UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA

DIPLOMOVÁ PRÁCE

Opakované vězňovo dilema

Katedra matematické analýzy a aplikací matematiky
Vedoucí diplomové práce: **Mgr. Pavel Holeček, Ph.D.**
Vypracovala: **Bc. Hana Krakovská**
Studijní program: N1103 Aplikovaná matematika
Studijní obor: Aplikace matematiky v ekonomii
Forma studia: prezenční
Rok odevzdání: 2019

BIBLIOGRAFICKÁ IDENTIFIKACE

Autor: Bc. Hana Krakovská

Název práce: Opakované vězňovo dilema

Typ práce: Diplomová práce

Pracoviště: Katedra matematické analýzy a aplikací matematiky

Vedoucí práce: Mgr. Pavel Holeček, Ph.D.

Rok obhajoby práce: 2019

Abstrakt: Tato práce se zabývá opakovaným vězňovým dilematem, problémem z oblasti teorie her. Je rozčleněná do dvou hlavních částí, teoretické a praktické. Teoretická část uvede čtenáře nejdříve do problematiky vězňova dilematu jednokolového a následně opakovaného. Představí nejznámější strategie a turnaje opakovaného vězňova dilematu. Část praktická vychází z reálných dat získaných z turnaje, který byl zorganizován autorkou práce na Přírodovědecké fakultě Univerzity Palackého v Olomouci. Cílem turnaje bylo sledovat jak chování hráčů, kterými nebyly naprogramované algoritmy, nýbrž studenti, tak i souvislosti mezi naším turnajem a turnajem Roberta Axelroda, amerického vědce, který s těmito turnaji začal. Využito bylo metod pro klasifikaci hráčů a analýzu průběhu hry. Tyto navržené metody byly otestovány na reálných datech získaných během turnaje.

Klíčová slova: vězňovo dilema, opakované vězňovo dilema, teorie her, strategie, hráči, výplatní matice, rovnovážný bod, optimum, turnaj

Počet stran: 72

Počet příloh: 2

Jazyk: český

BIBLIOGRAPHICAL IDENTIFICATION

Author: Bc. Hana Krakovská

Title: Iterated prisoner's dilemma

Type of thesis: Master's

Department: Department of Mathematical Analysis and Application of Mathematics

Supervisor: Mgr. Pavel Holeček, Ph.D.

The year of presentation: 2019

Abstract: This thesis deals with the topic of the iterated prisoner's dilemma – a problem coming out of the area of game theory. The thesis is divided into two main parts – – theoretical and practical. Firstly, the theoretical part will introduce a reader to the problems of the prisoner's dilemma. Subsequently, the iterated prisoner's dilemma will be introduced. The practical part of the thesis is based on real data gathered from a tournament, which had been organized by author of this thesis at the Faculty of Science, Palacky University in Olomouc. The aim of this tournament was to observe behavior of the players who were not programmed algorithms but students themselves. Also, relations of our tournament and Robert Axelrod's tournament were contemplated. We concentrated on methods for classification of players as well as on analysis of game process. These designed methods were tested on real data gathered during the tournament.

Key words: prisoner's dilemma, iterated prisoner's dilemma, game theory, strategy, players, payoff matrix, Nash equilibrium, optimum, tournament

Number of pages: 72

Number of appendices: 2

Language: Czech

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracovala samostatně pod vedením pana Mgr. Pavla Holečka, Ph.D. a všechny použité zdroje jsem uvedla v seznamu literatury.

V Olomouci dne ……………………   ….................................

podpis

Obsah

Poděkování

Speciální poděkování patří mému vedoucímu diplomové práce, panu Mgr. Pavlu Holečkovi, PhD., který mi byl vždy ochoten pomoci. Kromě času, který mi věnoval na konzultacích, mi také vytvořil software, zaznamenávající průběh turnaje, k praktické části mé diplomové práce.

Dále bych ráda poděkovala své rodině, příteli a nejlepší kamarádce, kteří mi byli velkou psychickou oporou.

Děkuji také studentům Univerzity Palackého, kteří se zúčastnili turnaje, neboť právě díky jejich ochotě se mi podařilo získat veškerá data pro analýzu.

Úvod

Dva podezřelí, A a B jsou vyslýcháni policií. Neexistují žádní svědci a ani důkazy, které by mohly pomoct k vyřešení případu. Podezřelí jsou vyslýcháni zvlášť. Neví, co řekl druhý z nich a mohou buď mlčet, nebo mluvit. Jestli půjdou do vězení, a na kolik let případně, závisí na tom, jak se rozhodnou oba podezřelí.

Poznamenejme, že mlčení (odmítnutí komunikovat s policií) zde značí spolupráci s druhým podezřelým, v případě, že podezřelý udá druhého podezřelého, považuje se toto jeho chování za zradu.

Pokud nebude ani jeden z podezřelých mluvit s policií, pak budou oba odsouzeni na 1 rok. Pokud naopak by se oba udali vzájemně, pak by si ve vězení odseděli 5 let. Ovšem v případě, že bude mluvit pouze jeden z nich a druhý bude mlčet, potom ten, který mlčel, si ve vězení odsedí 10 let a druhý podezřelý bude volný.

	B mlčí (spolupracuje)	**B mluví (zrazuje)**
A mlčí (spolupracuje)	Oba podezřelí si odsedí 1 rok.	B bude volný, A si odsedí 10 let.
A mluví (zrazuje)	A bude volný, B si odsedí 10 let.	Oba jsou odsouzeni na 5 let.

Tabulka 1: tabulka počtu let strávených ve vězení pro jednotlivé podezřelé

Situace uvedená v příkladu je známá jako vězňovo dilema. Tato práce pojednává o opakovaném vězňově dilematu, nelze se ovšem do něj pouštět bez nastínění „klasické“ jednokolové varianty. Přestože opakované vězňovo dilema z jednokolového vychází, nelze říci, že se dá tento rozhodovací problém řešit stejným způsobem.

U verze bez opakování hra po odehrání jednoho kola končí a jediné, co hráči mohou na jejím konci udělat je být nespokojeni se svým rozhodnutím, které učinili.

Co když mají ale proti sobě odehrát několik kol za sebou? Co udělají, když je soupeř v předchozím kole zradí? Budou si hráči zrady oplácet? Anebo zradí se vůbec? Pokud se bavíme o odlišnostech mezi těmito dvěma problémy, stálo by ještě za zmínku, že na rozdíl od vězňova dilematu bez opakování, kde lze určit optimální strategii, u opakovaného vězňova dilematu to mnohdy nejde. Tam je pak potřeba rozlišit, zda hráči předem znají nebo neznají počet kol hry.

Od vzniku samostatného rozhodovacího problému dodnes uplynula již poměrně dlouhá doba, během níž vznikla spousta algoritmů od jednodušších až po složité, které se účastnily různých turnajů vězňova dilematu, a tedy existuje celá řada doporučení, vlastností úspěšných strategií apod., kterými se mohou hráči vězňova dilematu řídit. U opakovaného vězňova dilematu záleží také na tom, jaký soupeř je proti hráči postaven a podle toho by se také měla odvíjet případná strategie hráče.

Opakované vězňovo dilema je spojováno především se jménem **Robert Marshall Axelrod**, což je americký politolog, profesor působící na univerzitě v Michiganu, známý především pro svou interdisciplinární práci [1], ve které popisuje výsledky počítačových turnajů opakovaného vězňova dilematu, které také uspořádal a zajímavosti, kterých si u algoritmů všímal.

Ukázalo se, že si daleko lépe vedou strategie, které nemají sklony za každou cenu zrazovat a spíše se snaží o udržení spolupráce [2].

Čtenář se mj. seznámí s různými algoritmy od těch nejjednodušších až po složitější, ať už z turnajů Roberta Axelroda nebo některých dalších, které na jeho turnaje navázaly.

Druhá část práce je praktická a zabývá se metodami popisu chování hráčů, dvaceti studentů Univerzity Palackého v Olomouci. Metody byly testovány na reálných datech, získaných uspořádáním vlastního turnaje opakovaného vězňova dilematu. Tím, že hráči byli studenti, nikoliv naprogramované algoritmy, je potřeba počítat s tím, že se jejich chování mohlo v průběhu hry měnit, od čehož se odvíjel i způsob analýzy hry. Samotná příprava i průběh turnaje jsou popsány na začátku **kapitoly 7**.

1 Úvod do teorie her a základní pojmy z teorie her

Teorie her, oblast aplikované matematiky a ekonomie, je poměrně novým tématem. V porovnání se samotnou matematikou, jejíž kořeny sahají až do pravěku, první zmínka o teorii her pochází až z 20. století. Roku 1928 publikoval **John von Neumann** fundamentální teorém her dvou hráčů s nulovým součtem jejich výplat. Počátky teorie her jsou často spojovány se jmény **John von Neumann** a **Oskar Morgenstern** a jejich dílem [10] z roku 1944. Od r. 1944 se o teorii her začala zabývat spousta vědců, a stávala se čím dál tím populárnější oblastí. Našla si využití téměř ve všech oblastech každodenního života, což byl jeden z hlavních důvodů. Antropologie, psychologie, ekonomie, politika, biologie, filozofie - to je jen základní výčet oblastí, kde našla teorie her své využití.

Teorie her řeší rozhodovací problémy více účastníků (hráčů), problémy takové, kde je více než jedna varianta řešení a je potřeba zvážit, jak tyto problémy řešit, což často není úplně jednoduché.

V jednoduchosti se seznámíme s tím, jak lze hry dělit. Blíže se zaměříme zejména na části, které se budou týkat tématu diplomové práce, tedy vězňova dilematu. Pojmy a definice jsou převzaty z [9], [11], [15].

1.1 Dělení her

Jedno ze základních dělení her je rozdělení na **hry 2 hráčů** a **hry více hráčů**. Hrami více hráčů se nebudeme dále zabývat, neboť vězňovo dilema se většinou chápe[1] jako hra dvou hráčů. Dalším dělení her je na hry **kooperativní** a **nekooperativní**, tedy, kdy spolu hráči mohou[2] a kdy nemohou komunikovat během hry. Toto dělení bude mít vliv zejména na volby strategií hráčů, přičemž strategií budeme rozumět nějaký plán, kompletní popis, jak hrát hru [11]. U vězňova dilematu budeme dále předpokládat, že se nemohou předem domluvit, jakou strategii každý z nich zvolí. Další dělení je na **hry v normálním** a **hry v rozvinutém tvaru.**

[1] a my jej tak budeme uvažovat v práci
[2] Mohou se domluvit, jak budou hru hrát za účelem co nejlepšího výstupu

Hra v normálním tvaru ([11], [16]) je charakterizována třemi množinami – množinou hráčů, prostorů strategií a výplatních funkcí. Pro **hru 2 hráčů** je dána uspořádanou trojicí (X, Y, f), kde:

(1) *X* je neprázdná množina, soubor strategií hráče *A*

$$X = \{x_1, x_2, x_3, \dots x_m\}$$

(2) *Y* je neprázdná množina, soubor strategií hráče *B*

$$Y = \{y_1, y_2, y_3, \dots y_n\}$$

(3) f obsahuje výplatní funkce obou hráčů:

($f_1: X \times Y \to \mathbb{R}$ - výplatní funkce 1. hráče a $f_2: X \times Y \to \mathbb{R}$ výplatní funkce 2. hráče)

$$f = \{f_1, f_2\}$$

Pozn.: Funkce f bývá nejčastěji zadávána ve formě výplatní matice[3] nebo výplatní dvojmatice[4]. Hráč A zvolí libovolné[5] $x_i \in X$, hráč B $y_j \in Y, i = 1, 2, 3, \dots, m;\ j = 1, 2, 3, \dots, n$.

Je potřeba rozlišit, zda se jedná o **hru s nulovým součtem** výplat hráčů nebo **hru s nenulovým součtem.** Pokud dává součet výplat pro každou volbu strategií nulu (tedy to, co jeden hráč získá, to druhý hráč ztratí), pak jde o hry s **nulovým součtem.** Pro hru dvou hráčů s nulovým součtem platí:

$$\boldsymbol{f_1(x_i, y_j) + f_2(x_i, y_j) = 0}$$

$$\forall i \in \{1, 2, 3, \dots, m\}, \forall j \in \{1, 2, 3, \dots, n\}$$

V případě her s nulovým součtem můžeme výplaty prvního hráče zapsat ve formě (*<u>obecná matice hry s nulovým součtem</u>*) takto:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}$$

kde: $$\boldsymbol{f_1(x_i, y_j) = x_{ij}}$$

[3] V případě her s nulovým součtem. Dvojmatice – matice, která má jako prvky uspořádané dvojice, první člen uspořádané dvojice představuje výplatu hráče A, druhý člen výplatu hráče B.
[4] V případě her s nenulovým součtem
[5] Hráč A vybírá z řádků, hráč B volí mezi sloupci.

$$f_2(x_i, y_j) = y_{ij}$$

Navíc[6]:

$$y_{ij} = -x_{ij}$$

$$\forall i \in \{1, 2, 3, \dots, m\}, \forall j \in \{1, 2, 3, \dots, n\}$$

Obecně x_{ij} představuje výplatu hráče A, pokud by hráč A zvolil i-tou strategii a hráč B j - - tou strategii. Hráč B by ve hře s nulovým součtem obdržel výplatu $-x_{ij}$, proto se jeho výplaty do matice nezapisují.

Pokud součet výplat pro hru 2 hráčů pro jednotlivé situace nedává vždy nulu, tj.:

$$f_1(x_i, y_j) + f_2(x_i, y_j) \neq 0$$

$$\exists i \in \{1, 2, 3, \dots, m\}, \exists j \in \{1, 2, 3, \dots, n\}$$

pak hovoříme o **hrách s nenulovým součtem**.

Výplaty hráčů bychom opět zapsali maticově s tím rozdílem, že nyní už je potřeba zapisovat výplaty obou hráčů a tak místo klasické matice budeme pracovat s *dvojmaticí* (*obecná matice hry s nenulovým součtem*):

$$\begin{pmatrix} (x_{11}, y_{11}) & (x_{12}, y_{12}) & \cdots & (x_{1n}, y_{1n}) \\ \vdots & \vdots & \ddots & \vdots \\ (x_{m1}, y_{m1}) & (x_{m2,}, y_2) & \cdots & (x_{mn}, y_{mn}) \end{pmatrix}$$

kde (x_{ij}, y_{ij}) je uspořádaná dvojice (výplata hráče A, výplata hráče B), pokud by hráč A zvolil i-tou strategii a hráč B j-tou strategii. Tedy, první prvek v závorce na příslušné pozici matice představuje výplatu hráče A, 2. prvek výplatu druhého hráče B.

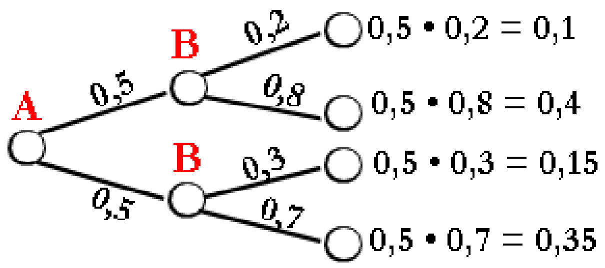
platí:

$$f_1(x_i, y_j) = x_{ij}$$

$$f_2(x_i, y_j) = y_{ij}$$

Pozn.: Kromě her v normálním tvaru se můžeme bavit o **hrách v rozvinutém (explicitním) tvaru**, kdy hráči hrají za sebou, tah po tahu, takže se navzájem ovlivňují, na rozdíl od her v normálním tvaru, kde se hráči rozhodují vždy ve stejný okamžik. Tahy se dají zakreslit do *tzv. stromu hry*. Každá větev stromu má přiřazenou pravděpodobnost, s jakou hráč vybere právě tuto možnost, vycházející z konkrétního uzlu.

[6] Tato rovnost platí pouze pro hry s nulovým součtem



Obrázek 1: Příklad stromu hry

Def.(dominance) [9]***:*** *Strategie S dominuje strategii T, pokud výplaty hráčů v případě volby strategie S jsou nejméně tak vysoké jako výplaty, které by získali hráči, kdyby zvolili strategii T a nejméně pro jednoho z hráčů je výplata v případě zvolené strategie S vyšší, než kdyby hráči zvolili strategii T.*

Princip dominance [9]***:*** Racionální hráč nikdy nebude volit dominovanou strategii, tedy takovou, která má všechny výplaty horší než některá další strategie.

Hlavním cílem u teorie her je najít optimální strategii pro hráče, tedy doporučení, jak hru hrát za účelem dosažení co nejlepšího výstupu. Princip dominance nám pomůže zjednodušit některé úlohy, ve kterých se vyskytují dominované strategie, což si můžeme ukázat na tomto příkladu.

Příklad 1: Máme matici 3×3. Naším cílem je podívat se, zda se v matici nevyskytují nějaké dominované strategie pro hráče a pokud ano, pak zredukovat dimenzi matice.

$$\begin{pmatrix} 1 & 3 & 4 \\ 3 & 2 & 4 \\ 2 & 7 & 6 \end{pmatrix}$$

Navržený postup řešení: Jedná se o hru s nulovým součtem, tedy, pro hráče A, jehož strategie jsou řádky matice, budeme upřednostňovat co největší hodnoty, pro hráče B, jehož strategie jsou sloupce matice, budeme upřednostňovat co nejmenší hodnoty, neboť jemu připadají výplaty opačné, tedy pro tuto matici výplaty se záporným znaménkem. Pokud se podíváme blíže na matici, můžeme si všimnout, že $x_{11} < x_{31} \wedge x_{12} < x_{32} \wedge x_{13} < x_{33}$ [7], tedy 1. řádek matice je dominován

[7] Mezi prvky matice nemusí být ostré nerovnosti, může se jednat i o neostré nerovnosti.

třetím řádkem a můžeme ho z matice vyškrtnout, jelikož racionální hráč by nikdy nevolil strategii, která je horší než některá jiná pro všechny možné volby strategie protihráče. Dostáváme matici:

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & 7 & 6 \end{pmatrix}$$

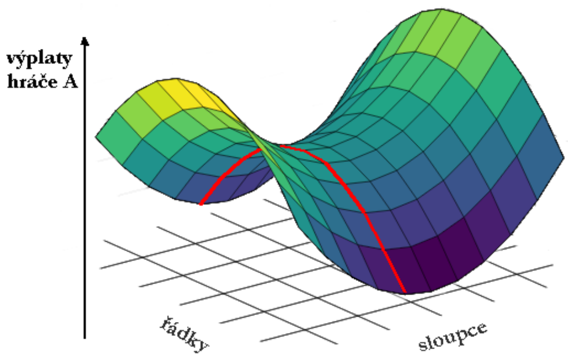
Nyní je na čase podívat se na hráče B, tedy na sloupce matice, zda se i u něj nevyskytují nějaké strategie, které jsou dominovány jinými. Výplaty hráče B by nyní tedy mohly být $\{-3, -2, -4, -7, -6\}$. V matici máme opačné hodnoty, proto nás momentálně budou zajímat co nejnižší kladná čísla, která zajistí hráči B co nejmenší prohru. Vidíme, že sloupec 3 je v tomto případě dominován 1. sloupcem a proto 3. sloupec z matice vyškrtneme. Dostáváme tuto redukovanou matici:

$$\begin{pmatrix} 3 & 2 \\ 2 & 7 \end{pmatrix}$$

Zde už se nenachází žádné dominované strategie, a tudíž matice už nemůže být více zredukována a použije se v této podobě pro hledání optimálních strategií hráčů. Pro případ hry dvou hráčů s nulovým součtem bude řeč o *sedlových bodech.*

Def. (Sedlový bod) [9]***:*** *Výstup v maticové hře (s výplatami hráče, jehož výplaty jsou zapisovány po řádcích*[8]*) je nazýván sedlovým bodem, jestliže hodnota na této pozici matice je menší nebo rovna všem ostatním hodnotám v tomto řádku a zároveň větší nebo rovna v porovnání s ostatními hodnotami v sloupci, ve kterém se tato porovnávaná hodnota nachází.*

Princip sedlového bodu [9]***:*** *Pokud má matice sedlový bod, pak by hráči měli hrát strategii, která zahrnuje tento sedlový bod.*



Obrázek 2: Sedlový bod

[8] Bereme v potaz hry s nulovým součtem

Termín „sedlový bod“ vychází z podmínky – hodnota nejmenší v řádku a největší ve sloupci. Pokud nakreslíme 3D graf – oblast okolo sedlového bodu připomíná sedlo. Pokud se v matici nachází více sedlových bodů, potom mají všechny stejnou hodnotu a nachází se v rozích pomyslného obdélníku.

