VU Einführung in Artificial Intelligence

SS 2024

Hans Tompits

Institut für Logic and Computation
Forschungsbereich Wissensbasierte Systeme

www.kr.tuwien.ac.at

Making Simple Decisions

Decision Theory

- Decision-theoretic agent:
  - combines *utility theory* with *probability theory*
  - makes rational decisions based on *beliefs* and *desires* in contexts of uncertainty and conflicting goals
  - has a continuous measure of outcome quality
    - ➥ in contrast to *goal-based agents* that have only a binary distinction between good (goal) and bad (non-goal) states.
- Decision theory:
  - In its simplest form, deals with choosing among actions based on the desirability of their *immediate outcomes*.
  - Thereby, the environment is assumed to be *episodic*, i.e.,
    - an agent's experience can be divided into atomic episodes such that succeeding episodes do not depend on actions taken in previous episodes.
    - ➥ This is in contrast to *sequential* environments, where current decisions influence future decisions.

Outcomes and Utilities

- We furthermore deal with *nondeterministic, partially observable environments*.
  - Possible outcome states are represented in terms of *random variables*:
    - $\text{RESULT}(a)$ denotes a random variable whose values are the possible outcome states for taking action a.
  - The probability of outcome s', given evidence observations $\mathbf{e}$, is written as

$$P(\text{RESULT}(a) = s' | a, \mathbf{e}),$$

    where a stands for the event that action a is executed.

- The agent's preferences are expressed by a *utility function* $U(s)$
  - assigns a single number to a state to express its desirability

Expected Utility

- The *expected utility* of an action a given evidence $\mathbf{e}$, denoted $EU(a|\mathbf{e})$, is the average utility value of the outcomes, weighted by the probability that the outcome occurs:

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

- *Principle of maximum expected utility* (MEU):
  - a rational agent should choose the action that maximises the agent's expected utility:

$$action = \underset{a}{argmax}\ EU(a|\mathbf{e}).$$

Preferences

- The MEU principle can be derived from general conditions that a rational agent should have.
- We use the following notation to describe an agent's preferences:
  - $A \succ B$: the agent prefers A over B;
  - $A \sim B$: the agent is indifferent between A and B;
  - $A \succsim B$: the agent prefers A over B or is indifferent between them.
- What sort of things are A and B?
  - States of the world, *but*: uncertainty about what is really being offered.
    - E.g., if you are an airline passenger and are offered pasta or chicken, you do not really know what lurks beneath the tinfoil cover.
  - ➥ The set of outcomes for each action can be seen as a *lottery*, where each action is a ticket.

Lottery

- A lottery, L, with possible outcomes $S_1, \ldots, S_n$ that occur with probabilities $p_1, \ldots, p_n$ is written as

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

  - Each S_i is either an atomic state or another lottery.

- Primary issue of utility theory:
  - How do preferences between complex lotteries relate to preferences between the underlying states in those lotteries?

➥ To address this issue, we list some conditions that we require that any reasonable preference relation should obey.

Axioms of Utility Theory

- *Orderability:* Given any two lotteries, a rational agent cannot avoid deciding which one it prefers, or whether it is indifferent between them.

  Exactly one of $(A \succ B), (B \succ A),$ or $(A \sim B)$ holds.

- *Transitivity:*

$$(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C).$$

- *Continuity:* If some lottery B is between A and C in preference, then:
  - there is some probability p for which the agent will be indifferent between getting B for sure and the lottery that yields A with probability p and C with probability $1 - p$.

$$A \succ B \succ C \Rightarrow \exists p \, [p, A; 1 - p, C] \sim B.$$

Axioms of Utility Theory (ctd.)

- *Substitutability:* If an agent is indifferent between A and B, then it is indifferent between two more complex lotteries that are the same except that B is substituted for A.

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

  This holds for $\succ$ instead of $\sim$ as well.

- *Monotonicity:* Suppose two lotteries have the same possible outcomes A and B.
  - If an agent prefers A to B, then the agent must prefer precisely the lottery that has a higher probability for outcome A.

