Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Vojtěch Luhan

Opakované hry s neúplnou informací

Katedra matematické analýzy

Vedoucí bakalářské práce: Doc. RNDr. Stanislav Hencl, Ph.D.
Studijní program: Matematika
Studijní obor: Obecná matematika

2010

Na tomto místě bych rád poděkoval všem, kdož mi jakýmkoli způsobem usnadnili vytváření této práce, zejména pak rodině a respondentům prezentovaného průzkumu.
Největší poděkování však chci za příjemnou spolupráci věnovat vedoucímu práce Stanislavu Henclovi, který se s velkým nasazením (obzvláště v posledních obdobích časového limitu) věnoval mé bakalářské práci a ochotně poskytoval dlouhé online konzultace v době, kdy setkání z důvodu zahraniční cesty nemohlo být uskutečněno osobně.

Prohlašuji, že jsem svou bakalářskou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne Vojtěch Luhan

Obsah

1 Úvod do tématu **5**

1.1 Motivační příklad 5
1.2 Popis hry 6
1.3 Základní rozdělení 7
1.4 Příklady 8

2 Vězňovo dilema a malý průzkum lidského chování **12**

2.1 Vězňovo dilema 12
2.2 Jednokolová hra 13
2.3 Vícekolová hra 13
2.4 Experiment 14

3 Opakované hry dvou hráčů s nulovým součtem **18**

3.1 Definice základních pojmů 18
3.2 Popis hry 19
3.3 Určování optimální strategie 20
3.4 Návrat k úvodním příkladům 29

Literatura **33**

Název práce: Opakované hry s neúplnou informací
Autor: Vojtěch Luhan
Katedra (ústav): Katedra matematické analýzy
Vedoucí bakalářské práce: Doc. RNDr. Stanislav Hencl, Ph.D.
e-mail vedoucího: hencl@karlin.mff.cuni.cz

Abstrakt: Maticovou hrou pro 2 hráče rozumíme souboj dvou účastníků, ve kterém jeden vybírá z dané matice řádek a druhý sloupec, přičemž se oba snaží, aby dle svých priorit dohromady vybrali pro ně co možná nejvýhodnější prvek. V této bakalářské práci uvažujeme převážně situaci, kdy jeden z hráčů má určitou informaci o dané matici, zatímco druhý nikoliv. Cílem práce pak je popsat strategie jednotlivých hráčů a najít mezi nimi rovnovážný stav. Pro srozumitelnost co možná největší skupině čtenářů je text prokládán příklady ze života (zejména Kapitola 1) a prezentuje výsledky autorova malého průzkumu chování lidí v případě speciální maticové hry – tzv. Vězňově dilematu (Kapitola 2), stěžejní částí jsou ovšem hlubší výsledky z dané problematiky (Kapitola 3).

Klíčová slova: teorie her, opakovaná hra, hra s nulovým součtem, minimax, optimální strategie, Vězňovo dilema

Title: Repeated games of incomplete information
Author: Vojtěch Luhan
Department: Department of mathematical analysis
Supervisor: Doc. RNDr. Stanislav Hencl, Ph.D.
Supervisor's e-mail address: hencl@karlin.mff.cuni.cz

Abstract: A matrix game for two players is a match between two participants, one choosing the row and the second one the column of a stated matrix. Each of them thinks with regard to the selection of the most valuable element according to his own preferences. In this bachelor's work, we mostly consider the situation in which one player has a certain information whereas the second one does not. The aim of the work is to describe the strategies of individual players and to find the balanced state between them. To make the text comprehensible to as big number of readers as possible, there are real-life examples, mostly in Chapter 1, and a presentation of author's survey results on the human behaviour in Chapter 2. Yet the principal part is Chapter 3, where the nontrivial results are introduced.

Keywords: game theory, repeated game, zero-sum game, minimax, optimal strategy, Prisoner's dilemma

Kapitola 1

Úvod do tématu

1.1 Motivační příklad

Příklad 1.1.1 Středoškoláci Alois a Bonifác spolu seděli v lavici zrovna, když se psal test z matematiky. Ani jeden z nich ale látku pořádně neuměl, a tak od sebe opisovali. Při opravování ale bylo zkušenému učiteli vše jasné. Zavolal si oba hříšníky, každého do jiného pokoje, a jednoho po druhém se zeptal, kdo opisoval. Pokud se shodnou na jednom, zůstane nevinnému z nich 1 z chování, zatímco viník bude okamžitě vyloučen ze školy (řekněme 4 z chování – ač se neudílí, dobře charakterizuje fakt, že je tento čin horší než dostat z chování 3). Pokud oba řeknou, že opisoval ten druhý, dostanou z chování oba 3, a přiznají-li se oba dva, bude to pro ně polehčující okolnost a oběma bude udělena 2 z chování. Tedy možnosti lze zapsat do matice o rozměrech 2x2, kde v každém prvku bude po řadě známka z chování pro Aloise a známka pro Bonifáce.

	Alois udal sebe	Alois udal Bonifáce
Bonifác udal sebe	2,2	1,4
Bonifác udal Aloise	4,1	3,3

Každý z výtečníků se primárně snaží, aby jako známku dostal co nejnižší hodnotu, podle toho také vybírá – Alois sloupec a Bonifác řádek. Alois přemýšlí, co může říci Bonifác. Pokud udal Aloise, vyplatí se mu, aby na oplátku udal Bonifáce (3 z chování je lepší než vyloučení). Pokud však udal sám sebe, pak je pro Aloise opět výhodnější udat Bonifáce (tímpádem může vyjít s čistým štítem, a to je lepší než 2). Zcela analogicky však uvažuje i Bonifác, a tak se navzájem udají. Oba tedy dostanou 3 z chování, přitom ideálním rozhodnutím by zřejmě bylo, kdyby se oba navzájem kryli, udali sebe a dostali z chování 2.

Hra se nyní dostala do tzv. Nashovy rovnováhy, kdy žádný hráč nemůže jednostrannou změnou svého rozhodnutí zlepšit svoji situaci. Rovnováha sice představuje optimální řešení pro každého z hráčů, globálně optimálním řešením (tzv. Paretovým optimem) ale, jak jsme právě viděli, být nemusí.

Tento speciální případ tzv. vězňova dilematu je jedním z příkladů praktického využití teorie zkoumané v této bakalářské práci. Samotná práce se bude zabývat zejména teoretickou stránkou dané problematiky, a to v širším záběru, občas však bude poukázáno i na možné použití v reálném životě.

Tak tomu bude např. v Kapitole 2, ve které jsou prezentovány výsledky malého průzkumu mezi veřejností ohledně jejich rozhodování v opakovaném vězňově dilematu. Jak ostatně název napovídá, objektem zájmu této práce budou, narozdíl od právě uvedeného příkladu, zejména hry opakované.

1.2 Popis hry

Velké množství racionálních rozhodování v běžném životě je možné přirovnat k hraní nějakého speciálního případu námi zkoumané hry, stejně jako by se pro mnohé případy hry daly vymyslet problémy, které by právě ony řešily. Uveďme si proto nyní naši hru v plné obecnosti, zatím alespoň neformálně (přesný matematický model, byť jen pro 2 hráče, je popsán v Kapitole 3).

Popisovaná hra v plné obecnosti umožňuje účast libovolného (konečného) počtu n hráčů. Každý z nich si nejprve sedne na jednu z připravených židlí $1, ..., n$ za stůl, nad nímž se vznáší jedna či více n-dimenzionálních matic (pro 2 hráče mají tvar obdélníku, pro 3 kvádru atd. – všechny ale mají stejné rozměry). Každá z nich má jakožto prvky vektory o délce n, kde v i-té složce nalezne hráč sedící na i-té židli počet korun (komu se nezamlouvá hazardní podtext hry, může místo peněz uvažovat třeba lentilky), které získá, bude-li vybrán daný prvek matice. Černá skříňka, umístěná uprostřed stolu, náhodně zvolí jednu z matic, se kterou nyní budou hráči *všechna kola* hrát. Podle předem určeného algoritmu prozradí některým hráčům, které matice určeny nebyly, tedy zúží množinu matic, o kterých daný hráč uvažuje, zpravidla jim však přímo neříká, kterou konkrétně vybral. Nyní je na hráčích, aby určili prvek matice, podle kterého budou všichni v tomto kole ohodnoceni. Vesměs bez přesné znalosti, se kterou maticí se hraje, určí v každém kole i-tý hráč hodnotu i-té souřadnice, tedy po výběru všech n hráčů bude vybráno všech n prvků souřadnic a vektor ohodnocení hráčů bude tedy jednoznačně určen. Vycházíme přitom z toho, že se všichni hráči chovají maximálně sobecky a vše, co dělají, dělají pouze ve snaze získat co největší výhru v průběhu hry. Nyní již černá skříňka zná zisky z tohoto kola, nicméně *nikomu je neříká* (pouze si je zapamatuje), a tím kolo skončí. Přejde se ke kolu dalšímu, ve kterém matice zůstává nezměněna (ač hráčům vesměs není známa její podoba) a hráči, zůstávající na svých židlích, znovu vybírají hodnoty své souřadnice, přičemž ta může být odlišná od té, kterou zvolili v kole předchozím. Na konci kola se opět vyhodnotí vybraný prvek a takto pokračujeme až do dohodnutého počtu kol (obecně může být i neznámý), po němž skříňka hráčům oznámí jejich celkové zisky (či ztráty), načež si hráči rozdělí výhry, hra skončí a účastníci pokračují v hraní od začátku. Po odehrání určeného počtu her černá skříňka vyhlásí celkovou úspěšnost jednotlivých hráčů, načež ti vezmou své dosud získané výhry a rozejdou se.

