THE: Mechanism design

Martin Hrubý

Brno University of Technology
Brno
Czech Republic

November 13, 2019

Úvod

Čerpáno z:

- McCarthy, N., Mierowitz, A.: Political Game Theory: An Introduction, Cambridge University Press, 2007
- Nisan, N. et al.: Algorithmic Game Theory, Cambridge University Press, 2007
- Magdaléna Hykšová: Přednášky z teorie her, Fakulta dopravní ČVUT

Social/public choice. Spravedlnost, férovost, bez závisti (envy-freeness), optimálnost.

Úvod

Zatím jsme studovali "hotové situace" a hledali racionalitu v reakci hráčů na situaci. Tzn., umíme zkoumat reakci hráčů na situaci.

- Mechanism design je opačnou úlohou.
- Také nazýváno "Reverse game theory" – Teorie her naopak.
- Zabývá se studiem situací s privátní (tajnou) informací – cílem je formou návrhu hry dosáhnout pravidel, kdy je pro hráče racionálním chováním "odtajnit" svoje privátní informace.
- Stále se předpokládá individuální racionalita hráčů.
- Společnost potřebuje občas provést kolektivní rozhodnutí (jako výsledek invididuální racionality zúčastněných).

Mechanismy s penězi. Mechanismy bez peněz.

Mechanism design (MD)

MD nachází uplatnění v těchto třídách problémů:

- Volby (teorie veřejné volby) – každý volič má své preference a zkoumáme volební mechanismy vedoucí k veřejně optimální volbě.
- Trhy – zkoumáme mechanismy ekonomiky. Dokonalý trh zřejmě neexistuje. Cílem je správná realokace zboží a peněz.
- Aukce – speciální případ trhu, kde je pouze jeden prodávající. Hledáme pravidla pro určení vítěze a způsobu vedení aukce. Kombinatorické aukce.
- Dělení (dělitelné/nedělitelné zboží, kvóty) – spravedlivé rozdělení, envy-free, efektivní.
- Veřejné záležitosti, pravidla a zákony – regulace, daně, alokace (kvót, statků, ...), ...

Klíčovým ukazatelem mechanismu je odolnost vůči strategické manipulaci (strategy-proof, incentive compatible, truthfull).

Teorie veřejné volby (Social, Public Choice)

Specifikace problému:

- Skupina voličů $I = \{i_1, ..., i_N\}$ chce zvolit jednoho z množiny A kandidátů.
- Každý volič má na množině A své preference (úplné uspořádání – ostré/neostré).
- Volíme člověka ve volbách do funkce, volíme alternativu veřejného rozhodnutí nebo jenom třeba film pro večerní promítání.
- Vždy máme alternativy, voliče a potřebu se společně rozhodnout.
- Očekáváme, že každý volič chce dosáhnout svého cíle – racionálně se rozhoduje a vnímá užitek z výsledku rozhodnutí.

Volební pravidla (způsob určení vítěze) = mechanismus.

Strategická manipulace

A – množina alternativ. Nechť S je množina permutací na A, tj. soubor všech možných preferenčních uspořádání na A.

$S_i \in S$ nechť je (pravdivá) preference hráče i (preference nad prvky A). $\succ_i$.

$A^* = F(S_i, S_{-i})$ nechť je vítěz voleb (výsledek volebního mechanismu).

Pokud pro i existuje $S_i' \neq S_i$ taková, že $A^{*'} = F(S_i', S_{-i})$ a $A^{*'} \succ_i A^*$, pak má i důvod deklarovat svou preferenci nepravdivě. Manipuluje.

Co když mohou manipulovat všichni?

Odolnost vůči strategické manipulaci

Esenciální otázka: proč je to důležité? Tj., proč je důležité zjistit pravdu?

- Volby (teorie veřejné volby) – skutečně pravá společensky optimální volba.
- Trhy a Aukce – zjištění optimální ceny.
- Dělení (dělitelné/nedělitelné zboží, kvóty) – spravedlnost jako společenská norma, stabilita systému.
- Veřejné záležitosti, pravidla a zákony – zajištění požadovaného chování hráčů.

Pokud je možnost manipulovat mechanismus, hráči to udělají.
Pokud manipulují všichni, systém nefunguje.
Mnohdy je větší diskuse kolem volby mechanismu (meta-volby) než nad samotným předmětem voleb. Kvórum. Formulace otázky referenda.