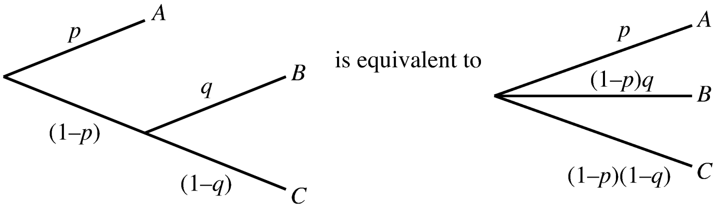
  $$A \succ B \Rightarrow (p > q \Leftrightarrow [p, A; 1-p, B] \succ [q, A; 1-q, B]).$$

Axioms of Utility Theory (ctd.)

- *Decomposability:* Compound lotteries can be reduced to simpler ones using the laws of probability.

  $[p, A; 1-p, [q, B; 1-q, C]] \sim [p, A; (1-p)q, B; (1-p)(1-q), C].$

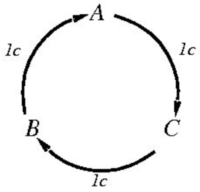
  This is known as the "no fun in gambling" rule:
  - two consecutive lotteries can be compressed into a single equivalent lottery.



Axioms of Utility Theory (ctd.)

- These conditions are known as the *axioms of utility theory*.
- Each axiom can be motivated by showing that an agent violating it will exhibit irrational behaviour.
- Consider, e.g., an agent with *intransitive preferences* $A \succ B \succ C \succ A$ can be induced to give away all its money:

  1. If the agent has A, we could offer to trade C for A plus one cent. The agent prefers C, so is willing to make the trade.
  2. We then offer B for C, extracting another cent, as the agent prefers B over C.
  3. Finally, we trade A for B. We are back to 1 except that the agent gave us 3 cents.
  4. We continue until the agent has no money.



$\Longrightarrow$ Clearly, the agent behaves irrationally.

Existence of Utility Function

As shown by von Neumann and Morgenstern (1944), the axioms of utility theory imply the following:

- *Existence of Utility Function:* Given an agents preferences that satisfy the axioms of utility theory, there exists a real-valued function U such that

$$\begin{aligned} U(A) > U(B) &\quad\Leftrightarrow\quad A \succ B \\ U(A) = U(B) &\quad\Leftrightarrow\quad A \sim B \end{aligned}$$

- *Expected Utility of a Lottery:* The utility of a lottery is the sum of the probability of each outcome times the utility of that outcome.

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

Utility Scales and Assessment

The preceding results show that a utility function exists, but they do not imply that it is *unique*.

- It can be shown that an agent's behaviour does not change if its utility function $U(S)$ is replaced by

  $$U'(S) = aU(S) + b,$$

  where a and b are constants and $a > 0$.

➥ U is determined up to *linear (affine) transformations*.

Utility Scales and Assessment (ctd.)

- In deterministic environments, where there are states and no lotteries, the behaviour of an agent is unchanged by an application of any *monotonic* transformation.
  - For instance, we could apply the square root to all utilities without changing the priority order of states.
  - One says:
    - An agent in a deterministic environment has a *value function* or *ordinal utility function*,
    - i.e., such functions just provide a *preference ranking on states*—the numbers do not matter.

- How to work out an agents utility function?
  - Present choices to an agent and use observed preferences to pin down the underlying utility function.
  - This process is called *preference elicitation*.

Utility Scales

- As we have seen, there is no *absolute* scale for utilities but it is useful to establish *some* scale for any particular problem.
- How to establish a scale?
  - Fix the utilities of any two particular outcomes.
  - Typically, we fix the utility of a "best possible prize" S_b at $U(S_b) = u_\top$ and a "worst possible catastrophe" S_w at $U(S_w) = u_\bot$.
  - *Normalized utilities* use a scale with $u_\bot = 0$ and $u_\top = 1$.

Utility Scales: Examples

- Some attempts have been made to find out the value that people place on their own lives.
- One common "currency" in medical and safety analysis is the *micromort:*
  - the event of a one-in-a-million chance of death.
- If people are asked how much they would pay to avoid a risk of a one-in-a-million chance of death they will respond with very large numbers, but their actual behaviour reflects a much lower monetary value for a micromort.
  - E.g., driving in a car for 370 km incurs a risk of one micromort; for a car with, say 150.000 km, that's about 400 micromorts.
  - People appear to be willing to pay about 10.000 Dollars more for a safer car that halves the risk of death (i.e., to incur 200 micromorts instead of 400), or about 50 Dollar per micromort.
- In general, studies on a large number of people showed that one micromort amounts to ca. 20 Dollars (in 1980s money).

