UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

OPAKOVANÉ HRY. NEDOKONALÉ MONITOROVANIE.

DIPLOMOVÁ PRÁCA

2018

Bc. Matej Švec

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Opakované hry. Nedokonalé monitorovanie.

DIPLOMOVÁ PRÁCA

Študijný program: ekonomicko-finančná matematika a modelovanie
Študijný odbor: 1114 Aplikovaná matematika
Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky
Vedúci práce: doc. RNDr. Ján Pekár, PhD.

Bratislava 2018 **Matej Švec**

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:	Bc. Matej Švec
Študijný program:	ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor:	aplikovaná matematika
Typ záverečnej práce:	diplomová
Jazyk záverečnej práce:	slovenský
Sekundárny jazyk:	anglický

Názov: Opakované hry. Nedokonalé monitorovanie.
Repeated games. Imperfect monitoring.

Cieľ: Cieľom práce je zosumarizovať poznatky o opakovaných hrách s nedokonalou iínformáciou a vyriešiť niekoľko problémov.

Vedúci:	doc. RNDr. Ján Pekár, PhD.
Katedra:	FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry:	prof. RNDr. Daniel Ševčovič, CSc.
Dátum zadania:	26.01.2017

Dátum schválenia: 27.01.2017

prof. RNDr. Daniel Ševčovič, CSc.
garant študijného programu

.... študent

.... vedúci práce

Poďakovanie Touto cestou sa chcem poďakovať svojmu vedúcemu diplomovej práce doc. RNDr. Jánovi Pekárovi, PhD., ktorý ma ochotne počas konzultácíi usmerňoval v mojej práci, poskytoval mi odborné rady a venoval svoj čas.

Osobitná vďaka patrí aj mojej rodine a priateľom za ich podporu a trpezlivosť.

Abstrakt v štátnom jazyku

ŠVEC, Matej: Opakované hry. Nedokonalé monitorovanie [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: doc. RNDr. Ján Pekár, PhD., Bratislava, 2018, 69 strán.

Témou našej diplomovej práce je alternatívna väzňova dilema. Je to variant opakovej hry, v ktorej hráči na rozdiel od klasickej väzňovej dilemy ťahajú asynchrónne, teda striedajú sa po každom ťahu. Práca je rozdelená na dve časti, a to na teoretickú a praktickú. Teoretická časť sa skladá z prvých dvoch kapitol, kde sme si zadefinovali kritéria optimality pre hry s rozličným informačným šumom a následnej sme sa bližšie pozreli na Firm But Fair a Firm Pavlov triedy stratégií. V poslednej kapitole sme nasimulovali round robin turnaj, kde sme overili platnosť teoretických predpokladov z prvých kapitol.

Kľúčové slová: Alternatívna väznova dilema, kritéria optimality, herné stratégie

Abstract

ŠVEC, Matej: Repeated games. Imperfect monitoring [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: doc. RNDr. Ján Pekár, PhD., Bratislava, 2018, 69 pages.

The topic of our master thesis is alternative prisoner's dilemma. Alternative prisoner's dilemma is a variant of repated game where participants make their decisions in asynchrony. The thesis is divided into two parts, theoretical and practical. The theoretical part consists of two chapters, where we define optimal criteria for games with variuos noisy level and take closer look on Firm But Fair and Firm strategy classes. In the last chapter we simulate round robin turnament, where we confirm theoretical assumptions from the privious chapters.

Keywords: Alternative prisoner's dilemma, optimal criteria, games strategies

Obsah

Úvod

S procesom rozhodovania sa človek stretáva každý deň. Niekedy urobiť isté rozhodnutie môže byť náročné a ovplyvňuje ho veľa faktorov. V rozhodovacom procese si vyberáme z viacerých možností, pričom každá z nich nám prinesie istý úžitok. V spoločnosti sa veľa rozhodnutí vykoná len ako reakcia na rozhodnutie niekoho iného. Môžeme povedať, že naše rozhodnutie často nezávisí len od nás, ale je ovplyvňované predošlými rozhodnutiami iných ľudí. Toto platí nie len v súkromnom živote, ale s podobnou situáciou sa môžeme stretnúť napríklad aj na finančnom trhu. Tu je zväčša situácia ešte zložitejšia, nakoľko na ňom vládne obrovská konkurencia.

