Rational decisions under uncertainty

Jiří Kléma

Department of Computer Science,
Czech Technical University in Prague

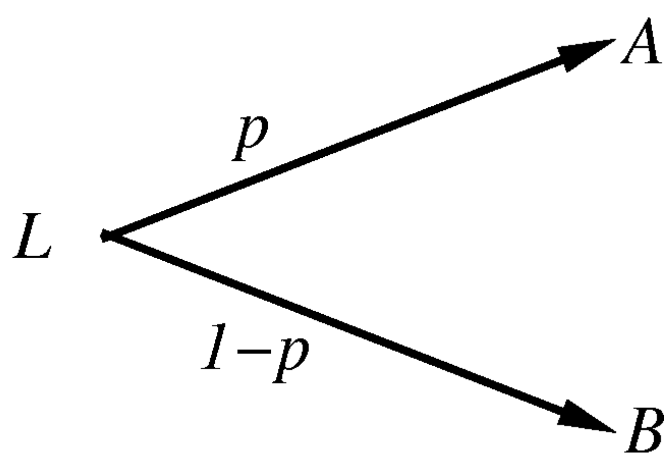
http://cw.felk.cvut.cz/doku.php/courses/a4b33zui/start

Outline

- Artificial agents with deliberate and effective decision making = **rational**
  - how to define them,
  - how to cope with uncertain action results,
  - decision theory + utility theory,
  - concepts: prize, lottery, utility function, preference,
- rational and deliberate human decision making
  - do we behave rationally when making decisions?
  - money as an example of ordinal utility measure,
- multiattribute utility
  - each and every state cannot be separately assessed,
  - preference and utility derived from its attributes,
- value of information
  - when does it pay off to make an effort to obtain a piece of information?

Task formalization, basic terms

- agent chooses among possible **states** of the world $\{s_1, \ldots, s_n\}$,
- every state can be assigned a **prize** $\{A, B, \ldots\}$,
- agent reaches states by performing **actions** $\{a_1, \ldots, a_m\}$
  - actions are stochastic, the outcome=state is not certain,
  - action a leading with prob p to state s_1 with prize A and with prob $p-1$ to state s_2 with prize B can be defined as **lottery** $L_a = [p, A; (1-p), B]$,
  - a deterministic lottery (no random element) is equal to a prize,
- rationality: the agent's goal is to apply action resulting in the highest prize.



Preferences

- How can we define prizes?
  - in general, they do not have to be numerical,
  - it suffices to define symbolic prizes with **preference** relations

    $A \succ B$ … A preferred to B,

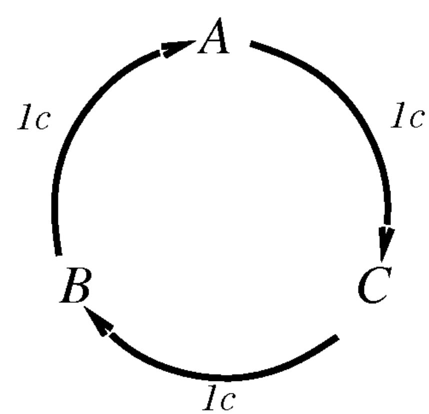
    $A \sim B$ … A and B indifferent,

    $A \succeq B$ … B not preferred to A,

- a rational agent has to implement preferences with certain **constraints**
  - orderability: $(A \succ B) \vee (B \succ A) \vee (A \sim B)$,
  - transitivity: $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$,
  - continuity: $A \succ B \succ C \Rightarrow \exists p\ [p, A;\ 1-p, C] \sim B$,
  - 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: $[p, A;\ 1-p, [q, B;\ 1-q, C]] \sim [p, A;\ (1-p)q, B;\ (1-p)(1-q), C]$.

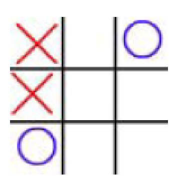
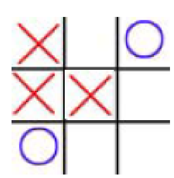
Transitivity as necessary condition of rationality

- Violating the constraints leads to irrationality,
- example: intransitive agent can give away all his money
  - assume an agent with preferences $A \succ B$, $B \succ C$, $C \succ A$,
  - it is willing to pay (say) 1 cent to exchange its C for somebody else's B,
  - consequently, it pays 1 cent to exchange its B for somebody else's A,
  - finally, it exchanges A for C and pays 1 extra cent again,
  - it owns C again, but has got 3 cents less.



