# Rozhodování za neurčitosti – jednotlivá rozhodnutí

**Jiří Kléma**

Katedra kybernetiky,
FEL, ČVUT v Praze

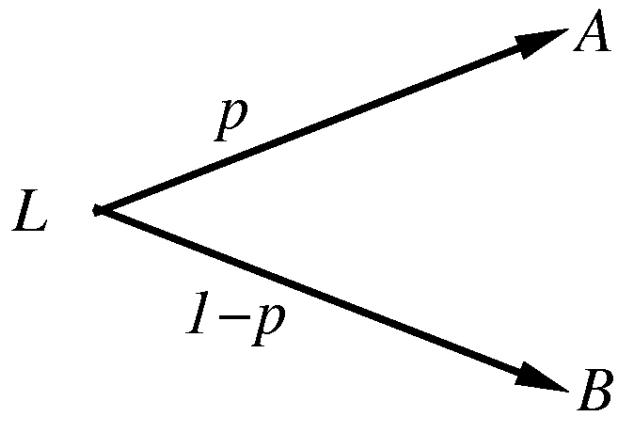
http://cw.felk.cvut.cz/doku.php/courses/a4b33zui/start

# Plán přednášky

- Umělí agenti s promyšleným a účelným rozhodováním (**racionální**)
  - jak je definovat,
  - jak se vyrovnat s nejistými výsledky akcí,
  - teorie rozhodování v kombinaci s teorií užitku,
  - pojmy: prémie, loterie, funkce užitku, přednost,
- rozumné a promyšlené rozhodování lidí
  - chovají se lidé při rozhodování rozumně?
  - peníze jako příklad ordinální míry užitku,
- multiatributový užitek
  - v obecném případě nelze odděleně hodnotit každý stav,
  - přednost a užitek se odvozuje od příznaků/atributů,
- hodnota informace
  - kdy se vyplatí vynaložit úsilí na její získání.

# Formalizace světa, základní pojmy

- agent rozlišuje mezi **stavy** světa $\{s_1, \ldots, s_n\}$,
- každý stav světa lze ohodnotit **prémií** $\{A, B, \ldots\}$,
- stavy světa agent ovlivňuje provedením **akcí** $\{a_1, \ldots, a_m\}$,
  - akce jsou stochastické, nelze předem jednoznačně určit výsledný stav,
  - aplikaci akce $a$ vedoucí s pstí $p$ do stavu $s_1$ ohodnoceného prémií $A$, s $p - 1$ do stavu $s_2$ ohodnoceného prémií $B$ lze popsat jako **loterii** $L_a = [p, A; (1 - p), B]$,
  - loterie bez prvku náhody odpovídá přímo prémii,
- racionalita: cílem agenta je zvolit akci vedoucí k nejvyšší prémii.



# Přednosti

- Jaký je význam prémie?
  - v obecném případě nemusí být číselná,
  - stačí symbolické prémie s relacemi **přednosti** (preferences)

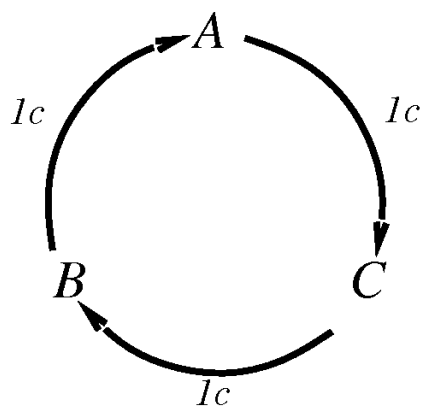
    $A \succ B \ldots A$ má přednost před $B$,

    $A \sim B \ldots A$ a $B$ jsou nerozlišitelné,

    $A \succeq B \ldots B$ nemá přednost před $A$,
- aby byl agent rozumný, relace přednosti musí splnit jistá **omezení** (constraints)
  - uspořádatelnost: $(A \succ B) \vee (B \succ A) \vee (A \sim B)$,
  - tranzitivitu: $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$,
  - spojitost: $A \succ B \succ C \Rightarrow \exists p\, [p, A;\ 1-p, C] \sim B$,
  - nahraditelnost: $A \sim B \Rightarrow [p, A;\ 1-p, C] \sim [p, B; 1-p, C]$,
  - monotónnost: $A \succ B \Rightarrow (p > q \Leftrightarrow [p, A;\ 1-p, B] \succ [q, A;\ 1-q, B])$,
  - rozložitelnost: $[p, A;\ 1-p, [q, B;\ 1-q, C]] \sim [p, A;\ (1-p)q, B;\ (1-p)(1-q), C]$.

