Overview

Making Simple Decisions

- Combining Beliefs and Desires under Uncertainty
- The Basis of Utility Theory
- Utility Functions
- Multiattribute Utility Functions
- Decision Networks
- The Value of Information
- Unknown Preferences

Expected Utility

- Uncertainty for an agent to make a decision, i.e. to perform an action a in possible states s to reach state s‘ with the transition model P(s‘|s,a):
  - $P(\text{Result}(a) = s') = \sum_s P(s)\, P(s'|s,a)$
- In its simplest form, we are interested in the utilities of the immediate outcomes of actions
  - The agents prefences are expressed by a utility function for a state U(s)
- The expected utility of an action EU(a) is the average utility of the possible outcomes, weighted by their probabilities:
  - $EU(a) = \sum_{s'} P(\text{Result}(a) = s')\, U(s')$
- **Maximum expected utility (MEU)**: action = $\text{argmax}_a\, EU(a)$
  - ➢ Implements the performance measure of a rational agent

Bus-Example for MEU

- Which bus should we choose to the central station, if we want to catch a train?
  - Every 20 minutes a bus drives to the central station and a train drives every hour.
  - Bus1 arrives 5 minutes before departure of train, Bus2 25 minutes
  - Goal: Minimize total waiting time.
- **Assignment of probabilities and utilities:**
  - P(Bus1 is late for more than 4 minutes) = 20%; Waiting Costs for Bus1 = 0
  - P(Bus2 is late for more than 24 minutes) = 1%; Waiting Costs for Bus2 = -20
  - U(Bus late) -60 // Waiting for the next train; U(Bus in time) = 0
- **Computation of MEU**
  - EU (Bus) = Costs (Bus) + P(Bus late) * U(Bus late) + P(Bus in time)*U(Bus in time)
  - EU(action = Bus1) = 0 + (0.2 * - 60) + 0 = -12
  - EU (action = Bus2) = -20 + (0.01 * -60) + 0 = -20.6

→ **take Bus1**

- **Sensitivity analysis:** Compute „tipping points“ for probability and utility: P (Bus1 late) > 35%;
  U(Bus late) > -105

Basis of Utility Theory

- Lottery metapher: The outcomes S_i for each action can be viewed as lottery (each action is a ticket) with corresponding probabilities p_i: $L = [p_1, S_1; p_2, S_2; \ldots\ p_n, S_n]$
  - S_i can be a state or another lottery
  - An agent prefers a lottery over another one or is indifferent: $(A \succ B)$, $(A \sim B)$, $(A \succsim B)$
- We require six constraints (axioms) for any reasonable preference relation:
  - **Orderability:** For two lottery, exactly one relation holds: $(A \succ B)$ or $(A \sim B)$ or $(A \succsim B)$
  - **Transitivity:** $(A \succ B) \land (B \succ C) \Rightarrow (A \succ C)$
  - **Continuity:** $A \succ B \succ C \Rightarrow \exists p\ [p,A; 1-p,C] \sim B$
  - **Substitutability:** $A \sim B \Rightarrow [p,A; 1-p,C] \sim [p,B; 1-p,C]$
  - **Monotonicity:** $A \succ B \Rightarrow (p > q \Leftrightarrow [p,A; 1-p,B] \succ [q,A; 1-q,B]$
  - **Decomposability:** $[p,A; 1-p,[q,B; 1-q,C]] \sim [p,A; (1-p)q,B;\ (1-p)(1-q),C]$ „no fun in gambling“

Utility Function

- If an agent‘s preferences obey the axioms, then there exist a function U, such that
  $U(A) > U(B) \Leftrightarrow A \succ B$ and
  $U(A) = U(B) \Leftrightarrow A \sim B$
- The expected utility of a lottery is the sum of the probabilities of each outcome times its utility
  $U([p_1, S_1; p_2, S_2; \ldots\ p_n, S_n]) = \sum_i p_i\, U(S_i)$
- An agent‘ behaviour would not change for linear transformations of the utility function like
  $U‘(S) = a\, U(S) + b$ // a and b are constants with $a > 0$