Maximizing expected utility

- von Neumann-Morgenstern theorem
  - given constrained preferences there exists a real-valued function U s.t.
    $U(A) \geq U(B) \Leftrightarrow A \succeq B$
    (we keep the preferences in prizes),
    $U([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i U(S_i)$
    (lottery utility computed as expected utility of the individual outcomes),
- maximizing expected utility, the principle
  - take an action (corresponding lottery) that maximizes expected utility,
- introduction of (explicit) utility is not a necessary condition of rationality
  - ex.: agent has its strategy in the form of look-up table.

State	Best action
	

From preferences towards utility function

- utility function maps prizes (and thus states) on real numbers
  - the linear ordering given by preferences must be preserved,
  - there is an infinite number of functions with the identical behavior of agent,
  - in deterministic environments (without lotteries) it is the only condition

    $A \prec B \sim C \preceq D$ agrees both with U_1 and U_2,

    the behavior of agent does not change.

	A	B	C	D
U_1	1	2	2	3
U_2	-1	2	2	1000

From preferences towards utility function

- with lotteries there is one more condition,
- behavior does not change with **linear** utility function transformations only,

  $$\forall k_1 > 0 \;\; U_2(x) = k_1 U_1(x) + k_2,$$

	A	B	C	D
U_1	1	2	2	3
U_2	-1	2	2	1000

  – U_1 and U_2 interchange preferences in lotteries $[0.5, A; 0.5, B]$ and $[0.9, A; 0.1, D]$,

- standardize by normalized utility
  – best possible prize $u_\top = 1.0$, worst possible catastrophe $u_\perp = 0.0$,
  – any intermediate prize A matches p set such that $A \sim [p, u_\top; (1-p), u_\perp]$.

People as "rational" money-driven agents

- St. Petersburg paradox (Bernoulli, 1738)
  - how much would you pay as an entry fee for the following game?
    - adversary repeatedly tosses a (standard) coin until the first head,
    - the number of coin tosses $n \rightarrow$ your gain 2^n Kč $\rightarrow$ game over,

People as "rational" money-driven agents

- St. Petersburg paradox (Bernoulli, 1738)
  - how much would you pay as an entry fee for the following game?
    - ∗ adversary repeatedly tosses a (standard) coin until the first head,
    - ∗ the number of coin tosses $n \rightarrow$ your gain 2^n Kč $\rightarrow$ game over,
  - provided that money directly represent utility function, you shall be willing to pay an arbitrary finite fee
    - ∗ let us apply von Neumann-Morgenstern theorem

      $U(pbgh) = U([p(h_1), U(h_1); p(h_2), U(h_2); \ldots]) =$

      $= \sum_{i=1}^{\infty} \frac{1}{2^i} 2^i = 1 + 1 + \cdots = \infty$
  - this conclusion does not seem to be truly rational
    - ∗ Bernoulli solved the paradox by log transform of money utility

      $U(k) = \log_2 k$

      $U(pbgh) = \sum_{i=1}^{\infty} \frac{1}{2^i} \log_2 2^i = \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \cdots = 2$
    - ∗ reverse transformation gives the real game fee: $2 = \log_2 k \Rightarrow k = 4$ Kč.

People as “rational” money-driven agents

- Tversky and Kahneman experiment (1982)
  - choose one of the lotteries L_1 and L_2, then one of the lotteries L_3 and L_4

Choice 1	Choice 2
$L_1 = [0.8, 80000Kc; 0.2, 0]$	$L_3 = [0.2, 80000Kc; 0.8, 0]$
$L_2 = [1, 60000Kc]$	$L_4 = [0.25, 60000Kc; 0.75, 0]$

