Fundamentals of Artificial Intelligence – Rational Decisions

Matthias Althoff

TU München

Winter semester 2023/24

Organization

The content is covered in:

- S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach", section "Making Simple Decisions"
- D. Barber, "Bayesian Reasoning and Machine Learning"
- R. D. Shachter, "Evaluating Influence Diagrams", Operations Research, Vol. 34, No. 6, pp. 871-882, 1986

Learning Outcomes

- You understand the principle of *maximum expected utility*.
- You understand the required constraints for *rational preferences*.
- You understand that preferences lead to utility.
- You understand that utility is individual and know why it is helpful to normalize it.
- You can explain *strict dominance* and *stochastic dominance* for multiattribute utilities.
- You can select value functions for deterministic and stochastic preference structures.
- You can create *decision networks* for a given decision problem.
- You can compute optimal policies for decision networks.
- You can compute the *value of information*.

Overview of Probabilistic Methods

This lecture focuses on actions in static environments.

	Static environment	**Dynamic environment**
Without actions	Bayesian networks (lecture 9)	Hidden Markov models (lecture 10)
With actions	Decision networks (lecture 11)	Markov decision processes (lecture 12)

In this lecture, we assume an episodic environment $\Rightarrow$ *immediate* outcome.

Basic Idea

$$\boxed{\text{decision theory} = \text{probability theory} + \text{utility theory.}}$$

Probability theory

- We denote the probabilistic outcome of an action a as $\texttt{Result}(a)$, which is a random variable.
- The probability of an outcome is

$$P(\texttt{Result}(a) = s') = \sum_{s} P(s)P(s'|s, a). \qquad (1)$$

Utility theory

- We capture agents' preferences with **utility functions** $U(s)$ of state s.
- The expected utility $EU(a)$ is

$$EU(a) = \sum_{s'} P(\texttt{Result}(a) = s')U(s'). \qquad (2)$$

$$P(\text{Result}(a)=s') = P(s' \mid a)$$
$$= \sum_{s} P(s', s \mid a)$$
$$= \sum_{s} P(s' \mid s, a) \cdot P(s)$$

$$EV(a) = \sum_{s'} P(\text{Result}(a)=s') \cdot U(s')$$

$$\text{action} = \arg\max_{a} EV(a)$$
$$= \arg\max_{a} \sum_{s'} P(\text{Result}(a)=s') \cdot U(s')$$

Maximum Expected Utility

When following the principle of **maximum expected utility**, the best action is

$$action = \arg\max_a EU(a)$$
$$\stackrel{(2)}{=} \arg\max_a \sum_{s'} P(\texttt{Result}(a) = s')U(s').$$

- The above maximization can be seen as the ultimate goal of artificial intelligence.
- Obstacles in practice:
  - Estimating the state s of the world requires perception, learning, knowledge representation, and inference.
  - Computing $P(\texttt{Result}(a) = s')$ requires a complete causal model of the world and NP-hard inference in (very large) Bayesian networks.
  - Computing $U(s')$ often requires searching or planning, because one may only know how good a state is if one knows where to get from there.

Helpful Rules for Computing Expected Utilities

Let us introduce the random variables X and Y. We continue the previously introduced notation, e.g., for $P(X = x)$ we write $P(x)$.

- Obviously, we have that:

$$EU(x,y) = P(x,y)U(x,y) \tag{3}$$

$$EU(x) = \sum_y EU(x,y) \tag{4}$$

$$EU(x|y) = \frac{EU(x,y)}{P(y)} \tag{5}$$

- Inserting (3) in (4) results in

$$EU(x) = \sum_y P(x,y)U(x,y) \tag{6}$$

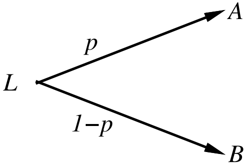
- Starting with (5), we have the following equivalences:

$$EU(x|y) = \frac{EU(x,y)}{P(y)} \overset{(3)}{=} P(x|y)U(x,y) \overset{(4)}{=} \sum_z P(x,z|y)U(x,y,z) \tag{7}$$

Preferences

- Utility is based on preferences.
- An agent chooses among **prizes** (A, B, etc.) and **lotteries**, i.e., situations with uncertain prizes:

Lottery $L = [p, A;\ (1-p), B]$
(pairs of prizes and probabilities)



We introduce preferences between prizes, which is denoted by

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

Rational Preferences: Constraints

Idea: preferences of a rational agent must obey constraints.
Constraints (also known as axioms of utility theory):

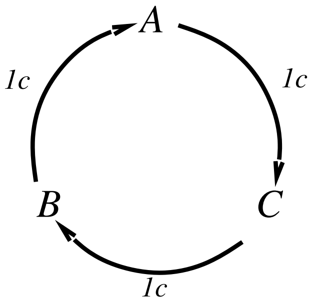
- Orderability (The agent cannot avoid deciding)
  $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
- Transitivity
  $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
- Continuity:
  $A \succ B \succ C \Rightarrow \exists\, p \quad [p, A;\ 1-p, C] \sim B$
- Substitutability (Also holds if we substitute $\succ$ for $\sim$)
  $A \sim B \Rightarrow [p, A;\ 1-p, C] \sim [p, B; 1-p, C]$
- Monotonicity
  $A \succ B \Rightarrow (p > q \Leftrightarrow [p, A;\ 1-p, B] \succ [q, A;\ 1-q, B])$
- Decomposability
  $\big[p, A; 1-p, [q, B; 1-q, C]\big] \sim [p, A; (1-p)q, B; (1-p)(1-q), C]$

Rational Preferences: Violation

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 preferences, we can derive the following consequences (for the proof see Neumann and Morgenstern, 1944):

- **Existence of Utility Function**: There exists a function U such that

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

- **Expected Utility of a Lottery**: The utility of a lottery is

$$U([p_1, s_1; \ \dots \ ; p_n, s_n]) = \sum_i p_i U(s_i).$$

The preceding theorems establish that a utility function *exists*, but not that it is *unique*. An agent's behavior would not change for the new utility

$$U'(s) = aU(s) + b, \qquad a \in \mathbb{R}^+, b \in \mathbb{R}.$$

Note: in a deterministic setting one often uses the term *value function* or *ordinal utility function* instead of *utility function*.

Tweedback Question

How much would you pay to avoid playing Russian roulette with a million-barreled revolver?