Concrete Utility Functions

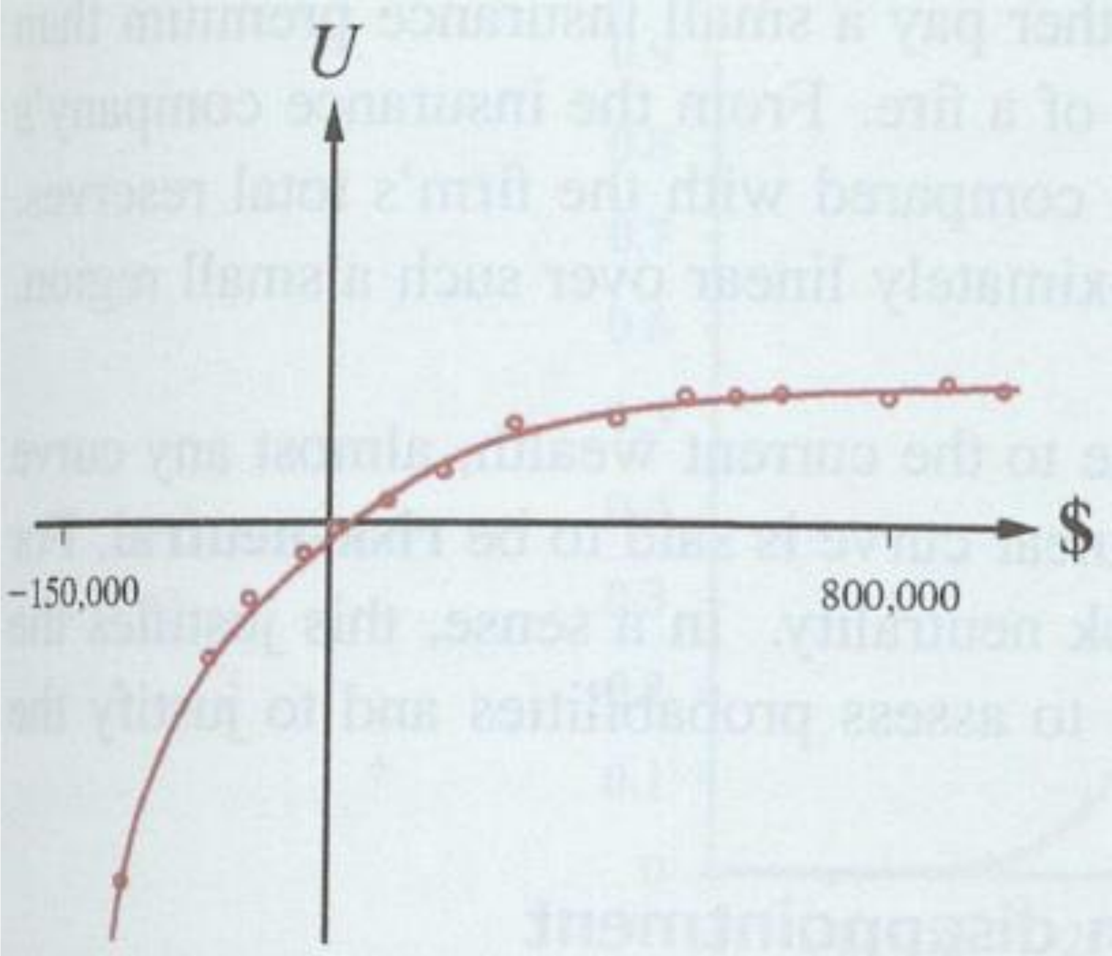
- Common utility funcitions:
  - Money
  - Time
  - QALY (Quality Adjusted Live Year):
    - Patients are willing to accept a shorter life expectancy to avoid disability
    - Example: Kidney patients on average are indifferent between living two years on dialysis or one year at full health
  - Micromort (costs for a one in a million chance of death)

Non-linear Utility of Money

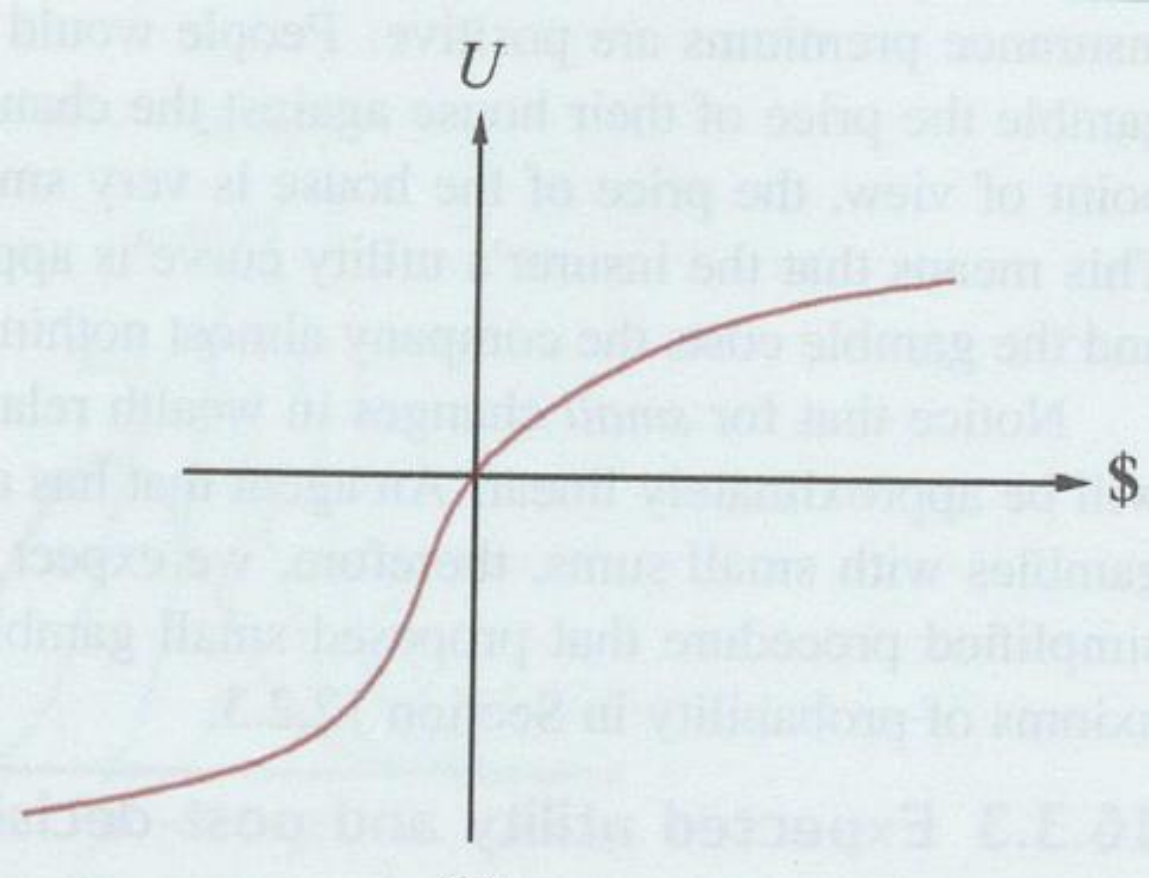
- For most people, the utility of gaining or loosing money is nonlinear
- Examples:
  - Assume, you won in a television show and can choose, whether you get 5 million € for sure or 15 million € with a chance of 50%
  - How much money should you give for participating in a lottery, where a coin is flipped n+1 times with a gain of 2^n € with n = number of coins showing uninterrupted „emblem“
  - Which lottery option would you prefer for lottery 1 and 2?
    - Lottery1: 4000 € with 80% vs. 3000 € with 100%
    - Lottery2: 16000 € with 20% vs. 12000 € with 25%

