THE: Dymické hry s úplnou informací
Opakované hry
Dynamic Games (Extensive Form Games), Repeated Games

Martin Hrubý

Brno University of Technology
Brno
Czech Republic

November 6, 2017

War and peace (R. Aumann, Nobel prize lecture)

- Teorie her není schopna dodat magickou formuli, která vyřeší všechny konflikty. Neexistuje akademická disciplína, která by toto dokázala.
- Cílem je porozumět konfliktu. Když konfliktu porozumíme, jsme schopni hledat jeho řešení.
  - Je to jako s rakovinou: A) konvenční léčba (řešit konkrétní projevy/následky), B) snaha pochopit vnitřní principy (DNA, procesy v těle).
  - Na válku se pohlíží jako na iracionální chování a to je chyba. Válka není iracionální.
  - Jedinec je racionální, pokud koná to nejlepší dle jeho informací.
  - Hraje zřejmě nekooperativní ekvilibrium.
- Kooperace – předpokládáme, že ve hře existuje efektivnější profil než je NE.
- Profil je kooperativní, pokud žádný hráč není schopen pro sebe garantovat lepší výsledek.

Repetition Enables Cooperation. Folk theorem.

Demo: dopravní model

Jet do práce cestou *a* nebo *b*? Když pojedou stejnou, cestovní doba se protáhne.

Petr/Jan	a	b
a	-10,-10	-5,-1
b	-1,-5	-20,-20

$PNE = \{(a, b), (b, a)\}$, $MNE = ((\frac{5}{8}, \frac{3}{8}), (\frac{5}{8}, \frac{3}{8}))$
Jak interpretovat dvě možná PNE? Jak Petr přesvědčí Jana k (b, a)? Signalizace (sekvenční hraní).

Kooperace. Může být kooperace (zdánlivá iracionalita) racionální?

Opakování hry a trest.

Sekvenční hry. Úvod

Čerpáno z:

- Myerson, R. B.: Game theory: Analysis of Conflict, Harvard University Press, 2004
- *Rasmunsen, E.: Games and Information: An Introduction to Game Theory, Blackwell Publishing, 2007*
- Mailath, G. J., Samuelson, L.: Repeated Games and Reputation, Oxford University Press, 2006
- *Basar, T., Olsder, G.J.: Dynamic Noncooperative Game Theory , Society For Industrial And Applied Mathematics, 1999*
- Magdaléna Hykšová: Přednášky z teorie her, Fakulta dopravní ČVUT
- Asu Ozdaglar: přednášky "Repeated Games"
- Ratliff, J.: A Folk Theorem Sampler

Připomenutí statických her

Hry v normální formě (statické hry).

- Hráč strategicky promýšlí možné tahy svoje a protivníka.
- Asi by mu hodně usnadnilo přemýšlení, kdyby věděl, co protivník bude táhnout.
- Kdyby se mu podařilo (tajně) vyzvědět (v čase libovolně blízkém uzávěrce rozhodování), co protihráč odevzdal za rozhodnutí tomu pomyslenému nezávislému arbitrovi, tak táhne velmi efektivně.
- Ve statických hrách toto nepřipouštíme.

V sekvenčních hrách předpokládáme, že se hráči střídají v tazích a tuto skutečnost si uvědomují. Přesnější název je ovšem *hry v rozšířené formě* – rozšíření není pouze v sekvenčnosti.

Stackelbergův model duopolu/oligopolu

Heinrich Freiherr von Stackelberg: Market Structure and Equilibrium (Marktform und Gleichgewicht) in 1934

- Bertrandův a Cournotův model předpokládá normální přístup ke hře. Hráči jsou (přibližně) stejně silní.
- Tento předpoklad neodpovídá mnohým reálným situacím s jedním extrémně silným hráčem (vůdcem, leader) a skupinou slabších hráčů (následovníků, followers).
- Stále doufáme, že všichni hráči jsou svým rozhodnutím schopni ovlivnit cenu produktu (tzn. není to monopol).
- Vznikají ovšem dvě role a dvě podoby (odlišného) chování: vůdce a následovníci.
- Vůdce: je si vědom svého postavení. Ví, že následovníci se budou inspirovat jeho rozhodnutím.
- Následovníci: čekají na rozhodnutí vůdce, pak se přizpůsobí.

Z principu tohoto rozložení sil neproběhne hra v jednom tahu, ale ve dvou.

Stackelbergův model oligopolu

Příklady ze života:

- Předpoklad: pohybujeme se v systému s nedokonalou konkurencí.
- Výroba silové elektřiny v ČR: ČEZ – vůdce, IPP (Independent Power Producers) – následovníci.
- IPP čekají na cenové rozhodnutí ČEZu, pak nastaví cenu za MWh o tři CZK nižší.
- IPP tudíž svou produkci vždy prodají (o 3 CZK/MWh levněji).
- Vůdce si tuto skutečnost uvědomuje a zahrnuje ji do svého rozhodování.