People as “rational” money-driven agents

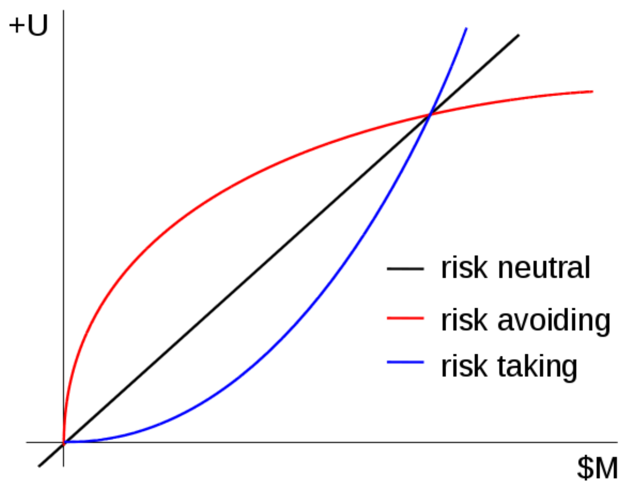
- Tversky and Kahneman experiment (1982)
  - choose one of the lotteries L_1 and L_2, then one of the lotteries L_3 and L_4

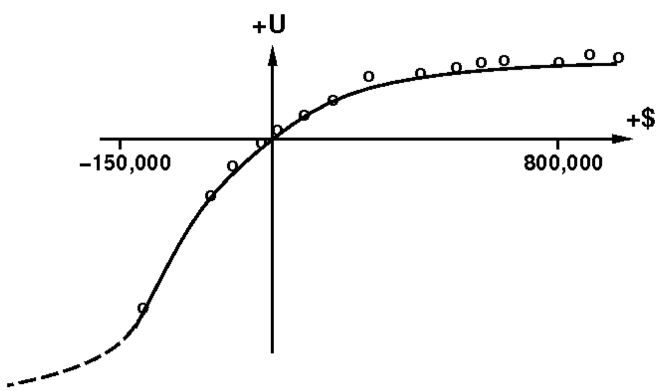
Choice 1	Choice 2
$L_1 = [0.8, 80000Kc; 0.2, 0]$	$L_3 = [0.2, 80000Kc; 0.8, 0]$
$L_2 = [1, 60000Kc]$	$L_4 = [0.25, 60000Kc; 0.75, 0]$

  - most people prefer lottery L_2 to L_1 and L_3 to L_4
    * does not seem rational, provided that $U(0Kc) = 0$ it holds
      choice 1: $0.8U(80000Kc) < U(60000Kc)$,
      choice 2: $0.8U(80000Kc) > U(60000Kc)$,
    * there is no utility function consistent with both choices,
  - possible explanations
    * people are irrational,
    * the analysis disregards regret when loosing a very likely reward ad L_2,
    * that is why people avoid/take risk in probable/unlikely events.

People as "rational" money-driven agents

- money is not the direct utility function
  - people often do not maximize monetary expected utility, $U([p_1, S_1; \ldots; p_n, S_n]) \neq \sum_i p_i U(S_i)$
  - and tend to avoid the risk, i.e., lotteries, $U([p_1, S_1; \ldots; p_n, S_n]) < \sum_i p_i U(S_i)$
- **utility curve** non-linearly transforms money to utility
  - we search for probability p, for which a given person does not distinguish prize x and lottery $[p, \$M; (1-p), \$0]$, $\$M$ is large