# Tranzitivita přednosti jako nutná podmínka rozumnosti

- Porušení omezení vede k nerozumnému chování,
- příklad: agent s netranzitivní předností přijde o všechny své peníze
  - mějme agenta s přednostmi $A \succ B$, $B \succ C$, $C \succ A$,
  - ten je ochoten ke směně své prémie $C$ za cizí $B$ a přidá (libovolně) malý doplatek,
  - následně smění svou prémii $B$ za cizí $A$ a opět přidá malý doplatek,
  - po směně $A$ za cizí $C$ s doplatkem je ve výchozím stavu bez tří malých doplatků.



# Maximalizace středního užitku

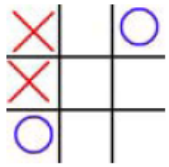
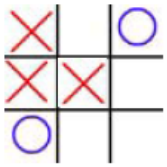
- von Neumann-Morgensternův teorém
  - pro přednosti splňující omezení existuje reálná funkce **užitku** (utility) $U$ taková, že

    $U(A) \geq U(B) \Leftrightarrow A \succeq B$

    (zachováme uspořádání prémií),

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

    (užitek loterie vyčíslíme jako střední užitek),
- princip maximalizace středního užitku
  - z více akcí (odpovídajících loterií) volíme tu, která maximalizuje střední užitek,
- (explicitní) zavedení užitku není nutnou podmínkou racionality
  - např. strategie agenta jako vyhledávací tabulka.

| Stav | Optimální akce |
|---|---|
|  |  |

# Od předností k funkci užitku

- fce užitku mapuje prémie (a tím i stavy) na reálná čísla
  - musí zachovat jejich lineární uspořádání dané přednostmi,
  - lze najít nekonečné množství funkcí vedoucích k identickému chování,
  - u deterministických prostředí (bez loterií) jde o jedinou podmínku

    $A \prec B \sim C \preceq D$ lze převést na $U_1$, $U_2$ s identickým chováním agenta

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

  - u loterií je chování agenta neměnné při lineární transformaci fce užitku

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

    * $U_1$ a $U_2$ z tabulky výše mj. zaměňují přednosti loterií $[0.5, A; 0.5, B]$ a $[0.9, A; 0.1, D]$,
- stadardizace zavedením normalizovaného užitku
  - nejlepší možný užitek $u_\top = 1.0$, nejhorší katastrofa $u_\perp = 0.0$,
  - mezilehlou prémii $A$ ohodnotíme užitkem $p$ tak, aby

    $A \sim [p, u_\top; (1-p), u_\perp]$.

# Lidé jako "racionální" agenti řízení penězi

- petrohradský paradox (Bernoulli, 1738)
  - kolik jste ochotni zaplatit jako vstupní poplatek za následující hru?
    - protihráč opakovaně hází standardní mincí dokud nepadne první hlava,
    - $n$ je počet hodů, získáte $2^n$ Kč, hra tím končí,

# Lidé jako "racionální" agenti řízení penězi

- petrohradský paradox (Bernoulli, 1738)
  - kolik jste ochotni zaplatit jako vstupní poplatek za následující hru?
    - protihráč opakovaně hází standardní mincí dokud nepadne první hlava,
    - $n$ je počet hodů, získáte $2^n$ Kč, hra tím končí,
  - pokud jsou peníze přímo fcí užitku, měli byste být ochotni vložit jakoukoli konečnou sumu
    - použijeme von Neumann-Morgensternův teorém
      $U(petrohrad) = U([p(h_1), U(h_1); p(h_2), U(h_2); \dots]) = \sum_{i=1}^{\infty} \frac{1}{2^i} 2^i = 1 + 1 + \dots = \infty$
  - toto je ale racionální pro málokoho
    - Bernoulli paradox vyřešil tím, že užitek peněz transformoval logaritmicky
      $U(k) = \log_2 k$
      $U(petrohrad) = U([p(h_1), U(h_1); \dots]) = \sum_{i=1}^{\infty} \frac{1}{2^i} \log_2 2^i = \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \dots = 2$
    - zpětnou transformací získáváme cenu hry: $2 = \log_2 k \Rightarrow k = 4$ Kč.