A €10

B €100

C €1,000

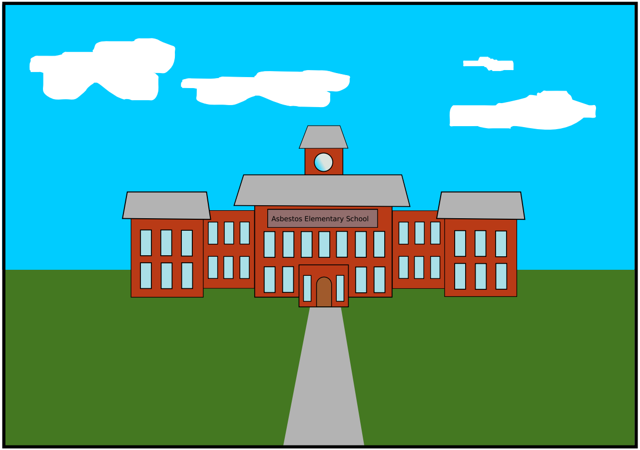
D €10,000

E €100,000

F €1,000,000

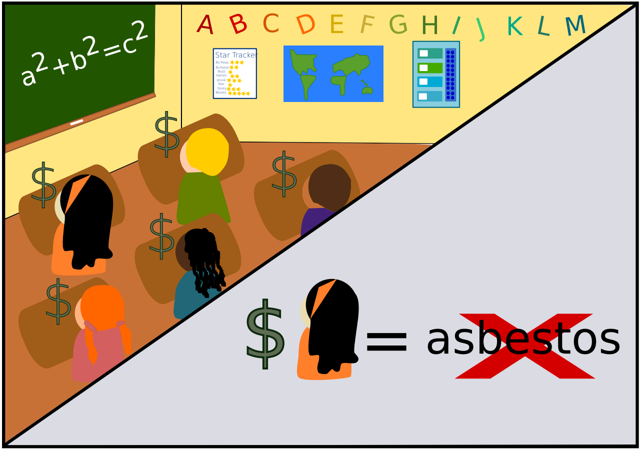
G more than €1,000,000

Utility: Prize on Life (1)



Ross Shachter relates an experience with a government agency that commissioned a study on removing asbestos from schools.

Utility: Prize on Life (2)

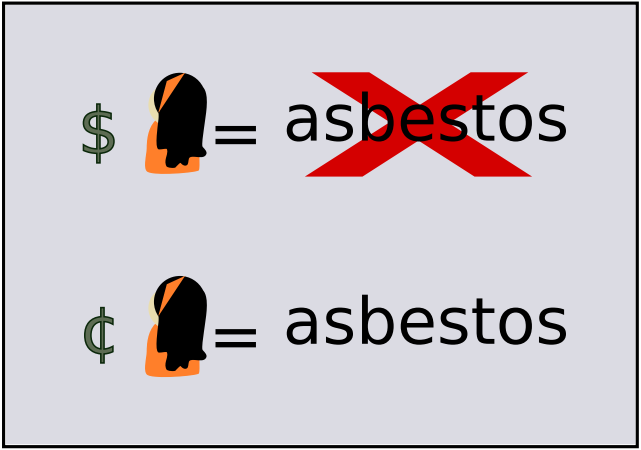


The decision analysts performing the study assumed a particular dollar value for the life of a school-age child, and argued that the rational choice under that assumption was to remove the asbestos.

Utility: Prize on Life (3)

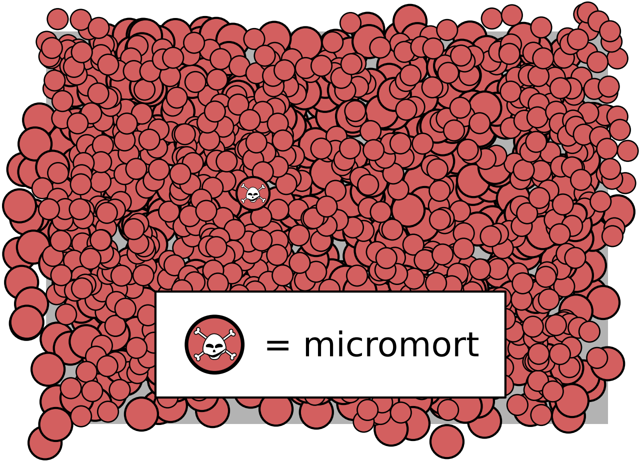
The agency, morally outraged at the idea of setting the value of a life, rejected the report out of hand.

Utility: Prize on Life (4)



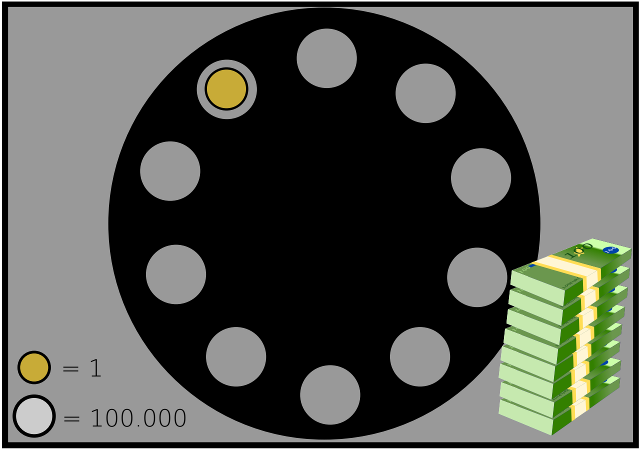
It then decided against asbestos removal – implicitly asserting a lower value for the life of a child than that assigned by the analysts.

Utility: Prize on Life (5)



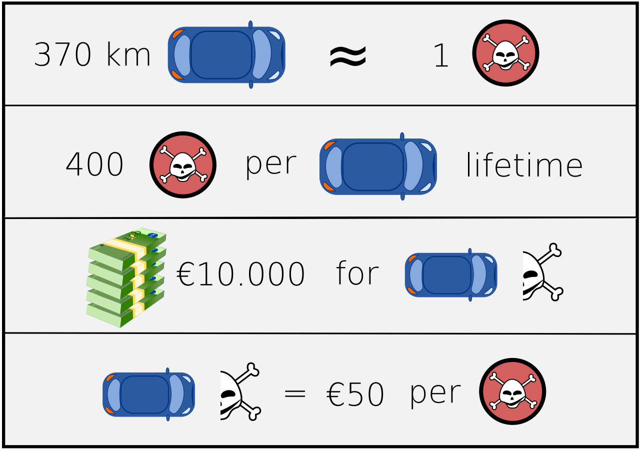
A common "currency" for life used in medical and safety analysis is the **micromort**, a one in a million chance of death.

