Decision Graphs - Influence Diagrams
Descriptive Decision Theory tries to simulate human behavior in finding the right or best decision for a given problem

Example:

- Company can choose one of two places for a new store
- Option 1: 125,000 EUR profit per year
- Option 2: 150,000 EUR profit per year

Company should take Option 2, because it maximized the profit.
Decisions under Uncertainty

In real world not every thing is known, so there are uncertainties in the model.

Example:
- There are plans for restructure the local traffic, which changes the predicted profit.
- Option 1: 125.000 EUR profit per year
- Option 2: 80.000 EUR profit per year

With modification Option 1 is the better one and without modification Option 2 is the better one.

To model these variations in the environment we use so called Decision Tables.

<table>
<thead>
<tr>
<th></th>
<th>( z_1 ) (no modification)</th>
<th>( z_2 ) (restructure)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_1 ) (Option 1)</td>
<td>( 125.000 = e_{11} )</td>
<td>( 125.000 = e_{12} )</td>
</tr>
<tr>
<td>( a_2 ) (Option 2)</td>
<td>( 150.000 = e_{21} )</td>
<td>( 80.000 = e_{22} )</td>
</tr>
</tbody>
</table>
In many cases probabilities could be assigned to each option

**Objective Probabilities** based on mathematic or statistic background

**Subjective Probabilities** based on intuition or estimations

Example:
- The management estimates the probability for the restructure to 30%

The decision can be chosen by expectation value

<table>
<thead>
<tr>
<th></th>
<th>$z_1$ (no modification)</th>
<th>$z_2$ (restructure)</th>
<th>Expectation Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1 = 0.7$</td>
<td>$p_2 = 0.3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_1$ (Option 1)</td>
<td>125.000 = $e_{11}$</td>
<td>125.000 = $e_{12}$</td>
<td>125.000</td>
</tr>
<tr>
<td>$a_2$ (Option 2)</td>
<td>150.000 = $e_{21}$</td>
<td>80.000 = $e_{22}$</td>
<td>129.000</td>
</tr>
</tbody>
</table>

Option 2 has the higher expectation value and should be used
An alternative $a_1$ dominates $a_2$ if the value of $a_1$ is always greater of (or equal to) the value of $a_2$

$\forall_j e_{1j} \geq e_{2j}$

Example:

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>150.000 = $e_{11}$</td>
<td>90.000 = $e_{12}$</td>
</tr>
<tr>
<td>$a_2$</td>
<td>125.000 = $e_{21}$</td>
<td>80.000 = $e_{22}$</td>
</tr>
</tbody>
</table>

Alternative $a_2$ could be dropped
Some more alternatives:

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>$z_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>0</td>
<td>20</td>
<td>10</td>
<td>60</td>
<td>25</td>
</tr>
<tr>
<td>$a_2$</td>
<td>-20</td>
<td>80</td>
<td>10</td>
<td>10</td>
<td>60</td>
</tr>
<tr>
<td>$a_3$</td>
<td>20</td>
<td>60</td>
<td>20</td>
<td>60</td>
<td>50</td>
</tr>
<tr>
<td>$a_4$</td>
<td>55</td>
<td>40</td>
<td>60</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>$a_5$</td>
<td>50</td>
<td>10</td>
<td>30</td>
<td>5</td>
<td>20</td>
</tr>
</tbody>
</table>

- $a_3$ dominated $a_1$
- $a_4$ dominated $a_5$

Alternatives $a_1$ and $a_5$ could be dropped
Probability Domination

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>0.3</td>
<td>0.2</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>$a_1$</td>
<td>20</td>
<td>40</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>$a_2$</td>
<td>60</td>
<td>30</td>
<td>50</td>
<td>20</td>
</tr>
</tbody>
</table>

Probability Domination means that the cumulated probability for the payout for is always higher.

**Algorithm:**
- Order payout by value in a decreasing order
- Cumulate probabilities

**Example:**
- $a_1$ : 50(0.1) 40(0.2) 20(0.3) 10(0.4)
- $a_2$ : 60(0.3) 50(0.4) 30(0.2) 20(0.1)
Probability Domination

Example:

- $a_1$: 50(0.1) 40(0.2) 20(0.3) 10(0.4)
- $a_2$: 60(0.3) 50(0.4) 30(0.2) 20(0.1)

