TDT4171 Artificial Intelligence Methods

Lecture 5 – Rational Agents

Norwegian University of Science and Technology

Helge Langseth
IT-VEST 310
helgel@ntnu.no

NTNU

Outline

1. Summary from last time
2. Chapter 16: Rational Agents
   - Rational preferences
   - Utilities
   - Money
   - Multiattribute utilities
   - Decision networks
   - Dominance
   - Value of information
3. Summary

Summary from last time

- **Temporal models** — variables replicated over time
- **Markov assumptions** and **stationarity assumption**, so we need
  - Transition model $\mathbf{P}(\mathbf{X}_t|\mathbf{X}_{t-1})$
  - Sensor model $\mathbf{P}(\mathbf{E}_t|\mathbf{X}_t)$
- Tasks are filtering, prediction, smoothing, most likely sequence; **all done recursively with constant cost per time step**
- Classes of models we consider:
  - **Hidden Markov models** have a single discrete state variable; used for speech recognition
  - **Dynamic Bayes nets** subsume HMMs – exact update intractable; approximations exist

Remember:
No lecture next week.

Chapter 16 – Learning goals

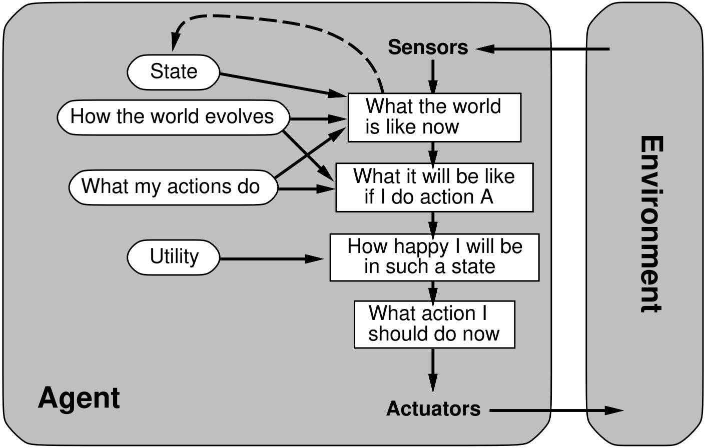
Understanding the relationship between

1. **Rational behaviour** – *"doing what is expected to maximize goal achievement, given the available information"*
2. **Preference structures**
3. **Utilities**

Being familiar with:

- **Utility functions** – Their foundation and definition
- **Utility elicitation**
- **Influence diagrams**

The utility-based agent



Preferences

An agent chooses among **prizes** (A, B, etc.)

Notation:

- $A \succ B$ — A preferred to B
- $A \sim B$ — indifference between A and B
- $A \succsim B$ — A preferred to B or indifference between A and B

Assumption:

An agent will always be able to compare to prizes A and B.

$\Rightarrow$ **No indecisiveness**.

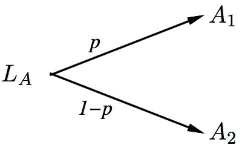
Preferences

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

Notation:

- $L_A \succ L_B$ — L_A preferred to L_B
- $L_A \sim L_B$ — indifference between L_A and L_B
- $L_A \succsim L_B$ — L_A preferred to L_B or indifference

Lotteries:
$$L_A = [p, A_1;\ (1-p), A_2]$$
$$L_B = [p, B_1;\ (1-p), B_2]$$



Again:
It is **not an option** to “chicken out”; a relation between L_A and L_B can always be established.

A small "quiz"

Which of the following two lotteries would you prefer?

- Lottery A: $[1, \$10\text{mill}]$,
- Lottery B: $[0.1, \$50\text{mill}; 0.89, \$10\text{mill}; 0.01, \$0]$.

What about these two:

- Lottery C: $[0.11, \$10\text{mill}; 0.89, \$0]$,
- Lottery D: $[0.1, \$50\text{mill}; 0.9, \$0]$.

Do you make **rational** choices if you "follow your heart"?
. . . **and what does that even mean?**

Discuss with your neighbour for a couple of minutes.