Utility: Prize on Life (6)



People would pay huge amounts (typically more than €10,000) to avoid playing Russian roulette with a million-barreled revolver.

Utility: Prize on Life (7)



Driving a car for 370 km is approximately a micromort, which are about 400 micromorts for the lifetime of a car. People pay about €10, 000 for a safer car that halves the risk of death corresponding to €50 per micromort.

Tweedback Question

You are a participant of a game show.

The game show master offers you to

A win €1,000,000, or

B flip a coin to potentially win €2,500,000.

What is your preference?

Utility of Money (1)

- Money is an obvious candidate for a utility function due to its versatility.
- Does money behave as a utility function?

Television game show

You have the following choice:

- Win €$1,000,000$, or
- flip a coin to potentially win €$2,500,000$.

Most people would take €$1,000,000$. Is this irrational?

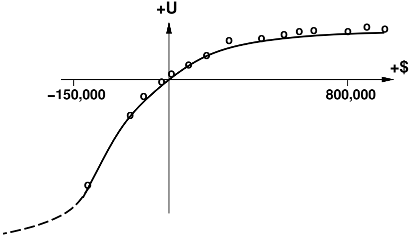
Expected value:
$0.5 \cdot 0 + 0.5 \cdot 2,500,000 =$ €$1,250,000 >$ €$1,000,000$.

Utility of Money (2)

Introduce: $s_n \hat{=}$ possessing €n.

Expected utility: $EU(Accept) = 0.5U(s_k) + 0.5U(s_{k+2,500,000})$,
$$EU(Decline) = U(s_{k+1,000,000})$$

Utility is not directly proportional to money for most individuals, e.g.,



Assume: $U(s_k) = 5$, $U(s_{k+1,000,000}) = 8$, $U(s_{k+2,500,000}) = 9$.

Result: $EU(Accept) = 7$, $EU(Decline) = 8 \rightarrow$ Decision is rational!
Money is not necessarily a utility function.

Multiattribute Utility

- How can we handle utility functions of many attributes $\mathbf{X} = X_1, \ldots, X_n$?
  E.g., what is $U(\mathit{Deaths}, \mathit{Noise}, \mathit{Cost})$ when choosing a site for an airport?
- A complete vector of assignments is denoted by $\mathrm{x} = \langle x_1, \ldots, x_n \rangle$. We assume that higher values correspond to higher utilities.
- How can complex utility functions be assessed from preference behavior?

Solution strategies

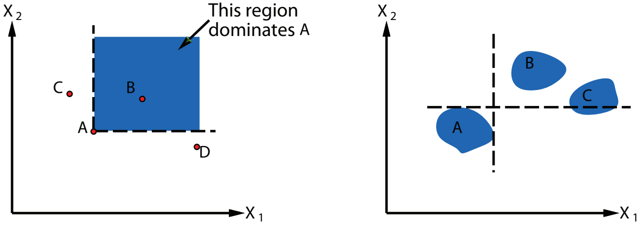
- Idea 1 (dominance): identify conditions under which decisions can be made without complete identification of $U(x_1, \ldots, x_n)$
- Idea 2 (preference structure): identify various types of **independence** in preferences and derive consequent canonical forms for $U(x_1, \ldots, x_n)$

Strict Dominance

Strict dominance

Choice B strictly dominates choice A iff

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

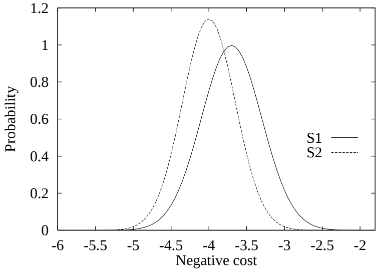


Deterministic attributes | Uncertain attributes

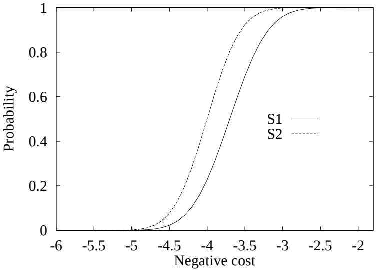
Strict dominance seldom holds in practice.

Stochastic Dominance (1)

probability distribution



cumulative distribution



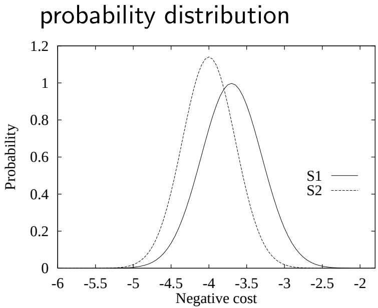
Stochastic dominance

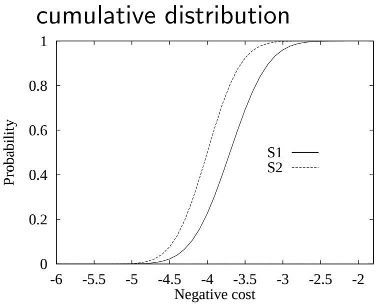
Random variable S_1 stochastically dominates random variable S_2 (corresponding distributions are p_1, p_s) iff

$$\forall\, t \quad \int_{-\infty}^{t} p_1(x)dx \leq \int_{-\infty}^{t} p_2(x)dx, \tag{8}$$

i.e., the cumulative distribution of p_1 is always smaller than that of p_2.

