# Přednášky z Teorie her (THE)
# Game Theory

Martin Hrubý

Brno University of Technology
Brno
Czech Republic

September 21, 2020

# Úvod - Matematická teorie her

- Vědní disciplína zkoumající racionální lidské chování při *rozhodování* (jsou i další složky/projevy inteligence).
- Multi-oborová disciplína - matematika, ekonomie, sociologie, politologie, biologie, informatika, ...
- Z mnoha důvodů je snaha chování lidí studovat, analyzovat a modelovat
- cílem je chování pochopit a predikovat (dle možností – nebudeme věštit)
  - v druhé části semestru budeme zkoumat mechanismy (a jejich návrh), které způsobí, že se racionální jedinec bude chovat tak, jak chceme my

Poznámka: Teorie her není nauka o programování počítačových her.

# Z historie modelů rozhodování

Talmud(0-500AD) – návody a přikázání v obchodu, právu a běžném životě. Proslavil se zejména problém "marriage contract problem" – muž měl tři ženy, v určitém okamžiku zemřel a zanechal odkaz $E$. Manželky měly kontrahována pro tento případ různá dědictví (100, 200, 300).

Talmud doporučuje spravedlivé dělení v případě, že $E$ je menší než suma kontraktů.

V případě $E = 100$ je to $(33.3, 33.3, 33.3)$, $E = 200$ pak (50,75,75), $E = 300$ pak $(50, 100, 150)$.

Až v roce 1985 se ukázalo, že řešení Talmudu odpovídá poznatkům o kooperativních hrách (nucleolus).

Brams, S. J.: The Win-Win Solution: Guaranteeing Fair Shares to Everybody, v knihovně

# Z historie (poznatky, výsledky)

1713 James Waldegrave ukázal první známé řešení ve smíšených strategiích pro 2-hráčové hry (hra Le Her).

1785 Condorcetův paradox

1838 Augustin Cournot publikoval model oligopolu, který značně předstihl dobu. Odpovídá dnešnímu Nashovu ekvilibriu.

1871 V knize "The Descent of Man, and Selection in Relation to Sex" Charles Darwin publikoval první (zatím intuitivní) herně-teoretickou argumentaci pro evoluční biologii.

1913 Zermelo's Theorem

1921-27 Emile Borel – publikoval čtveřici článků o strategických hrách. Položil první formální definici smíšených strategií v minimax řešení her.

1928 John von Neumann dokázal minimax theorem v článku Zur Theorie der Gesellschaftsspiele.

# Z historie (poznatky, výsledky)

1930 F. Zeuthen: Problems of Monopoly and Economic Warfare. V kapitole IV naznačil řešení "bargaining problemu", které později Harsanyi ukázal ekvivalentní k Nash's bargaining solution.

1944 John von Neumann and Oskar Morgenstern: Theory of Games and Economic Behavior. Považována za jednu z nejvýznamnějších matematických knih 20. století.

1950 Melvin Dresher and Merrill Flood (Rand Corporation): the Prisoner's Dilemma.

1950-53 John F. Nash publikoval své řešení v nekooperativních hrách a v teorii vyjednávání (bargaining).

1952-53 L. S. Shapley (Rand Corporation): jádro v kooperativních hrách.

Pak se to značně rozjelo...

# Současný stav udělených Nobelových cen za teorii her

1972 J.R. Hicks, K.J. Arrow
1994 J.C. Hersanyi, J.F. Nash, R. Selten
1996 J.A. Mirrlees, W. Vickrey
2005 R.J. Aumann, T.C. Schelling
2007 L. Hurwicz, E. S. Maskin, R.B. Myerson
2012 A. E. Roth, L. S. Shapley
2014 J. Tirole



# Teorie her pro informatiky na FITu

- Teorie her je disciplína zapadající do *inteligentních systémů* (UI).
- Umožní nám formalizovat a modelovat specifickou část inteligence.
- V mnoha ohledech nám ulehčí pochopení lidského chování – ovšem vždy na úrovni modelů.
- PSYCHOvědy: psychologie, psychiatrie, sociologie, politologie, ekonomie jsou z pohledu informatiky soft-vědy.
  - Výjimky: experimentální ekonomie/teorie her.
- Nás zajímá matematika, modelování a simulace, algoritmizace, datové struktury, extrémně náročné výpočty, ...

# Cíl předmětu Teorie her (THE)

- Rozšířit matematické vzdělání a přemýšlení.
- Naučit se analyzovat rozhodovací situace, hledat v nich možná řešení.
- Předmět je úvodem do problematiky a má mít co možná nejširší záběr.
- Teorie her má taky své pod-obory, zkusíme projít všechny (většinu).
- Student by měl být schopen analyzovat zadanou situaci a vytvořit její matematický/počítačový model. Na základě modelu by měl být schopen poznat chování modelovaného systému, případně predikovat pravděpodobné chování hráčů.
- Výsledek studia her...