Různé přístupy v situacích, kdy je následovníků více (vícehráčovo Stackelbergovo ekvilibrium, ...).

Princip zpětné indukce, zatím neformálně

Backward induction.

- Vůdce tedy nemůže nechat vypůsobit následovníky a pak se rozhodnout (oni právě čekají na jeho tah).
- Vůdce experimentuje s možnou reakcí následovníků a hledá strategii, která mu přinese maximální zisk
  - Ta strategická úvaha je na vůdci.
  - Následovníci pouze zareagují na tah vůdce.
  - Předpokládáme (i vůdce předpokládá), že následovníci potáhnou BR.
- Zpětná indukce vede k NE, je to ovšem podmnožina možných PNE, kde nejsou zahrnuty iracionální varianty (hrozby).

Pozn.: Můžeme zavést komplikovanější chování, jako například systém vůdce versus oligopolistické chování následovníků.
Pozn.: Taková perlička - Unexpected hanging paradox.

Stackelbergův model oligopolu: design modelu

Model je v množstevních strategiích.
Firma 1 zvolí množství $q_1 \geq 0$,
Firma 2 to sleduje, zhodnotí a volí $q_2 \geq 0$.

Výplata firmy i je pak určena

$$u_i(q_i, q_j) = q_i[P(q) - c]$$

kde $P(q) = M - q$, $q = q_1 + q_2$.
c jsou výrobní náklady; $P(q)$ je cena na trhu při dodaném množství q.

R_1, R_2 budou rozhodnutí firem o množstevní strategii.

Zopakujme $u_i(q_i, q_j) = q_i[P(q) - c]$.

Backwards-induction: $R_2(q_1)$:

$$\max_{q_2 \geq 0} u_2(q_1, q_2) = \max_{q_2 \geq 0} q_2[M - q_1 - q_2 - c]$$

Hledáme tedy maximum funkce s proměnnou q_2. Matematická analýza: vyšetřování extrémů funkcí.

Hledáme extrém: $\frac{\partial f(q_2)}{\partial q_2}; f(q_2) = q_2 M - q_2 q_1 - q_2^2 - cq_2$
Bod extrému: $M - q_1 - 2q_2 - c = 0$

$$q_2 = R_2(q_1) = \frac{M - q_1 - c}{2}; q_1 < M - c$$

$R_2(q_1)$ je analyticky vyjádřené množstevní rozhodnutí následovníka vztažené k rozhodnutí vůdce. Vůdce ví, že takto se rozhodne následovník a s tímto vědomím volí jeho optimální strategii.

$$[q_1 \geq 0] \quad max \quad u_1(q_1, R_2(q_1)) = \quad max \quad q_1[M - q_1 - R_2(q_1) - c]$$
$$max \quad q_1(M - q_1 - c)/2$$

$$f(q_1) = \frac{1}{2}q_1 M - \frac{1}{2}q_1^2 - \frac{1}{2}q_1 c$$

$$\frac{1}{2}M - 2\frac{1}{2}q_1 - \frac{1}{2}c = 0$$

$$M - 2q_1 - c = 0$$

$$q_1^* = \frac{M - c}{2}; \quad q_2^* = R_2(q_1^*) = \frac{M - c}{4}$$

Ekvilibrium je $q^* = (q_1^*, q_2^*) = \left(\frac{M-c}{2}, \frac{M-c}{4}\right)$.

Pokud tedy vůdce dodá na trh jiné (např. větší) množství než $\frac{M-c}{2}$, rozhodně jeho zisk bude nižší. Jako domácí cvičení toto prověřte. To samé platí pro následovníka. Následovník ovšem má vždy možnost volit BR a tuto volbu táhnout.

Modely oligopolu

Známe již tři pohledy na věc:

- Cournot – oligopol v množstevních strategiích, statická hra.
- Bertrand – oligopol v cenových strategiích, statická hra.
- Stackelberg – oligopol v množstevních strategiích, sekvenční hra.

Tyto modely jsou common knowledge pro každého herního teoretika.

Dnes víme, že se hráči nenechají strhnout k cenové hře, která může vést k *cenové válce*. Po cenové válce je totiž rozložení vlivu na trhu přibližně podobné, jako na začátku, pouze hráči méně trží.

Princip zpětné indukce

Pedro/Juana	Opera	Fotbal
Opera	3,2	0,0
Fotbal	0,0	2,3

Pedro táhne první. Ví, že Juana zvolí BR. Proto se podívá, co by byly ty BR (Oo, Ff). Z možností (Oo,Ff) volí Oo.
Čím to, že je tu naprosto jasné unikátní PNE?