Stochastic Dominance (2)





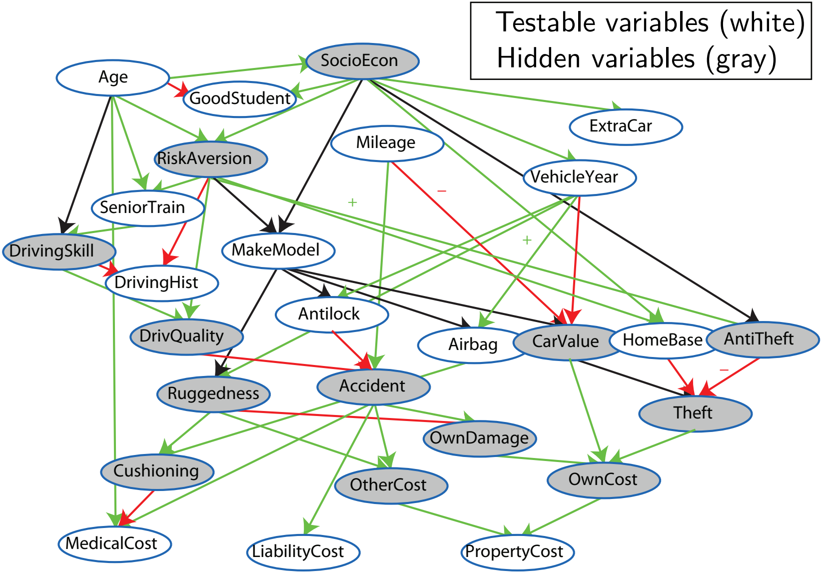
If U is monotonic in x, then the expected utility of S_1 with distribution p_1 is at least as high as the expected utility of S_2 with distribution p_2:

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

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

Example: Cost increases with distance from city $\Rightarrow$ closer airport more likely to meet any given budget limit, i.e. (8) holds $\Rightarrow$ stochastic dominance.

Label the arcs + or –



Preference Structure: Deterministic

Preference independence

Two attributes X_1 and X_2 are **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.

Example: airport problem with $\langle Noise, Cost, Safety \rangle$:
⟨20,000 suffer, €4.6 billion, x deaths/mpm⟩ vs.
⟨70,000 suffer, €4.2 billion, x deaths/mpm⟩ does not depend on x.

Mutual preference independence (Leontief, 1947)

If every pair of attributes is preferentially independent of its complement, then this also applies to every subset of attributes.

Theorem (Debreu, 1960)

Mutual preference independence $\Rightarrow \exists$ **additive** value function:
$V(x_1, \ldots, x_n) = \sum_i V_i(x_i)$

$\rightarrow$ assess n times $V_i(x_i)$; often a good approximation otherwise.

Preference Structure: Stochastic

Consider preferences over lotteries.

Utility independence

A set of attributes X is **utility-independent** of Y iff preferences over lotteries in X do not depend on attributes in Y.

Mutual utility independence

Each subset is utility independent of its complement
$\Rightarrow \exists$ **multiplicative** utility function (Keeney, 1974).
Example: The function for three attributes is

$$U = k_1 U_1 + k_2 U_2 + k_3 U_3 + k_1 k_2 U_1 U_2 + k_2 k_3 U_2 U_3 + k_3 k_1 U_3 U_1 + k_1 k_2 k_3 U_1 U_2 U_3$$

For conciseness, we use U_i to mean $U_i(x_i)$.

Decision Trees

Decision trees are a method for graphically organizing sequential decision processes.

Components of a decision tree:

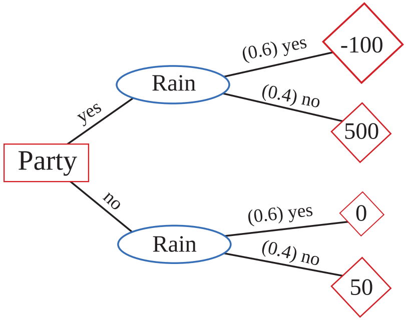
- **Decision nodes** (rectangles): Decision nodes have branches for each alternative decision.
- **Utility nodes** (diamonds): Utility nodes are leaf nodes and represent the utility value of each branch.
- **Chance nodes** (ovals): Chance nodes represent random variables.

Expected utility of any decision: weighted summation of all branches from the decision to all reachable leaves from the decision.

Decision Trees: Example (1)

Should you go ahead with a fund-raising garden party or not?

- If it rains during the party, people will leave and you will lose money.
- If you won't go ahead with the party and it doesn't rain, you can do something else fun.
- The probability of rain is $P(\texttt{rain}) = 0.6$.
- The utilities are:
  $U(\texttt{party}, \texttt{rain}) = -100$,
  $U(\texttt{party}, \texttt{noRain}) = 500$,
  $U(\texttt{noParty}, \texttt{rain}) = 0$,
  $U(\texttt{noParty}, \texttt{noRain}) = 50$.



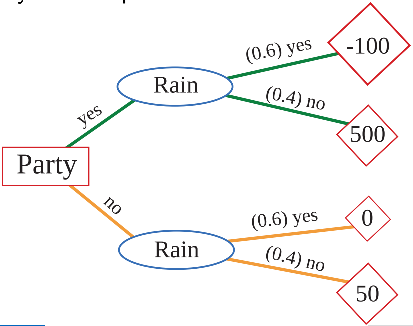
Decision Trees: Example (2)

After using (6) from slide 7 and the fact that $P(r, party) = P(r, noParty) = P(r)$ since Party is a decision variable, we obtain

$$EU(\texttt{party}) = \sum_{r \in \{\texttt{rain}, \texttt{noRain}\}} P(\texttt{r})U(\texttt{r}, \texttt{party}) = 0.6 \cdot (-100) + 0.4 \cdot 500 = 140$$

$$EU(\texttt{noParty}) = \sum_{r \in \{\texttt{rain}, \texttt{noRain}\}} P(\texttt{r})U(\texttt{r}, \texttt{noParty}) = 0.6 \cdot 0 + 0.4 \cdot 50 = 20$$

Thus, throwing the party is the optimal choice.



Decision Trees: Discussion

Benefits:

- General method.
- Explicit encoding of utilities and probabilities associated with each decision and event.
- Especially helpful for small, sequential decision processes.

But:
Representing the tree can become exponentially complex with increasing number of sequential decisions.

Solution:
Decision networks (introduced later) enable a more compact description of the decision problem.

Motivation for Evaluating Information

- Usually, there are costs when carrying out tests to acquire information about the value of a random variable.
- So far, we have not taken into account if it is worth obtaining a specific piece of information.
- One of the most important aspects in decision making is to ask the right questions.

  Example: a doctor has to carefully select the diagnostic tests and questions most important to the patient.

- **Information value theory** guides an agent to choose what information to acquire.