Rational preferences

Idea: Preferences of a rational agent must obey constraints.

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

We hope:

Rational preferences follow some "rules"

$\Rightarrow$ Behavior can be described using a mathematical formulation.

$\Rightarrow$ Behavior can be implemented in an intelligent agent.

Rational preferences contd.

Violating the constraints leads to self-evident irrationality

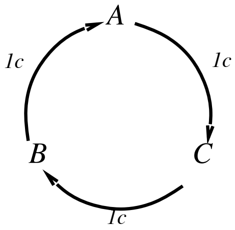
Example: An agent with intransitive preferences can be induced to give away all its money!

Assume he has preferences $A \succ B \succ C \succ A$ and see what happens.

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 – The foundation for the *Utility function*

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

1. $U(A) \geq U(B) \quad \Leftrightarrow \quad A \succsim B$
2. $U([p_1, S_1; \ \ldots \ ; p_n, S_n]) = \sum_i \ p_i \cdot U(S_i) = \mathbb{E}_S[U(S)]$

This gives rise to the MEU principle:
To be rational, the agent must choose the action that maximizes expected utility!

Action selection:

$$\mathbb{E}U(A|\mathbf{e}) = \sum_i P\left(\mathtt{Result}_i(A)|\mathbf{do}(A), \mathbf{e}\right) \cdot U\left(\mathtt{Result}_i(A)\right)$$

Maximizing expected utility

Theorem – The foundation for the *Utility function*

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

1. $U(A) \geq U(B) \quad \Leftrightarrow \quad A \succsim B$
2. $U([p_1, S_1; \ \ldots \ ; p_n, S_n]) = \sum_i \ p_i \cdot U(S_i) = \mathbb{E}_S[U(S)]$

This gives rise to the MEU principle:
To be rational, the agent must 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! (For example, a lookup table for perfect tic-tac-toe)

Being rational vs. being “all-seeing” vs. being lucky

Rational decision making?

Play movie at `./../../poker.mp4`

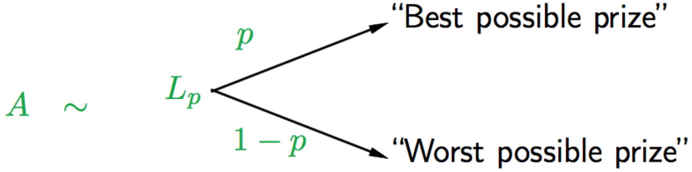
Utilities – and how to quantify them. . .

Utilities map states to real numbers – but 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 $A \sim L_p$;
  $U(A) \leftarrow p \cdot u_\top + (1-p) \cdot u_\bot$

(This makes sense, as we already think of probabilities in terms of accepting bets. . .)



Utilities – and how to quantify them...

Utilities map states to real numbers – but 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 $A \sim L_p$;
  $U(A) \leftarrow p \cdot u_\top + (1-p) \cdot u_\perp$

Note: Behavior is **invariant** w.r.t. linear transformation

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

Therefore, it is natural use **normalized utilities**; $u_\top = 1.0$, $u_\perp = 0.0$, and we get $U(A) = p$ in the procedure above.

Are you rational? – Example continued

Recall:

- Lottery A: $[1, \$10\text{mill}]$,
- Lottery B: $[0.1, \$50\text{mill}; 0.89, \$10\text{mill}; 0.01, \$0]$.
- Lottery C: $[0.11, \$10\text{mill}; 0.89, \$0]$,
- Lottery D: $[0.1, \$50\text{mill}; 0.9, \$0]$.

Are you rational? – Example continued

Recall:

- Lottery A: $[1, \$10\text{mill}]$,
- Lottery B: $[0.1, \$50\text{mill}; 0.89, \$10\text{mill}; 0.01, \$0]$.
- Lottery C: $[0.11, \$10\text{mill}; 0.89, \$0]$,
- Lottery D: $[0.1, \$50\text{mill}; 0.9, \$0]$.