# Úvod – Struktura a organizace předmětu

- Přednášky – 12 přednášek, 13. opakování.
- Projekt (nutný pro zápočet)
  - Studijní – nastudování vybrané kapitoly nad rámec přednášek, zpráva (originální, vlastní výklad problematiky).
  - Aplikační – implementace konkrétního modelu, simulační studie.
  - Infiltrační – návštěva existující instituce tematicky příbuzné předmětu, zpráva.
  - Termín – viz organizace předmětu.
- Zkouška – obsah, termín (lze dohodnout).
- Konzultace, zpětná vazba (reflexe a vyjednávání).

Studijní literatura:

- Knihy ve fakultní knihovně (grant FRVŠ FR0110/2010/F1).
- Studiní texty Teorie her (pracovní stránka kurzu).

# Přednášky 1/3

Jedna přednáška vždy znamená jedno téma (tzn. některé jsou informačně hustější, jiné volnější).

1. Úvod – organizace předmětu, úvod do nezbytných matematických pojmů, základy teorie rozhodování.
2. Hry v normální formě (s nenulovým součtem) – základní pojmy, modelování rozhodovacích situací, základy analýzy strategických her, základní modely oligopolu (Cournotův a Bertrandův model – analytické a simulační řešení).
3. Hry v normální formě (s nulovým součtem) – věta o Mini-maxu, sedlový bod, řešení. Lineární optimalizace.
4. Algoritmy pro řešení strategických her – algoritmy výpočtu řešení hry, věta o existenci řešení (ekvilibria) a její *důkaz*.

Cíl: modelovat strategické situace. Poznat herně-teoretickou matematiku.

# Přednášky 2/3

7. Sekvenční (tahové) hry – základní pojmy, modely, řešení, Stackelbergův model oligopolu. Důvěryhodná hrozba. SPNE.
8. Kooperativní hry a vyjednávání (bargaining) – pojmy, modely, řešení. Jádro hry, Shapley value, Voting power, Nash bargaining solution.
9. Opakované hry (repeated games) – vliv opakování hry na strategické rozhodování hráčů. Opakování a kooperace. Korelované ekvilibrium – motivace, definice, aplikace a výpočet.

Cíl: jaký vliv má spolupráce na efektivnost našeho života?

# Přednášky 3/3

10. Mechanism design – základní pojmy, návrh pravidel, strategická manipulace. Teorie veřejné volby (Public Choice) – volební mechanismy, Condorcetův paradox, Arrowův paradox (a další).
11. Teorie aukcí – mechanismy aukcí, návrh mechanismu aukce, modely.
12. Evoluční biologie – herní podstata naší evoluce, samoorganizace v populacích jedinců. Stejné principy jsou ovšem platné např. v modelování zatížení kom. sítí.
13. Aplikace, případové studie – aplikace v modelování energetických trhů.

Cíl: poznat, modelovat a *navrhovat pravidla* ve strategických situacích. Případová studie.

# Úvod do matematických pojmů

V THE budeme formálně definovat pojmy, algoritmy a postupy.

- Diskrétní matematika.
- Algebra.
- Matematická analýza.
- Pravděpodobnost. Statistika.
- Operační výzkum.

Opakujeme a aplikujeme již získané matematické znalosti.
THE je matematika našeho života.

# Úvod do matematických pojmů

- Pojem množina – matematický objekt $A$, pro který platí, že jsme schopni pro každý ($\forall$) objekt $o$ jednoznačně určit, zda-li objekt $o$ je nebo není prvkem $A$
  - $o \in A$, $o \notin A$
  - Množiny obvykle zapisujeme velkým písmenem.
- Množina není seznam. Každý prvek je tam pouze jednou (na rozdíl od multi-množiny).
- $A = \{a_1, a_2, ..., a_x\}$. Prázdná množina $\emptyset$.
- $|A|$ značí počet prvků množiny (kardinalita množiny). $|A| = x, x \in \mathbb{N}^+ \cup \{0\}$.
- Definice konkrétní množiny: $B = \{a \in A | condition(a)\}$
- Podmnožina: $\subset, \subseteq$
- Potenční množina: $2^A$. Kolik má prvků? Kolik je $|2^\emptyset|$?

# Množinové operace

- $C = A \cup B \Leftrightarrow \forall c \in C : c \in A \vee c \in B$
- $C = A \cap B \Leftrightarrow \forall c \in C : c \in A \wedge c \in B$
- $C = A \setminus B \Leftrightarrow \forall c \in C : c \in A \wedge c \notin B$
  $C = \{c \in A | c \notin B\}$

Operátory $\cup, \cap, \sum$ budeme používat s iterační proměnnou:

$$C = \bigcup_{i \in N} fun(i)$$

$$I = \sum_{i=1}^{N} fun(i)$$

$$I = \sum_{i}^{N} fun(i)$$

# Kartézský součin (angl. product), relace

- Kartézský součin (značíme $\times$ nebo $\prod$) množin $A$ a $B$ je množina uspořádaných dvojic $C = A \times B = \{(a, b) | a \in A \wedge b \in B\}$.
- Relace – toto lidské slovo má opravdu hodně významů...
- Pro matematiku je binární relace podmnožinou kartézského součinu množin $A$ a $B$ (tzn. je to množina uspořádaných dvojic). Definujeme i n-ární relace $R \subseteq A \times B \times C \times \ldots$.
- $R \subseteq A \times B$, $R \subseteq A \times A$, $R \subseteq A^2$.
- Notace: $x = (a, b)$, $x \in R$, $aRb \Leftrightarrow x \in R$.
- Různé vlastnosti relací.

