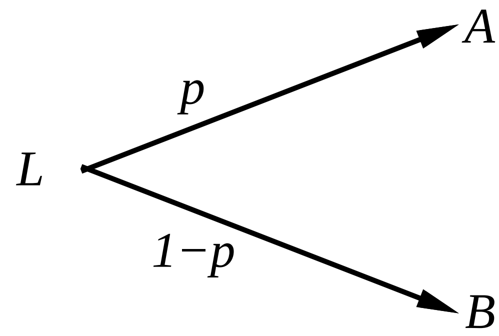
Rational decisions

Preferences

Consider an agent who chooses among alternatives (A, B, etc.), sometimes called states, outcomes, or prizes.

Agent may also choose among 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

Constraints on preferences

Preferences of a rational (sensible) agent must obey certain 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

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

Constraints on preferences

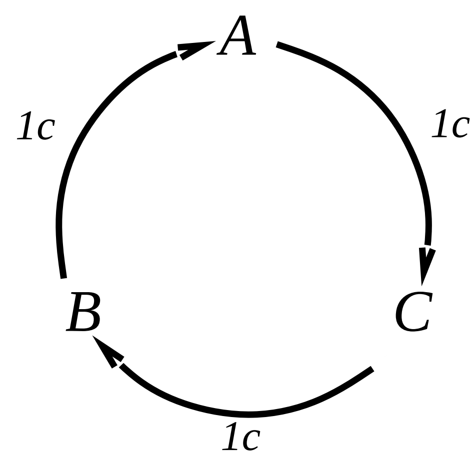
Violating the constraints leads to self-evident irrationality

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 previous constraints, there exists a real-valued function U such that

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

Choose the action that achieves the maximum expected utility (MEU).

An agent that chooses according to MEU is rational:

If the utility function reflects the performance measure by which the agent is judged, it will achieve the highest possible performance.

MEU principle still applies when environment is uncertain
(expectation is over all forms of uncertainty).

Note: An agent can be rational without ever representing or manipulating utilities and probabilities (e.g., look up table behavior for tic-tac-toe).

Utilities

Utilities map states to real numbers. Which numbers?

Standard approach to assessment of human utilities:

Define "best possible prize," having utility $u_\top$.

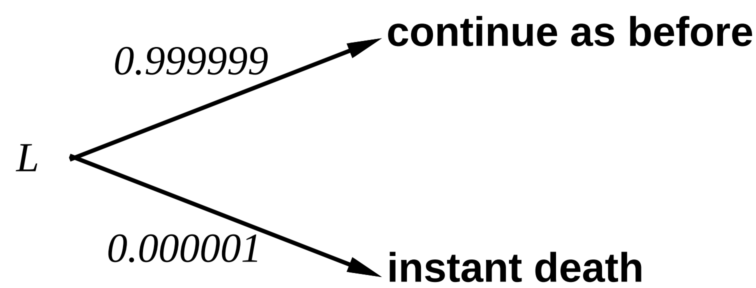
Define "worst possible catastrophe," having utility $u_\bot$.

To evaluate utility of some state A, compare it to lottery

$L = [p, u_\top; (1-p), u_\bot]$.

Adjust p until $A \sim L$, yielding $U(A) = p(u_\top - u_\bot) + u_\bot$.

pay $30 ~



Utility scales

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

Micromorts: one-millionth chance of death
useful for assessing product risks

QALYs: quality-adjusted life year (one year in good health)
useful for medical decisions involving substantial risk

Note: behavior is **invariant** with respect to positive linear transformation

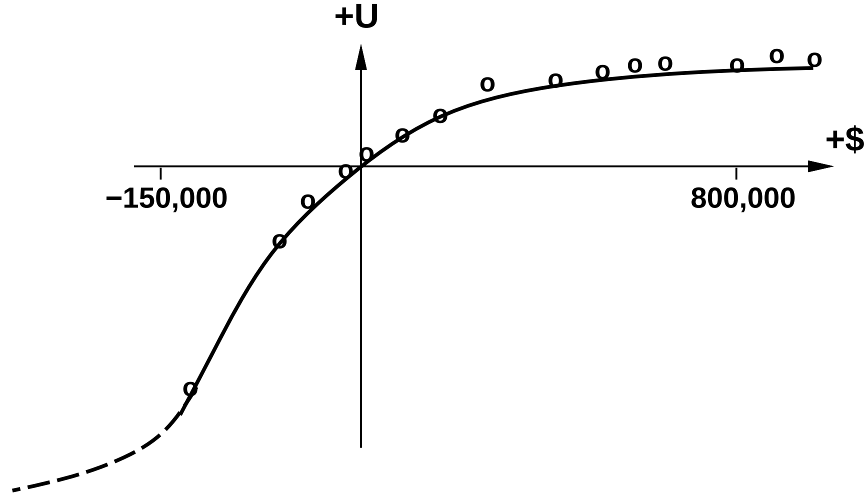
$$U'(x) = k_1 U(x) + k_2 \quad \text{where } k_1 > 0$$

Utility of money

Money does not translate directly to utility: $U(\$m) \neq m$.