Může se také stát, že hra s nulovým součtem nebude mít ani jeden sedlový bod. V takovém případě bychom nemohli hledat řešení hry v čistých, nýbrž ve smíšených strategiích, které si zde ale popisovat nebudeme, jelikož se netýkají tématu této práce.

Rovnovážný bod [9]*, často nazýván také jako* ***Nashova rovnováha*** *(u her s nenulovým součtem) koresponduje se sedlovými body u her s nulovým součtem. Jedná se o situaci, kdy si žádný z hráčů nepřilepší tím, že by se rozhodl zvolit jinou strategii než tu, která je pro něj rovnovážnou.*

Paretovo optimum [9] je pojmenováno po italském ekonomovi Vilfredovi Paretovi. V teorii her je chápán výstup hry jako paretovsky optimální v případě, že ve hře není žádný jiný výstup, který by byl pro všechny hráče minimálně stejně dobrý a alespoň pro jednoho hráče lepší, tj. pokud by se někteří hráči rozhodli odchýlit od tohoto bodu a získali vyšší výplatu, pak by to bylo na úkor jiných hráčů, kterým by výplata klesla.

Běžně užívané zkratky v souvislosti s vězňovým dilematem:

- **C** – z angl.. „cooperation“ – spolupráce
- **D** – z angl.. „defection“ – zrada
- **PD** – z angl. „prisoner’s dilemma“ – – vězňovo dilema
- **IPD** – z angl. „iterated prisoner’s dilemma“ – – opakované vězňovo dilema

Pozn.: Často rovnovážný bod hry není *paretovsky optimální*, neboť bývá často dominován jiným bodem, jako např. u vězňova dilematu – bod (D, D) [9] je dominován bodem (C, C). Bod (C, C) je paretovsky optimální, ale rovnovážným bodem je (D, D).

[9] (D, D) – hráči se vzájemně zradili, 1. prvek dvojice odpovídá 1. hráči (hráči A), 2. prvek 2. hráči (hráči B)

2 Vězňovo dilema – hra bez opakování

Vězňovo dilema je problematika z oblasti teorie her. V předchozí kapitole jsme si hry rozdělili do několika kategorií. Jedním z nejznámějších problémů nekooperativních her s nenulovým součtem, které herně-teoretické modely řeší, je vězňovo dilema.

Vězňovo dilema je již od samotného vzniku záhadou. Tisíce matematiků, psychologů, politologů, filozofů a ekonomů se zamýšlelo nad řešením tohoto problému. Přesto zůstává stále stejně tajemné jako v roce 1950, kdy tento rozhodovací problém vznikl. Za tvůrce lze považovat Merrill Floodovou a Melvina Dreshera z RAND Corporation, významné americké výzkumné stanice, kteří se zabývali výzkumy z oblasti teorie her. Jméno dal tomuto problému ale až Albert W. Tucker (Albert Tucker, 1950 – E. Rasmusen popisuje ve svém díle [8]), který ho roku 1951 prezentoval ve formě krátkého detektivního příběhu [5].

A i dnes je tento rozhodovací problém vysvětlován příběhem, nejčastěji s dvěma podezřelými (vězni), kteří mají možnost svým rozhodnutím, které může být buď u výslechu mlčet (spolupracovat s druhým podezřelým) nebo mluvit (zradit druhého podezřelého), rozhodnout o „svém vlastním osudu“. Pokud se rozhodne jeden podezřelý mluvit a druhý mlčet, pak ten, co mluvil (udal druhého podezřelého), bude volný a ten, co mlčel, si odsedí několik let ve vězení. V případě, že budou oba vězni mluvit, odsedí si příslušný počet let ve vězení oba[10], v opačném případě, tedy když budou oba mlčet, půjdou do vězení rovněž oba, ale na menší počet let, než když by vězni spolu spolupracovali. Jak by se měli podezřelí zachovat, aby ve vězení strávili co nejmíň let, už bude popsáno dále.

V praxi se ale vězňovo dilema většinou chápe jako **hra dvou hráčů,** kteří za svá rozhodnutí **obdrží** příslušnou **výplatu**, nikoli počet let, který mají strávit ve vězení. Vzhledem k tomu, že předpokládáme, že se hráči chovají racionálně, je jejich cílem maximalizovat své skóre. Hráči získávají výplaty podle předem stanovené výplatní dvojmatice (výplatní tabulky), na každé pozici matice reprezentuje první člen výplatu prvního hráče, druhý člen výplatu druhého hráče. Výplatní tabulky mohou být pro různé hry různé, jen musí být pro jejich členy splněny určité podmínky, které budou představeny dále.

Nicméně protože je v posledních letech vězňovo dilema velmi aktuální téma, které zajímá spoustu vědců, uvažují se např. i hry více než dvou hráčů a více možností voleb hráče, než je jen spolupráce a zrada. V této kapitole autorka vycházela především ze zdrojů [2] a [9].

[10] Vycházíme z obecné výplatní tabulky (**Tabulka 2**), ve které musí být splněny určité podmínky **mezi výplatami jednotlivých hráčů**

2.1 Vězňovo dilema – obecná definice problému

Jak již bylo zmíněno výše, vězňovo dilema bývá nejčastěji vysvětlováno na příkladu dvou podezřelých (vězňů), jejichž cílem je odsedět si ve vězení co nejmenší počet let, nicméně chápáno bývá vězňovo dilema většinou jako hra dvou hráčů, kteří za svá rozhodnutí[11] obdrží příslušné výplaty, které se snaží maximalizovat. Modelový příklad byl popsán v samotném úvodu, nyní se podíváme na podmínky hry a také na nalezení optima hry.

Vězňovo dilema je obecně hra 2 hráčů daná následující tabulkou:

		Hráč B	
		Spolupráce	**Zrada**
Hráč A	**Spolupráce**	(R, R)	(S, T)
	Zrada	(T, S)	(P, P)

Tabulka 2: Obecná tabulka výplat (převzato z [2])

nebo analogicky, ve formě dvojmatice:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (R,R) & (S,T) \\ (T,S) & (P,P) \end{pmatrix} \end{array}$$

kde $R, S, T, P \in \mathbb{R}$ jsou libovolné konstanty, které, aby se dalo hovořit o vězňově dilematu, musí splňovat dvě následující podmínky [2]:

1.) $T > R > P > S$

2.) $R > \frac{(S+T)}{2}$

kde R je zkratka z anglického „Reward“ – odměna za společnou spolupráci, S „Sucker payoff“ – nulová výplata za oškubání se protihráčem, T „Temptation“ – pokušení za zradu spolupracujícího soupeře, P „Punishment“ – potrestání za vzájemnou zradu.

[11] Mezi spoluprací a zradou

Když se oba hráči rozhodnou spolupracovat, pak za vzájemnou spolupráci obdrží odměnu R. Pokud jeden z hráčů zradí a druhý spolupracuje, pak zrazující hráč obdrží výplatu T a hráč, který spolupracoval, ale byl zrazen, obdrží nulovou výplatu S. Jestliže ale oba hráči zradí, poté získají výplatu P, což značí trest pro oba hráče. Přestože se výplatní tabulky mohou lišit, obecně vypadají jako **Tabulka 2**.

Proč musí být splněny tyto nerovnosti mezi konstantami R, S, T, P pro vězňovo dilema?

1.) Pokud by např. hráč obdržel vyšší výplatu při spolupráci než při oškubání[12] soupeře, pak by hráči neměli důvod se zrazovat. Naopak, pokud by obdrželi více za vzájemnou zradu než spolupráci, pak by se neustále bez přemýšlení zrazovali. Takto bychom si mohli jednoduše odůvodnit význam všech nerovností.
2.) Druhá nerovnost musí být splněna proto, aby hráči byli motivováni spolupracovat. Při porušení této nerovnosti by se hráčům vyplatilo se střídat tak, že by v jednom kole zrazoval 1. hráč a spolupracoval 2. hráč, ve 2. kole 1. hráč spolupracoval a druhý zrazoval atd. Průměrná výplata za zradu spolupracujícího a spolupráci se zrazujícím tedy musí být menší než výplata za vzájemnou spolupráci.

2.2 Optimum hry

Vraťme se nyní ještě na chvíli k modelovému příkladu, který byl uveden v úvodu práce. Jak by se měl podezřelý zachovat, aby ve vězení strávil co nejkratší dobu a zároveň nezapomínal ale na to, že takto bude uvažovat i druhý podezřelý?

Předpokládejme, že si podezřelý A bude myslet, že jeho oponent B bude spolupracovat (tj. bude mlčet). Pokud bude spolupracovat i sám A, odsoudí oba podezřelé na 1 rok (viz **Tabulka 1** v úvodu). Pokud A zradí, tj. obviní z činu podezřelého B, bude A volný a B si odsedí ve vězení 10 let, tudíž pokud si bude A myslet, že B bude spolupracovat, měl by A zradit. Ale co když si A bude myslet, že ho podezřelý B zradí? Pokud on sám bude spolupracovat, tj. mlčet, pak si odsedí ve vězení 10 let. Pokud zradí (tj. oba se vzájemně obviní), pak oba půjdou sedět na 5 let, tedy, opět by měl A zradit. Odtud vyplývá, že by měl podezřelý zradit bez ohledu na to, jak se zachová druhý podezřelý. Samozřejmě, stejně uvažujeme i v případě druhého podezřelého.

A tedy, oba podezřelí by se měli vzájemně zradit, tedy udat druhého podezřelého. Pokud se ale A a B zradí, půjdou si odsedět 5 let, přestože by pro ně bylo výhodnější spolupracovat (mlčet)

[12] Hráč zrazuje spolupracujícího soupeře – „oškubává“ ho.

a jít sedět jen na 1 rok [2]. Bod (zrada, zrada) je rovnovážným bodem této hry. Pokud by se jeden z podezřelých rozhodl namísto zrady spolupracovat, pak by si místo pěti let ve vězení odseděl 10 let, zatímco druhý podezřelý by byl volný. Hráč, který by se rozhodl odchýlit[13] od zrady, by si nepřilepšil, ale naopak by svou situaci zhoršil. Tedy, bod (zrada, zrada) je rovnovážným bodem vězňova dilematu bez opakování.

Můžeme zkusit ověřit, zda se v případě vězňova dilematu bez opakování jedná o paretovské optimum. Pokud bychom se opět podívali na **Tabulku 1,** můžeme vidět, že přestože je bod (zrada, zrada) rovnovážným bodem, není paretovským optimem, protože kdyby podezřelí spolu spolupracovali, odseděli by si ve vězení oba po jednom roku namísto pěti let, což by byla pro oba podezřelé lepší situace.

Proč tedy podezřelí spolu nespolupracují? Důvod je jednoduchý, nemohou si vzájemně důvěřovat, a proto raději sází na jistotu.

2.3 Výplatní tabulka

Jak už bylo zmíněno výše, vězňovo dilema je nejčastěji chápáno jako hra dvou hráčů, kteří za svou volbu spolupráce nebo zrady obdrží příslušnou výplatu.

Pokud hráč hraje vězňovo dilema, musí se rozhodnout, zda se svým protihráčem chce spolupracovat či nikoli. Hráč poté v závislosti na své a protihráčově volbě obdrží výplatu podle výplatní tabulky (dvojmatice). Ta se může lišit podle použité literatury (např.[2], [14], [18]), my ovšem budeme předpokládat hodnoty, které ve svých turnajích, popsaných dále, využil Robert Axelrod. Tyto hodnoty uvádí **Tabulka 3,** která byla poprvé představena v roce 1959 v článku [19], populární se stala v 80. letech 20. století, kdy ji Robert Axelrod využil ve svých turnajích v letech 1980 a 1984 (**podkapitola 3.2**), v současnosti popisována jako ta „klasická“ matice, která se pro vězňovo dilema využívá.

[13] Odchýlit se od strategie – zvolit jinou strategii se snahou dosažení lepšího výsledku

		Hráč B	
		Spolupráce	**Zrada**
Hráč A	**Spolupráce**	Oba hráči získají po 3 bodech.	Hráč A nezíská nic, hráč B získá 5 bodů.
	Zrada	Hráč A získá 5 bodů, hráč B nezíská nic.	Oba hráči získají po 1 bodu.

Tabulka 3: výplaty jednotlivých hráčů vězňova dilematu (v souladu s Axelrodovými turnaji)

Situaci prezentované v **Tabulce 3** odpovídá následující dvojmatice:

$$
\begin{array}{cc}
 & \begin{array}{cc} \;C & \quad D \end{array} \\
\begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (3,3) & (0,5) \\ (5,0) & (1,1) \end{pmatrix}
\end{array}
$$

Pokud tedy hráči hrají vězňovo dilema bez opakování (jednokolové vězňovo dilema), není obtížné se rozhodnout, kterou strategii zvolit. V takovém případě je totiž rovnovážným bodem bod (zrada, zrada). Můžeme se ale zkusit zamyslet, zda by se hráči měli tohoto doporučení držet i v případě, že si odehrají více kol vězňova dilematu než jen jedno, tedy hrají opakované vězňovo dilema. Tuto problematiku probereme v následující kapitole.

3 Opakované vězňovo dilema

V předchozí kapitole jsme si popsali, co to je vězňovo dilema bez opakování a jak by se měl hráč, který vězňovo dilema hraje, zachovat. Je však daleko zajímavější zkoumat, jak se budou hráči chovat, pokud si odehrají několik kol za sebou, tedy, opakované vězňovo dilema. V takovém případě už totiž nemusí být nejvhodnější zradit svého soupeře, a tedy často neexistuje jednoznačné doporučení pro hráče, jak dosáhnout nejvyššího skóre. Nicméně pokud hráč sleduje chování protihráče, může své skóre do jisté míry optimalizovat. Jelikož se jedná o opakované vězňovo dilema, hráči si mohou oplácet své předchozí tahy. U opakovaného vězňova dilematu je potřeba rozlišovat 2 hlavní situace, a to, zda je nebo není hráčům předem znám počet kol hry.

Pokud by hráči znali dopředu počet kol hry, pak zde platí to samé, co v případě vězňova dilematu bez opakování [2]. Optimální strategií pro oba hráče je zradit v každém kole hry. Toto doporučení vyplývá z toho, že hráč by měl určitě zradit v posledním kole, neboť mu toto chování nemá jeho soupeř jak vrátit. Takto samozřejmě uvažuje i sám hráč, což znamená, že se oba hráči zradí v posledním kole hry a protože vědí[14], že je protihráč zradí v posledním kole tak či tak, rozhodnou se ho zradit už v předposledním kole. Opět toto chování za předpokladu, že se hráči chovají racionálně, můžeme očekávat i u protihráče. Touto úvahou se postupně dostaneme k tomu, že hráč by měl zradit už v prvním kole hry.

Nicméně, takováto úvaha neplatí pro hry, kdy hráči předem počet kol neznají, takže nevědí, kdy mohou zradit, aby jim jejich tah nemohl soupeř vrátit zpět. [2]

Uskutečnila se řada her a turnajů opakovaného vězňova dilematu, kde se ukázalo, že ochota hráčů spolupracovat se svým soupeřem byla většinou úplně minimální [2]. Vědcům z oblasti teorie her trvalo velmi dlouho, než nalezli takové strategie, které by přinutily hráče spolupracovat. Lze totiž sledovat, že výplaty hráčů přes nekonečný počet kol by byly vyšší v případě, kdyby hráči spolupracovali po celou dobu hry, než pokud by se hráči pokoušeli vzájemně zrazovat a doufali, že jim tyto zrady bude protihráč tolerovat. Hlavním zdrojem pro tvorbu této kapitoly byla kniha [3].

Pozn.: Nutno poznamenat, že většina turnajů a her opakovaného vězňova dilematu, které se uskutečnily, probíhaly způsobem, kdy lidé vytvořili a zaslali algoritmy a celá hra, popř. turnaj, probíhala už jen jako počítačová simulace.

Literatury, týkající se her vězňova dilematu s lidmi, je podstatně méně (viz např. [14], [17], [18], [19]).

[14] Předpokládáme, že hráči se chovají racionálně.

Ve zbytku kapitoly se budeme zabývat prvním případem, tedy naprogramovanými algoritmy. **Kapitola 5** se naopak zaměří na druhý případ, hry s lidskými hráči. Stejně tak se budeme hrami s lidskými hráči zabývat i v praktické části práce, protože jsou podstatně méně prozkoumanou oblastí.

3.1 Nejznámější algoritmy

V této části si popíšeme nejznámější algoritmy, které se vyskytovaly opakovaně v turnajích.

Výpis některých známých algoritmů a popis, jak fungují:

- **TFT** (Tit for tat) - „Oko za oko" - Začíná spoluprací a dále kopíruje protihráčův předchozí tah.
- **AllC** (Always cooperates) - Vždy spolupracuje.
- **AllD** (Always defects) – Vždy zrazuje.
- **RAND** (Random) – Strategie, která spolupracuje s pravděpodobností $\frac{1}{2}$. Nereaguje na volbu protihráče, její tahy jsou zcela náhodné.
- **Pavlov**[15] – Strategie, která spolupracuje, pokud v předchozím kole zvolili oba hráči stejnou strategii, jinak zradí. Přestože se většinou Pavlov řadí mezi milé strategie, tedy ty, které nezrazují jako první, lze se v některé literatuře setkat i s dělením na PavlovC, PavlovD, které se liší ve volbě spolupráce/zrady v prvním kole hry.
- **GRIM trigger** – Strategie, která začíná spoluprací, ale v případě, že ji zradí protihráč, zrazuje neustále[16].

[15] Často známé také jako „win-stay, lose-shift", což značí, že pokud hráč v předchozím kole „vyhrál", pak by měl hrát stejnou strategii, jako v předchozím kole, v opačném případě, tedy, pokud prohraje, tak by měl změnit strategii (buď ze spolupráce na zradu, nebo ze zrady na spolupráci)

[16] Nejčastěji se jedná o první zradu, ale slovo „trigger" znamená v překladu spouštěč, takže impulsem pro neustálé zrazování může být i nějaký větší počet zrad, než je 1, např. 10.

Za zmínku stojí i strategie, které vznikly modifikací tak úspěšné TFT strategie a jejich autoři doufali, že si povedou lépe než samotná TFT. Mezi ně patří např.:

- **GTFT**[17]- strategie funguje podobně jako TFT, nicméně s nějakou malou pravděpodobností (např. 0,1) odpouští zradu.
- **STFT**[18]– chová se jako TFT, kromě prvního kola, ve kterém protihráče zradí.
- **TFTT**[19]– Oproti TFT je tato strategie více odpouštějící, protože zrazuje až po dvou zradách protihráče. Výhodou je, že odpouští izolované zrady, které mohly např. vzniknout chybou programu (šumem). Naopak nevýhodou této strategie je to, že pokud protihráč tuto strategii identifikuje, může ji celkem snadno vykořisťovat tím způsobem, že bude střídavě zrazovat a spolupracovat s touto strategií.
- **TTFT**[20] - Tato strategie je oproti předchozí strategii (i oproti klasickému TFT) „přísnější“, za jednu zradu protihráče zradí dvakrát.

  Pozn.: Spousta tvůrců algoritmů se domnívalo, že toto zpřísnění TFT, které tkví v menším odpouštění zrad, bude mít pozitivní vliv na výplaty této strategie. U výsledků některých turnajů, viz např. [2], kdy byl 10krát opakován turnaj s 10 algoritmy, se ale ukázal pravý opak, tedy, že čím je TFT víc zpřísněno, tím horších výsledků dosáhne. Lépe než TFT si v tomto turnaji vedla jeho více odpouštějící verze TFTT a naopak hůře přísnější verze TTFT (více v této knize na str. 97 – tabulka 4.5). Rozhodující je vždy samozřejmě to, v prostředí jakých algoritmů se daný algoritmus vyskytuje.

Pozn.: Kromě těchto jednoduchých strategií existují i složitější, které využívají celou historii protihráčových tahů a na jejímž základě se snaží odhadnout, o jaký typ protihráče se jedná. Dvě nejznámější z nich budou popsány v **podkapitolách 4.3** a **4.4.**

---

[17] Generous tit for tat – „Velkorysé oko za oko“
[18] Suspicious tit for tat -„Podezíravé oko za oko“
[19] Tit for two tats - „Oko za dvě oka“
[20] Two tits for tat – „Dvě oka za oko“

3.2 Turnaje Roberta Axelroda

V roce 1984, Robert Axelrod (* 27. 5. 1943), významný americký politolog a profesor, působící na univerzitě v Michiganu, uveřejnil ve své knize Evolution of Cooperation výsledky turnajů, týkajících se opakovaného vězňova dilematu, kde vyzval experty z oblasti teorie her k zaslání počítačových algoritmů. Tato kniha [1] se stala bohatou inspirací pro mnoho výzkumů z oblasti teorie her a také vědcům při psaní jejich vědeckých článků. Získala si celou škálu čtenářů, od špičkových vědců až po širokou veřejnost.