# Lidé jako "racionální" agenti řízení penězi

- experiment Tverskyho a Kahnemana (1982)
  - volte mezi loteriemi $L_1$ a $L_2$, následně mezi loteriemi $L_3$ a $L_4$

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

# Lidé jako "racionální" agenti řízení penězi

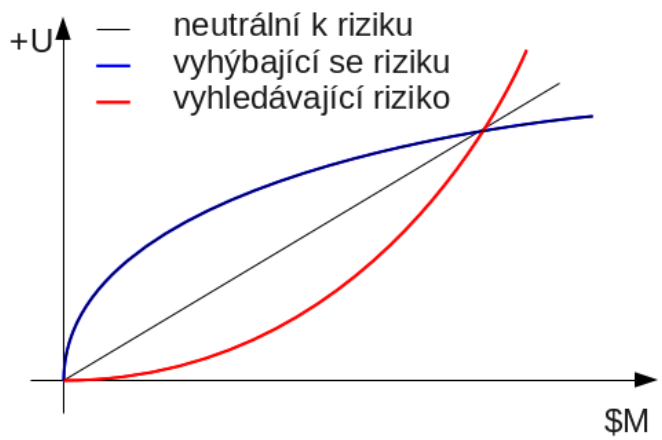
- experiment Tverskyho a Kahnemana (1982)
  - volte mezi loteriemi $L_1$ a $L_2$, následně mezi loteriemi $L_3$ a $L_4$

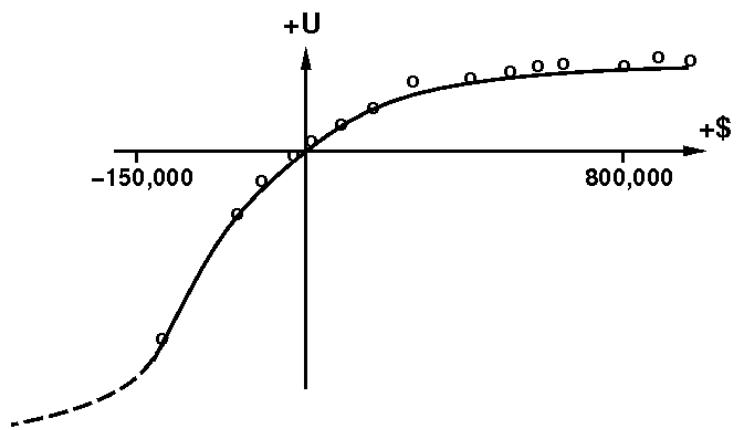
| Volba 1 | Volba 2 |
|---|---|
| $L_1 = [0.8, 80000Kc; 0.2, 0]$ | $L_3 = [0.2, 80000Kc; 0.8, 0]$ |
| $L_2 = [1, 60000Kc]$ | $L_4 = [0.25, 60000Kc; 0.75, 0]$ |

  - většina lidí upřednostní loterii $L_2$ před $L_1$ a $L_3$ před $L_4$
    - z hlediska racionality podivné, za předpokladu $U(0Kc) = 0$ platí současně
      volba 1: $0.8U(80000Kc) < U(60000Kc)$,
      volba 2: $0.8U(80000Kc) > U(60000Kc)$,
    - neexistuje žádná funkce užitku konzistentní s oběma volbami,
  - možné závěry
    - lidé jsou iracionální,
    - daná analýza nebere v úvahu lítost nad ztrátou velmi pravděpodobného zisku ad $L_2$,
    - proto se lidé snaží vyhnout riziku u pravděpodobných událostí,
    - a naopak riskují hodně u málo pravděpodobných událostí.

# Lidé jako “racionální” agenti řízení penězi

- peníze nejsou přímou funkcí užitku
  - jejich použití lidmi často nemaximalizuje střední užitek, $U([p_1, S_1; \ldots; p_n, S_n]) \neq \sum_i p_i U(S_i)$
  - obvykle pozorujeme snahu vyhnout se riziku, tedy loterii, $U([p_1, S_1; \ldots; p_n, S_n]) < \sum_i p_i U(S_i)$
- **křivka užitku**
  - nelineárně převádí peníze na užitek,
  - hledáme pst $p$, pro kterou daná osoba nerozlišuje mezi cenou $x$ a loterií $[p, \$M; (1-p), \$0]$, $\$M$ je velké



