GHENT UNIVERSITY

FACULTY OF ENGINEERING AND ARCHITECTURE

E016350 - Artificial Intelligence

Lecture 16

Decisions & Action

Rational decisions

Aleksandra Pizurica

Ghent University

Fall 2024

Overview

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

[R&N], Chapter 16

These slides are based on: S. Russel and P. Norvig: *Artificial Intelligence: A Modern Approach*, (Fourth Ed.), http://aima.cs.berkeley.edu/, the corresponding slides of S. Russel and the slides of D. Klein and P. Abbeel (course Introduction to Artificial Intelligence), http://ai.berkeley.edu/

Decision-theoretic agents

A decision-theoretic agent makes rational decisions based on
what it believes and **what it wants**.

Combines beliefs and desires **under uncertainty**.

While a logical agent can only make binary distinction between good and bad (i.e., goal and non-goal) states, a decision-theoretic agent has
a **continuous measure of outcome quality**.

The agent's preferences are captured by a **utility** function $U(s)$, which assigns a single number to a state s describing how desirable that state is.

Maximizing Expected Utility

The principle of Maximum Expected Utility (MEU): a rational agent chooses the action that maximizes the agent's expected utility given the evidence **e**:

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

The expected utility of an action given the evidence is the averaged utility of the outcomes, weighted by the probability that the outcome occurs:

$$EU(a|\mathbf{e}) = \sum_s P(\textsc{Result}(a) = s|a, \mathbf{e})U(s)$$

Conditioning on a in $P(\textsc{Result}(a) = s|a, \mathbf{e})$ stands for the event that action a is effectively executed.

Maximizing Expected Utility - equivalent formulations

The expected utility of the best possible action given the evidence:

$$EU(\alpha|\mathbf{e}) = \max_a \sum_s P(\textsc{Result}(a) = s|a, \mathbf{e})U(s)$$

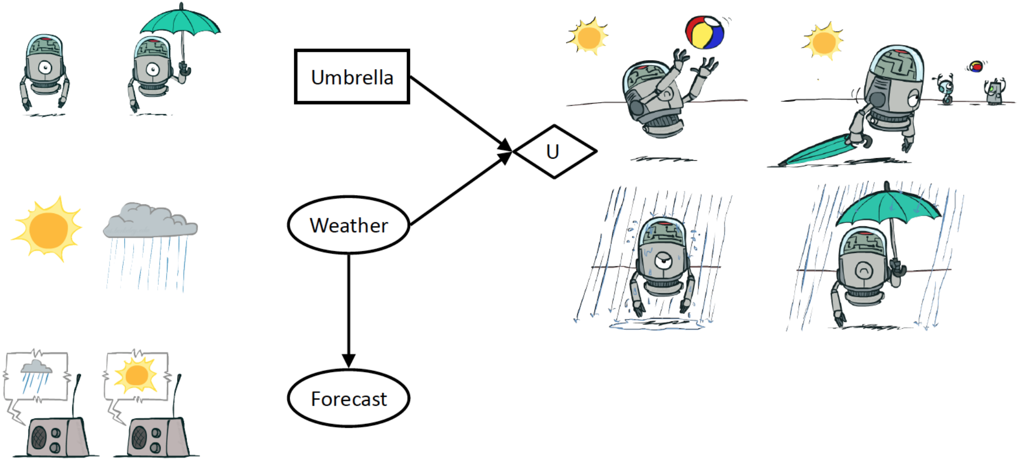
is the Maximum Expected Utility (MEU) of any action under the evidence $\mathbf{e}$.
Also referred to as the MEU of the evidence $\mathbf{e}$

$$MEU(\mathbf{e}) = \max_a EU(a|\mathbf{e}) = EU(\alpha|\mathbf{e})$$

and often expressed as:

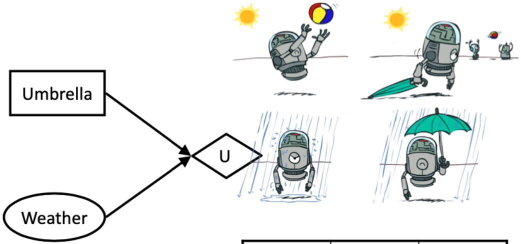
$$MEU(\mathbf{e}) = \max_a \sum_s P(s|\mathbf{e})U(s, a)$$