Příklad:
$A = \{1, 2, 3\}; B = \{x, y\}$
$A \times B = \{(1, x), (1, y), (2, x), (2, y), (3, x), (3, y)\}$
$R \subseteq A \times B$ např. $R = \{(1, x), (1, y)\}$

# Další matematické znalosti

Bylo by dobré umět/chápat:

- Přečíst/pochopit matematický zápis.
- Úpravy algebraických výrazů (odvozování, zjednodušování).
- Vyřešení lineární, kvadratické a jednoduché diferenciální rovnice. Základní řešení soustav lineárních rovnic.
- Derivovat. Hledání extrémů funkcí.

Nová znalost (možná) bude: lineární programování.

# Základní pojmy THE: model situace

- Hra – hra je strategická interakce dvou a více hráčů.
- Hráč – hráč je účastník hry.
- Strategie (akce) – varianty možného chování hráče.
- Zisk/Užitek – hráčův výnos ze hry (užitek je míra uspokojení ze zisku).
- Preference – rozlišení strategií nebo užitků.
- Racionalita hráče – základní předpoklad pro řešitelnost situace.
- Hraní hry – jakým způsobem hráči realizují své tahy.
- Řešení hry – pravděpodobný způsob a výsledek chování hráčů.

Hráč provádí rozhodovaní nad strategiemi s cílem dosáhnout *individuálního optimálního výsledku* (užitku) ve hře.

Selfish (sobecký) agent/hráč. Altruismus.

# Hra a hráč

Hra v pojetí TH je *model* situace – tzn. je to zjednodušení reality. Realitu abstrahujeme na hráče, pravidla hry, strategie, preference, užitek a herní koncept.

- Hra má pravidla – víme, jak se hra hraje.
- Hráči mohou mít o hře různou míru informace.
- Nezkoumáme, jestli je hráč ve hře dobrovolně, jestli chce hrát, jaké má pocity.
- Všechno, co hráč ví, co zkoumá, co preferuje, všechno je obsaženo v modelu té situace (protivníci, strategie, preference, užitek, řešení hry).
- Hry proti přírodě – příroda je hráč bez strategického chování (o jeho chování nelze říct nic víc, než že svou strategii volí náhodně).

# Pojmy model a simulace, opakování

- Model systému $A$ je jiný systém $B$, který je mu jaksi podobný.
- Model je vždy zjednodušení reálné situace/systému.
- Modelování je *proces zdůvodněného zjednodušování* sys. $A$.
- Simulace je proces experimentování s modelem ($A \rightsquigarrow B$) jehož cílem je získat nové znalosti o $A$ prostřednictvím $B$.

Závěr: budeme zkoumat inteligentní rozhodování lidí prostřednictvím modelů. Musíme se vyrovnat s faktem, že naše modely budou pouze zjednodušovat realitu s tou nadějí, že zdůrazní alespoň ty aktuálně relevantní aspekty.

Význam pojmu *simulace* v našem pojetí: analýza hry a predikce výsledku hry. U počítačových modelů herních situací je to opět provedení experimentu s konkrétními vstupy a konkrétním cílem.

# Strategie hráče

- Hráč provádí rozhodování.
- Musí poznat své možnosti – identifikace jeho možností je součást hráčových analytických schopností a taky jeho inteligence.
  - (v realitě): Poznání těchto možností je de facto hráčovo know-how.
  - V počítačovém modelování rozhodovací situace obvykle musí modelář sám strategie určit.
- Všechny strategie hráče $i \in Q$ ve hře tvoří jeho *množinu strategií* – $S_i$.
- Množiny strategií hráčů jsou obecně jiné (až disjunktní, tzn. $A \cap B = \emptyset$).

Budeme zkoumat tzv. množinu strategických profilů

$$S = \prod_{i \in Q} S_i$$

# Užitek a preference

Užitek:

- Užitek (angl. utility) je vztažen ke každé hráčově strategii. Na základě užitku může hráč provádět rozhodování.
- Výsledek (řešení, výstup, outcome) hry je společný užitek všech hráčů.
- Zisk a užitek (později vše sloučíme do jednoho pojmu).

Preference:

- Preference je fenomén veškerého modelování rozhodovacích situací.
- Preference není předmětem rozhodování. Je to *vstup* do modelu. Každý má jiné preference (co to znamená?).

**Preference hráče nemusíme chápat, ale musíme je poznat a správně modelovat.**

Preference masochisty.

# Vyjádření užitku

- Obvykle velmi dobře chápeme užitek finanční – proč??? Chápeme jasně hodnotu peněz a jsme schopni srovnávat dva různé užitky (1 CZK versus 1000 CZK).
- Užitek je (subjektivní) míra uspokojení plynoucí ze spotřeby statků.