Utility of Money

Empirical data over a limited range:



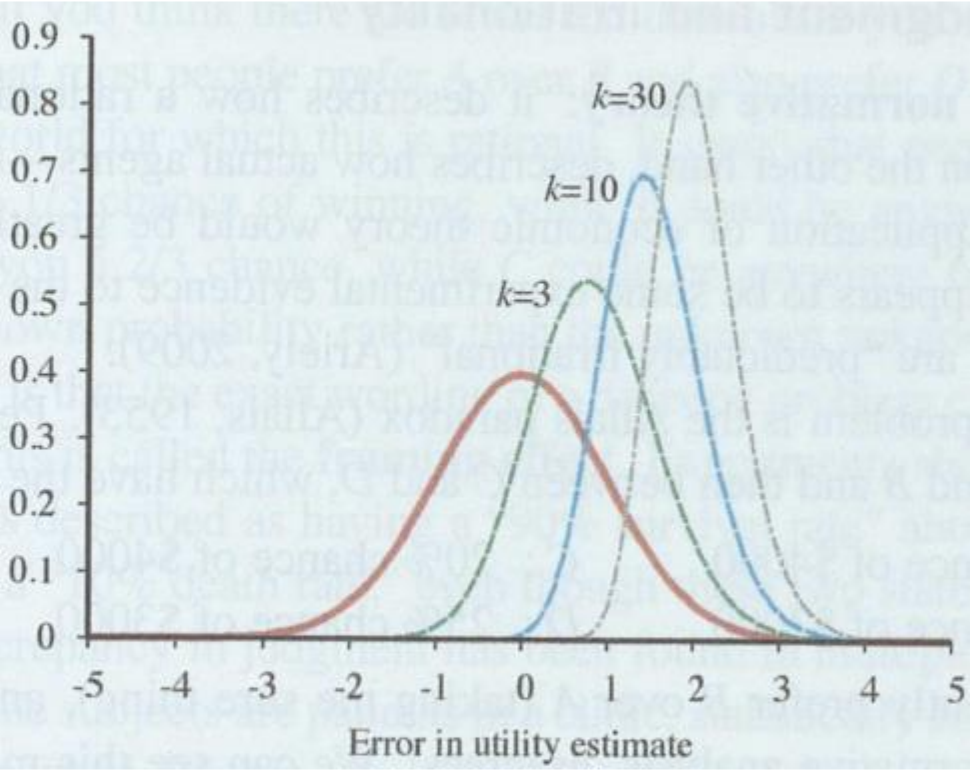
Typical curve for the full range:



(b)

Expected Utility and Post-Decision Disappointment

- We choose the best action a* with the maximal expected utility: **$a^* = \text{argmax}_a\ EU(a)$**
- Usually, the expected utility is quite complex and must be a estimated
  - Each estimate has an error rate, e.g. given by a normal distribution with mean and standard deviation.
  - If we choose the best action from many similar good actions, we probably choose an action in the upper end of its distribution, i.e. the estimation is too good by chance
    - ➢ **Optimizer‘s curse:** Overestimation of utility of best action
    - ➢ Calculation of degree of unjustified optimism depending on standard deviation of the estimates and number of available actions
      - ➢ right: expected unjustified optimism when choosing from k actions with true utility of zero
        - ➢ for k = 30, the utility of the best action is estimated ca. two standard deviations better than the real utilty