V našej diplomovej práci sa budeme venovať opakovaným hrám s nedokonalým monitorovaním. Nedokonalé monitorovanie v tomto prípade znamená, že do hry vstupuje takzvaný informačný šum. Informačný šum je pravdepodobnosť, že hráčove rozhodnutie bolo chápané nesprávne. Čo i len jedno zle chápané rozhodnutie môže výrazným spôsobom ovplyvniť ďalší priebeh hry. Už z názvu je jasné, že hráči sa v hre nebudú rozhodovať jednorazovo, ale opakovane. Ako príklad takejto situácie sme si zvolili alternatívnu väzňovu dilemu.

Práca je rozdelená do troch kapitol. V prvej si zadefinujem problém alternatívnej väzňovej dilemy a bližšie sa zoznámime s kritériami optimality v tejto hre pri rozličných úrovniach informačného šumu. V druhej kapitole si postupne rozšírime známe stratégie s 2- ťahovou pamäťou a zadefinujeme si dve nové skupiny stratégií. V poslednej tretej kapitole overíme naše teoretické predpoklady nasimulovaním round robin turnaja, kde otestujeme výkonnosť jednotlivých stratégií.

Cieľom našej diplomovej práce je nájsť stratégie, ktoré budú spĺňať podmienky optimality v alternatívnej väzňovej dileme. Následne poto pomocou simulácii overiť ich optimalitu v hre proti ostatným známym stratégiám.

1 Väzňová dilema

Väzňova dilema je simulácia klasického problému, ktorý sa vyskytuje v teórii hier, politických vedách, ale aj v teóriách celulárnych automatov. Táto hra bola sformulovaná v roku 1950 Merrillom Floodom a Melvinom Dresherom a je príkladom hry s nenulovým súčtom výplat. Hra je formulovaná nasledovne:

Dvaja väzni spolu vykradli banku, ale boli zadŕžaní políciou. Po zadŕžaní boli umiestnení do samostatných ciel bez možnosti komunikácie. Obaja dostanú ponuku od polície na spoluprácu (zradiť svojho komplica), za čo môžu získať nižší trest.

Tabuľka výplat je definovaná nasledovne:

		P1	
		C	D
P2	C	R/R	T/S
	D	S/T	P/P

Obr. 1: Tabulka výplat

Ak obidvaja budú navzájom spolupracovať (ani jeden nezradí toho druhého), obaja dostanú výplatu R (reward). Naopak ak obidvaja sa navzájom zradia, dostanú výplatu P (punishment). V prípade, že jeden hráč zradí, zatiaľ čo druhý bude spolupracovať, ten čo zradil, dostane výplatu T (templation) a hráč, čo spolupracoval, dostane výplatu S (sucker). Výplatou sa v tejto hre rozumie počet ušetrených rokov vo väzení. Vo väzňovej dileme platí $T > R > P > S$ a zároveň $2R > T + S$, z čoho vyplýva, že spolupráca sa oplatí viac ako alternovanie medzi výplatami T a S. [6]

1.1 Iterovaná väzňova dilema

Iterovaná väzňova dilema (Iterated prisoner's dilemma = IPD) je rozšírením štandardnej hry o opakovanie sa celej situácie, pričom ani jeden z hráčov/väzňov nevie, koľko krát sa hra bude opakovať, čiže nemôže plánovať viac ako jeden ťah dopredu. Stratégiou pre IPD je algoritmus, ktorý sa rozhoduje, či dané kolo spolupracovať alebo nespolupracovať na základe výsledkov z minulých kôl. Je očividné, že niektoré stratégie dosiahnu lepšie skóre ako ostatné. Ak hráč spolupracuje príliš často, druhý hráč toho môže ľahko využiť a naopak často zrádzať. Ak hráč naopak často zrádza a stratégia, proti ktorej hrá, si to všimne, vedie to k nízkemu profitu oboch hráčov. Pri iterovanej väzňovej dileme už nevieme tak jednoznačne určiť, ktorá stratégia je optimálna, nakoľko úspech stratégie závisí aj od stratégie druhého hráča a na špecifických parametroch hry[8].