E.g., would you rather have a prize of \$5000, or enter a lottery $[.5, \$10000; .5, \$0]$?

Empirical utility curve: For what p are you indifferent between prize $\$m$ and a lottery $[p, \$M;\ (1-p), \$0]$, where M is some fixed large number.

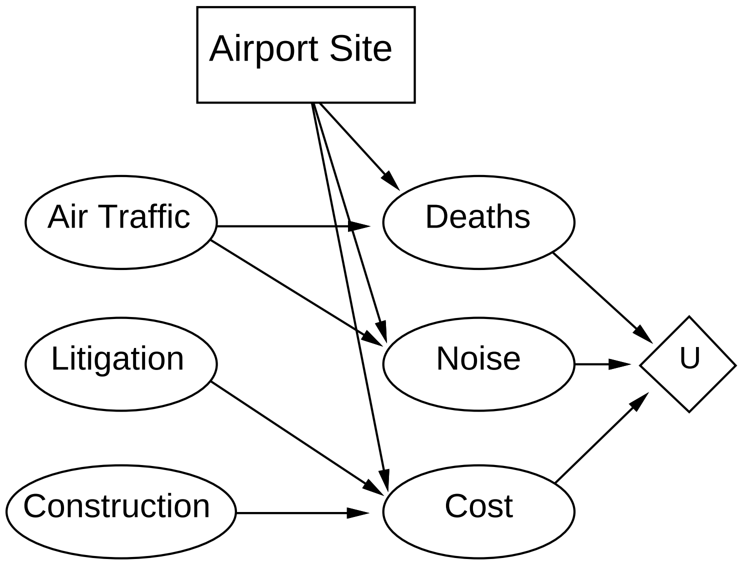


People are risk averse when it comes to gains, risk prone when it comes to losses (extrapolation).

Decision networks

Decision network = belief network + utility nodes + action nodes

Complete model for one-shot rational decision making



Algorithm:

For each value of action node

compute expected value of utility node given action and evidence

Return action yielding MEU

Multiattribute utility

How can we handle utility functions of many variables $X_1 \dots X_n$?
E.g., what is $U(Deaths, Noise, Cost)$?

Need an arbitrary look up table in the worst case, with $O(d^n)$ entries.

How can we simplify these complex utility functions?

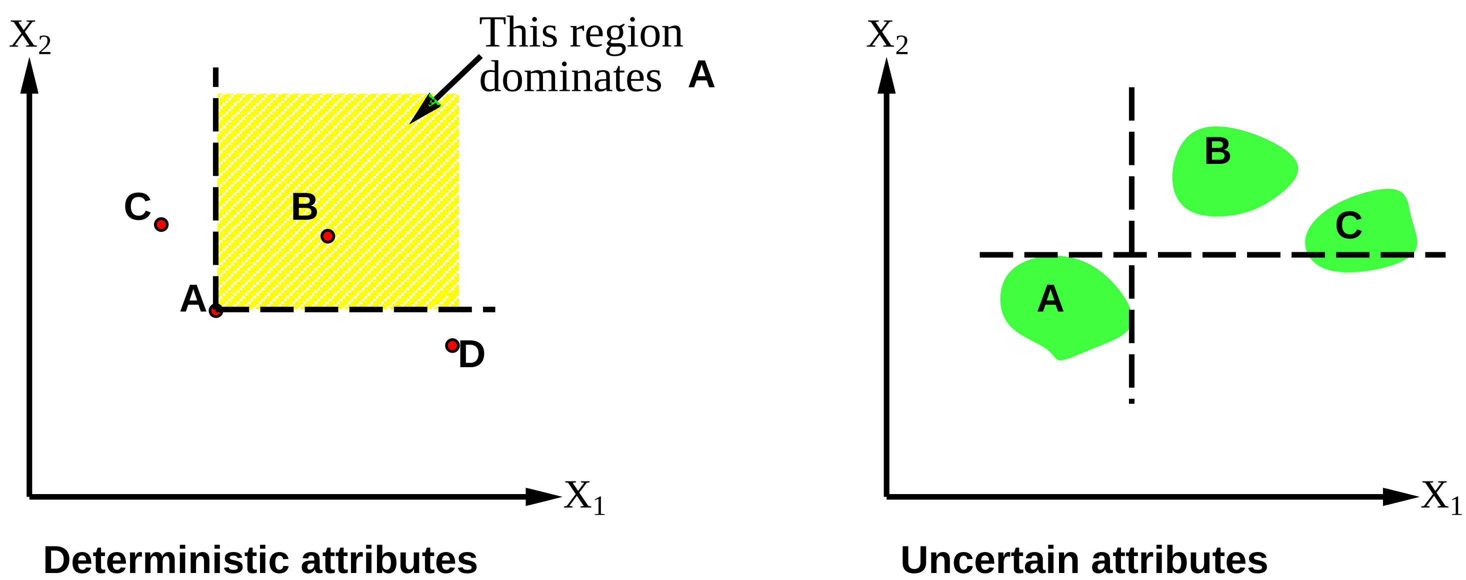
Scheme 1: identify conditions under which decisions can be made without complete identification of $U(x_1, \dots, x_n)$