Value of Information: Another Example

Problem

- Oil company buys one of n indistinguishable blocks of ocean-drilling rights.
- One block contains oil worth €C, while the others are worthless.
- The asking price of each block is €C/n.
- A seismologist can tell if oil is in block 3.
- How much should the company pay for this service?

oil : $C - \frac{C}{n}$ $\Leftarrow p = \frac{1}{n}$

no oil : $\frac{1}{n-1} C - \frac{C}{n}$ $\Leftarrow p = \frac{n-1}{n}$

Solution

- With probability $1/n$, the survey indicates oil in block 3. The company buys the block and makes a profit of $C - C/n = (n-1)C/n$.
- With probability $(n-1)/n$, the survey says that there is no oil. Buying another block increases the chances to $1/(n-1)$ so that the expected profit is $C/(n-1) - C/n = C/(n(n-1))$.
- The expected profit is $\frac{1}{n}\frac{(n-1)C}{n} + \frac{n-1}{n}\frac{C}{n(n-1)} = \frac{C}{n}$: Maximum payment for the seismologist should be C/n.

General Formula (1)

Basic idea

expected value of information
$=$ expected value of best action given the information at no charge
$-$ expected value of best action without information.

Value of information

The phrase **value of information** (VOI) refers to the value of evidence of a random variable E_j, that is, we discover $E_j = e_j$.

Given the initial evidence e, the value of the current best action α is

$$MEU(\alpha|\mathrm{e}) = \max_a \sum_{s'} P(\mathtt{Result}(a) = s'|\mathrm{e})U(s')$$

and the value of the new best action α_{e_j} (after new evidence $E_j = e_j$) is

$$MEU(\alpha_{e_j}|\mathrm{e}, e_j) = \max_a \sum_{s'} P(\mathtt{Result}(a) = s'|\mathrm{e}, e_j)U(s')$$

General Formula (2)

E_j is a random variable whose value is *currently* unknown.

To determine the value of discovering E_j, we must average over all possible values e_{jk}, using our *current* beliefs about its value:

$$\boxed{VOI_{\mathrm{e}}(E_j) = \left(\sum_k P(E_j = e_{jk}|\mathrm{e}) MEU(\alpha_{e_{jk}}|\mathrm{e}, E_j = e_{jk})\right) - MEU(\alpha|\mathrm{e}).}$$

General Formula: Oil Example (1)

a_i:	buy rights of block i
$s' \in \{oil, noOil\}$:	state models whether oil has been found

We choose block 1 without loss of generality when no survey is bought:

$$
\begin{aligned}
MEU(\alpha|\mathrm{e}) =& \max_a \sum_{s'} P(\texttt{Result}(a) = s'|\mathrm{e})U(s') \\
=& \sum_{s'} P(\texttt{Result}(a_1) = s'|\mathrm{e})U(s') \\
=& P(\texttt{Result}(a_1) = oil|\mathrm{e})U(oil) \\
& + P(\texttt{Result}(a_1) = noOil|\mathrm{e})U(noOil) \\
=& \frac{1}{n}(C - \frac{C}{n}) + \frac{n-1}{n}(-\frac{C}{n}) = 0
\end{aligned}
$$

General Formula: Oil Example (2)

a_i:	buy rights of block i
$s' \in \{oil, noOil\}$:	state models whether oil has been found
e_1:	oil in block 3
e_2:	no oil in block 3

When there is oil in block 3, we choose block 3:

$$\begin{aligned} MEU(\alpha_{e_1}|\mathrm{e}, e_1) =& \max_a \sum_{s'} P(\mathtt{Result}(a) = s'|\mathrm{e}, e_1)U(s') \\ =& \sum_{s'} P(\mathtt{Result}(a_3) = s'|\mathrm{e}, e_1)U(s') \\ =& P(\mathtt{Result}(a_3) = oil|\mathrm{e}, e_1)U(oil) \\ &+ P(\mathtt{Result}(a_3) = noOil|\mathrm{e}, e_1)U(noOil) \\ =& 1(C - \frac{C}{n}) + 0(-\frac{C}{n}) = C - \frac{C}{n} \end{aligned}$$

General Formula: Oil Example (3)

a_i:	buy rights of block i
$s' \in \{oil, noOil\}$:	state models whether oil has been found
e_1:	oil in block 3
e_2:	no oil in block 3

When there is no oil in block 3, we choose any other block (here: block 1):

$$
\begin{aligned}
MEU(\alpha_{e_2}|\mathrm{e}, e_2) &= \max_a \sum_{s'} P(\texttt{Result}(a) = s'|\mathrm{e}, e_2)U(s') \\
&= \sum_{s'} P(\texttt{Result}(a_1) = s'|\mathrm{e}, e_2)U(s') \\
&= P(\texttt{Result}(a_1) = oil|\mathrm{e}, e_2)U(oil) \\
&\quad + P(\texttt{Result}(a_1) = noOil|\mathrm{e}, e_2)U(noOil) \\
&= \frac{1}{n-1}\left(C - \frac{C}{n}\right) + \frac{n-2}{n-1}\left(-\frac{C}{n}\right)
\end{aligned}
$$

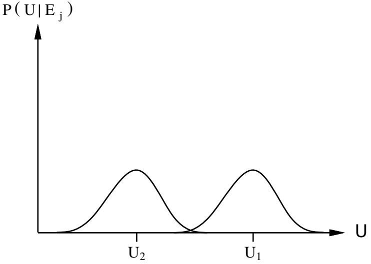
General Formula: Oil Example (4)

Value of information (VOI):

$$
\begin{aligned}
VOI_{\mathrm{e}}(E) &= \left(\sum_k P(E = e_k|\mathrm{e})MEU(\alpha_{e_k}|\mathrm{e}, E = e_k)\right) - \underbrace{MEU(\alpha|\mathrm{e})}_{=0} \\
&= \quad P(E = e_1|\mathrm{e})MEU(\alpha_{e_1}|\mathrm{e}, E = e_1) \\
&\quad + P(E = e_2|\mathrm{e})MEU(\alpha_{e_2}|\mathrm{e}, E = e_2) \\
&= \frac{1}{n}\left(C - \frac{C}{n}\right) + \frac{n-1}{n}\left(\frac{1}{n-1}\left(C - \frac{C}{n}\right) + \frac{n-2}{n-1}\left(-\frac{C}{n}\right)\right) \\
&= \frac{Cn - C}{n^2} + \frac{Cn - C + (n-2)(-C)}{n^2} \\
&= \frac{Cn - C + C}{n^2} = \frac{C}{n}
\end{aligned}
$$

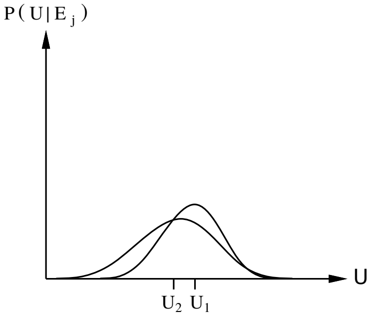
General Formula: Road Example (1)

- Suppose a_1 and a_2 are two different routes through some mountains:
  - a_1 is a straight highway through a low pass.
  - a_2 is a winding dirt road over the top.
- a_1 is clearly preferable although both are likely blocked by avalanches.
- Expected utility U_1 is therefore clearly greater than U_2.
- Satellite reports E_j on road conditions result in new expectations U_1' and U_2'.
- Satellite reports in this case are not worth much since it is unlikely that the new information will change the plan.



