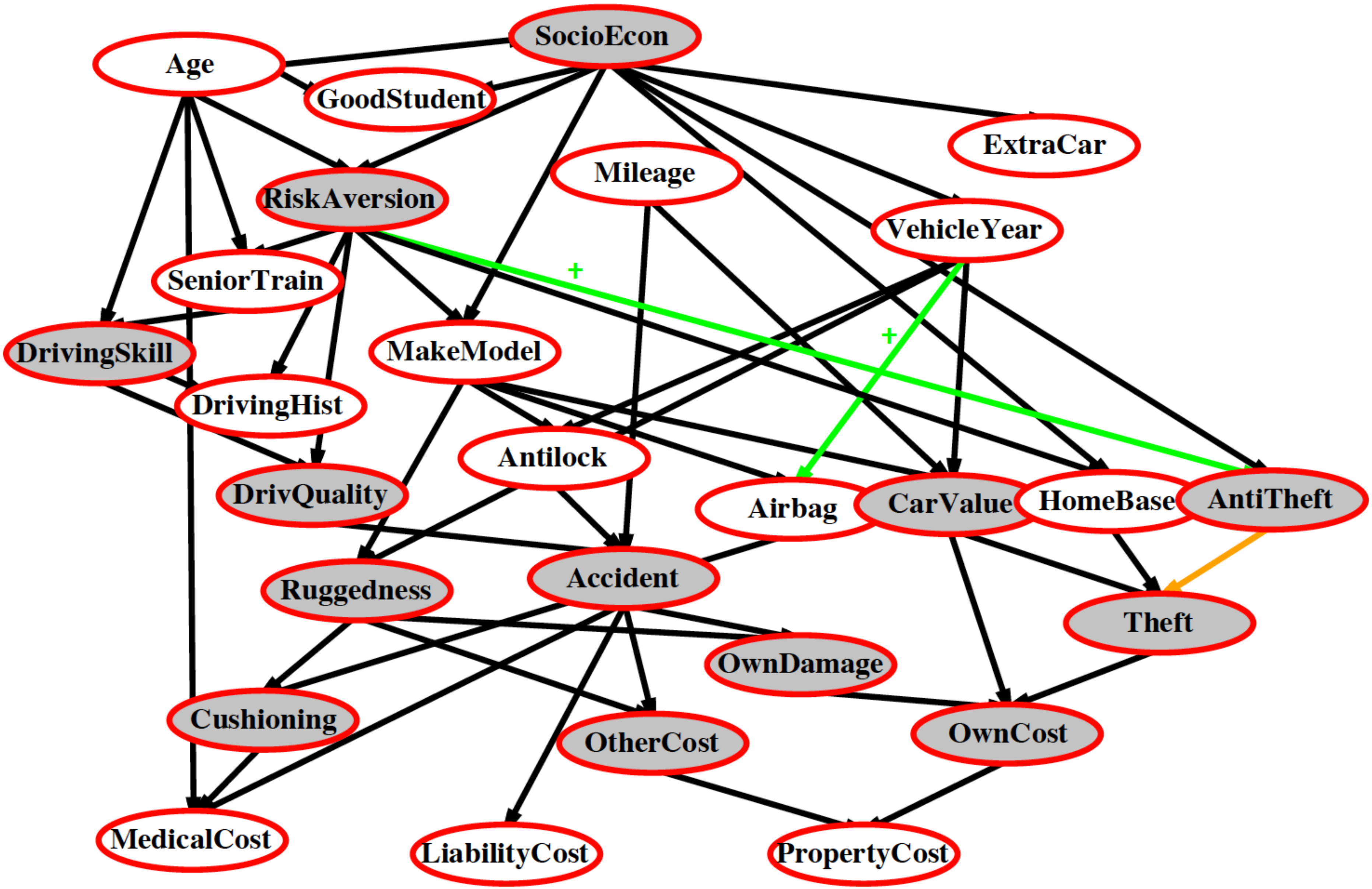
Label the Arcs + or -



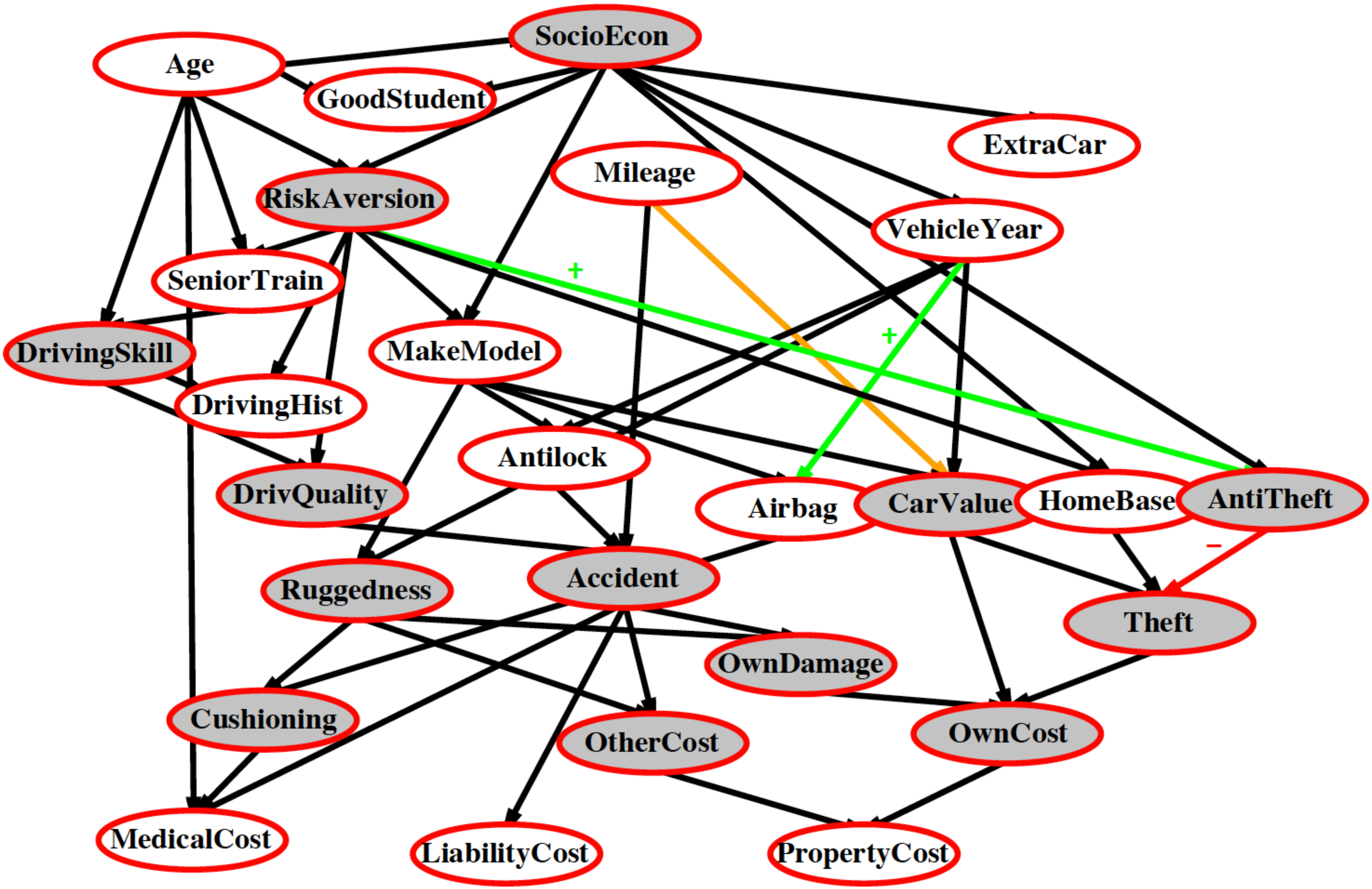