Většinová volba

Základním mechanismem voleb je **většinová volba**.

- Sečteme hlasy pro jednotlivé kandidáty, vítězí kandidát s nejvyšším počtem hlasů.
- Máme-li dva kandidáty $A = \{a_1, a_2\}$, je volba jednoduchá: sečteme hlasy pro a_1 a hlasy pro a_2.
- První problém: co uděláme, pokud je počet hlasů shodný? Pokud volby zopakujeme, pak celá činnost nedává smysl, protože nemůžeme očekávat, že voliči v druhém kole *změní svoje preference*.
- Pokud je kandidátů více, může to dapadnout ještě horším způsobem.

Více kandidátů a mechanismus většinové volby (duely) – pak silně záleží na pořadí rozhodování o jednotlivých dvojicích.

Preference

- Máme množinu alternativ A.
- Preference na množině A. Opět budeme definovat na relaci slabé preference $R \subseteq A \times A$.
- Obvyklým způsobem chápeme relaci striktní preference a indiference.
- Slabou preferenci budeme zapisovat:
  - $\succeq_i$ vztaženou ke hráči i
  - $\succeq$ vztaženou k celkové společnosti
- Striktní preferenci obdobně: $\succ_i$, $\succ$
- Připomeňme, že očekáváme jistou logiku v preferencích – *tranzitivitu* preferenční relace.

Chceme zjistit, co se stane při preferencích jednotlivců $\succ_i$, jaká bude preference (a potom volba) společnosti $\succ$.

Smysl preference ve volbách

- Pokud je $|A| = 1$, volby se nekonají.
- Pokud je $|A| = 2$, intuitivně chápeme, že hráč preferuje jednu z alternativ (přeje si jednu a druhou nechce).
- Pokud je $|A| \geq 3$, volič stále preferuje jednu alternativu (jeho maximum), ale musí mít názor i na preference nad zbytkem alternativ.
- Pak píšeme $a_1 \succ_i a_2 \succ_i ... \succ_i a_x$.

V teorii veřejné volby nepracujeme s maximy hráčů, ale s jejich preferencemi.

Většinová volba: manipulovatelnost při $|A| = 2$

Většinová volba se dvěma alternativami je z principu **nemanipulovatelná**.

- Hráč i preferuje a_1 nad a_2. Tj., preference je $a_1 \succ_i a_2$.
- Užitek $U_i(a_1) = 1$, $U_i(a_2) = 0$.
- Hráč **nezvýší** svůj užitek (ani slabě), pokud bude prezentovat jinou preferenci – $a_2 \succ_i a_1$.
- Při $|A| = 3$ to může být jinak.

Modelová situace voleb – I.

Preference/Voliči:	i_1	i_2	i_3
1.	A	A	B
2.	B	C	C
3.	C	B	A

Tabulka modelu preference voličů $I = \{i_1, i_2, i_3\}$.
Volič i_1: $A \succ_i B \succ_i C$.

Při většinové volbě zvítězí alternativa A. Celková preference je $A \succ B \succ C$.

Modelová situace voleb – II.

Preference/Voliči:	i_1	i_2	i_3
1.	A	C	B
2.	B	A	C
3.	C	B	A

Tabulka modelu preference voličů $I = \{i_1, i_2, i_3\}$.
Volič i_1: $A \succ_i B \succ_i C$.

První názor

Preference/Voliči:	i_1	i_2	i_3
1.	A	C	B
2.	B	A	C
3.	C	B	A

Vítěz by měl být schopen zvítězit v duelu s každým protikandidátem.

- Duel A:B (vítěz A, dva hlasy), tzn. celek $A \succ B$
- Duel A:C (vítěz C, dva hlasy), tzn. celek $C \succ A$
- Duel B:C (vítěz B, dva hlasy), tzn. celek $B \succ C$

1. kolo – vítězí A, 2. kolo (A versus C) – vítězí C. Je tedy C vítěz voleb?

Vítěze nevidíme, ke všemu je zmatek v celkové preferenci voličů: $A \succ B \succ C \succ A$.

Condorcetův paradox

Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet (markýz de Condorcet, 1743–1794).