Example - Take or Leave Umbrella



Credit: D. Klein and P. Abbeel, Introduction to Artificial Intelligence), http://ai.berkeley.edu/

Example - Take or Leave Umbrella



W	P(W)
sun	0.7
rain	0.3

A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

Umbrella=leave:

$$EU(leave) = \sum_w P(w)U(leave, w) = 0.7 \cdot 100 + 0.3 \cdot 0 = 70$$

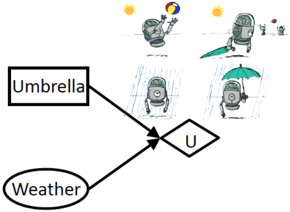
Umbrella=take:

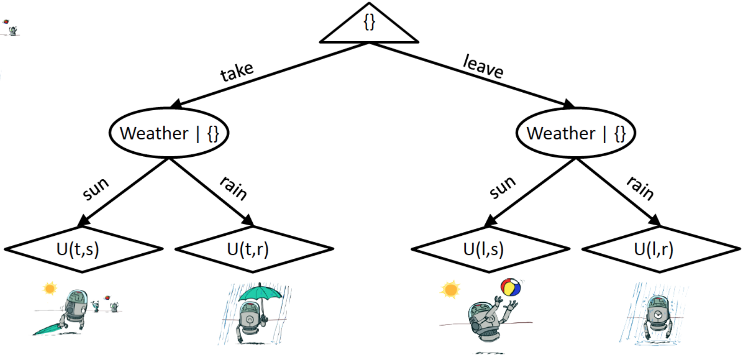
$$EU(take) = \sum_w P(w)U(take, w) = 0.7 \cdot 20 + 0.3 \cdot 70 = 35$$

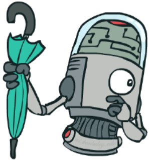
Optimal decision = leave

$$MEU(\{\}) = \max_a EU(a) = 70$$

Example - Take or Leave Umbrella







- Almost exactly like expectimax / MDPs
- What's changed?

Credit: D. Klein and P. Abbeel, Introduction to Artificial Intelligence), http://ai.berkeley.edu/

Preferences, prizes and lotteries

Notation:

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

A and B could be some states of the world, or some prizes

There is most often uncertainty about what is really being offered
(e.g., think of choices "the pasta dish" or "chicken" on a flight)

So, we can think of the set of outcomes of each action as a lottery

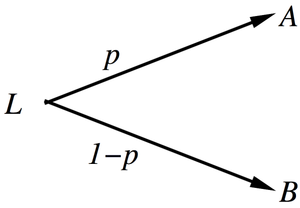
Lottery $\Leftrightarrow$ situation with uncertain prizes

Action outcome as a lottery

Denote lottery L with possible outcomes $S_1, ..., S_n$ that occur with probabilities $p_1, ..., p_n$ as:

$$L = [p_1, S_1; p_2, S_2; ...p_n, S_n]$$

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



Rational preferences

Rational preferences $\implies$

behavior describable as maximization of expected utility

Constraints on rational prefereces, also called **axioms of utility theory**:

Orderability

$$(A \succ B) \vee (B \succ A) \vee (A \sim B)$$

Transitivity

$$(A \succ B) \wedge (B \succ C) \implies (A \succ C)$$

Continuity

$$A \succ B \succ C \implies \exists p, \quad [p, A;\ 1-p, C] \sim B$$

Substitutability

$$A \sim B \implies [p, A;\ 1-p, C] \sim [p, B;\ 1-p, C]$$

Monotonicity

$$A \succ B \implies (p > q \Leftrightarrow [p, A;\ 1-p, B] \succ [q, A;\ 1-q, B])$$

Rational preferences, contd.