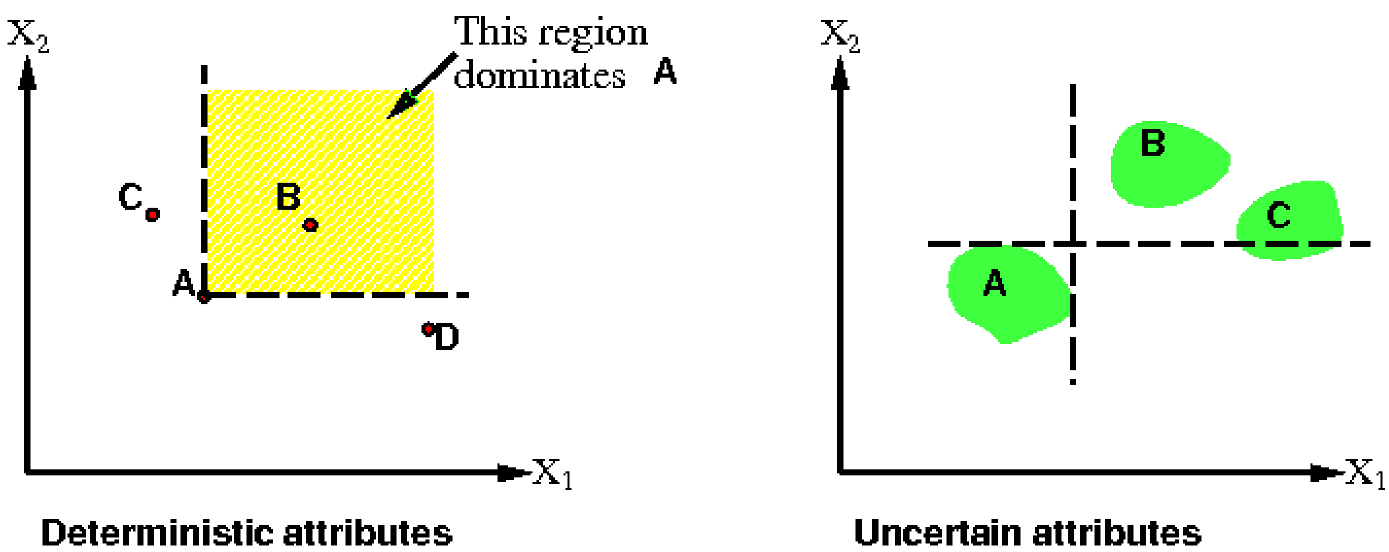


Multiattribute utility functions

- Often we cannot assign a prize to every state
  - too many states or infinite state space,
  - states usually described by features
    (airport locality selection – safety, noise level, land prize),
- utility function has several parameters then
  - $U(X_1, \ldots, X_n)$ (parameters resp. attributes instead of state),
  - n attributes with m distinct values define m^n states,
  - utility function can be simplified by assumption of preference **regularity**
    * preference monotonicity when changing single attribute
      $x \geq y \Rightarrow U(X_1, \ldots, X_i = x, \ldots, X_n) \geq U(X_1, \ldots, X_i = y, \ldots, X_n)$,
    * relationships of independence among attributes wrt preferences
      state defs: $A \sim (x_1, y_1)$, $B \sim (x_2, y_1)$, $C \sim (x_1, y_2)$, $D \sim (x_2, y_2)$
      preference independence: $(A \succ B \Rightarrow C \succ D) \wedge (A \succ C \Rightarrow B \succ D)$
  - preference regularities correspond to a simplified utility function
    * $U(x_1, \ldots, x_n) = f[f_1(x_1), \ldots, f_n(x_n)]$, f is simple, e.g., addition.

Strict dominance

- assumption: U monotonously increasing in all attributes,
- choice B strictly dominates choice A iff
  - $\forall i \; X_i(B) \geq X_i(A) \Rightarrow f_i(X_i(B)) \geq f_i(X_i(A)) \Rightarrow U(B) \geq U(A)$
  - one airport location safer, less noisy with cheaper land than others,
- rarely applicable in practice
  - utility further decreased by uncertainty in estimation of attribute values.