1.3 Základní rozdělení

Naši hru můžeme rozlišovat dle několika kritérií:

1. *Podle počtu hráčů*
    - hry se dvěma hráči
    - hry s více hráči
2. *Podle zdroje peněz*
    - hry s nulovým součtem (při vybrání jakéhokoli prvku je součet zisků všech hráčů vždy 0 – hráči tedy své zisky získávají na úkor ostatních) – např. poker
    - hry s nenulovým součtem (někdo nehrající 'přispěl do banku') – např. Příklad 1.1.1
3. *Podle počáteční informovanosti*
    - hry s úplnou informací (všichni hráči vědí, s jakou maticí se hraje)
    - hry s neúplnou informací
        - hry s úplnou informací u jednoho či více hráčů
        - hry s neúplnou informací na všech stranách

V případě her s neúplnou informací lze hry dělit také na ty, ve kterých černá skříňka na začátku některým hráčům (pravdivě) řekla, které matice jistě nevybrala, a ty, ve kterých tak neučinila. Další možné skupiny her jsou ty, ve kterých je hráčům povoleno mezi sebe vyslat signály s nápovědou (ty již pravdivé být nemusí), a ty, ve kterých tak činit nesmí.

Obecný případ uvedený v sekci 1.2 vypadá bez zavedeného aparátu značení poněkud těžce prostupný, uveďme proto pár jednoduchých příkladů z [1], jak mohou jednotlivé hry vypadat a co především na nich budeme studovat.

1.4 Příklady

Příklad 1.4.1 Dva hráči (nazvěme je Hráč I a Hráč II) hrají opakovaně hru s nulovým součtem danou maticemi

$$G_1^1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad G_2^1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Jednotlivé prvky značí hodnoty, které v případě zvolení prvku dá Hráč II Hráči I. Tedy Hráč I se snaží o zvolení co možná největších hodnot, jež by dostal od Hráče II, který se naopak snaží jeho zisky co možná nejvíc zmenšit. Hodnotu x v prvcích matic G_1^1 a G_2^1 tedy budeme chápat jako výplatní vektor $(x, -x)$, což budeme pro zkrácení zápisu mlčky používat i v následujícím textu.

Černá skříňka (mluvme o ní prostě jako o Rozhodčím) na začátku každé hry *pevně a pro celou hru neměnně* zvolí, se kterou maticí se bude hrát, a Hráči I prozradí, která to je. Druhý hráč tuto informaci nemá (ví ale, že jí disponuje soupeř), považuje tedy zvolení každé z matic za stejně pravděpodobné. Následně v každém kole současně a veřejně Hráč I zvolí řádek a Hráč II sloupec, ve kterých se bude hrát. Rozhodčí zaznamená jejich zisky a, *aniž by je zveřejnil*, přistoupí k dalšímu kolu. Hráč I tedy na základě znalosti matice i tahů obou hráčů (uvažujeme tzv. zcela pozorovatelnou hru) ví, kolik na konci dostane, tento fakt však není znám jeho soupeři. Každá hra sestává z n kol (předpokládáme n velké), po nichž Rozhodčí vyhlásí výsledky a Hráč II předá výhru Hráči I. Aby byla hra nezávislá na počtu kol, bude nás obvykle zajímat zejména průměrný zisk hráčů v každém kole, tedy jejich celková výhra vydělená počtem kol.

Hráč I tedy má výhodu, že zná matici – jak tohoto faktu ale maximálně využít?

Bude-li hrát *zcela odhalující strategii* a druhý hráč to zjistí, bude jeho průměrný zisk velmi nízký, resp. a/n, kde a je počet kol, po kolika tuto strategii Hráč II prokoukne. Vzhledem k tomu, že nás zajímají hlavně velmi dlouhé hry, je v tomto případě zisk $\lim_{n\to\infty} a/n = 0$ ($a \in \mathbf{N}$), tedy asymptoticky nulový.

Pokud se Hráč I rozhodne 'blafovat' (např. tvrdit, že hraje zcela odhalující strategii, zatímco ve skutečnosti byla vybrána matice opačná), může si pro právě probíhající hru zajistit až zisk blížící se 1, do dalších her však ztratí důvěru soupeře, který mu nadále může a nemusí věřit, což vyústí v náhodné vybírání sloupců ze strany Hráče II, což je to samé, jako kdyby náhodně vybíral řádky už Hráč I, a tyto varianty tedy můžeme sloučit.

Další možností je tedy hrát *nic neodhalující strategii*, to znamená nedat najevo svou soukromou informaci a hrát tak, jako kdyby Hráč I nevěděl, která matice byla zvolena (hrát zcela náhodně si můžeme představit třeba tak, že si Hráč I před každým tahem hodí spravedlivou mincí a vybere řádek podle toho, která strana padla). Tuto hru tedy můžeme považovat za ekvivalentní hraní hry s maticí

$$G_3^1 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix},$$

která informovanému hráči přináší průměrný zisk v hodnotě 1/4 bez ohledu na to, která matice byla na začátku vybrána.

Tedy výhodnější pro Hráče I paradoxně je vůbec neodhalit svou znalost matice (hrát nic neodhalující strategii), což mu při velmi dlouhých hrách přinese průměrný zisk 1/4 za kolo oproti 0 při strategii zcela odhalující.

Poznámka 1.4.2 Rád bych na tomto místě upozornil, abychom nezaměňovali termíny *využít* a *odhalit*. Zatímco využít svou informaci znamená zvolit svou strategii podle toho, která matice byla vybrána, odhalit svou znalost matice značí, že hráčova strategie je jiná, než kdyby počáteční informací nedisponoval. Tedy v právě rozebraném příkladě Hráč I svou strategii *využil* k tomu, aby zvolil nejvhodnější strategii, kterou bylo svou znalost matice *neodhalit* – tedy hrát stejně, jako kdyby žádnou výhodu neměl. Vzhledem k tomu, že nyní uvažujeme hru, ve které hráči navzájem sledují své tahy (tzv. *zcela pozorovatelná hra*), je každý tah informovaného hráče nějakým signálem pro soupeře – tedy i tím, že svou informaci neuplatní, může dát jisté znamení druhému hráči o tom, která matice byla vybrána.

Příklad 1.4.3 Uvažujme nyní zcela stejnou hru, pouze s rozdílem v maticích, které mají tentokrát tvar

$$G_1^2 = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad G_2^2 = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

Opět je to Hráč I, který ví, se kterou maticí se hraje. Jenže nyní je to on, jenž se snaží minimalizovat zisky Hráče II.

Logickým řešením je hraní *zcela odhalující strategie*, při které bude hrát stále spodní řádek v případě G_1^2 a stále horní řádek u G_2^2. Ať už Hráč II zvolí jakýkoli sloupec, vždy mu bude vyplaceno 0 korun a Hráč I docílil svého nejlepšího výsledku.

Pro srovnání, *neodhalující strategie* vede ke hře s maticí

$$G_3^2 = \begin{pmatrix} -1/2 & 0 \\ 0 & -1/2 \end{pmatrix},$$

která Hráči I přináší ztrátu průměrně 1/4 na kolo.

Toto je tedy příklad hry, ve které je nejvýhodnější hrát zcela ohalující strategii.

Z uvedených příkladů se zdá, že nejvýhodnější strategií je buď svou znalost matice zcela odhalit, nebo naopak neodhalovat vůbec, existují ale i příklady, kdy se nejlepší řešení nachází 'někde mezi' – tzv. *částečně odhalující (smíšené) strategie*:

Příklad 1.4.4 Uvažujme nyní, že náhoda v každé hře pro naše 2 hráče vybírá jednu z těchto matic:

$$G_1^3 = \begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix}, \qquad G_2^3 = \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}.$$

Hráč I je opět seznámen s výběrem, Hráč II nikoliv.

Hraní zcela odhalující strategie, tedy volení dominantního řádku (horní u G_1^3 a spodní u G_2^3) vede u velmi dlouhých her k průměrnému zisku téměř 0, neboť Hráč II po odhalení strategie (na konci každé hry jsou zveřejněny výsledky) bude volit prostřední sloupec u G_1^3 a levý v případě G_2^3.

Hraní nic neodhalující strategie prozměnu směřuje k hraní s maticí

$$G_3^3 = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix},$$

též s průměrným ziskem téměř 0.