- Francouzský filosof a matematik.
- Poukázal na koncepční nedostatek většinové volby (zmíněný paradox): předpokládáme-li tranzitivitu u preferencí voličů, lze dosáhnout netranzitivní celkové preference.
- Condorcetův vítěz je alternativa X s tou vlastností, že pro každou alternativu Y je počet voličů preferujících X před Y větší, než počet voličů preferujících Y před X.
- Příklad: Většina prefuje C před B, C před A, B před A. Závěr: A je poraženo B i C, většina preferuje C nad B, vítězí C.
- Condorcetův vítěz nemusí existovat.
- Další studie: Condorcet's jury theorem.

Strategická manipulace

Preference/Voliči:	i_1	i_2	i_3
1.	A	C	B
2.	B	A	C
3.	C	B	A

- Jsem třetí volič.
- Vidím, že můj kandidát B není oblíbený, ale nechci připustit zvolení A.
- Budu proto prezentovat svoji preferenci zmanipulovanou: $C \succ_3 B \succ_3 A$ a ROZHODNU o vítězi voleb.

Situace, kdy podání nepravdivé informace o preferencích mohou ovlivnit výsledek volby jsou nežádoucí. Hledáme mechanismy, které jsou zabezpečeny proti strategické manipulaci.

Borda count

- Jean Charles de Borda (1733 – 1799).
- Kritizoval volební mechanismy ve Francouzské akademii.
- Navrhl rozdělit "body" kandidátům: každý volič přidělí $|A|$ svému prvnímu kandidátovi, $|A| - 1$ druhému, ..., jeden poslednímu. Vítěz je kandidát s nejvyšší sumou bodů.
- Zdá se to jako ideální metoda.
- Mnohdy vede k řešení, je ovšem strategicky manipulovatelná.

Demo: Borda count

Preference/Voliči:	5 voličů	4 voliči	3 voliči
1.	A	B	C
2.	C	C	B
3.	B	A	A

Většinová volba je $A \succ B \succ C$ (5:4:3). Jak dopadnou duely?

- A má $3 \cdot 5 + 1 \cdot 4 + 1 \cdot 3 = 22$ bodů.
- B má $1 \cdot 5 + 3 \cdot 4 + 2 \cdot 3 = 23$ bodů.
- C má $2 \cdot 5 + 2 \cdot 4 + 3 \cdot 3 = 27$ bodů. Je to Bordův vítěz.
- C vítězí i Condorcetovsky.

Borda count, jiný příklad

	1	2	3	4	5	6	7
1.	X	A	B	X	A	B	X
2.	C	X	A	C	X	A	C
3.	B	C	X	B	C	X	B
4.	A	B	C	A	B	C	A

X=22, A=17, B=16, C=15

X odstoupí. V tom okamžiku je výsledek: A=13, B=14, C=15

Borda count není odolný proti **irelevantním alternativám**.

Duncan Black (1908–1991)

Vítězem je:

- Condorcetův vítěz, pokud existuje.
- Bordův vítěz v ostatních případech.

Zkuste sami navrhnout volební mechanismus pro 3 a více kandidáty (např. různé kolové (sekvenční) formy voleb dvou protikandidátů).

Kenneth Arrow ukázal, že spravedlivé volby nejsou možné.

Pozor: Obecně je nelze garantovat. Neznamená to, že žijeme v nespravedlivém světě.

Profil

- Nechť A je množina alternativ. Budeme formalizovat obor hodnot všech možných preferencí.
- Množina $\mathcal{R}$ je principem totožná s S_i ve strategických hrách. Podobně $\mathcal{R}^N$ je S.

Definition

Mějme množinu alternativ A. Množina $\mathcal{R}$ obsahuje všechna možná uspořádání na množině A (je izomorfní s permutacemi na A).

Definition

Mějme množinu alternativ A a n voličů. Profilem ρ uspořádání preferencí budeme rozumět uspořádanou n-tici

$$\rho = (R^1, R^2, ..., R^n) \in \mathcal{R}^n$$

Social welfare function

Také pod názvem: Social Agregation Function, Preference Agregation Rule, ...

Definition

Funkcí společenského blahobytu (social welfare function) budeme rozumět funkci

$$F : \mathcal{R}^n \to \mathcal{R}$$

Preference jednotlivce budeme zapisovat $x \succ_i y$ (vždy vztaženo k nějakému profilu ρ). Preference společnosti $x \succ y$.

Funkce F je mechanickým arbitrem (řídí se pravidly), která na zadané preference voličů sestaví preferenci celé společnosti (základ pro volbu optima).

Základní předpoklady

Je-li $|A| = 2$, pak je jediným rozumným mechanismem většinová volba.