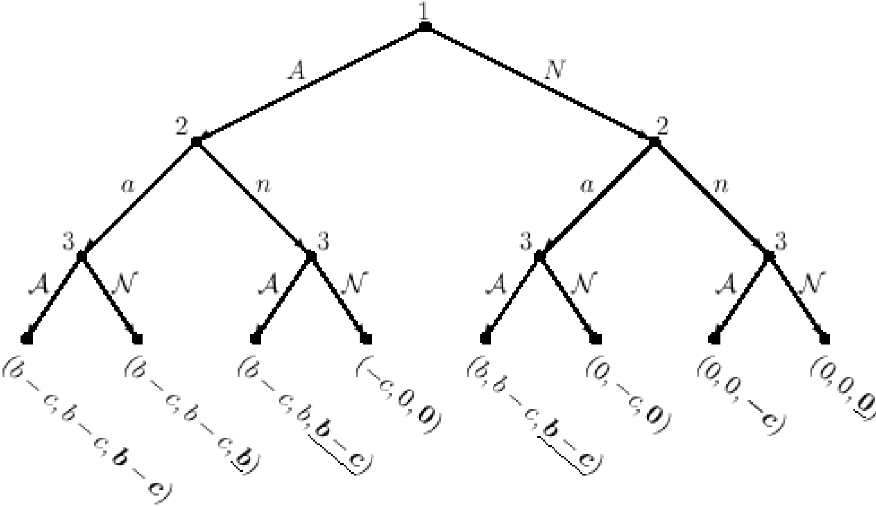
Pořadí tahů

Je výhodné táhnout první?

Příklad s poslanci: Poslanci chtějí odhlasovat zákon o zvýšení svých platů. Zvýšení přinese každému poslanci individuální užitek $b \geq 0$, ovšem poslanec hlasující "pro" sníží u věřejnosti svou popularitu, což odpovídá ztrátě $c \geq 0$ (naštěstí $b > c$). Poslanci jsou tři, k odhlasování zákona potřebujeme dva hlasy pro. K hlasování přistupují sekvenčně a hlasování je veřejné. Kdo půjde první?

- Člověk bez znalosti THE by se možná ostýchal jako první hlasovat o tak choulostivé věci veřejně.
- Je evidentní, že zákon projde.
- Pokud budu mít to štěstí, že budu moci jít hlasovat první, radostně budu hlasovat "proti" s přidáním nějaké zásadní řeči o nestoudných mzdových požadavcích některých poslanců.

Pořadí tahů


1
A
N
2
2
a
n
a
n
3
3
3
3
A
N
A
N
A
N
A
N
(b − c, b − c, b − c)
(b − c, b − c, b)
(b − c, b, b − c)
(−c, 0, 0)
(b, b − c, b − c)
(0, −c, 0)
(0, 0, −c)
(0, 0, 0)

Důveryhodná/nedůveryhodná hrozba

Credible/Incredible threat.

Uvažujme situaci, kdy hráč A přijde k hráči B s tím, že má bombu a pokud mu B nedá jistý obnos (nemá smysl zkoumat jeho výši), tak bombu odpálí (způsobí smrt A i B).
Hráč B očekává u A racionalitu a zkoumá, zda-li je hrozba "odpálit" důvěryhodná. Očekává, že A je racionální (dá přednost životu). Pak nemá smysl hrát strategii "Odevzdat". Paradoxně, většina lidí však peníze zaplatí.

Hráč zkoumá, zda-li má racionálně zahrnout hrozbu protivníka do svého rozhodování (stále předpokládá racionalitu protivníka).

Nedůvěryhodnou hrozbu hráč ohlašuje (signalizuje), pokud doufá, že ji nebude muset naplnit. Aby se hrozba stala důvěryhodnout, hráč mnohdy sobě způsobí škodu (pálení lodí).

Důveryhodná/nedůveryhodná hrozba

A/B	left	right
top	2,1	0,0
bottom	1,2	1,2

- PNE: (top,left), (bottom, right)
- Ikdyž je right slabě dominovaná strategií left, přesto není iracionální. Tvoří hrozbu.
- Musí se však A obávat right?
- Ve statické hře sloupcový volí left.
- Sekvenční chování může potlačit vliv hrozby.

Informace ve hře

- Hry s úplnou/neúplnou informací (complete/incomplete) – hráčům je známa/neznáma struktura hry (hráči, strategie, zisky).
- Hry s dokonalou/nedokonalou informací (perfect/imperfect) – hráčům je známa historie hry.
- Nás budou zajímat hry s úplnou informací.
- Dále budeme zkoumat perfektnost/imperfektnost informace.

Hry proti přírodě – první tah má příroda (zdroj nedokonalosti v informaci).

Historie znamená informaci o aktuálním uzlu ve stromu hry. Pokud je znalost uzlu jistá (informační množina je jednoprvková – singleton), pak je perfektní informace. Nejistotu modelujeme víceprvkovou informační množinou.

Perfect recall – hráč je schopen si pamatovat vlastní tahy.