Label the Arcs + or -



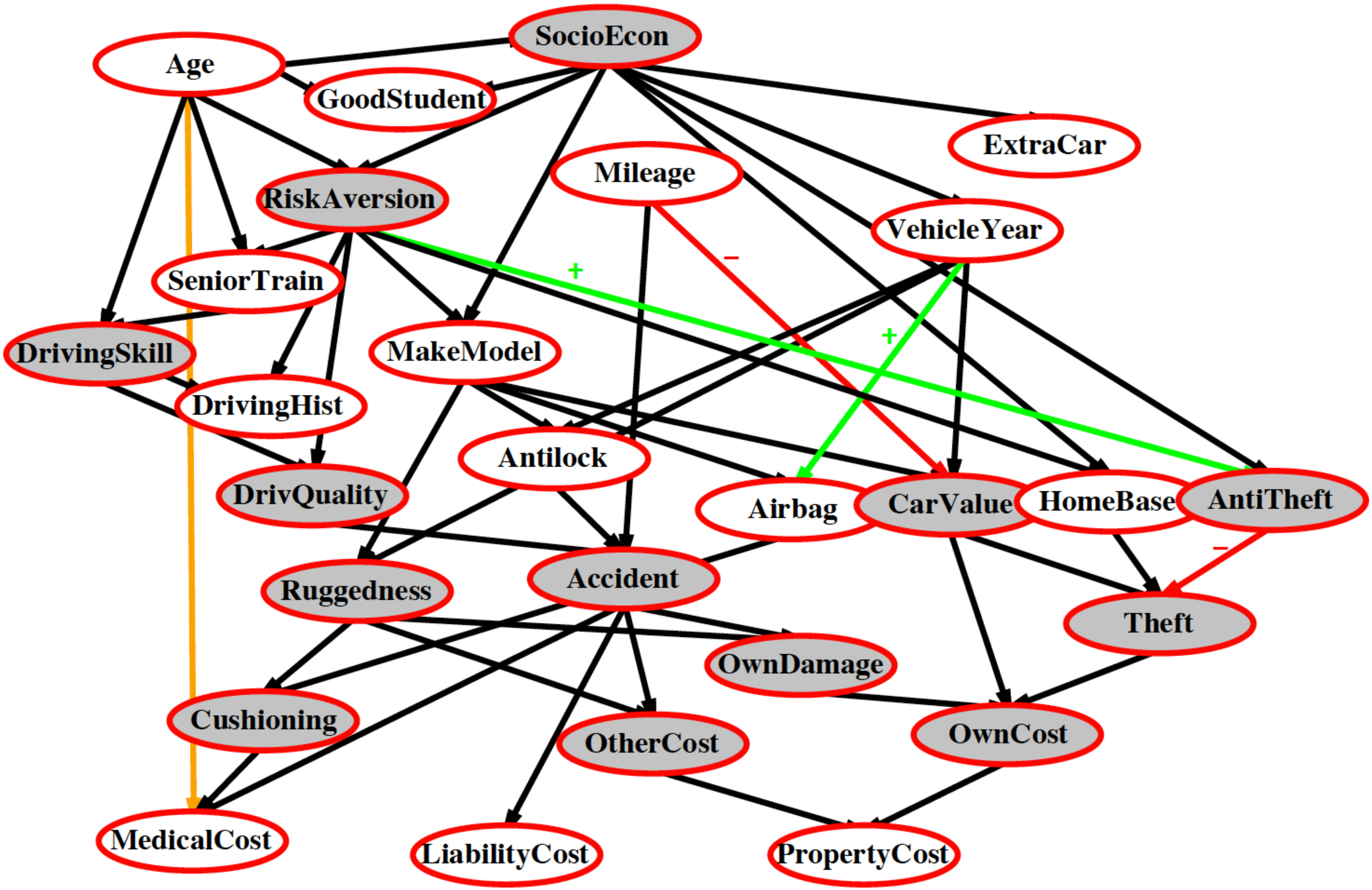
Label the Arcs + or -