$a_2$ dominates $a_1$. 
### Multi Criteria Decisions - Example

<table>
<thead>
<tr>
<th>Sales $e_1$</th>
<th>Profit $e_2$</th>
<th>Environment Pollution $e_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>800</td>
<td>7000</td>
</tr>
<tr>
<td>$a_2$</td>
<td>600</td>
<td>7000</td>
</tr>
<tr>
<td>$a_3$</td>
<td>400</td>
<td>6000</td>
</tr>
<tr>
<td>$a_4$</td>
<td>200</td>
<td>4000</td>
</tr>
</tbody>
</table>

### Efficient Alternatives
- Only focus on alternatives which are not dominated by others
- Example: Drop $a_4$

### Finding a decision
- If multiple alternatives are effective we need an algorithm to choose the preferred one
- Simplest algorithm: Chose one target (most important, alphabetical) and optimize for this value
Goal find a function $U(e_1, e_2, \ldots, e_n)$ as a combination of all targets, which could be optimized

**Linear combination**
- Simplest variant: Linear combination of all targets
- $U(e_1, e_2, \ldots, e_i) = \sum_{i=1}^{n} \omega_i \cdot e_i$

**Example**
- $\omega_1 = 10$, $\omega_2 = 1$, $\omega_3 = 500$

<table>
<thead>
<tr>
<th></th>
<th>Sales $e_1$</th>
<th>Profit $e_2$</th>
<th>Environment Pollution $e_3$</th>
<th>$U(e_1, e_2, e_3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>800</td>
<td>7000</td>
<td>-4</td>
<td><strong>13000</strong></td>
</tr>
<tr>
<td>$a_2$</td>
<td>600</td>
<td>7000</td>
<td>-2</td>
<td>12000</td>
</tr>
<tr>
<td>$a_3$</td>
<td>400</td>
<td>6000</td>
<td>0</td>
<td>10000</td>
</tr>
</tbody>
</table>
Think about, how you would decide!

**Decision Rules**
- Maximin - Rule
- Maximax - Rule
- Hurwicz - Rule
- Savage-Niehans - Rule
- Laplace - Rule
Maximin - Rule

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>60</td>
<td>30</td>
<td>50</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>$a_2$</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>140</td>
<td>10</td>
</tr>
<tr>
<td>$a_3$</td>
<td>-30</td>
<td>100</td>
<td>120</td>
<td>130</td>
<td>-30</td>
</tr>
</tbody>
</table>

Chose the one with the highest minimum

**Contra:** To pessimistic, only focus on one column

Example

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>0.99</td>
<td>1,000,000</td>
<td>0.99</td>
</tr>
<tr>
<td>$a_2$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Maximax - Rule

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>60</td>
<td>30</td>
<td>50</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>$a_2$</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>140</td>
<td><strong>140</strong></td>
</tr>
<tr>
<td>$a_3$</td>
<td>-30</td>
<td>100</td>
<td>120</td>
<td>130</td>
<td>130</td>
</tr>
</tbody>
</table>

Chose the one with the highest maximum

**Contra**: To optimistic, only focus on one column

Example

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
</tr>
<tr>
<td>$a_2$</td>
<td>1,000,001</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td><strong>1,000,001</strong></td>
</tr>
</tbody>
</table>
Hurwicz - Rule

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>Max</th>
<th>Min</th>
<th>$\Phi(a_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>60</td>
<td>30</td>
<td>50</td>
<td>60</td>
<td>60</td>
<td>30</td>
<td>0.4 \cdot 60 + 0.6 \cdot 30 = 42</td>
</tr>
<tr>
<td>$a_2$</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>140</td>
<td>140</td>
<td>10</td>
<td>0.4 \cdot 140 + 0.6 \cdot 10 = 62</td>
</tr>
<tr>
<td>$a_3$</td>
<td>-30</td>
<td>100</td>
<td>120</td>
<td>130</td>
<td>130</td>
<td>-30</td>
<td>0.4 \cdot 130 + 0.6 \cdot (-30) = 34</td>
</tr>
</tbody>
</table>

Combination of Maximin and Maximax - Rule

$\Phi(a) = \lambda \cdot \max(e_i) + (1 - \lambda) \cdot \min(e_i)$

$\lambda$ represents readiness to assume risk

**Contra:** Only focus on two column

Example ($\min(a_1) < \min(a_2), \max(a_1) < \max(a_2) \Rightarrow$ chose $a_2$)

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>0.99</td>
<td>1,000,000</td>
<td>0.99</td>
</tr>
<tr>
<td>$a_2$</td>
<td>1,000,001</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1,000,001</td>
<td>1</td>
</tr>
</tbody>
</table>
Savage-Niehans - Rule