Definice sekvenční hry (Extensive Form Game)

Definition

Konečná sekvenční hra s perfektní informací Γ^E je čtveřice (Q, H, p, U), kde:

- Q je množina hráčů,
- H je množina historií (kontextů, uzlů). H^T označuje množinu terminálních historií. $H^0 = \emptyset$ označuje počáteční uzel.
- $p(h) : H \setminus H^T \to Q$ přiřazuje každé neterminální historii h hráče, který je aktuálně na tahu.
- U je vektor užitkových funkcí $U_i(h) : H^T \to \mathbb{R}^1$ pro všechny $i \in Q$.

Pozn.: Mininum pro definici hry s perfektní informací je (Q, H, p).
Pozn.: Zavedeme: $A(h)$ je množina akcí pro každé $h \in H \setminus H^T$, tzn. množina proveditelných akcí po dosažení historie h.

Co je ta historie?

Budeme chtít vyjádřit, že hráč neví, v jakém stavu (uzlu) stromu je.

- Historie $h \in H$ je sekvence akcí $(a_k)_{k=1,\ldots}$, kde každá akce a_j představuje tah hráče
- Prázdná sekvence je H^0, počátek hry.
- Je-li $(a_k)_{k=1,\ldots,K} \in H$ $(K \in \mathbb{N})$ a $L < K$, pak $(a_k)_{k=1,\ldots,L} \in H$.
- Každý prvek $h \in H$ se nazývá historie, každá složka h se nazývá akce.

Množina H dává představu o způsobu hraní hry, je to množina všech cest od počátku k libovolnému uzlu stromu hry. Předpokládáme tedy, že v $h = (a_k)_{k=1,\ldots,K}, j \in \{1, \ldots, K\}$ táhl v j-tém okamžiku hráč $p(h')$ akci a_j, kde $h' = (a_k)_{k=1,\ldots,j-1}$.

Co je ta historie?

- Je-li H konečná množina, je i příslušná hra konečná.
- Je-li nejdelší prvek $h \in H$ konečně dlouhý, má hra konečný horizont.
- Je-li h historií délky k, pak (h, a) je historií délky $k + 1$. Skládá se pak z h následovaného a.
- Z toho plyne definice $A(h) = \{a | (h, a) \in H\}$.

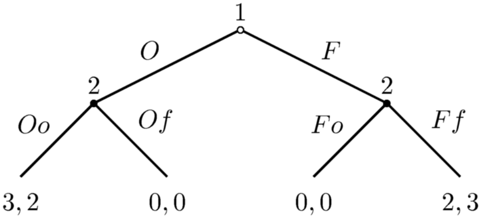
Zpětná indukce může být použita pouze na konečné hry.

Informační množiny typu $I \subseteq H \setminus H^T$

Mějme rozklad (relace, která je reflexivní, symetrická a tranzitivní) na množině $H \setminus H^T$, tzn. množinu množin, které dohromay tvoří $H \setminus H^T$.

- Jednu informační množinu budeme označovat I.
- Z principu relace "rozklad na množině" plyne, že každá $h \in H$ je právě v jedné množině rozkladu.
- je-li $h \in I$, pak je hráč $p(h)$ nejistý, zda-li je v uzlu h nebo v jiném $h' \in I$.
- $\forall h, h' \in I : p(h) = p(h')$.
- jsou-li všechny I jednoprvkové, pak je hra s dokonalou informací (perfect information), jinak nedokonalou (imperfect)

H, I



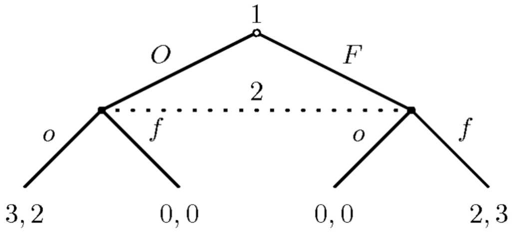
$$H = \{(), (O), (F)\} \cup H^T$$

$$H^T = \{(O, o), (O, f), (F, o), (F, f)\}$$

$$p(()) = 1, p((O)) = p((F)) = 2, A(()) = \{O, F\}$$

$$Is = \{\{\emptyset\}, \{O\}, \{F\}\}$$

H, I



$$Is = \{\{\emptyset\}, \{O, F\}\}$$

Pozn.: Hráč 2 ve svém okamžiku tahu neví, co hrál protivník (v jakém je on uzlu). Fakticky to odpovídá situaci hry v normální formě.

Zermelův teorém

Theorem

Každá konečná hra dokonalé informace má PNE, které je dosažitelné zpětnou indukcí.
Pokud žádný hráč nemá v žádných dvou terminálních historiích stejné užitky, je ve hře pouze jedno PNE.

Garance existence PNE!