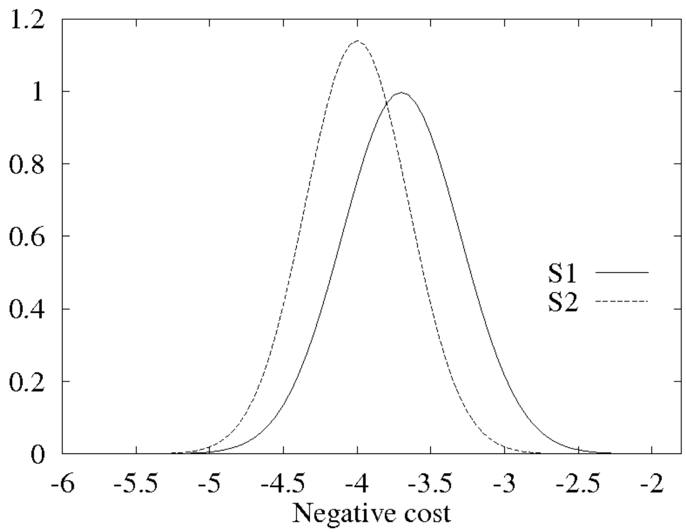


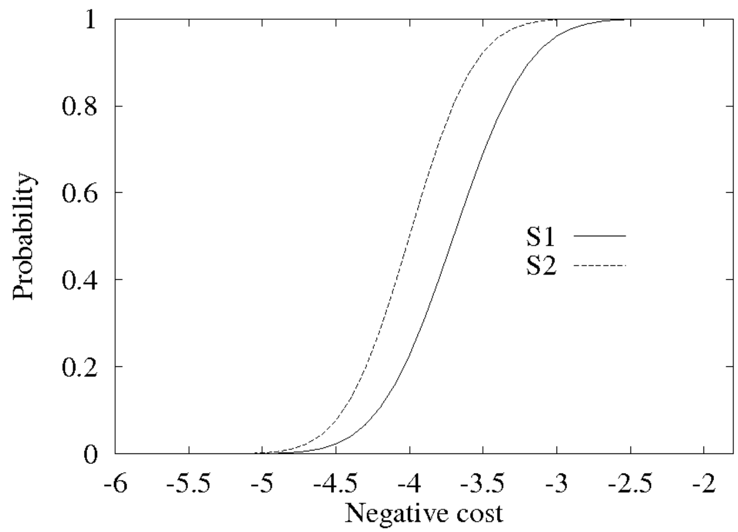
Stochastic dominance

- do not compare the worst possible attribute value in the first state with the best possible in the second,
- rather compare cumulative distribution functions of the attributes,
- distribution p_1 stochastically dominates distribution p_2 if
  – $\forall t \ \int_{-\infty}^{t} p_1(x)dx \leq \int_{-\infty}^{t} p_2(x)dx$,
- for U monotonously increasing with x it necessarily holds
  – $\int_{-\infty}^{\infty} p_1(x)U(x)dx \geq \int_{-\infty}^{\infty} p_2(x)U(x)dx$,
- for multiple attributes require stochastic dominance of a state in all attributes,

Stochastic dominance – example

- S1: the airport cost at location 1 3.7 ± 0.4mld,
- S2: the airport cost at location 2 4.0 ± 0.35mld,
- choose S1.





Value of information

- Agent rarely has complete information at its disposal
  - what questions shall it ask?
  - question $\rightarrow$ information with both value and costs (for test, time of an expert, etc.),
  - agent sorts questions by the difference between value and costs,
  - negatively valued questions not asked, actions taken based on the current information,
  - agent typically myopic – greedy decisions, disregards interactions between questions.
- How to compute the value of information?
  - has the given piece of information potential to change the current plan?
  - can be a modified plan significantly better than the current one?

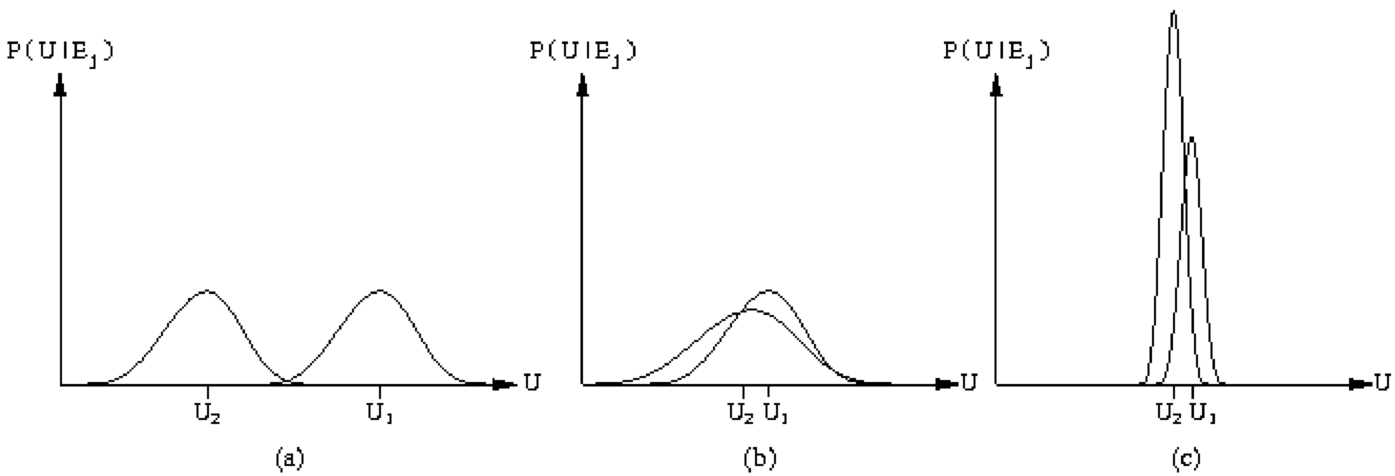
Value of information – qualitative distinctions

- 3 examples: actions A_1 and A_2, their expected utility U_1 and U_2,
- the utility distributions known a priori, E_j will bring the precise action utility,

  (a) choice is obvious, information worth little,

  (b) choice is unclear, information worth a lot,

  (c) choice is unclear, information worth little.