Scheme 2: identify various types of **independence** in preferences and derive resulting canonical forms for $U(x_1, \dots, x_n)$

Strict dominance

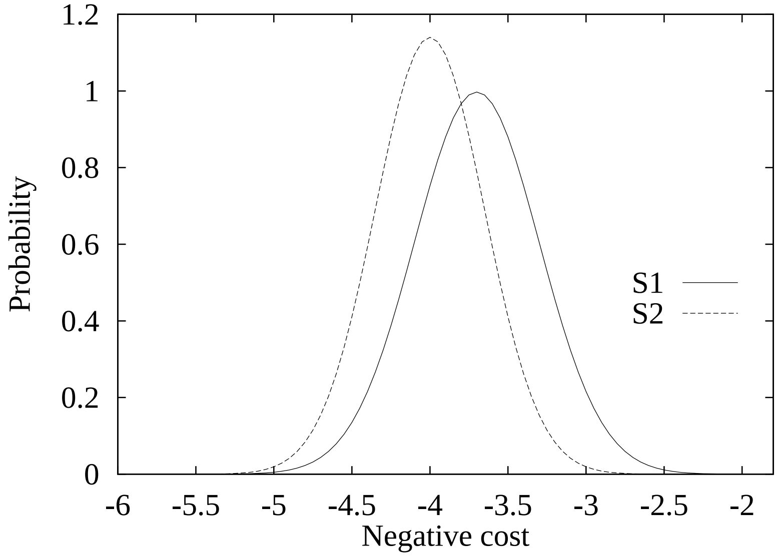
Define variables $X_1 \ldots X_n$ such that U is monotonic in each

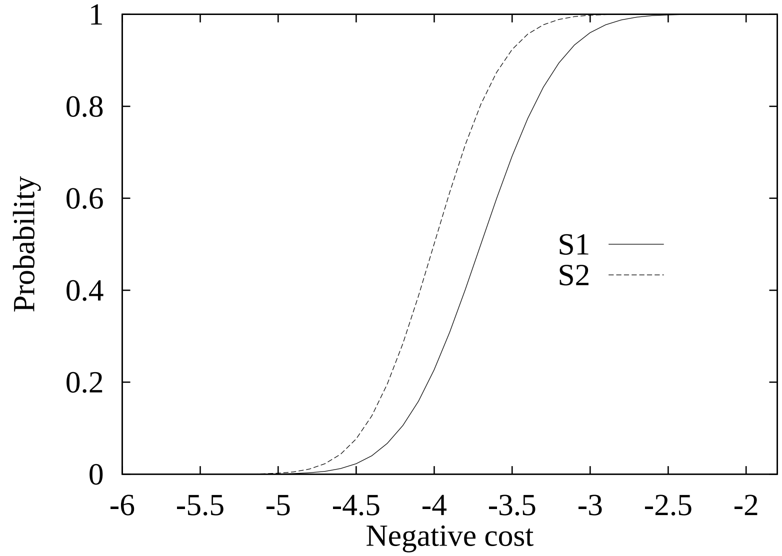
Strict dominance: choice B strictly dominates choice A iff
$\forall\, i \;\; X_i(B) \geq X_i(A)$ (and hence $U(B) \geq U(A)$)



Strict dominance seldom holds in practice

Stochastic dominance: Univariate example





Choice S_1 stochastically dominates choice S_2 if $\forall\, x$

$P(S_1 > x) > P(S_2 > x)$ or equivalently, $P(S_1 < x) < P(S_2 < x)$

If S_1 and S_2 have outcome distributions p_1 and p_2,

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

If U is monotonic in x, $\int_{-\infty}^{\infty} p_1(x)U(x)dx \geq \int_{-\infty}^{\infty} p_2(x)U(x)dx$

Stochastic dominance cont.