- Počet alternativ je $|A| \geq 3$.
- Funkce společenského blahobytu je definována pro všechny profily individuálních preferencí (tzn., arbitr je schopen sestavit společenskou preferenci v libovolné situaci preferencí voličů).
- Společnost obsahuje alespoň dva voliče.

Podmínka tranzitivity

Definition

Funkce společenského blahobytu F je tranzitivní, je-li $F(\rho)$ tranzitivní pro všechny $\rho \in \mathcal{R}^N$.

Tranzitivitu považujeme za minimální projev logiky rozhodování (např. pro nalezení maxima).

Arbitr je schopen sestavit tranzitivní společenskou preferenci ve všech možných situacích preferencí jednotlivců.

Podmínka nediktátorství

Definition

Funkce společenského blahobytu F je nediktátorská, pokud neexistuje $i \in Q$ takový, že pro všechny $\rho \in \mathcal{R}^N$ a každé dvě alternativy $x, y \in A$ platí, že pokud $x \succ_i y$, pak $x \succ y$.

Tato podmínka ukládá, aby žádný hráč i nezpůsobil svou preferencí preferenci celku, bez ohledu na to, co si přeje celek. Zatím nemluvíme o strategické manipulaci (předkládá nepravdivé preference).

Uvidíme, že s touto podmínkou je zásadní problém. Navíc, diktátorem není rozuměn "zlý člověk" – prostě jsou situace, kdy rozhodnutí závisí na jednotlivci.

Podmínka Pareto optimality

Definition

Funkce společenského blahobytu F je slabě Pareto optimální, pokud $x \succ_i y$ pro všechny $i \in Q$ způsobí, že $x \succ y$ pro libovolné $x, y \in A$.

Tato podmínka je velmi přímočará: pokud si každý zvlášť preferuje x před y, pak i celek preferuje x před y.

Pozn.: Existuje ještě striktně Paretovská optimalita v teorii veřejné volby, podmínka je ovšem značně silnější (možno dostudovat). Tzn., slabě/silně Paretovská optimalita v tomto případě nesouvisí se slabostí/striktností preference.

Podmínka Nezávislosti na irelevantních alternativách (IIA)

Independence of irrelevant alternatives (IIA) je přítomna ve všech částech THE posuzujících preference. Formálně je definovaná:

Definition

Funkce společenského blahobytu F je nezávislá na irelevantních alternativách, pokud platí pro libovolné $x, y \in A$ a libovolné dva profily $\rho, \rho' \in \mathcal{R}^N$, že $x \succeq_i y \Leftrightarrow x \succeq_i' y$ pro všechny $i \in Q$ implikuje $x \succeq y \Leftrightarrow x \succeq' y$.

Máme dva pohledy na věc: $\rho, \rho' \in \mathcal{R}^N$, pokud v obou všichni hráči preferují x před y, pak i celek musí v obou případech preferovat x před y.

Jinak řečeno, pokud v obou profilech mají hráči stejný názor na preferenci na $\{x, y\}$, měly by být výsledné společenské preference konzistentní v otázce preference mezi $\{x, y\}$.

Demonstrace IIA

Uvažujme Borda count, která způsobí: A – 11 bodů, D – 8 bodů, tedy $A \succ D$.

	V	W	X	Y	Z
1.	A	A	B	B	C
2.	B	C	C	C	B
3.	C	B	D	D	D
4.	D	D	A	A	A

Omezíme-li se pouze na alternativy $\{A, D\}$, pak: A – 7 bodů, D – 8 bodů, tedy $D \succ A$.

	V	W	X	Y	Z
1.	A	A	D	D	D
2.	D	D	A	A	A

Arrowův teorém

- Nazýván též: Arrow's paradox, Arrow's impossibility theorem.
- Kenneth Arrow (Nobelova cena, 1972).
- Nejvýznamnější teorém ve společenských vědách.
- Arrow ukázal, že funkce společenského blahobytu, která je tranzitivní, Pareto optimální a IIA, je diktátorská.
- Taky lze říct, že předchozí 4 podmínky (tranzitivita, Pareto optimalita, IIA, nediktátorství) jsou vzájemně nekonzistentní – nemohou platit současně.
- Následky: vždy jedna vlastnost vypadne

Arrow's Theorem

Theorem

Je-li A konečná množina alternativ s alespoň třemi alternativami, pak neexistuje společenská agregační funkce $F : \mathcal{R}^N \to \mathcal{R}$*, která je tranzitivní, nediktátorská, slabě Pareto optimální a nezávislá na irelevantních alternativách.*

Existuje několik ekvivalentních formulací Arrowova theoremu a stejně tak několik jeho důkazů. Předvedeme si důkaz založený na zkoumání množiny *rozhodující* uspořádanou dvojici (x, y).