Let $U(\$50\text{mill}) = 1$, $U(\$0) = 0$, $U(\$10\text{mill}) = u$. If you prefer Lottery A over Lottery B we get

$$u > 0.1 + 0.89u \qquad \Leftrightarrow \qquad u > \frac{10}{11}.$$

Are you rational? – Example continued

Recall:

- Lottery A: $[1, \$10\text{mill}]$,
- Lottery B: $[0.1, \$50\text{mill}; 0.89, \$10\text{mill}; 0.01, \$0]$.
- Lottery C: $[0.11, \$10\text{mill}; 0.89, \$0]$,
- Lottery D: $[0.1, \$50\text{mill}; 0.9, \$0]$.

Let $U(\$50\text{mill}) = 1$, $U(\$0) = 0$, $U(\$10\text{mill}) = u$. If you prefer Lottery A over Lottery B we get

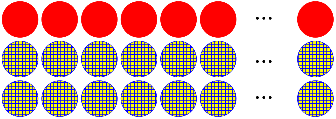
$$u > 0.1 + 0.89u \qquad \Leftrightarrow \qquad u > \frac{10}{11}.$$

Hence, $$\mathbb{E}U(C) = 0.11u > 0.11\frac{10}{11} = 0.1 = \mathbb{E}U(D),$$

and **Lottery C** $\succ$ **Lottery D** for a rational agent when $L_A \succ L_B$.

Are you rational? – New example

- 60 balls in an urn
- 20 are RED, 40 are BLUE **or** YELLOW
- We don't know how many are BLUE or YELLOW



- A: You receive \$100 if you draw a RED ball,
- B: You receive \$100 if you draw a BLUE ball
- C: You receive \$100 if you draw a RED or YELLOW ball
- D: You receive \$100 if you draw a BLUE or YELLOW ball

Which is better?

- Lottery A or Lottery B?
- Lottery C or Lottery D?

Are you rational? – New example

- 60 balls in an urn
- 20 are **RED**, 40 are **BLUE or YELLOW**
- We don't know how many are **BLUE** or **YELLOW**

- A: You receive \$100 if you draw a **RED** ball,
- B: You receive \$100 if you draw a **BLUE** ball

Let $R = P(\text{Red ball})$, $B = P(\text{Blue ball})$ $Y = P(\text{Yellow ball})$.
Let $U(\$0) \equiv 0$ and $U(\$100) \equiv 1$.

If Lottery A $\succ$ Lottery B, it means that

$$R \cdot 1 + (B+Y) \cdot 0 > B \cdot 1 + (R+Y) \cdot 0 \;\Rightarrow\; R > B$$

Are you rational? – New example

- 60 balls in an urn
- 20 are **RED**, 40 are **BLUE** **or** **YELLOW**
- We don't know how many are **BLUE** or **YELLOW**

- C: You receive \$100 if you draw a **RED** or **YELLOW** ball
- D: You receive \$100 if you draw a **BLUE** or **YELLOW** ball

Let $R = P(\text{Red ball})$, $B = P(\text{Blue ball})$ $Y = P(\text{Yellow ball})$.
Let $U(\$0) \equiv 0$ and $U(\$100) \equiv 1$.

If Lottery A $\succ$ Lottery B, it means that

$$R \cdot 1 + (B+Y) \cdot 0 > B \cdot 1 + (R+Y) \cdot 0 \quad \Rightarrow \quad R > B$$

If Lottery D $\succ$ Lottery C, it means that

$$R \cdot 0 + (B+Y) \cdot 1 > B \cdot 0 + (R+Y) \cdot 1 \quad \Rightarrow \quad B > R$$

Human are sometimes irrational – Consequences?

Why are people irrational (sometimes)?

Are we . . .

- not obeying the MEU principle (but which constraints are unreasonable?)
- lacking computational power?
- focussing on "the lottery itself", e.g. **regret**?

Consequences for AI:

- The choices a rational agent can make are **only as good** as the preferences they are based on.
- If the agent is given conflicting preference judgements, it is **not possible** for the rational agent to understand (or mimic) them.
- Acting rationally (the point of the agents this course) is **not** the same as acting like a human!