V roku 1980 Robert Axelrod zorganizoval počítačový turnaj, v ktorom proti sebe bojovalo 14 stratégií. Tieto stratégie boli navrhnuté poprednými vedcami v oblasti teórií hier. Víťazná stratégia bola Tit for Tat, ktorú vymyslel Anatol Rapoport. Táto stratégia spolupracuje na začiatku a následne vždy opakuje posledný ťah súpera. Bohužiaľ táto stratégia prestáva byť efektívna, pokiaľ sa do IPD vnesie informačný šum, teda pravdepodobnosť, že hráčove rozhodnutie bude chápané nesprávne. V iterovanej väzňovej dileme medzi dvoma stratégiami Tit for Tat, len jedno zradenie vedie k nekonečnému opakovaniu tejto akcie a teda k nižším vyplátam. Preto pre takúto hru s informačným šumom boli vymyslené nové skupiny stratégií, z ktorých sa veľmi dobre ukázala Pavlova stratégia. Táto stratégia spolupracuje vždy, okrem prípadu, že v minulom kole obdržala výplatu S (sucker), vtedy zrádza. Ako sa ukázalo, táto stratégia je schopná, aj po náhodnom zradení, v pomerne krátkom čase zas dosiahnuť spoluprácu v hre so svojím klonom[7].

1.2 Alternatívna väzňova dilema

Alternatívna väzňova dilema (APD) je variant Iterovanej väzňovej dilemy (IPD) s tým rozdielom, že hráči neťahajú simultánne, ale ťahajú postupne jeden za druhým. Predstavme si túto hru medzi dvoma hráčmi A a B. Akcie hráča A budeme označovať veľkými písmenami (C alebo D) a akcie hráča B budeme označovať malými písmenami (c alebo d). Výplaty v tejto hre sú vyplácané po každej akcii hráča, čo znamená, že každý hráč dostane každé kolo dve výplaty. Predstavme si nasledujúcu APD hru:

1) Hráč A zahrá akciu C. Keďže je to úplne prvý ťah, žiadny hráč nedostane výplatu.

2) Hráč B zahrá akciu d. Nakoľko posledná akcia hráča A bola C, hráč A dostane výplatu S a hráč B dostane výplatu T.

3) Hráč A zahrá akciu D. Posledná akcia hráča B bola d, čiže obidvaja dostanú

výplatu P.

4) Hráč B zahrá akciu c. Posledná akcia hráča A bola D a teda hráč A dostane výplatu T a hráč B dostane výplatu S.

Môžeme vidieť, že po sérii akcií $CdDc$, každý hráč bude mať hodnotu výplat $T + P + S$. Tento proces pokračuje až do konca hry s tým, že výplaty sa určujú stále na základe posledných dvoch ťahov[9, 4].

Na prvý pohlaď sa môže zdať, že IPD a APD sú z matematického hľadiska rovnaké, ale súboj dvoch stratégii v týchto dvoch hrách vedie k signifikantne rozličným výsledkom. Napríklad v IPD jedna chyba medzi dvoma Tit for Tat stratégiami vedie k neustálemu striedaniu výplat T a S, zatiaľ čo v APD jedna chyba vedie k opakujúcej sa výplate P. Z toho vyplýva, že stratégie, ktoré sú úspešné v IPD, nemusia byť úspešné v APD a naopak. Ďalším takýmto príkladom je Pavlova stratégia. V IPD vie po malej chybe rýchlo napraviť vzájomnú spoluprácu ($CDDDCCCC\ldots$), ale v APD malá chyba vedie k 6 kolovému cyklu striedania výplat T, S a P ($CdDcDdCdDcDd$). Z toho vyplýva, že Pavlová stratégia je "self - cooperating" (schopná spolupracovať sama so sebou) v IPD, ale nie v APD. Stratégia Firm but Fair (FBF), ktorá spolupracuje vždy okrem výnimky, keď obdrží výplatu S, je zase naopak „self - cooperating" v APD, ale nie v IPD. V APD malá chyba týchto dvoch stratégií vedieť k pomerne rýchlej obnove spolupráce ($CdDcCc\ldots$), ale v IPD sa po chybe ich spolupráca už neobnoví a budú si striedať výplaty T a S ($CDDCCDDC\ldots$). Z týchto príkladov je jasné, že IPD a APD sú dva rozličné matematické problémy a "optimálne" stratégie v jednej hre môžu dosahovať veľmi slabé výsledky v hre druhej.