Multivariate case: stochastic dominance on all attributes $\Rightarrow$ optimal

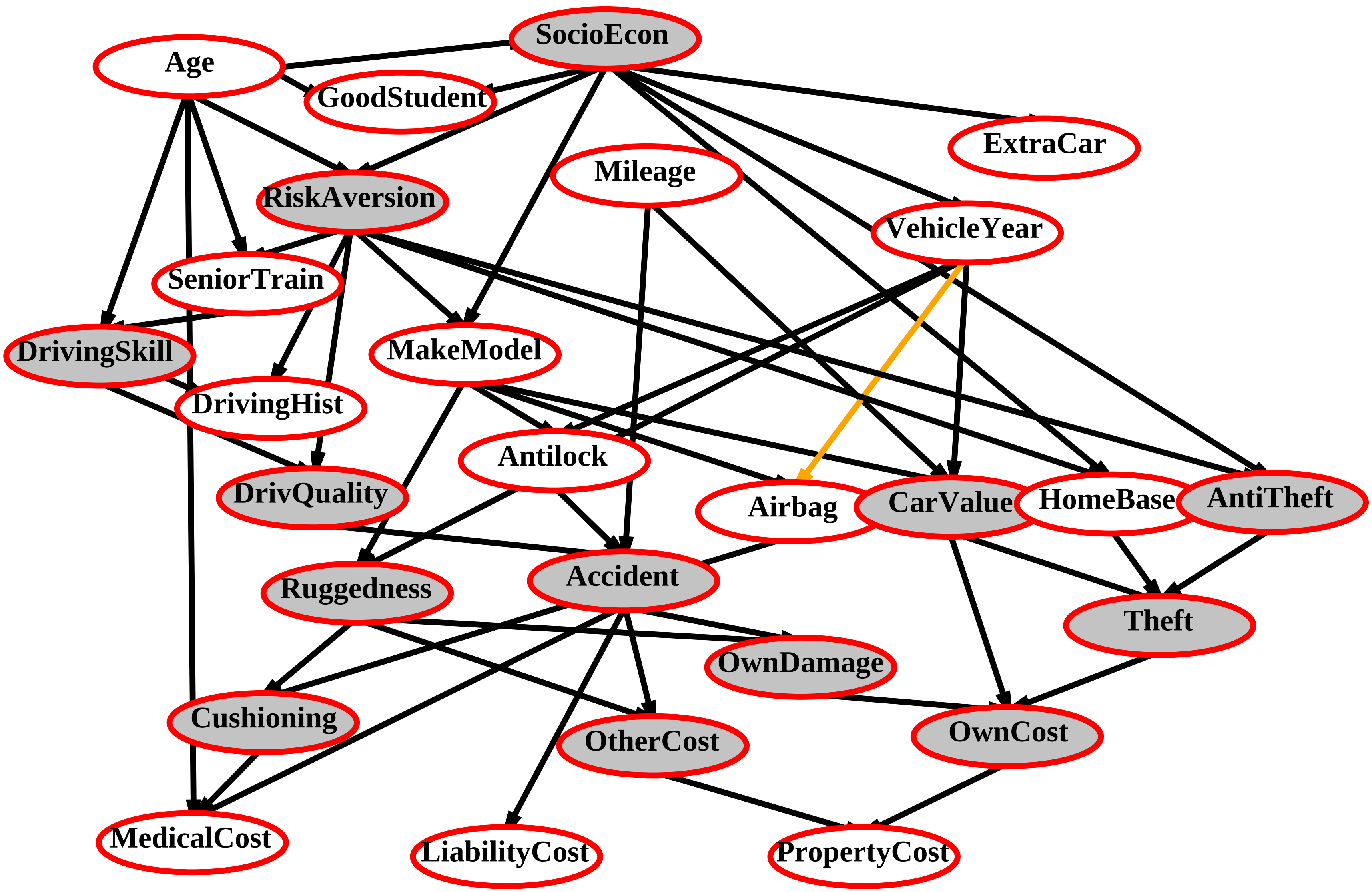
Stochastic dominance can often be determined without
exact distributions using **qualitative** reasoning

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

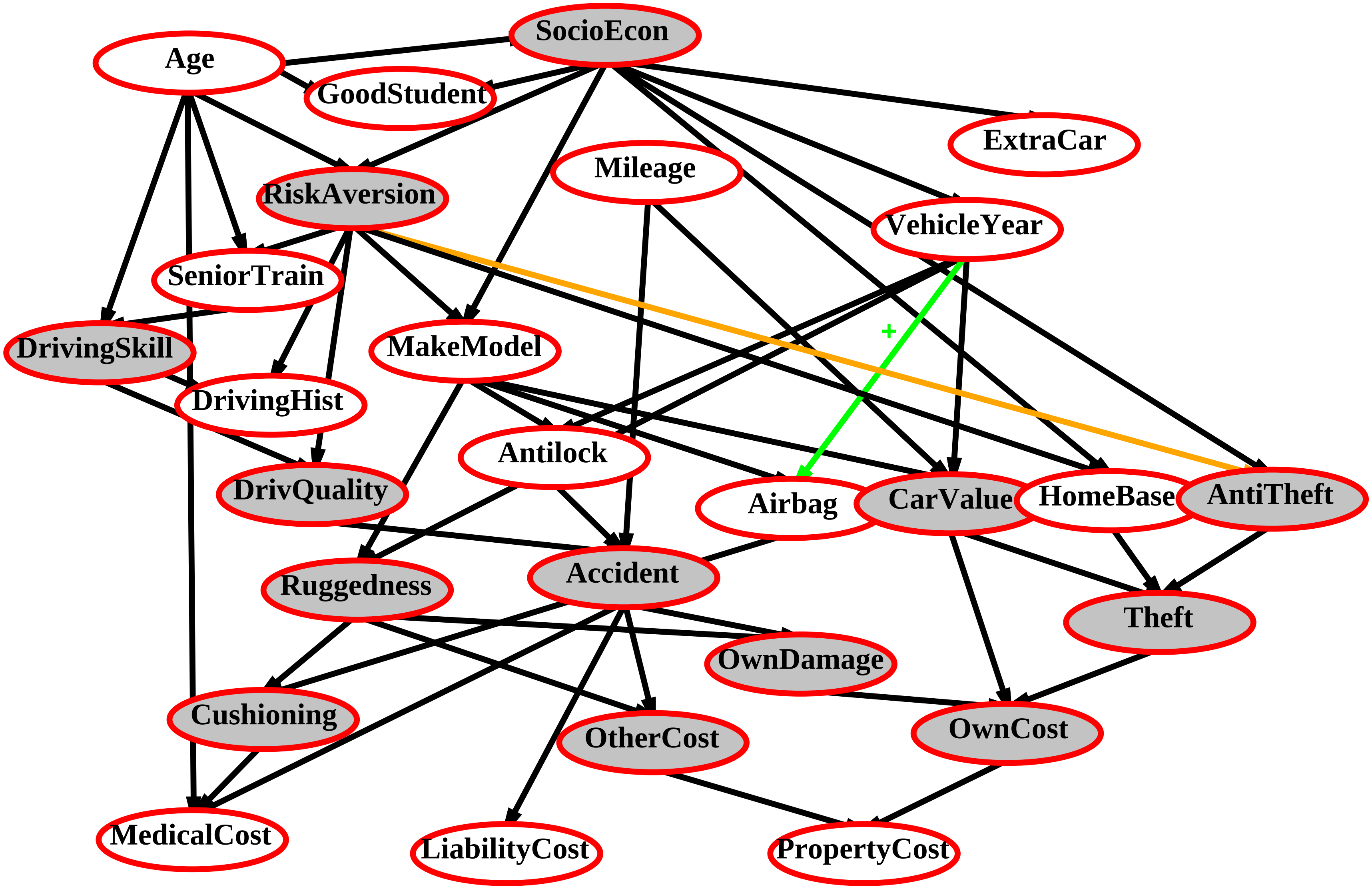
E.g., injury increases with collision speed