Money and utilities – An example

- You pay a fixed fee M to enter a game.
- A fair coin is tossed repeatedly until a "`tail`" appears, ending the game. You win 2^k, where k is the number of "`heads`" you have seen prior to the "`tail`".

What would be a fair price to pay for entering the game?

- How can a fair price be found?
- What is it in this case?

Discuss with your neighbour for a couple of minutes.

Money and utilities – An example

- You pay a fixed fee M to enter a game.
- A fair coin is tossed repeatedly until a "`tail`" appears, ending the game. You win 2^k, where k is the number of "`heads`" you have seen prior to the "`tail`".

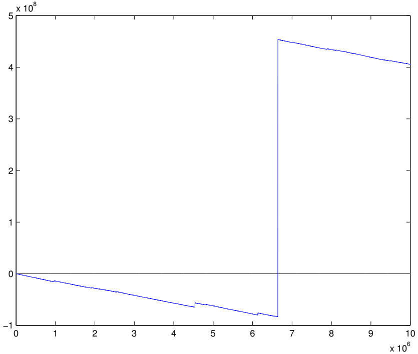
$$
\begin{aligned}
\mathbb{E}[\text{Payout}] &= -M + \sum_{i=0}^{\infty} P(k=i) \cdot \text{Payout}(i) \\
&= -M + \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{i+1} \cdot 2^i \\
&= -M + \sum_{i=0}^{\infty} \frac{1}{2} = \infty
\end{aligned}
$$

Any finite entry cost is fair!

Money and utilities – An example

- You pay a fixed fee M to enter a game.
- A fair coin is tossed repeatedly until a "`tail`" appears, ending the game. You win 2^k, where k is the number of "`heads`" you have seen prior to the "`tail`".

Example: Pay 25 to take part

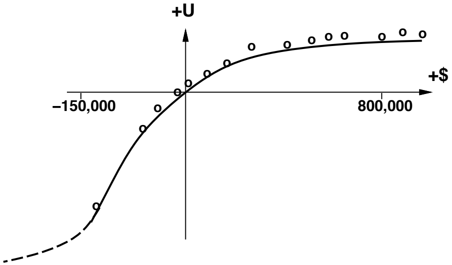


Money

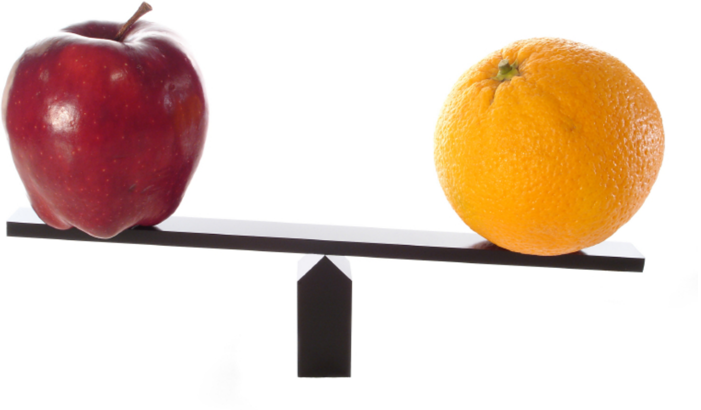
Money does **not** behave as a utility function

Given a lottery L with expected monetary value $\mathtt{EMV}(L)$, usually $U(L) < U(\mathtt{EMV}(L))$, i.e., people are **risk-averse**.

Typical empirical data:

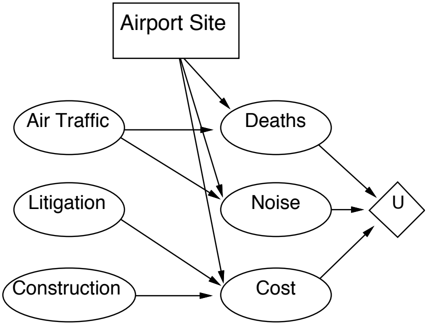


Multiattribute utility theory



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**

Decision networks – in GeNIe

Must define the utility of each combination of consequences

- `Deaths`
- `Noise`
- `Cost`

GeNIe-demo: `airport.xdsl`

Multiattribute utility theory

How can we handle utility functions of **many variables** $X_1 \ldots X_n$, e.g., what is $U(\texttt{Deaths}, \texttt{Noise}, \texttt{Cost})$?