Value of information – general description

- current evidence E, current best action α
  possible outcomes of the action S_i, possible future observation E_j
- expected utility without knowing the value of E_j:
  $EU(\alpha|E) = \max_a \sum_i U(S_i)\ P(S_i|E, a)$
- if we knew that $E_j = e_{jk}$, then we would choose a different action $\alpha_{e_{jk}}$
- expected utility when knowing the value of E_j:
  $EU(\alpha_{e_{jk}}|E, E_j = e_{jk}) = \max_a \sum_i U(S_i)\ P(S_i|E, a, E_j = e_{jk})$
- when assessing the value of information, the value of E_j is unknown
  expected utility must aggregate over all possible values of E_j
  $VPI_E(E_j) = \left(\sum_k P(E_j = e_{jk}|E) EU(\alpha_{e_{jk}}|E, E_j = e_{jk})\right) - EU(\alpha|E)$
- VPI = value of perfect information
  - exact evidence about E_j can be obtained.

Value of information – characteristics

- VPI is always non-negative

$$\forall j, E \quad VPI_E(E_j) \geq 0,$$

  – even though it can lead into a state with a lower utility eventually,

- VPI is not additive

$$VPI_E(E_j, E_k) \neq VPI_E(E_j) + VPI_E(E_k),$$

- VPI is order-independent

$$VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,E_j}(E_k) = VPI_E(E_k) + VPI_{E,E_k}(E_j),$$

- the agent inquires information if: $\exists E_j \; VPI_E(E_j) > Cost(E_j)$,
- consequence
  – evidence gathering becomes a **sequential decision problem**.

Value of information – investment example

:: There are three types of investment opportunity (I): stocks (s), funds (f) and state bonds (b). Investment profit depends on whether markets (M) grow (↑), stay at the same level (resp. grow with inflation, →) or fall down (↓). Based on the values in table below compute the value of information about future market change.

M	$Pr(M)$	U(s,M)	U(f,M)	U(b,M)
$\uparrow$	0.5	1500	900	500
$\rightarrow$	0.3	300	600	500
$\downarrow$	0.2	-800	-200	500

Value of information – investment example

$$EU(\alpha|\{\}) = \max_{I\in\{s,f,b\}} \sum_{M\in\{\uparrow,\rightarrow,\downarrow\}} U(I,M)\ Pr(M) =$$

$$= \max(.5\times 1500 + .3\times 300 - 0.2\times 800,$$

$$.5\times 900 + .3\times 600 - 0.2\times 200, 500) =$$

$$= \max(680, 590, 500) = 680$$

$$EU(\alpha_{\uparrow}|\{\uparrow\}) = \max_{I\in\{s,f,b\}} U(I,\uparrow) = 1500\ \ (EU(\alpha_{\rightarrow}|\{\rightarrow\}) = 600, EU(\alpha_{\downarrow}|\{\downarrow\}) = 500)$$

$$VPI_{\{\}}(M) = \Big[\sum_{M\in\{\uparrow,\rightarrow,\downarrow\}} Pr(M) EU(\alpha_M|M)\Big] - EU(\alpha|\{\}) =$$

$$= .5\times 1500 + .3\times 600 + 0.2\times 500 - 680 = 1030 - 680 = 350$$

Summary

- rational agent takes action leading to the best expected result,
- its decisions can be based on three types of theory
  - probability – how to cope with observations in uncertain world,
  - utility – how to describe what to strive for, how to formulate goal,
  - decision making – actions to take based on stochastic model and goals,
- how to define utility function, what it is good for
  - complex worlds, states defined by attribute vectors, dominance decisions,
  - pieces of information to prefer, when to ask for them,
- people are just “approximately” rational
  - in complex worlds we must employ instincts and heuristics
    * automatic system that decides quickly, but imprecisely,
    * reflexive human system approaches the ideal view of rationality,
  - AI – both ideally rational agents and agents behaving like people.