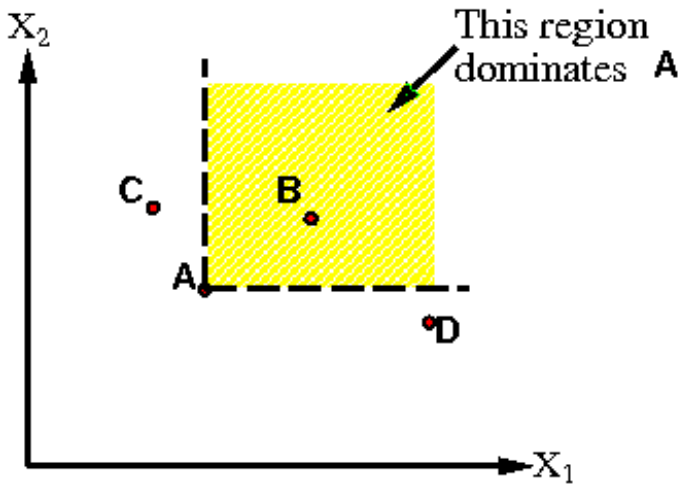



# Vícepříznakové funkce užitku

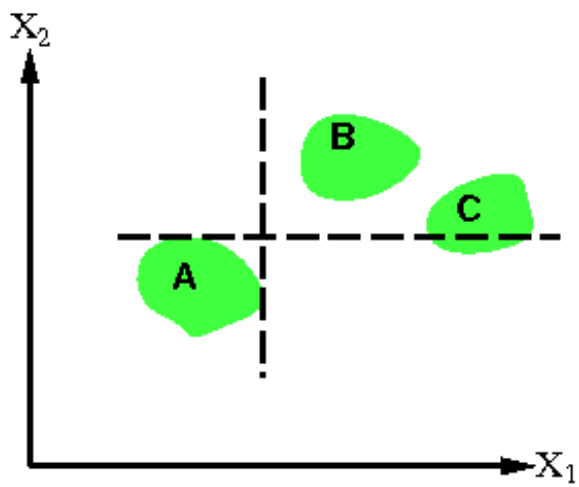
- Často nelze přímo přiřadit prémii každému stavu
  - stavů může být příliš resp. nekonečně mnoho,
  - stavy ale obvykle lze popsat příznaky
    (bezpečnost, hlučnost, cena při výběru místa pro nové letiště),
- func užitku má pak více parametrů
  - $U(X_1, \ldots, X_n)$ (parametry resp. příznaky namísto stavu),
  - $n$ atributů s $m$ hodnotami definuje $m^n$ stavů,
  - vyčíslení užitku lze ale zjednodušit předpoklady **pravidelností** v přednostech
    * monotonicity předností při změně jediného příznaku, ostatní příznaky konstantní,
      $x \geq y \Rightarrow U(X_1, \ldots, X_i = x, \ldots, X_n) \geq U(X_1, \ldots, X_i = y, \ldots, X_n)$,
    * různých typů nezávislosti mezi příznaky ve vlivu na přednosti,
      definice stavů: $A \sim (x_1, y_1)$, $B \sim (x_2, y_1)$, $C \sim (x_1, y_2)$, $D \sim (x_2, y_2)$
      přednosti při nezávislosti: $(A \succ B \Rightarrow C \succ D) \wedge (A \succ C \Rightarrow B \succ D)$
  - pravidelnosti v přednostech mohou odpovídat zjednodušené fci užitku
    * $U(x_1, \ldots, x_n) = f[f_1(x_1), \ldots, f_n(x_n)]$,
    * $f$ je obvykle jednoduchá, např. sčítání.

# Ostrá dominance

- předpoklad: $U$ monotónně rostoucí ve všech příznacích,
- ostrá dominance stavu $B$ před stavem $A$
  - $\forall i \; X_i(B) \geq X_i(A) \Rightarrow f_i(X_i(B)) \geq f_i(X_i(A)) \Rightarrow U(B) \geq U(A)$
- př.: jedno z míst pro nové letiště je bezpečnější, méně hlučné, i cenově náročné než jiné,
- jen zřídka použitelná v praxi
  - použitelnost snižuje mj. nejistota při odhadu hodnot příznaků.