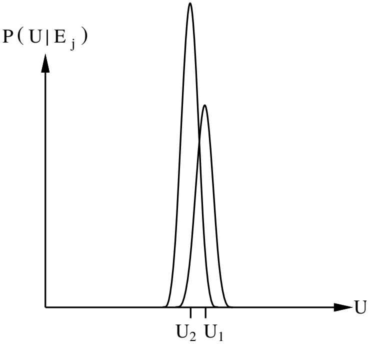
General Formula: Road Example (2)

- Now suppose a_1 and a_2 are similar winding dirt roads, where one is only slightly shorter.
- U_1 and U_2 are quite close, but the utility distributions are fairly broad.
- The difference in utilities will be high, given the information whether a road is blocked or not.
- Satellite reports in this case are very valuable.



General Formula: Road Example (3)

- Now suppose a_1 and a_2 are similar winding dirt roads, where one is only slightly shorter.
- The probability of road blocking is low for both routes.
- U_1 and U_2 are quite close and the utility distributions are fairly narrow.
- Satellite reports in this case are not valuable since the utility difference will be small.



Properties of VOI

- **Nonnegative** – in *expectation*

$$\forall\, \mathrm{e}, E_j \quad VOI_{\mathrm{e}}(E_j) \geq 0$$

- **Nonadditive** – consider, e.g., obtaining E_j twice

$$VOI_{\mathrm{e}}(E_j, E_k) \neq VOI_{\mathrm{e}}(E_j) + VOI_{\mathrm{e}}(E_k)$$

- **Order-independent**

$$VOI_{\mathrm{e}}(E_j, E_k) = VOI_{\mathrm{e}}(E_j) + VOI_{\mathrm{e},e_j}(E_k) = VOI_{\mathrm{e}}(E_k) + VOI_{\mathrm{e},e_k}(E_j)$$

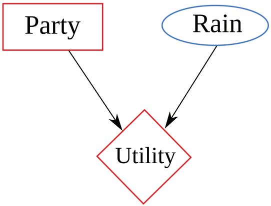
Decision Networks (aka Influence Diagrams)

Problem: Decision trees can become very large.

Idea: Make use of conditional independence as for Bayesian networks.

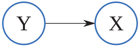
Solution: Add decision nodes and utility nodes to Bayesian networks to enable rational decision making:

- **Decision nodes** (rectangles): Decision maker has a choice of actions.
- **Utility nodes** (diamonds): Represent the agent's utility function, where the parents directly influence the value.
- **Chance nodes** (ovals): Represent random variables as in Bayesian networks.



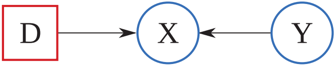
Decision Networks: Syntax (1)

Links to Random Variables:



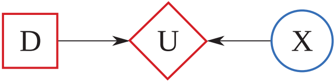
: Random variable X conditionally depends on the value of parental random variable Y.

: Value of random variable X will be revealed as the decision D is taken.



Links to Utility Nodes:

- The utility function depends on the parents of the utility node.
- Parents of the utility node can be random variables and decision nodes.
- We assume that there is at most one utility node in a decision network.

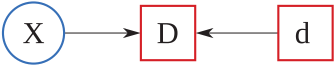


Decision Networks: Syntax (2)

Information Links (links to decision nodes):

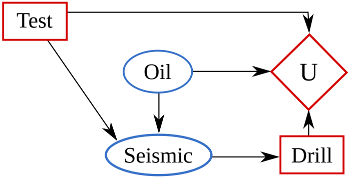
: State of variable X will be known *before* the decision D is taken.

: Decision d is known *before* the decision D is taken.



Example

- Oil company wants to buy ocean-drilling rights.
- First, the company has to decide to carry out a seismic test. Its result is represented by the variable `Seismic`, and depends on whether there is oil.
- Based on this result, the company has to decide whether or not to drill.



Partial Ordering of the Nodes

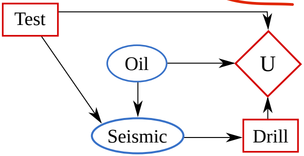
Decision networks define a *partial ordering* (no order implied amongst variables within X_n) of their nodes (exceptions exist):

$$X_0 < D_1 < X_1 < D_2, \dots, X_{n-1} < D_n < X_n,$$

with X_k being the variables revealed between decisions D_k and D_{k+1}.

Obtaining Partial Orders

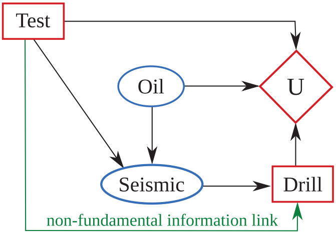
1. Identify the first decision D_1 and all variables X_0 to make that decision.
   Oil-drilling example: `Test`.
2. Identify the next decision D_2 and the variables X_1 that are revealed after decision D_1 and before decision D_2, etc. to obtain $X_0 < D_1 < X_1 < D_2, \dots$
   Oil-drilling example: `Test` < `Seismic` < `Drill`.
3. Place any unrevealed variables at the end of the ordering.
   Oil-drilling example: `Test` < `Seismic` < `Drill` < `Oil`.