Existuje ovšem i částečně odhalující strategie, která vede nejen k nenulovým ziskům, ale zaručuje při mnohokrát opakovaných velmi dlouhých hrách Hráči I dokonce zisky v hodnotě alespoň 1:

Nechme Hráče I vyrobit pravidelný čtyřstěn a obarvit jednu z jeho stěn. Hráč I následně tímto tetraedrem na začátku každé hry hodí a podívá se, která stěna leží vespodu. Je-li to ta obarvená, pak bude *ve všech kolech hry* volit spodní řádek při zvolení G_1^3, nebo horní řádek v případě G_2^3. Stojí-li tetraedr na neobarvené stěně (to nastane s pravděpodobností 3/4), bude Hráč I volit přesně naopak, tedy ve všech kolech horní řádek u G_1^3 a spodní u G_2^3. Ať už tedy byla vybrána jakákoliv matice, informovaný hráč s pravděpodobností 3/4 volí po celou hru svůj dominantní řádek, zatímco 'méně výhodnému' řádku připadá pravděpodobnost 1/4. Použitím Bayesovy věty (viz [7]) zjistíme i opačný směr, totiž že pokud Hráč I zahrál horní řádek, potom hrajeme s pravděpodobností 3/4 s maticí G_1^3 a symetrickým použitím stejné věty dostaneme obdobné tvrzení pro dolní řádek a G_2^3.

Hráč I soupeři svou strategii klidně může prozradit, pouze mu neukazuje, co padlo na čtyřstěnu. V 1. kole pak oba zjistí, zda bude Hráčem I zahráván po celou hru spodní nebo horní řádek.

Řekněme, že bude zahrávat řádek horní. Pak se matice G_1^3 redukuje na horní řádek, tedy matici rozměrů 1×3 tvaru $(\,4 \quad 0 \quad 2\,)$ a G_2^3 na $(\,0 \quad 4 \quad -2\,)$, přičemž G_1^3 byla na začátku vybrána s pravděpodobností 3/4 a s pravděpodobností 1/4 to byla G_2^3. Tedy při mnohokrát opakovaném hraní této hry o velmi velkém počtu kol se hra převádí na výběr sloupce Hráčem II z matice

$$3/4\,(\,4 \quad 0 \quad 2\,) + 1/4\,(\,0 \quad 4 \quad -2\,) = (\,3 \quad 1 \quad 1\,),$$

která Hráči I dává v každém případě zisk blížící se alespoň 1.

V případě, že je v prvním kole zahrán spodní řádek, dostaneme se symetrickou úvahou k převodu na matici

$$1/4\,(\,4 \quad 0 \quad -2\,) + 3/4\,(\,0 \quad 4 \quad 2\,) = (\,1 \quad 3 \quad 1\,),$$

která opět zaručuje zisk v limitní hodnotě minimálně 1.

Kapitola 2

Vězňovo dilema a malý průzkum lidského chování

2.1 Vězňovo dilema

Vězňovo dilema je matematická úloha, která zkoumá vůli spolupracovat s kolegou bez toho, abychom věděli, jak bude kooperovat on. Jelikož pro tento problém nemůže být (alespoň za současných podmínek) nalezena optimální strategie, která by zaručovala nejvyšší zisky v reálném životě, je tato problematika zkoumána nejen matematiky, ale také psychology, programátory (snaha vyvinout nejúspěšnější algoritmus), biology (vztahy v přírodě, vývoj druhů, úspěšnost civilizací s jednotlivými měrami spolupráce), politiky (kooperace mezi státy, např. ve zbrojení), ekonomy, sportovci (podvádění, např. doping) a dalšími.

Problém vězňova dilematu začali jako první zkoumat Merrill Flood a Melvin Dresher v polovině 20. století, v 90. letech pak Albert W. Tucker problém formalizoval takto:

Policie zadržela 2 podezřelé ze spáchání vážného trestného činu. Nemá proti nim však dostatek důkazů – k usvědčení potřebuje vždy svědectví druhého z nich. Oba podezřelé zadržuje v oddělených celách bez možnosti vzájemné komunikace. Za každým zvlášť pak vyšetřovatel přijde a řekne mu, že musí oznámit viníka incidentu – buď může udat kolegu, nebo zůstat mlčet (žádná jiná alternativa se nepřipouští). V případě, že jeden udá druhého a ten přitom mlčí, bude mlčící usvědčen, půjde na 10 let do vězení a druhý podezřelý bude propuštěn. Pokud se udají navzájem, trest si 'rozdělí' – každý půjde *sedět na 5 let. V případě, že oba budou mlčet, soud nenashromáždí dostatek důkazů a oba pachatele se podaří odsoudit jen na 1 rok za drobnější delikty. Předpokládáme přitom, že oba podezřelí se snaží pouze minimalizovat svůj trest, bez ohledu na to, jak dopadne druhý. Jak je nejvýhodnější se rozhodnout?* (doslovně přeloženo z [4])

2.2 Jednokolová hra

V případě, že tento problém řešíme pouze 'jednokolově', tedy obdobně jako bylo uvedeno v trochu jiném kontextu v Příkladu 1.1.1, je nasnadě, kterak by si odsouzení vybrali. Jelikož je pro ně v každém případě (tedy ať už kolega bude udávat nebo mlčet) výhodnější udat toho druhého, skončí to tím, že budou oba odsouzeni na 5 let, ač to zjevně není globálně nejvýhodnější řešení.

2.3 Vícekolová hra

Zajímavější situace však nastane, pokud hrajeme tuto hru vícekolově, tedy pokud je zde možnost 'oplatit' zradu z předchozího kola. Z racionálního pohledu je i zde jasné řešení: udat kolegu v každém kole. Totiž:

V posledním kole se kolegu vyplatí udat, protože to už nám nemůže vrátit (viz minulý odstavec). Podobně uvažuje jistě i on, tedy poslední kolo skončí zradou z obou stran. Je tedy výhodné zradit už v předposledním kole, neboť v posledním se tak jako tak oba udají. Takto uvažují oba a postupují až k prvnímu kolu, tedy se udají vždy.

Výzkumy (zejména Roberta Axelroda z 80. let 20. století [3]) prováděné soupeřením jednotlivých algoritmů však zjistily, že nejvyšších zisků (tzn. nejkratší úhrnné doby strávené ve vězení), dosahovaly programy, které splňovaly 4 podmínky, které byly v rozporu s právě uvedeným racionálním chováním:

- byly hodné (nebyly první, kdo zradil)
- oplácely zradu (v příštím kole s velkou pravděpodobností zradily také)
- dokázaly zapomenout (to zabraňuje, aby se série udávání zacyklila)
- nezáviděly (nesnažily se získat víc než soupeř)

V experimentech s lidským chováním bylo dosahováno výsledků ještě překvapivějších a různorodějších, rozhodl jsem se proto, že drobný průzkum chování v rámci bakalářské práce jako zajímavost provedu i já.

2.4 Experiment

Vyrobil jsem jednoduchý program (viz příloha), který se 30-krát zeptal, jestli chce uživatel se spoluviníkem spolupracovat, nebo ho udat, přičemž od druhého kola i připojoval, jak se v minulém kole rozhodl program a vypsal tresty pro oba podezřelé. Tresty jsem použil stejné jako v popisu dilematu v sekci 2.1, pouze jsem vyměnil roky za měsíce, aby nebyla většina uživatelů v úhrnu odsouzena v podstatě na doživotí. Matice trestů (kde v první složce prvků je trest pro uživatele a ve druhé pro program) tedy vypadala následovně:

	Program spolupracoval	Program nespolupracoval
Uživatel spolupracoval	1,1	10,0
Uživatel nespolupracoval	0,10	5,5.

Algoritmus programu byl prostý:

- v prvním kole program na 90% spolupracoval,
- pokud v předchozím kole uživatel nezradil, pak na 90% spolupracoval,
- pokud v předchozím kole uživatel zradil, potom v tomto kole na 50% zradil prozměnu program.

Průzkumu se zúčastnilo celkem 120 lidí, kteří odehráli dohromady 155 her. Jejich průměrné výsledky (průměr míněn aritmetický) v závislosti na určitých faktorech byly následující (vzhledem k tomu, že byli respondenti v duchu hry žádáni pouze o minimalizování vlastních trestů, jsou tresty programu v závorkách uvedeny spíše informativně):

	Počet	**Průměr**	
Všechny hry	155	67,61	(121,99)
První hry respondentů	120	68,25	(123,42)
Další hry respondentů	35	65,40	(117,11)
Uživatel zradil jako první	80	70,74	(161,86)
Program zradil jako první	61	61,69	(68,74)
Uživatel zradil v 1. kole	71	73,30	(167,66)

Většina respondentů hrála nepravidelně (nebo ze 30-ti kol nebylo možné odhalit jejich strategii), dvě pětiny hráčů ale využily jednu ze tří nejpoužívanějších strategií ve vězňově dilematu:

- *holubice* – hráč v každém kole spolupracuje bez ohledu na soupeře,
- *jestřáb* – opak holubice – hráč nikdy nespolupracuje,
- *tit for tat* – strategie 'oko za oko, zub za zub' – hráč začne spolupracovat, ale zradu vždy v příštím kole opětuje, poté se vrací ke spolupráci.