Algorithm:
- Find the maximal value for every column
- Subtract value from maximal value
- Use alternative with the lowest regret

Regret Table:

<table>
<thead>
<tr>
<th></th>
<th>( z_1 )</th>
<th>( z_2 )</th>
<th>( z_3 )</th>
<th>( z_4 )</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_1 )</td>
<td>60 - 60 = 0</td>
<td>70</td>
<td>70</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>( a_2 )</td>
<td>60 - 10 = 50</td>
<td>90</td>
<td>110</td>
<td>0</td>
<td>110</td>
</tr>
<tr>
<td>( a_3 )</td>
<td>60 - (-30) = 90</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>90</td>
</tr>
</tbody>
</table>
Another example

we chose $a_1$

Regret Table:

\[
\begin{array}{cccccc}
\text{ } & z_1 & z_2 & z_3 & z_4 & \text{Max} \\
\hline
a_1 & 1 & 0 & 0 & 0 & 1 \\
a_2 & 0 & 1,000,000 & 1,000,000 & 1,000,000 & 1,000,000 \\
\end{array}
\]
Same example, but we add alternative $a_3$

Now we chose $a_2$

Regret Table:

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>1,999,000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1,999,000</td>
</tr>
<tr>
<td>$a_2$</td>
<td>1,998,999</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,998,999</td>
</tr>
<tr>
<td>$a_3$</td>
<td>0</td>
<td>2,000,000</td>
<td>2,000,000</td>
<td>2,000,000</td>
<td>2,000,000</td>
</tr>
</tbody>
</table>
Chose the one with the highest mean value

**Contra:**
- Not every condition has the same probability
- Duplication of one condition could change the result

Most people would also chose $a_3$ in this example
The following axioms should be fulfilled by the rules

**Addition to a column**
The decision should not be changed, if a fixed value is added to a column

**Additional rows**
The preference relation between two alternatives should not be changed, if a new row is added

**Domination**
If $a_1$ dominates $a_2$, $a_2$ could not be optimal

**Join of equal columns**
The preference relation between two alternatives should not change, if two columns with the same outcomes are joined to a common column
### Decision Rules Conclusion

<table>
<thead>
<tr>
<th>Rule</th>
<th>Example Result</th>
<th>Addition to a row</th>
<th>Additional Rows</th>
<th>Domination</th>
<th>Join of equal Rows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximin</td>
<td>$a_1$</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Maximax</td>
<td>$a_2$</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Hurwicz</td>
<td>$a_2$</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Savage-Niehans</td>
<td>$a_1$</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Laplace</td>
<td>$a_3$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

No Rule fulfills all axioms $\Rightarrow$ no perfect rule

Common usage: Remove duplicate Columns and use Laplace

Better: Define subjective probabilities and use them
A *preference ordering* $\succeq$ is a ranking of all possible states of affairs (worlds) $S$

- these could be outcomes of actions, truth assignments, states in a search problem, etc.

- $s \succeq t$: means that state $s$ is *at least as good as* $t$

- $s \succ t$: means that state $s$ is *strictly preferred to* $t$

We insist that $\succeq$ is

- reflexive: i.e., $s \succeq s$ for all states $s$

- transitive: i.e., if $s \succeq t$ and $t \succeq w$, then $s \succeq w$

- connected: for all states $s,t$, either $s \succeq t$ or $t \succeq s$
Preference Orderings

Note that transitivity is not always given in decision making.

Consider the following set of dice (Efron Dice):

- Die A has sides: 2, 2, 4, 4, 9, 9
- Die B has sides: 1, 1, 6, 6, 8, 8
- Die C has sides: 3, 3, 5, 5, 7, 7

The probability that A rolls a higher number than B, the probability that B rolls higher than C, and the probability that C rolls higher than A are all \( \frac{5}{9} \), so this set of dice is nontransitive. In fact, it has the even stronger property that, for each die in the set, there is another die that rolls a higher number than it more than half the time.
Why Impose These Conditions?