Vězňovo dilema a opakované vězňovo dilema se staly bohatým zdrojem výzkumných prací od 50. let 20. století. Nicméně, publikování Axelrodovy knihy [1] v 80. letech 20. století mělo také velký vliv na rozšíření této problematiky i do jiných oblastí, než je jen samotná teorie her, jako je např. evoluční biologie, informatika, psychologie atd. V následujících dvou podkapitolách si představíme dva turnaje opakovaného vězňova dilematu, které uspořádal právě Robert Axelrod.

Turnajem zde máme na mysli počítačovou simulaci, kdy soupeří proti sobě počítačové algoritmy, přihlášené do turnaje jejich tvůrci.

3.2.1 První turnaj

V roce 1979 uspořádal Robert Axelrod turnaj, týkající se opakovaného vězňova dilematu a vyzval některé experty z oblasti teorie her k navržení svých strategií, počítačových algoritmů (programů). Tyto algoritmy fungovaly tak, že v každém kole hry volily mezi spoluprací a zradou, ať už s přihlédnutím na historii hry (poslední kolo nebo více předchozích kol) anebo se algoritmus rozhodoval bez ohledu na volbu protihráče jako např. AllD (z anglického „Always Defect“), který zrazuje v každém kole hry.

Axelrod obdržel 14 strategií a do turnaje zahrnul jednu extra, která byla založena na tom, že zrazovala a spolupracovala se stejnou pravděpodobností, $\frac{1}{2}$, která je známá pod zkratkou RAND[21]. Strategie byly postaveny každá proti každé, včetně sebe samé a každá hra měla přesně 200 kol. Poznamenejme ještě, že ve skutečnosti byl turnaj pro dosažení stabilnějších odhadů skóre opakován pětkrát a výsledky byly nakonec zprůměrovány. Celkem tedy Axelrod získal 240 000 dat – jednotlivých voleb hráčů. Autory navrhovaných PC strategií byli profesoři politologie, matematiky, psychologie, informatiky a ekonomie.

Vítězem, který dosáhl nejvyššího průměrného skóre[22], se stal překvapivě velmi jednoduchý algoritmus, známý pod zkratkou TFT, z anglického „Tit for tat“, do češtiny překládán jako „Oko

[21] z anglického „Random“ neboli náhodný
[22] Skóre se průměrovalo z pěti opakování turnaje

za oko“, který navrhl Anatol Rapoport, profesor psychologie na univerzitě v Torontu. Tato strategie funguje na principu, že v prvním kole hry vždy spolupracuje a poté se chová stejně, jako její protihráč v předchozím kole.

Na základě pozorování, které strategie si v turnaji vedly dobře, bylo vypozorováno několik společných vlastností úspěšných strategií:

- **Milé** – Nikdy nezrazují jako první.
- **Odpouštějící** – Oplácí zradu, ale vrací se ke spolupráci, pokud protihráč nepokračuje ve zradě, mají tzv. „krátkou paměť“.
- **Nezávistivé** – Jde jim o vlastní zisk, nikoli o porážku soupeře.
- **Vyprovokovatelné** – Nechají se vyprovokovat zradou, tedy nejsou až přespříliš optimistické jako např. strategie známá pod zkratkou AllC (Always Cooperates), která spolupracuje za jakýchkoli podmínek. Jinak řečeno, nenechají se „oškubávat“ strategiemi, které nejsou milé.

3.2.2 Druhý turnaj

Druhý turnaj překonal první v počtu algoritmů, kdy jich R. Axelrod obdržel 62, což naprosto předčilo veškerá jeho očekávání. Tentokrát měli zúčastnění výhodu, protože si mohli prostřednictvím různých článků v počítačových časopisech nastudovat tuto problematiku a také obdrželi výsledky prvního turnaje[23]. Pozváni byli i účastníci prvního turnaje, jinak byli různí, od desetiletého nadšence do počítačů, až po profesory informatiky, fyziky, ekonomie, psychologie, matematiky, sociologie, politologie a evoluční biologie. Přijeli celkem ze šesti zemí, Spojených států, Kanady, Velké Británie, Norska, Švýcarska a Nového Zélandu. Druhý turnaj měl za cíl jednak ověřit si tvrzení, která vyvodil Axelrod po skončení prvního turnaje a také hledat nové důvody, které vysvětlují úspěšnost nebo naopak selhání některých algoritmů. Hlavním cílem bylo pokusit se najít strategii, která si v turnaji povede lépe, než si vedla TFT v prvním turnaji.

Včetně algoritmu RAND tedy proti sobě soupeřilo 63 algoritmů, každý s každým, čímž Axelrod získal přes milion tahů jednotlivých hráčů. Z 63 algoritmů bylo 39 milých[24], většina z nich také odpouštějících[25].

[23] Všichni zúčastnění obdrželi ještě před turnajem zprávu, kde mohli vidět, jak si vedly algoritmy v prvním turnaji.
[24] Nikdy nezrazují jako první.
[25] Odpouštějící – Oplácí zradu, ale vrací se ke spolupráci, pokud protihráč nepokračuje ve zradě, mají tzv. „krátkou paměť“

Trochu zarážející po skončení turnaje byl samotný fakt, že stejně jako v prvním turnaji, zvítězila nejjednodušší strategie, TFT a i přesto, že o jejím úspěchu v prvním turnaji všichni zúčastnění věděli, se jim nepodařilo nalézt žádnou strategii, která by jí dokázala konkurovat.

Pozn.: Zatím jsme se seznámili jen s typem her, kdy hráči hrají hru „každý s každým". Od dob Axelroda se ale vyvíjejí dva typy přístupů, které testují robustnost[26] strategií a dále odvozují optimální strategie:

(1) **Round - robin turnaje** – „každý proti každému"
(2) **Evoluční přístup** – Některé strategie se rozmnožují na úkor druhých, tj. ty slabší vymírají, ty lepší se rozšiřují, až je dosaženo stability.

Pokud je postavena každá strategie proti každé další, pak lze po skončení hry ukázat, jakou výplatu získala za celou hru v porovnání s ostatními strategiemi, zatímco evoluční přístup ukazuje strategie z jiného úhlu pohledu, a to z hlediska počtu potomků nebo přežití v určitém prostředí. V rámci těchto přístupů bylo vyvinuto a analyzováno mnoho nových strategií [2].

Pozn.: Zajímavý způsob optimální interakce se soupeřem se podařil týmu ze southamptonské univerzity vedených profesorem **N. Jenningsem** na jednom z turnajů opakovaného vězňova dilematu z roku 2014[27]. Jak už víme, hráči vězňova dilematu mezi sebou nemohou komunikovat. Přesto se ale tomuto týmu podařilo vyhrát turnaj způsobem, kdy spolu hráči vzájemně komunikovali. Tento tým představil skupinu strategií, které fungovaly tak, že se dokázaly pomocí 5-10 specifických tahů na začátku hry rozpoznat a dále spolu optimálně interagovat. Pokud se tyto strategie rozpoznaly po několika prvních tazích, poté se jedna začala chovat jako pán a druhá jako otrok. Pán vždy zrazoval a otrok vždy spolupracoval. Tím si pán v takovýchto hrách přišel na velmi pěkné průměrné skóre za hru. Pokud strategie zjistila, že nehraje proti některé další ze skupiny, pak vždy zrazovala. Tyto skupinové strategie obsadily první 3 příčky na tomto turnaji a tím porazily všechny neskupinové strategie[28].

[26] Robustnost = invariantnost vůči malým odchylkám.
[27] K 20. výročí her opakovaného vězňova dilematu
[28] Tedy např. i včetně strategie TFT

4 Identifikace protihráče a optimální interakce s ním

Představme si, že by hráč v IPD (často užívaná zkratka pro opakované vězňovo dilema – – z angl. „iterated prisoner's dilemma") věděl, jak se budou chovat všichni jeho protihráči. Pak by vždy zrazoval strategie jako AllC a AllD a naopak spolupracoval se strategiemi jako je GRIM nebo TFT za účelem maximalizace výplat[29]. Jinými slovy, volil by optimální protistrategii v závislosti na tom, proti komu by hráč byl postaven. Přestože ale hráč předem nezná strategii soupeře, může se pokusit ji během hry identifikovat.

Např., pokud strategie spolupracovala se svým oponentem v předchozích 10 kolech, zatímco jeho soupeř stále zrazoval, lze z tohoto chování usuzovat, že bude strategie spolupracovat i nadále. Dá se tedy očekávat, že se bude jednat o AllC a ideálně za účelem co nejvyššího skóre by měl oponent takového hráče zrazovat i po zbytek hry. Právě z důvodu identifikace protihráče za účelem maximalizace výplat mají některé algoritmy v sobě zabudovaný identifikační mechanismus, který umožňuje hráči určit chování jeho protihráče, a tedy optimálně s ním interagovat[30]. Každý algoritmus identifikuje chování soupeře odlišným způsobem. Kromě obecného popisu identifikace protihráče si představíme ve stručnosti také dvě strategie, které identifikační mechanismus využívají. Hlavními zdroji pro vytvoření této podkapitoly byly knihy [1] a [2].

Identifikace protihráče a následná optimální interakce s ním se jeví jako velmi dobrá schopnost některých algoritmů, bohužel, s analýzou strategií jsou však spojeny také problémy, které je potřeba vzít v úvahu. Největšími nedostatky identifikace jsou tyto:

1.) Identifikovat lze pouze strategie spadající do nějakého předem dohodnutého konečného souboru strategií (např. AllC, AllD, Pavlov atd.). Tento konečný soubor zahrnuje ale jen malou část všech možných strategií, což je problém. Na druhou stranu, čím je algoritmus propracovanější a čím více protistrategií dokáže identifikovat, tím déle[31] mu bude trvat tyto strategie rozpoznat, takže identifikace pak nebude mít žádný, nebo jen minimální, účinek. Všechny strategie ale identifikovat ani nelze z toho důvodu, že stále vznikají nové a nové strategie.

2.) Existuje zde určitá šance, že se strategie s identifikačním mechanismem připraví díky identifikaci o spolupráci s některými protistrategiemi. Aby strategie dokázala identifikovat

[29] Jak už bylo zmíněno, tohle ale zvládne jen hráč, který je schopen měnit chování během hry.
[30] Nemusí se vždy ale nutně jednat o optimální chování, jak uvidíme dále.
[31] Ve smyslu počtu kol, které stráví sbíráním dat potřebných pro identifikaci.

svého soupeře, často ho musí také zradit. Například, za účelem rozlišení AllC od GRIM[32] by musel protihráč nejméně jedenkrát zradit, čímž by ale v případě, že by se jednalo o strategii GRIM, připravila o možnost vzájemné spolupráce s touto strategií v budoucnu. Tzn., taková akce může mít negativní efekt na výplaty hráče v budoucnu.

3.) Problémovými strategiemi pro strategie s identifikačním mechanismem jsou strategie, které mohou své chování kdykoli měnit. Dobrým příkladem může být třeba strategie Adaptivní Pavlov, jíž se budeme věnovat v **podkapitole 4.3**.

Pozn.: V praktické části práce se budeme také zabývat identifikací protivníka, nicméně je zde rozdíl oproti identifikaci, jež je nejčastěji zkoumána v literatuře. Obvykle je identifikací myšlena situace, kdy se hráč snaží odhadnout chování soupeře z toho důvodu, aby mohl s ním co nejlépe interagovat a maximalizoval tak svou výhru. Identifikace v praktické části je prováděna zpětně, díváme se na ní z role pozorovatele turnaje. První situace umožňuje hráči chovat se tak, aby byl schopen identifikovat své soupeře, např. zkusí ho zradit a bude zkoumat jeho reakci na tuto zradu. Takové možnosti ale my jako pozorovatelé nemáme, a proto je tato „zpětná identifikace" hráčů náročnější. Některé informace, které by se nám pro identifikaci hodily, bohužel často nejsou dostupné, nejsme schopni ovlivnit chování hráčů za účelem získání všech potřebných dat. V literatuře se tato situace hojně nevyskytuje, ale přesto je prakticky velmi užitečná, neboť tato identifikace může posloužit hráčům např. při zopakování tohoto turnaje se stejnými hráči. Získáme cenné informace o chování hráčů, jejich slabých nebo naopak silných stránkách, které by mohly být dále hráčům velmi užitečné.

4.1 Hledání optimální protistrategie ke strategii soupeře

Ve hře IPD[33] optimální strategie závisí na strategiích možných oponentů. Za předpokladu, že by hráč znal algoritmus/strategii protihráče, pak by si mohl odvodit, jak hrát, aby dosáhl co nejlepších výsledků. Pokud by např. věděl, že se jeho protihráč bude chovat jako AllC, pak by se měl hráč chovat jako AllD a neustále ho zrazovat za dosažením co nejvyšší výplaty.

Protože strategie AllC, AllD a RAND nezávisí na chování protihráče, jejich optimální protistrategií je chovat se jako AllD. Vzhledem k tomu, že GRIM[34], TFT, STFT[35] a TTFT se

[32] Jak už bylo zmíněno, GRIM je strategie, která po jedné zradě zrazuje neustále.
[33] Velmi často užívaná zkratka pro opakované vězňovo dilema (z angl.. „iterated prisoner's dilemma")
[34] Strategie, která začíná spoluprací, ale v případě, že ji zradí protihráč, zrazuje neustále
[35] TFT začínající hru zradou

pomstí, jakmile je soupeř zradí, optimální strategie pro soupeře je vždy spolupracovat kromě posledního kola, ve kterém zradí[36]. TFTT je velkorysejší a odpouští jednokolovou zradu, její oponent tedy může maximalizovat výplatu střídavou volbou zrady a spolupráce. Pokud Pavlov začíná spoluprací, pak jeho soupeř by měl vždy spolupracovat kromě posledního kola; v opačném případě by soupeř měl začít zradou a pak vždy spolupracovat kromě posledního kola. **Tabulka 4** ukazuje optimální protistrategii ke každé strategii z těchto devíti vybraných strategií.

Strategie	Doporučení pro protihráče (optimální protistrategie)
AllC	Vždy zraď. (AllD)
AllD	Vždy zraď. (AllD)
RAND	Vždy zraď. (AllD)
GRIM	Vždy spolupracuj kromě posledního kola, ve kterém zraď.
TFT	Vždy spolupracuj kromě posledního kola, ve kterém zraď.
TFTT	Začni zradou a poté střídej spolupráci a zradu po 1 kole.
STFT	Vždy spolupracuj kromě posledního kola, ve kterém zraď.
TTFT	Vždy spolupracuj kromě posledního kola, ve kterém zraď.
Pavlov	Pokud tě Pavlov na začátku zradil (PavlovD), pak začni také zradou a poté spolupracuj kromě posledního kola, ve kterém zraď. Pokud Pavlov začal spoluprací (PavlovC), pak vždy spolupracuj kromě posledního kola, ve kterém zraď.

Tabulka 4: Doporučené chování pro hráče proti známé strategii (převzato z [2])

4.2 Hledání optimální protistrategie v populaci několika algoritmů / strategií

Během turnaje IPD hraje hráč většinou s více než jedním protihráčem, a tak je na místě si uvědomit, že by naše strategie (či algoritmus) měla optimálně interagovat se všemi ostatními strategiemi.

Pokud se jedná o lidského hráče, který je schopen měnit strategii během hry, pak např. v případě, že by jeho soupeři byli hráči, kteří se chovají jako některá strategie z **Tabulky 4** a hráč by je byl schopen identifikovat, pak by se měl držet doporučení z tabulky, aby optimálně interagoval

[36] V případě, že by znali předem počet kol hry

s každým z nich a dosáhl tak co nejvyššího skóre. Některé algoritmy mají v sobě zabudovaný identifikační mechanismus, který jim pomůže rozpoznat[37] protihráčovu strategii a optimálně s ní interagovat.

Může být užitečné vědět dopředu, kdo se turnaje účastní - podle toho se mohou hráči rozhodnout, jakou strategii zvolit, aby v celkovém součtu skóre turnaj vyhrála.

Strategie Always Cooperate (AllC) je dominována[38] strategií Always Defect (AllD), a tedy AllD by byla optimální strategií u populace[39] sestávající pouze ze strategií AllD a AllC.

Nicméně, v populaci sestávající z AllD, AllC a TFT, nemusí být AllD nutně optimální strategií. V takovéto situaci bude záležet také na počtu jednotlivých strategií, vyskytující se v populaci. AllD sice zvítězí ve všech hrách s AllC i TFT a bude remízovat ve hrách s ostatními AllD, ovšem AllC a TFT mezi sebou udrží spolupráci.

Pozn.: Kromě strategií, které zohledňují jen poslední tah (nebo posledních pár tahů) se můžeme setkat se strategiemi propracovanějšími, které zohledňují více tahů protihráče, někdy i celou historii hry. Asi není nutno vysvětlovat, že složitější strategie nemusí nutně znamenat lepší, ale na druhou stranu je pravda, že některé z nich jsou vcelku dobře propracované a v turnajích, po vzoru turnajů Axelrodových, se jim vedlo daleko lépe než např. strategii TFT, vítězi obou turnajů R. Axelroda.

V následujících dvou podkapitolách si představíme dvě takové úspěšné strategie. Ty fungují na principu identifikace soupeře, jsou však od sebe odlišné způsobem, jakým identifikaci provádí. Hlavním zdrojem pro tvorbu těchto podkapitol byla kniha [2].

4.3 Adaptivní Pavlov

Algoritmus **Adaptivní Pavlov** uplatňuje jednoduchý mechanismus pro odlišení spolupracujících strategií a několika málo reprezentativních zrazujících strategií a podle této identifikace uzpůsobuje své chování.

Algoritmus Adaptivní Pavlov (APavlov) se poprvé objevil na turnaji v r. 2005, který kopíroval původní turnaj R. Axelroda. Konkrétně to bylo v jedné části turnaje, zvané „Competition 4“. Přihlášeno bylo celkem 50 hráčů včetně 8 standardních strategií. Autorem tohoto algoritmu je Jiawei Li, čínský učitel informatiky a důvod, proč je zde zmíněn, je ten, že se umístil na prvním místě v této soutěži.

[37] Např. pomocí několika specifických tahů na začátku hry
[38] Být dominován jinou strategií – za jakékoli situace být horší než dominující strategie. Při hledání optima můžeme strategie, které jsou dominovány jinými vyřadit z možných optimálních strategií.
[39] Proti sobě je postaveno více algoritmů AllD a AllC

Princip algoritmu [25] :

APavlov funguje na principu rozdělení hry do období po šesti kolech jdoucích za sebou. Pro jednotlivá období v průběhu hry uplatňuje různé taktiky identifikace:

- V prvním období se APavlov vždy chová jako strategie TFT, tedy, začne spoluprací v 1. kole a po dalších pět kol kopíruje protihráčův předchozí tah.
- Na konci každého období APavlov klasifikuje svého protihráče do jedné z 5 kategorií: **kooperativní strategie[40], STFT[41], PavlovD, AllD** nebo **RAND** a pozmění svou strategii v novém období podle chování protihráče v období předchozím.

Klasifikace přitom probíhá následovně:

- Pokud protihráč **spolupracoval** v předchozím období po **celých 6 kol**, zařadí ho APavlov do skupiny **kooperativních strategií.**
- Pokud protihráč **zrazoval** APavlova v předchozím období **celých 6 kol**, pak je zařazen do kategorie **AllD**
- Pokud hrál postupně za sebou tahy **D, C, D, C, D, C**, pak je jeho chování vyhodnoceno jako **STFT**
- Pokud hrál postupně za sebou tahy **D, D, C, D, D, C**, pak je jeho chování vyhodnoceno jako **PavlovD**
- Pokud nenastala ani jedna z předchozích situací, pak je vyhodnocen jako **RAND**

Na základě výsledků této klasifikace zvolí APavlov optimální strategii pro další období:

- Jestliže byl protihráč v předchozím období klasifikován jako **RAND** nebo **AllD**, pak se APavlov bude v následujícím období chovat jako **AllD.**
- Pokud bylo chování protihráče klasifikováno jako **STFT**, pro další období se bude APavlov chovat jako TFTT.
- Pokud protihráč byl klasifikován jako **kooperativní strategie**, pak se v dalším období bude APavlov chovat jako **TFT.**

[40] Ta strategie, která ho v žádném kole v předchozím období nezradila.
[41]Strategie, která se v předchozím období chovala jako TFT kromě 1. kola období, kdy ho zradila.