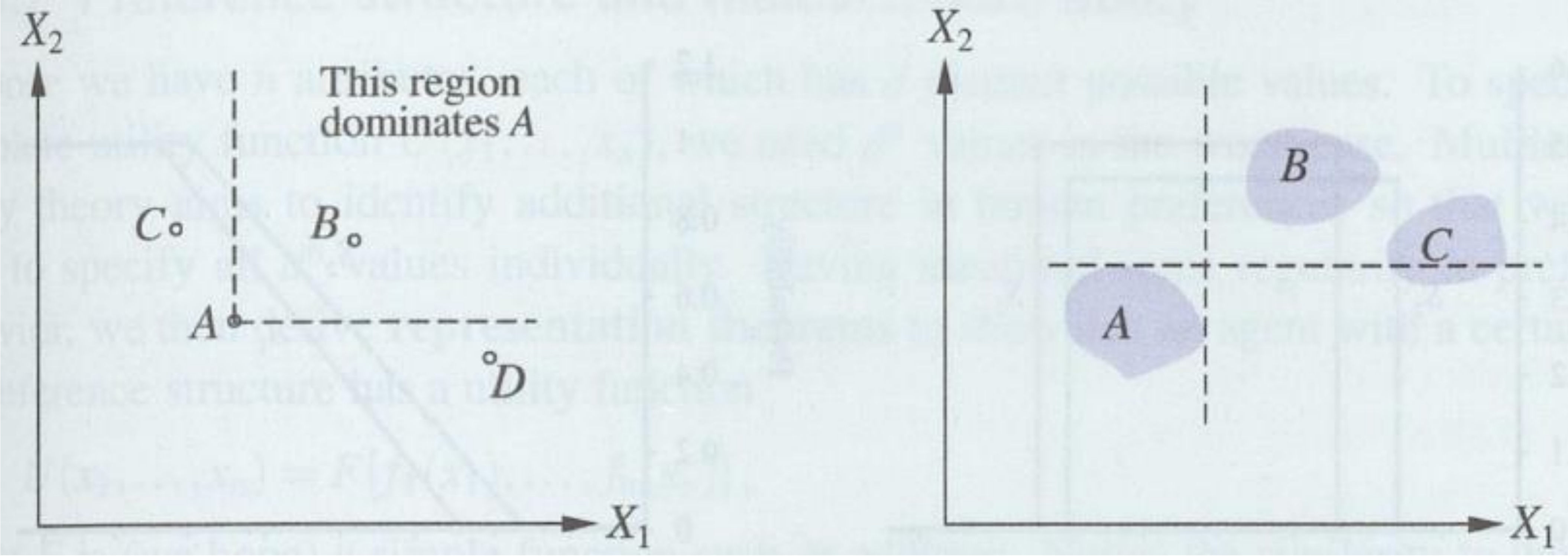


Human Judgement and Irrationality

- Psychological research has shown, that people act apparently irrationally in certain situations:
  - **Certainty effect:** People prefer a sure gain to an unsure but higher gain expectation, e.g. 3000 € for sure compared to 80% chance of gaining 4000 €
    - Explanations: (1) Strong regret if loosing, (2) distrust in stated probabilities, (3) computational burden
    - Generalization: Aversion against ambiguity
  - **Framing effect**: The wording of choices can have a big impact on choices, e.g. 90% success rate versus 10% chance of failure
  - **Anchoring effect**: A first impression is corrected insufficiently with opposite evidence, e.g. if a restaurant offers wine for 100 €, people will not buy it, but a bottle for 40 € seems like a bargain.
- There are **evolutionary explanations** for this behaviour (e.g. first impression on people is often appropriate) and it can be corrected with reflection.

Multiattribute Utility Functions

- The utility of many decisions depends on more than one attribute
  - e.g. localisation of an airport: costs, reachability, noise, safety issues etc.
  - **$U(x_1, \ldots x_n) = F[f_1(x_1), \ldots , f_n(x_n)]$** with F being hopefully a simple function like addition
- Qualitative analysis:
  - An option ist **strictly dominant** to another, if it is better in each attribute (below left)
  - An option is **stochastically dominant**, if it is better for each probability niveau (below right)

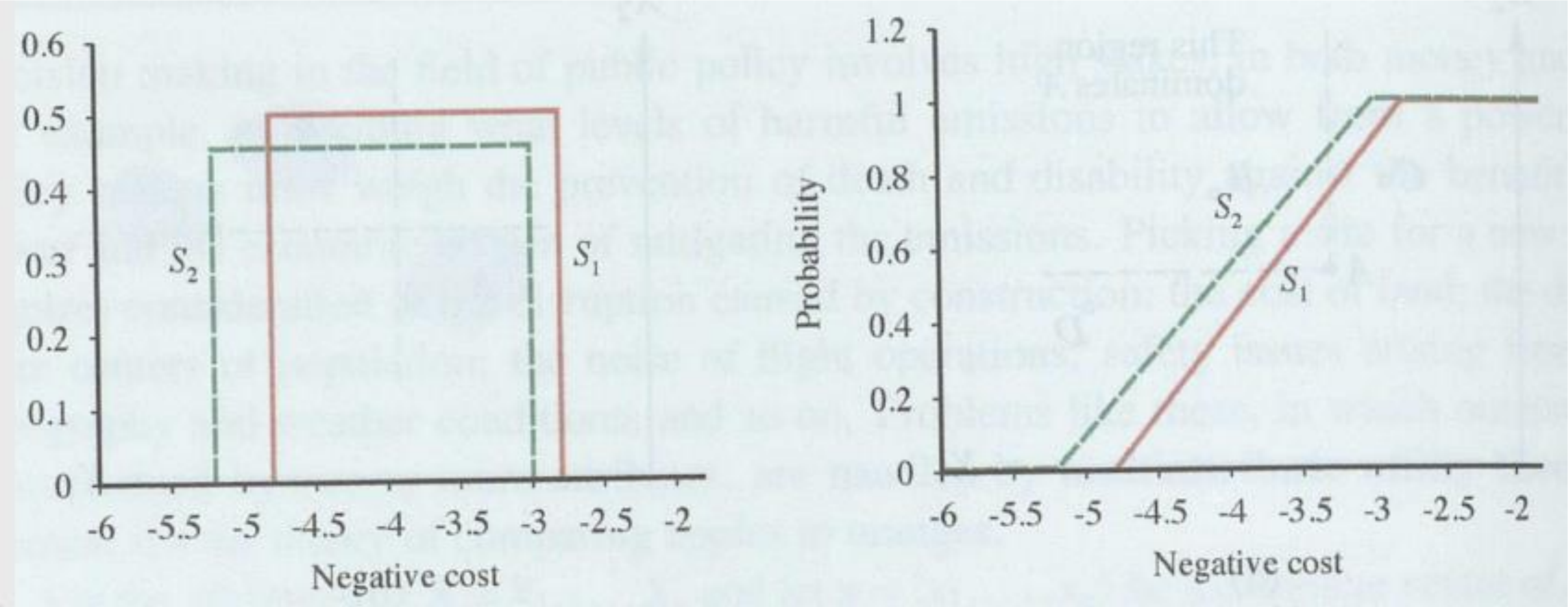


Stochastic Dominance

- Stochastic dominance can be computed by comparing small slices of the probability curve (i.e. for every probability level the utility of one action is better than the other) or by computing the integral:

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

- Example:



Preference Independance

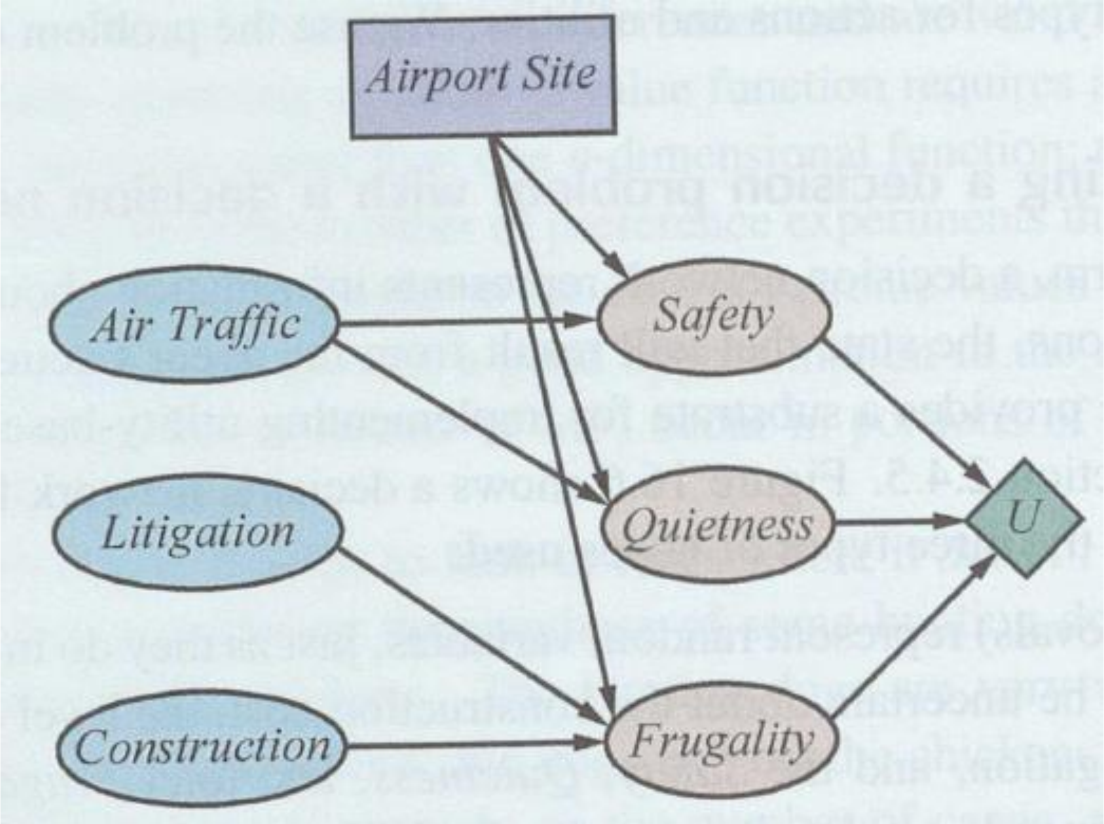
- Without strict or stochastic dominance the decision is more difficult:
  - With m attributes and n values per attribut, there are m^n states
  - We need some structure to handle such situations
- **Preference Independance:** Two attributes X_1 and X_2 are preferentially independant to a third attribut X_3, if the preference between outcomes $\langle x_1, x_2, x_3 \rangle$ and $\langle x_1', x_2', x_3 \rangle$ does not depend on the particular value for attribute X_3
  - If attributes X_i are **mutually preferentially independant (MPI)** their values V_i can be added:

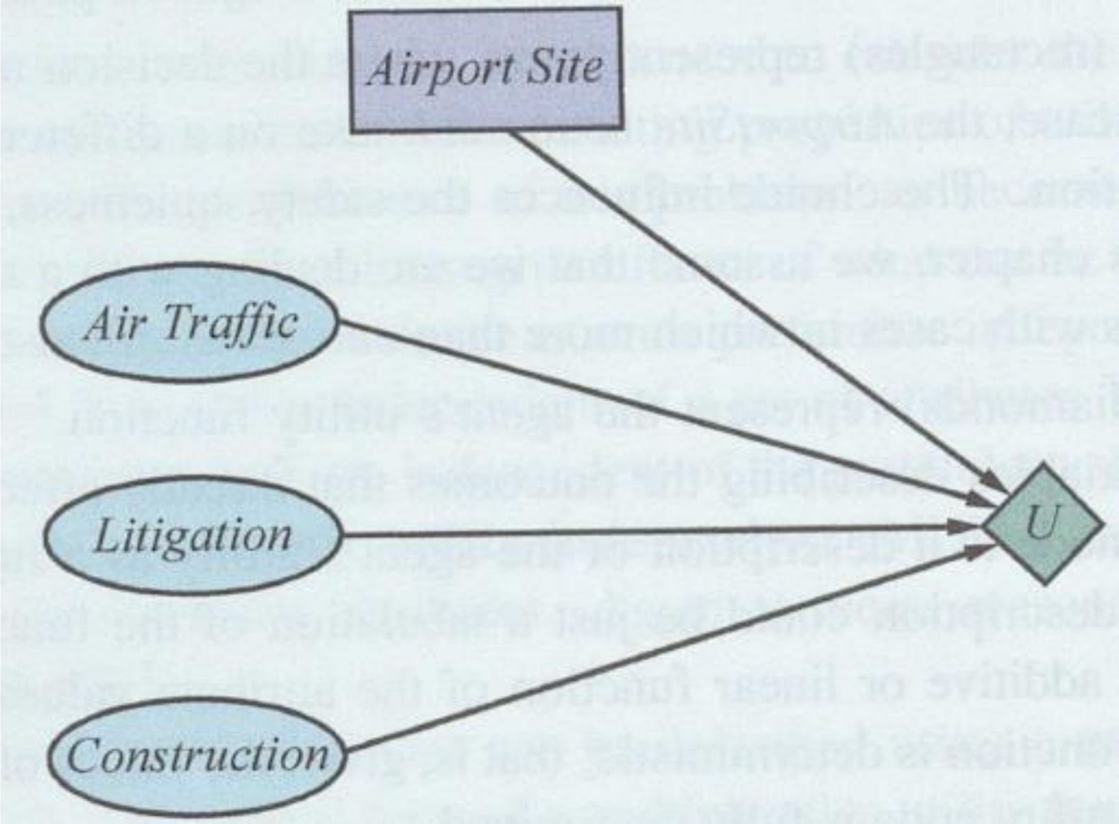
$$V(x_1 \ldots x_n) = V_1(x_1) + V_2(x_2) + \ldots + V_n(x_n) = \sum_i V_i(x_i)$$

- **Without MPI and with uncertainty: More complicated** (may need additional assumptions to add their utilities)
  - e.g. you are in a medieval market considering the purchase of some hunting dogs, some chickens and some cages for chickens. Without cages, the hunting dogs will eat the chickens. There is uncertainty about the prices of hunting dogs, chickens and cages.

Decision Networks

- Extension of Bayesian networks with chance nodes (oval) and additional node types for actions (rectangles) and utilities (diamonds)
  - Optional: Simplification without intermediate chance nodes (right)
- Example: Airport site decision



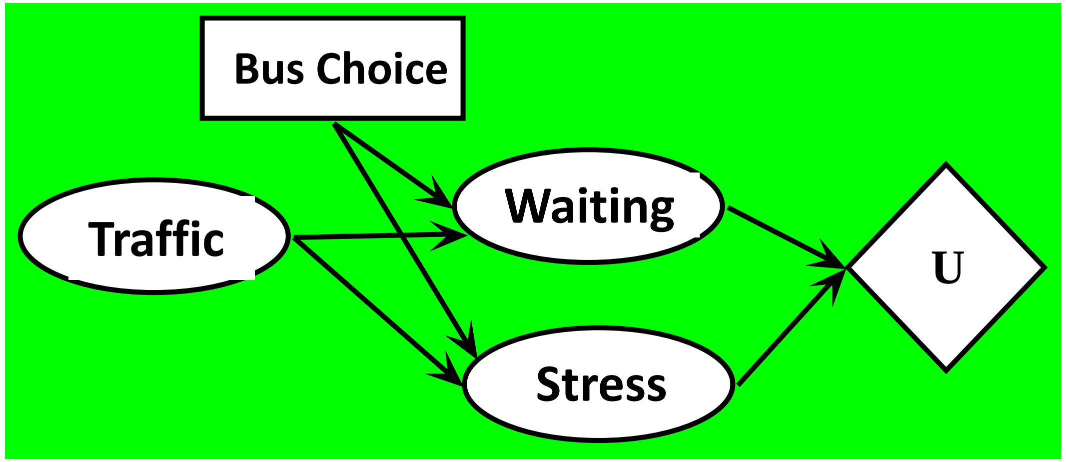


Evaluating Decision Networks

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 probability 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

Extension of Bus-Example

- Which bus should we choose to the central station, if we want to catch a train?
- Extensions:
  - Include stress as second attribute to waiting time
  - Waiting time and stress depend on traffic situation (high or normal; P(Traffic=high) = 0.1)
- Decision network:



CPTs (**T**raffic: **H**igh or **N**ormal; **W**aiting; **S**tress)

Bus	Traffic	P(W)	P(S)
1	H	50%	95%
2	H	2%	5%
1	N	5%	20%
2	N	0.5%	1%

Utilities:

U(W)= -60; U(¬W) = 0

U(S) = -10; U(¬S) = 0

Cost(Bus1)= 0

Costs(Bus2)= -20

U = Cost(Bus) + U(W)*P(W) + U(¬W)*P(¬W) + U(S)*P(S) + U(¬S)*P(¬S)

T=H, Bus=1: 0 + -60*0.50 + 0 + -10*0.95 + 0 = -39.5

T=H, Bus=2: -20 + -60*0.02 + 0 + -10*0.05 + 0 = -21.7 ⇒ *Take Bus2 if traffic = high*

The Value of Information

- Usually, not all information for making a decision is available (e.g in medicine).
- Therefore, the agent must decide if and what information to gather, even if information gathering is costly in itself.
- Algorithm: Compute the average utility with additional information and compare it to the utility without information
  - Apriori-Probability of the additional information necessary

Oil-Example

An oil company knows, that exactly one block from an area of n indistinguishably block contains oil with a value of x € und wants to buy the rights for this block. A seismologist offers the company definite information about one block. How much should the company pay for this information?