Utility Scales: Examples (ctd.)

- Another measure is the *QALY* ("quality-adjusted life year"), useful for medical decisions involving substantial risks:
  - one QALY equates to one year in perfect health.
- The QALY is an indicator for the time-trade-off (TTO) to choose between remaining in a state of ill health for a period of time vs. being restored to perfect health but having a shorter life expectancy.
  - E.g., on average, kidney patients are indifferent between living two years on a dialysis machine and one year at full health.

The Utility of Money

- Money plays a significant role in human utility functions.
- Usually, an agent exhibits a *monotonic preference* for more money, all other things being equal ("ceteris paribus"), i.e., the agent prefers more money to less.
- But: this does not mean that money behaves as a utility function, because it says nothing about preferences between *lotteries* involving money.
- Example:
  - Suppose you have won in a game show and are offered a choice:
    - either take the $1,000,000 prize *or*
    - gamble it on the flip of a coin: the coin coming up heads means you end up with nothing, the coin coming up tails means you get $2,500,000.
  - How would you decide?

The Utility of Money (ctd.)

- Assuming the coin is fair, i.e., there is a 50:50 chance for coming up heads or tails, the *expected monetary value* (*EMV*) of the gamble is

$$\frac{1}{2} \cdot \$0 + \frac{1}{2} \cdot \$2,500,000 = \$1,250,000$$

$\Longrightarrow$ The EMV is more than the original \$1,000,000, but is accepting the gamble the better decision?

- Let S_n denote a state of possessing n Dollars, and say your current wealth is k Dollars.
  - ➥ The expected utilities of accepting and declining the gamble are

$$\begin{aligned} EU(Accept) &= \frac{1}{2} \cdot U(S_k) + \frac{1}{2} \cdot U(S_{k+2,500,000}) \\ EU(Decline) &= U(S_{k+1,000,000}). \end{aligned}$$

The Utility of Money (ctd.)

- How to define the utility?
  - The utility is not directly proportional to monetary value, because the utility for the first million is very high, but what about the utility for the next million?
- Assume you assign a utility of 5 to your current financial status S_k, 9 to the state $S_{k+2,500,000}$, and 8 to the state $S_{k+1,000,000}$.
- Then:

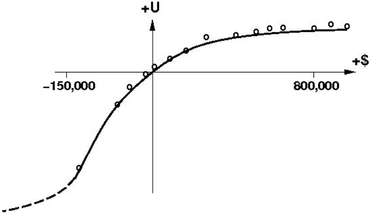
$$\begin{aligned} EU(Accept) &= \frac{1}{2} \cdot U(S_k) + \frac{1}{2} \cdot U(S_{k+2,500,000}) = \frac{5}{2} + \frac{9}{2} = 7 \\ EU(Decline) &= U(S_{k+1,000,000}) = 8. \end{aligned}$$

$\Longrightarrow$ the rational action would be to decline, because the expected utility of accepting is 7 and for declining 8.

- On the other hand, a billionaire would most likely have a utility function that is locally linear over the range of a few million more, and thus would accept.

The Utility of Money (ctd.)

- In a pioneering study of actual utility functions, Grayson (1960) found that the utility of money was almost exactly proportional to the *logarithm* of the amount.



- Preferences between different levels of debt can display a reversal of the concavity associated with positive wealth.
  - E.g., someone already \$10,000,000 in debt might well accept a gamble on a fair coin with a gain of \$10,000,000 for heads and a loss of \$20,000,000 for tails.

    $\Longrightarrow$ This leads to the S-shaped form of the curve.

Risks

- For a positive wealth, given a lottery L with expected monetary value $EMV(L)$, usually $U(L) < U(S_{EMV(L)})$, where $S_{EMV(L)}$ is the state of being handed the expected money of the lottery as the sure thing.
  - ➥ I.e., people are *risk-averse*—they prefer a sure thing with a payoff that is less than the expected monetary value of a gamble.
- On the other hand, when in large debt, the behaviour is *risk-seeking*.
- The value an agent will accept in lieu of a lottery is the *certainty equivalent* of the lottery.
  - Studies have shown that most people will accept about 400 Dollars in lieu of a gamble that gives 1000 dollars half the time and 0 Dollar the other half.
  - That is, the certainty equivalent of the lottery is 400 Dollars vs. the EMV of 500 Dollars.