Opakované hry

- Opakovaná hra je případ hry v rozšířené formě (sekvenční hry), která sestává z konečného/nekonečného opakování tzv. základní hry (stage game).
- V opakovaných hrách (specificky v těch nekonečně opakovaných) je racionální hrát *společensky optimální profil*, který se může značně lišit od NE.
- V opakovaných hrách s konečným (a předem známým) počtem opakování je racionální hrát v poslední hře NE a v průběhu ten koncept (NE, sociálně optimální profil), který je výhodnější s vědomím budoucího opakování hry. (př. chain-store paradox)
- Konečně-opakované hry lze řešit zpětnou indukcí.

Opakované hry, úvod

Zatím známe hry:

- Nekooperativní s nulovým/nenulovým součtem v normální (maticové) formě – hráči chápou situaci (jednotahovost) a strategicky analyzují množinu profilů.
- Nekooperativní hry v rozšířené formě – hráči se v tazích střídají, existuje víc tahů. Velmi záleží na stavu informace ve hře (dokonalá/nedokonalá informace o historii hry).
- Kooperativní hry definované charakteristickou funkcí – hráči hledají optimální rozložení koalic a optimální dělbu zisku přisouzené koalici (imputace).
  - Čím je podmíněna kooperativnost (garance, že protivník dodrží dohodu)?

Opakujeme tzv. "base game" (stage game) – základní hru.
Co je strategie v opakované hře? Návaznost na ESS.

Příhoda s Hodžou Nasredinem

Hodža Nasredin byl známý muslimský filosof a humorista. Jednou přišel do lázní, ale lazebník ho odbyl špatnými službami. Při odchodu z lázní Hodža zaplatil velkou mincí. Při příští návštěvě se lazebník překonával, ale Hodža při odchodu zaplatil malou mincí. Lazebník se divil: "Hodžo, posledně jsi zaplatil hodně za ... služby a za dnešní (lepší) služby málo. Proč?". Hodža odpověděl: "Posledně jsem platil za dnes a dnes jsem platil za posledně.".

Pokud lazebník pochopil Hodžovu strategii v opakované hře "návštěva lázní", pak budou od toho okamžiku oba hrát kooperativní profil.

Opakované hry nám umožňují modelovat sociální normy, vliv "trestu" na další opakování hry a přirozeně vnucenou spolupráci.

Opakované hry, úvod

Dále na úrovni nekooperativních her:

- Konfliktní situace se může opakovat. Pokud si hráči fakt opakování uvědomují, mluvíme o opakovaných hrách.
- Hráči se mohou poučit o chování protihráče, tzn. mají paměť.
- Je prostor pro vyjednávání. Je možno "trestat".
- Rozlišujeme dle opakování:
  - Konečné opakování – hra se opakuje T-krát, kde T je konkrétní předem známé číslo.
  - Nekonečné opakování – hra se opakuje bez ukončení. Není určen konec, tj. v každém kole předpokládám, že bude následovat další.
- Hráč na počátku hry ví o opakování a přizpůsobí tomu svůj první tah. Navazujeme tímto na sekvenční hry.

Hráč bere v úvahu, že jeho tah ovlivní bezprostřední výsledek hry, ale i budoucnost hry (vnímá svoji "reputaci").

Opakované vězňovo dilema

	C	D
C	1,1	-1,2
D	2,-1	0,0

Předpokládejme nyní strategii (nazvanou *grim trigger strategy*):

- Hraju C tak dlouho, dokud druhý hraje C.
- Pokud druhý zahraje D, od této doby hraju navždy D (trest).

Pokud někdo zahraje D, krátkodobě získá zisk 2, ale v dalších hrách už jenom 0 (namísto kooperativního 1). To znamená, že pro racionálně (citlivě k budoucnosti) uvažujícího hráče je zisk $(1, 1, 1, 1, ...)$ dominující nad $(..., 2, 0, 0, ...)$.

Strategie C je tedy best-response na *grim trigger strategy*. Tato strategie je NE v opakovaném Vězňově dilematu. Strategie (D, D) ovšem taky.

Strategie v opakované hře

Je nutno odlišovat dva pohledy na opakovou hru:

- Bázová (základní) hra má strategie dle $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$.
- Opakovaná hra má strategie definované nad celou posloupností opakování bázové hry.
- Situaci opakované hry modelujeme jako rozhodnutí hráče, jak se bude v rámci opakování her konzistentně chovat.

Příklady strategií v opakované hře:

- Vždy budu hrát ...
- Budu hrát strategii, kterou v předešlém kole hrál soupeř.
- Budu hrát kooperativní strategii, pokud mě soupeř zradí, budu ho trestat po dobu k kol hraním nekoop. strategie, a pak budu hrát opět kooperativní strategii (trest a pozdější odpuštění).

Soutěž v opakovaném Vězňově dilematu (Iterated P.D. Competition, R. Axelrod, 1979).