Poznámka 2.4.1 Holubice a tit for tat jsou strategie, při kterých hráč nesmí 'závidět' soupeři – až na povolenou zradu v posledním kole si totiž nikdy nemůže zajistit dřívější odchod z vězení. To je přesný opak jestřába, který naopak vždy zajistí větší tresty kumpánovi. Jak uvidíme v následujících statistikách, nemusí to však nic říkat o tom, kdo dostane menší tresty v porovnání mezi sebou.

Další strategie a jejich úspěšnost při vzájemném soupeření jsou poutavě popsány např. v [5]. Úspěšnost popsaných třech strategií (s případným udáním v posledním kole, které, jak jsme viděli v sekci 2.2, v hrách se známým počtem kol 'nemůže nic zkazit') byla v naprogramované hře následující:

	Počet	**Průměr**	
Holubice	24	53,08	(30,58)
Tit for tat	17	55,71	(57,47)
Jestřáb	21	73,81	(224,29)

Předpokládejme nyní, že jsou výsledky hráčů v jednotlivých kategoriích nezávislé, stejně rozdělené náhodné veličiny řídící se normálním rozdělením s příslušnou střední hodnotou a rozptylem, které jsou nezávislé i mezi kategoriemi navzájem.

Výsledky dvou strategií mezi sebou můžeme porovnávat tzv. *dvojvýběrovým t-testem* (jeho přesný algoritmus, stejně jako všechny potřebné definice a tvrzení, můžeme najít např. v [7]): na nějaké hladině $\alpha > 0$ můžeme potvrdit hypotézu, že je jedna strategie (jejíž výsledky tvoří náhodnou veličinu X) úspěšnější než druhá (náhodná veličina Y). Označme μ_1 střední hodnotu normálního rozdělení, kterým se řídí výsledky strategie, o níž chceme dokázat, že je lepší (aritmetický průměr této strategie $\overline{X}_{n_1}$ musí být vyšší než aritmetický průměr $\overline{Y}_{n_2}$), a μ_2 střední hodnotu pro druhou strategii. Můžeme potom vyslovit nulovou hypotézu $H_0 : \mu_1 = \mu_2$ a proti ní hypotézu alternativní $H_1 : \mu_1 > \mu_2$. V případě, že naměřená data padnou do *kritického oboru*, jsou v rozporu s hypotézou H_0, kterou proto zamítáme ve prospěch H_1, tedy dokážeme, že $\mu_1 > \mu_2$ na hladině α, tzn. první strategie je výhodnější s pravděpodobností alespoň $1 - \alpha$.

Říkáme přitom, že data padnou do kritického oboru, pokud

$$\frac{\overline{X}_{n_1} - \overline{Y}_{n_2}}{S^*_{X,Y}\sqrt{\frac{n_1+n_2}{n_1 n_2}}} \geq t_{1-\alpha;n_1+n_2-2},$$

kde n_1, resp. n_2 jsou počty měření u 1., resp. 2. strategie, $\alpha > 0$ pevně zvolená hladina testu, $t_{1-\alpha;n_1+n_2-2}$ kvantil Studentova t-rozdělení o n_1+n_2-2 stupních volnosti a

$${S^*_{X,Y}}^2 = \frac{1}{n_1 + n_2 - 2}\left((n_1 - 1)\, S_X^2 + (n_2 - 1)\, S_Y^2\right),$$

kde S_X^2, resp. S_Y^2 jsou výběrové rozptyly náhodných veličin X, resp. Y.

Takto jsem postupoval při ověřování některých hypotéz o lidských volbách strategií a jejich výhodnosti. Zde vybírám ty, které mi přišly pro naměřená data ze hry proti výše popsanému algoritmu nejzajímavější nebo nejpřekvapivější (hladinou testu v souladu s předchozím textem rozumíme maximální pravděpodobnost, že nesprávně zamítneme nulovou hypotézu – tedy že potvrdíme něco, co neplatí):

- Na hladině 0,1% hráči dosahují lepších výsledků, pokud jako první zradí počítač, ne oni.
- Na hladině 5% platí, že pokud hráč v prvním kole zradí, pak dosáhne horšího celkového výsledku.

- Na hladině 0,1% lze tvrdit, že hráči nechávají počítač odsedět alespoň o 3 roky delší trest (v průzkumu byli žádáni, aby dbali pouze vlastních zisků).
- Na hladině 5% nelze zamítnout hypotézu, že hráč, který celou dobu udával, odsedí stejnou dobu jako hráč preferující přátelštější algoritmus. Nelze však dokázat ani opak.
- Na hladině 1% lze tvrdit, že hrát tit for tat je ve srovnání se souhrnnými výsledky hráčů, kteří se rozhodovali podle jiných kritérií, výhodné.
- Na hladině 5% bylo potvrzeno, že stálou spoluprácí si hráč ušetří 8 měsíců vězení. Ze všech strategií má holubice vůbec nejnižší průměrný souhrnný trest, zdá se tedy být proti algoritmu z programu nejvýhodnější (sice být nemusí, ale proti algoritmu, který nemá tendence 'oškubat' příliš hodného protivníka s největší pravděpodobností i bude).
- Na hladině 5% lze ale potvrdit také hypotézu, že hraním holubice bude hráč pykat o 16 měsíců (tedy o více než polovinu trestu počítače) déle, než jeho soupeř.
- Na hladině 5% platí, že pokud hráč praktikuje strategii jestřába, pak si jeho oponent odsedí o celých 143 měsíců, tedy téměř 12 let, více.

Na závěr kapitoly bych pro zajímavost ještě rád uvedl, že nejlepším dosaženým výsledkem bylo (dvakrát) 30 měsíců (strategie holubice bez udání v posledním kole), nejhorším pak měsíců 114.

Kapitola 3

Opakované hry dvou hráčů s nulovým součtem

Tato kapitola kromě vlastní práce využívá informací prezentovaných zejména v [1], [2] a [4], zavádění jednotlivých pojmů potom bylo obvykle činěno s přihlédnutím k [7] a [8].

3.1 Definice základních pojmů

Ačkoli bychom na tomto místě mohli popsat hru v plné obecnosti, jako tomu bylo v sekci 1.2, zadefinujeme pro jednodušší orientaci pouze pojmy, které budeme v následujícím textu potřebovat, tedy zcela vynecháme zejména termíny *počáteční nápověda od Rozhodčího* a *signály mezi hráči.*

Základní pojmy pro opakované hry s nulovým součtem dvou hráčů s neúplnou informací na straně jednoho z nich (Hráče II) tedy definujeme takto:

- Přirozené číslo N – počet her,
- Přirozená čísla $n_1, \ldots, n_N$ – počet kol v jednotlivých hrách (budeme-li počítat s jedinou hrou, pak její délku budeme značit n),
- Konečná množina K – množina matic ke hře o rozměrech $k_1 \times k_2$,
- Konečná množina R – množina pravděpodobnostních rozdělení na K ($p \in R$ je pravděpodobnost zvolení jednotlivých prvků K),

- Konečné množiny T_i – množiny pravděpodobnostních rozdělení na $\{1, \ldots, k_i\}$, $i \in \{1, 2\}$ (tedy $t_i \in T_i$ jsou vektory délky k_i, jejichž prvky jsou nezáporné a jejich součet je vždy přesně 1, neboli $\sum_{j=1}^{k_i} t_i(j) = 1$),
- s_1^m – výběr prvku z množiny $\{1, \ldots, k_1\}$ podle pravděpodobnostního rozdělení $t_1 \in T_1$ – tah Hráče I v kole m
- s_2^m – výběr prvku z množiny $\{1, \ldots, k_2\}$ podle pravděpodobnostního rozdělení $t_2 \in T_1$ – tah Hráče II v kole m
- Konečná množina S_i – množina všech tahů s_i, $i \in \{1, 2\}$,
- Množina $S = \{(s_1, s_2) \mid s_1 \in S_1, s_2 \in S_2\}$ – množina možností, jak hráči v dané situaci zahrají,
- Množina Σ – množina strategií $\sigma_k : R \times K \to T_1$ Hráče I – zobrazení určujících na začátku hry na základě $p \in R$ a $k \in K$ rozdělení četnosti volby jednotlivých řádků během hry (nezáleží-li výběr strategie na k, budeme značit pouze σ),
- Množina Θ – množina strategií $\theta : R \to T_2$ Hráče II (Hráč II svou strategii určuje pouze podle $p \in R$, neboť nemá informaci o volbě k),
- Funkce $G_k : S \to \mathbf{R}$ – výplatní funkce, která každé dvojici tahů s přiřadí reálné číslo $G_k(s)$, což je zisk, který následně dostane Hráč I od Hráče II (uvažujeme hry s nulovým součtem).