1.3 N- ťahové stratégie v alternatívnej väzňovej dileme

V tejto časti si zadefinujeme a predstavíme stratégie, ktoré sa často vyskytujú v literatúre týkajúcej sa alternatívnej väzňovej dilemy. N- ťahová história hry je reťazec skladajúci sa z písmen C a D $H = h_n, h_{n-1}, ..., h_2, h_1$, kde $h_k = C$ ak hráč spolupracoval k ťahov dozadu a $h_k = D$ ak hráč zradil k ťahov dozadu. Nepárne k reprezentuje akciu súpera a párne k reprezentuje akciu hráča. Stále zachováme konvenciu zápisu, súperove ťahy označujeme malými písmena a hráčove veľkými. Na príklad 3- ťahová história hry cCd znamená, že súperov posledný ťah bol zrada, hráčov posledný ťah bol spolupráca a súperov predošlý ťah bol spolupráca. Vidíme, že reťazec H môže obsahovať 2^n možnosti: $H_1 = C \dots Cc$, $H_2 = C \dots Cd, \dots$, $H_{2^n-1} = D \dots Dc$, $H_{2^n} = D \dots Dd$. N- ťahová stratégia pre APD sa rozhoduje či spolupracovať alebo nespolupracovať na základe n- rozmerného reťazca H. Táto stratégia je definovaná číslami 2^n $(a_1, \dots, a_{2^n})$, kde $0 \leq a_i \leq 1$ pre všetky i a každé a_i predstavuje pravdepodobnosť spolupráce ak $H = H_i$. Ako príklad si uvedieme 2- ťahovú stratégiu $(1, 0, 1, 0.2)$. Táto stratégia spolupracuje s pravdepodobnosťou 1 ak $H = H_0$ alebo ak $H = H_2$, čo je v prípade ak posledné dva ťahy boli Cc alebo Dc. Zrádza, ak posledné dva ťahy boli Cd a spolupracuje s pravdepodobnosťou 0.2, ak posledné ťahy boli Dd. Takýmto spôsobom vieme ľahko zapísať známe 1 aj 2- ťahové stratégie. Príklad 1- ťahových: $ALLC(1,1)$, $ALLD(0,0)$, TFT(1,0) a náhodná stratégia $(\frac{1}{2}, \frac{1}{2})$. Príklad 2- ťahových: $Pavlov(1,0,0,1)$, Firm but Fair$(1,0,1,1)$ a stratégie $g1(1,1/3,1,1/2)$, $g2(1,1/2,1,1/2)$, $g3(1,2/3,1,1/2)$. 2- ťahovú stratégiu vieme dostať pomocou triviálneho rozšírenia 1- ťahovej stratégie. 2- ťahová stratégia (a,b,a,b) sa správa rovnako ako 1- ťahová (a,b)[2, 10]. Vo všeobecnosti vieme povedať, že n- ťahová stratégia S_n je ekvivalentná n+1- ťahovej stratégii, ktorá je spojená sama so sebou: $S_{n+1} = S_n \oplus S_n$.

Pre úplné zadefinovanie n- ťahových stratégií, by sme si mali zadefinovať ich správanie na začiatku hry (ťahy 1…n). Vo všeobecnosti predpokladáme, že $h_k = C$ pre $k \geq$i, čo znamená, že história "ťahov" pred tým než sa hra začala, bola spolupráca. Napríklad, keď sa pri prvom ťahu hry pozeráme na 3- ťahovú históriu reťazca H, dostávame: $H = cCc$. Respektíve môžeme si určiť ako sa bude stratégia na začiatku správať, alebo urobiť si priemer zo všetkých počiatočných možností. Vo všeobecnosti správanie n- ťahových stratégií na začiatku hry je relevantné iba pre prípad hry bez informačného šumu[9].