How can complex utility functions be **assessed** from preference **behaviour**?

Idea 1: Identify various types of **independence** in preferences and derive consequent canonical forms for $U(x_1, \ldots, x_n)$

Idea 2: Identify conditions under which decisions can be made without complete identification of $U(x_1, \ldots, x_n)$

Preference structure: Dependence

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., $\langle \texttt{Noise}, \texttt{Cost}, \texttt{Safety} \rangle$:

- $\langle$20,000 suffer, \$4.6 billion, x deaths/mpm$\rangle \succ$
- $\langle$70,000 suffer, \$4.2 billion, x deaths/mpm$\rangle$

for any number of deaths $\Rightarrow$ `Safety` P.I. of `Noise` and `Cost`

Theorem

If every pair of attributes is P.I. then there is an **additive** value function:

$$U(x_1, \ldots, x_n) = \sum_{i=1}^{n} U_i(x_i)$$

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

Decision networks – in GeNIe

Must define the utility of each of the consequences **separately**

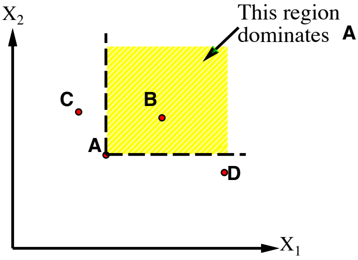
- `Deaths`
- `Noise`
- `Cost`

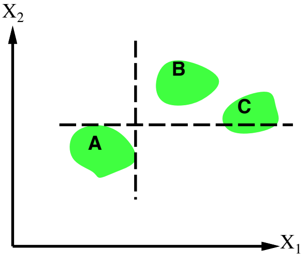
GeNIe-demo: `airportAdditive.xdsl`

Strict dominance

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

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





Strict dominance seldom holds in practice

Stochastic dominance

Consider the two lotteries where I flip a coin:

- **Lottery A:** You get $\$1$ if I get `heads` and $\$2$ if I get `tails`
- **Lottery B:** You get $\$3$ if I get `heads` and $\$1$ if I get `tails`

It is obvious which lottery is the better one for you...

Stochastic dominance

Consider the two lotteries where I flip a coin:

- **Lottery A:** You get $\$1$ if I get `heads` and $\$2$ if I get `tails`
- **Lottery B:** You get $\$3$ if I get `heads` and $\$1$ if I get `tails`

It is obvious which lottery is the better one for you...

- Independent of lottery, there is a 50% probability you are unlucky (get `heads` in Lottery A, `tails` in Lottery B).
  Outcome is worth $\$1$ in either case
- Independent of lottery, there is a 50% probability you are lucky (get `tails` in Lottery A, `heads` in Lottery B).
  Outcome is worth $\$2$ and $\$3$ respectively.
- **Seems appropriate to say that Lottery B is better...**

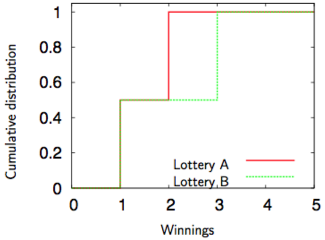
Stochastic dominance

Consider the two lotteries where I flip a coin:

- **Lottery A:** You get $\$1$ if I get `heads` and $\$2$ if I get `tails`
- **Lottery B:** You get $\$3$ if I get `heads` and $\$1$ if I get `tails`

It is obvious which lottery is the better one for you...

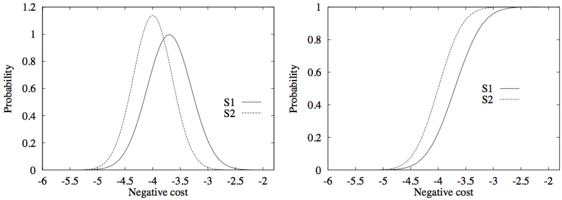
Cumulative distribution over winnings:



Whether lucky or unlucky, it is never silly to choose Lottery B.