3.2 Popis hry

Na začátku jsou určena čísla N a $n_1, \ldots, n_N$, pro přesnost výsledků obvykle velká. Pro každé $M = 1, \ldots, N$ je sled akcí následující: Nejprve je náhodně zvoleno $k \in K$ v závislosti na pravděpodobnostním rozdělení $p \in R$, načež pro každé $m = 1, \ldots, n_M$ hráči podle svých strategií $\sigma_k \in \Sigma$, resp. $\theta \in \Theta$ volí své tahy $s_i^m \in S_i$ představující složky $s^m \in S$, čímž je zvolena výplata $G_k(s^m)$. Po odehrání kola n_M je vyhlášen průměrný zisk

$$G_k^M := \frac{1}{n_M} \sum_{m=1}^{n_M} G_k(s^m)$$

a hodnota M se zvýší o 1. Takto se postupuje pro každé M až do $M = N$, po kterém se vyhlásí průměrný zisk za všechna kola

$$G := \frac{1}{\sum_{M=1}^{N} n_M} \sum_{M=1}^{N} n_M G_k^M = \frac{1}{\sum_{M=1}^{N} n_M} \sum_{M=1}^{N} \sum_{m=1}^{n_M} G_k(s^m),$$

který se Hráč I po celou dobu snažil maximalizovat a Hráč II minimalizovat.

Vzhledem k tomu, že uvažujeme strategie, podle kterých hráči volí konkrétní řádky a sloupce se stejnou pravděpodobností po celou hru, je průměrný zisk za hru stejný, jako průměrný zisk za jediné kolo. Ten můžeme určit v závislosti na strategiích hráčů při zvolení matice $k \in K$ takto:

$$G_k^m(\sigma_k, \theta) := \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \; t_1^{\sigma_k}(i) \; t_2^{\theta}(j) \; G_k(i, j),$$

kde $t_1^{\sigma_k}$ je pravděpodobnostní rozdělení na $\{1, \ldots, k_1\}$ určené strategií $\sigma_k \in \Sigma$, tedy $t_1^{\sigma_k}(i)$ je pravděpodobnost, že Hráč I zvolí i-tý řádek (obdobně pro t_2^{θ}). Pro pravděpodobnostní rozdělení p je potom průměrný výplatní vektor přes všechny matice $k \in K$

$$G^p(\sigma, \theta) := \sum_{k \in K} p_k G_k^m(\sigma, \theta),$$

ve kterém je zohledněna i pravděpodobnost výběru té které matice $k \in K$, neboť zisky jednotlivých hráčů nejsou u všech matic stejné. Můžeme tuto hru (uchovávající informaci o K a p) označit $\Gamma(p)$.

Poznámka 3.2.1 Ač o ní Hráč II nemá přesnou informaci, je matice $k \in K$ zvolena na začátku hry pevně a zůstává neměnná pro všechna kola. Mění se až na začátku další hry.

3.3 Určování optimální strategie

Zaveďme nejprve pojmy zaručení a ubránění hodnoty:

Definice 3.3.1 Řekneme, že Hráč I může *zaručit* α, jestliže

$$\forall \varepsilon > 0 \;\; \exists \sigma \in \Sigma \;\; \exists m_0 \in \mathbf{N} \;\; \forall \theta \in \Theta \;\; \forall m \geq m_0 \qquad G^p(\sigma, \theta) \geq \alpha - \varepsilon.$$

Hráč II naopak může *ubránit* α, jestliže

$$\forall \varepsilon > 0 \;\; \forall \sigma \in \Sigma \;\; \exists \theta \in \Theta \;\; \exists m_0 \in \mathbf{N} \;\; \forall m \geq m_0 \qquad G^p(\sigma, \theta) \leq \alpha + \varepsilon.$$

Poznámka 3.3.2 Definice tedy říká, že Hráč I může zaručit zisk hodnoty α, pokud pro každé kladné ε dovede zvolit strategii, že od nějakého kola N do konce hry získává průměrně alespoň $\alpha - \varepsilon$, ať už Hráč II hraje jakkoli.

Druhá část definice pak říká, že Hráč II může ubránit ztrátu 'pouze α', jestliže pro libovolné kladné ε a libovolnou strategii Hráče I dokáže vymyslet strategii, že od nějakého kola N již neztrácí průměrně za kolo více než $\alpha + \varepsilon$.

Poznámka 3.3.3 Termín *zisk* používáme v širším slova smyslu – ziskem mohou být i nekladné hodnoty, tedy ztrátu ztotožňujeme (ostatně jako v běžném životě) se zápornými zisky.

Definice 3.3.4 Řekneme, že hra $\Gamma(p)$ má *hodnotu* $v(p)$, jestliže tuto hodnotu může Hráč I zaručit a současně Hráč II ubránit.

Funkce $G^{\cdot}(\cdot,\cdot)$ je tedy funkce třech proměnných – pravděpodobnostního rozdělení p na K, strategie Hráče I $\sigma \in \Sigma$ a strategie Hráče II $\theta \in \Theta$.

Předpokládáme-li p pevně zvolené na začátku, pak se z $G^p(\cdot,\cdot)$ stává funkce dvou proměnných určující sílu strategie jednoho hráče proti strategii druhého.

Bude-li Hráč I na $G^p(\cdot,\theta)$ ale pohlížet jako na funkci jedné proměnné, pak zjistí, která strategie je za dané situace nejvýhodnější proti soupeřově strategii θ. Určí tím tedy $\sup_\sigma G^p(\sigma,\theta)$. Hráč II se ovšem bude snažit svou ztrátu minimalizovat – svou strategii bude vybírat tak, aby i při nejlepší hře Hráče I dosáhl

$$\overline{w}(p) := \inf_{\theta\in\Theta} \sup_{\sigma\in\Sigma} G^p(\sigma,\theta).$$

Při racionálním hraní obou hráčů, kdy se starají pouze o vlastní zisky (těmto hrám říkáme *antagonistické* – opakem jsou tzv. *hry proti přírodě*, kdy je soupeř lhostejný ke svým výsledkům), se pak $\overline{w}(p)$ stává největším ziskem Hráče II, který může ubránit (tzv. *horní hodnota hry*).

Zcela analogickou úvahou dojdeme k tomu, že

$$\underline{w}(p) := \sup_{\sigma\in\Sigma} \inf_{\theta\in\Theta} G^p(\sigma,\theta)$$

je při racionálním uvažování nejvyšší hodnotou, kterou může zaručit Hráč I (*spodní hodnota hry*). Uvažují-li oba hráči podle právě popsaného schématu, potom říkáme, že se řídí *minimaxovým principem.* V případě her proti přírodě, kdy soupeř pouze náhodně 'dělá naschvály', je obvykle výhodnější jiná strategie, nicméně vše právě popisované pro ně platí úplně stejně.

Věta 3.3.5 Pro jakékoli pravděpodobnostní rozdělení p platí $\overline{w}(p) \geq \underline{w}(p)$.

Důkaz. Z definice suprema a infima víme, že pro všechna $\sigma_0 \in \Sigma$ a $\theta_0 \in \Theta$ platí $\inf_\theta G^p(\sigma_0, \theta) \leq G^p(\sigma_0, \theta_0)$ a $G^p(\sigma_0, \theta_0) \leq \sup_\sigma G^p(\sigma, \theta_0)$, z čehož nám plyne

$$\inf_\theta G^p(\sigma_0, \theta) \leq \sup_\sigma G^p(\sigma, \theta_0).$$

Levá strana nerovnosti nezávisí na θ, proto platí také

$$\inf_\theta G^p(\sigma_0, \theta) \leq \inf_\theta \sup_\sigma G^p(\sigma, \theta).$$

V poslední nerovnosti je pravá strana nezávislá na σ, tedy platí pro všechny strategie Hráče I, speciálně

$$\sup_\sigma \inf_\theta G^p(\sigma, \theta) \leq \inf_\theta \sup_\sigma G^p(\sigma, \theta),$$

tedy $\overline{w}(p) \geq \underline{w}(p)$, což jsme chtěli dokázat. □

Definice 3.3.6 *Smíšenou optimální strategií* nazveme takovou dvojici strategií (σ_0, θ_0), pro kterou platí

$$\sup_{\sigma\in\Sigma} G^p(\sigma, \theta_0) \leq G^p(\sigma_0, \theta_0) \leq \inf_{\theta\in\Theta} G^p(\sigma_0, \theta).$$

Navíc o ní řekneme, že je *čistá*, jestliže $\sigma_0(i) = 0$ pro $i \in \{1, \ldots, k_1\}, i \neq i_0$, $\sigma_0(i_0) = 1$ a $\theta_0(j) = 0$ pro $j \in \{1, \ldots, k_2\}, j \neq j_0$, $\theta_0(j_0) = 1$, neboli pokud je σ_0 vektor délky k_1 složený z $k_1 - 1$ nul a jedné jedničky a obdobně θ_0 obsahuje $k_2 - 1$ nul a jednu jedničku.

Existuje-li čistá smíšená optimální strategie (častěji používáme zkráceně čistá optimální strategie) (σ_0, θ_0), pak říkáme, že má hra v prvku (i_0, j_0) *sedlový bod.*

Poznámka 3.3.7 Definice jinými slovy říká, že smíšená optimální strategie je jakákoli strategie taková, jejíž zisky nemůže ani jeden z hráčů vylepšit ve svůj prospěch jednostrannou změnou své strategie. Je-li tohoto optima dosaženo tím, že hráči volí stále stejný prvek, pak o strategii říkáme, že je čistá a o tomto prvku mluvíme jako o sedlovém bodě.