1.4 Kritéria optimality

Predtým, než si zadefinujeme kritéria optimality, pokúsime sa načrtnúť, čo znamená byť "úspešný" pre stratégiu v alternatívnej väzňovej dileme. Predstavme si jednoduchú APD hru medzi dvoma stratégiami X a Y s konečným počtom kôl N. Cieľom každej stratégie je maximalizovať priemernú hodnotu výplat,

$$w = \frac{1}{N-1}\sum_{i=2}^{N} p(i) \tag{1}$$

kde $p(i)$ predstavuje výplatu dosiahnutú v $i - tom$ ťahu. Podobne vieme tento vzorec zapísať aj pre APD hru s nekonečným počtom kôl,

$$w = \lim_{N\to\infty}\frac{1}{N-1}\sum_{i=2}^{N} p(i) \tag{2}$$

Očakávanú hodnotu výplat stratégie X oproti stratégii Y v nekonečnej hre budeme značiť ako $w(X|Y)$. Taktiež si zadefinujeme očakávanú hodnotu výplat w v hre, kde hrá stratégia X proti svojmu klonu ako $w(X|X)$. V niektorých prípadoch (napríklad pri hre dvoch deterministických stratégií bez informačného šumu) môže hodnota výplat závisieť od toho, ktorá stratégia ide prvá. V týchto prípadoch preto spriemerujeme obe možnosti. Predpokladáme, že stratégia X bude postupne čeliť rozličným stratégiám Y. Cieľom tejto stratégie X bude maximalizovať hodnotu výplat proti celému spektru stratégií Y. Predpokladajme konečnú množinu protihráčov $Y = Y_1, \ldots, Y_M$ potom:

$$w(X) = \frac{1}{M}\sum_{i=1}^{M} w(X|Y_i) \tag{3}$$

V populácii oponentov Y_i s meniacou sa frekvenciou $f(Y_i)$, kde $\sum_i f(Y_i) = 1$:

$$w(X) = \sum_i f(Y_i) w(X|Y_i) \tag{4}$$

A v spojitom "priestore" oponentov S:

$$w(S) = \int_S f(s) w(X|s) ds \tag{5}$$

Relatívny úspech dvoch stratégii X a Y teda vieme porovnať cez $w(X)$ a $w(Y)$. Stratégiu s vyšším w považujeme za úspešnejšiu. Jasne ale vidíme, že úspech danej stratégie v APD hre závisí od toho, proti akej stratégii hrá[8]. Našim cieľom je nájsť stratégiu, ktorá dosahuje dobré výsledky proti celému spektru stratégií a určiť kritériá optimality, ktoré silno korelujú s vysokou priemernou hodnotou výplat. Vo všeobecnosti by stratégia X mala spolupracovať so stratégiou Y, ak ich spolupráca v tomto kole napomáha spolupráci aj v tom ďalšom. Predstavme si, že stratégia Y je 1- ťahová stratégia (a, b), čiže s pravdepodobnosťou a spolupracuje po tom, čo stratégia X spolupracovala a s pravdepodobnosťou b zrádza po tom, čo stratégia X zradila. Potom stratégia X pri stálej spolupráci dosiahne priemernú hodnotu výplat $a * R + (1 - a) * S$ a pri stálom zrádzaní $b * T + (b - 1) * P$. Teda vieme povedať, že stála spolupráca má zmysel ak:

$$aR + (1-a)S > bT + (1-b)P \Rightarrow a > \left(\frac{T-P}{R-S}\right)b + \frac{P-S}{R-S} \tag{6}$$