Popsaná verze APavlov se také označuje **APavlov2006**. O 5 let později byl APavlov modifikován **APavlov2011** – při klasifikaci byla vypuštěna třída PavlovD a upraven způsob identifikace pro třídy AllD a STFT:

- 6 spoluprací / zrad ze strany soupeře v rámci 1 období – na konci období chování protihráče zařazeno do kategorie kooperativních strategií / AllD
- V případě 4 a více zrad ze strany protihráče za jedno období hry je jeho chování klasifikováno do kategorie **AllD**.
- V případě 3 zrad během 1 období je protihráčovu chování přiřazena kategorie STFT
- Jinak[42] je klasifikováno chování jako **RAND**

Pozn.: Optimální reakce APavlova na chování protihráče stejná jako pro APavlov2006, neboť rozřazuje chování do stejných kategorií[43].

Procesy interakce APavlov[44] se spolupracujícími strategiemi, AllD, STFT a PavlovD v prvních šesti kolech hry[45] jsou uvedeny na **Obrázku 3.**

Například, pokud dojde k takové interakci, jak je znázorněno na **Obrázku 3(c)**, bude soupeř označen jako STFT a APavlov bude v příštím období spolupracovat dvakrát za účelem dosažení spolupráce.

(a)

	1	2	3	4	5	6
APavlov	C	C	C	C	C	C
Kooper[46]	C	C	C	C	C	C

(c)

	1	2	3	4	5	6
APavlov	C	D	D	D	D	D
AllD	D	D	D	D	D	D

(b)

	1	2	3	4	5	6
APavlov	C	D	C	D	C	D
STFT	D	C	D	C	D	C

(d)

	1	2	3	4	5	6
APavlov	C	D	D	C	D	D
PavlovD	D	D	C	D	D	C

Obrázek 3: Identifikace soupeře strategií APavlov na základě 6 kol hry (převzato z [2]):
(a) APavlov spolupracuje s každou spolupracující strategií **(b)** Vždy zrazuje AllD
(c) Pokud strategie střídá D a C, zatímco je postavena proti TFT, pak ji APavlov vyhodnotí jako STFT.
(d) Pokud strategie střídá tahy D, D, C, zatímco hraje proti TFT, pak je vyhodnocena jako RAND.
Pokud je proces interakce odlišný od procesu na **Obrázku 3**, soupeř bude označen jako RAND.

---

42 Tj. při jedné nebo 2 zradách za období
43 S vynecháním PavlovD algoritmu
44 Hovoříme o APavlov2006
45 APavlov se tedy chová jako TFT.
46 Jakákoli spolupracující strategie

Pozn.:

- Na turnaji v praktické části práce byla použita tato novější varianta algoritmu APavlov
- Namísto úpravy některých parametrů, jak to dělá většina adaptivních strategií, používá APavlov identifikační mechanismus, který funguje jako expertní systém. Znalost různých soupeřů je vyjádřena pak formou *báze pravidel* „Pokud…, pak…“. Např.: „Pokud soupeř spolupracuje v 6 kolech, pak je vyhodnocen jako AllC“.
- Identifikace soupeře APavlov provádí v každém období během turnaje, aby se odstranila případná chybná identifikace a také, aby se APavlov uměl vypořádat s hráči, kteří během hry mění své strategie.
- Je zřejmé, že identifikační soubor adaptivních strategií může být rozšířen tak, aby zahrnoval více strategií, které lze identifikovat; avšak při rostoucím počtu prvků identifikačního souboru se zvýší počet výpočtů. Je třeba udělat kompromis mezi tím, kolik strategií je algoritmus schopen identifikovat a tím, jak rychle bude schopen chování protihráče identifikovat.

4.4 MyStrategy (MS)

Algoritmus *MyStrategy* (MS) používá identifikační mechanismus k identifikaci svého soupeře, ovšem trochu jiným způsobem.

Princip algoritmu:

MyStrategy předpokládá, že soupeř je jeden z následujících algoritmů (**TFT, AllC, AllD, STFT, PavlovC, PavlovD, TFTT, TTFT, GRIM** nebo **RAND**). Hru začíná zradou. Pokud se jeho soupeř v prvním kole rozhodne pro zradu, MS zvolí spolupráci v druhém kole, v opačném případě MS zvolí zradu. MS vždy ve třetím kole vybírá spolupráci. Tímto způsobem dokáže MS určit strategii z předem dané množiny strategií už po třech kolech hry.

Předpokládejme, že volby MS a jejího oponenta v prvních 3 kolech jsou uvedeny v **Tabulce 6**. Protože soupeř začíná hru zradou, musí to být jedna ze strategií AllD, STFT, RAND a Pavlov. Vzhledem k tomu, že MS zradil v prvním kole a soupeř ve 2. kole spolupracoval, nemůže se jednat ani o AllD, ani o STFT. Protože MS a soupeř spolupracují ve druhém kole, soupeř by neměl ve třetím kole zradit, kdyby byl Pavlov. Proto musí být protivník RAND. Optimální protistrategií k RAND je AllD, tudíž MS se bude chovat jako AllD v následujících kolech hry. Tímto způsobem lze strategii identifikovat po několika kolech hry a potom uplatnit optimální strategii. Způsob, jakým MS identifikuje protihráče, je ukázán v **Tabulce 5**.

Hráči	Možné tahy hráčů	Výsledek identifikace
MyStrategy protihráč	D C C D C C	**PavlovD**
MyStrategy protihráč	D C C D D D	**AllD**
MyStrategy protihráč	D C C D D C C	**STFT**
MyStrategy protihráč	D D C C C C C	**AllC**
MyStrategy protihráč	D D C C C D C	**TFTT**
MyStrategy protihráč	D D C C D C C	**PavlovC**
MyStrategy protihráč	D D C C C D D C C	**TFT**
MyStrategy protihráč	D D C C C C D D D C	**TTFT**
MyStrategy protihráč	D D C C C C D D D D	**GRIM**

Tabulka 5: Identifikace 9 strategií pomocí MyStrategy (převzato z [2], opraveny chyby)[47]

Pozn.: Pokud se protihráč nechová podle tabulky, pak MS přiřadí protihráči strategii RAND – – viz např. **Tabulka 6.**

V **Tabulce 6** vidíme, že protihráč začal hru zradou. Po prvním kole MS identifikuje, že se bude jednat buďto o nějaký z algoritmů, který začíná zradou (**PavlovD, AllD, STFT**), příp. o **RAND**. V druhém kole, protože protihráč začal zradou, MS zvolí spolupráci. Protihráč po vzájemné zradě rovněž spolupracoval, tedy určitě se nebude jednat o AllD ani STFT[48], může to být tedy buď **PavlovD**[49] nebo **RAND.** Ve 3. kole hry MS vždy volí spolupráci. Protihráč ale po předchozí vzájemné spolupráci MS zradí a tím pádem se nechová ani jako PavlovD. Výsledkem identifikace je tedy **RAND**.

Optimální protistrategií k RAND je AllD, tudíž MS se bude chovat jako AllD v následujících kolech hry. Tímto způsobem lze strategii identifikovat po několika kolech hry a potom uplatnit optimální strategii.

[47] Barevně rozlišeny tahy, které od sebe odlišují jednotlivé situace
[48] STFT se chová jako TFT, jen začíná hru zradou. Ve druhém kole by určitě oplatilo zradu.
[49] Pavlov, který začíná zradou, v ostatních kolech se chová klasicky-tedy, pokud oba hráči se zradí, nabídl by spolupráci.

	Kolo 1	Kolo 2	Kolo 3
MyStrategy	zrada	spolupráce	spolupráce
Protihráč	zrada	spolupráce	zrada

Tabulka 6: Možný průběh hry mezi strategií MyStrategy a neznámým oponentem [2]

Pozn.: Identifikační mechanismy některých strategií mohou mít nastavenu určitou hodnotu, kdy je soupeř ještě považován za identifikovatelného, tj. pokud chyba nepřekročí tuto hodnotu s určitou pravděpodobností, pak lze protihráčovu strategii identifikovat.

Zkusme ještě srovnat strategie MS a APavlov. Obě strategie, jak už bylo zmíněno, využívají identifikační mechanismus. Zatímco strategie APavlov si hru rozdělí na úseky, aby zamezila případné chybné identifikaci protihráče z důvodu změny jeho chování během hry, MS využívá identifikační mechanismus k odhalení chování hráče v několika málo kolech a podle toho se po celý zbytek hry snaží s tímto hráčem optimálně interagovat[50].

MS je aktivnější a "agresivnější" v určování protihráče (na začátku ho zkusmo zradí) než APavlov. To může být ovšem nevýhodou, např. u strategií jako je GRIM, která po zradě protihráče začne zrazovat neustále.

[50] Nevýhoda- není invariantní vůči změně strategie.

5 Hry s lidskými hráči

Tato kapitola je speciálně věnována hrám IPD, kde jako hráči vystupují lidé. Jak už bylo zmíněno, opakované vězňovo dilema s lidskými hráči je velmi specifická hra, většinou totiž v IPD hrách a turnajích vystupují naprogramované algoritmy, jejichž chování je předem dáno. Samozřejmě, jak jsme už uvedli, mohou mít algoritmy zabudovaný identifikační mechanismus, který do jisté míry zajišťuje optimální interakci s protihráči, nicméně lidé se mohou chovat nepředvídatelně a někdy jejich chování nelze přiřadit k žádnému algoritmu, což by i pro algoritmy s identifikačním mechanismem mohlo znamenat vážný problém.

Literatury, která se věnuje tomuto tématu, není mnoho. V této kapitole budou popsány výsledky experimentů ze tří vybraných článků na toto téma. Každý experiment fungoval trochu jinak a také záleželo na autorech samotných experimentů, co budou těmito experimenty sledovat.

Prvním z experimentů, který si zde představíme, je experiment popsaný v [19]. Hry se účastnilo 44 studentů Ohio State University. Výplatní matice byla totožná s **Tabulkou 3.** Hráči byli postupně usazováni po dvojicích[51] ke stolu s přepážkou (aby na sebe neviděli a nemohli spolu během hry komunikovat). Před sebou měli dvě tlačítka: **černé**, které měli zmáčknout v případě, že v daném kole chtějí se soupeřem spolupracovat, a **červené**, pokud ho chtějí zradit. Předem byli seznámeni s pravidly a k dispozici jim byla po celou dobu hry výplatní tabulka a taky sedící rozhodčí, nad nímž se také vždy po skončení kola rozsvítila tabule, zaznamenávající volby hráčů, takže oba hráči věděli, jaký tah zvolil soupeř. Hráči spolu odehráli 50 kol.

Závěr experimentu: Výsledky ukázaly, že ve většině případů převažovala kombinace tlačítek červené - červené (u 17 z 22 dvojic), převaha kombinace tlačítek černé - černé se vyskytla jen u dvou dvojic. Pokud by hráči hráli náhodně, pak by se každá ze 4 kombinací tlačítek[52] vyskytla v průměru 12,5krát za1 hru. Tento průměr byl překročen pro kombinaci červené – červené[53] u dvaceti případů. Pouze dvě dvojice z 22 během hry spolupracovaly. Dále byl také proveden mediánový test (jak test funguje – viz [24]), který srovnával počet kombinací tlačítek červené – červené v první a druhé polovině hry a ukázalo se, že v druhé části hry je významná převaha kombinace červené – červené tlačítko, tedy hráči měli větší tendenci se zrazovat v druhé části hry.

[51] Každý hráč hrál jen s jedním soupeřem
[52] Tj. červené - červené, červené - černé, černé - červené, černé - černé
[53] Zrada - zrada

Pozn.: V tomto článku byly dále popsány dva další případy, jeden z nich na principu modifikace výplatní tabulky a druhý s původní tabulkou, ale po 25 odehraných kolech nechali hráče na 2 minuty spolu komunikovat o dalším průběhu hry. Ukázalo se, že v důsledku komunikace mezi hráči klesl počet vzájemných zrad. Bližší informace, týkající se experimentu, jsou k nalezení ve zmiňovaném článku [19].

Dalším experimentem byl experiment popsaný v [17] a [18]. Účelem experimentu bylo zjistit, zda lidé používají předpovídané druhy kooperativních strategií[54], takže byli požádáni studenti biologie bernské univerzity o to, aby si zahráli IPD.

Milinski a Wedekind zkoumali vliv pracovní paměti na spolupráci tak, že jednu polovinu hráčů nechali hrát klasické IPD a druhou polovinu hráčů nechali mezi každým kolem hry IPD odehrát 1 kolo sekundární hry na paměť, konkrétně pexesa, kdy každý z hráčů otočil po každém odehraném kole vězňova dilematu 2 kartičky pexesa, aby snížili jejich schopnost pamatovat si výsledek předchozích kol IPD.

Poté ještě nechali hráče odehrát 20 kol proti některé z 2 strategií – buď AllC, která ale v 16. kole zradila nebo proti strategii, která se po prvních 6 kol chovala jako TFT, od 7. kola po zbytek hry jako AllD. Pro identifikaci chování hráče (jako GTFT nebo jako Pavlov[55]) byly spočítány podmíněné pravděpodobnosti spolupráce P_{CC}[56], P_{CD}[57], P_{DC}, P_{DD}.

Závěr experimentu: Autoři článku zjistili, že hráči, kteří nemuseli mezi koly IPD hrát pexeso volili „složitější strategie“, jako např. Pavlov (který reaguje nejen na soupeřovu volbu, ale na kombinaci voleb obou hráčů, tedy jak protihráčovu, tak svou), zatímco hráči, kteří mezi koly IPD hráli pexeso, používali jednodušší strategie typu TFT, GTFT.

Pozn.: Zkusme porovnat tento experiment s naším experimentem z **kapitoly 7**. Nezkoušeli jsme sice snížit „schopnost pamatovat si“ u jedné skupiny hráčů, na druhou stranu, chování hráčů jsme porovnávali s větší množinou strategií než jen GTFT a Pavlov, navíc jsme předpokládali, že chování hráčů během hry je proměnlivé.

---

54 Konkrétně GTFT a Pavlov

55 V článku uvažují pouze dvě skupiny (GTFT, Pavlov), do kterých je hráč při této identifikaci zařazen

56 $P_{CC} = P(C|CC)$ … s jakou pravděpodobností bude hráč spolupracovat v případě, že v předchozím kole oba hráči spolupracovali

57 $P_{CD} = P(C|CD)$ …pravděpodobnost, že hráč (A nebo B) bude spolupracovat v případě, že v předchozím kole tento spolupracoval a soupeř ho zradil.

Posledním experimentem, který si zde uvedeme je Flood - Dresher experiment a je podrobně popsán v [14]. Tento experiment byl proveden roku 1950, přítomen u něj byl i **Robert C. Tucker**, který tomuto problému dal jméno a interpretoval ho příběhem s dvěma vězni tak, jako známe dnes.

V r. 1950 M. M. Flood a M. Dresher pozvali 2 kamarády, ekonoma Armena Alchiana (hráč A) a matematika Johna Williamse (hráč B), aby si zahráli vězňovo dilema. Výplatní matice vypadala takto:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (0.5, 1) & (-1, 2) \\ (1, -1) & (0, 0.5) \end{pmatrix} \end{array}$$

Jednotkami byly centy dolarů. Matice není symetrická, což nebývá moc obvyklé u IPD, ale obecně stačí, aby byly splněny podmínky pro R, S, T, P (viz **kapitola 2.1**). Lze ale vidět, že hráč B měl výhodu oproti hráči A, neboť mohl obdržet jen vyšší (nebo minimálně stejně vysoké) výplaty. Hra měla 100 kol, což hráči předem věděli. Nemohli spolu komunikovat během hry, ale na konci každého kola byli informováni o tazích soupeře a o příslušných výplatách. V článku je k dispozici tabulka se všemi koly hry a volbami jednotlivých hráčů, včetně jejich komentářů, které měli během hry. Po 100. kole byly spočítány výplaty hráčů a hráči obdrželi příslušný počet centů.

Závěr experimentu: Ve hře převažovala volba vzájemné spolupráce – v 60 % případů, naopak vzájemná zrada se objevila jen ve 12 % případů. Přitom v prvních 10 kolech se objevila vzájemná zrada hned 5krát, poté už jen zřídka. Nejdelší „období vzájemných zrad“ trvalo 3 kola. Hráči se také vzájemně zradili v posledním 100. kole hry, přitom hráč A zradil už o kolo dříve. Hráč A začal zradou také v prvním kole, neboť takové chování očekával i od hráče B.

6 Možné modifikace opakovaného vězňova dilematu

V práci jsme se omezili pouze na základní[58] verzi hry, tzn., uvažujeme pouze dva hráče, kteří mohou volit mezi spoluprací a zradou a neuvažuje se ani možnost šumu. Ovšem z důvodu, že se v některé literatuře, týkající se vězňova dilematu, mohou objevit i tyto modifikace, je vhodné se o nich alespoň ve stručnosti zmínit.

Zavedení šumu

V některé literatuře ale se šumem počítají, neboť může být součástí her IPD. Šum zde chápeme tak, že hráč může dostat chybnou informaci o tom, jakou strategii jeho protihráč zvolil. Např., hráč A spolupracuje s hráčem B, ale hráč B obdrží od hráče A chybnou odpověď – zradu. Tato „drobná" chyba může mít ale markantní vliv na celou hru a hlavně na skóre obou hráčů.

Pozn.: Kdybychom např. proti sobě postavili dva algoritmy TFT, které za normálních okolností spolu spolupracují po celou dobu hry a uvažovali šum, pak by se právě v případě šumu (kdy TFT zvolí sice spolupráci, ale vlivem šumu je změněna na zradu), dostaly algoritmy do koloběhu neustálého oplácení zrad. Došlo by ke střídání (C, D), (D, C), (C, D), (D, C) …, takže ve výsledku by skóre hráčů byla o dost menší. Je tedy vhodné, aby v případě šumu byl schopen hráč na tuto situaci reagovat. Doporučuje se např. odpouštět zradu s nějakou malou pravděpodobností, jako to např. dělá algoritmus GTFT, nebo se chovat jako strategie Pavlov, která umí rychle zareagovat na nízké skóre získané v předchozím kole [21].

Více hráčů

Problematika opakovaného vězňova dilematu je většinou chápána jako hra dvou hráčů. Darwen a Yao [22] popisují v článku IPD s *n* hráči, tvrdí, že většina reálných problémů (především z oblasti ekonomie) se nedá modelovat pomocí opakovaného vězňova dilematu 2 hráčů, nýbrž pomocí IPD *n* hráčů. Jejich experimenty mj. ukázaly, že spolupráce je ve velké skupině hráčů méně pravděpodobná, než když jsou ve vězňově dilematu proti sobě postaveni dva hráči.

---

[58] Ale zároveň také nejběžnější.

Více voleb

Také můžeme uvažovat, že hráči mají možnost více voleb, než jen spolupracovat nebo zradit. Buď by mohlo být do hry zahrnuto např. 5 voleb hráče, mezi nimi např. možnost spolupráce na $50~\%$ apod. anebo bychom například mohli zahrnout celý interval $\langle 0; 1 \rangle$, kde:

- extrémní hodnoty: 0 – úplná zrada, 1 – úplná spolupráce,
- ostatní hodnoty: $x \in \ (0; 1) :$ částečná spolupráce z $x \cdot 100~\%$

7 Praktická část

V praktické části se budeme zabývat analýzou turnaje, uspořádaného autorkou práce. Robert Axelrod, a i většina dalších vědců, kteří uskutečňovali turnaje IPD, stavěli proti sobě většinou naprogramované algoritmy, tento experiment byl ale založen na analýze chování studentů, kteří se účastnili turnaje fyzicky a své volby proti soupeři v jednotlivých kolech hry zadávali do počítačového softwaru, který jim vždy po skončení kola ukázal volbu[59] soupeře v daném kole, výplaty obou hráčů během hry a historii tahů celé právě hrané hry. Sesbíraná data se dále analyzovala a výstupem bylo množství tabulek a grafů (viz **podkapitoly 7.3, 7.4, 7.5 a 7.6**). Protože ale lidé nefungují jako algoritmy, které mají přesně naprogramován postup, jak se budou v dané situaci chovat, mohou své chování v průběhu hry nebo napříč hrami libovolně měnit. To byl také důvod, proč nebyla analýza prováděna pro jednotlivé hry vcelku, ale v rámci jednotlivých her vždy pro úseky, které jsme si vymezili. K tomu nám posloužilo posuvné okno s fixně zvolenou délkou. Jak jsme hledali nejbližší strategii a další důležité náležitosti jsou popsány v **podkapitolách 7.2**, **7.3** a dále také na přiloženém CD, kde lze najít mj. veškerá data získaná z turnaje.