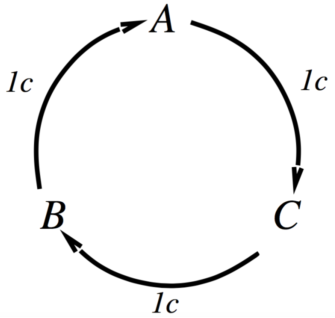
Violating the constraints leads to **self-evident irrationality**

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

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



Preferences lead to utility

From the axioms of utility, the following consequences can be derived:

Existence of utility function:

$$U(A) > U(B) \;\; \Leftrightarrow \;\; A \succ B$$
$$U(A) = U(B) \;\; \Leftrightarrow \;\; A \sim B$$

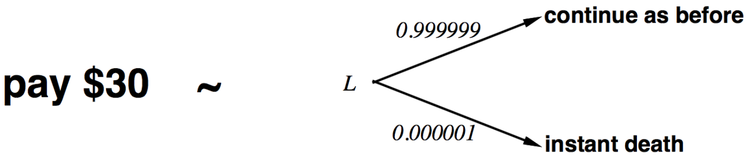
Expected utility of a lottery:

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

Utility functions

Utilities map from lotteries to real numbers and must obey certain utility axioms.
We cannot say much more – an agent can have any preferences it likes,
no matter how unusual they can be.

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: the utility function is not unique: the agent's behaviour wouldn't change if its utility function were transformed according to an affine 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

The utility of 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

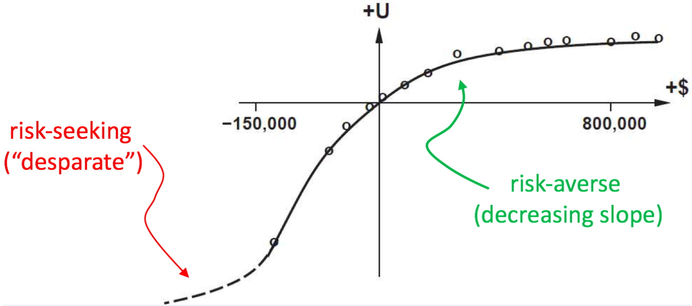
The utility of money

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

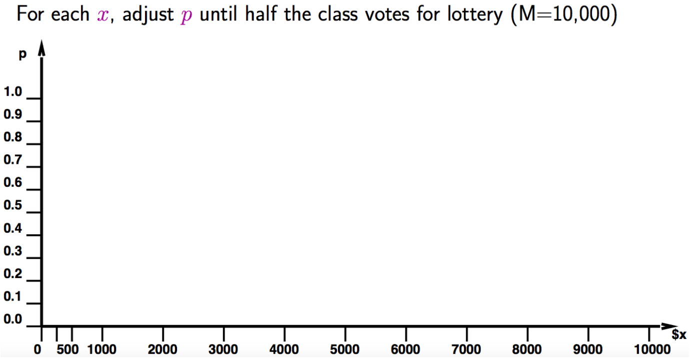
Utility curve: for what probability p am I indifferent between a prize x and a lottery $[p, \$M;\ (1-p), \$0]$ for large M?

Typical empirical data, extrapolated with risk-prone behaviour:



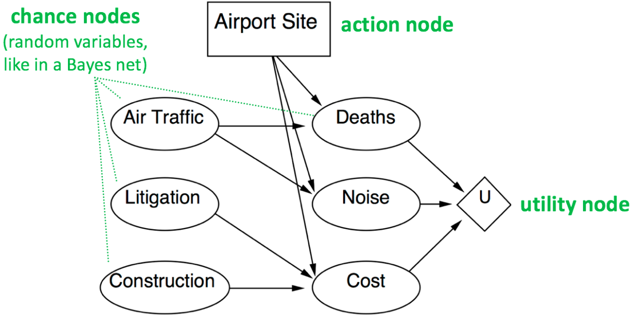
Student group utility

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

Multi-attribute utility

How can we handle utility functions of many variables $X_1 ... X_n$?

E.g., siting a new airport requires considerations of safety issues/death risks (arising from local topography), noise (how many people suffer), cost (of the land) etc.

What is $U(Deaths, Noise, Cost)$?

How can complex utility functions be assessed from preference behaviour?