Přístupy:

- Kardinalistická teorie – jsme schopni vnímat velikost rozdílu mezi užitky.
- Ordinalistická teorie – nejsme schopni porovnávat užitky metrikou, ale alespoň poznáme lepší užitek.
- Zkoumáme preference na množině strategií nebo na množině možných užitků. Mnohdy je to zaměnitelné.
- Výrazné problémy s preferencí začnou až u multi-kriteriálního rozhodování.

# Racionalita hráče

Racionalita je pojem, který pořádně nechápou ani zkušení herní teoretici.

- Racionalita hráče je základní předpoklad pro modelovatelnost situace (racionalita je deterministická).
- Pojem s rozsáhlým významem a množstvím definic a pohledů.
- Racionální jedinec je schopen identifikovat své cíle a podniká kroky k jejich dosažení.
- ... je schopen promyslet situaci v celé její komplexnosti (nedělá ukvapená rozhodnutí).
- ... neřídí se "citem (intuicí)", ale jeho postoje jsou matematicky zdůvodnitelné (logická reakce na vstupy a stav situace).
- Racionalitu budeme formálně definovat v Teorii volby/užitku: hráč je schopen posoudit dvě strategie.
- Racionalita versus morálka.

# Racionalita hráče

- Racionální jedinec maximalizuje svůj užitek.
- Racionalita mezi hráči musí být common knowledge (všichni si o sobě navzájem uvědomují, že jsou racionální a současně ví, že ví...).
- Racionalita je jistota, je to logické chování.

Firma/jedinec neřekne, že už má dost a dále neprodukuje. Pokud to přesto řekne, pak je součástí jejího užitku i jiný faktor (např. sledování kvality života šéfa a zaměstnanců). Pak je užitek vícesložkový – (tržba, pohoda).

- ... pak ovšem musíme definovat operátor preference

$Trzby = \mathbb{R}[CZK], Pohoda = \mathbb{R}[???]$
$Rozhodovani \subseteq Trzby \times Pohoda$

Klademe si otázku $\forall r_1, r_2 \in Rozhodovani$ : která možnost je lepší?

Je lepší $(1000, 20)$ než $(5000, 4)$??? A co srovnání s $(20000, -10)$?

Pokud řekneme, že náš užitek je dán funkcí $u : Rozhodovani \to \mathbb{R}$ tak, že $1 pohoda = 100 CZK$, pak maximalizujeme funkci:

$$u(trzba, pohoda) = trzba + 100 \cdot pohoda$$

Možná podoba rozhodovací množiny *Rozhodovani* (zřejmě je ohraničená).

# Preference, volba, užitek a racionalita

- Potřebujeme formalizovat vztah mezi volbou a užitkem. Formalizace=model.
- Ekonomické teorie.
- Zkoumání křivek užitku.
- Mírná skepse nad tím, zda-li to vůbec funguje.

Prozření prognostika: výsledek modelu je vždy *expertní názor* na to, jak to může fungovat. Pokud má někdo *lepší názor*, pak sem s ním...!

# Teorie volby - Úvod

Čerpáno z:

1. McCarthy, N., Mierowitz, A.: Political Game Theory: An Introduction, Cambridge University Press, 2007
2. Magdaléna Hykšová: Přednášky z teorie her, Fakulta dopravní ČVUT

- Studium racionality je základem studia racionálního rozhodování.
- Naším základním předpokladem je fakt, že studujeme racionálního jedince (opakem jsou hry proti přírodě).
- Racionalitu není možno zaměňovat s nějakou formou morálky – racionální jedinec nekoná dobro, pouze sleduje svoje preference.
- Jedinec identifikuje svoje možnosti a hledá mezi nimi optimum (maximum). Musí definovat význam optima.
- Otázka: je schopen odhalit optimum?

# Teorie užitku, příklady

- Pro začátek intuitivně...
- Organizujeme svatbu. Možnosti jsou:
  - v chatě (krásná chata, +1000 bodů); nebo v sále (+800 bodů).
  - Při nepříznivnivé počasí v chatě $\Rightarrow$ 0 bodů, v sále nevadí.
  - Pravděpodobnost nepříznivého počasí je 25 procent.
- Očekáváný užitek z chaty je $0.75 \cdot 1000 + 0.25 \cdot 0 = 750$.
- Očekávaný užitek ze sálu je 800, tzn. větší.
- Volím sál.

Důležité !!! Pokud po tomto zjištění stále váhám, pak to znamená, že bud' (něco je v modelu chybně):

- Jsem chybně modeloval pravděpodobnost nepříznivého počasí.
- Jsem chybně modeloval užitek při realizaci v chatě nebo sále.
- V modelu chybí nějaký další aspekt užitku (např. štěstí nevěsty, rodičů, hostů, ...).
- *Nejsem* racionálně uvažující jedinec.
- Riziko – specifický aspekt. Změna užitkové funkce.

