Contents

Chapter 1

Repeated games

1.1 Úvod

V předchozích kapitolách jsme se zabývali různými formami nekooperativních her, tedy her, kde hráči nechtějí, nesmí nebo neumí kooperovat. Koncepce nekooperativních her prostě ve svých pravidlech hry nepřipouštěla vyjednávání a omezovala se pouze na individuální racionalitu.

V kooperativních hrách již byla možnost vyjednávat připuštěna. Předpokládáme však stále, že hráči při dosažení dohody o kooperativním profilu svůj slíbený tah musí splnit. Kooperativní hru jsme chápali stále jako jednotahovou a musela být tehdy zavedena důvěra v plnění závazků.

Nekooperativní profil je v reálných situacích mnohdy neefektivní v porovnání s kooperativním profilem, ale individuální racionalita prostě koná svoje dílo. Připomeňme například slavné Vězňovo dilema a jeho aplikace v reálném životě (zbrojení, zbytečné investice do reklamy, mezinárodní dohody a podobně). V této kapitole si ostatně Vězňovo dilema opět vypůjčíme pro demonstraci následující myšlenky.

> Pokud se strategická interaktivní situace opakuje, začnou si hráči uvědomovat, že je lepší kooperovat – tj. oželet možnost krátkodobého zisku a doufat v dlouhodobý profit.

Toto poznání hezky dokumentuje veselá příhoda ze života Hodžy Nasredina[1].

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.

[1]Leonid Solovjov: Nasredinova dobrodružství, Nakladatelství Svoboda, Praha, 1976

Není proto divu, že nám opakované hry umožňují modelovat sociální normy, vliv "trestu" na další opakování hry a přirozeně vnucenou spolupráci. *Uvidíme, že kooperace v opakované hře je výsledek racionální úvahy hráče.* Ve hrách prostě neočekáváme filantropické chování[2], ale sobeckost, kterou jsme si již dříve vysvětlili jako ryzí individuální racionalitu.

Pro herní teoretiky je pochopitelně výzvou matematicky modelovat tento fenomén a následně odvodit konkrétní podmínky, za kterých ryzí individuální racionalita jedince vede k jeho kooperativnímu jednání.

Takže nyní předpokládejme strategickou interaktivní situaci – budeme ji nazývat *bázová hra* (angl. stage game). Hráči zvolí své tahy a obdrží výsledek hry, který si samozřejmě uvědomují a, jak budeme dále předpokládat, si výsledek hlavně zapamatují. Pokud nejsou spokojeni s výsledkem a hra se má opakovat, tuto svou zkušenost promítnou do svého dalšího rozhodnutí. Nabízí se myšlenka, že takto mohou konvergovat do stabilního bodu, který my nazíváme Nashovým ekvilibriem. Pak se nabízí otázka, proč zde mluvíme o kooperativním profilu, který má vyplynout z opakování a nemluvíme o Nashově ekvilibriu, do kterého také hráči mají dokonvergovat buď opakováním hry nebo ve vlastních úvahách.

Zásadní rozdíl mezi opakováním vynucenou kooperací a Nashovým ekvilibriem je poznání faktu ukončení opakování, lépe řečeno počtu opakování hry. Uvidíme, že je značný rozdíl mezi hrou s konečným počtem opakování a hrou s nekonečným počtem opakování. Co může znamenat nekonečný počet opakování hry v našem rozhodně konečném životě? Za nekonečnou opakovanost budeme považovat stav, kdy hráči prostě hrají navždy a konec nevidí.

Představme si populární turistickou destinaci s velkou fluktuací turistů. V této destinaci předpokládejme restauraci, kde majitel předpokládá každý den naprosto čerstvé hosty s absolutně nulovou informací o kvalitě jeho podniku. Pak je nanejvýš jisté, že majitel se bude snažit hodně "nekooperovat" a nezáleží mu na výsledném názoru hostů. Předpokládejme ovšem turistu, který mu hned na začátku oznámí, že v místě setrvá měsíc a každý den plánuje někam zajít na oběd a večeři. K tomuto hostu pojme majitel restaurace úplně jinou strategii, která může mít charakter kooperativity. Host totiž svým oznámením vysloví hrozbu, že už jeho podnik nenavští, pokud nebude se službami spokojen. Jeho hrozba je samozřejmě nedůvěryhodná, pokud je v lokalitě pouze jedna restaurace a míra důvěryhodností roste se stavem služeb v okolních podnicích. Nikoho by ovšem nemělo překvapit, že si ušlý zisk vynahradí na hostově poslední večeři před hostovým odjezdem z dovolené.