Can annotate belief networks with stochastic dominance information:
$X \xrightarrow{+} Y$ (X positively influences Y) means that
$\forall\, x_1, x_2\;\; x_1 \geq x_2 \Rightarrow \mathbf{P}(Y|x_1, \mathbf{z})$ stochastically dominates $\mathbf{P}(Y|x_2, \mathbf{z})$
for every value $\mathbf{z}$ of Y's other parents $\mathbf{Z}$

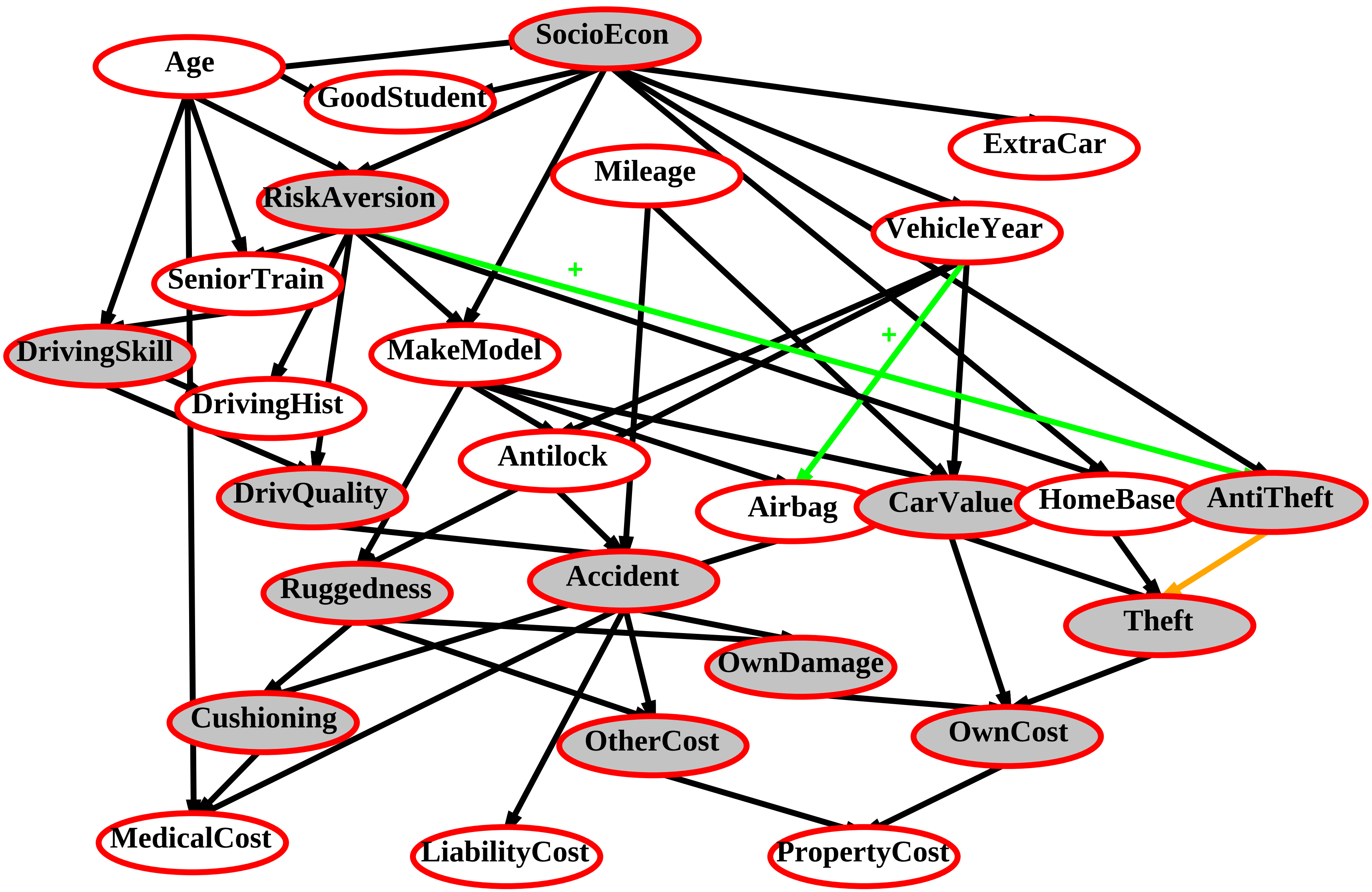
Label the arcs + or −



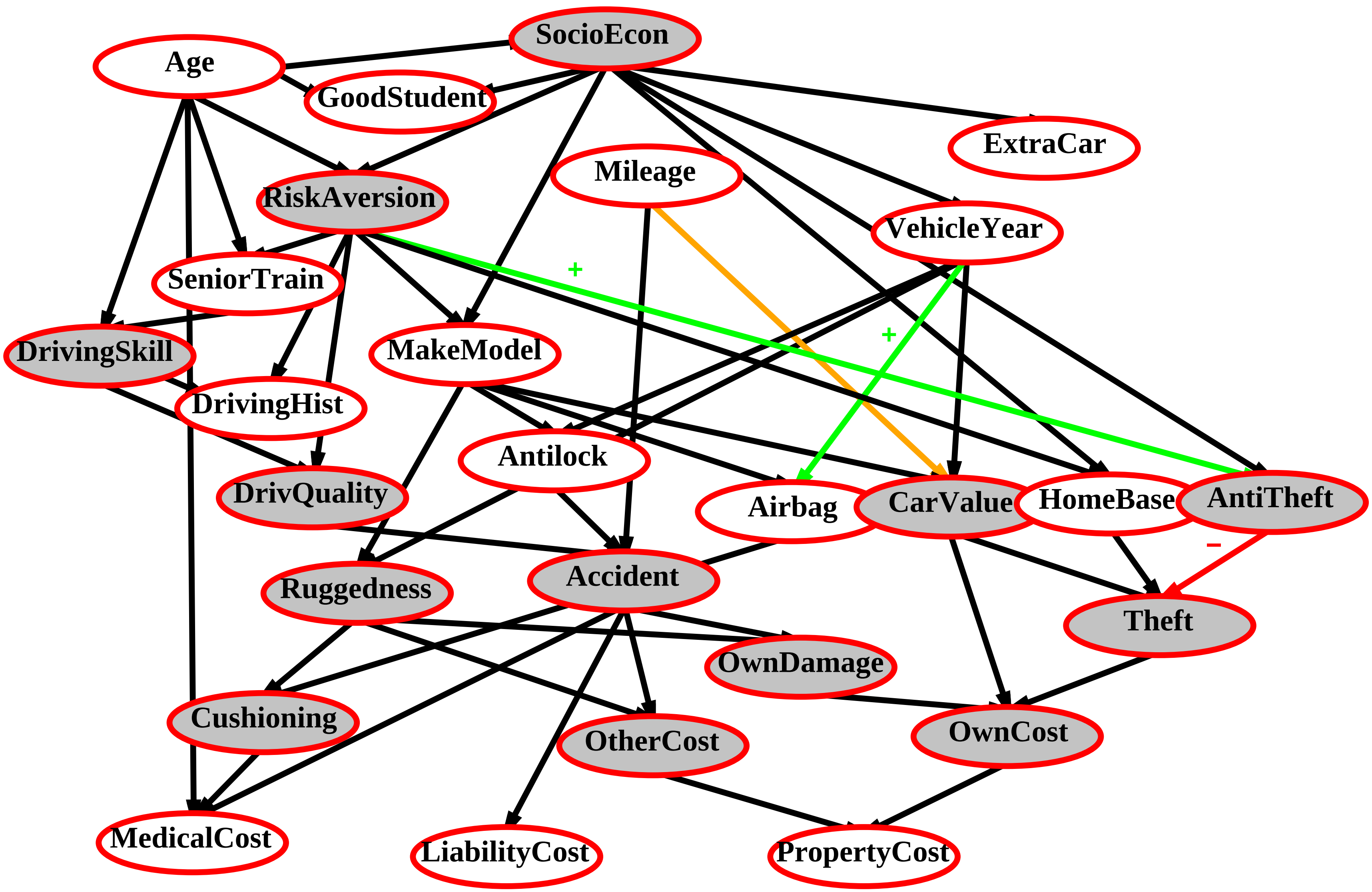
Label the arcs + or −



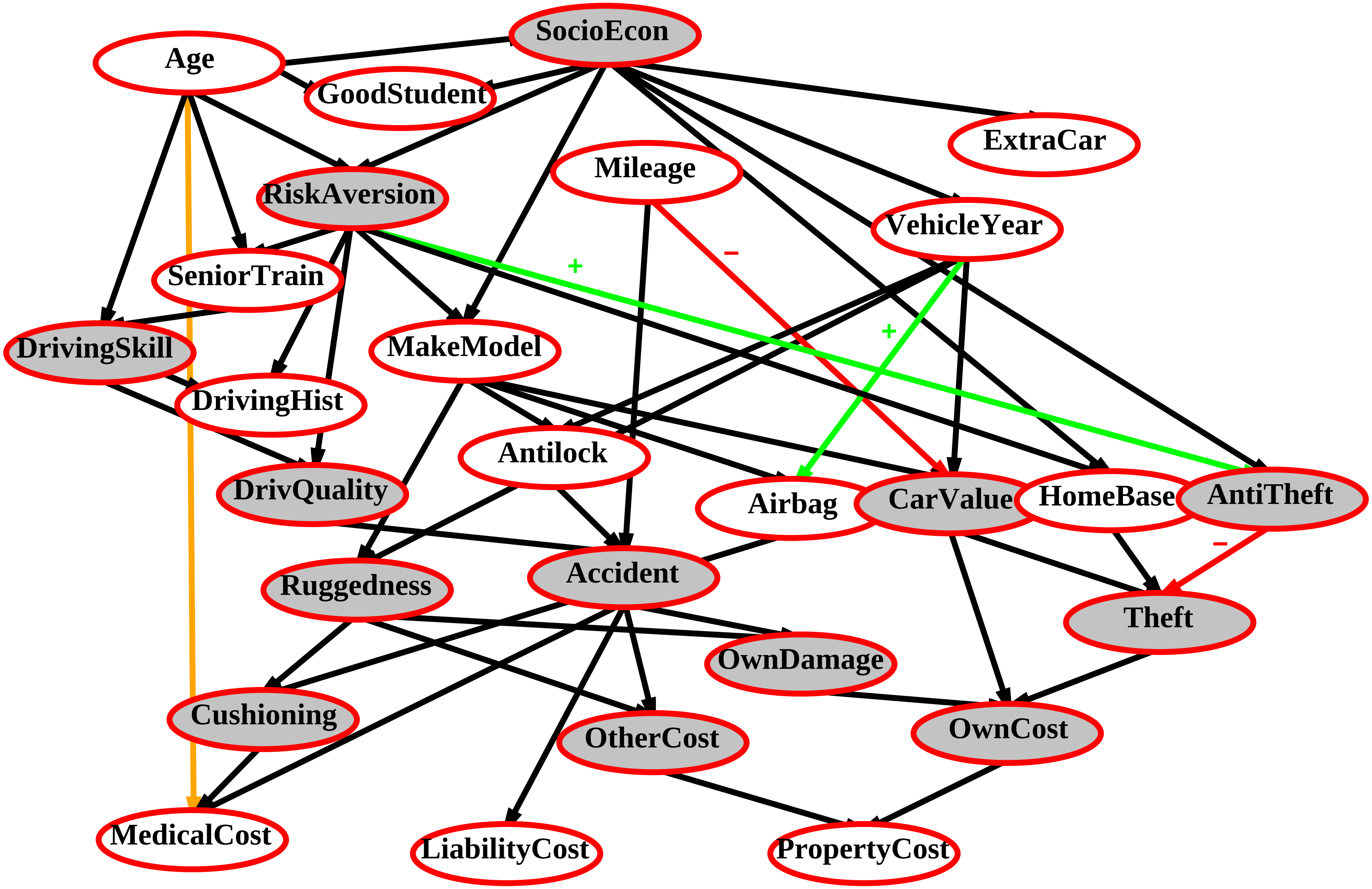
Label the arcs + or −



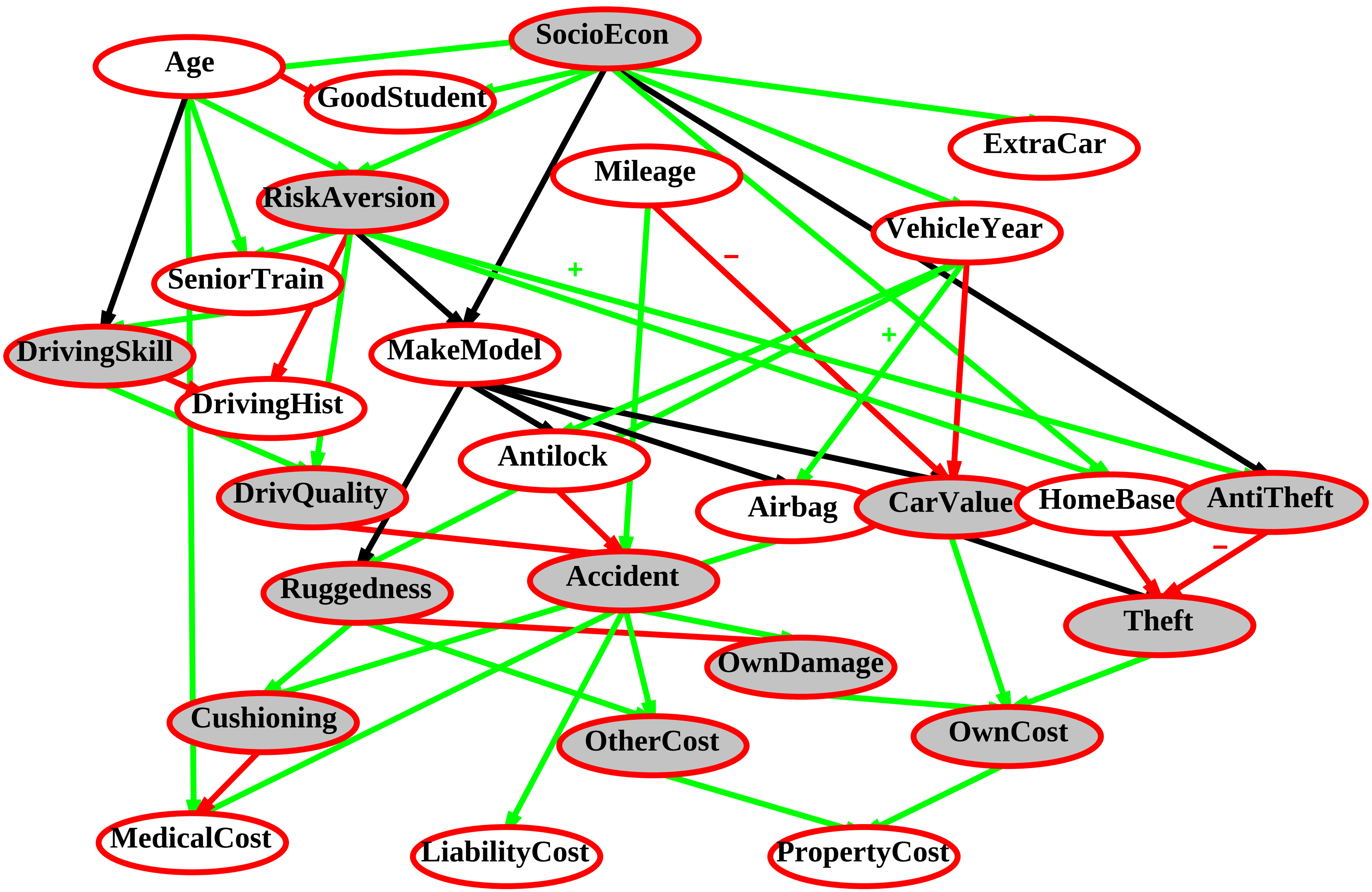
Label the arcs + or −



Label the arcs + or −