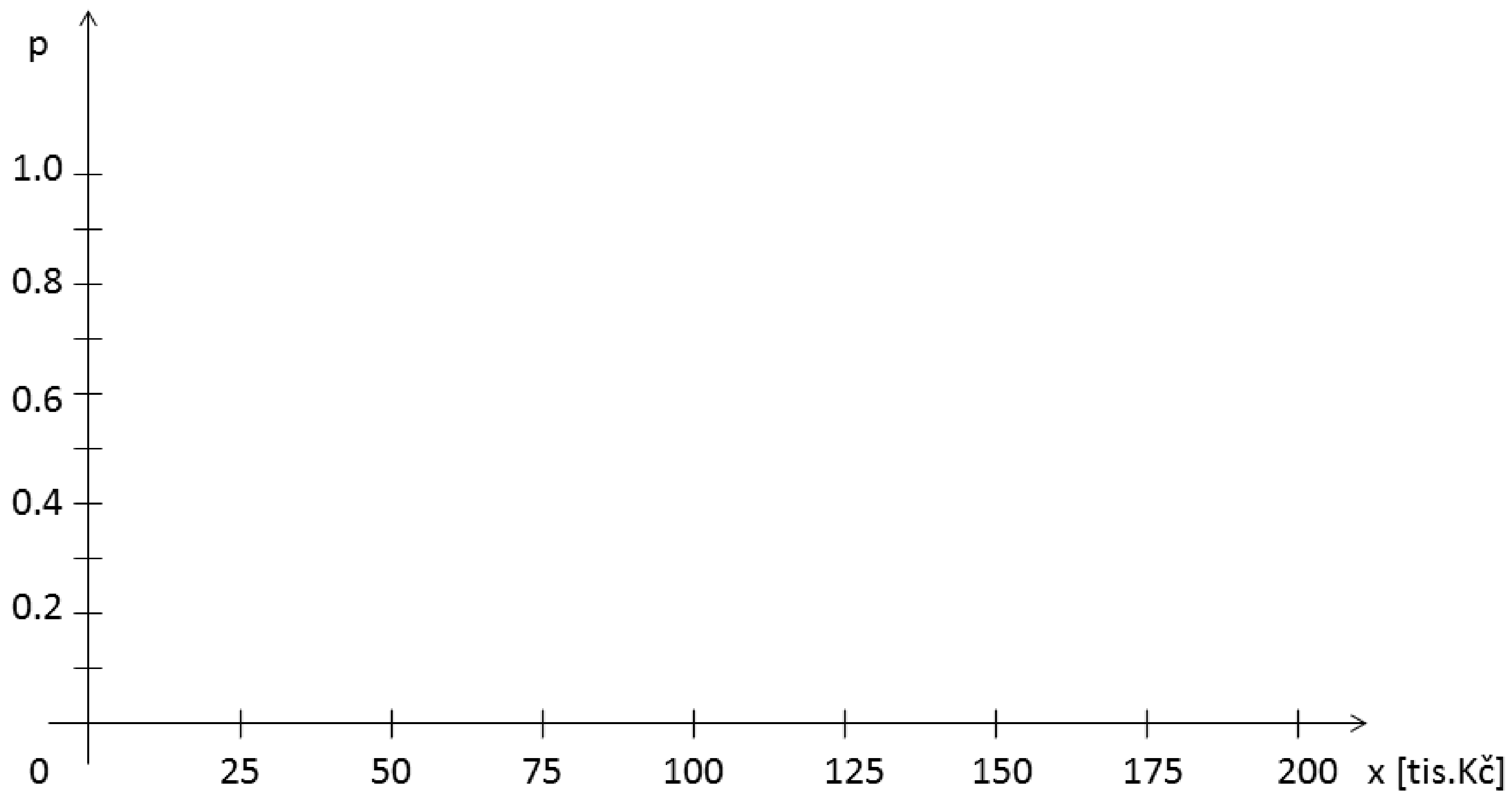
Recommended reading, lecture resources

:: Reading

- Russell, Norvig: **AI: A Modern Approach**, Rational Decisions
  - chapter 16, http://aima.eecs.berkeley.edu/slides-pdf/chapter16.pdf
  - book online on Google books (limited access): http://books.google.com/books?id=8jZBksh-bUMC.

Experimental ZUI utility curve

- For each x adjust p such that
  - half the students chooses lottery $[p, 200000Kc; 1-p, 0]$, half prefers x,
- what is the relationship between the curve and risk taking?



Utility and insurance

- On the concave curve the rational motivation for insurance can be shown.