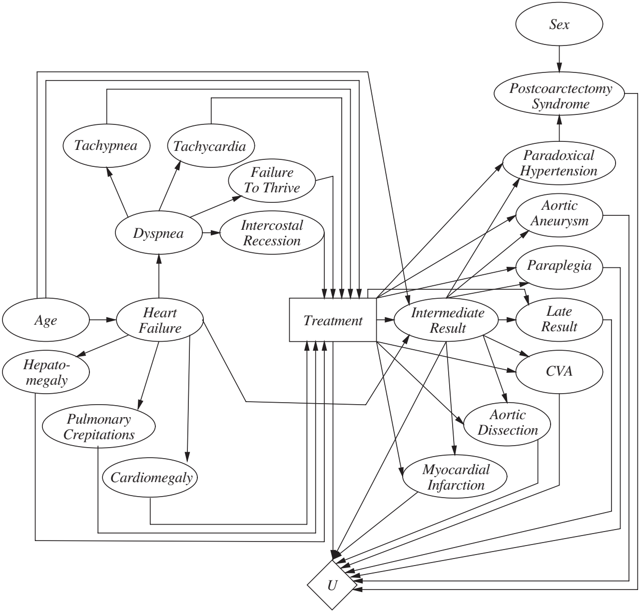
Fundamental Information Links

- An information link is *fundamental*, if its removal would change the partial ordering.
- *No forgetting assumption*: all past decisions and revealed variables are available at the current decision.
- Due to the *no forgetting assumption*, we only need to draw *fundamental* information links.

Example:



Decision Network for Aortic Coarctation



Evaluating Decision Networks through Policies

We have to find a decision maximizing the expected utility for each decision variable:

$$\pi^*(\mathrm{e}) = \arg\max_{d_i} EU(d_i|\mathrm{e}).$$

Given that we have as evidence the previous decisions $d_{1:i-1}$ and revealed chance nodes $x_{1:i-1}$ (random variable X_i is revealed after decision variable D_i), we obtain

$Eu(x) = \sum_y p(x,y) \cdot V(x,y)$

$$\begin{aligned}
&\pi^*(x_{1:i-1}, d_{1:i-1}) \\
&= \arg\max_{d_i} EU(d_i|x_{1:i-1}, d_{1:i-1}) \\
&\overset{(7)}{=} \arg\max_{d_i} \sum_{x_i} \underbrace{P(\mathtt{Result}(d_i) = x_i|x_{1:i-1}, d_{1:i-1})}_{=P(x_i, d_i|x_{1:i-1}, d_{1:i-1})} U(x_{1:i}, d_{1:i}).
\end{aligned}$$

We refer to $\pi^*(x_{1:i-1}, d_{1:i-1})$ as the policy for decision variable d_i.

Evaluating Decision Networks: Backwards Evaluation

Using the partial ordering, we can compute the joint probability as

$$P(x_{1:n}, d_{1:n}) = \underbrace{P(x_{1:n}|d_{1:n})}_{=\frac{P(x_{1:n},d_{1:n})}{P(d_{1:n})=1}} = \prod_{i=1}^{n} P(x_i|\underbrace{x_{1:i-1}, d_{1:i}}_{\text{“parents”}}). \tag{9}$$

Solving a decision network for a utility $U(x_{1:n}, d_{1:n})$ corresponds to finding

$$\begin{aligned} MEU(d_{1:n}) &= \max_{d_1}\sum_{x_1}\dots\max_{d_n}\sum_{x_n} P(x_{1:n}, d_{1:n})U(x_{1:n}, d_{1:n}) \\ &\overset{(9)}{=} \max_{d_1}\sum_{x_1}\dots\max_{d_n}\sum_{x_n}\prod_{i=1}^{n} P(x_i|x_{1:i-1}, d_{1:i})U(x_{1:n}, d_{1:n}). \end{aligned}$$

We solve the problem backwards by constructing a reduced decision network for $X_{1:i-1}$ and $D_{1:i-1}$:

$$\max_{d_1}\sum_{x_1}\dots\max_{d_{n-1}}\sum_{x_{n-1}}\prod_{i=1}^{n-1} P(x_i|x_{1:i-1}, d_{1:i})\underbrace{\max_{d_n}\sum_{x_n} P(x_n|x_{1:n-1}, d_{1:n})U(x_{1:n}, d_{1:n})}_{=:\tilde{U}(x_{1:n-1},d_{1:n-1})}$$

which is a reduced decision diagram for the modified utility $\tilde{U}(x_{1:n-1}, d_{1:n-1})$.

Evaluating Decision Networks: Considering Structure

The reduction has to be performed until the final decision is made.

The presented approach does not yet consider structure. Without loss of generality, we can split the utility into one part that is independent of X_n, D_n and one part that depends on them:

$$U(x_{1:n}, d_{1:n}) = U_a(x_{1:n-1}, d_{1:n-1}) + U_b(x_{1:n}, d_{1:n})$$

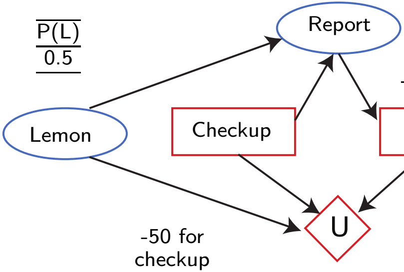
Then the computational effort for updating the utility function when eliminating X_n, D_n is reduced to

$$\begin{aligned}\tilde{U}(x_{1:n}, d_{1:n}) =& U_a(x_{1:n-1}, d_{1:n-1}) \\ & + \max_{d_n} \sum_{x_n} P(x_n|x_{1:n-1}, d_{1:n}) U_b(x_{1:n}, d_{1:n}).\end{aligned}$$

This can be done analogously for other variables and decisions.

Example: Car Buyer

P(L)
0.5



L	C	$P(\text{good}\|L, C)$	$P(\text{bad}\|L, C)$	$P(\text{none}\|L, C)$
T	T	0.2	0.8	0
T	F	0	0	1
F	T	0.9	0.1	0
F	F	0	0	1

B	L	$U(B, L)$
T	T	-600
T	F	1000
F	T	-400
F	F	-400

- You want to buy a used car for your business, which could be a "lemon" (i.e., unreliable); however, without a car you lose business.
- Before deciding to buy it, you can take it to a mechanic for checkup. S/he will give you a report, labelling the car "good" or "bad".
- The checkup costs €50.
- Should you risk buying the car without the report?