Pozn.:

1) Od dob Roberta Axelroda uvažovány 2 typy turnajů – *každý s každým* a *evoluční přístup*. V našem turnaji ale nepracujeme ani s jedním přístupem. Každý student hrál jen proti několika náhodně vylosovaným soupeřům.
2) Jak již bylo zmíněno výše, většina turnajů byla založena na tom, že byly proti sobě stavěny naprogramované algoritmy, přesto se ale uskutečnilo několik turnajů s lidskými hráči (viz **kapitola 5**)

7.1 Průběh a cíle turnaje

Dne 27. 3. 2018 se zúčastnilo 20 dobrovolníků z řad studentů Univerzity Palackého v Olomouci turnaje opakovaného vězňova dilematu. Tento turnaj byl uspořádán autorkou práce a studenti využili software, který pro tyto účely vytvořil vedoucí práce, pan Mgr. Pavel Holeček, PhD. Hráči byli přítomni fyzicky, tedy neprogramovali algoritmy, jako tomu bylo např. v případě Axelrodových turnajů. Samotnému turnaji předcházela krátká prezentace, kde byla studentům vysvětlena pravidla hry. Studenti byli poté rozsazeni k počítačům s náhodně přiřazeným číslem

[59] Spolupráci či zradu

a s připraveným softwarem a hráli hry proti některým z dalších přítomných studentů[60]. Protihráči jim byli rovněž vybráni náhodně a podle čísla nebyli schopni identifikovat, o kterého hráče se jedná a tudíž ani sledovat jeho reakce během hry. Hráči nevěděli, kolik kol bude mít hra. V každém kole každé hry měli možnost volby spolupráce nebo zrady, podle čehož pak po každém kole obdrželi příslušnou výplatu. Během jednotlivých her mohli také na svém počítači sledovat skóre své i skóre svého protihráče a historii provedených tahů.

Výplatní tabulka, která byla pro turnaj využita, se shoduje s tabulkou, kterou pro své turnaje použil Robert Axelrod (**Tabulka 7**). Sledovali jsme chování hráčů a podobnost s některými známými algoritmy. Počet kol jednotlivých her nebyl fixní, ale pohyboval se v rozmezí 20 až 23, což však hráči předem nevěděli. Důvodem proměnlivého počtu kol byl fakt, že pokud by hráči znali počet kol dopředu, pak by se minimálně v posledním kole měli rozhodnout zradit, protože jim soupeř nemůže tuto zradu vrátit (viz **kapitola 3**).

Každý ze studentů odehrál 4 hry s některými z dalších studentů a poté 2 hry proti některým algoritmům. Množina vybraných algoritmů byla různorodá. Pro každý z algoritmů přidejme rovněž důvod, proč byl vybrán pro tento turnaj. Chování všech těchto čtyř algoritmů už bylo popsáno dříve, konkrétněji v podkapitolách **3.1** a **4.3**.

Byly vybrány následující algoritmy:

- **AllC** – jedná se o jeden z nejjednodušších algoritmů; dalo se očekávat, že hráči budou schopni odhadnout, o jaký algoritmus se jedná a přijít si na vysoké skóre.
- **Adaptivní Pavlov** – algoritmus propracovanější, během hry obvykle mění své chování a tak dokáže dobře reagovat na „nepředvídatelné“ lidské chování.
- **Pavlov** – úspěšný algoritmus v IPD turnajích
- **TFT** – rovněž velmi úspěšný algoritmus v IPD turnajích

Doplňme, že všechny tyto algoritmy mají společnou vlastnost, jsou milé, a tudíž nezačaly zrazovat jako první, tzn., hráč, který byl ochoten spolupracovat, s nimi udržel spolupráci během celé hry.

Protože počet kol pro jednotlivé hry byl různý, sledovala autorka průměrnou výhru za jedno kolo, a to odděleně pro hry člověk vs. algoritmus a pro hry člověk proti člověku. Výstupem jsou jak tabulky s výplatami hráčů, tak i bližší analýza pro pět náhodně vybraných her (viz **podkapitola 7.3** a **7.6**)

[60] I proti některým algoritmům, popsáno dále.

		Hráč 2	
		Spolupráce	**Zrada**
Hráč 1	**Spolupráce**	(3, 3)	(0, 5)
	Zrada	(5, 0)	(1, 1)

Tabulka 7: Výplatní tabulka využitá při turnaji

7.2 Analýza získaných dat

V této části si popíšeme, jakým způsobem byla zpracována data získaná z turnaje a jaké metody byly použity pro analýzu. Tato data nám poskytl software[61], ve kterém hráči hru hráli. Celkem bylo za turnaj nasbíráno 3456 oddělených voleb hráčů, tj. 1728 „řádků" hry. Data byla poté použita k analýze – hledání nejbližší strategie.

Vzhledem k náročnosti[62] této analýzy, ve které je potřeba přepočítávat pro každé kolo hry podmíněné pravděpodobnosti, vzdálenosti apod., bylo podrobně zanalyzováno pouze 5 náhodně vybraných her. Soubor s daty a analýzou je k nalezení na přiloženém CD.

Pozn.: Cílem ani nebylo zanalyzovat všechny hry všech hráčů v turnaji, ale navrhnout vhodnou metodu pro analýzu dat a otestovat její funkčnost na vzorku reálných dat.

Analýzu můžeme rozdělit do tří hlavních fází[63]:

1) **Klasifikace hráče** - ke kterému z vybraných referenčních algoritmů (AllC, AllD, TFT, Pavlov, RAND) je jeho chování nejblíže. Počítá se s tím, že chování hráčů se může měnit v průběhu hry. Proto sledujeme chování za určitý časový (myšleno za určitý počet po sobě jdoucích kol) úsek a výsledkem je klasifikace hráče v rámci tohoto úseku. Je tedy i vidět, jak se chování hráče měnilo v průběhu hry.

[61] Software byl vytvořen vedoucím diplomové práce, p. Mgr. Pavlem Holečkem, Ph.D.
[62] A také množství dat
[63] Blíže jsou tyto fáze popsány v **podkapitole 7.3**

2) **Grafická reprezentace** – pro usnadnění orientace v průběhu hry a chování obou hráčů.
3) **Sumarizace** - jak se vyvíjel zisk obou hráčů v průběhu hry. Užitečné např. v případě, kdy hráč měnil své chování (strategii) během hry -zda se hráči tato změna se mu vyplatila či nikoliv.

Pro všechny hry hráč vs. hráč byl na základě kumulativních výplat za hru sestrojen jeden společný graf (viz **Obrázek 7**, **Obrázek 8**), kde lze sledovat vývoje výplat v rámci jednotlivých her a porovnání se situacemi, kde by hráči spolu spolupracovali a kdy by se vzájemně jen zrazovali. Podívejme se nyní na jednu z pěti blíže analyzovaných her.

7.3 Výsledky analýzy pro vybrané hry

7.3.1 Klasifikace hráče

Snažíme se o klasifikaci hráče v závislosti na tom, ke kterému z vybraných referenčních algoritmů (vybráno bylo 5 možných výsledků klasifikace – **AllC**, **AllD**, **TFT**, **Pavlov**, **RAND**[64]) má jeho chování nejblíže.

Počítáme přitom s tím, že chování se v průběhu hry může měnit. To byl také důvod, proč jsme sledovali chování za určitý časový úsek[65] a výsledkem byla klasifikace hráče v rámci tohoto úseku. Je tedy i vidět, jak se chování hráče měnilo v průběhu hry.

Všechny výpočty byly prováděny v rámci oken s délkou 8 kol. Oknem rozumíme po sobě následující kola hry, ze kterých vždy počítáme výstupy. Tuto délku okna jsme zvolili proto, neboť se při zkoušení různých délek jevila jako nejvíce vyhovující, dává nám nejlepší výsledky. Je také kompromisem mezi hodně malou délkou okna, ze které by bylo složité něco zanalyzovat, příp. by analýza neměla velkou vypovídací hodnotu, a příliš velkou délkou okna, kdy by pak počet oken v rámci 1 hry byl zase až moc malý a mohly by nám uniknout některé změny v chování hráčů během hry.

Začínáme vždy od prvního kola hry, u délky okna 8 zahrneme prvních 8 kol hry (prvních 8 tahů obou hráčů). Poté okno posuneme o řádek níže a to samé provedeme s řádky 2 – 9. Výsledky analýzy v rámci okna vždy zapíšeme k poslednímu řádku v okně[66].

[64] Množina referenčních algoritmů by se pochopitelně dala rozšířit
[65] Určitý počet po sobě jdoucích kol
[66] Tedy, z analýzy v rámci oken budeme mít první výstup až v 8. řádku hry.

Nyní už budeme pracovat s reálnými daty. Začneme analýzou prvního okna první náhodně vylosované hry. První řádek je zabarven šedě, což nám značí, že volby hráčů v tomto kole se neuplatní ve výpočtech, neboť je potřeba znát výsledky předchozího kola v rámci tohoto okna. Jednotlivé řádky hry jsou zabarveny jednou ze čtyř barev v závislosti na tom, jaká byla kombinace voleb obou hráčů. (viz např. **Tabulka 8**)

Např., pokud se hráči vzájemně zradili, pak je toto pole zabarveno červenou barvou. Toto barevné značení jednak usnadnilo výpočty, neboť vše bylo počítáno ručně, ale také se lze ve hře lépe zorientovat a hned vidět, jaké měli hráči tendence. Z tohoto osmiřádkového okna budeme nyní počítat podmíněné pravděpodobnosti spolupráce, které následně využijeme k hledání algoritmu, jehož chování je nejbližší chování hráče v rámci okna.

kolo	hráč A	hráč B
1	C	D
2	C	D
3	C	C
4	D	D
5	D	D
6	D	D
7	D	C
8	C	D

Tabulka 8: První okno hry

Pracovat budeme se čtyřmi podmíněnými pravděpodobnostmi spolupráce P_{CC}, P_{CD}, P_{DC}, P_{DD}, kde:

- P_{CC} = P (C|CC) … s jakou pravděpodobností v okně bude hráč spolupracovat v případě, že v předchozím kole oba hráči spolupracovali
- P_{CD} = P (C|CD) … pravděpodobnost, že hráč[67] bude spolupracovat v případě, že v předchozím kole hráč A spolupracoval a hráč B zradil[68].
- P_{DC} = P (C|DC) … pravděpodobnost, že hráč bude spolupracovat v případě, že v předchozím kole hráč A zradil a hráč B spolupracoval.
- P_{DD} = P (C|DD) … s jakou pravděpodobností v okně bude hráč spolupracovat v případě, že v předchozím kole se hráči zradili.

Pozn.: Nutno počítat s tím, že ne vždy budeme mít k dispozici všechny 4 pravděpodobnosti. V takovém případě budeme počítat jen s pravděpodobnostmi, které máme vždy k dispozici.

[67] Hráč A či B v závislosti na tom, pro koho pravděpodobnost počítáme

[68] U podmíněných pravděpodobností **1. písmeno indexu značí tah hráče A, 2. písmeno tah 2. hráče**

V případě, že je dostupná např. jen 1 pravděpodobnost, není tak jednoznačné, ke kterému algoritmu přiřadit chování hráče jako v případě, že známe podmíněných pravděpodobností více.

Pravděpodobnosti počítáme zvlášť pro A a B hráče a potom je zapíšeme do tabulky, která vypadá takto:

	hráč A	hráč B
P_{CC}	0,0000	0,0000
P_{CD}	1,0000	0,5000
P_{DC}	1,0000	0,0000
P_{DD}	0,0000	0,3333

Tabulka 9: Vypočtené podmíněné pravděpodobnosti pro první okno hry

Tyto podmíněné pravděpodobnosti jsou nejprve spočítány pro první okno a poté jsou aktualizovány s každým posunem okna. Jejich výpočet (konkrétně tento) vychází z **Tabulky 8**, a pokud např. budeme chtít určit pravděpodobnost P_{DD}, pak vyjdeme z kol, která se následují po kole, kdy se oba hráči zradili (tedy jedná se vždy o řádek pod každým tím řádkem, který je zabarven červeně[69]), neboť nás zajímá pravděpodobnost spolupráce poté, co se hráči vzájemně zradili.

Pokud se podíváme na kola 5, 6 a 7 u hráče A, můžeme vidět, že hráč A vždy zradil poté, co se hráči v předchozím kole vzájemně zradili, tedy jeho pravděpodobnost spolupráce po zradě obou hráčů, tj. P_{DD}, je nulová. Hráč B ale v 7. kole spolupracoval, tedy v 1 ze 3 případů, a proto je tato jeho podmíněná pravděpodobnost rovna $\frac{1}{3}$. Takto bychom si mohli napočítat všechny podmíněné pravděpodobnosti pro toto okno.

Když už máme podmíněné pravděpodobnosti, musíme je porovnat s podmíněnými pravděpodobnostmi pro námi vybrané algoritmy AllC, TFT, RAND, Pavlov a AllD. Díky tomu, že se jedná o algoritmy, které berou v potaz maximálně předchozí tah[70], jsou pro ně podmíněné pravděpodobnosti pevně dané (viz **Tabulka 10 –** podmíněné pravděpodobnosti pro algoritmy převzaty z [2])

	AllC	AllD	TFT	Pavlov	RAND
P_{CC}	1	0	1	1	0,5
P_{CD}	1	0	0	0	0,5
P_{DC}	1	0	1	0	0,5
P_{DD}	1	0	0	1	0,5

Tabulka 10: Podmíněné pravděpodobnosti spolupráce pro vybrané algoritmy (pro hráče A)

[69] Konkrétně pro tuto hru by se jednalo o kola 5,6 a 7.
[70] Tedy, nedívají se hlouběji do historie

Tabulka 10 zobrazuje podmíněné pravděpodobnosti pro hráče A. Pro hráče B by tabulka vypadala dost podobně, pouze u algoritmu TFT by došlo k prohození pravděpodobností pro P_{CD} a P_{DC} z toho důvodu, že 1. písmeno v indexu u podmíněných pravděpodobností značí tah hráče A a 2. písmeno tah hráče B. V případě TFT algoritmu by hráč A určitě nespolupracoval, kdyby ho soupeř v předchozím kole zradil, tudíž P_{CD} je zde rovno nule, naopak hráč B by v tomto případě spolupracoval, protože by jakožto TFT neměl důvod zradit hráče A, který s ním v předchozím kole spolupracoval.

AllC strategie má všechny tyto podmíněné pravděpodobnosti rovny 1, neboť spolupracuje za jakýchkoli předchozích podmínek, naopak AllD má všechny tyto hodnoty nulové, protože nespolupracuje nikdy.

Teď nás tedy bude zajímat, jak je si chování hráčů A a B podobné s chováním těchto algoritmů. Pro každý z referenčních algoritmů (tj. v našem případě $\forall j = 1, \dots 5$, kde j ... počet referenčních algoritmů) hledáme vzdálenost mezi podmíněnými pravděpodobnostmi hráče a j – tého referenčního algoritmu. Využili jsme dva typy vzdáleností[71]:

- ***euklidovskou vzdálenost:***

$$d\_euklid(j) = \sqrt{\sum_i (x_i - y_{ji})^2}$$

$$\forall j = 1, \dots 5$$

$$i = 1, 2, 3, 4 \dots v\ případě, že\ známe\ všechny\ 4\ podmíněné\ pravděpodobnosti^{72}$$

$$a\ kde\ x_i \dots i - tá\ podmíněná\ pravděpodobnost\ spolupráce\ pro hráče, pro\ kterého\ je\ analýza\ prováděna$$

$$y_{ji} \dots i - tá\ podmíněná\ pravděpodobnost\ j - tého\ referenčního\ algoritmu^{73}$$

Výsledkem klasifikace s pomocí této vzdálenosti je referenční algoritmus[74] *k*, pro který platí:

$$d\_euklid(k) = \min_{j = 1, 2, \dots, 5} d\_euklid(j)$$

- ***maximální vzdálenost*:**

$$d\_\max(j) = (\max_i |x_i - y_{ji}|)$$

[71] Šla by ale použít jakákoliv jiná vzdálenost

[72] Index $i = 1$ odpovídá P_{CC}, $i = 2$ odpovídá P_{CD}, $i = 3$ pro P_{DC}, $i = 4$ pro P_{DD}. Pokud známe jen některé z podmíněných pravděpodobností, pracujeme jen s některými x_i.

[73] Viz **Tabulka 9**

[74] Obecně jich může být i více

Výsledkem klasifikace s pomocí této vzdálenosti je referenční algoritmus[75] *k*, pro který platí:

$$d\_max(k) = \min d\_max(j)$$
$$\forall j = 1, \dots 5$$
$$i = 1, 2, 3, 4 \dots v\ případě, že\ známe\ všechny\ 4\ podmíněné\ pravděpodobnosti$$
$$a\ kde\ x_i \dots i - tá\ podmíněná\ pravděpodobnost\ spolupráce\ prohráče, pro\ kterého\ je\ analýza\ prováděna$$
$$y_{ji} \dots i - tá\ podmíněná\ pravděpodobnost\ j - tého\ referenčního\ algoritmu$$

Uvedená situace, kdy známe všechny x_i nemusí samozřejmě vždy nastávat. Ke konci hry už hráči nebyli tolik ochotni spolupracovat, a proto neznáme jejich pravděpodobnosti P_{CC} (**viz např. Tabulka 11**). V takovémto případě výpočet vzdálenosti bude vycházet jen z dalších tří pravděpodobností a tu čtvrtou do výpočtu nezahrneme (tedy pracovali jsme pouze s $i = 1, 2, 3$) protože nemáme informace, jak by se hráči v této situaci zachovali.

kolo	hráč A	hráč B
13	D	D
14	D	D
15	C	D
16	D	D
17	D	C
18	D	C
19	D	D
20	D	D

Tabulka 11: 13. okno ve hře

Podívejme se nyní na tabulku pro obě tyto vzdálenosti v prvním okně hry (**Tabulka 12**). V řádcích tabulky jsou jednotlivé referenční algoritmy a zkoumáme, ke kterému má chování daného hráče nejblíže. Ve sloupcích tabulky se nacházejí vypočtené vzdálenosti od daných referenčních algoritmů, zvlášť pro hráče A a B. Pro srovnání jsou všechny hodnoty uvedeny s použitím obou vzdáleností.

Žlutě zabarvená políčka jsou minimální hodnoty pro dané sloupce a v řádcích odpovídajících těmto hodnotám nalezneme algoritmy nejbližší chování hráče v rámci okna). Tedy, z prvního okna bychom s pomocí euklidovské vzdálenosti (a i s pomocí d_max) řekli, že chování hráče A cmá nejblíže k chování algoritmu RAND[76], hráči B by byl přiřazen jako nejbližší algoritmus

[75] Obecně jich může být i více
[76] Z naší množiny referenčních algoritmů

v případě použití obou typů vzdáleností algoritmus AllD, jen v případě využití d_max navíc i algoritmus RAND, neboť vzdálenosti pro AllD a RAND zde vyšly totožné.