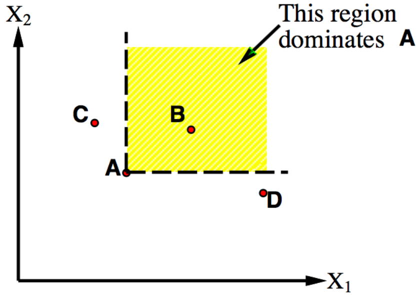
Idea 1: identify conditions under which decisions can be made without complete identification of $U(x_1, \ldots, x_n)$

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

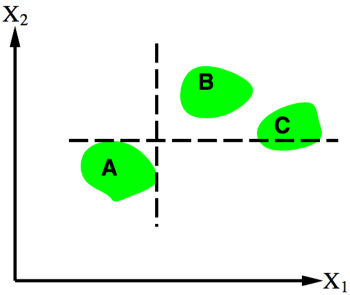
Strict Dominance

Typically define attributes such that U is monotonic in each

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

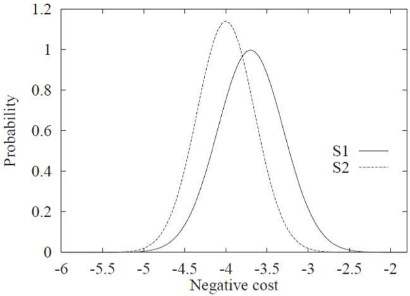


Deterministic attributes

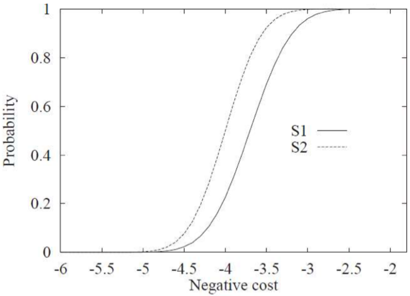


Uncertain attributes

Stochastic dominance



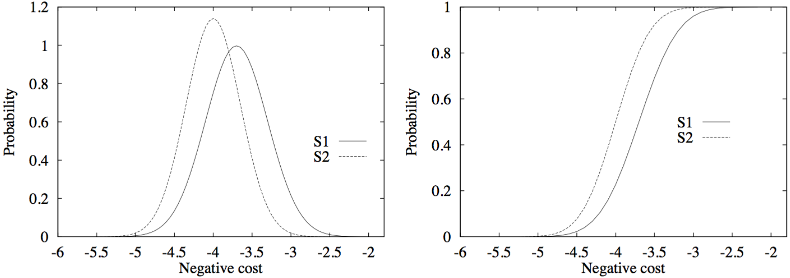
Distribution of the cost for siting the airport at S1 or S2



The corresponding cumulative distributions

Utility decreases with cost $\rightarrow S_1$ stochastically dominates S_2 (i.e., S_2 can be discarded)

Stochastic dominance contd.



Distribution p_1 stochastically dominates distribution p_2 iff

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

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

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

Multi-attribute case: stochastic dominance on all attributes $\implies$ optimal

Stochastic dominance contd.

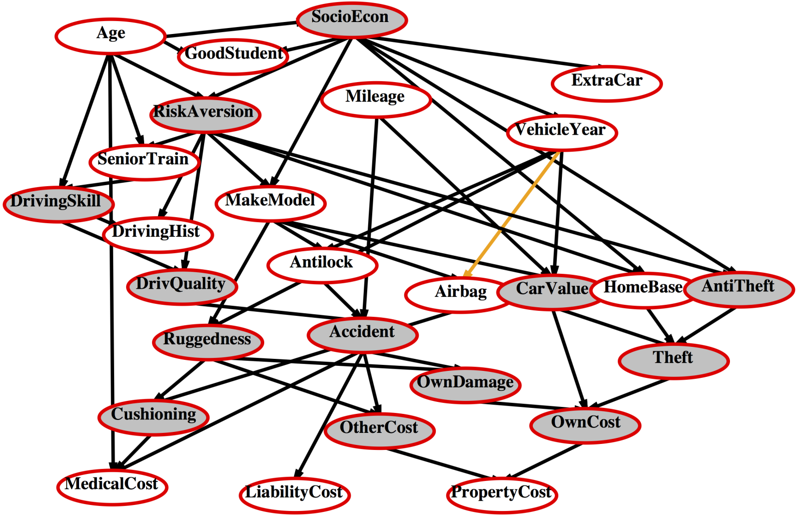
Stochastic dominance can often be determined without exact distributions using **qualitative** reasoning