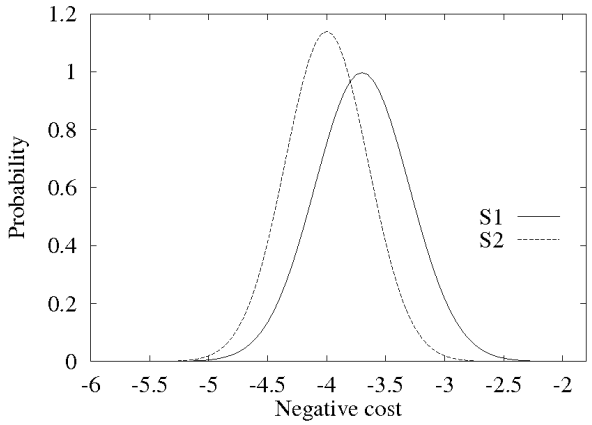

Deterministic attributes

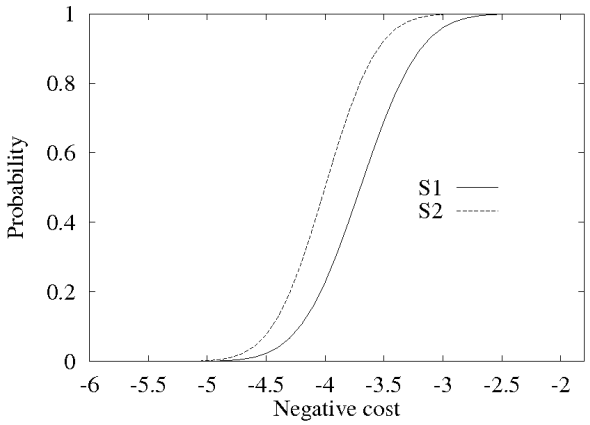



Uncertain attributes

# Stochastická dominance

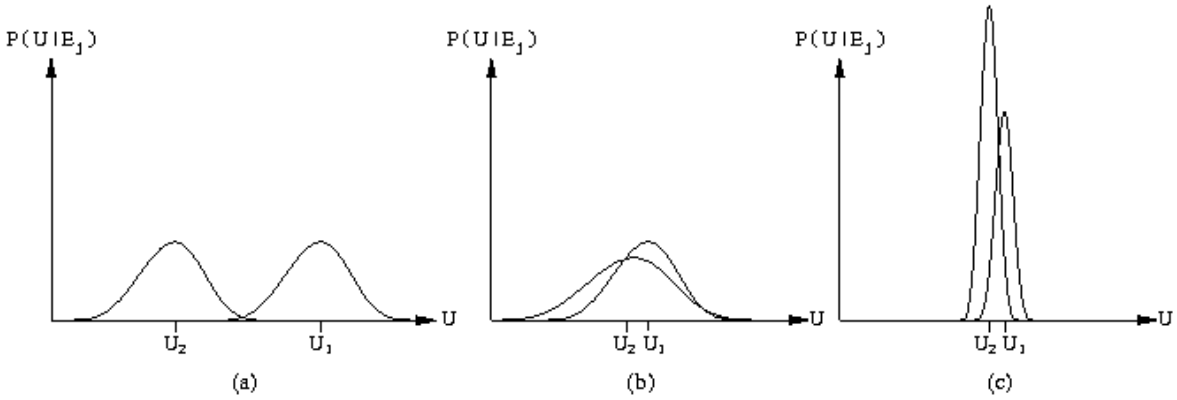
- nesrovnáváme nejhorší možnou hodnotu příznaku v jednom stavu s nejlepší ve druhém,
- vycházíme ze srovnání kumulativních distribučních funkcí příznaků,
- distribuce $p_1$ stochasticky dominuje distribuci $p_2$ tehdy, pokud
  - $\forall t \ \int_{-\infty}^{t} p_1(x)dx \leq \int_{-\infty}^{t} p_2(t)dt$,
- pro $U$ monotónně rostoucí s $x$ pak nutně platí, že
  - $\int_{-\infty}^{\infty} p_1(x)U(x)dx \geq \int_{-\infty}^{\infty} p_2(x)U(x)dx$,
- pro více atributů vyžadujeme stochastickou dominanci stavu na všech atributech,
- př.: cena letiště na místě 1 je $3.7 \pm 0.4$mld, na místě 2 je $4.0 \pm 0.35$mld – volíme 1.



# Hodnota informace

- Zřídkakdy jsou informace agenta úplné
  - jaké otázky má agent klást?
  - otázka $\rightarrow$ informace s hodnotou, ale i náklady (za test, čas experta apod.),
  - agent řadí otázky dle rozdílu hodnota-náklady,
  - záporně hodnocené otázky neklade a volí akci na základě dané informace,
  - typicky je krátkozraký – neuvažuje interakce mezi otázkami.
- Jak vyčíslit hodnotu informace?
  - má informace potenciál změnit stávající plán?
  - může být nový plán výrazně lepší než ten stávající?