Pozn.: Tato metoda nemusí hráči vždy přiřadit jen jeden algoritmus. Např. v případě, kde máme v rámci okna samá C, tedy když hráči spolupracují, tato metoda by jejich chování přirovnala k AllC a zároveň i TFT a Pavlov, neboť hráč ještě neměl možnost (v rámci okna) ukázat, jak by se zachoval v případě, kdyby ho jeho soupeř zradil.

	d_euklid			d_max	
Referenční algoritmy	hráč A	hráč B		hráč A	hráč B
A a AllC	1,4142	1,6415		1,0000	1,0000
A a AllD	1,4142	**0,6009**		1,0000	**0,5000**
A a TFT	1,4142	1,1667		1,0000	1,0000
A a Pavlov	2,0000	1,3017		1,0000	1,0000
A a RAND	**1,0000**	0,7265		**0,5000**	**0,5000**

Tabulka 12: Vypočtené vzdálenosti z prvního okna hry

Po obdržení těchto hodnot jsme algoritmy, kterým odpovídaly tyto minimální hodnoty, přiřadili jednotlivým hráčům, a to k poslednímu řádku okna. Tedy v případě, že máme délku okna 8, budeme mít první nejbližší strategii až v 8. řádku (kole) hry.

kolo	Výsledek klasifikace s použitím d_euklid					Výsledek klasifikace s použitím d_max			
	hráč A		hráč B			hráč A		hráč B	
. . .	. . .		. . .			. . .		. . .	
8	RAND	C	D	AllD		RAND	C	D	AllD, RAND

Tabulka 13: Zápis výsledných nejbližších algoritmů z 1. okna (**Tabulky 10)** do výsledné tabulky

Toto samé provedeme pro druhé okno hry, tedy:

kolo	Výsledek klasifikace s použitím d_euklid					Výsledek klasifikace s použitím d_max			
	hráč A		hráč B			hráč A		hráč B	
. . .	. . .		. . .			. . .		. . .	
8	RAND	C	D	AllD		RAND	C	D	AllD, RAND
9	RAND	D	D	AllD		RAND	D	D	AllD, RAND

Tabulka 14: Výsledky ze dvou oken hry

Takto postupně získáme nejbližší algoritmy pro všechna kola hry, počínaje 8. kolem. Kola, která předcházejí 8., budeme muset určit jiným způsobem – např. oknem menší délky.

Klasifikace hráče byla provedena i pro počáteční kola hry, kdy ještě nebylo odehráno 8 tahů (kola 1 až 7). Tím pádem pro počáteční tahy hry do klasifikace vstupují data za menší počet kol, než je stanovených 8, a proto výsledek klasifikace nemusí být tak přesný. V tabulkách jsou proto výsledky zaznačeny šedou barvou místo černé (viz např. **Obrázek 4**).

Například, pokud hráč A v prvním kole zvolil spolupráci, pak jeho chování určitě nebude blízké strategii AllD, všechny ostatní strategie z naší množiny referenčních algoritmů ale nelze vyřadit.

kolo	hráč A	hráč B
1	C	D
2	C	D
3	C	C
4	D	D
5	D	D
6	D	D
7	D	C

Tabulka 15: Prvních 7 kol hry

Pokud se podíváme na tabulky s nejbližšími algoritmy, poměrně často zde vystupuje strategie RAND, což může mj. značit to, že hráči se neblížili k žádné z dalších uvedených strategií. Protože všechny 4 její podmíněné pravděpodobnosti jsou rovny 0,5, tak do ní spíše spadnou hráči, kteří nemají jasně stanovenou strategii, ale spíše zkouší.

Ve hře lze sledovat poměrně velké množství červeně zabarvených řádků, které znamenají vzájemnou zradu hráčů, a konkrétně v této hře bylo jen několik málo kol, kde hráči vzájemně spolupracovali.

7.3.2 Grafická reprezentace hry

Ze samotného výpisu průběhu hry je obtížné něco říct o chování hráče v daný okamžik[77]. Proto jsme navrhli následující grafickou reprezentaci, která usnadňuje orientaci v průběhu hry a chování hráčů. Znázorňuje chování hráče v závislosti na tom, jak se k němu v předchozím kole zachoval protihráč. Barvy znázorňují jak moc „hodný“ je hráč na svého protihráče. Od situací, kdy protihráči v následujícím kole odpustil zradu z předchozího kola až po opačný extrém, kdy svého protihráče zradil bezdůvodně (tj. i když s ním protihráč v předchozím kole spolupracoval). Podívejme se na **Obrázek 5**, který barevně rozlišuje 4 druhy chování jednotlivých hráčů:

- **nevyprovokovaná zrada =** zrada, které nebyla vyprovokována ze strany soupeře, tj. soupeř v předchozím kole zvolil spolupráci
- **oplacená zrada =** zrada, která byla odplatou za protihráčovu předchozí zradu
- **spolupráce =** spolupráce, které předcházela spolupráce soupeře
- **odpuštěná zrada =** spolupráce hráče po soupeřově zradě za účelem pokusu o navázání spolupráce v dalších kolech hry

[77] Jedná se jen o posloupnosti písmen C a D.

ID: 15

kolo	hráč A	hráč B
1	C	D
2	C	D
3	C	C
4	D	D
5	D	D
6	D	D
7	D	C
8	C	D
9	D	D
10	D	D
11	C	C
12	D	C
13	D	D
14	D	D
15	C	D
16	D	D
17	D	C
18	D	C
19	D	D
20	D	D

d_max			
AllC, TFT, Pavlov, RAND	C	D	AllD, RAND
AllC	C	D	AllD, Pavlov
AllC	C	C	RAND
RAND	D	D	AllD, RAND
RAND	D	D	AllD, RAND
RAND	D	D	AllD, RAND
RAND	D	C	AllD, RAND
RAND	C	D	AllD, RAND
RAND	D	D	AllD, RAND
RAND	D	D	AllD
TFT	C	C	AllD
RAND	D	C	Pavlov, RAND
AllD, RAND	D	D	Pavlov
AllD, RAND	D	D	RAND
AllD, RAND	C	D	RAND
AllD, RAND	D	D	RAND
AllD, RAND	D	C	Pavlov, RAND
AllD	D	C	RAND
AllD	D	D	AllD
AllD	D	D	AllD, RAND

d_euklid			
AllC, TFT, Pavlov, RAND	C	D	AllD, RAND
AllC	C	D	AllD, Pavlov
AllC	C	C	RAND
RAND	D	D	AllD, RAND
RAND	D	D	AllD, RAND
RAND	D	D	AllD, RAND
RAND	D	C	AllD, RAND
RAND	C	D	AllD
RAND	D	D	AllD
AllD, TFT, RAND	D	D	AllD
TFT	C	C	AllD
RAND	D	C	Pavlov
AllD	D	D	Pavlov
AllD	D	D	Pavlov
AllD	C	D	Pavlov
AllD	D	D	Pavlov
AllD	D	C	Pavlov
AllD	D	C	RAND
AllD	D	D	AllD
AllD	D	D	AllD, RAND

tabulka výplat (kumul.)	
0	5
0	10
3	13
4	14
5	15
6	16
11	16
11	21
12	22
13	23
16	26
21	26
22	27
23	28
23	33
24	34
29	34
34	34
35	35
36	**36**

Obrázek 4: Výsledná analýza jedné vybrané hry z turnaje

kolo	hráč A	hráč B
1	C	D
2	C	D
3	C	C
4	D	D
5	D	D
6	D	D
7	D	C
8	C	D
9	D	D
10	D	D
11	C	C
12	D	C
13	D	D
14	D	D
15	C	D
16	D	D
17	D	C
18	D	C
19	D	D
20	D	D

nevyprovokovaná zrada
oplacená zrada
spolupráce
odpuštěná zrada

Obrázek 5: Jiný způsob analýzy hry pomocí 4 druhů chování

Z **Obrázku 5** lépe vidíme, jak se vyvíjela hra. Ze všech zrad, ke kterým během hry došlo, byla zhruba třetina těch nevyprovokovaných a už na začátku hry hráč B zkoušel, zda mu budou procházet zrady u jeho soupeře. Hráč A začal spoluprací a snažil se ji udržet i přesto, že hráč B ho zrazoval. Od 4. kola jej ale začal zrazovat také. Mezi velkým množstvím zrad od obou hráčů, které už poté byly spíše reakcí na protihráčovo přechozí kolo, lze sledovat i pár kol, ve kterých hráči zkoušeli s protihráčem spolupracovat, ovšem tato spolupráce netrvala dlouho. Pokud se ještě vrátíme k **Obrázku 4**, můžeme sledovat, že většina algoritmů, které vyšly jako nejbližší k chování hráčů, byly strategie nemilé a neodpouštějící, které si nevedly zrovna nejlépe v Axelrodových turnajích.

Skóre, které hráči obdrželi po skončení hry, bylo 36 bodů pro každého hráče, což při počtu dvaceti kol není zrovna nejlepší výsledek[78]. Při vzájemné spolupráci by si hráči mohli přijít na 60 bodů.

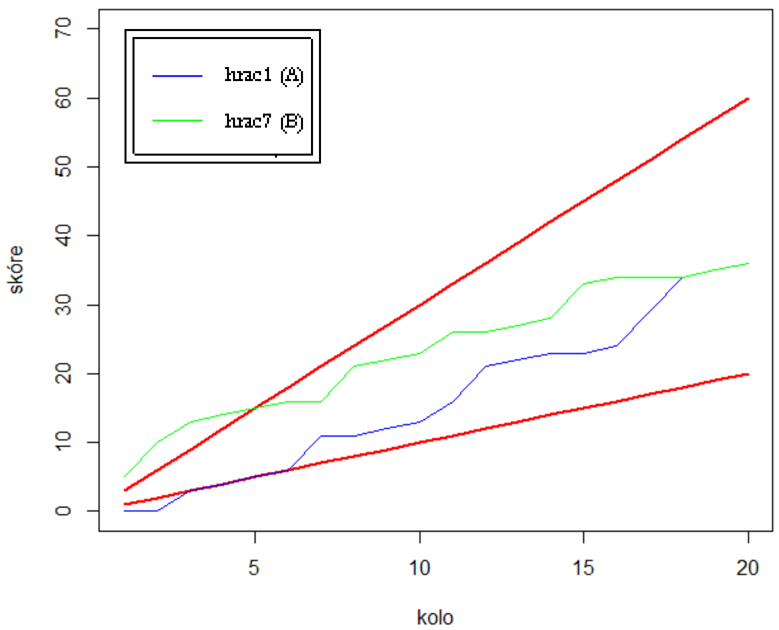
[78] Průměrné skóre - 1, 8 bodu / 1 kolo (max. je 5 bodů /1 kolo při jednostranné zradě, 3 body za vzáj. spolupráci.

7.3.3 Sumarizace

Kromě samotného chování hráče může být užitečné sledovat, jestli se mu toto chování vyplatilo, nebo ne. Proto je analýza průběhu hry doplněna o následující grafy, ze kterých je vidět vývoj výher obou hráčů v průběhu hry. Pokud došlo k razantní změně v chování hráče v jednom z tahů, lze tak snadno ověřit, jaký vliv měla tato změna na výši jeho výhry.

Podívejme se nyní na graf (viz **Obrázek 6**), ve kterém jsou zakresleny jednak křivky výplat jednotlivých hráčů, a jednak pro porovnání 2 červené přímky. Horní přímka zobrazuje, jaké skóre by hráči během hry dosahovali, kdyby zvolili vzájemnou spolupráci. Naopak dolní přímka ukazuje, jakého skóre by dosáhli, pokud by se neustále zrazovali.[79]

Z grafu lze sledovat, že na začátku si vedl zřetelně lépe hráč B. To je způsobeno vstřícností hráče A v prvních 3 kolech, kdy se marně pokoušel nabídnout spolupráci, ale hráč B mu na jeho první dva pokusy odpověděl nevyprovokovanou zradou. Touto zradou získal hráč B náskok, který si udržuje až zhruba do 17 kola. V 18. kole dochází ke zlomu a po řadě vzájemných zrad se pokouší naopak hráč B o vstřícné chování k hráči A a nabízí spolupráci. Hráč A na toto ale odpovídá zradou a tím srovnává skóre se svým soupeřem. **Kapitola 7.6** obsahuje analýzu dalších vybraných her turnaje. Celá analýza včetně výpočtů a výsledků je součástí přílohy, na přiloženém CD.



Obrázek 6: Grafické znázornění analyzované hry

[79] Řešení doporučované teorií her pro jednokolové vězňovo dilema a IPD, kdy je počet kol oběma hráčům předem znám (viz **kapitola 3**); zde ale hráčům počet kol nebyl znám.

7.4 Celkové výsledky turnaje

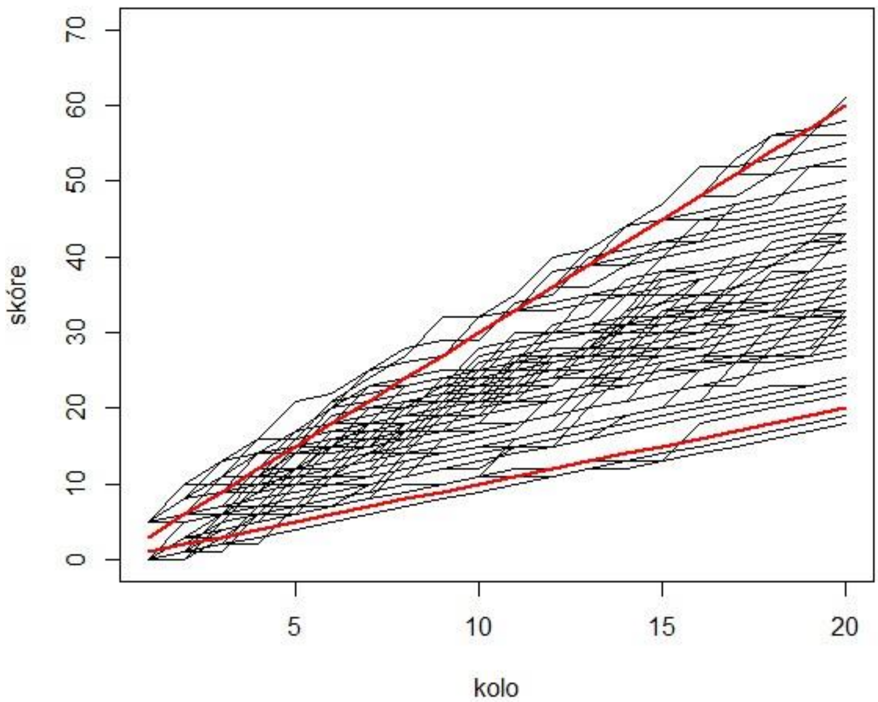
V **Tabulce 15** můžeme vidět výsledky turnaje pro hry člověk vs. člověk.

7.4.1 Hry člověk vs. člověk

Hry, ve kterých hráli proti sobě 2 lidi, dopadly obecně hůře než ty hry, které byly hrány proti algoritmům. Skóre se kvůli proměnlivým počtům kol na hru přepočítávalo na průměr za 1 kolo. Nejlepší hráč si přišel na průměrné skóre 2, 4341; nejhorší hráč na 1, 3166 (Připomeňme, že nejvyšší možné průměrné skóre za kolo je 5 (viz **Tabulka 7**). Naopak nejhorší možné je 0). Rozdíl mezi těmito hodnotami je poměrně velký. Pouze jedna čtvrtina hráčů měla průměrné skóre větší než 2. Pokud bychom spočítali průměrné skóre na 1 hru pro všechny hráče zároveň, vyšla by hodnota 1, 8843.

tabulka her hráč - hráč							**počet výher, proher a remíz**		
hráči	**průměrné skóre za 1 kolo hry**				**celk. prům. skóre**	**pořadí**	**# výher**	**# proher**	**# remíz**
hrac1	1,8000	1,5000	2,3500	1,1905	**1,7101**	**16.**	2	1	1
hrac2	1,7727	2,6000	1,1364	0,9524	**1,6154**	**18.**	1	1	2
hrac3	1,5455	1,4545	2,5909	2,0000	**1,8977**	**11.**	1	3	
hrac4	1,9565	2,0500	1,4500	2,2273	**1,9209**	**10.**	1		3
hrac5	3,0000	2,7826	1,7391	1,6190	**2,2852**	**2.**		2	2
hrac6	1,8095	2,0000	1,5714	1,6190	**1,7500**	**15.**	2	1	1
hrac7	1,8000	1,6818	1,4500	1,5652	**1,6243**	**17.**	1	1	2
hrac8	1,7727	2,0500	1,9565	2,0000	**1,9448**	**9.**	2		2
hrac9	2,0000	3,0000	1,6000	3,1364	**2,4341**	**1.**	3	1	
hrac10	1,9565	0,9130	2,3636	2,6818	**1,9788**	**7. - 8.**		3	1
hrac11	3,0000	1,7273	1,8095	1,4348	**1,9929**	**6.**	2		2
hrac12	1,0952	1,6000	1,1364	1,4348	**1,3166**	**20.**		2	2
hrac13	1,8000	1,5500	2,3333	1,7500	**1,8583**	**12.**	1	2	1
hrac14	1,7619	1,5909	2,5000	2,1905	**2,0108**	**5.**		3	1
hrac15	1,1818	1,4286	1,3500	1,7826	**1,4357**	**19.**		3	1
hrac16	1,9565	1,9130	1,5455	2,5000	**1,9788**	**7. - 8.**	1	3	
hrac17	2,0500	2,1304	1,6000	2,4286	**2,0523**	**4.**	4		
hrac18	2,0000	1,6667	2,0000	1,7500	**1,8542**	**13.**	3		1
hrac19	1,6364	1,8000	2,7273	2,7273	**2,2227**	**3.**	4		
hrac20	1,7391	1,5909	2,0952	1,7826	**1,8020**	**14.**		2	2

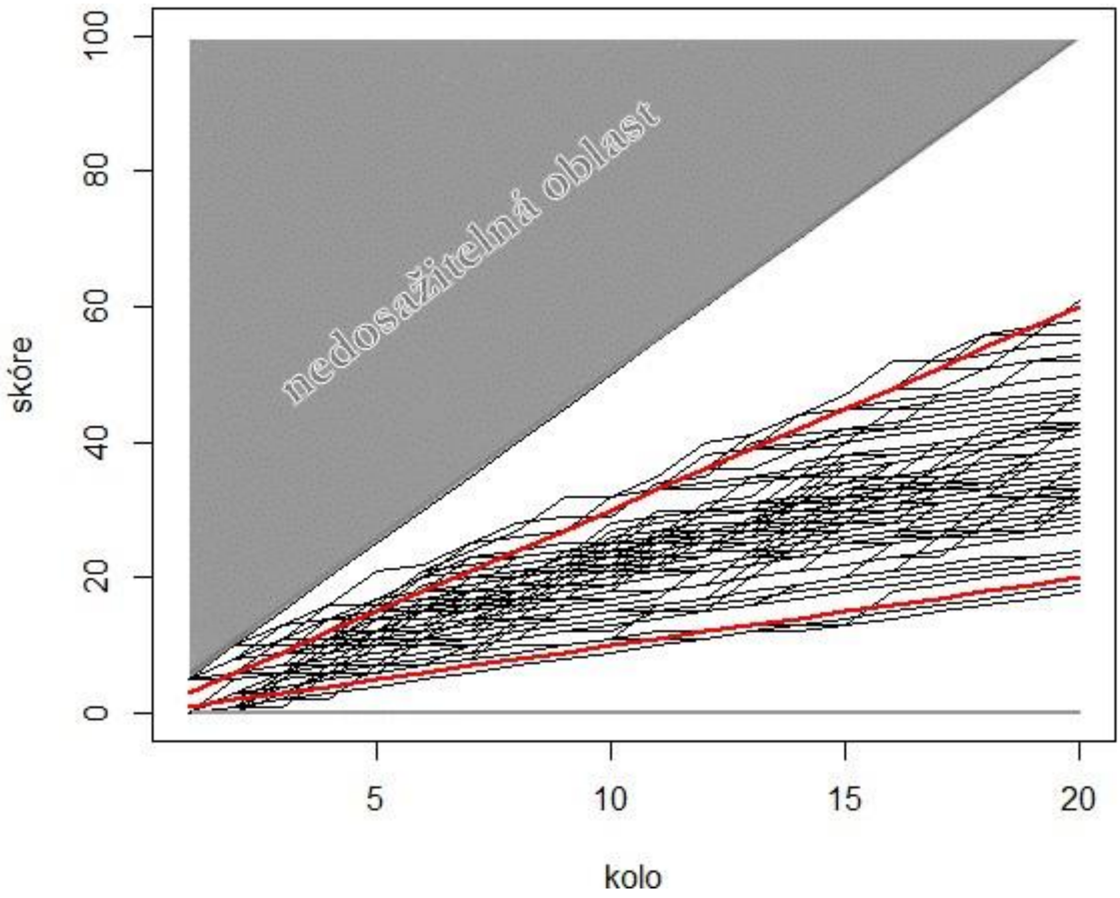
Tabulka 16: Výsledky turnaje her hráč vs. hráč



Obrázek 7: Souhrnný graf vývoje výplat pro všechny hry hráč vs. hráč

Na **Obrázku 7** můžeme sledovat, jak se vyvíjely výplaty hráčů během jednotlivých kol her. Oblast je vymezena dvěma červenými přímkami – horní pro případ, kdy by oba hráči spolupracovali, dolní pro případ, kdy by se oba hráči zrazovali v každém kole hry. Protože hráči odehráli v každé hře mezi 20 a 23 koly, utnuli jsme jejich hodnoty pro účely grafického zpracování po 20. kole. Mimo jiné si můžeme na obrázku všimnout také toho, že většina hodnot leží mezi těmito dvěma červenými přímkami, až na pár výjimek, kdy se někteří hráči dostali nad nebo pod tyto křivky z důvodu, že jejich protihráči se po nějaký čas nechali jimi „oškubávat".