# Definice racionality

- Je-li racionální jedinec konfrontován se dvěma možnostmi $x$ a $y$, pak je schopen jednoznačně rozhodnout, jestli nepreferuje $x$ před $y$ nebo nepreferuje $y$ před $x$ nebo nepreferuje ani jednu možnost.
  - Jestli možnosti splňují tuto podmínku, pak mluvíme o úplnosti.
- V kladném vyjadřování: pokud řeknu, že nepreferuju $x$ před $y$, pak tím míním, že $y$ je stejně dobrá nebo lepší možnost.
- Je-li konfrontován se třemi možnostmi $x, y, z$ a nepreferuje-li $y$ před $x$ a současně nepreferuje $z$ před $y$ (tzn. $x$ je asi lepší než $y$ a $y$ je asi lepší než $z$), pak nemůže preferovat $z$ před $x$.
  - Jestli možnosti splňují toto, pak mluvíme o tranzitivitě.
- Úplnost a tranzitivita nás budou zajímat. Bez nich není racionálního rozhodování.

Mám $A$ množinu alternativ. Pro všechny $a_1, a_2 \in A$ jsem schopen posoudit preferenci a dokonce i na úrovni tranzitivity. Co z toho plyne pro rozhodování?

# Konečné množiny akcí (strategií, možností) a výstupů

- Jedinec volí z množiny akcí $A = \{a_1, a_2, ..., a_k\}$.
- Předpokládáme, že jedinec má kompletní informaci (complete information) o situaci a tudíž je schopen jednoznačně predikovat následky svých akcí (jistota, certainty) – tzn. rozhodování *probíhá za jistoty*.
- Pak definujeme množinu výstupů/důsledků $X = \{x_1, x_2, ..., x_n\}$.
- Z předpokladu jistoty plyne, že pro každou akci $a \in A$ je jednoznačně přiřazen právě jeden výstup $x \in X$ (zobrazení je *jednoznačné*).
- Formálně, existuje funkce

$$u : A \to X$$

- Zkoumání užitkových funkcí pro nás bude významné.

# Konečné množiny akcí (strategií, možností) a výstupů

Formálně, existuje funkce $u : A \to X$.

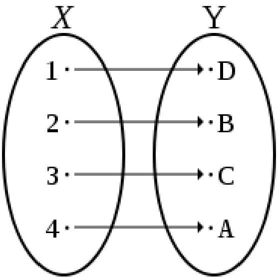
- Dále předpokládáme, že všechny prvky množiny výstupů jsou dosažitelné (feasible) – tzn., každý $x \in X$ je důsledek nějaké akce $a \in A$ ($u$ je zobrazení *na množinu*).
- Formálně: $x_i$ je dosažitelný, pokud $\exists a \in A : u(a) = x_i$.
- Z předpokladu dosažitelnosti a jednoznačnosti plyne, že je jedno, zkoumáme-li jedincovy preference nad akcemi nebo nad výstupy (zobrazení je vzájemně jednoznačné, bijektivní).

Co by znamenalo, kdyby pro dvě různé akce $a_1, a_2 \in A$ platilo $u(a_1) = u(a_2)$?

# Demo

| $A$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| $u$ | $x_1$ | $x_2$ | $x_3$ |

Říkáme tedy, že pokud je zobrazení $u$ vzájemně jednoznačné, pak je jedno, jestli zkoumáme preference na akcích nebo na užitcích.



# Rozhodování

Rozhodování je akt výběru akce $a \in A$.

- Počet zkoumaných charakteristik: mono-kriteriální rozhodování, více-kriteriální rozhodování
- Počet rozhodovatelů: jeden versus více $\Rightarrow$ rozhodování, jehož výsledek je ovlivněn alespoň dvěma racionálními účastníky (HRA)

# Preference formálně

Předpokládejme, že existuje relace $R \subseteq X \times X$ (nebo $R \subseteq A \times A$) obsahující "názory" jedince na posouzení jeho možností. Zkoumáme, zda-li je jedinec schopen na základě $R$ provést rozhodnutí (racionálně).

- Slabá preference (neostrá, weak preference) je definována binární relací $R$, kde $x_i R x_j$ značí, že $x_j$ není preferováno před $x_i$. Říkáme tím:
  - $x_i$ je lepší nebo stejná možnost
  - $x_j$ není preferováno nad $x_i$, v nejlepším případě může být stejná, spíše horší možnost
  - připouštíme, že $x_i$ a $x_j$ mohou být stejně dobré.
- Analogie relace $R$ je binární relace $\geq$ například nad čísly (pak ovšem píšeme dle zvyklostí $x_i R x_j \Leftrightarrow x_i \geq x_j$).

# Striktní preference a indiference

Máme množinu alternativ $A$ (nebo užitků $X$).

Pomocí $R$ můžeme definovat dvě významné relace **striktní preference** (strict preference) a **indiference** (indifference, nerozlišitelnost).

### Definition

Pro každé $x, y \in A$, $xPy$ ($x$ je striktně preferováno před $y$) právě tehdy, pokud $xRy \wedge \neg yRx$. Podobně, $xIy$ ($x$ je indiferentní s $y$) právě tehdy, pokud $xRy \wedge yRx$.

Poznámka: relace $P$ je zaměnitelná s $>$. Podobně je $I$ a $=$.

Říkáme: hráč strikně preferuje $x$ nad (před) $y$, hráč je indiferentní mezi $x$ a $y$.

# Volba jedince

Notace: hvězdičkou bývá v TH označována konečná volba hráče/hráčů.