Samozrejme, neustála spolupráca alebo neustále zrádzanie nie sú jediné možnosti stratégií. Práve naopak, daná stratégia proti väčšine stratégií dosiahne najvyššie skóre práve striedaním týchto dvoch akcií. Bez ohľadu na tento fakt platí všeobecný princíp. Stratégia by mala byť schopná nadviazať bezpodmienečnú spoluprácu ak vidí, že druhá stratégia je ochotná spolupracovať a naopak mala by byť schopná zabrániť "vykorisťovaniu", ak druhá stratégia príliš často zrádza. Jedna z možností, ako si zadefinovať kritéria optimality stratégie, je porovnať skóre proti stratégiám $ALLC$ a $ALLD$ [4, 8, 1]. Dostávame tri základné charakteristiky, ktoré by mala optimálna stratégia spĺňať:

1) $Self-cooperating$ schopnosť dosiahnuť spoluprácu so svojim klonom
2) $C-exploiting$ schopnosť využívať spoluprácu súpera (ALLC)
3) $D-unexploitable$ schopnosť odolávať zrádzaniu (ALLD)

1.5 Kritéria optimality pre APD hru bez informačného šumu

V APD hre bez informačného šumu budeme funkciu výplat $w(X|Y)$ značiť ako $w_0(X|Y)$, aby sme ju rozlíšili od hry s informačným šumom.

Ako prvú si zadefinujeme vlastnosť $Self-cooperating$, kde hodnota výplat $w_0(X|X)$ musí byť medzi P a R vrátane.

Definícia č.1 Stratégia X je $Self-cooperating$, ak $w_0(X|X) > P$ a je úplne $Self-cooperating$, ak $w_0(X|X) = R$.

Druhú vlastnosť $C-exploiting$ si zadefinujeme pomocou výplaty proti stratégii $ALLC$. $w_0(X|ALLC)$ musí byť medzi R a T vrátane.

Definícia č.2 Stratégia X je $C-exploiting$, ak $w_0(X|ALLC) > R$ a je úplne

$C - exploiting$, ak $w_0(X|ALLC) = T$.

Poslednú, tretiu vlastnosť zadefinujeme pomocou výplaty proti stratégii $ALLD$. $w_0(X|ALLD)$ musí byť medzi S a P vrátane.

Definícia č.3 Stratégia X je $D - unexploitable$, ak $w_0(X|ALLD) > S$ a je úplne $D - unexploitable$, ak $w_0(X|ALLD) = P$.

Vo všeobecnosti vieme zapísať "relatívnu úspešnosť" stratégie X oproti stratégii Y ako:

$$\sigma_Y(X) = \frac{w(X|Y) - inf_Z w(Z|Y)}{sup_Z w(Z|Y) - inf_Z w(Z|Y)} \tag{7}$$

Z toho vyplýva, že $\sigma_Y(X) = 1$, ak stratégia X dosiahla najvyššie možné skóre proti stratégii Y a naopak $\sigma_Y(X) = 0$, ak dosiahla najnižšie možné. Teraz si zadefinujeme rovnakú "relatívnu úspešnosť" stratégie X ale proti stratégiám $ALLC$ a $ALLD$ a budeme ju značiť σ_C respektíve σ_D:

$$\sigma_C = \sigma_{ALLC}(X) = \frac{w(X|ALLC) - w(ALLC|ALLC)}{w(ALLD|ALLC) - w(ALLC|ALLC)} = \frac{w(X|ALLC) - R}{T - R} \tag{8}$$

$$\sigma_D = \sigma_{ALLD}(X) = \frac{w(X|ALLD) - w(ALLC|ALLD)}{w(ALLD|ALLD) - w(ALLC|ALLD)} = \frac{w(X|ALLD) - S}{P - S} \tag{9}$$

Teda stratégia je $C - exploiting$ ak $\sigma_C > 0$ a úplne $C - exploiting$ ak $\sigma_C = 1$. Rovnako stratégia je $D - unexploitable$ ak $\sigma_D > 0$ a úplne $D - unexploitable$ ak $\sigma_D = 1$. Taktiež si zadefinujeme "relatívnu úspešnosť" stratégie samej proti sebe:

$$\sigma_S = \frac{w(X|X) - w(ALLD|ALLD)}{w(ALLC|ALLC) - w(ALLD|ALLD)} = \frac{w(X|X) - P}{R - P} \tag{10}$$