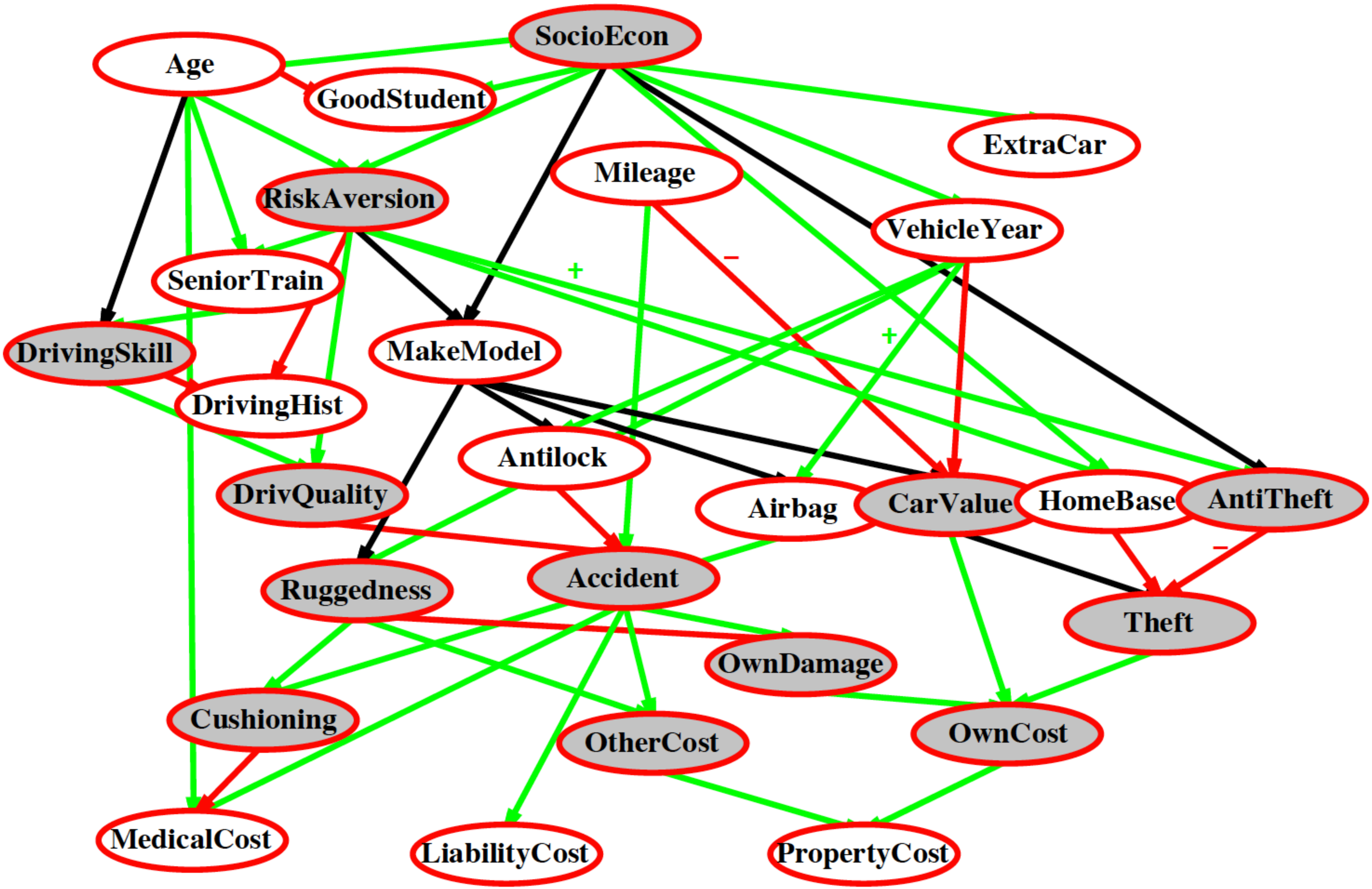
Label the Arcs + or -



Label the Arcs + or -



Label the Arcs + or -



Preference Structure: Deterministic

X_1 and X_2 are preferentially independent of X_3 iff:

- Preference between $\left\{ \begin{array}{c} \langle x_1, x_2, x_3 \rangle \\ \langle x_1', x_2', x_3 \rangle \end{array} \right\}$ does not depend on x_3

Example, ⟨Noise, Cost, Deaths⟩:

- ⟨20,000 suffer, \$4.6 billion, 0.06 deaths/mpm⟩ vs
- ⟨70,000 suffer, \$4.2 billion, 0.06 deaths/mpm⟩

Theorem (Leontief '47): Mutual P.I.

- If every pair of attributes is P.I. of its complement
- Then every subset of attributes is P.I. of its complement

Theorem (Debreu '60): Mutual P.I.

- $\exists$ additive-value-function $V(x_1,\ldots,x_n) = \sum_i V_i(x_i)$
- V_i value function for att X_i. $V(n,c,d) = -n \times 10^4 - c - d \times 10^{12}$

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

- Implies the existance of a utility function U where $U_i = U_i(x_i)$

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

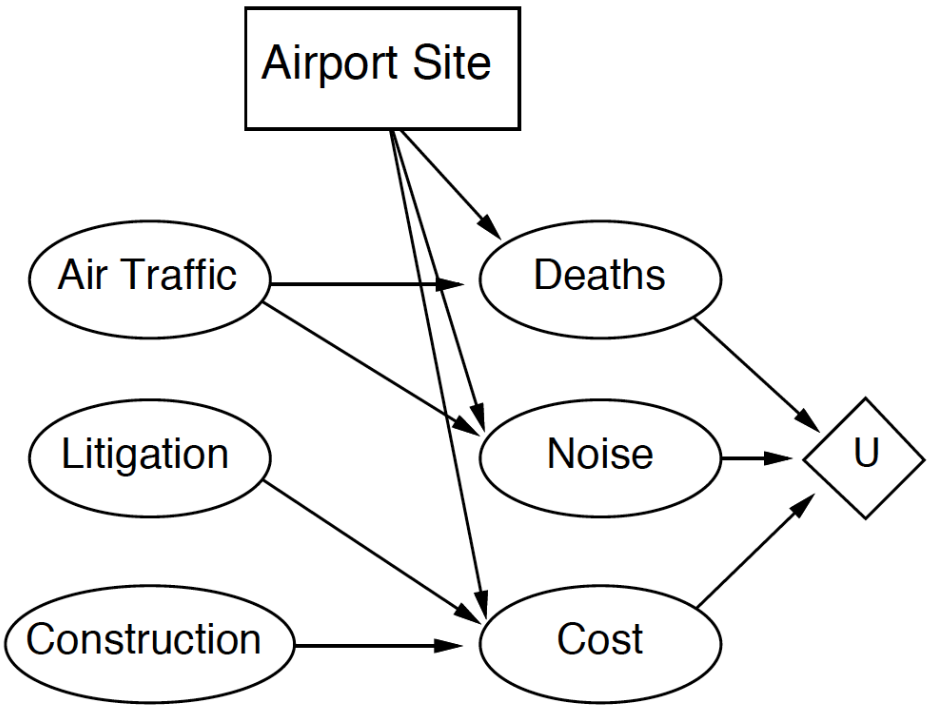
Routine procedures/software for generating preference

- Tests identifying various canonical utility-functions families

Decision Networks

Add action nodes and utility nodes to belief networks

- To enable rational decision making



Algorithm:

foreach value(action-node)