    $\Longrightarrow$ The difference is called the *insurance premium*.

Risks (ctd.)

- Risk aversion is the basis for the insurance industry, because it means that insurance premiums are positive.
- People would rather pay a small insurance premium than gamble the price of their house against the chance of a fire.
  - The price of a house is small compared with the insurance company's total reserves.
  - The insurance company's utility curve is approximately linear over such a small region, and the gamble costs the company almost nothing.
- Note:
  - for small changes in wealth compared to the current wealth, almost any curve will be approximately linear.
  - An agent that has a linear curve is said to be *risk-neutral*.

Human Judgment and Irrationality

- Decision theory is a *normative theory*, i.e., it describes how a rational agent *should* act.
- A *descriptive theory*, on the other hand, describes how actual agents (e.g., humans) *really do* act.
- Evidence suggests that these two kinds of theories do not coincide $\Longrightarrow$ humans appear to be "predictably irrational".

Allais Paradox

- Assume that there is a choice between lotteries A and B and then between C and D, which have the following prizes:
  - A: 80% chance of winning \$4000
  - B: 100% chance of winning \$3000
  - C: 20% chance of winning \$4000
  - D: 25% chance of winning \$3000
- Most people prefer B over A (taking the sure thing), and C over D (taking the higher EMV).
- However, the normative analysis yields a different result:
  - Assume, without loss of generality, a utility function with $U(\$0) = 0$.
  - Then, $B \succ A$ implies $U(\$3000) > 0.8 \cdot U(\$4000)$, and $C \succ D$ implies $0.2 \cdot U(\$4000) > 0.25 \cdot U(\$3000)$.
  - From the latter we obtain
$$U(\$3000) < \frac{0.2}{0.25} U(\$4000) = 0.8 \cdot U(\$4000).$$
  - ➥ There is no utility function consistent with theses choices!

Allais Paradox (ctd.)

- One possible explanation for the apparent irrational preferences is the *certainty effect*, i.e., people are strongly attracted to gains that are certain.
- Why is that?

Allais Paradox (ctd.)

- Possible answers:
  1. People may choose to reduce their computational burden: by choosing the certain outcomes, there is no need to estimate probabilities.
  2. People may mistrust the legitimacy of the stated probabilities (in particular, if stated by people with a vested interest in the outcomes).
  3. People may account their emotional state as well as their financial state.
     - People know they would experience *regret* if they gave up a certain reward (B) for an 80% chance of a higher reward and then lost.
     - I.e., in choosing A, there is a 20% chance of getting no money and *feeling like a complete idiot*, which is worse than just getting no money.

➥ Choosing B over A and C over D may not be irrational: just willing to give up $200 EMV to avoid a 20% chance of feeling like an idiot.

Ellsberg Paradox

- Prizes have an equal value, but probabilities are underconstrained.
- Payoff depends on the color of a ball chosen from an urn.
- You are told that the urn contains 1/3 red balls, and 2/3 either black or yellow balls, but you do not know how many black and how many yellow.
- Then, you are asked to choose between A and B, and then between C and D:
  - A: \$100 for a red ball
  - B: \$100 for a black ball

  - C: \$100 for a red or a yellow ball
  - D: \$100 for a black or yellow ball
- If you think there are more red than black balls, you should prefer A over B and C over D, and the opposite otherwise.
- *But* most people prefer A over B and D over C!
- ➥ People have *ambiguity aversion*.

Ellsberg Paradox (ctd.)

Ambiguity aversion (ctd.):

- A: \$100 for a red ball
- B: \$100 for a black ball

- C: \$100 for a red or a yellow ball
- D: \$100 for a black or yellow ball

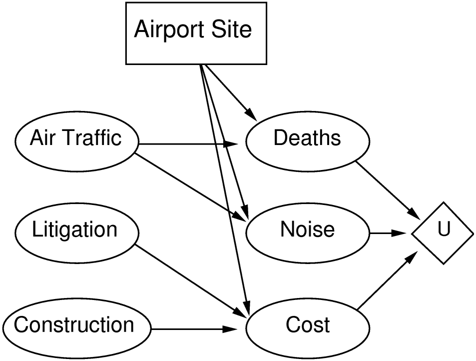
➤ A gives you a $1/3$ chance of winning, while B could be anywhere between 0 and $2/3$.