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

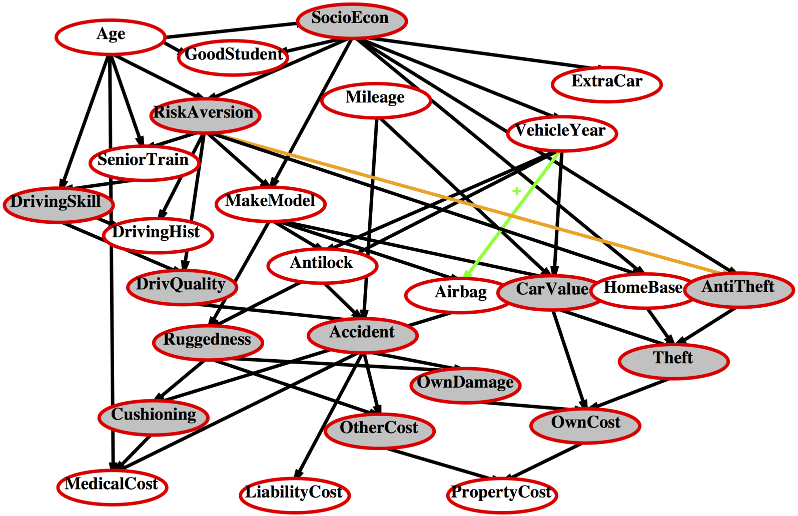
E.g., injury increases with collision speed