# Hodnota informace – obecné vyjádření

- současná pozorování $E$, současná nejlepší akce $\alpha$
  možné výsledky akce $S_i$, možné další pozorování $E_j$
- očekávaný užitek bez zjištění $E_j$:
  $EU(\alpha|E) = \max_a \sum_i U(S_i)\, P(S_i|E, a)$
- budeme-li znát $E_j = e_{jk}$, pak můžeme volit jinou akci $\alpha_{e_{jk}}$
- očekávaný užitek při zjištění $E_j$:
  $EU(\alpha_{e_{jk}}|E, E_j = e_{jk}) = \max_a \sum_i U(S_i)\, P(S_i|E, a, E_j = e_{jk})$
- při rozhodování o hodnotě informace neznáme skutečnou hodnotu $E_j$
  očekávaný zisk musíme počítat přes všechny stavy náhodné proměnné $E_j$
  $VPI_E(E_j) = \left(\sum_k P(E_j = e_{jk}|E) EU(\alpha_{e_{jk}}|E, E_j = e_{jk})\right) - EU(\alpha|E)$
  - VPI je vždy nezáporná, i když reálně může nová informace vést do stavu s nižším užitkem,
  - VPI není aditivní, $VPI_E(E_j, E_k) \neq VPI_E(E_j) + VPI_E(E_k)$,
  - hodnota VPI ale nezávisí na pořadí
    $VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,E_j}(E_k) = VPI_E(E_k) + VPI_{E,E_k}(E_j)$,
  - k dotazu dojde, pokud: $\exists E_j\, VPI_E(E_j) > Cost(E_j)$.

# Hodnota informace – investiční příklad

:: Existují tři typy investiční příležitosti (I): akcie (a), smíšený fond (f) a státní dluhopisy (d). Ziskovost investice závisí na tom, zda trhy (T) porostou ($\uparrow$), zůstanou na stejné úrovni (resp. budou růst s inflací, $\rightarrow$) nebo klesnou ($\downarrow$). Na základě hodnot v tabulce určete hodnotu informace o chování trhů.

| $T$ | $Pr(T)$ | U(a,T) | U(f,T) | U(d,T) |
|---|---|---|---|---|
| $\uparrow$ | 0.5 | 1500 | 900 | 500 |
| $\rightarrow$ | 0.3 | 300 | 600 | 500 |
| $\downarrow$ | 0.2 | -800 | -200 | 500 |

# Hodnota informace – investiční příklad

:: Existují tři typy investiční příležitosti (I): akcie (a), smíšený fond (f) a státní dluhopisy (d). Ziskovost investice závisí na tom, zda trhy (T) porostou ($\uparrow$), zůstanou na stejné úrovni (resp. budou růst s inflací, $\rightarrow$) nebo klesnou ($\downarrow$). Na základě hodnot v tabulce určete hodnotu informace o chování trhů.

| $T$ | $Pr(T)$ | U(a,T) | U(f,T) | U(d,T) |
|---|---|---|---|---|
| $\uparrow$ | 0.5 | 1500 | 900 | 500 |
| $\rightarrow$ | 0.3 | 300 | 600 | 500 |
| $\downarrow$ | 0.2 | -800 | -200 | 500 |

$$
\begin{aligned}
EU(\alpha|\{\}) &= \max_{I\in\{a,f,d\}} \sum_{T\in\{\uparrow,\rightarrow,\downarrow\}} U(IT)\ Pr(T) = \\
&= \max(.5\times 1500 + .3\times 300 - 0.2\times 800, .5\times 900 + .3\times 600 - 0.2\times 200, 500) = \\
&= \max(680, 590, 500) = 680 \\
EU(\alpha_{\uparrow}|\{\uparrow\}) &= \max_{I\in\{a,f,d\}} U(I,\uparrow) = 1500 \ \ (EU(\alpha_{\rightarrow}|\{\rightarrow\}) = 600, EU(\alpha_{\downarrow}|\{\downarrow\}) = 500) \\
VPI_{\{\}}(T) &= \Big(\sum_{T\in\{\uparrow,\rightarrow,\downarrow\}} Pr(T)EU(\alpha_T|T)\Big) - EU(\alpha|\{\}) = \\
&= .5\times 1500 + .3\times 600 + 0.2\times 500 - 680 = 1030 - 680 = 350
\end{aligned}
$$

# Shrnutí

- racionální agent volí akci, která vede k nejlepšímu očekávanému výsledku,
- může se přitom opírat o tři teorie
  - pravděpodobnosti – jak nakládat s pozorováními v nejistém světě,
  - užitku – jak popsat o co usilovat, jak by mohl vypadat cíl,
  - rozhodování – jaké akce volit na základě stochastického modelu a cílů,
- jak definovat funkci užitku, k čemu ji lze využít
  - složitější světy, popis stavů atributy, rozhodnutí o dominanci,
  - kdy získávat informaci, na jakou informaci se zaměřit,
- lidé jsou pouze "přibližně" racionální
  - ve složitém světě musí rozhodovat i na základě instinktů, heuristik,
    - automatický systém rozhoduje rychle, ale velmi přibližně,
    - reflexivní systém se blíží ideální racionalitě,
  - UI jako obor modelující ideální racionální agenty vs agenty chovající se jako lidé.