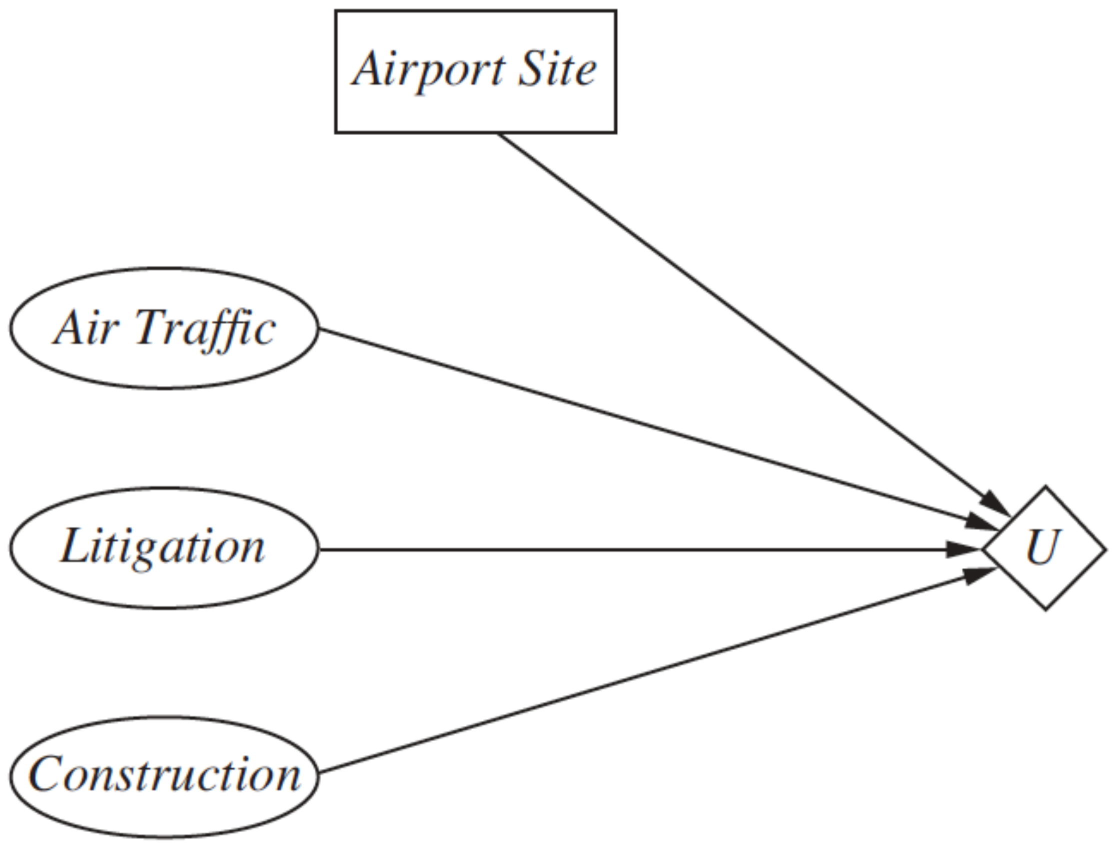
compute expected value of utility node given action, evidence