Vo všeobecnosti nemusí ale nutne platiť, že σ_S sa rovná $\sigma_X(X)$. Predstavme si stratégiu $B(0,1)$. Pre túto stratégiu platí: $inf_Z w(Z|B) = w(ALLC|B) = S$,

$sup_Z w(Z|B) = w(ALLD|B) = T$ a $w(B|B) = (S+T)/2$. Potom pre tabuľku výplat s hodnotami(T=5,R=3,P=1,S=0) dostávame:

$$\sigma_S = \frac{w(B|B) - P}{R - P} = \frac{2.5 - 1}{3 - 1} = \frac{3}{4} \tag{11}$$

$$\sigma_B(B) = \frac{w(B|B) - S}{T - S} = \frac{2.5 - 0}{5 - 0} = \frac{1}{2} \tag{12}$$

Teraz si určíme aké σ hodnoty musí stratégia nadobúdať, aby sme ju považovali za „optimálnu“. Dokonale optimálna stratégia X musí nadobúdať maximálne možné skóre proti všetkým stratégiám Y, teda $\sigma_Y(X) = 1$. Bohužiaľ, jednoduchým argumentom ale vieme ukázať, že takáto stratégia neexistuje[12, 20].

Veta č.1 Dokonale optimálna stratégia neexistuje.

Dôkaz. Predstavme si stratégiu $GRIM$, ktorá spolupracuje až kým protihráč prvý krát nezradí, potom už len zrádza. Žiadna stratégia nevie dosiahnuť maximálne možné skóre aj proti stratégii $GRIM$ aj proti stratégii $ALLC$. Ak stratégia aspoň raz zradí po vzájomnej spolupráci, dosiahne nízke skóre proti stratégii $GRIM(\sigma_{GRIM}(X) = 0)$ a naopak ak nikdy po vzájomnej spolupráci nezradí, dosiahne nízke skóre proti stratégii $ALLC(\sigma_C = 0)$. □

Teraz budeme uvažovať stratégie, ktoré sú optimálne v interakcii proti spolupráci a proti zrádzaniu. Na tento zápis použijeme Maynard Smith kritéria [13]:

$w(X|ALLC) \leq w(ALLC|ALLC)$

$w(ALLC|X) \geq w(ALLC|ALLC)$

$w(X|ALLD) \geq w(ALLD|ALLD)$

$w(ALLD|X) \leq w(ALLD|ALLD)$

Aby stratégia X porážala stratégiu $ALLC$, musí platiť $w(X|ALLC) > w(ALLC|ALLC)$, čo implikuje $\sigma_C > 0$. Podobne, aby stratégia X porážala stratégiu $ALLD$, musí platiť $w(X|ALLD) = w(ALLD|ALLD)$ a $w(X|X) > W(ALLD|X)$, čo implikuje $\sigma_D = 1$ a $\sigma_S > 0$.

Na to, aby stratégia X zabránila invázií od stratégii $ALLC$, musí platiť $w(X|X) > w(ALLC|X)$. Nahradením $w(X|X)$ za σ_s a $w(ALLC|X)$ za σ_C dostávame:

$$R - (R - S)\sigma_C \leq P + (R - P)\sigma_S \rightarrow \sigma_C \geq \frac{R - P}{R - S}(1 - \sigma_S) \tag{13}$$

Na to, aby stratégia X zabránila invázií od stratégii $ALLD$, musí platiť $w(X|X) > w(ALLD|X)$. Nahradením $w(X|X)$ za σ_s a $w(ALLD|X)$ za σ_D dostávame:

$$T - (T - P)\sigma_D < P + (R - P)\sigma_S \rightarrow \sigma_D > 1 - \frac{R - P}{T - P}\sigma_S \tag{14}$$

Pomocou týchto vzorcov si zadefinujeme "silno optimálnu" stratégiu, ktorá vie porážať a zároveň zabraňovať "vykorisťovaniu" od iných stratégií, vie dosiahnuť spoluprácu sama so sebou a vie zabrániť invázii od stratégií $ALLC$ a $ALLD$. A "slabo optimálnu" stratégiu, ktorá spĺňa rovnaké vlastnosti až na stupeň spolupráce samej so sebou, ktorý je na nižšej úrovni[4, 6].