Optimální strategie přitom nemusí přinášet (a kromě hry s nulovými maticemi ani nepřináší) potenciálně největší zisky oběma hráčům, je to ovšem nejlepší řešení, které nevyžaduje žádnou kooperaci s protihráčem.

Uvažme pro jednoduchost příklad hry s jedinou maticí:

Příklad 3.3.8 Mladý Dmitrij byl vážně uražen Igorem, a tak ho vyzval na souboj. Pravidla jsou jednoduchá – oba si stoupnou proti sobě, s jediným nábojem v pistoli, na určitou vzdálenost od sebe. Zamíří, a když se řekne pal, mají oba najednou možnost vystřelit. Pokud ani jeden nevystřelil, nebo se ani jeden netrefil, jdou o krok blíž a střelba se opakuje. Takto nastanou 3 kola, ovšem během nich může každý z nich vystřelit jen jedinkrát.

Předpokládejme, že jsou Dmitrij i Igor stejně dobří střelci, na základní vzdálenost se trefí s pravděpodobností 20%, o krok blíž na 40% a při posledním pokusu na 50%.

Můžeme do matice zapsat rozdíl pravděpodobnosti v procentech, že Igor zabije Dmitrije, a pravděpodobnosti, že Dmitrij zabije Igora:

$$G^4 = \begin{pmatrix} 0 & -12 & -20 \\ 12 & 0 & 10 \\ 20 & -10 & 0 \end{pmatrix},$$

kde řádek určuje, na kolikátém stanovišti vystřelil Igor, a sloupec, na kolikátém Dmitrij.

Z matice nyní vidíme, že čistou optimální strategií pro Igora je vybrat vždy druhý řádek, tedy vystřelit po prvním přiblížení, neboť nejnižší prvek (0) z něj je vyšší, než jsou nejnižší prvky řádků ostatních (-20, resp. -10). Podobně uvažuje i Dmitrij a vystřelí také po 1. kroku, neboť prostřední sloupec má nejvyšší hodnotu 0. Čistá optimální strategie Igora tedy je $(0, 1, 0)$, stejně jako u Dmitrije.

Ve výsledku tedy oba vystřelí na 2. stanovišti, což je sedlový bod celé hry, a jeho prvek (0) je její hodnota v.

Věta 3.3.9 Z existence sedlového bodu plyne $\overline{w}(p) = \underline{w}(p)$.

Důkaz. Budiž (σ_0, θ_0) čistá optimální strategie hry s výplatní funkcí $G^p(\sigma, \theta)$. Potom všechny strategie $\sigma \in \Sigma$ Hráči I (pokud Hráč II hraje strategii θ) přinášejí zisk v nejlepším případě stejný, tedy $G^p(\sigma, \theta_0) \leq G^p(\sigma_0, \theta_0)$. Obdobně Hráči II odchylka od optimální strategie přináší větší ztráty (tedy Hráči I větší zisky), tudíž $G^p(\sigma_0, \theta_0) \leq G^p(\sigma_0, \theta)$. Z toho nám vyplývá, že pro všechny strategie $\sigma \in \Sigma$ a $\theta \in \Theta$ jest

$$G^p(\sigma, \theta_0) \leq G^p(\sigma_0, \theta_0) \leq G^p(\sigma_0, \theta).$$

Odtud vyplývá

$$\sup_{\sigma} G^p(\sigma, \theta_0) \leq G^p(\sigma_0, \theta_0) \leq \inf_{\theta} G^p(\sigma_0, \theta),$$

z čehož s přihlédnutím k nerovnostem

$$\inf_\theta \sup_\sigma G^p(\sigma,\theta) \leq \sup_\sigma G^p(\sigma,\theta_0),$$

$$\inf_\theta G^p(\sigma_0,\theta) \leq \inf_\theta \sup_\sigma G^p(\sigma,\theta)$$

dostáváme

$$\inf_\theta \sup_\sigma G^p(\sigma,\theta) \leq G^p(\sigma_0,\theta_0) \leq \sup_\sigma \inf_\theta G^p(\sigma,\theta),$$

tedy $\overline{w}(p) \leq \underline{w}(p)$. Ovšem s ohledem na Větu 3.3.5 musí platit i opačná nerovnost, z čehož nám již plyne $\overline{w}(p) = \underline{w}(p)$.

$\square$

Platí ovšem i mnohem silnější tvrzení, které říká, že ve smíšených strategiích existuje rovnovážný stav pro libovolnou hru 2 hráčů s nulovým součtem. Zformulujme v této chvíli (bez důkazu – ten lze nalézt například v [6]) tuto základní větu celé teorie her:

Věta 3.3.10 (von Neumannova o minimaxu) Pro každou hru dvou hráčů s nulovým součtem o konečném počtu kol existují strategie $\sigma_0 \in \Sigma$ a $\theta_0 \in \Theta$ vedoucí k výplatě V, že

- pro strategii $\theta_0 \in \Theta$ je V nejvyšší zisk, který může Hráč I zaručit,
- pro strategii $\sigma_0 \in \Sigma$ je V nejnižší ztráta, kterou může Hráč II ubránit.

Tvrzení 3.3.11 Zvolme pevně $p \in R$, podle p také $k \in K$ a označme Σ_C množinu všech čistých strategií Hráče I (tedy všech vektorů délky k_1 s $(k_1 - 1)$ nulami a jednou jedničkou) a Θ_C množinu čistých strategií Hráče II ve hře $\Gamma(p)$. Pokud nyní platí

$$\inf_{\theta\in\Theta_C} \sup_{\sigma\in\Sigma_C} G^p(\sigma,\theta) = \sup_{\sigma\in\Sigma_C} \inf_{\theta\in\Theta_C} G^p(\sigma,\theta),$$

potom v $\Gamma(p)$ existuje sedlový bod.

Důkaz. Nechť $\sigma_0 \in \Sigma_C$ a $\theta_0 \in \Theta_C$ jsou takové strategie, pro které platí

$$\inf_{\theta \in \Theta_C} G^p(\sigma_0, \theta) = \sup_{\sigma \in \Sigma_C} \inf_{\theta} G^p(\sigma, \theta),$$

$$\sup_{\sigma \in \Sigma_C} G^p(\sigma, \theta_0) = \inf_{\theta \in \Theta_C} \sup_{\sigma} G^p(\sigma, \theta).$$

Jejich existence plyne z konečného množství čistých strategií (konkrétně jich je $|\Sigma_C|\,|\Theta_C| = k_1 k_2$. Předpokládáme, že platí rovnost pravých stran v právě uvedených vzorcích, proto musí platit i rovnost levých:

$$\inf_{\theta \in \Theta_C} G^p(\sigma_0, \theta) = \sup_{\sigma \in \Sigma_C} G^p(\sigma, \theta_0).$$

Využijeme-li znalost o supremu, že $\sup_{\sigma \in \Sigma_C} G^p(\sigma, \theta_0) \geq G^p(\sigma_0, \theta_0)$, dostaneme

$$\inf_{\theta \in \Theta_C} G^p(\sigma_0, \theta) = \sup_{\sigma \in \Sigma_C} G^p(\sigma, \theta_0) \geq G^p(\sigma_0, \theta_0),$$

tedy pro všechna $\theta \in \Theta_C$

$$G^p(\sigma_0, \theta) \geq G^p(\sigma_0, \theta_0).$$

Obdobnou úvahou dospějeme také k

$$G^p(\sigma, \theta_0) \leq G^p(\sigma_0, \theta_0)$$

pro všechna $\sigma \in \Sigma_C$ – tedy pro všechna $\sigma \in \Sigma_C$ a $\theta \in \Theta_C$ platí

$$G^p(\sigma, \theta_0) \leq G^p(\sigma_0, \theta_0) \leq G^p(\sigma_0, \theta).$$

Strategie (σ_0, θ_0) jsou tedy sedlovým bodem funkce $G^p(\sigma, \theta)$. □

Tvrzení 3.3.12 Maticová hra může mít více sedlových bodů (a tedy čistých optimálních strategií), všechny ale mají stejnou hodnotu.

Důkaz. (i) Dokažme nejprve, že hra může mít více sedlových bodů:

Uvažme hru s jedinou maticí (oba hráči tedy mají přesné informace o jejím tvaru), a to

$$G^5 = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 2 & 1 \end{pmatrix}.$$

Tvrdíme nyní, že prvky $(2,1)$ i $(2,2)$ jsou sedlovými body, tedy že jak $(\sigma_1, \theta_1) := ((0,1,0),(1,0,0))$, tak $(\sigma_2, \theta_2) := ((0,1,0),(0,1,0))$ jsou čisté optimální strategie. Čistota je zřejmá, nepatrně těžší je dokázat, že jsou optimální. Dokažme optimalitu pouze pro (σ_1, θ_1), neboť pro (σ_2, θ_2) by se prováděla zcela analogicky.