Computation:

- Apriori-Value of one block x/n €
- Case distinction:
  - If offered block contains oil: P = 1/n $\rightarrow$ Gain = x - x/n
  - If offered block contains no oil: P = (n-1)/n $\rightarrow$ Gain = value of remaining blocks rises to x/(n-1) - x/n
- Total gain: $\frac{1}{n}(x - \frac{x}{n}) + \frac{n-1}{n}(\frac{x}{n-1} - \frac{x}{n}) = \frac{xn - x}{n^2} + \frac{xn - (xn - x)}{n^2} = \frac{xn}{n^2} = \frac{x}{n}$

➢ *The information about a block is worth the value for the rights of a block!*

Deriving the Value of Information

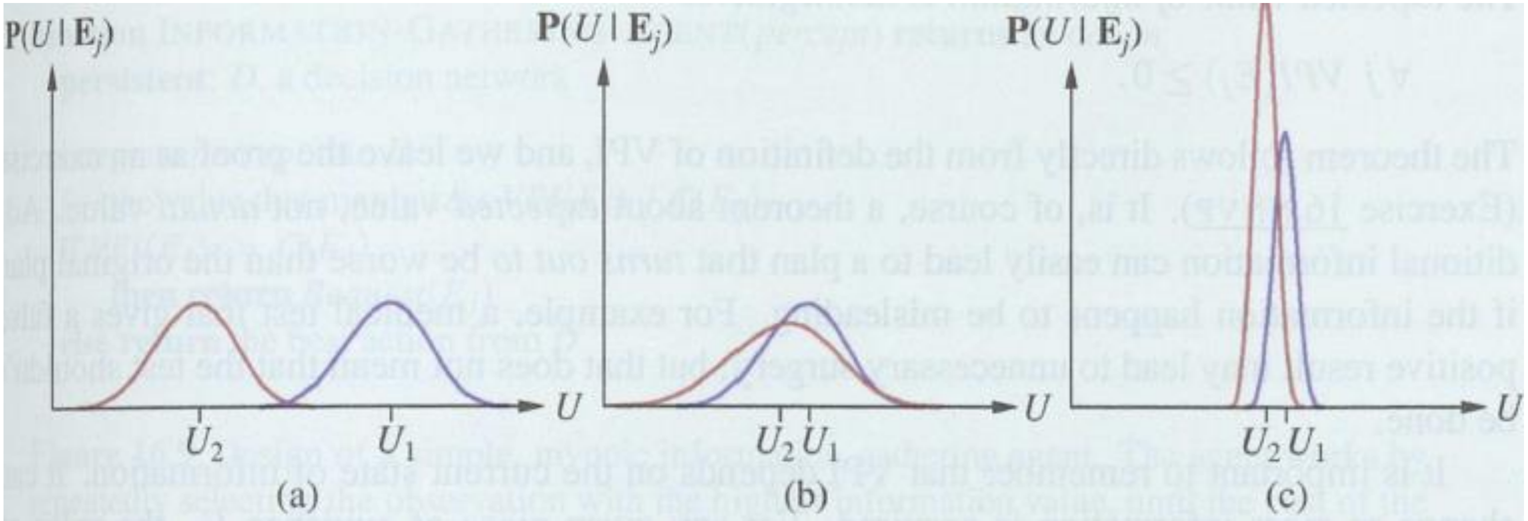
- Compare the expected utility (EU) of the best action α with and without obtaining a new evidence E_j ($E_j = e_j$), where e_j is currently unknown.
  - Without: $EU(\alpha) = \max_a \sum_{s'} P(Result(a) = s')\ U(s')$
  - With: $EU(\alpha_{e_j} | e_j) = \max_a \sum_{s'} P(Result(a) = s' | e_j) U(s')$
- Since E_j is currently unknown, we must average over all possible values e_j weighted with the current probabilty of e_j, to compute the value of perfect information (VPI):
  - $VPI(E_j) = \left(\sum_{e_j} P(E_j{=}e_j)\ EU(\alpha_{e_j} | E_j = e_j) \right) - EU(\alpha)$

Simple example with two actions α_1 and α_2 with expected utilities U_1 and U_2:

a) U_1 better than U_2, no information needed

b) Information cruical for decision

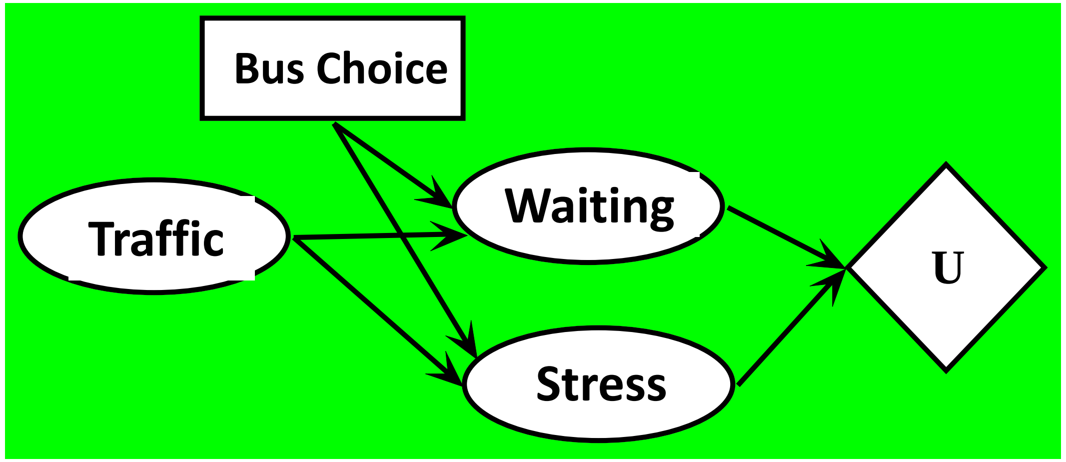
c) U_1 similar to U_2, no information needed