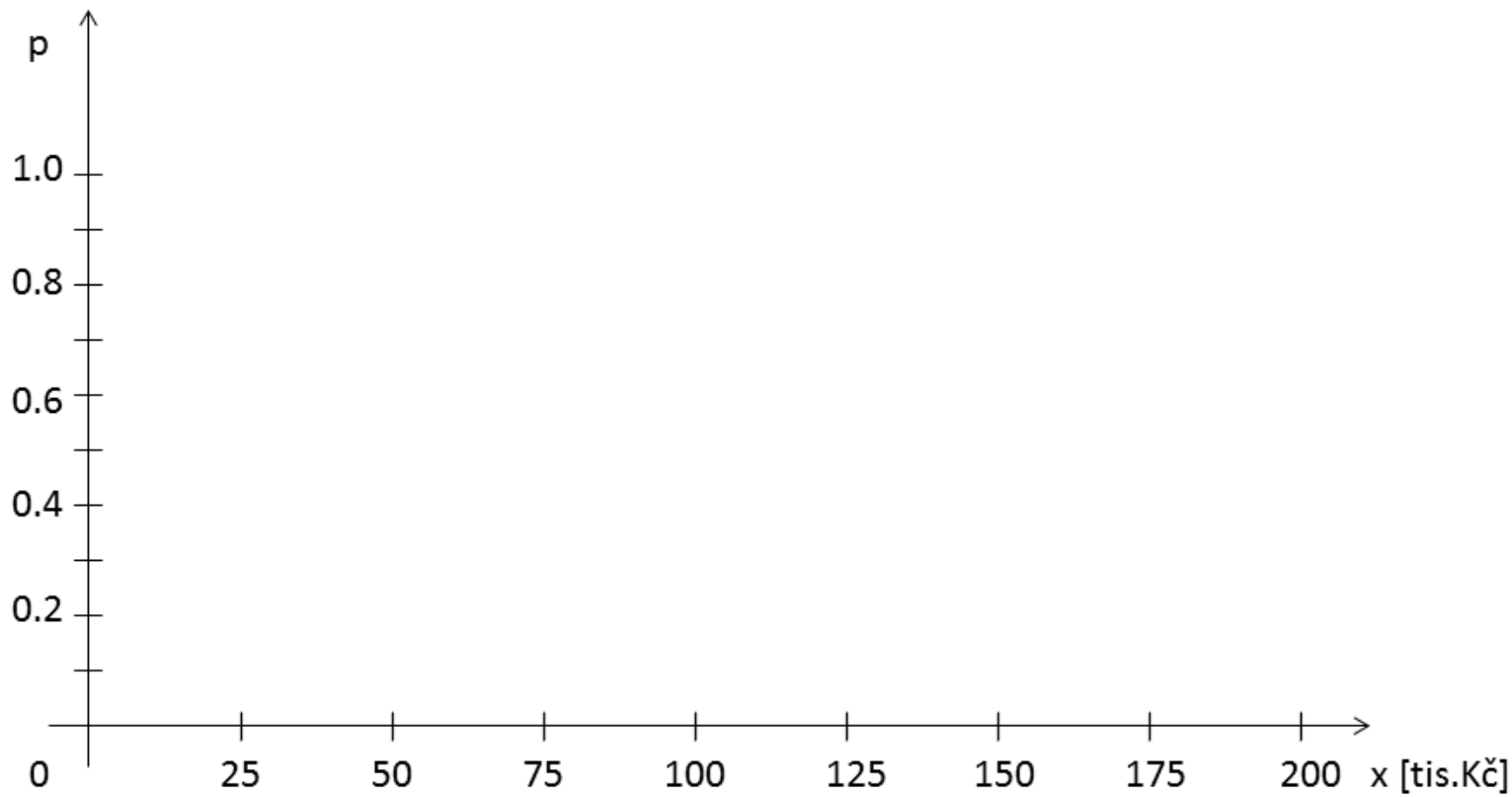
# Doporučené doplňky – zdroje přednášky

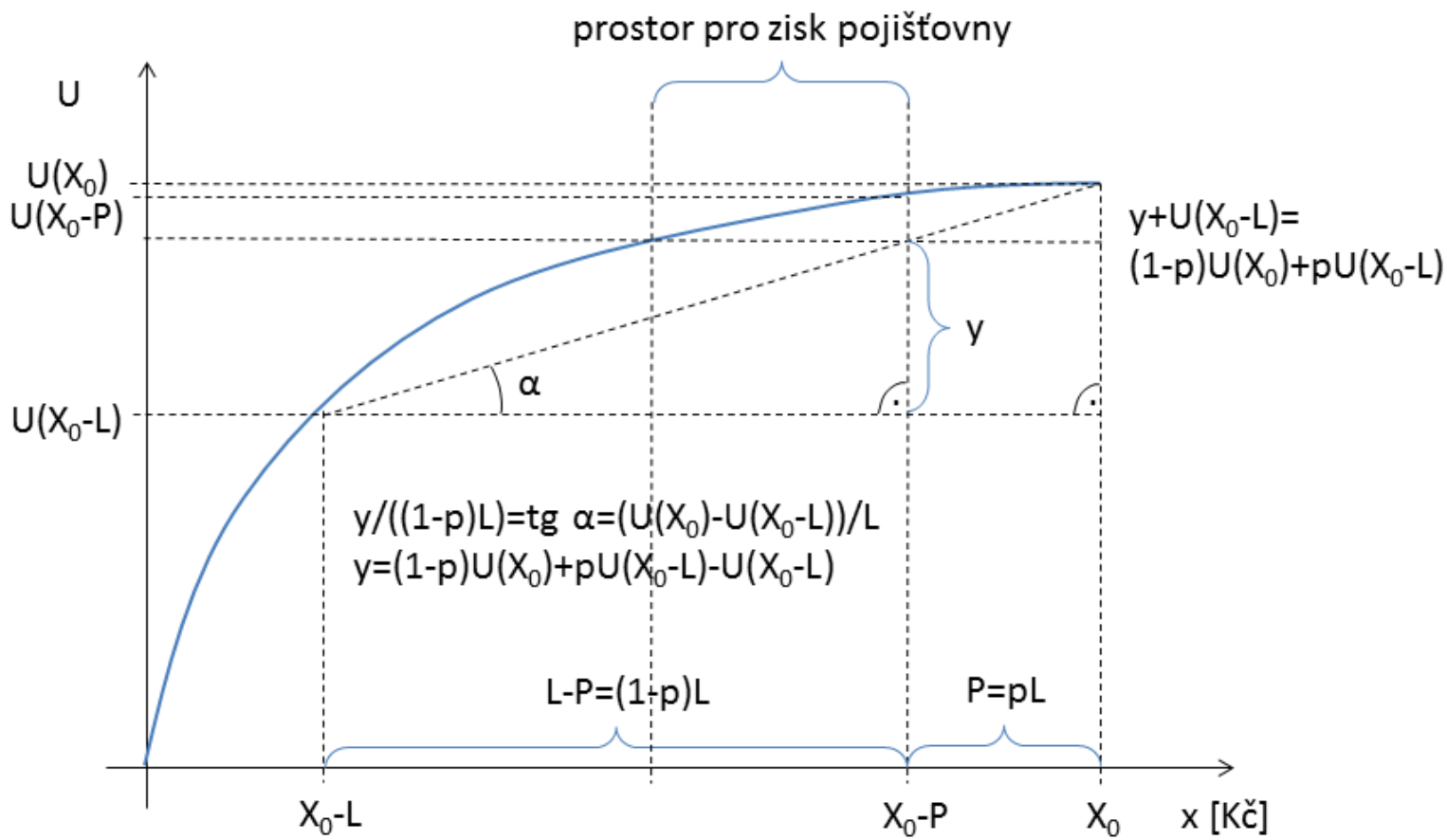
:: Četba

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

# Experimentální ZUI užitek

- Pro každé $x$ přizpůsobte $p$ tak, aby
  - polovina kurzu volila loterii $[p, 200000Kc; 1-p, 0]$ a druhá polovina $x$,
- jak průběh souvisí s ochotou riskovat?



# Užitek a pojištění

- Na konkávní křivce lze ukázat důvod pro pojištění.