$\Rightarrow$ **Lottery B stochastically dominates Lottery A**

Stochastic dominance - Mathematics



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:

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

Note! This is true for any monotonic U!

Multiattribute case: stoch.dom. on all attributes $\Rightarrow$ optimal

Stochastic dominance contd.

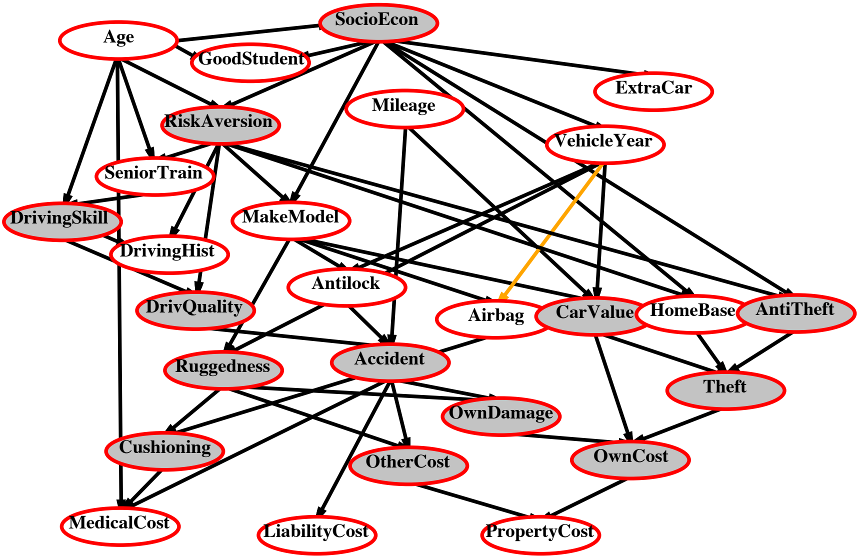
Stochastic dominance can often be determined without exact distributions using **qualitative** reasoning

- 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
- Injury increases with collision speed

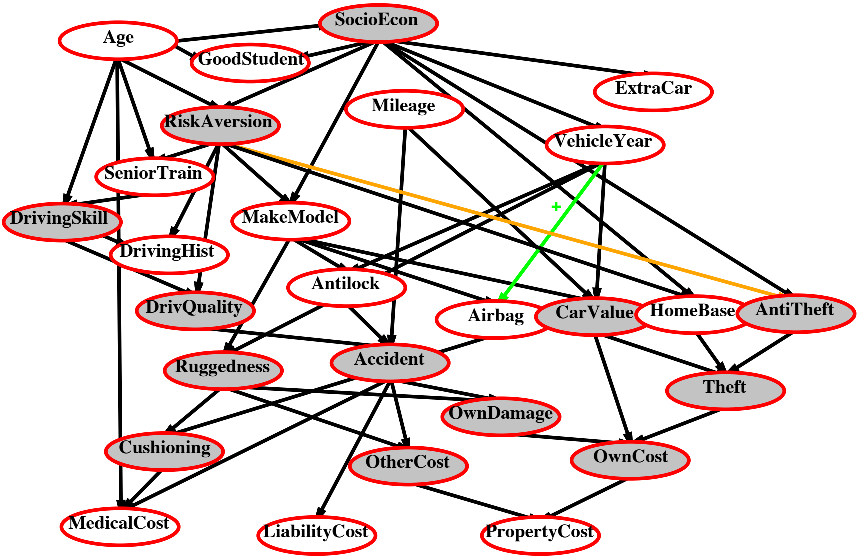
How about this plan for reasoning about dominance:

- Annotate BN with information about stochastic dominance
- $X \xrightarrow{+} Y$ (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 $\mathbf{P}(Y|x_2, \mathbf{z})$
- Dramatically simplifies the BN building process; allows qualitative inferences like "*Chances of cancer increases if you smoke*" (without quantification of the increase)