**Jedinec bude zřejmě volit takovou volbu $x^* \in A$, pro kterou $x^* R y$ pro všechna $y \in A$.**

Zkoumáme, zda-li je jedinec **schopen volby na základě relace** $R$. Respektive, definujeme takové vlastnosti $R$, aby taková volba byla možná.

# Volba jedince, maximální množina

**Definition**

Pro relaci slabé preference $R$ a množinu voleb $A$ definujeme maximální množinu $M(R, A) \subseteq A$ tak, že

$$M(R, A) = \{x \in A | xRy; \forall y \in A\}$$

Jedinec bude zřejmě volit z maximální množiny. Ta by ovšem měla obsahovat alespoň jeden prvek.

Otázka: existuje taková množina $M(R, A)$? Co to může ovlivnit?

# Volba jedince, maximální množina

Předpoklad: pokud je $|M(R, A)| > 1$, pak jedinec volí jednu z $M(R, A)$ náhodně, protože je indiferentní mezi shodnými alternativami, tzn. $\forall x_1, x_2 \in M(R, A) : x_1 I x_2$

Schopnost přijmout indiferenci nad strategiemi $M$.

Pokud jedinec nesouhlasí s výše uvedeným postojem, pak je buď chybná jeho preference nebo jedinec není racionální.

Zjištění maximální množiny zřejmě nebude možné, pokud pro nějaká dvě $x, y \in A$ neplatí ani $xRy$ ani $yRx$ (narušení úplnosti).

# Nezávislost na irelevantních alternativách

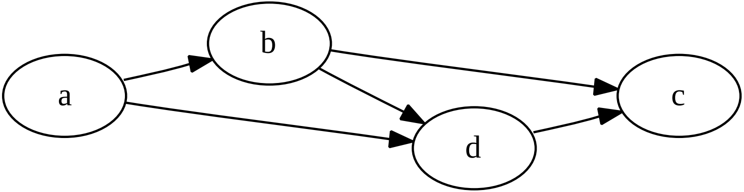
Později budeme navíc vyžadovat, aby jedincova preference byla nezávislá na irelevantních alternativách (např. v Teorii veřejné volby).

### Definition

Mějme množinu alternativ $A = \{a, b\}$ a preferenci $R$ na $A$. Předpokládejme, že $aRb$. Pokud chceme zachovat nezávislost na irelevantních alternativách, tak přidání alternativy $c$ do $A$ nesmí změnit preferenci mezi $a$ a $b$ (tzn. $aRb$).

# Demo

Mějme množinu užitků $A = \{a, b, c, d\}$. Relaci slabé preference $R$ zobrazíme graficky:




Je relace úplná?

$M(R, A) = \{a\}$, protože pouze pro $a$ platí $aRy; \forall y \in A$.

Pozn.: předpokládáme, že $xRx; \forall x \in A$ platí automaticky.

# Pomocné definice

## Definition

Binární relace $R$ na $X$ (tzn. $R \subseteq X^2$) je:

1. úplná, pokud $\forall x, y \in X, x \neq y : xRy \vee yRx$.
2. reflexivní, pokud $\forall x \in X : xRx$.
3. tranzitivní, pokud $xRy \wedge yRz \Rightarrow xRz \;\; \forall x, y, z \in X$.
4. kvazi-tranzitivní (definováno nad relací $P$), pokud $xPy \wedge yPz \Rightarrow xPz \;\; \forall x, y, z \in X$ (pomocí $R$ je ekvivalentní se zápisem $xRy \wedge \neg(yRx) \wedge yRz \wedge \neg(zRy) \Rightarrow xRz \wedge \neg(zRx)$).
5. anti-symetrická, pokud $xRy$ a $yRx$ značí $x = y$ (jsou totožné)
6. asymetrická, pokud $xRy$ pak neplatí $yRx$

Poznámka: úplnost vyjadřuje schopnost jedince *mít názor* na preferenci nad každými dvěma akcemi (se zachováním tranzitivity). *Musíme s tímto počítat i přesto, že mnoho ekonomů a psychologů o této schopnosti člověka v některých situacích pochybuje.*

# Demo; paradox

Příklad: mějme množinu $X = \{b_1, ..., b_{1000}\}$ tisíce různých lahví piva definovanou takto: $b_1$ je lahev, kde jedna kapka piva byla nahrazena kapkou vody, $b_2$ dvě kapky až $b_{1000}$ tisíc kapek.

Jistě by každý prohlásil, že $b_1 I b_2$, $b_2 I b_3$, ..., $b_{999} I b_{1000}$ (je mi jedno, jestli zvolím lahev ředěnou 999 kapkami vody nebo 1000 kapkami).

Protože $xIy$ implikuje $xRy$ (z definice), pak musí platit i $b_{1000} R b_{999}, ..., b_2 R b_1$.

Pokud akceptujeme tranzitivitu, pak $b_{1000} R b_1$. Vedle toho ovšem $b_1 P b_{1000}$. Kde se stala chyba?

# Úplné neostré uspořádání

### Definition

Mějme (neprázdnou) množinu voleb $A$. Úplné neostré uspořádání na množině $A$ je binární relace $R \subseteq A \times A$, která je úplná, reflexivní a tranzitivní.