Proto je pro nás nekooperativní hra de fakto opakovaná hra s jedním opakováním. Její ekvilibrium známe z předešlých kapitol. Pokud se hra zopakuje dvakrát nebo třikrát, je to stále stejná nekooperativní situace. Rozhodně bude hra nekooperativní ve svém posledním opakování. My v této kapitole

[2]Wikipedia: Filantropie (z řec. filein, milovat a anthrópos, člověk - láska k člověku) znamená humanisticky motivovanou dobročinnost, dávání peněz, zboží, času nebo úsilí pro podporu obecně prospěšného účelu, zpravidla v delším časovém horizontu a s jasně definovanými cíli. V obecnější poloze lze filantropii pojmout jako jakýkoli altruistický počin, který směřuje k podpoře dobra nebo zlepšování kvality života. Lidé, kteří jsou známi pro své filantropické počiny, se někdy nazývají filantropové.

uvidíme vliv nekonečného opakování. Aby ovšem toto mělo nějaký smysl, musí si hráči pamatovat předchozí výsledky – toto bude druhý parametr opakování.

1.2 Opakované vězňovo dilema

Pro účely této kapitoly budeme dilema vězňů Johna a Petera prezentovat odlišným způsobem (obrázek 1.1), který je ovšem strategicky ekvivalentní s předchozím zápisem užitkové matice. Strategií C budeme označovat kooperativní strategii (nevypovídat) a strategií D nekooperativitu (zradu spoluvězně).

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

Figure 1.1: Bázová hra Vězňova dilematu

Opakovanou hru lze vyjádřit jako nadhru bázové hry. Má proto také svoje strategie, něco jako dlouhodobý plán chování. 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 $(..., 2, 0, 0, ...)$. Stejně jako předpokládáme úplnost informace v bázové hře, předpokládáme i úplnost informace v opakované hře – hráči proto navzájem znají své dlouhodobé koncepce (strategie nadhry).

Další jiné příklady strategií v opakované hře:

- Vždy budu hrát strategii x.
- 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 odpuštění).

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

1.3 Preference a očekávaný užitek v opakované hře

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

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

Pozor, δ_i^t je $t-$mocnina δ_i. Interpretaci ohodnocovacího faktoru můžeme chápat jako hledání balance mezi dnešním příjmem a zítřejším. Při $\delta = 0.5$ řekneme, že hráč vnímá dnešních 100 peněz stejnou hodnotou jako zítřejších 200 peněz.

Co znamená faktor δ ve svých extrémních stavech?

- 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).
- Pokud je δ_i blízko 1, hráč i vnímá budoucí zisky stejně intenzivně jako ty dnešní. Pokud by byl faktor $\delta_i = 1$, pak by byl jeho zisk v nekonečných hrách nekonečný (nekonvergoval by) a proto tuto hodnotu vyjímáme z oboru hodnot faktoru.

Notace: Budeme nadále klást $\delta_i = \delta; \forall i \in Q$.

Nyní 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 posloupností W a posloupností zisků $(c, c, ...)$?

Suma nekonečné posloupnosti $(c, c, ...)$ je $\frac{c}{1-\delta}$ (suma geometrické řady). Hráč je indiferentní mezi W a $(c, c, ...)$, pokud $c = (1-\delta)V$. Proto nazýváme hodnotu $\pi_i = (1-\delta)V$ jako průměrný zisk hráče, tzn.

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

1.4 Definice opakované hry

Definice 1. *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* $(1-\delta)\sum_{t=1}^{T} \delta^{t-1} U_i(s^t)$.

Nekonečně opakované Vězňovo dilema

Vrátíme se k Vězňově dilematu na obrázku 1.1 a na něm ukážeme vliv opakování na chování hráčů.

Pokud se Vězňovo dilema hraje se T-krát, kde T je:

- $T = 1$, pak je ve hře jasné nekooperativní ekvilibrium $NE = (D, D)$.

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

Jak se přístup ke hře změní, pokud se situace opakuje ∞-krát?

Proposition 1. *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 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.

□

Při $\delta \geq \frac{1}{2}$ tedy racionální hráč pochopí, že je racionální hrát kooperativní strategii. Nikdo ho k tomu nemusí nutit, hráč vnímá ztrátu při trestání protihráčem a je citlivý k této budoucí ztrátě – jeho citlivost k budoucnosti je dostatečně velká, aby ji dokázal ve své úvaze zaznamenat.