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

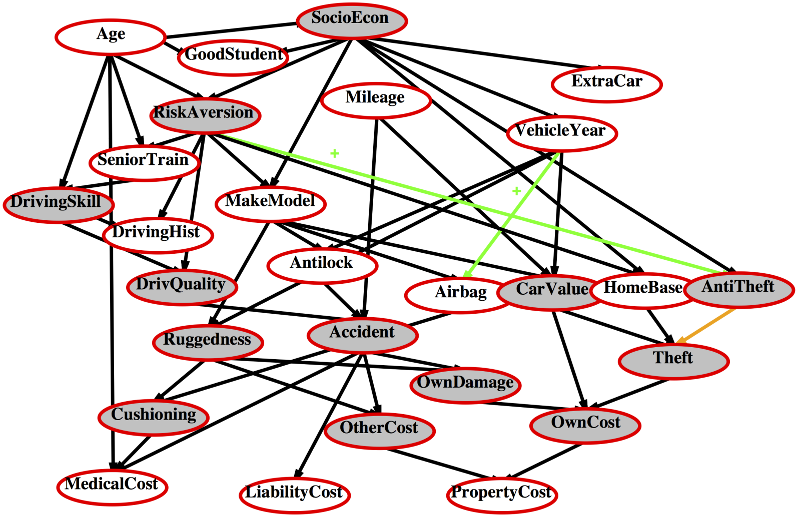
Example: label the arcs in belief nets



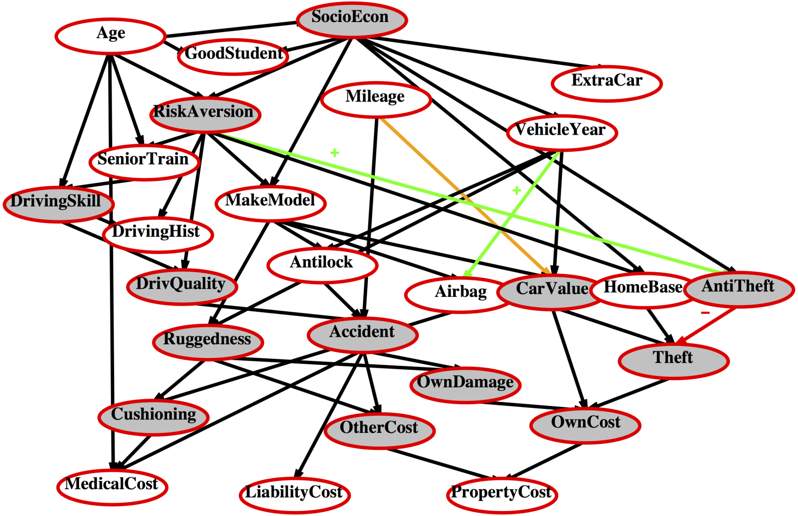
Example: label the arcs in belief nets



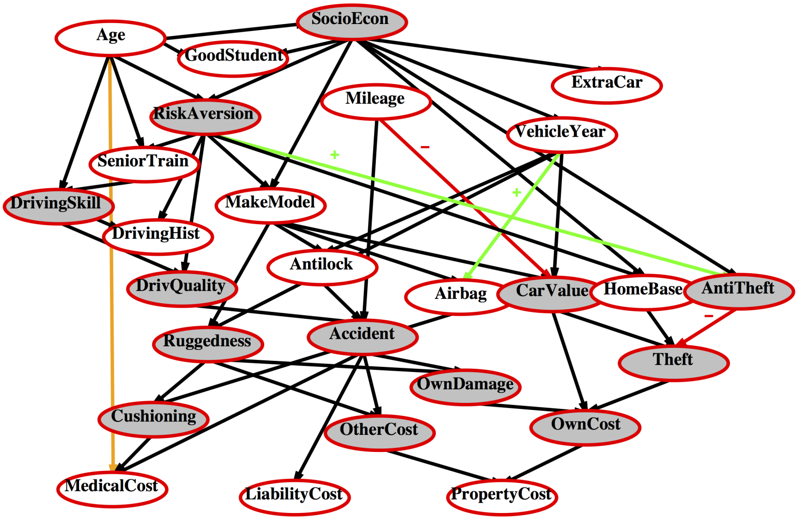
Example: label the arcs in belief nets



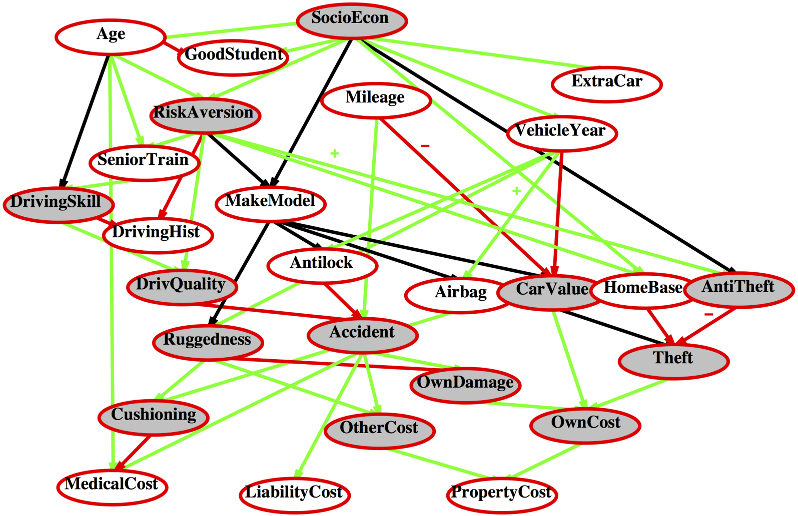
Example: label the arcs in belief nets



Example: label the arcs in belief nets



Example: label the arcs in belief nets



Preference structure

Suppose we have n attributes, each with d distinct possible values.
Specification of $U(x_1, ..., x_n)$ requires d^n values in the worst case

The worst case $\leftrightarrow$ the agent's preferences have no regularity at all

Multi-attribute theory is based on the supposition that the preferences of typical agents have some structure

Basic approach: identify regularities in the expected behaviour (some preference structure) and express the agent's utility function as

$$U(x_1, ..., x_n) = F[f(x_1), ..., f(x_n)]$$

Preference structure

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⟩ (mpm – million passenger miles)

The set of attributes $\langle Noise, Cost, Safety \rangle$ exhibits
mutual preferential independence (MPI)

If attributes $X_1, ..., X_n$ are MPI, then the agent's preference behaviour
can be described as additive value function: $V(x_1, ..., x_n) = \sum_i V(X_i)$

Preference structure: stochastic

Uncertainty $\rightarrow$ now we need to consider preferences over lotteries:

X is utility-independent of **Y** iff
preferences over lotteries in **X** do not depend on **y**

A set of attributes is mutually utility independent (MUI) if each of its subsets
is utility-independent of the remaining attributes