Poznámka: Anti-symetrii zde nezavádíme (by nám narušovala indiferenci). Proč reflexivita?

### Theorem

*Je-li $A$ konečná množina a $R$ úplné neostré uspořádání, pak $M(R, A) \neq \emptyset$.*

Máme garanci, že pokud je preference definována *správně (pořádně)* a množina alternativ je konečná, pak existuje zvolitelné optimum.

# Důkaz věty o existenci maxima na množině alternativ 1/3

Nechť $A$ je konečná (a neprázdná) množina a $R$ je úplná, reflexivní a tranzitivní relace. Důkaz bude proveden matematickou indukcí. Připomeňme:

$$M(R, A) = \{x \in A | xRy; \forall y \in A\}$$

**Krok 1**: Je-li $A$ jednoprvková množina, tedy $A = \{a\}$, pak z reflexivity plyne $aRa$ a proto $M(R, A) = \{a\}$.

**Krok 2**: Ukážeme, že je-li tvrzení pravdivé pro $A'$ s $n$ prvky a relaci $R'$ na $A'$, pak musí být pravdivé i pro libovolnou $A$ s $n + 1$ prvky a uspořádání $R$.

tzn. přidáním alternativy se podmínky preference nenaruší

# Důkaz věty o existenci maxima na množině alternativ 2/3

**Důkaz kroku 2**: Budeme pracovat s množinami $A$ a $A'$, kde

$$A = A' \cup \{a\}$$

Na množinách $A$ a $A'$ jsou definována uspořádání $R$ a $R'$ tak, že $R'$ je $R$ omezená na $A'$, tedy

$$R' = R \cap (A' \times A')$$

Dle našich předpokladů (a dle postupu matematické indukce) je

$$M(R', A') \neq \emptyset$$

Pak z předpokladů úplnosti preferenční relace plyne, že pro libovolné $y \in M(R', A')$ platí buď $yRa$ nebo $aRy$ nebo platí oboje. Budeme proto zkoumat dvě varianty preferencí $y$ a nově přidané alternativy $a$.

# Důkaz věty o existenci maxima na množině alternativ 3/3

$$A = A' \cup \{a\}$$

$\forall y \in M(R', A')$ platí buď $yRa$ nebo $aRy$ nebo platí oboje.

1. Platí $yRa$ neboli prvky maxima na $A'$ jsou preferovány nad novým prvkem $a$. Pak tedy $yRz$ pro všecny $z \in A' \cup \{a\}$ (z definice maximální množiny) a proto $y \in M(R, A)$. A tím dokazujeme krok 2.
2. Platí $aRy$. Pokud tedy $y \in M(R', A')$, pak $yRz$ pro libovolné $z \in A'$. Vycházíme z $aRy$ a víme, že $yRz$ pro všechny $z \in A'$. Z tranzitivity $R$ plyne $aRz$ pro všechny $z \in A'$. Obecně to implikuje $aRw$ pro všechny $w \in A$ a proto $a \in M(R, A)$ a to opět dokazuje krok 2.

Principem matematické indukce jsme dokázali tuto větu pomocí kroků 1 a 2. **Q.E.D.**

# Teorie užitku

Zkoumání volby provádíme na užitkové funkci s číselným výstupem $u : A \rightarrow \mathbb{R}$.

- Můžeme plně využít zvyklostí s operátorem $\geq$ (naše preference).
- $u(x) \geq u(y) \Leftrightarrow xRy$
- $u(x) > u(y) \Leftrightarrow xPy$
- $u(x) = u(y) \Leftrightarrow xIy$

Stále můžeme rozlišovat kardinalistický (chápeme velikost rozdílu mezi $u(x) - u(y), xPy$) a ordinalistický přístup k užitku (pouze chápeme preferenci).
V mnoha algoritmech TH ani nepotřebujeme znát *míru preference* $x$ nad $y$ (uvidíme např. v definici best-response, dominance).

# Definice funkce užitku

## Definition

Mějme $A$ a $R \subseteq A^2$. Řekneme, že užitková funkce $u : A \rightarrow \mathbb{R}^1$ reprezentuje $R$, jestliže pro všechna $x, y \in A, u(x) \geq u(y) \Leftrightarrow xRy$.

Snadno jsme schopni definovat $>$ a $=$.

Rozhodovatel-maximalista pak volí z množiny

$$M(R, A) = arg \ \max_{x \in A}[u(x)]$$

Poznámka: ve hrách tomu začneme říkat Best-response.

# Rozhodování bez jistoty

- Rozhodování za rizika: výsledková funkce přiřazuje každému rozhodnutí pravděpodobnostní rozložení na množině výsledků.
- Rozhodování za neurčitosti: $u : A \to 2^X$ (Pravděpodobnostní rozložení ovšem není známé nebo je nezjistitelné).
- Postoj k riziku: risk-neutral, risk-averse, risk-seeking.

Očekávaný užitek.