➤ Likewise, D gives you a $2/3$ chance, while C could be anywhere between $1/3$ and $3/3$.

➥ Most people *elect the known probability* rather than the unknown one.

Decision Networks

- *Decision networks* (or *influence diagrams*) are a general framework for supporting rational decisions.
- They contain information about an agent's current state, its possible actions, the state that will result from the agent's action, and the utility of that state.
- Example of a decision network for the *airport siting problem:*



Decision Networks (ctd.)

Decision network uses three types of nodes:

- *Chance nodes* (ovals): represent random variables.
  - E.g., the agent is uncertain about construction costs, the level of air traffic, the potential for litigation.
  - There are also the *Deaths*, *Noise*, and *Cost* variables, depending on the site chosen.
  - Chance nodes are associated with a conditional probability distribution that is indexed by the state of the parent nodes.
- *Decision nodes* (rectangles): represent points where a decision maker has a choice of actions; e.g., the choice of an airport site influences the cost, noise, etc.
- *Utility nodes* (diamonds): represent the agent's utility function.
  - It has as parents all variables describing the outcome that directly affect utility.

Evaluating Decision Networks

- Algorithm for evaluating decision networks:
  1. Set the evidence variables for the current state.
  2. For each possible value of the decision node:
     a) Set the decision node to that value.
     b) Calculate the posterior probabilities for the parent nodes of the utility node, using a standard probabilistic inference algorithm.
     c) Calculate the resulting utility for the action.
  3. Return the action with the highest utility.

☞ Decision networks are an extension of *Bayesian networks*, in which only chance nodes occur.

The Value of Information

- In the decision network analysis it is assumed that *all relevant information* is available before making a decision.
- In practice this is hardly ever the case:
  - ☞ *One of the most important parts of decision making is knowing what questions to ask.*
- *Information value theory* enables an agent to choose what information to acquire.
- Basic assumption:
  - the agent can acquire the value of any observable chance variables.
- These observation actions affect only the *belief state*, not the external physical state.
- The value of an observation derives from the *potential* to affect the agent's eventual physical action $\Longrightarrow$ this potential can be estimated directly from the decision model itself.

The Value of Information: Example

A simple example:

- An oil company plans to buy one of n indistinguishable blocks of ocean-drilling rights.
- One of the blocks contains oil worth C dollars, while all other are worthless.
- The price for each block is C/n Dollars.
- If the company is *risk neutral*, then it is indifferent between buying a block and not buying one.
- Now assume that the company can buy information (results of a survey) that says definitively whether block 3 contains oil or not.
- How much should the company be willing to pay for this information?

Example (ctd.)

To answer this question, we examine what the company would do if it had the information:

- With probability $1/n$, the survey will indicate oil in block 3.
  - In this case, the company will buy block 3 for C/n dollars and make a profit of $C - C/n = (n-1)C/n$ dollars.
- With probability $(n-1)/n$, the survey will show that block 3 contains no oil, hence the company will buy a different one.
  - Now, the probability of finding oil in one of the other blocks changes from $1/n$ to $1/(n-1)$, so the expected profit is $\frac{C}{(n-1)} - \frac{C}{n} = \frac{C}{n(n-1)}$ Dollars.
- Then, the resulting expected profit, given the survey information is

$$\frac{1}{n} \cdot \frac{(n-1)C}{n} + \frac{n-1}{n} \cdot \frac{C}{n(n-1)} = \frac{C}{n}.$$

➡ The company should be willing to pay up to C/n Dollars
$\Longrightarrow$ the information is worth as much as the block itself!

Remarks

- The value of information derives from the fact that *with* the information, one's course of action can be changed to suit the *actual* situation.
- One can discriminate according to the situation:
  - without the information, one has to do what is *best on average* over the possible situations.
- In general, the value of a given piece of information is defined to be the difference in expected value between the best actions before and after an information is obtained.

The Value of Perfect Information

- Assumption:
  - *Exact evidence* about the value of a random variable E_j can be obtained (i.e., we learn $E_j = e_j$).