Preference v opakované hře, ohodnocovací faktor δ

Každý hráč má v každém kole t hry užitek $U_i(s^t), s^t \in S$ a definuje ohodnocovací faktor (discounting factor) $\delta_i \in \langle 0, 1)$, kterým ohodnocuje postupné výhry při hraní profilů $(s^1, s^2, ..., s^T)$, tedy $s^j \in S$:

$$U_i(s^1) + \delta_i U_i(s^2) + ... + \delta_i^{T-1} U_i(s^T) = \sum_{t=1}^{T} \delta_i^{t-1} U_i(s^t)$$

Pozor, δ_i^t je $t-$mocnina δ_i.

- Pokud je δ_i blízko 0, hráč i je velmi necitlivý k budoucnosti a naopak.
- Pokud je $\delta_i = 0$, pak hráč i neví o budoucnosti (jednotahová hra).
- Př.: $\delta_i = \frac{1}{2}$, pak dnešních 100 CZK je ekvivalentní zítřejším 200 CZK
- Budeme nadále klást $\delta_i = \delta; \forall i \in Q$.

Preference v opakované hře

Chceme ohodnotit posloupnost zisků a *vyjádřit ji jedním číslem*. Pomocí něho budeme porovnávat posloupnosti zisků (preference nad posloupnostmi).

Posloupnost $W = (w^1, w^2, ..., w^T)$ má sumu $V = \sum_{t=1}^{T} \delta_i^{t-1} w^t$. Existuje nyní takové číslo c, že by byl hráč indiferentní mezi W a $(c, c, ...)$?

Suma pro $(c, c, ...)$ je $\frac{c}{1-\delta}$ (suma geometrické řady). Hráč je indiferentní mezi W a $(c, c, ...)$, pokud $c = (1-\delta)V$. Proto nazýváme $(1-\delta)V$ průměrným ziskem, tzn.

$$\pi_i = (1-\delta) \sum_{t=1}^{T} \delta_i^{t-1} w^t$$

Definice hry v rozšířené formě (opakování)

Definition

Konečná sekvenční hra s dokonalou informací je Γ^E je čtveřice (Q, H, p, U), kde:

- Q je množina hráčů,
- H je množina historií (kontextů, uzlů). H^T označuje množinu terminálních historií. $H^0 = \emptyset$ označuje počáteční uzel.
- $p(h) : H \setminus H^T \to Q$ přiřazuje každé neterminální historii h hráče, který je aktuálně na tahu.
- U je vektor užitkových funkcí $U_i(h) : H^T \to \mathbb{R}^1$ pro všechny $i \in Q$.

Pozn.: Miminum pro definici hry s dokonalou informací je (Q, H, p). Pozn.: Zavedeme: $A(h)$ je množina akcí pro každé $h \in H \setminus H^T$, tzn. množina proveditelných akcí po dosažení historie h.

Definice opakované hry

Definition

Mějme $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$. T-krát opakovaná hra Γ^T s bázovou hrou Γ a ohodnocovacím faktorem δ je hra v rozšířené formě s dokonalou informací a současnými tahy hráčů, kde:

- množina terminálních historií je množina sekvencí $(s^1, s^2, ..., s^T)$ profilů hry Γ.
- $p(h) = Q$ pro všechny historie $h = (s^1, s^2, ..., s^t)$ $(\forall t \in \{1, ..., T-1\})$.
- $A_i(h) = S_i$ pro všechny historie $h = (s^1, s^2, ..., s^t)$ $(\forall t \in \{1, ..., T-1\})$.
- každý hráč i ohodnocuje každou terminální historii $(s^1, s^2, ..., s^T)$ dle svého průměrného zisku $(1-\delta)\sum_{t=1}^{T} \delta^{t-1} U_i(s^t)$.

Poznámka k definici: nekonečně opakované hry

V případě nekonečně opakované hry Γ^∞:

- Terminální historie jsou nekonečné sekvence $(s^1, s^2, ...)$.
- Průměrný zisk je $(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} U_i(s^t)$.

V obou případech (konečné, nekonečné opakování) je terminální historie nazývána taky "outcome path" (výsledková posloupnost).

Path ve smyslu cesta grafem hry v rozšířené formě.

Zpátky k vězňům

	C	D
C	1,1	-1,2
D	2,-1	0,0

Hraje se T-krát:

- $T = 1$, pak je $NE = (D, D)$
- $T > 1$ a T je konečné, zpětnou indukcí získáme SPNE (Subgame Perfect NE), které je opět $NE = (D, D)$

$T = 2$ - v druhém kole mě zradí. Mám důvod věřit tomu, že při mé strategii C hra v prvním kole skončí v (C, C)?

Zpátky k vězňům

Hraje se ∞-krát:

Proposition

Je-li $\delta \geq \frac{1}{2}$, pak má ∞-krát opakované Vězňovo dilema (dle předchozího zadání) SPNE (C, C), které se hraje v každém tahu.