Structure of preference ordering imposes certain “rationality requirements” (it is a weak ordering)

E.g., why transitivity?
- Suppose you (strictly) prefer coffee to tea, tea to OJ, OJ to coffee
- If you prefer X to Y, you will trade me Y plus $1 for X
- I can construct a “money pump” and extract arbitrary amounts of money from you
Rather than just ranking outcomes, we are often able to quantify our degree of preference

A utility function $U : S \rightarrow \mathbb{R}$ associates a real-valued utility with each outcome.
- $U(s)$ measures the degree of preference for $s$

Note: $U$ induces a preference ordering $\succeq_U$ over $S$ defined as: $s \succeq_U t$ iff $U(s) \geq U(t)$
- $\succeq_U$ will be reflexive, transitive, connected
Under conditions of uncertainty, each decision $d$ induces a distribution $Pr_d$ over possible outcomes
- $Pr_d(s)$ is probability of outcome $s$ under decision $d$

The *expected utility* of decision $d$ is defined

$$EU(d) = \sum_{s \in S} Pr_d(s)U(s)$$

The *principle of maximum expected utility (MEU)* states that the optimal decision under conditions of uncertainty is that with the greatest expected utility.
A decision problem under uncertainty is:

- a set of decisions $D$
- a set of outcomes or states $S$
- an outcome function $Pr : D \rightarrow \Delta(S)$
  
  $\Delta(S)$ is the set of distributions over $S$ (e.g., $Pr_d$)
  
  - a utility function $U$ over $S$

A solution to a decision problem under uncertainty is any $d^* \in D$ such that $EU(d^*) \geq EU(d)$ for all $d \in D$
Where do utilities come from?
- underlying foundations of utility theory tightly couple utility with action/choice
- a utility function can be determined by asking someone about their preferences for actions in specific scenarios (or “lotteries” over outcomes)

Utility functions needn’t be unique
- if I multiply $U$ by a positive constant, all decisions have same relative utility
- if I add a constant to $U$, same thing
- $U$ is unique up to positive affine transformation
Complications

Outcome space is large
  ○ like all of our problems, states spaces can be huge
  ○ don’t want to spell out distributions like $Pr_d$ explicitly
  ○ Solution: Bayes nets (or related: influence diagrams)

Decision space is large
  ○ usually our decisions are not one-shot actions
  ○ rather they involve sequential choices (like plans)
  ○ if we treat each plan as a distinct decision, decision space is too large to handle directly
  ○ Solution: use dynamic programming methods to construct optimal plans (actually generalizations of plans, called policies... like in game trees)
**Decision Networks**

*Decision networks* (also known as *influence diagrams*) provide a way of representing sequential decision problems

- basic idea: represent the variables in the problem as you would in a BN
- add decision variables – variables that you “control”
- add utility variables – how good different states are
Sample Decision Network
Decision Networks: Chance Nodes

**Chance nodes**
- random variables, denoted by circles
- as in a BN, probabilistic dependence on parents

\[
\begin{align*}
Pr(\text{flu} | \text{flu}) &= 0.5 \\
Pr(\text{flu} | \text{mal}) &= 0.3 \\
Pr(\text{flu} | \text{none}) &= 0.05 \\
Pr(\text{flu}) &= 0.3 \\
Pr(\text{mal}) &= 0.1 \\
Pr(\text{none}) &= 0.6
\end{align*}
\]

\[
\begin{align*}
Pr(\text{pos} | \text{flu}, \text{bt}) &= 0.2 \\
Pr(\text{neg} | \text{flu}, \text{bt}) &= 0.8 \\
Pr(\text{null} | \text{flu}, \text{bt}) &= 0 \\
Pr(\text{pos} | \text{mal}, \text{bt}) &= 0.9 \\
Pr(\text{neg} | \text{mal}, \text{bt}) &= 0.1 \\
Pr(\text{null} | \text{mal}, \text{bt}) &= 0 \\
Pr(\text{pos} | \text{no}, \text{bt}) &= 0.1 \\
Pr(\text{neg} | \text{no}, \text{bt}) &= 0.9 \\
Pr(\text{null} | \text{no}, \text{bt}) &= 0 \\
Pr(\text{pos} | \neg \text{D}, \neg \text{bt}) &= 0 \\
Pr(\text{neg} | \neg \text{D}, \neg \text{bt}) &= 0 \\
Pr(\text{null} | \neg \text{D}, \neg \text{bt}) &= 1
\end{align*}
\]
Decision networks

- variables decision maker sets, denoted by squares
- parents reflect *information available* at time decision is to be made

In example decision node: the actual values of Chills and Fever will be observed before the decision to take test must be made
- agent can make different decisions for each instantiation of parents (i.e., policies)
Value node

- specifies utility of a state, denoted by a diamond
- utility depends *only on state of parents* of value node
- generally: only one value node in a decision network