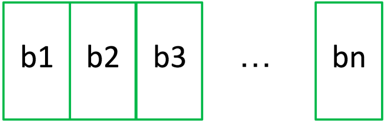
MUI implies that the agent's behaviour can be described using a
multiplicative utility function

For the case of three attributes:

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

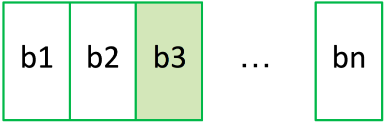
n-attribute problem with MUI needs n single-attribute functions and n constants

The value of information: example



- Only one block contains oil, and this oil is worth C dollars
- The price per block is C/n dollars

The value of information: example

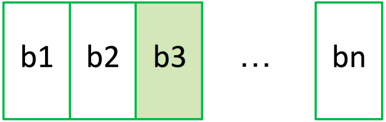


- Only one block contains oil, and this oil is worth C dollars
- The price per block is C/n dollars

A seismologist offers the results of a survey for block b3, which indicates definitely whether that block contains oil.

– How much should the company pay for this information?

The value of information: example



- Only one block contains oil, and this oil is worth C dollars
- The price per block is C/n dollars

A seismologist offers the results of a survey for block b3, which indicates definitely whether that block contains oil.

– How much should the company pay for this information?

To answer this, calculate the expected profit given the survey:

$$\frac{1}{n} \times (C - C/n) + \frac{n-1}{n} \times (C/(n-1) - C/n) = \frac{C}{n}$$

General formula

The value of the current best action a:

$$EU(\alpha|\mathbf{e}) = \max_a \sum_s P(\textsc{Result}(a) = s|a, \mathbf{e})U(s)$$

The value of the new best action (after obtaining the new evidence $E_j = e_j$) will be

$$EU(\alpha_{e_j}|\mathbf{e}, e_j) = \max_a \sum_s P(\textsc{Result}(a) = s|a, \mathbf{e}, e_j)U(s)$$

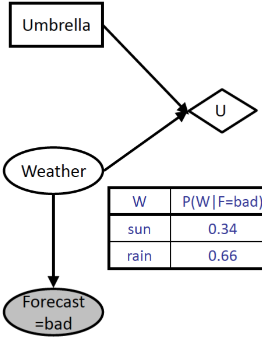
But the current value of E_j is unknown

$\rightarrow$ To determine the value of discovering E_j, given $\mathbf{e}$, we must average over all possible values of e_{jk} that we might discover for E_j:

$$VPI(E_j|\mathbf{e}) = \left(\sum_k P(E_j = e_{jk}|\mathbf{e})EU(\alpha_{e_{jk}}|\mathbf{e}, E_j = e_{jk})\right) - EU(\alpha|\mathbf{e})$$

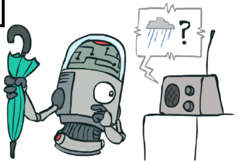
"Perfect" information because assumes obtaining exact evidence about the value of E_j

Example - Take or Leave Umbrella



A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

W	P(W|F=bad)
sun	0.34
rain	0.66



Umbrella=leave:

$$\begin{aligned} EU(leave|bad) &= \sum_w P(w|bad)U(leave,w) \\ &= 0.34 \cdot 100 + 0.66 \cdot 0 = 34 \end{aligned}$$

Umbrella=take:

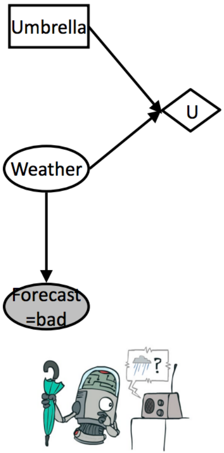
$$\begin{aligned} EU(take|bad) &= \sum_w P(w|bad)U(take,w) \\ &= 0.34 \cdot 20 + 0.66 \cdot 70 = 53 \end{aligned}$$