Mnoho partii MD je založeno na hledání **minimální vítězné koalice**. Skupinová racionalita.

Pomocná definice

Definition

Pro funkci společenského blahobytu F, je množina/podmnožina $W \subset Q$:

- semi-rozhodující dvojici (x, y), platí-li $x \succ y$ pro všechny $\rho \in \mathcal{R}^N$, kde $x \succ_i y; \forall i \in W$ a $y \succ_j x; \forall j \in Q \setminus W$. Rozdělení společnosti na dva tábory.
- rozhodující dvojici (x, y), platí-li $x \succ y$ pro všechny $\rho \in \mathcal{R}^N$, kde $x \succ_i y; \forall i \in W$.
- rozhodující, pokud je rozhodující dvojici (x, y) pro všechny $x, y \in A$.

Základ leží v semi-rozhodující množině $W \subset Q$ pro x proti y:

$$(\forall i \in W : x \succ_i y \wedge \forall j \in Q \setminus W : y \succ_j x) \wedge x \succ y$$

Pomocné lemma (field expansion lemma)

Lemma

Je-li F tranzitivní funkce společenského blahobytu, která je slabě Paretovská a IIA, pak $W \subset Q$ je rozhodující, pokud je W semi-rozhodující dvojici (x, y) pro některá $x, y \in A$.

Důkaz jako domací příprava (projekt). Požadavky lemma na F jsou (téměř) nesplnitelné.

Plyne z toho, že pokud je nějaká skupina voličů semi-rozhodující pro nějaký pár alternativ (x,y), pak je skupina rozhodující. Pro funkce F, které jsou slabě Paretovské a IIA, skupina, která semi-rozhoduje nějaké (x,y), rozhoduje nad všemi alternativami.

Dále z toho plyne, že skupina rozhoduje vše nebo nerozhoduje vůbec nic. Z definice, W musí být maximální. Diktátorství skupiny není problém (to je demokracie).

Semi-rozhodující skupina

Předpokládejme dva tábory lidí A a B.

- A: $x >_A y$ a $x >_A z$
- B: $y >_B x$ a $x \geq_B z$
- Skupina A nechť je při F semi-rozhodující x nad y.
- Pak A je rozhodující x nad z. Implikuje, je semi-rozhodujici x nad z.
- V důsledku A semi-rozhoduje libovolné páry.
- A je pak rozhodující.

Důkaz Arrowse

Ukážeme, že buď je jednočlenná skupina rozhodující (tzn. porušíme nediktátorství) nebo je celá společnost nerozhodující (a tím porušíme podmínku slabé Paretovskosti). Tj., pokud se celá společnou nedohodne na lib. (x, y), pak se **nedohodne na ničem**. Pokud si dokáže společnost odhlasovat nějaké (x, y), pak je rozhodující, ale pravděpodobně obsahuje diktátora.

Celkově vzato ukážeme, že stanovené podmínky jsou vzájemně neslučitelné.

Velmi důležité: zkoumáme celý prostor $\mathcal{R}^N$ a všechny možné F. Tzn., důkaz založíme na situaci, kterou lze považovat za hypotetickou.

Důkaz Arrowse

Předpokládejme konečnou množinu alternativ A, kde $|A| \geq 3$. Pro potřeby *sporu* předpokládejme, že F je tranzitivní, nediktátorská, slabě Paretovská a IIA (najdeme libovolný případ, kdy to neplatí a tím prokážeme spor s předpokladem).

Z předchozího lemma plyne, že každá skupina $W \subset Q$ je buď rozhodující nebo nerozhoduje žádné (x, y). Tvrzení: pouze jedna (maximální) skupina může být rozhodující.

Předpokládejme dvě disjunktní množiny (prázdný průnik) $J, K \subset Q$, které nejsou semi-rozhodující pro libovolné x,y (a tím nerozhodují vůbec). Nechť $L = Q \setminus (J \cup K)$. Protože $N > 2$ a neexistuje singleton (jednoprvková množina) $\{i\}$, která by byla rozhodující, pak J,K,L mohou existovat.