Utility depends only on disease and drug

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>U(fludrug, flu)</td>
<td>20</td>
</tr>
<tr>
<td>U(fludrug, mal)</td>
<td>-300</td>
</tr>
<tr>
<td>U(fludrug, none)</td>
<td>-5</td>
</tr>
<tr>
<td>U(maldrug, flu)</td>
<td>-30</td>
</tr>
<tr>
<td>U(maldrug, mal)</td>
<td>10</td>
</tr>
<tr>
<td>U(maldrug, none)</td>
<td>-20</td>
</tr>
<tr>
<td>U(no drug, flu)</td>
<td>-10</td>
</tr>
<tr>
<td>U(no drug, mal)</td>
<td>-285</td>
</tr>
<tr>
<td>U(no drug, none)</td>
<td>30</td>
</tr>
</tbody>
</table>
Decision nodes are totally ordered
- decision variables $D_1, D_2, \ldots, D_n$
- decisions are made in sequence
- e.g., BloodTst (yes,no) decided before Drug (fd,md,no)

**No-forgetting property**
- any information available when decision $D_i$ is made is available when decision $D_j$ is made (for $i < j$)
- thus all parents of $D_i$ are parents of $D_j$
Let $Par(D_i)$ be the parents of decision node $D_i$

- $Dom(Par(D_i))$ is the set of assignments to parents

A policy $\delta$ is a set of mappings $\delta_i$, one for each decision node $D_i$

- $\delta_i : Dom(Par(D_i)) \rightarrow (D_i)$

- $\delta_i$ associates a decision with each parent assignment for $D_i$

For example, a policy for BT might be:

$\delta_{BT}(c, f) = bt$
$\delta_{BT}(c, \sim f) = \sim bt$
$\delta_{BT}(\sim c, f) = bt$
$\delta_{BT}(\sim c, \sim f) = \sim bt$
Value of a policy $\delta$ is the expected utility given that decision nodes are executed according to $\delta$

Given associates $x$ to the set $X$ of all chance variables, let $\delta(x)$ denote the assignment to decision variables dictated by $\delta$

- e.g., assigned to $D_1$ determined by it’s parents’ assignment in $x$
- e.g., assigned to $D_2$ determined by it’s parents’ assignment in $x$ along with whatever was assigned to $D_1$
- etc.

Value of $\delta$:

$$EU(\delta) = \sum_{X} P(X, \delta(X))U(X, \delta(X))$$

An optimal policy is a policy $\delta^*$ such that $EU(\delta^*) \geq EU(\delta)$ for all policies $\delta$
Computing the Best Policy

We can work backwards as follows

First compute optimal policy for Drug (last decision)
- for each assignment to parents (C,F,BT,TR) and for each decision value (D = md,fd,none), compute the expected value of choosing that value of D
- set policy choice for each value of parents to be the value of D that has max value
- eg: $\delta_D(c, f, bt, pos) = md$
Next compute policy for BT given policy $\delta_D(C, F, BT, TR)$ just determined for Drug.

- since $\delta_D(C, F, BT, TR)$ is fixed, we can treat Drug as a normal random variable with deterministic probabilities.

- i.e., for any instantiation of parents, value of Drug is fixed by policy $\delta_D$.

- this means we can solve for optimal policy for BT just as before.

- only uninstantiated variables are random variables (once we fix its parents).
You want to buy a used car, but there’s a good chance it is a “lemon” (i.e., prone to breakdown). Before deciding to buy it, you can take it to a mechanic for inspection. S/he will give you a report on the car, labelling it either “good” or “bad”. A good report is positively correlated with the car being sound, while a bad report is positively correlated with the car being a lemon.

The report costs $50 however. So you could risk it, and buy the car without the report.

Owning a sound car is better than having no car, which is better than owning a lemon.
Car Buyer’s Network

Bayesian Networks

Report: good, bad, none

<table>
<thead>
<tr>
<th></th>
<th>g</th>
<th>b</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>l i</td>
<td>0.2</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>~l i</td>
<td>0.9</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>l ~i</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>~l ~i</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Utility for Lemon and Buy

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>~l</th>
</tr>
</thead>
<tbody>
<tr>
<td>buy l</td>
<td>-600</td>
<td></td>
</tr>
<tr>
<td>buy ~l</td>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>~buy l</td>
<td>-300</td>
<td></td>
</tr>
<tr>
<td>~buy ~l</td>
<td>-300</td>
<td></td>
</tr>
</tbody>
</table>

-50 if inspect

\[ \sim = \text{not} \]

\[ \begin{array}{cc}
 l & \sim l \\
 0.5 & 0.5 \\
\end{array} \]
Evaluate Last Decision: Buy (1)

\[ EU(B|I, R) = \sum_L P(L|I, R, B)U(L, B) \]