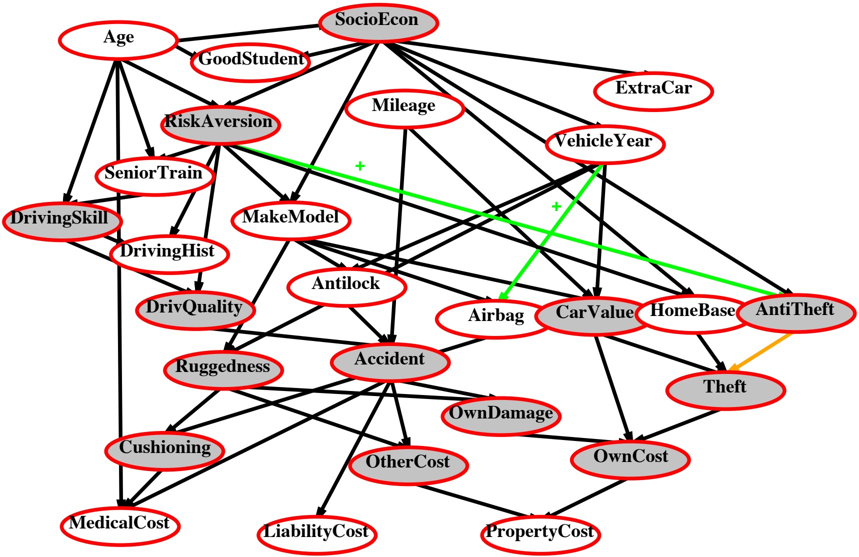
Label the arcs + or –



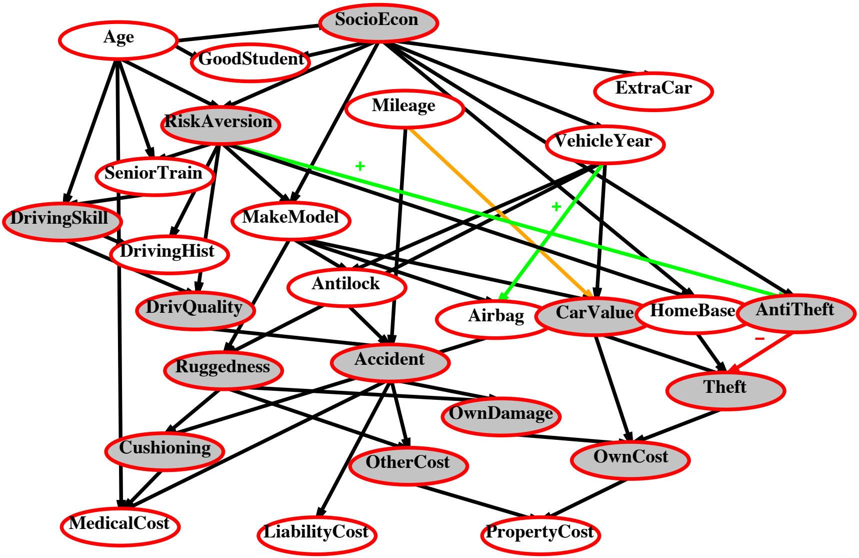
Label the arcs + or –



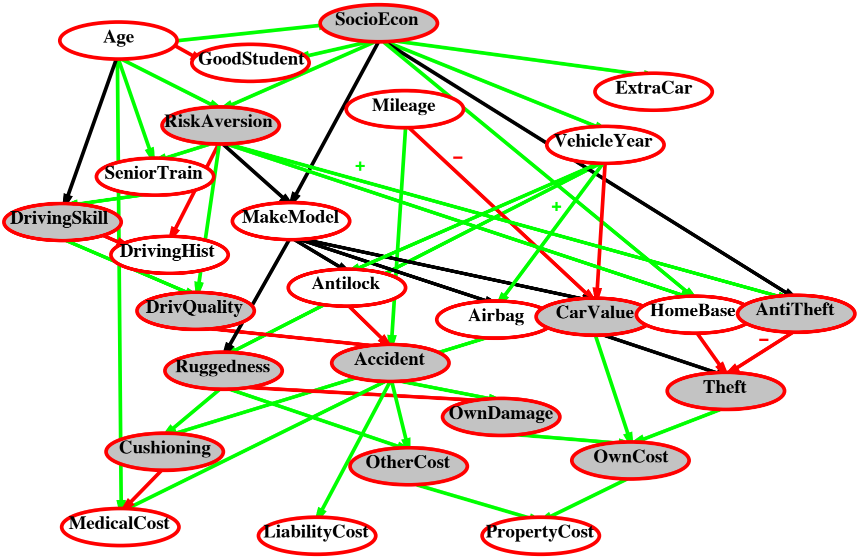
Label the arcs + or –



Label the arcs + or –



Label the arcs + or −



Value of information

Idea: Compute value of acquiring each possible piece of evidence.

Example: Buying oil drilling rights:

- Two blocks A and B, exactly one has oil, worth k
- Prior probabilities $1/2$ each, mutually exclusive.
- Current price of each block is $k/2$.
- A consultant offers accurate survey of A. **Fair price?**

Solution: compute expected value of information (VOI)

VOI = expected value of best action **given** the information
− expected value of best action **without** information

Value of information

Idea: Compute value of acquiring each possible piece of evidence.

Example: Buying oil drilling rights:

- Two blocks A and B, exactly one has oil, worth k
- Prior probabilities $1/2$ each, mutually exclusive.
- Current price of each block is $k/2$.
- A consultant offers accurate survey of A. **Fair price?**

Solution: compute expected value of information (VOI)

- Survey may say "oil in A" or "no oil in A", **prob. 0.5 each**

$$\begin{aligned}\text{VOI} &= [0.5 \times \text{value of "buy A" given "oil in A"} \\ &\quad + 0.5 \times \text{value of "buy B" given "no oil in A"}] - 0 \\ \text{VOI} &= (0.5 \times k/2) + (0.5 \times k/2) - 0 = k/2\end{aligned}$$

General formula for VPI: Value of Perfect Information

- Current evidence $\boldsymbol{E} = \boldsymbol{e}$, current best action α
- Possible action outcomes S_i, potential new evidence E_j