Bus-Example Adapted

- **Should we ask for traffic information to choose a bus?**

Decision network:



CPTs (**T**raffic: **H**igh or **N**ormal; **W**aiting; **S**tress)

Bus	Traffic	P(W)	P(S)
1	H	50%	95%
2	H	2%	5%
1	N	5%	20%
2	N	0.5%	1%

P(Traffic=high) = 0.1

Utilities:

U(W)= -60; U(¬W) = 0

U(S) = -10; U(¬S) = 0

Cost(Bus1)= 0

Costs(Bus2)= -20

U = Cost(Bus) + U(W)*P(W) + U(¬W)*P(¬W) + U(S)*P(S) + U(¬S)*P(¬S)

T=H, Bus=1: 0 + -60*0.50 + 0 + -10*0.95 + 0 = -39.5

T=N; Bus=1: 0 + -60*0.05 + 0 + -10*0.20 + 0 = -5

T=H, Bus=2: -20 + -60*0.02 + 0 + -10*0.05 + 0 = -21.7

T=N, Bus=2: -20 + -60*0.005 + 0 + -10*0.01 + 0 = -20.4

Without traffic info (TI): EU (Bus1) = -8.45

Bus1 = 0.1 * -39.5 + 0.9 * -5 = -8.45

Bus2 = 0.1 * -21.7 + 0.9 * -20.4 = -20.53

With TI: EU = 0.1 * -21.7 + 0.9 * -5 = -6.67

T=H: prefer Bus2: -21.7

T=N: prefer Bus1: -5

⇒ **Traffic info is useful (value gain: 1.78 U)!**

Properties of the Value of Information

- The expected value of information (VPI) is nonnegative: $\forall$ **j VPI(E_j) $\geq$ 0**
  - However the actual value might be negative (e.g. if the outcome of a medical test leads to a plan with unnessary treatment surgery)
- VPI is not additive: **VPI(E_j, E_k) $\neq$ VPI(E_j) + VPI(E_k)** (in general)
- VPI is not order dependant: **VPI(E_j, E_k) = VPI(E_j) + VPI($E_k | E_j$) = VPI(E_k) + VPI($E_j | E_k$) = VPI(E_k, E_j)**

Algorithm for an Information-Gathering Agent

- A greedy information gathering could iteratively select the most useful information, if its costs are less than its expected benefit

function INFORMATION-GATHERING-AGENT(*percept*) **returns** an *action*
  persistent: *D*, a decision network

  integrate *percept* into *D*
  $j \leftarrow$ the value that maximizes $VPI(E_j)\ /\ C(E_j)$
  if $VPI(E_j) > C(E_j)$
    then return $Request(E_j)$
  else return the best action from *D*

- However in some cases, planning a sequence of information gathering actions is better
  - Example: Planning of radiologic examinations (X-Ray $\rightarrow$ CT $\rightarrow$ MRT vs. MRT immediately)

Sensititity Analysis and Robust Decisions

- In practice, probability and utility values (parameters) are often potential incorrect estimates
- **Sensitivity analysis** tries to assess that kind of uncertainty
  - For a model:
    - Compute for each parameter derivative of expected utility to find most important one
  - For an actual decision:
    - Compute for each parameter the **tipping point** (see slide 4)
    - Compute for all **reasonable pertubations of parameters**, if decision remains stable
  - For getting a robust decision: Select the best decision under the assumption of the worst reasonable pertubations of parameters (let θ stand for all parameters in the model):

    $\mathbf{a^* = argmax_a\ min_\theta\ EU(a; \theta)}$ [control theory, decision analysis, risk management]
  - Bayesian approach:
    - Model the uncertainty of all parameters with additional „**hyperparameters**“
  - Difficult: Dealing with structural uncertainty (i.e. whether the model is correct)
    - Possible solution: **Ensemble of models**

Unknown Preferences

- Uncertainty about parameters can be represented with additional parameters
  - If an agent does not know its utilities, e.g. in a restaurant ordering a known or an unknown dessert?
    - For the unknown dessert, probabilities for different utilities can be stated and summarized to an expected utility
  - If an agent should act for another agent (e.g. human), but does not know its utilities, it must decide, whether to act nevertheless or ask the human or do nothing
    - Can be calculated, if the costs for asking and the probabilities of the different utilities of the humans are known
    - In addition, asking enables learning of better estimates in the future

Knowledge Engineering Process

- It is useful to separate probabilities and utilities in a model
- **Knoweldge engineering process:**
  1. Sufficient understanding of the domain
  2. Development of causal model
     - Terminology (variables)
     - Causal relations (links)
     - Optional simplification by removing, aggregating oder seperating variables
  3. Assignment of probabilities
  4. Assignment of utilities
  5. Model evaluation and refinement with cases (gold standard)
  6. Sensitivity analysis of model (optionally add hyperparameters based on parameter ranges)
  7. Application to problem
  8. Sensitivity analysis of result (optionally with „worst case analysis“ based on ranges)