1.5 Folk theorems

Folk theorémy[3] jsou skupina vět (solution concepty) v opakovaných hrách. V opakovaných hrách ukazují, že kterýkoliv výsledek hry je dosažitelné řešení, jestli je pro hráče alespoň tak dobrý jako jejich minimax zisk. Také můžeme říct, že Folk theorémy ukazují, za jakých okolností jsou hráči schopni dosáhnout společně zvoleného profilu.

[3] R. B. Myerson, jeden z nositelů Nobelovy ceny (za Mechanism design) preferuje nazývat Folk theorémy jako General feasibility theorems. Název Folk theorémy stále ovšem přetrvává.

Definujeme nyní minimax zisk hráče. Ten představuje přijatelné dno zisků ve hře a tedy jakousi sociální normu.

Definice 2. *Mějme hru $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$. Minimax zisk hráče $i \in Q$ ve hře Γ je:*

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

Hodnota $\underline{v_i}$ je tedy nejnižší dosažitelné BR_i hráče (výraz v závorce je přece best-response na s_{-i} hráče i). Předpokládáme, že hráči jsou schopni si kolektivně uvědomit míru přijatelného výsledku hry (feasible outcome).

Definice 3. *Řekneme, že výplatní vektor $v \in \mathbb{R}^N$ je striktně individuálně racionální, pokud platí pro všechny hráče $i \in Q : v_i > \underline{v_i}$.*

Výplatní vektor v je užitkový vektor v nějakém profilu $s \in S$, kde všichni hráči získávají více, než jsou jejich globální minima ve hře. Takový profil se může stát předmětem dohody mezi hráči a být tím kooperativním profilem, který se hráči dohodnou společně hrát. Jaké jsou však ty trestací strategie hráčů?

Definice 4. *Minimax strategií proti hráči i nazveme sub-profil m^i_{-i} (ve dvouhráčové hře strategii protivníka), která hráči i přinese $\underline{v_i}$. Je to:*

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

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

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 – tedy pouhou racionalitou bez "silových zásahů" do konání hráčů.

Věta 2 (Folk 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.*

Uvědomme si nejdřív sílu tohoto výroku: absolutně všechny profily hry jsou pro hráče dosažitelné (jsou schopni se na nich dohodnout a pak je vymáhat). Ještě lépe: pro všechny profily ve hře existuje $\delta < 1$ takové, že jsou tyto profily vymahatelné.

Proof. 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 $\underline{v_i}$ hráči i, který vybočil).

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:

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

Nekonečnou posloupnost počínající kolem $t+1$ nahradíme jejím součtem. Ve výrazu nám zůstávají tři hlavní prvky zisku hráče i, který v kole t zradí: 1) zisk za období do kola t, kdy hráči kooperovali, 2) zisk hráče v kole, kdy zradil a 3) zisky hráče, který je až do konce věčnosti trestán.

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

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

Druhý člen:

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

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, že v_i je slabě lepší než průměrný zisk ve hře, kdy v kole t hráč zradil:

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

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

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

Protože $\delta \in \langle 0, 1 \rangle$, pak toto bude platit, pokud:

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

Přelomová hodnota, při které jsou obě strany nerovnice stejné, je:

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

A zbývá nám ověřit, že $\delta < 1$ a tudíž odpovídá definici ohodnocovacího faktoru:

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

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.

□

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).

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

Pro ohodnocovací faktory $\delta > \frac{1}{7}$ je jisté, že hráči budou racionálně nuceni hrát profil (U, L). Hráčům zbývá přesvědčit své protihráče, že jsou této citlivosti k budouctnosti schopni.

1.6 Jak lidé hrají v opakovaných hrách

Podle laboratorních experimentů[4] lidé hrají v opakovaných vězňových dilematech kooperativněji, než by se očekávalo. Dokonce i u konečně opakovaných hrách. Jak si to vysvětlit, když teorie předpokládá u konečně opakovaných her nekooperativní jednání? Víme, že 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.

[4] Čerpáno z přednášek Asu Ozdaglara

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

Jaký je tedy závěr ze studia opakovaných her? Je dobré ctít altruismus a férovost, ale jediný důvod, proč dosahujeme stability v kooperativním profilu je trestání protihráče, který uhne z kooperativního profilu.