$$\mathbb{E}U(\alpha \mid \boldsymbol{e}) = \max_a \sum_i P(S_i \mid a, \boldsymbol{E} = \boldsymbol{e}) \cdot U(S_i)$$

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

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

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

$$\mathtt{VPI}_{\boldsymbol{E}}(E_j) = \overbrace{\sum_k \underbrace{P(E_j = e_{jk} \mid \boldsymbol{e})}_{\text{Prob. for } E_j = e_{jk}.} \cdot \underbrace{\mathbb{E}U(\alpha_{e_{jk}} \mid \boldsymbol{e}, E_j = e_{jk})}_{\text{Expected utility when } E_j = e_{jk}.}}^{\text{Expected utility over uncertain } E_j.} - \mathbb{E}U(\alpha \mid \boldsymbol{e})$$

Luckily, we can do this in GeNIe instead!

Value Of Information – in GeNIe

GeNIe-demo: `oil.xdsl`

GeNIe-demo: `VOI.xdsl`

Properties of VPI

Nonnegative in expectation:

$$\forall\, j, E \;\; \mathsf{VPI}_E(E_j) \geq 0$$

Order-independent

$$\begin{aligned} \mathsf{VPI}_E(E_j, E_k) &= \mathsf{VPI}_E(E_j) + \mathsf{VPI}_{E,E_j}(E_k) \\ &= \mathsf{VPI}_E(E_k) + \mathsf{VPI}_{E,E_k}(E_j) \end{aligned}$$

Nonadditive; consider, e.g., obtaining E_j twice

$$\mathsf{VPI}_E(E_j, E_k) \neq \mathsf{VPI}_E(E_j) + \mathsf{VPI}_E(E_k)$$

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. It is **NP-complete** in general.

Summary

- Rational agents can always use **utilities** to make decisions
- The **MEU principle** tells us how to behave
- It can be quite laborious to elicit preference structures from domain experts
  - $\Rightarrow$ **structured approaches** are available
- **Value of Information** helps focus information gathering for rational agents
- **Influence diagrams** are extensions to BNs that let us make rational decisions.

Remember:
No lecture next week.