Label the arcs + or −



Independence: Deterministic Environments

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, q deaths/mpm⟩ vs.
⟨70,000 suffer, \$4.2 billion, q deaths/mpm⟩

$X_1 \ldots X_n$ are mutually preferentially independent if each pair of variables is preferentially independent of each other variable.

Mutual preferential independence $\Rightarrow$ existence of additive value function:

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

Independence: Stochastic environments

Need to consider preferences over *lotteries*:
Set of variables $\mathbf{X}_1$ is utility-independent of $\mathbf{X}_2$ if
preferences over lotteries in $\mathbf{X}_1$ do not depend on $\mathbf{X}_2$.

Set of variables $\mathbf{X}$ is mutually utility independent if each subset of its variables is utility-independent of the remaining variables.

Mutual UI implies existence of a multiplicative utility function:

$$U(\mathbf{X}) = k_1 U_1(\mathbf{X}) + k_2 U_2(\mathbf{X}) + k_3 U_3(\mathbf{X})$$
$$+ k_1 k_2 U_1(\mathbf{X}) U_2(\mathbf{X}) + k_2 k_3 U_2(\mathbf{X}) U_3(\mathbf{X}) + k_3 k_1 U_3(\mathbf{X}) U_1(\mathbf{X})$$
$$+ k_1 k_2 k_3 U_1(\mathbf{X}) U_2(\mathbf{X}) U_3(\mathbf{X})$$

Note: N component single-variable utility functions and N free parameters.

Value of information

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

Example: hidden money
Holding \$1 coin in one closed hand
$P(left) = P(right) = .5$
You can pay \$0.50 to guess which hand has coin.
Liz offers to tell whether left hand contains coin. Fair price?

Solution: compute expected value of information
$=$ expected value of best action given the information
minus expected value of best action without information
$=$ 0.50 - 0.00 $=$ 0.50

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)

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