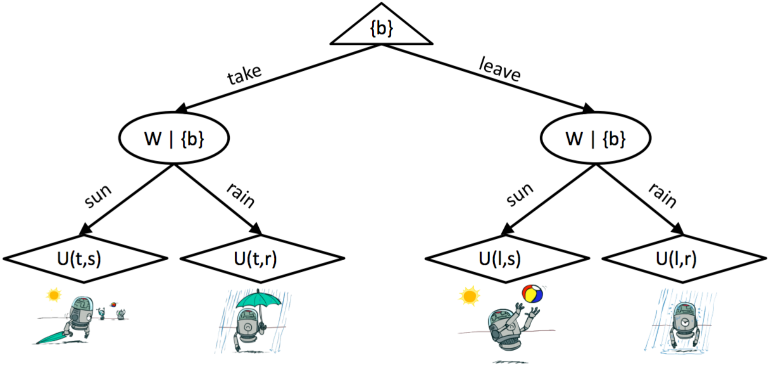
Optimal decision = take

$$MEU(F = bad) = \max_a EU(a|bad) = 53$$

Credit: D. Klein and P. Abbeel, Introduction to Artificial Intelligence), http://ai.berkeley.edu/

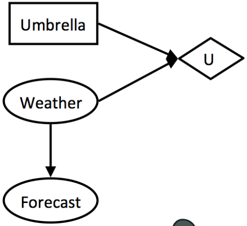
Example - Take or Leave Umbrella



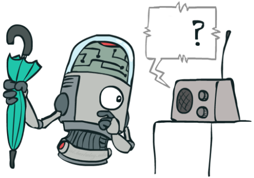


Credit: D. Klein and P. Abbeel, Introduction to Artificial Intelligence), http://ai.berkeley.edu/

Example - Take or Leave Umbrella



A	W	U
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70



MEU with no evidence:

$$MEU(\{\}) = \max_a EU(a) = 70$$

MEU if forecast is bad:

$$MEU(F = bad) = \max_a EU(a|bad) = 53$$

MEU if forecast is good:

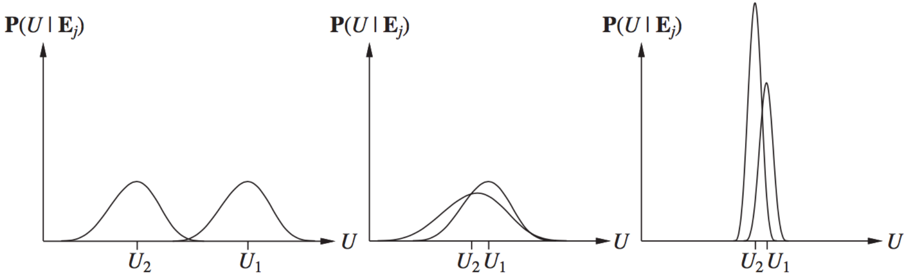
$$MEU(F = good) = \max_a EU(a|good) = 95$$

Forecast distribution: $P(F = good) = 0.59$

$$VPI(E'|e) = \sum_{e'} P(e'|e) MEU(e, e') - MEU(e) = 0.59 \cdot 95 + 0.41 \cdot 53 - 70 = 7.8$$

Qualitative behaviour

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**

$$\forall \mathbf{e}, E_j \quad VPI_{\mathbf{e}}(E_j) \geq 0$$

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

$$VPI_{\mathbf{e}}(E_j, E_k) \neq VPI_{\mathbf{e}}(E_j) + VPI_{\mathbf{e}}(E_k)$$

Order-independent

$$VPI_{\mathbf{e}}(E_j, E_k) = VPI_{\mathbf{e}}(E_j) + VPI_{\mathbf{e},e_j}(E_k) = VPI_{\mathbf{e}}(E_k) + VPI_{\mathbf{e},e_k}(E_j)$$

Note: when more than one piece of evidence can be gathered,
maximizing VPI for each to select one is not always optimal
$\Longrightarrow$ evidence-gathering becomes a **sequential** decision problem

Summary

Probability theory $\rightarrow$ what an agent should believe
Utility theory $\rightarrow$ what an agent wants, and
Decision theory (puts the two together) $\rightarrow$ what an agent should do

Rational agent – chooses the decision that leads to the best expected outcome (principle of maximum expected utility)

Lotteries – choices among uncertain prizes

Multi-attribute utility theory deals with utilities that depend on several distinct attributes of states

Decision networks – a natural extension of Bayesian networks, with action nodes and utility nodes

The value of information – the expected improvement in utility compared with making a decision without the information