Důkaz Arrowse

Dále předpokládejme profil $\rho^- \in \mathcal{R}^N$ takový, že:

$$\begin{array}{ll} x \succ_i^- y \succ_i^- z & \forall i \in J \\ z \succ_j^- x \succ_j^- y & \forall j \in K \\ y \succ_t^- z \succ_t^- x & \forall t \in L \end{array}$$

Protože J a K jsou ne-semi-rozhodující pro jakýkoliv pár, plyne z toho celkově $z \succeq^- x$ a $y \succeq^- z$. Pokud je F opravdu tranzitivní, znamená to navíc $y \succeq^- x$.

Z toho plyne, že skupina $J \cup K$ je ne-semi-rozhodující pro (x, y), a tím že $J \cup K$ je celkově nerozhodující. Sjednocení dvou nerozhodujících skupin tvoří tedy opět nerozhodující množinu.

Důkaz Arrowse

Z předpokladu nediktátorství plyne, že žádná singleton skupina nemůže být rozhodující.

Závěrem tedy žádná skupina voličů voličů (včetně Q) nemůže být rozhodující a to porušuje předpoklad slabé Paretovskosti.

□

Pozor: našli jsme hypotetickou konfiguraci, kdy může nastat, že celá skupina Q může být ne-semi-rozhodující jeden pár alternativ, což implikuje, že nerozhodne nic. Jiné teorémy zjemňují Arrowův paradox zavedením omezení, která vylučují takové blokující situace.

Pozn.: Existují formulace důkazu Arrow's Impossibility Theoremu, které prokáží existenci diktátora (dále v projektech).

Funkce veřejné volby

Zatím jsme konstruovali funkci, která vytvářela kolektivní preferenci. Zjednodušíme problém na volbu jednoho vítěze (už nebudeme potřebovat konstruovat kompletní preferenci).

Definition

Funkce veřejné volby $G : \mathcal{R}^N \rightarrow A$ zobrazuje dané preferenční profily na množinu alternativ.

Funkce G je funkce na množinu, tzn. každá volba $a \in A$ má preferenční profil, který vede na a.

Strategická manipulovatelnost

Definition

$G(\rho)$ je manipulovatelná v profilu ρ, pokud pro některého $i \in Q$ existuje profil $\rho' = (R_1, ..., R_{i-1}, R'_i, R_{i+1}, ...)$ takový, že $G(\rho') \succ_i G(\rho)$. G je nemanipulovatelná (incentive compatible), pokud není manipulovatelná ve všech $\rho \in \mathcal{R}^N$.

Funkce veřejné volby je tedy manipulovatelná, pokud některý hráč může změnou prezentace své preference dosáhnout výsledku, který preferuje.

Paralela: preference hráče je jeho strategie, profil je strategický profil. Užitek hráče je dán umístěním vítěze v hráčově preferenci. Pokud hráč dokáže změnou své strategie posunout svého preferovaného kandidáta výše k výhře, zvyšuje svůj užitek.

Gibbard-Satterthwaitův teorém

Definition

Volič i je diktátor v rámci G, jestliže pro všechny $\rho \in \mathcal{R}^N$ a $\forall b \in A : b \neq a : a \succ_i b \Rightarrow G(\rho) = a$. G se nazývá diktatura, pokud existuje i takový, že je diktátor.

Diktátor se tedy prosadí v každém profilu. Zatím jsme uvažovali manipulovatelnost.

Theorem

Pokud existují tři a více alternativy, a G je nemanipulovatelná, pak existuje diktátor.

Co z toho plyne? Můžeme zajistit, aby G byla odolná proti strategické manipulaci, diktátor se však stejně projeví.

Gibbard-Satterthwaitův teorém

Zjednodušeně důkaz tvrzení: Máme množinu alternativ A. Pak můžeme sestrojit tranzitivní a IIA funkci veřejné volby, která zvolí vítěze $v_1 = G(A)$. V dalším kole ubereme z A prvek v_1 a hledáme $v_2 = G(A \setminus \{v_1\})$ a tak dál. Vytvoříme tedy preferenční uspořádání $v_1 \succ v_2 \succ ...$, které je ovšem slabě Paretovské.

Dostali jsme se tak k vytvoření společenské preference (funkce společenského blahobytu) a z Arrowova teorému plyne, že proces musel ovlivnit diktátor.