Proof.

Předpokládejme grim trigger strategii.
Pokud hráč i bude hrát stále C, je jeho průměrný zisk $(1-\delta)[1+\delta+\delta^2+...] = 1$. Zahraje-li D v libovolném, např. prvním, tahu, pak je jeho zisk $(1-\delta)[2+0+0+...] = 2(1-\delta)$. Pokud je $\delta = \frac{1}{2}$, pak je to srovnatelný zisk, pokud $\delta > \frac{1}{2}$, pak je horší.

$\square$

Co je Nashovo ekvilibrim v opakovaných hrách?

- Obecný význam NE známe (je to profil, kde žádný hráč změnou své strategie nezíská pro sebe lepší výsledek).
- Definice strategie v opakovaných hrách byla dána (je to chování hráče přes všechna opakování hry).
- Zisk hráče v opakované hře je suma všech dosažených zisků přes všechna opakování.
- Hráči hledají dlouhodobou strategii, která jim přinese optimální zisk – v NE.

Vytvoříme superhru (supergame), která je pro nás obvyklá nekooperativní hra v normální formě.

Ukázka superhry nad Vězňovým dilematem

Uvažujeme superhru Γ nad Vězňovým dilematem, které bude T-krát/∞-krát opakováno, kde $S_i = \{trigger, C, D\}$.

Průměrné užitky jsou:

A/B	trigger	C	D
trigger	(1,1)	(1,1)	(0,0)
C	(1,1)	(1,1)	(-1,2)
D	(0,0)	(2,-1)	(0,0)

Folk theorems (General feasibility theorems, R. Myerson)

Forma "solution concept" v opakovaných hrách.

Folk theorémy jsou třída vět, které v opakovaných hrách ukazují, že kterýkoliv výsledek hry je přijatelné řešení, jestli je pro hráče alespoň tak dobrý jako jejich minimax zisk.

Definition

Mějme hru $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$. Minimax zisk hráče $i \in Q$ ve hře Γ je:

$$\underline{v_i} = \min_{s_{-i} \in S_{-i}} \left(\max_{s_i \in S_i} [U_i(s_i, s_{-i})] \right)$$

$\underline{v_i}$ je tedy nejnižší BR_i. Předpokládáme, že hráči jsou schopni si kolektivně uvědomit míru přijatelného výsledku hry (feasible outcome). Ten představuje jakousi sociální normu.

Minimax strategie

Důkaz je založen na použití *grim trigger* strategie.

Definition

Řekneme, že výplatní vektor v je striktně individuálně racionální, pokud platí pro všechny hráče $i \in Q : v_i > \underline{v_i}$.

Definition

Minimax strategií proti hráči i nazveme sub-profil m^i_{-i} (ve dvouhráčové hře strategii protivníka), která mu přinese $\underline{v_i}$. Je to:

$$m^i_{-i} = arg \min_{s_{-i} \in S_{-i}} \left[\max_{s_i \in S_i} U_i(s_i, s_{-i}) \right]$$

Plyne z toho, že $BR_i(m^i_{-i}) = s'_i$ taková, že $U_i(s'_i, m^i_{-i}) = \underline{v_i}$.

Folk theorem

Doplňme, že $V^* \subset \mathbb{R}^N$ je množina přijatelných (stejných jako $\underline{v_i}$) a striktně individuálně racionálních zisků.

Folk theorems ukazují, že pokud jsou hráči dostatečně citliví k budoucnosti (mají dostatečně vysoké δ_i), pak jsou schopni dosáhnout libovolného dosažitelného stritně individuálně racionálního zisku formou NE.

Theorem

Pro všechny $v = (v_1, v_2, ..., v_N) \in V^*$ existuje $\underline{\delta} < 1$ takové, že pro všechny $\delta > \underline{\delta}$, existuje NE v $\Gamma^\infty(\delta)$ s vektorem zisků v.

Když hrajete s člověkem, jehož $\delta_i \to 0$, nemá smysl doufat v trestání.

Folk theorem, důkaz

Předpokládejme nejdříve, že existuje profil $s = (s_i)_{i \in Q}$ takový, že $U_i(s) = v_i$, $(v_i)_{i \in Q} \in V^*$ (jinak bychom museli zavést smíšené strategie, jde to, ale to bychom si to značně zkomplikovali).

Zavedeme si trigger strategii:

1. Hrej od nultého kola s_i tak dlouho, dokud protihráč hraje s_{-i}. Pokud protihráč(i) vybočí, přesuň se do stavu 2.
2. Hrej už jenom m_i^j (strategie, která garantuje $\underline{v_i}$).