return MEU action

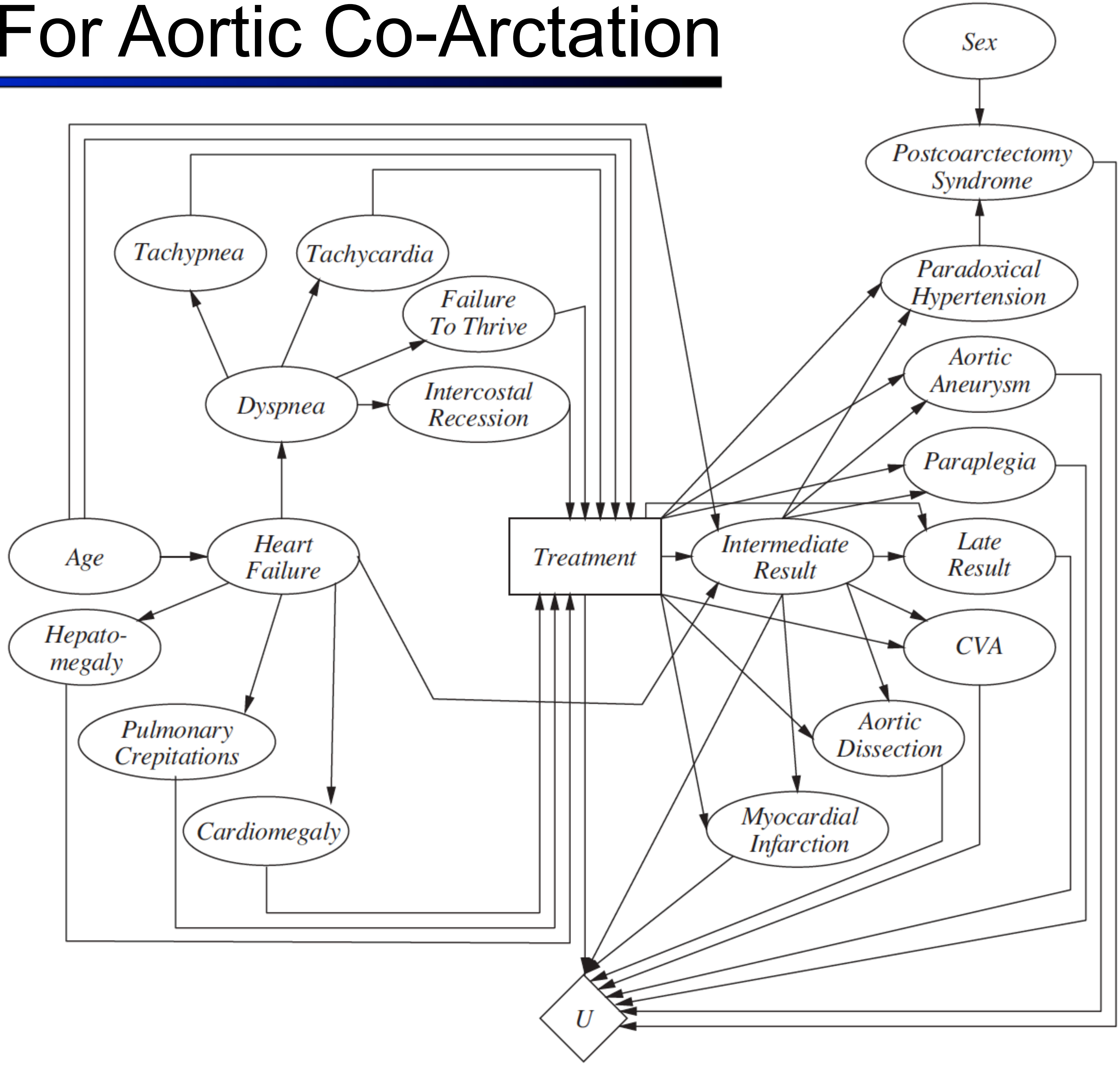
Simplified Decision Networks

Action-utility function: Chance outcome states omitted

- Q-function expected utility associated with each action



DN For Aortic Co-Arctation



The Value of Information

So far we have assumed that

- All relevant information is available to make a decision
- In practice this is hardly ever the case
- Most important is knowing what questions to ask

Example: A doctor cannot expect all investigations

- When the patient enters the consulting room
- Investigations are expensive and sometimes hazardous
- Importance: Lead to better treatment, how likely results are

Information theory

- Allows an agent to choose what information to acquire
- Prior to selecting an action agent acquires the value of it
- A simplified form of sequential decision making
- Observation actions afect only agent's belief state (not env)

Value of Information

Idea : Compute the 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 probablilities 0.5 each, mutually exclusive
- Current price of each block is k/2 each
- Consultant offers accurate survey of A. Fair price?