Evaluate Last Decision: Buy (1)

Last decision: should we buy the car? Using (7) from slide 7 results in

$$EU(B|C,R) = \sum_{\ell} \underbrace{P(\ell, B|C,R)}_{=P(\ell|C,R)} \underbrace{U(B,\ell,C,R)}_{\substack{=U(B,\ell) \text{ and} \\ -50 \text{ for checkup}}}$$

We use variable elimination as in Bayesian networks to obtain $P(L|C,R)$:

- **Case $C = c$, $R =$ good:**
  $P(L|c,\text{good}) = \alpha P(L,c,\text{good}) = \alpha P(L)P(\text{good}|L,c) = \alpha\langle P(\ell)P(\text{good}|\ell,c), P(\neg\ell)P(\text{good}|\neg\ell,c)\rangle = \alpha\langle 0.5 \cdot 0.2, 0.5 \cdot 0.9\rangle = \langle 0.18, 0.82\rangle$.
- **Case $C = c$, $R =$ bad:**
  $P(L|c,\text{bad}) = \alpha P(L,c,\text{bad}) = \alpha P(L)P(\text{bad}|L,c) = \ldots = \langle 0.89, 0.11\rangle$
- **Case $C = \neg c$, $R =$ none:**
  $P(L|\neg c,\text{none}) = \alpha P(L,\neg c,\text{none}) = \alpha P(L)P(\text{none}|L,\neg c) = \ldots = \langle 0.5, 0.5\rangle$

Evaluate Last Decision: Buy (2)

The expected utilities for buying and not buying are (the cases (c, none), $(\neg c, \text{good})$, and $(\neg c, \text{bad})$ do not occur):

- **Case $C = c$, $R =$ good:**
  $EU(b|c, \text{good}) = P(\ell|c, \text{good})U(b, \ell) + P(\neg\ell|c, \text{good})U(b, \neg\ell) - 50 = 0.18 \cdot (-600) + 0.82 \cdot 1000 - 50 = 662$
  $EU(\neg b|c, \text{good}) = P(\ell|c, \text{good})U(\neg b, \ell) + P(\neg\ell|c, \text{good})U(\neg b, \neg\ell) - 50 = 0.18 \cdot (-400) + 0.82 \cdot (-400) - 50 = -450$
  $\rightarrow \pi^*(B|c, \text{good}) = b$
- **Case $C = c$, $R =$ bad:**
  $EU(b|c, \text{bad}) = \ldots = -474$
  $EU(\neg b|c, \text{bad}) = \ldots = -450$
  $\rightarrow \pi^*(B|c, \text{bad}) = \neg b$
- **Case $C = \neg c$, $R =$ none (no checkup cost):**
  $EU(b|\neg c, \text{none}) = \ldots = 200$
  $EU(\neg b|\neg c, \text{none}) = \ldots = -400$
  $\rightarrow \pi^*(B|\neg c, none) = b.$

Evaluate First Decision: Checkup

First decision: should we check the car? Using (6) from slide 7 results in

$$\underbrace{EU(\pi^*(B|C,r),C)}_{=EU(C)} = \sum_{r,\ell} \underbrace{P(\pi^*(B|C,r),r,\ell,C)}_{=P(r,\ell|C)=P(r|\ell,C)P(\ell|C)} \underbrace{U(\pi^*(B|C,r),r,\ell,C)}_{\substack{=U(\pi^*(B|C,r),\ell)\text{ and}\\ -50\text{ for checkup}}}$$

	$P(R,L\|c)$	$\pi^*(B\|c,R)$	$U(\pi^*(B\|c,R),L)$
good, ℓ	$0.2\cdot 0.5=0.1$	b	$-600-50=-650$
bad, ℓ	$0.8\cdot 0.5=0.4$	$\neg b$	$-400-50=-450$
good, $\neg\ell$	$0.9\cdot 0.5=0.45$	b	$1000-50=950$
bad, $\neg\ell$	$0.1\cdot 0.5=0.05$	$\neg b$	$-400-50=-450$

$EU(c) = (\textit{use combinations of table above}) =$
$0.1\cdot(-650)+0.4\cdot(-450)+0.45\cdot 950+0.05\cdot(-450)=160$

$EU(\neg c) = P(\ell|\neg c,\text{none})U(b,\ell)+P(\neg\ell|\neg c,\text{none})U(b,\neg\ell)=$
$0.5\cdot(-600)+0.5\cdot 1000=200$

$\rightarrow \pi^*(C)=\neg c.$

Value of Information for the Car Buyer

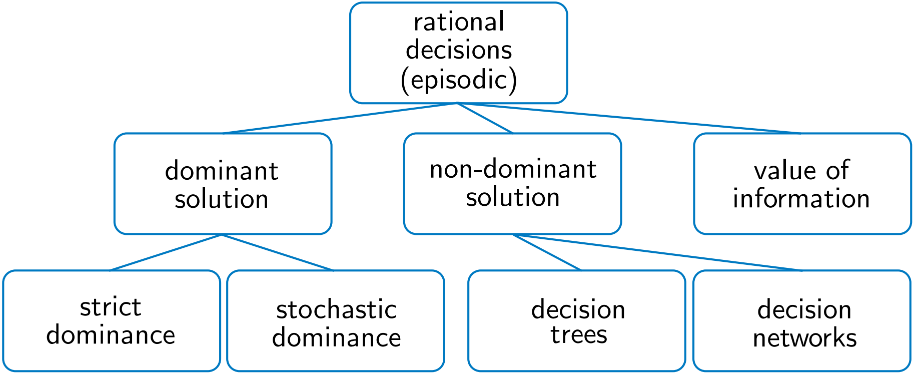
- So optimal policy is: no checkup, buy the car.
- The value of information (see slide 37) directly follows from intermediate results:

$$VOI_{\mathrm{e}}(E_j) = \underbrace{\left(\sum_k P(E_j = e_{jk}|\mathrm{e}) MEU(\alpha_{e_{jk}}|\mathrm{e}, E_j = e_{jk})\right)}_{=EU(c)+50} - \underbrace{MEU(\alpha|\mathrm{e})}_{=EU(\neg c)}$$

$$= 210 - 200 = 10.$$

- When the checkup would cost less than €10, choosing a checkup would result in a better expected utility.

Overview of Making Rational Decisions in Episodic Environments



Summary

- **Decision theory** puts together probability theory and utility theory.
- Utility theory shows that an agent with consistent preferences can be described as possessing a utility function.
- A **rational agent** selects actions that maximize the expected utility.
- **Stochastic dominance** helps making unambiguous decisions.
- **Decision trees** and **decision networks** provide a simple formalism for expressing and solving decision problems.
- The **value of information** supports the decision for gathering more information or not.