\( I = i, R = g: \)

\[ EU(\text{buy}) = P(l|i, g)U(l, \text{buy}) + P(\sim l|i, g)U(\sim l, \text{buy}) - 50 \]
\[ = 0.18 \cdot (-600) + 0.82 \cdot 1000 - 50 = 662 \]
\[ EU(\sim \text{buy}) = P(l|i, g)U(l, \sim \text{buy}) + P(\sim l|i, g)U(\sim l, \sim \text{buy}) - 50 \]
\[ = -300 - 50 = -350 (-300 \text{ indep. of lemon}) \]

So optimal \( \delta_{\text{Buy}}(i, g) = \text{buy} \)

\( I = i, R = b: \)

\[ EU(\text{buy}) = P(l|i, b)U(l, \text{buy}) + P(\sim l|i, b)U(\sim l, \text{buy}) - 50 \]
\[ = 0.89 \cdot (-600) + .11 \cdot 1000 - 50 = -474 \]
\[ EU(\sim \text{buy}) = P(l|i, b)U(l, \sim \text{buy}) + P(\sim l|i, b)U(\sim l, \sim \text{buy}) - 50 \]
\[ = -300 - 50 = -350 (-300 \text{ indep. of lemon}) \]

So optimal \( \delta_{\text{Buy}}(i, b) = \sim \text{buy} \)
Evaluate Last Decision: Buy (2)

\( I = \sim i, R = n \) (note: no inspection cost subtracted):

\[
EU(\text{buy}) = P(l| \sim i, n)U(l, \text{buy}) + P(\sim l| \sim i, n)U(\sim l, \text{buy})
\]
\[
= 0.5 \cdot (-600) + 0.5 \cdot 1000 = 200
\]
\[
EU(\sim \text{buy}) = P(l| \sim i, n)U(l, \sim \text{buy}) + P(\sim l| \sim i, n)U(\sim l, \sim \text{buy}) - 50
\]
\[
= -300 - 50 = -350 (-300 \text{ indep. of lemon})
\]

So optimal \( \delta_{Buy}(\sim i, g) = \text{buy} \)

So optimal policy for Buy is:

- \( \delta_{Buy}(i, g) = \text{buy}; \delta_{Buy}(i, b) = \sim \text{buy}; \delta_{Buy}(\sim i, g) = \text{buy} \)

Note: we don’t bother computing policy for \((i, \sim g), (\sim i, g), \text{ or } (\sim i, b)\), since these occur with probability 0
Evaluate First Decision: Inspect

\[ EU(I) = \sum_{L,R} P(L, R|I)U(L, \delta_{\text{Buy}}(I, R)) , \]

where \( P(R, L|I) = P(R|L, I)P(L|I) \)

\[ EU(i) = 0.1 \cdot (-650) + 0.4 \cdot (-300) + 0.45 \cdot 1000 + 0.05 \cdot (-300) - 50 \]
\[ = 187.5 \]
\[ EU(\sim i) = P(l| \sim i, n)U(l, \text{buy}) + P(\sim l| \sim i, n)U(\sim l, \text{buy}) \]
\[ = .5 \cdot -600 + .5 \cdot 1000 = 200 \]

So optimal \( \delta_{\text{Inspect}}(\sim i) = \text{buy} \)

| \( P(R, L|I) \) | \( \delta_{\text{Buy}} \) | \( U(L, \delta_{\text{Buy}}) \) |
|----------------|----------------|-------------------|
| \( g, l \)     | 0.1            | \text{buy}        | -600 − 50 = -650 |
| \( g, \sim l \) | 0.45           | \text{buy}        | 1000 − 50 = 950  |
| \( b, l \)     | 0.4            | \sim \text{buy}   | -300 − 50 = -350 |
| \( b, \sim l \) | 0.05           | \sim \text{buy}   | -300 − 50 = -350 |
So optimal policy is: don’t inspect, buy the car

- $EU = 200$

- Notice that the EU of inspecting the car, then buying it iff you get a good report, is 237.5 less the cost of the inspection (50). So inspection not worth the improvement in EU.

- But suppose inspection cost $25: then it would be worth it ($EU = 237.5 - 25 = 212.5 > EU(\sim i)$)

- The expected value of information associated with inspection is 37.5 (it improves expected utility by this amount ignoring cost of inspection). How? Gives opportunity to change decision ($\sim buy$ if bad).

- You should be willing to pay up to $37.5 for the report