Solution: Compute expected value of information, that is

- Expected value of best action given the information
- Minus expected value of best action without information
- Survey may say oil in A with probability 0.5

$$[0.5 \times (\text{value of buy A given oil in A}) + 0.5 \times (\text{value of buy B given no oil in A})] - 0$$

$$[0.5 \times k / 2 + 0.5 \times k / 2] - 0 = k / 2$$

General Formula

Given :

- Current evidence E, Current best action α
- Possible action outcomes S_i, Potential new evidence E_j

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

Suppose we knew:

- $E_j = e_{jk}$, then we would choos $\alpha_{e_{jk}}$ such that

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

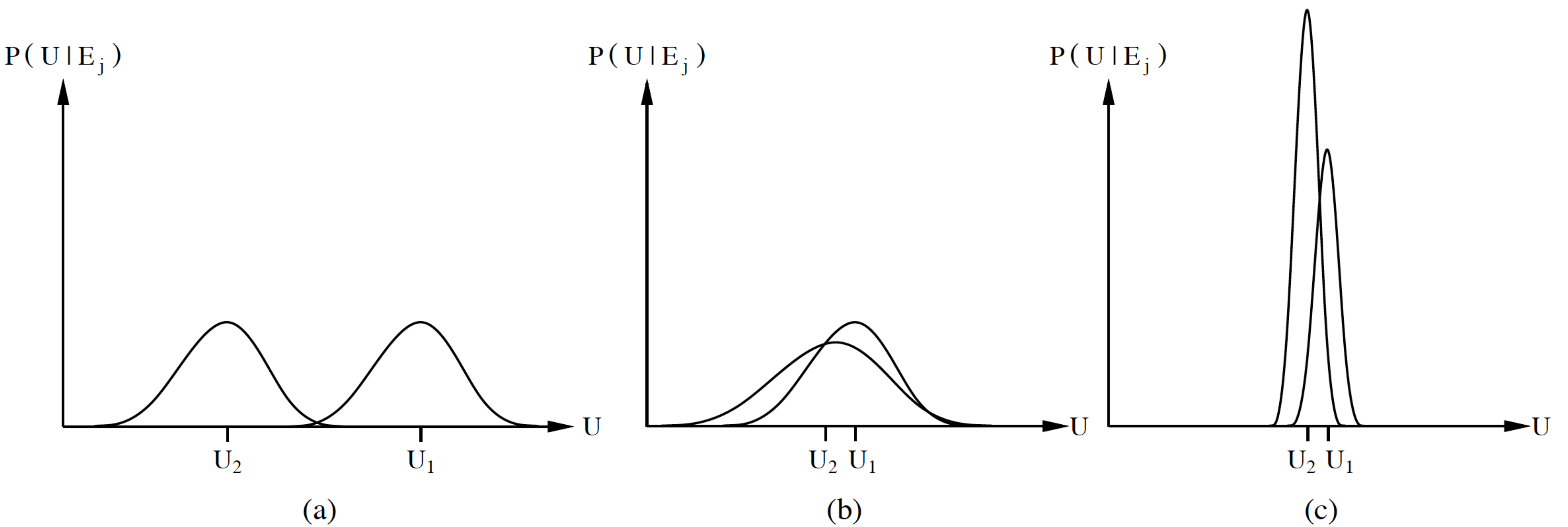
- E_j is a random variable whose value is currently unknown
- Must compute expected gain over all possible values

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

VPI $=$ Value of perfect information

Qualitative Behaviors

(a) Choice is obvious: information is worth little
(b) Choice is non-obvious: information is worth a lot
(c) Choice is non-obvious: information is worth little



Information has value: may cause the choice of better plan

Properties of VPI

Non-negative: In expectation, not post hoc

$$\forall j,E.\ VPI_E(E_j) \geq 0$$

Non-additive: consider for example, obtaining E_j twice

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

Order-independent:

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

When more than one piece of evidence can be gathered

- Maximizing VPI for each to select one is not always optimal
- Evidence gathering becomes a sequential decision problem

Information-Gathering Agent

function Information-Gathering-Agent(pecept) **return** action

persistent D // decision network

integrate percept into D

$j = \operatorname{argmax}_k VPI(E_k) / Cost(E_k)$

if $(VPI(E_j) > Cost(E_j))$

return $Request(E_j)$

return the best action from D

$Request(E_j)$: The next percept provides the value of E_j

Myopic (greedy): Calculate VPI with only one evidence variable