Stable Matchings problem

- Model zpárování podle preferencí (muži a ženy, studenti a koleje, zaměstnanci a firmy).
- Množina $M = \{m_1, m_2, ...\}$ mužů a množina $W = \{w_1, w_2, ...\}$.
- Zavedeme preference obvyklým způsobem, předpokládáme $|M| = |W|$.
- Zpárování je vytvoření $R \subset M \times W$, že každý muž je spárován právě s jednou ženou a naopak

Stable Matchings problem

Ukázka dělení 1:1 nedělitelného zboží bez postranních plateb (finanční kompenzace). Existují další modely.

Definition

Mějme Stable-Matchings problém. Zpárování se nazývá nestabilní, pokud jsou dva muži m, m' a dvě ženy w, w', že:

- m je zpárováno s w
- m' je zpárováno s w'
- $w' \succ_m w$ a $m \succ_{w'} m'$

Pár (m, w') se nazývá blokující pár. Zpárování, kde není blokující pár, se nazývá stabilní. Skupinová racionalita (skupina se může odtrhnout a zařídit se po svém).

Demo

$\succ_{m_1}$	$\succ_{m_2}$	$\succ_{m_3}$	$\succ_{w_1}$	$\succ_{w_2}$	$\succ_{w_3}$
w_2	w_1	w_1	m_1	m_3	m_1
w_1	w_3	w_2	m_3	m_1	m_3
w_3	w_2	w_3	m_2	m_2	m_2

Zpárování $\{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ je nestabilní, protože (m_1, w_2) je blokující pár. Zpárování $\{(m_1, w_1), (m_3, w_2), (m_2, w_3)\}$ je stabilní.

Otázka: je vždy možné najít stabilní zpárování?

Algoritmus pro zpárování–mužská iniciativa

1. Každý muž zašle návrh své nejpreferovanější ženě.
2. Každá žena, která obdržela více návrhů, vybere z nich nejpreferovanějšího muže a zbytek odmítne.
3. Odmítnutí muži zašlou návrh nejpreferovanější ženě, která je ještě neodmítla.
4. Opakuje se, dokud existuje nezpárovaný muž.

Theorem

Algoritmus pro zpárování–mužská varianta končí ve stabilním zpárování.

Důkaz při domácí přípravě ([Nisan]).

Algoritmus pro zpárování–mužská iniciativa

- Označme zpárování jako μ. Žena w přiřazená v párování μ muži m je označena $\mu(m)$. Podobně muž: $\mu(w)$.
- Zpárování μ je výhodnější pro muže (male-optimal), pokud neexistuje jiné zpárování ν:
  - $\forall m \in M : \nu(m) \succ_m \mu(m)$
  - $\forall m \in M : \nu(m) = \mu(m) \wedge \exists j \in M : \nu(j) \succ_j \mu(j)$
  - tzn. slabá preference ν nad μ
- Podobně lze definovat zpárování výhodnější pro ženy (female-optimal).

Theorem

Algoritmus pro zpárování–mužská varianta je výhodnější pro muže (male-optimal).

Důkaz při domácí přípravě ([Nisan]).

Algoritmus pro zpárování

Může existovat zpárování, které by nezvýhodňovalo jednu skupinu? Nelze...
Může existovat skupina $S \subset M \cup W$, která by měla zájem na vlastní dohodě? Formálně:

Definition

Zpárování μ' dominuje nad μ, pokud existuje skupina $S \subset M \cup W$ taková, že pro všechny $m, w \in S$ platí současně:

- $\mu'(m), \mu'(w) \in S$
- $\mu'(m) \succ_m \mu(m)$ a $\mu'(w) \succ_w \mu(w)$

Stabilita je zvláštní případ dominance, kde se omezíme pouze na množiny S obsahující pouze jeden pár. Množina nedominovaných zpárování se nazývá jádro (core) párovací hry.

Algoritmus pro zpárování–závěr

Theorem

Jádro párovací hry je tvořeno množinou stabilních zpárování.

Zatím jsme předpokládali, že preference hráčů jsou common knowledge. Řekneme, že funkce společenské volby je strategy-proof, pokud je pro hráče slabě dominantní strategií prezentovat své preference pravdivě.

Theorem

Algoritmus pro zpárování–mužská varianta je ze strany mužů strategy-proof.

Důkaz, že muži říkají pravdu je na straně 258 [Nisan].

Příště

Mechanismy s penězi.

- Aukce.
- Vickrey-Clarke-Groves mechanismus.
- Revenue Equivalence Theorem.