Definice 3.3.6 nám říká, že musíme ověřit, že pro každou dvojici (q_1, q_2), $q_1, q_2 > 0$ a $q_1 + q_2 = 1$ platí

$$G^p((q_1, q_2), \theta_0) \leq G^p(\sigma_0, \theta_0) = 2,$$

tedy že Hráč I nemůže při stálém volení levého sloupce Hráčem II zvýšit zisky tím, že by přerozdělil pravděpodobnost zvolení horního, resp. spodního řádku. Jeho zisk je při strategiích (σ_0, θ_0) roven 2, při strategiích $((q_1, q_2), \theta_0)$ pak

$$1q_1 + 2q_2 \leq 2q_1 + 2q_2 = 2(q_1 + q_2) = 2,$$

tedy v nejlepším případě stejný.

Obdobně je třeba pro každou trojici (p_1, p_2, p_3), $p_1, p_2, p_3 > 0$ a $p_1 + p_2 + p_3 = 1$ dokázat

$$G^p(\sigma_0, (p_1, p_2, p_3)) \leq G^p(\sigma_0, \theta_0) = 2,$$

tedy že své zisky nemůže zvýšit ani Hráč II. To již dopočítáme snadno:

$$2p_1 + 2p_2 + 1p_3 \leq 2p_1 + 2p_2 + 2p_3 = 2(p_1 + p_2 + p_3) = 2.$$

(ii a) Jednoznačnost hodnoty dokažme sporem:

Nechť existují dva sedlové body s hodnotami $x_1 < x_2$, bez újmy na obecnosti uvažujme, že každý leží v jiném řádku (pokud by se lišily pouze sloupcem, provedeme důkaz analogicky pro Hráče II). Potom jsou v řádku, kde leží x_2, všechny hodnoty alespoň rovny x_2, což je ostře víc, než x_1, tedy přechodem z řádku, na kterém je x_1, na řádek s x_2 může Hráč I zlepšit svoji minimaxovou strategii, takže jeho optimální strategií není stálé volení řádku s x_1, a tudíž zde není ani sedlový bod, což je spor.

(ii b) Jednoznačnost lze dokázat i přímo:

Jestliže existuje sedlový bod s hodnotou $v(p)$, pak tuto hodnotu může Hráč I zaručit (v každém kole bude od Hráče II dostávat alespoň $v(p)$) a současně Hráč II ubránit (nebude Hráči I dávat více než $v(p)$). Pak ale z faktu, že se jedná o hru s nulovým součtem, jednoznačnost hodnoty sedlového bodu přímo plyne.

□

Definice 3.3.13 Funkce více proměnných f je *konkávní* na krychli S, pokud pro každé 2 různé body $x, y \in S$ a každá reálná $\alpha, \beta > 0$ splňující $\alpha + \beta = 1$ platí

$$f(\alpha x + \beta y) \geq \alpha f(x) + \beta f(y).$$

Věta 3.3.14 Funkce (proměnné p)

$$\overline{w}(p) = \inf_{\theta \in \Theta} \sup_{\sigma \in \Sigma} G^p(\sigma, \theta) \qquad \text{a} \qquad \underline{w}(p) = \sup_{\sigma \in \Sigma} \inf_{\theta \in \Theta} G^p(\sigma, \theta)$$

jsou konkávní na množině pravděpodobnostních rozdělení na K.

Důkaz. Důvod je stejný pro obě funkce, dokažme tedy tvrzení jen pro $\overline{w}$.

Budiž E konečná množina indexů, $\{p_e\}_{e \in E}$ množina pravděpodobnostních rozdělení na K a $\alpha := (\alpha_e \in [0,1])_{e \in E}$ vektor délky $|E|$ splňující $\sum_{e \in E} \alpha_e p_e = p$. Abychom dokázali konkávitu $\overline{w}(p)$, musíme ukázat, že

$$\overline{w}\left(\sum_{e \in E} \alpha_e p_e\right) \geq \sum_{e \in E} \alpha_e \overline{w}(p_e).$$

Pro tento účel si zahrajeme 2 podobné hry. V obou je nejprve (podle rozdělení pravděpodobnosti α) zvolen index $e \in E$, v jedné je však Hráči II výsledek tohoto výběru oznámen, zatímco ve druhé nikoli (Hráč I ho ví pokaždé). V obou verzích je již pak standardně zvolena matice $k \in K$ podle pravděpodobnostního rozdělení p_e a hráči zvolí své strategie $\sigma_k \in \Sigma$, resp. $\theta \in \Theta$, které vedou k celkové výplatě $G^n_k(\sigma_k, \theta)$, kde n je délka hry. Ať už ale Hráči II zvolení e oznámeno bylo nebo ne, o výběru k informován není ani v jednom případě.

Uvažujme nyní hru, ve které je Hráč II informován o e. Pak se jedná o hru $\Gamma(p_e)$, tedy budeme-li uvažovat i pravděpodobnost zvolení jednotlivých prvků $e \in E$, je nejhorší případ, který může pro Hráče II nastat, roven

$$\sum_{e \in E} \alpha_e \inf_{\theta \in \Theta} \sup_{\sigma \in \Sigma} G^{p_e}(\sigma, \theta) = \sum_{e \in E} \alpha_e \overline{w}(p_e).$$

To je pro Hráče II jistě lepší případ (tedy menší hodnota) než v případě, že o e informován není, neboť potom nemůže optimalizovat strategii přímo podle pravděpodobnostního rozdělení p_e, podle kterého se vybírá matice k. Hráči I je jedno, zda zná výběr e nebo ne, neboť má v každém tahu konkrétní informaci přímo o hodnotě k – i kdyby tedy hrál, jako by e neznal, tak je hra ekvivalentní s $\Gamma\left(\sum_{e\in E}\alpha_e p_e\right)$, pro jejiž nejhorší případ platí

$$\inf_{\theta\in\Theta}\sup_{\sigma\in\Sigma} G^{\sum_{e\in E}\alpha_e p_e}(\sigma,\theta) = \overline{w}\left(\sum_{e\in E}\alpha_e p_e\right) = \overline{w}(p) \geq \sum_{e\in E}\alpha_e\overline{w}(p_e).$$

□

Důsledek 3.3.15 Pokud ve hře $\Gamma(p)$ může Hráč I zaručit $f(p)$, potom může zaručit i $Cav(f(p))$, kde $Cav(f)$ je konkávní funkce, pro niž platí, že $Cav(f(p)) \geq f(p)$ pro všechna rozdělení pravděpodobnosti p na K a současně není v žádném bodě větší než jakákoli funkce, která by to splňovala také.

Tvrzení 3.3.16 Budiž f funkce proměnné p. Potom konkávní funkce, pro niž platí, že $Cav(f(p)) \geq f(p)$ pro všechna rozdělení pravděpodobnosti p na K a současně není v žádném bodě větší než jakákoli funkce, která by to splňovala také, existuje.

Důkaz. Výplatní matice uvažujeme konečné, tedy Hráč I za kolo nikdy nemůže zaručit víc, než je maximální prvek (označme ho G_{max}). Potom $f_{max} \equiv G_{max}$ je konkávní majoranta funkce $f(p)$, a množina všech konkávních majorant funkce f (označme ji KM) tedy není prázdná. Potom

$$Cav(f(p)) := \inf_{f_{KM}\in KM} f_{KM}.$$

Zbývá ještě dokázat, že je $Cav(f)$ konkávní – tedy v souladu s Definicí 3.3.13 ukázat, že pro všechna $p_1, p_2 \in R$ a všechna $\alpha, \beta > 0$ splňující $\alpha + \beta = 1$ platí

$$Cav\left(f\left(\alpha p_1 + \beta p_2\right)\right) = \alpha\ Cav(f(p_1)) + \beta\ Cav(f(p_2)).$$

To plyne z konkávity funkcí f_{KM} a vlastností infima:

$$\begin{aligned} Cav\left(f\left(\alpha p_1 + \beta p_2\right)\right) &= \inf_{f_{KM}\in KM} f_{KM}\left(\alpha p_1 + \beta p_2\right) \geq \\ &\geq \inf_{f_{KM}\in KM}\left(\alpha f_{KM}(p_1) + \beta f_{KM}(p_2)\right) \geq \\ &\geq \inf_{f_{KM}\in KM}\alpha f_{KM}(p_1) + \inf_{f_{KM}\in KM}\beta f_{KM}(p_2) = \\ &= \alpha \inf_{f_{KM}\in KM} f_{KM}(p_1) + \beta \inf_{f_{KM}\in KM} f_{KM}(p_2) = \\ &= \alpha\ Cav(f(p_1)) + \beta\ Cav(f(p_2)). \end{aligned}$$

□

Poznámka 3.3.17 V případě, že Hráč I nevyužívá svoji počáteční výhodu (jeho strategie je stejná, ať už byla vybrána jakákoli matice), je množina jeho tahů pro každé kolo n

$$B = \left\{ x_1 \in S_1^{|K|} \mid \forall k, k' \in K : x_{1,k} = x_{1,k'} \right\},$$

tedy množina $|K|$-tic stejných prvků. Potom ale během hry ani není potřeba znát, která matice k byla na začátku vybrána, ba naopak loterie může proběhnout až na konci celé hry pouze z důvodu určení zisků obou hráčů.