# Rozhodování za nejistoty: Příklad s investorem

Investor přemýšlí, kam umístit svůj kapitál. Tři alternativy:

1. Zakoupit cenné papíry, které s absolutní jistotou nesou 5.5% zisk.
2. Zakoupit akcie, které ponesou:
   - 3% s pravděpodobností $\frac{1}{3}$.
   - 6% s pravděpodobností $\frac{1}{3}$.
   - 9% s pravděpodobností $\frac{1}{3}$.
3. Zakoupit akcie, které ponesou:
   - 4% s pravděpodobností $\frac{1}{2}$.
   - 8% s pravděpodobností $\frac{1}{2}$.

Z hlediska rozhodování musí počítat s určitou nejistotou, kterou zde pojmenováváme *loterie*. Jeho tři alternativy jsou tři loterie.

Loterie je volba jako ji známe z teorie volby. Definujeme relaci preference a další pojmy. Více v "Axiomatické teorii užitku".

# Očekávaný užitek (Expected Utility/Payoff): St. Petersbourg paradox

St. Petersbourg paradox (D. Bernoulli, 1738): mějme hru, kde účastník zaplatí vstupní poplatek $c$ a pak háže mincí tak dlouho, dokud nepadne hlava. Protistrana souhlasí, že mu zaplatí 1 dukát, pokud padne v prvním hodu, 2 dukáty v druhém hodu, 4 v třetím hodu, atd.

Očekávaný zisk hráče tedy musí být:

$$E = (-c) + \frac{1}{2}1 + \frac{1}{4}2 + \frac{1}{8}4 + \frac{1}{16}8 + \ldots + = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \ldots = \sum_{k=1}^{\infty} \frac{1}{2} = \infty$$

Problém je, že s tímto většina reálných lidí nesouhlasí a za svou účast by nezaplatili více než 20 dukátů, resp. radostně by prodali svou účast ve hře za 20 dukátů.

# Užitek

St. Petersbourg paradox vyvolal diskuzi, jaký je vlastně užitek z přijetí nějakého (např. finančního) vstupu.
Gabriel Cramer (v komunikaci s D. Bernoullim): lidé hodnotí částky podle užitku, který jim přinesou.
Předpoklad: jakákoliv částka přesahující $2^{24}$ dukátů člověku připadá stejná jako $2^{24}$ dukátů.
Pak je očekávaný užitek hry:

$$\frac{1}{2}1 + \frac{1}{4}2 + \frac{1}{8}4 + ... + \frac{1}{2^{24}}2^{23} + \frac{1}{2^{25}}2^{24} + \frac{1}{2^{26}}2^{24} + ... + =$$

$$= \frac{1}{2} + \frac{1}{2} + ... + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + ... = 12 + 1 = 13$$

Potom je realisticky očekávaný užitek ze hry 13 dukátů.

# Log utility

Počet jednotek užitku $u(x)$ z vlastnictví částky $x$. Při navýšení majetku z $x$ na $x + dx$ je přírustek užitku $du(x)$ přímo úměrný přírustku $dx$ a nepřímo úměrný dosavadnímu majetku $x$ ($\alpha$ je počáteční majetek, $b \in \mathbb{R}^+$ je nějaký koeficient vnímání přírustku).

$$du(x) = \frac{b\ dx}{x}$$

$$u(x) = b \ln x + c; c \in \mathbb{R}$$

$$u(x) = b \ln x - b \ln \alpha = b \ln \frac{x}{\alpha}$$

obecně: užitek je ln(po akci) − ln(před akcí).

# Log utility ve St. Petersbourg paradoxu

$$E = \sum_{n=1}^{\infty} \frac{1}{2^n} b \ln \frac{\alpha + 2^{n-1}}{\alpha} =$$

$$= b \ln [(\alpha + 1)^{\frac{1}{2}} (\alpha + 2)^{\frac{1}{4}} (\alpha + 4)^{\frac{1}{8}} ...] - b \ln \alpha$$

Částka $D$, jejíž přidání k počátečnímu majetku přinese stejný užitek:

$$b \ln \frac{\alpha + D}{\alpha} = b \ln \left[ (\alpha + 1)^{\frac{1}{2}} (\alpha + 2)^{\frac{1}{4}} (\alpha + 4)^{\frac{1}{8}} ... \right] - b \ln \alpha$$

z toho plyne, že takové $D$ je:

$$D = \left[ (\alpha + 1)^{\frac{1}{2}} (\alpha + 2)^{\frac{1}{4}} (\alpha + 4)^{\frac{1}{8}} ... \right] - \alpha$$

Pro nulové $\alpha = 0$ počáteční jmění je $D = \sqrt[2]{1} \cdot \sqrt[4]{2} \cdot \sqrt[8]{4} \cdots = 2$. Mít nulový počáteční majetek, tak nezaplatím za učast ve hře víc než dva dukáty.

# Příště

- Definice strategických her, resp. her ve strategickém tvaru.
- Interakce dvou a více jedinců při rozhodování.
- Zkoumání jejich užitku budeme provádět kardinalisticky na základě užitkových funkcí.
- Preference se pak budou budovat nad strategiemi v kontextu možných tahů protihráčů (budeme mluvit o dominantnosti strategií).
- Zavedeme koncept ekvilibria ve hře.

# Funguje Teorie her?