Pokud hráči hrají kooperativně, mají ve hře zisk v. Hráč i, který vybočí v kole t, aby získal $\overline{v_i} = BR_i(s_{-i})$, získá maximálně:

$$(1-\delta)\left[v_i + \delta v_i + \cdots + \delta^{t-1} v_i + \delta^t \overline{v_i} + \delta^{t+1} \underline{v_i} + \delta^{t+2} \underline{v_i} + \cdots\right]$$

$$(1-\delta)\left[v_i + \delta v_i + \cdots + \delta^{t-1} v_i + \delta^t \overline{v_i} + \delta^{t+1} \underline{v_i} + \delta^{t+2} \underline{v_i} + \cdots\right]$$

$$= (1-\delta)[v_i + \delta v_i + \cdots + \delta^{t-1} v_i + \delta^t \overline{v_i} + \frac{\delta^{t+1} \underline{v_i}}{1-\delta}] =$$

$$= (1-\delta) v_i + (1-\delta)[\delta v_i + \cdots + \delta^{t-1} v_i] + (1-\delta)\delta^t \overline{v_i} + \delta^{t+1} \underline{v_i}$$

Druhý člen:

$$(1-\delta)[\delta v_i + \cdots + \delta^{t-1} v_i] = v_i(\delta + \cdots + \delta^{t-1} - \delta^2 - \cdots + \delta^t) = v_i(\delta - \delta^t)$$

První a druhý člen:

$$(1-\delta) v_i + v_i(\delta - \delta^t) = v_i - \delta v_i + \delta v_i - v_i \delta^t = v_i(1-\delta^t)$$

Celkově (zisk hráče i, který v kole t vybočil z kooperativního profilu):

$$= (1-\delta^t)v_i + \delta^t(1-\delta)\overline{v_i} + \delta^{t+1}\underline{v_i}$$

Aby navrhovaná strategie (tzn. hrát kooperativně) byla optimální, musí platit:

$$v_i \geq (1-\delta^t)v_i + \delta^t(1-\delta)\overline{v_i} + \delta^{t+1}\underline{v_i}$$

$$v_i \geq v_i - \delta^t[v_i - \overline{v_i} + \delta\overline{v_i} - \delta\underline{v_i}]$$

$$v_i \geq v_i - \delta^t[v_i + (\delta-1)\overline{v_i} - \delta\underline{v_i}]$$

To bude platit, pokud:

$$v_i + (\delta-1)\overline{v_i} - \delta\underline{v_i} \geq 0$$

Přelomová je hodnota

$$\underline{\delta} = \max_{i \in Q} \left[\frac{\overline{v_i} - v_i}{\overline{v_i} - \underline{v_i}} \right]$$

$$\underline{v_i} < v_i < \overline{v_i} \Rightarrow \overline{v_i} - v_i < \overline{v_i} - \underline{v_i} \Rightarrow \underline{\delta} < 1$$

Pro všechny $\delta \geq \underline{\delta}$ je racionální hrát zvolenou strategii a neuhýbat. Resp. při $\delta = \underline{\delta}$ je deviující hráč na stejném zisku, jako by kooperoval, při $\delta > \underline{\delta}$ má striktně horší zisk.

$\square$

Tento Folk theorem ukazuje, že pro jistou míru paměti (nebo citlivosti k budoucnosti) je protihráč racionálně nucen ke kooperativní strategii.

Příklad

	L	R
U	6,6	0,-100
D	7,1	0,-100

$NE = (D, L)$, $\underline{v_1} = 0$, $\underline{v_2} = 1$. Minimax strategie hráče 2 je hrát R.

Nash Folk Theorem připouští hraní $(U, L) \Rightarrow (6, 6)$, ovšem hráč 2 musí hrát R kdykoliv hráč 1 uhne (a to je značně bolestivé).

$$\underline{\delta_i} = \frac{\overline{v_i} - v_i}{\overline{v_i} - \underline{v_i}} = \frac{7-6}{7-0} = \frac{1}{7}$$

Další Folk Theoremy v rámci projektů.

Jak lidé hrají v opakovaných hrách

Podle laboratorních experimentů (Asu Ozdaglar):

- Lidé hrají v opakovaných vězňových dilematech kooperativněji, než by se očekávalo.
- Dokonce i u konečně opakovaných hrách.
- Lidé jsou konfrontováni s maticí zisků, ale to není jejich veškerý užitek.
- Další složky užitku pramení ze sociálních preferencí.

Typy sociálních preferencí lidí:

- Altruismus – mají užitek z toho být hodní na druhé.
- Férovost – mají užitek z toho být fair na druhé.
- Pomstichtivost – lidé rádi trestají ty, co vybočí z férovosti nebo jiných sociálních norem.

Příště

Máme uzavřenu část definice her, herních konceptů a ekvilibrií.

Následují aplikace a druhá strana THE:

- Mechanism design – teorie veřejné volby a aukce.
- Evoluční teorie her.
- Rozbor rozsáhlého modelu z elektroenergetiky.
- Závěrečné opakování (???).