Z grafu lze ale vidět, že nejvíce výplaty kolísaly zhruba uprostřed mezi přímkami. Hráči jsou vesměs nepoučitelní a zrazují, i když se ve většině případů stane to, že nakonec dosáhnou oba hráči skóre menšího, než by získali vzájemnou spoluprací (horní červená přímka). Na **Obrázku 8** máme stejný graf, jako na předchozím obrázku s tím rozdílem, že je zde zakresleno, jakých až výplat mohli hráči dosahovat během hry a jak moc pod se pohybovali. Přímka, která kopíruje dolní hranici nedosažitelné oblasti, znázorňuje případ, kdy by se hráč nechával neustále zrazovat od svého protihráče.



Obrázek 8: Vývoj výplat hráčů v porovnání s přímkami vzájemné spolupráce a vzájemné zrady

7.4.2 Hry člověk vs. algoritmus

V hrách člověk vs. algoritmus si vesměs hráči vedli lépe, a proto byla skóre počítána zvlášť pro tyto dva druhy her. Všechny algoritmy, které byly vybrány do turnaje, byly milé a nezačaly zrazovat jako první, takže si s nimi hráči mohli udržet spolupráci během celé hry. Opakovaným zrazováním algoritmu si ale mohli přijít hráči na vysoké skóre v případě, že jim byl vylosován jako soupeř algoritmus AllC a oni ho neustále zrazovali[80]. Nejhorší odpověď na pokusy zrazovat algoritmus dostali zřejmě hráči, kteří byli postaveni proti algoritmu Adaptivní Pavlov, který si je v rámci svých úseků zanalyzoval jako některou ze strategií, které s ním nechtějí spolupracovat. V takovém případě je už potom těžké navázat s AP spolupráci. Jak funguje APavlov bylo podrobně popsáno v **podkapitole 4.3**.

[80] jako např. hráč č. 3, který si tímto chováním přišel na první místo v tabulce

tabulka her hráč - algoritmus									
hráči	**průměrné skóre za 1 kolo hry**							**celk. prům. skóre**	**pořadí**
hrac1	2,0952	3,4	2,7476					**2,7476**	**11.**
hrac2	3	1,0952	2,0476					**2,0476**	**20.**
hrac3	5	2,5909	3,7955					**3,7955**	**1.**
hrac4	2,6	3,087	2,8435					**2,8435**	**10.**
hrac5	3	3,1739	3,087					**3,087**	**3.**
hrac6	2,9545	3,087	3,0208					**3,0208**	**4.**
hrac7	2,2857	2,3182	2,3019					**2,3019**	**18.**
hrac8	3,7826	1,2174	2,5					**2,5**	**14.**
hrac9	2,7391	2,9545	2,8468					**2,8468**	**9.**
hrac10	3	3	3					**3**	**5. - 7.**
hrac11	3	3	3					**3**	**5. - 7.**
hrac12	1	2,8	1,9					**1,9**	**23.**
hrac13	2,0952	2,7143	2,4048					**2,4048**	**15.**
hrac14	3,0909	3,0909	3,0909					**3,0909**	**2.**
hrac15	3,4	1,3913	2,3957					**2,3957**	**16.**
hrac16	3	3	3					**3**	**5. - 7.**
hrac17	1,6522	3	2,3261					**2,3261**	**17.**
hrac18	1,3636	3	2,1818					**2,1818**	**19.**
hrac19	3,4	2,5714	2,9857					**2,9857**	**8.**
hrac20	3,0909	2,0435	2,5672					**2,5672**	**13.**
APavlov	1,85	2,5217	1,7143	1,3333	2,3182	1,4348	3	**2,0246**	**21.**
Pavlov	1,4091	0,9091	2,0476	1,7826	0,4783	1,55		**1,3628**	**24.**
AllC	0	1,8261	3	2,4	2,7391			**1,993**	**22.**
TFT	1,8571	3	3	2,3636	2,9545	3		**2,6959**	**12.**

Tabulka 17: Výsledky turnaje her hráč vs. algoritmus

Tabulka 17 obsahuje výsledky z části turnaje, kdy byli proti sobě postaveni studenti a algoritmy. Opět jsou skóre přepočítávána na jedno kolo (jako aritmetický průměr), kvůli nejednotným délkám u různých her.

Průměrné skóre na 1 hráče pro tyto hry bylo 2,5883. Samotné algoritmy postavené proti lidským hráčům si vedly vesměs velmi špatně, výjimkou byl jen algoritmus TFT. K velkému překvapení se algoritmus AllC neumístil na posledním místě. Hráči totiž mohli po pár kolech zjistit, že se jedná o algoritmus, který nikdy nevolí nikdy zradu a mohli ho tedy "oškubat" bez obavy z odplaty. Na posledním místě se kupodivu umístil Pavlov, který obecně v turnajích bez lidí nemívá špatné výsledky.

7.5 Souhrnné informace pro všechny hry hráč vs. hráč

Zkusme se ještě podívat, jakému algoritmu by odpovídalo chování během celého turnaje[81], uvážíme-li všechny hráč vs. hráč současně. Turnaj bychom brali jako jednu hru o 1646 kolech[82] a počítali z ní podmíněné pravděpodobnosti spolupráce, přičemž jsme ale vždy vynechali první kola jednotlivých her, neboť mezi posledním kolem předchozí hry a prvním kolem hry následující dochází ke změně hráčů a není tedy na místě počítat zde podmíněné pravděpodobnosti spolupráce při určitých volbách hráčů v předchozím kole, protože tato kola na sobě nezávisí.

V **Tabulce 18** můžeme vidět počty spoluprací po jednotlivých situacích v předchozím kole, tj.: po kole, kdy oba hráči spolupracovali, kdy jeden spolupracoval a druhý ne a po kole, kdy se hráči vzájemně zradili. Ze všech her hráč vs. hráč současně (s vynecháním prvních kol) jsme spočítali tyto počty a poté podmíněné pravděpodobnosti spolupráce pro tyto jednotlivé situace, viz **Tabulka 19**. Na **Obrázku 9** jsou tyto pravděpodobnosti zakresleny do grafu.

	počet spoluprací	z celk.počtu
pro P_{CC}	379	442
pro P_{CD}	54	215
pro P_{DC}	76	215
pro P_{DD}	128	774

Tabulka 18: Počet spoluprací po jednotlivých situacích hry (CC, CD, DC, DD) pro celý turnaj[83]

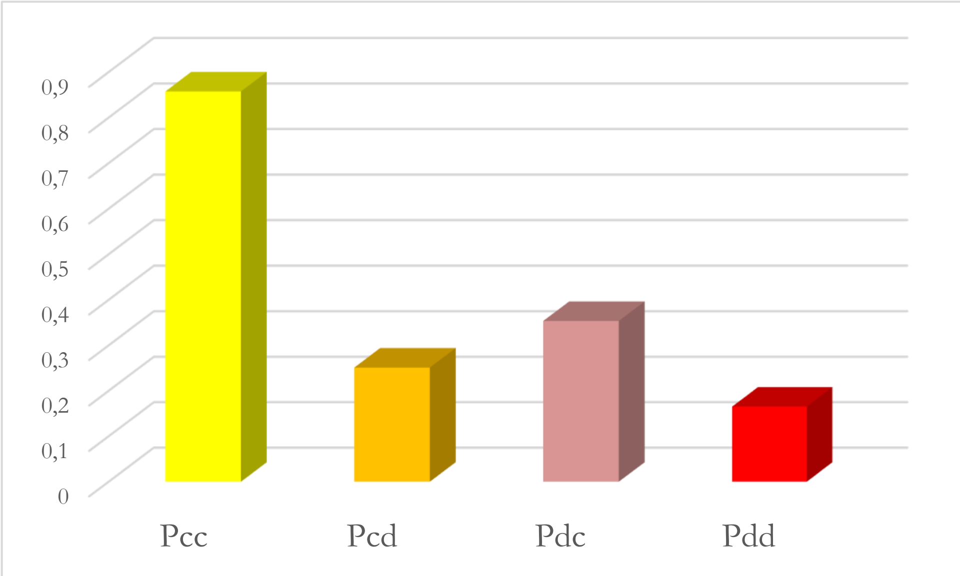
P_{CC}	0,8574661
P_{CD}	0,2511628
P_{DC}	0,3534884
P_{DD}	0,1653747

Tabulka 19: Podmíněné pravděpodobnosti spolupráce pro celý turnaj

V **Tabulce 20** vidíme analýzu chování hráčů během celého turnaje (her hráč vs. hráč). Opět jsme využili metodu nejmenší vzdálenosti. Jako nejbližší strategie nám pomocí obou vzdáleností vyšla strategie **RAND**.

Na **Obrázku 10**, máme graficky vyobrazen poměr mezi počtem spoluprací a zrad pro oddělené volby hráčů.

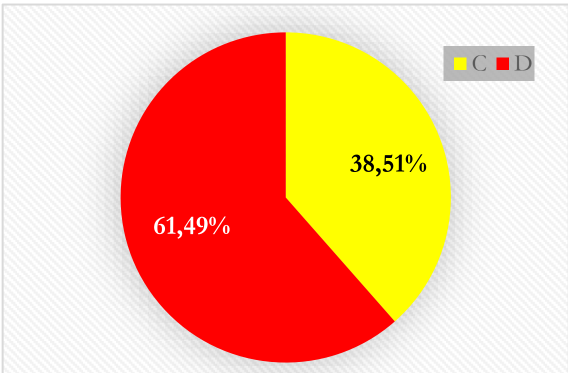
[81] Na základě minimální vzdálenosti
[82] Všechny řádky z her hráč vs. hráč jsme umístili pod sebe a analyzovali, jakoby se jednalo o 1 hru.
[83] Pro hry hráč vs. hráč



Obrázek 9: Souhrnný graf podmíněných pravděpodobností spolupráce pro všechny hry hráč vs. hráč

Referenční alg.	d_euklid	d_max
A a AllC	1,3022	0,8346
A a AllD	0,9750	0,8575
A a TFT	0,7271	0,6465
A a Pavlov	0,9513	0,8346
A a RAND	**0,5685**	**0,3575**

Tabulka 20: Hledání nejmenší vzdálenosti mezi chováním hráče a pěti algoritmy



Obrázek 10: Graf znázorňující poměr spolupráce a zrady pro hry hráč vs. hráč

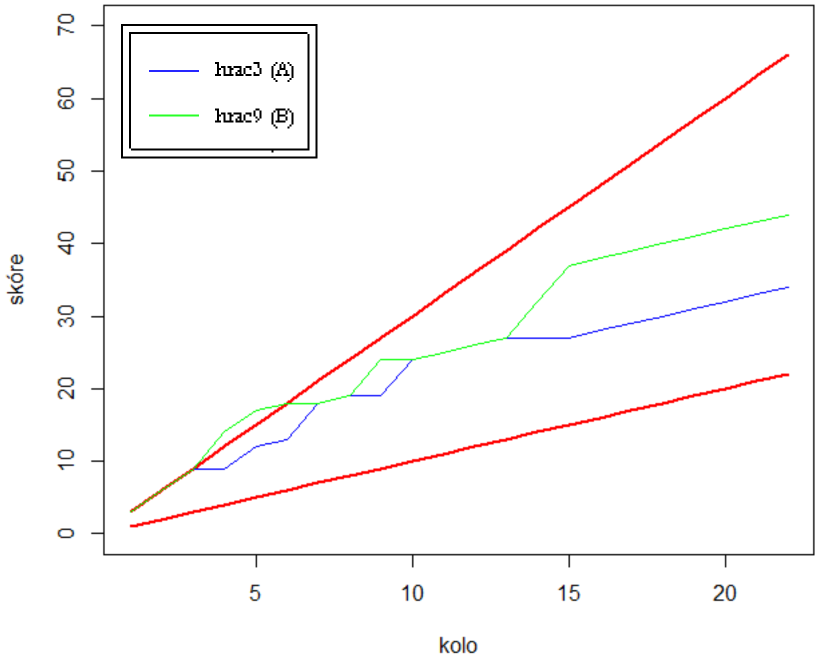
7.6 Výsledky analýzy dalších vybraných her turnaje

Na těchto pár obrázcích a grafech jsou ukázány výsledky analýzy zbylých 4 náhodně zvolených her. Analýza je k nalezení na přiloženém CD.

<u>**Analýza hry č. 2:**</u>

Hráči: hrac3, hrac9

ID této hry v souborech na přiloženém CD: 17



Obrázek 11: Graf vývoje výplat pro hru č. 2

Průběh hry č. 2 a také podrobnější analýza chování je vidět na další straně (**Obrázek 12**). **Obrázek 11** ukazuje, jak si vedli hráči během hry, opět navíc v porovnání se situacemi vzájemné spolupráce / vzájemného zrazování během celé hry.

Hra začala vzájemnou spoluprací hráčů. Ve třetím kole se hráč B rozhodl pro zradu svého soupeře, přičemž hráč i navzdory této zradě v dalším kole spolupracoval. To zřejmě vyvolalo u hráče B touhu zkusit zradit znovu. Nyní už se však jednalo o vzájemnou zradu. V 7. kole se hráč B pokoušel navázat spolupráci, ale marně. Dále se víceméně hráči už jen zrazovali. V 9. kole se hráč A rozhodl o spolupráci, bohužel neopětovanou. V dalším kole naopak spolupracoval hráč B, ale hráč A už ne. Ve 14. a 15. kole se dvakrát za sebou rozhodl spolupracovat hráč A, ale ani to nepomohlo navázat opět spolupráci. U hráče A došlo k jednostranné zradě 4 krát, u hráče B jen 2krát, tedy hráč A si přišel na skóre o 10 bodů vyšší.

ID: 17

kolo	hráč A	hráč B
1	C	C
2	C	C
3	C	C
4	C	D
5	C	C
6	D	D
7	D	C
8	D	D
9	C	D
10	D	C
11	D	D
12	D	D
13	D	D
14	C	D
15	C	D
16	D	D
17	D	D
18	D	D
19	D	D
20	D	D
21	D	D
22	D	D

d_euklid			
AllC, Pavlov, TFT, RAND	C	C	AllC, Pavlov, TFT, RAND
AllC, TFT, Pavlov	C	C	AllC, TFT, Pavlov
AllC, TFT, Pavlov	C	C	AllC, TFT, Pavlov
AllC, TFT, Pavlov	C	D	AllC, TFT, Pavlov
AllC	C	C	Pavlov
AllC	D	D	AllD, Pavlov, RAND
RAND	D	C	Pavlov
RAND	D	D	RAND
RAND	C	D	RAND
RAND	D	C	RAND
AllD, RAND	D	D	RAND
AllD	D	D	RAND
AllD	D	D	TFT
AllD, Pavlov	C	D	TFT
RAND	C	D	AllD, TFT
AllD	D	D	AllD
AllD, RAND	D	D	AllD
RAND	D	D	AllD
RAND	D	D	AllD
RAND	D	D	AllD
RAND	D	D	AllD
AllD, TFT	D	D	AllD

tabulka výplat (kumul.)	
3	3
6	6
9	9
9	14
12	17
13	18
18	18
19	19
19	24
24	24
25	25
26	26
27	27
27	32
27	37
28	38
29	39
30	40
31	41
32	42
33	43
34	**44**

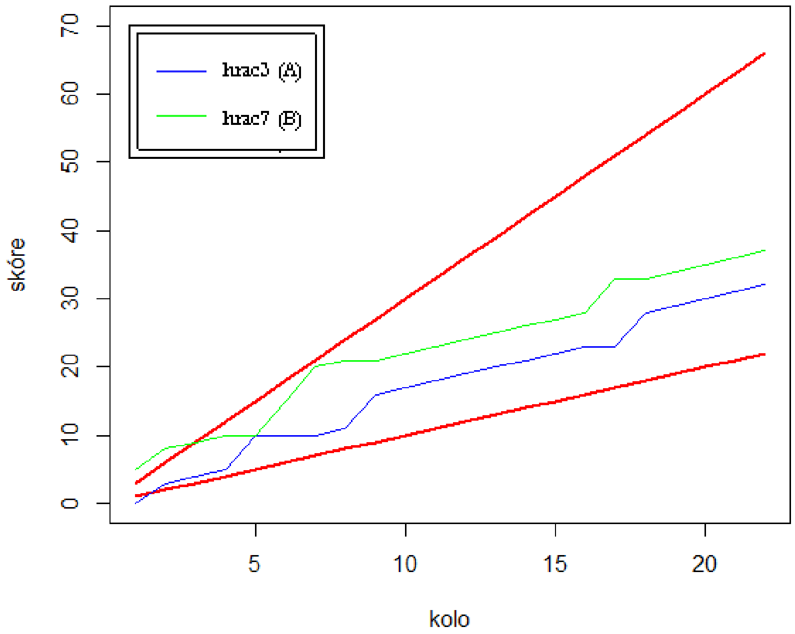
kolo	hráč A	hráč B
1	C	C
2	C	C
3	C	C
4	C	D
5	C	C
6	D	D
7	D	C
8	D	D
9	C	D
10	D	C
11	D	D
12	D	D
13	D	D
14	C	D
15	C	D
16	D	D
17	D	D
18	D	D
19	D	D
20	D	D
21	D	D
22	D	D

Obrázek 12: Tabulky pro hru č. 2 - vývoj hry, analýza nejbližších strategií v rámci oken, analýza pomocí 4 druhů chování

Analýza hry č. 3:

Hráči: hrac3, hrac7

ID této hry v souborech na přiloženém CD: 23



Obrázek 13: Graf vývoje výplat pro hru č. 3

Podrobnější analýza hry č. 3 je opět na další straně (**Obrázek 14**) a rovněž v příloze na CD. Průběh této hry vypadal zezačátku tak, že hráči se zkoušeli vzájemně střídavě zrazovat a čekali, zda jim toto chování „projde“. Od 10. kola už se víceméně zrazovali téměř neustále. V porovnání s červenými přímkami lze vidět, že hráči se od 10. kola drželi spíše v blízkosti dolní přímky, nelze sledovat jejich snahu o navázání vzájemné spolupráce.

Chování hráčů bylo zezačátku hry přirovnáváno především k algoritmu RAND, což lze poměrně pěkně vidět už ze samotné hry. Ke konci hry bylo naopak jejich chování přirovnáváno především k algoritmu AllD, což vidíme mj. i podle převládajícího počtu červených políček od 10. kola až k samotnému konci hry na **Obrázku 14**. Hráč A získal 32 bodů, hráč B 37 bodů během hry, což je z velké části zapříčiněno jejich neochotou spolupracovat.