3.4 Návrat k úvodním příkladům

Definice 3.5.1 Pro p pravděpodobnostní rozdělení na K definujme *hru bez znalosti* (označme ji $\Gamma_1(p)$) jako hru s nulovým součtem pro 2 hráče hranou s maticí

$$G(p) = \sum_{k \in K} p_k k,$$

kde k jsou jednotlivé matice z K. Jako $v_1(p)$ označíme hodnotu této hry bez znalosti $\Gamma_1(p)$.

Tvrzení 3.5.2 Hráč I může zaručit $v_1(p)$ ve hře $\Gamma(p)$ tím, že bude hrát optimální strategii v příslušné hře $\Gamma_1(p)$.

Důkaz. Hra $\Gamma_1(p)$ není nic jiného než $\Gamma(p)$, při které Hráči I nebylo sděleno, s jakou maticí se hraje. Bude-li tedy hrát stejně, jako kdyby tuto informaci neměl, bude tak muset hrát i Hráč II. Hra je pak tedy zcela ekvivalentní $\Gamma_1(p)$, v níž hraním optimální strategie (ta existuje z von Neumannovy Věty o minimaxu 3.3.10) Hráč I dosáhne $v_1(p)$. □

Tvrzení 3.5.3 Hráč I takto může zaručit i $Cav(v_1(p))$.

Důkaz. Plyne z Tvrzení 3.5.2 a Důsledku 3.3.15. □

Poznámka 3.5.4 Bez důkazu uveďme, že všechna tato tvrzení mají i duální znění pro situaci, kdy je informovaným hráčem Hráč II, zatímco Hráč I informaci o výběru k nezíská. Rozdíl ve znění tvrzení pak bude ale především ten, že Hráč II díky své informovanosti může zaručit $Vex(v_1(p))$, což je 'nejmenší konvexní funkce větší než v_1' (analogie Důsledku 3.3.15).

Tato nesymetrie je dána tím, že i když změníme informovaného hráče, bude se Hráč II stále snažit hodnotu výplaty minimalizovat.

Důsledek 3.5.5 V případě, že ani jeden z hráčů není informován o výběru k, může Hráč I zaručit zisk $v_1(p)$ a současně tuto ztrátu může Hráč II ubránit. Je to tedy hodnota hry s neúplnou informací na obou stranách. V případě, že je Hráč I informován o výběru matice, pak může využitím své znalosti zaručit i $Cav(v_1(p))$ (tím neříkáme, že by nemohl zaručit i víc), je-li informovaným hráčem Hráč II, pak lze zaručit prozměnu $Vex(v_1(p))$.

Příklad 1.4.1 (pokračování) Označme p pravděpodobnost, se kterou bude zvolena matice G_1^1. Pravděpodobnost zvolení G_2^1 pak bude $(1-p)$. Potom

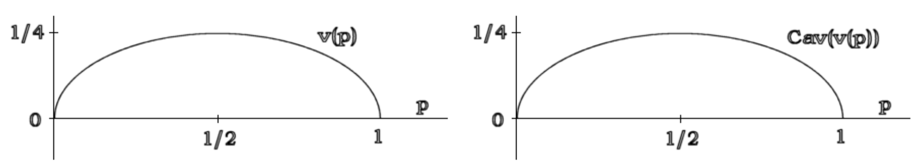
$$G(p) = \sum_{k \in K} p_k k = p \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (1-p) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}.$$

Uveďme bez výpočtu (případní zájemci si mohou ověřit v [1]), že hodnota hry je

$$v_1(p) = p(1-p).$$

Tím jsme tedy mimo jiné ověřili, že pro $p = 1/2$ může Hráč I zaručit $Cav(v_1(1/2)) = 1/4$, což bylo nastíněno již v úvodu.

Graf průběhu hodnoty hry v závislosti na p a hodnotu, kterou může Hráč I zaručit (v tomto případě je $v_1(p)$ konkávní, a tak se neliší), můžeme vidět na následujícím grafu:



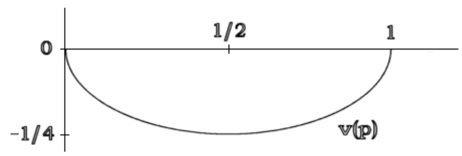
Příklad 1.4.3 (pokračování) Označme opět p pravděpodobnost, se kterou bude zvolena matice G_1^2 a $(1-p)$ pravděpodobnost zvolení G_2^2. Potom

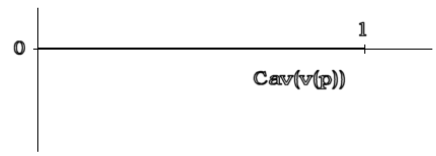
$$G(p) = \sum_{k \in K} p_k k = p \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + (1-p) \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -p & 0 \\ 0 & -(1-p) \end{pmatrix}$$

a dle [1]

$$v_1(p) = -p(1-p).$$

Nyní využijeme Tvrzení 3.5.3 ze kterého víme, že pokud může Hráč I zaručit $v_1(p)$, může při využití své znalosti zvolené matice zaručit i $Cav(v_1(p))$, tedy výběrem nulových řádků zamezit jakýmkoli ziskům soupeře (pak platí $Cav(v_1(p)) \equiv 0$). Obě funkce vypadají následovně:





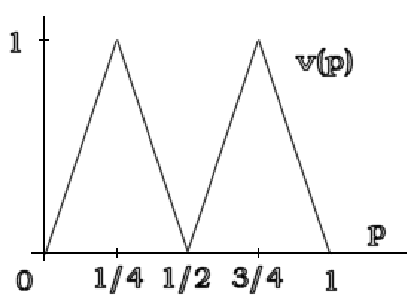
Příklad 1.4.4 (pokračování) V tomto příkladě platí:

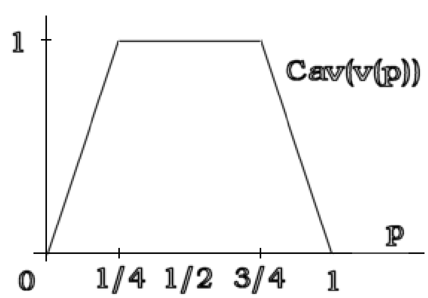
$$G(p) = p\begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix} + (1-p)\begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 4p & 4-4p & 4p-2 \\ 4p & 4-4p & 2-4p \end{pmatrix},$$

$$v_1(p) = \begin{cases} 4p, & p \in [0, 1/4], \\ 2-4p, & p \in [1/4, 1/2], \\ 4p-2, & p \in [1/2, 3/4], \\ 4-4p, & p \in [3/4, 1], \end{cases}$$

$$Cav(v_1(p)) = \begin{cases} 4p, & p \in [0, 1/4], \\ 1, & p \in [1/4, 3/4], \\ 4-4p, & p \in [3/4, 1]. \end{cases}$$

Grafy funkcí $v_1(p)$ a $Cav(v_1(p))$ tedy vypadají takto:





Pro $p = 1/2$ tedy vidíme, že při správně zvolené strategii plynoucí z povědomí o pravé matici může Hráč I opravdu zaručit i průměrný výdělek $Cav(v_1(1/2)) = 1$. Optimální strategie (graf funkce $v(p)$) pro hru s nevyužíváním znalosti matice, při které se nabývá hodnota 1 pouze pro $p = 1/4$ a $p = 3/4$, přitom ale při $p = 1/2$ nezaručuje zisky žádné, tedy závěr, že lze zaručit zisk 1, je poměrně netriviální výsledek.

Literatura

[1] Aumann R.J., Hart S.: *Handbook of Game Theory with Economic Applications, volume 1*, Elsevier Science Publishers – North Holland, Amsterdam, 1992, 110-132

[2] Turnovec F., Chobot M.: *Teória hier*, Slovenské pedagogické nakladateľstvo, Bratislava, 1967, 39-65

[3] Axelrod R.: *The evolution of cooperation*, Basic Books, New York, 1984

[4] Wikipedia.org: *Prisoner's dilemma*
http://en.wikipedia.org/wiki/Prisoner's_dilemma

[5] Malý M.: *Experimenty s algoritmy na osobním blogu*
http://maly.blog.sme.sk/c/64463/Vezeni-se-meni.html

[6] Goodrich M.A.: *Proof of the Minimax Theorem*, 6-8
http://students.cs.byu.edu/c̃s670ta/MinimaxTheoremProof.pdf

[7] Dupač V., Hušková M.: *Pravděpodobnost a matematická statistika*, Univerzita Karlova v Praze – Nakladatelství Karolinum, Praha, 2005; 94, 115, 131, 155

[8] Zajíček L.: *Vybrané partie z matematické analýzy pro 1. a 2. ročník*, MatfyzPress, Praha, 2003