  ➥ We use the phrase *value of perfect information* (VPI).
- Given initial evidence $\mathbf{e}$, the value of the current best action α is defined by

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

- The value of the new best action α_{e_j} after evidence $E_j = e_j$ is obtained is

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

- But the value of E_j is currently unknown, so to determine the value of discovering E_j, given current information $\mathbf{e}$, we average over all possible values e_{j_k} that might be discovered for E_j:

$$VPI_{\mathbf{e}}(E_j) = (\sum_k P(E_j = e_{j_k}|\mathbf{e})EU(\alpha_{e_{j_k}}|\mathbf{e}, E_j = e_{j_k})) - EU(\alpha|\mathbf{e}).$$

Some Properties of the VPI

- The expected value of information is *nonnegative*:

$$VPI_{\mathbf{e}}(E_j) \geq 0, \text{ for all } \mathbf{e} \text{ and all } E_j.$$

- VPI is *nonadditive*:

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

- VPI is *order independent*:

$$VPI_{\mathbf{e}}(E_j, E_k) = VPI_{\mathbf{e}}(E_k, E_j).$$

Decision-theoretic Expert Systems

- *Decision analysis* (evolved in the 1950s and 1960s) studies the application of decision theory to actual decision problems.
- Traditionally, there are two roles in decision analysis:
  - the *decision maker*, stating preferences between outcomes; and
  - the *decision analyst*, who enumerates possible actions and outcomes, and elicits preferences to determine the best course of action.
- Early expert system research concentrated on answering questions rather than on making decisions.
- The addition of *decision networks* allows expert systems to recommend optimal decisions, reflecting preferences as well as available evidence.

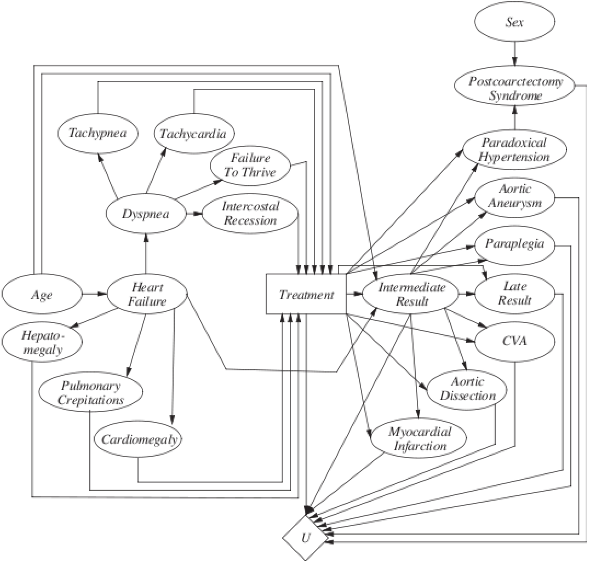
Decision-theoretic Expert Systems (ctd.)

The process of creating a decision-theoretic expert system, e.g., for selecting a medical treatment for congenital heart disease (aortic coarctation) in children:

1. create a causal model (e.g., determine symptoms, treatments, disorders, outcomes, etc.);
2. simplify to a qualitative decision model;
3. assign probabilities (e.g., from patient databases, literature studies, experts subjective assessments, etc.);
4. assign utilities (e.g., create a scale from best to worst outcome and give each a numeric value);
5. verify and refine the model, evaluate the system against correct input-output-pairs, a so called *gold standard*;
6. perform sensitivity analysis, i.e., check whether the best decision is sensitive to small changes in the assigned probabilities and utilities.

Influence Diagram Example

Influence diagram for aortic coarctation:



Summary

- *Decision theory* puts probability theory and utility theory together to describe what an agent *should do*.
- A *rational agent* makes decisions by considering all possible actions and choosing the one that leads to the best expected outcome.
- An agent whose preferences are consistent with a set of simple axioms possesses a *utility function*; furthermore, it selects actions as if maximising expected utility.
- The *value of information* is defined as expected improvement in utility compared with making a decision without the information.
- *Expert systems* that incorporate utility information are more powerful than pure inference systems:
  - they are able to make decisions and use the value of information to decide whether to acquire it, and
  - they can calculate their sensitivity to small changes in probability and utility assessments.