ID: 23

kolo	hráč A	hráč B
1	C	D
2	C	C
3	D	D
4	D	D
5	D	C
6	C	D
7	C	D
8	D	D
9	D	C
10	D	D
11	D	D
12	D	D
13	D	D
14	D	D
15	D	D
16	D	D
17	C	D
18	D	C
19	D	D
20	D	D
21	D	D
22	D	D

d_euklid			
AllC, Pavlov, TFT, RAND	C	D	AllD, RAND
AllC	C	C	AllC, TFT
RAND	D	D	RAND
RAND	D	D	RAND
RAND	D	C	RAND
RAND	C	D	RAND
RAND	C	D	AllD, RAND
RAND	D	D	AllD
RAND	D	C	AllD
RAND	D	D	Pavlov
RAND	D	D	Pavlov
RAND	D	D	AllD
AllD	D	D	AllD
AllD	D	D	AllD
AllD	D	D	AllD, TFT
AllD	D	D	AllD, TFT
AllD, TFT	C	D	AllD, TFT
AllD, TFT	D	C	AllD
AllD	D	D	TFT
AllD	D	D	TFT
AllD	D	D	TFT
AllD	D	D	TFT

tabulka výplat (kumul.)	
0	5
3	8
4	9
5	10
10	10
10	15
10	20
11	21
16	21
17	22
18	23
19	24
20	25
21	26
22	27
23	28
23	33
28	33
29	34
30	35
31	36
32	**37**

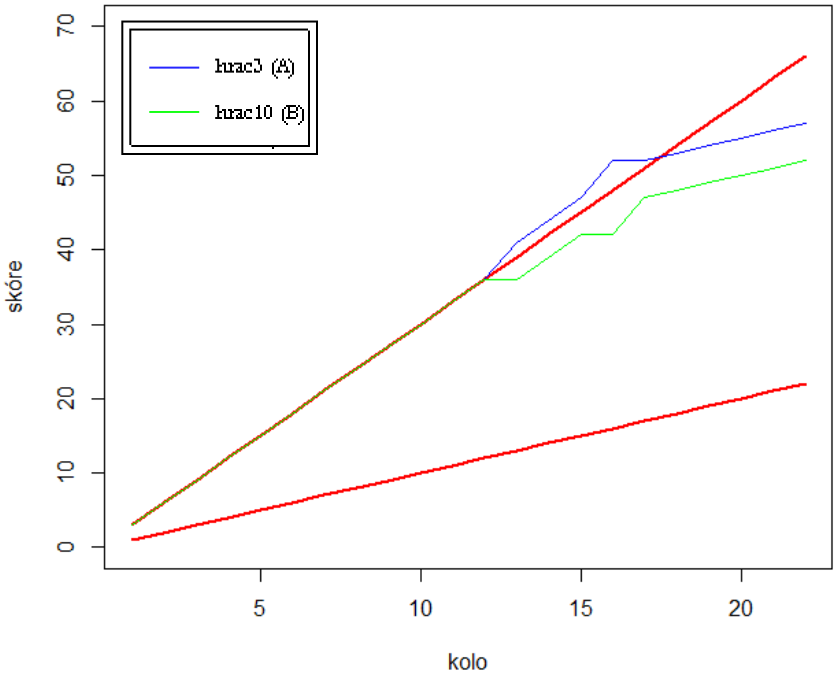
kolo	hráč A	hráč B
1	C	D
2	C	C
3	D	D
4	D	D
5	D	C
6	C	D
7	C	D
8	D	D
9	D	C
10	D	D
11	D	D
12	D	D
13	D	D
14	D	D
15	D	D
16	D	D
17	C	D
18	D	C
19	D	D
20	D	D
21	D	D
22	D	D

Obrázek 14: Tabulky pro hru č. 3 - vývoj hry, analýza nejbližších strategií v rámci oken, analýza pomocí 4 druhů chování

<u>Analýza hry č. 4:</u>

Hráči: hrac3, hrac10

ID této hry v souborech na přiloženém CD: 32



Obrázek 15: Graf vývoje výplat pro hru č. 4

Hra č. 4 oproti předchozím hrám je specifická v tom, že hráči první polovinu hry spolu vzájemně spolupracovali. Poté se dvakrát hráč A v 13. a 16. kole pokusil zradit, v 17. kole mu hráč B zradu oplatil a po zbylých 5 kol už se hráči jen zrazovali, možná kvůli očekávání blízkého konce hry. Oba hráči si přišli na poměrně pěkné výplaty, nicméně při vzájemné spolupráci během celé hry by hráči obdrželi vyšší skóre.

Pozn.: Výsledek klasifikace pro začátek hry je značně neurčitý. Jako algoritmy, ke kterým se hráči nejvíce blíží, jsou nejčastěji označeny zároveň AllC, TFT a Pavlov. To je způsobeno tím, že na začátku hry hráči stále spolupracují a klasifikační algoritmus nemá ještě informaci o tom, jak by se hráč zachoval, kdyby ho protihráč zradil. Proto jako možný výsledek označí všechny tři algoritmy, které na spolupráci odpovídají také spoluprací.

ID: 32

kolo	hráč A	hráč B
1	C	C
2	C	C
3	C	C
4	C	C
5	C	C
6	C	C
7	C	C
8	C	C
9	C	C
10	C	C
11	C	C
12	C	C
13	D	C
14	C	C
15	C	C
16	D	C
17	C	D
18	D	D
19	D	D
20	D	D
21	D	D
22	D	D

d_euklid			
AllC,TFT,Pavlov,RAND	C	C	AllC,TFT,Pavlov,RAND
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	D	C	AllC,TFT,Pavlov
AllC,TFT	C	C	AllC
AllC,TFT	C	C	AllC
AllC,TFT	D	C	AllC
AllC,TFT	C	D	AllC,TFT,Pavlov,RAND
TFT	D	D	Pavlov
TFT	D	D	RAND
TFT	D	D	RAND
TFT	D	D	AllD,TFT,Pavlov,RAND
AllD,TFT,RAND	D	D	AllD,TFT,Pavlov,RAND

tabulka výplat (kumul.)	
3	3
6	6
9	9
12	12
15	15
18	18
21	21
24	24
27	27
30	30
33	33
36	36
41	36
44	39
47	42
52	42
52	47
53	48
54	49
55	50
56	51
57	**52**

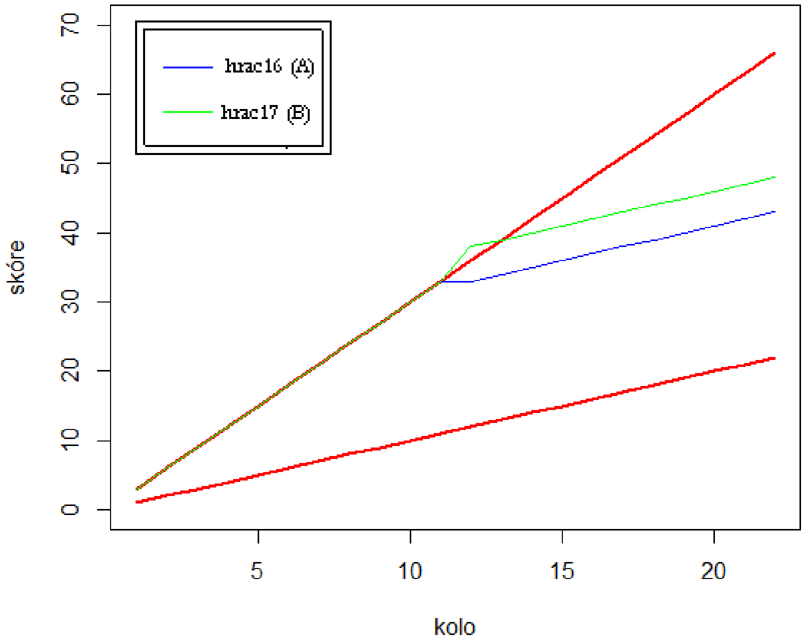
kolo	hráč A	hráč B
1	C	C
2	C	C
3	C	C
4	C	C
5	C	C
6	C	C
7	C	C
8	C	C
9	C	C
10	C	C
11	C	C
12	C	C
13	D	C
14	C	C
15	C	C
16	D	C
17	C	D
18	D	D
19	D	D
20	D	D
21	D	D
22	D	D

Obrázek 16 : Tabulky pro hru č. 4 - vývoj hry, analýza nejbližších strategií v rámci oken, analýza pomocí 4 druhů chování

Analýza hry č. 5:

Hráči: hrac16**,** hrac17

ID této hry v souborech na přiloženém CD: 77



Obrázek 17: Graf vývoje výplat pro hru č. 5

Poslední blíže analyzovaná hra je hra č. 5. Její průběh vidíme na následující stránce. Hráči opět začali vzájemnou spoluprací a spolupracovali spolu v prvních 11 kolech. Ve 12. kole se hráč B rozhodl vyzkoušet, jak bude hráč A reagovat na jeho zradu. Hráč A v dalším kole se rozhodl pomstít a až do konce hry se už hráči vzájemně zrazovali. Na **Obrázku 17** vidíme, že nejprve přímky výplat hráčů kopírovaly horní červenou přímku, poté se od této přímky začaly odchylovat, od 14. kola už byly kompletně pod horní „hranici".

ID: 77

kolo	hráč A	hráč B
1	C	C
2	C	C
3	C	C
4	C	C
5	C	C
6	C	C
7	C	C
8	C	C
9	C	C
10	C	C
11	C	C
12	C	D
13	D	D
14	D	D
15	D	D
16	D	D
17	D	D
18	D	D
19	D	D
20	D	D
21	D	D
22	D	D
23	D	D

d_euklid			
AllC,TFT,Pavlov,RAND	C	C	AllC,TFT,Pavlov,RAND
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	C	AllC,TFT,Pavlov
AllC,TFT,Pavlov	C	D	AllC,TFT,Pavlov
TFT,Pavlov	D	D	Pavlov
TFT	D	D	RAND
TFT	D	D	AllD, RAND
TFT	D	D	AllD
TFT	D	D	AllD
TFT	D	D	AllD
AllD, TFT	D	D	AllD
AllD, TFT	D	D	AllD, TFT
AllD, TFT	D	D	AllD, TFT
AllD, TFT	D	D	AllD, TFT
AllD, TFT	D	D	AllD, TFT

tabulka výplat (kumul.)	
3	3
6	6
9	9
12	12
15	15
18	18
21	21
24	24
27	27
30	30
33	33
33	38
34	39
35	40
36	41
37	42
38	43
39	44
40	45
41	46
42	47
43	48
44	**49**

kolo	hráč A	hráč B
1	C	C
2	C	C
3	C	C
4	C	C
5	C	C
6	C	C
7	C	C
8	C	C
9	C	C
10	C	C
11	C	C
12	C	D
13	D	D
14	D	D
15	D	D
16	D	D
17	D	D
18	D	D
19	D	D
20	D	D
21	D	D
22	D	D
23	D	D

Obrázek 18 : Tabulky pro hru č. 5 - vývoj hry, analýza nejbližších strategií v rámci oken, analýza pomocí 4 druhů chování

7.7 Poznatky při použití navrhovaných metod a možná zlepšení do budoucna

Výsledky turnaje závisí samozřejmě i na množině strategií. Při praktickém využití je nutné rozhodnout se, které strategie použít jako referenční.

Navrhovaná metoda klasifikuje hráče vždy. Dal by se také uvažovat přístup, který by odmítl hráče klasifikovat, pokud by jeho chování nebylo dostatečně blízké žádné z referenčních strategií. Opět by bylo ale potřeba uvážit, co ještě není, a co naopak už je, blízké některé z uvažovaných strategií, protože tato hranice může být velmi tenká.

Srovnejme nyní výsledky z našeho turnaje s některými výsledky turnajů, kdy byly proti sobě postaveny PC algoritmy. Nejlepší hráč v našem turnaji (ve hrách hráč vs. hráč – viz **Tabulka 11**), hrac9, získal průměrnou výplatu 2, 4341 bodu/kolo, což je téměř srovnatelná průměrná výplata s prům. výplatou TFT v 1. Turnaji R. Axelroda, tj. 2, 52 bodu/kolo.

Průměr celkových průměrných skóre za 1 kolo pro všechny hráče našeho turnaje byl 1,8843 bodu. Naopak nejhorší hráč, hrac12, si přišel na průměrné skóre ze všech her hráč vs. hráč na 1,3166 bodu a toto průměrné skóre lze přirovnat ke skóre algoritmu RAND v 1. Axelrodově turnaji. Opravdu lze sledovat u her, které hráč odehrál (viz Příloha na CD), že jeho chování lze považovat spíše za náhodné.

U her hráč vs. algoritmus (viz **Tabulka 17**) si nejlepší hráči přišli na vyšší skóre než nejlepší hráči u her hráč vs. hráč. Jedním z důvodů je, že všechny strategie postavené proti hráčům byli spolupracující strategie, AllC se dokonce nechalo oškubávat, proto hráč, který odhalil, že hraje proti tomuto algoritmu, mohl tento algoritmus oškubat a získat průměrné skóre za 1 kolo hry proti tomuto algoritmu 5 bodů.

Závěr

Cílem práce bylo čtenáře seznámit se základy, týkajícími se jak jednokolového, tak opakovaného vězňova dilematu, a dále, protože se jedná o poměrně rozsáhlé téma a nelze obsáhnout vše, se zaměřit na některé zajímavé pasáže z opakovaného vězňova dilematu. Protože se praktická část zaměřovala na reálná data získaná z turnaje opakovaného vězňova dilematu a využívala také některé známé algoritmy, bylo důležité čtenáře uvést do problematiky a představit mu některé známé algoritmy.

Nejznámějšími turnaji opakovaného vězňova dilematu, které kdy proběhly, byly ty první, turnaje Roberta Axelroda. Jeho dílo má stále obrovský vliv na řadu vědeckých článků a na vývoj nových algoritmů nebo modifikace algoritmů úspěšných. Protože většina turnajů, které kdy proběhly, a dá se o nich dočíst, byly turnaje mezi naprogramovanými algoritmy, chtěli jsme vyzkoušet, jak by si s takovýmto problémem poradili lidé, kdybychom je nechali hrát proti sobě.

Ukázalo se, že si většina lidí nevedla v turnajích moc dobře, což mohlo být způsobeno několika faktory. Mohlo se to týkat např. nedostatečné znalosti této problematiky některých lidí, nedůvěry vůči jiným hráčům apod. Když byli lidé postaveni proti algoritmům, vedli si zpravidla lépe, než když byli postaveni proti jinému člověku, neboť někteří dokázali chování algoritmu během hry identifikovat a mohli s ním poté optimálně interagovat, což není tak jednoduché u hry s člověkem, jehož strategie se může kolo od kola libovolně měnit. Přesto se našli lidé, kteří např. spolupracovali se strategií AllC, která s nimi spolupracovala i po několika zradách z jejich strany, přestože ji mohli oškubat a přijít si na průměrné skóre z jedné hry 5 bodů.

Tato diplomová práce mi hodně přinesla. Teorie her mě vždy zajímala. Problematika opakovaného vězňova dilematu, o kterém jsem mnoho před psaním této práce nevěděla, protože nebývá součástí kurzů ve škole ani obvykle součástí knih o teorii her, je velmi zajímavým tématem a rozhodně jsem za toto téma práce ráda.

Psaní práce bylo zajímavé, praktická část sice pracnější, zejména pak vymýšlení metod pro analýzu a ověřování funkčnosti na datech, neboť se nejedná o běžně zkoumanou problematiku, ale o to větší byla má motivace „vymyslet něco nového“.

Seznam použitých zdrojů

[1] AXELROD, R. M. *The evolution of cooperation*. Rev. ed. New York: Basic Books, 2006. ISBN 0465005642.

[2] KENDALL, G.; YAO, X.; SIANG, Y. C.: *The Iterated Prisoners' Dilemma*: 20 Years On. World Scientific Publishing, River Edge, NJ, USA, 2007. ISBN 9812706973.

[3] LUCE, R. D. a RAIFFA H. *Games and decisions: introduction and critical survey*. New York: Dover Publications, 1957. ISBN 0486659437.

[4] MAŇAS, M.: *Teorie her a její aplikace*, Praha, SNTL, 1991. ISBN 800300358X.

[5] MÉRŐ, L. *Moral calculations: game theory, logic, and human frailty*. New York, NY: Copernicus, c1998. ISBN 0387984194.

[6] OSBORNE, Martin J. *An Introduction to Game Theory*. New York: Oxford University Press, 2009. ISBN 0195322484.

[7] POUNDSTONE, W. Prisoner’s dilemma: *John Von Neumann, Game Theory and the Puzzle of the Bomb*, Anchor books, A division of random house, Inc., New York, 1992, ISBN 0385415675.

[8] RASMUSEN, Eric. *Games and information: an introduction to game theory*. 2nd ed. Cambridge, MA: B. Blackwell, 1994. ISBN 1557865027.

[9] STRAFFIN, P. D. *Game theory and strategy: by Philip D. Straffin*. Washington: Mathematical Association of America, 1993. ISBN 0883856376.

[10] VON NEUMANN, J., MORGENSTERN O.: *Theory of games & economic behavior, Princeton:* Princeton University Press, 1944, ISBN 0691003629.

[11] WILLIAMS, J. D. *The compleat strategyst being a primer on the theory of games of strategy*. New York: McGraw-Hill Book Company, 1966, ISBN 0486251011.

[12] AXELROD, R. Effective Choice in the Prisoner's Dilemma. *Journal of Conflict Resolution* [online]. 2016, **24**(1), 3-25 [cit. 2018-04-09]. DOI: 10.1177/002200278002400101. ISSN 0022-0027. Dostupné z: http://journals.sagepub.com/doi/10.1177/002200278002400101

[13] AXELROD, R. More Effective Choice in the Prisoner's Dilemma. *Journal of Conflict Resolution* [online]. 2016, **24**(3), 379-403 [cit. 2018-04-09]. DOI: 10.1177/002200278002400301. ISSN 0022-0027. Dostupné z: http://journals.sagepub.com/doi/10.1177/002200278002400301

[14] DE HERDT, Tom. *Cooperation and fairness: the Flood – Dresher experiment revisited.* Review of Social Economy [online]. 2003, **61**(2), 183-210 [cit. 2018-05-17]. DOI: 10.1080/0034676032000098219. ISSN 0034-6764. Dostupné z: http://www.tandfonline.com/doi/abs/10.1080/0034676032000098219

[15] FERGUSON, T. S., *Game theory, Class notes for math, Fall*, 2000 [online]. [cit. 2018-02-15]. Dostupné z: http://www.cs.cmu.edu/afs/cs/academic/class/15859-f01/www/notes/mat.pdf

[16] HYKŠOVÁ, M.. Teorie her [online]. České vysoké učení technické v Praze, Fakulta dopravní, 2007 [cit. 2017-09-29]. Dostupné z: http://physics.ujep.cz/~jskvor/MatematikaII/AplikovanaMatematikaCVUT/Teorie%20her/hry.pdf[5]

[17] MILINSKI, M. a WEDEKIND, C. Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat. *Proceedings of the National Academy of Sciences* [online]. 1996, **93**(7), 2686-2689 [cit. 2018-04-17]. DOI: 10.1073/pnas.93.7.2686. ISSN 0027-8424. Dostupné z: http://www.pnas.org/cgi/doi/10.1073/pnas.93.7.2686

[18] MILINSKI, M. a WEDEKIND, C. Working memory constrains human cooperation in the Prisoner's Dilemma. *Proceedings of the National Academy of Sciences* [online]. 1998, **95**(23), 13755-13758 [cit. 2018-04-17]. DOI: 10.1073/pnas.95.23.13755. ISSN 0027-8424. Dostupné z: http://www.pnas.org/cgi/doi/10.1073/pnas.95.23.13755

[19] SCODEL, A., MINAS, J. S., RATOOSH P., LIPETZ M. Some descriptive aspects of two-person non-zero-sum games. *Journal of Conflict Resolution* [online]. 2016, **3**(2), 114-119 [cit. 2018-04-19]. DOI: 10.1177/002200275900300203. ISSN 0022-0027. Dostupné z: http://journals.sagepub.com/doi/10.1177/002200275900300203

[20] SCODEL, A. Induced collaboration in some non-zero-sum games. *Journal of Conflict Resolution* [online]. 2016, **6**(4), 335-340 [cit. 2018-04-19]. DOI: 10.1177/002200276200600404. ISSN 0022-0027. Dostupné z: http://journals.sagepub.com/doi/10.1177/002200276200600404

[21] WU, J. a AXELROD R. How to Cope with Noise in the Iterated Prisoner's Dilemma. *Journal of Conflict Resolution* [online]. 2016, **39**(1), 183-189 [cit. 2018-05-06]. DOI: 10.1177/0022002795039001008. ISSN 0022-0027. Dostupné z: http://journals.sagepub.com/doi/10.1177/0022002795039001008

[22] YAO, X. a DARWEN, P. J. An experimental study of N-Person Iterated Prisoner's Dilemma games. YAO, Xin, ed. *Progress in Evolutionary Computation* [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995, 1995-6-2, s. 90-108 [cit. 2018-03-10]. Lecture Notes in Computer Science. DOI: 10.1007/3-540-60154-6_50. ISBN 3540601548. Dostupné z: http://link.springer.com/10.1007/3-540-60154-6_50

[23] *Dictionary of Game Theory Terms*, Game Theory .net, [online]. [cit. 2018-09-15]. Dostupné z: http://www.gametheory.net/dictionary

[24] *Median test* [online]. [cit. 2018-04-17]. Dostupné z: https://www.itl.nist.gov/div898/software/dataplot/refman1/auxillar/meditest.htm

[25] *Source code for axelrod.strategies.apavlov* [online]. [cit. 2019-01-07]. Dostupné z: https://axelrod.readthedocs.io/en/stable/_modules/axelrod/strategies/apavlov.html