Definícia č.4 Silno optimálna stratégia je stratégia s hodnotami $\sigma_D = 1$,$\sigma_S = 1$ a $\sigma_C > 0$.

Definícia č.5 Slabo optimálna stratégia je stratégia s hodnotami $\sigma_D = 1$,$\sigma_S > 0$ a $\sigma_C > \dfrac{R-P}{R-S}(1-\sigma_S)$.

Podľa týchto kriterií teraz porovnáme najznámejšie stratégie vo väzňovej dileme. Čísla v zátvorkach budú znamenať pravdepodobnosť spolupráce po výplatach R, S, T, P a g je "štedrosť" stratégií GTFT a FBF. V nasledujúcej tabuľke sú vypočítané hodnoty σ_S, σ_C a σ_D pre stratégie v APD hre bez informačného šumu.

Stratégia	σ_S	σ_C	σ_D
TFT (1 0 1 0)	1	0	1
GTFT (1 g 1 g)	1	0	1 -g
ALLC (1 1 1 1)	1	0	0
ALLD (0 0 0 0)	0	1	1
PAV(1 0 0 1)	1	0	1/2
FBF(1 0 1 g)	1	0	$\dfrac{1}{1+g}$

Tabuľka 1: σ hodnoty pre APD bez informačného šumu

Ako môžeme vidieť z tabuľky, žiadna z týchto stratégií nespĺňa ani slabé podmienky optimality. Stratégia TFT je k tomu ale najbližšie, nakoľko si dokáže poradiť s častým zrádzaním, ale nevie dosiahnuť bezpodmienečnú spoluprácu, čím nie je úplne efektívna voči stratégii $ALLC$. Treba si ale všimnúť, že všetky tieto stratégie si pamätajú len 2 a menej ťahov dozadu. V skutočnosti žiadna stratégia, ktorá reaguje na menej ako 3 ťahy dozadu nie je silno optimálna.

Lemma č.1 2- ťahovú stratégiu $X = (a, b, c, d)$ kde $d \neq 0$, vie stratégia $ALLD$ zneužiť s pravdepodobnosťou najmenej $\frac{d}{1+d}$.

Dôkaz. V akomkoľvek ťahu bude stratégia X buď spolupracovať (obdrží výplatu S) alebo zrádzať (obdrží výplatu P). Stratégia X spolupracuje s pravdepodobnosťou d, ak predošlí ťah bol zradenie a s pravdepodobnosťou b, ak predošlí ťah bola spolupráca. Keďže každé nenulové b len zvýši pravdepodobnosť zneužitia stratégie X môžeme položiť $b = 0$, čiže stratégia X vždy po spolupráci zradí. Teda stratégia X bude $1/d$ času tráviť zamietaním, z čoho vyplýva pravdepodobnosť zneužitia:

$$\frac{1}{1+1/d} = \frac{d}{1+d} \tag{15}$$

□

Veta č.2 Žiadna 0, 1 ani 2- ťahová stratégia nie je silno optimálna v APD hre bez informačného šumu

Dôkaz. Každú 0- ťahovú stratégiu $S_0 = (a)$ vieme zapísať ako 2- ťahovú stratégiu (a, a, a, a). Každú 1- ťahovú stratégiu $S_1 = (a, b)$ vieme zapísať ako 2- ťahovú stratégiu (a, b, a, b). Predpokladajme 2- ťahovú stratégiu (a, b, c, d) s pravdepodobnosťou počiatočnej spolupráce e. Ak $b = 1$, stratégia $ALLD$ vie túto stratégiu zneužiť v každom kole. Zároveň vieme, že ak $d \neq 0$, tak stratégia $ALLD$ vie túto stratégiu zneužiť minimálne s pravdepodobnosťou: $\frac{d}{1+d}$ a $\sigma_D \leq \frac{1}{1+d}$. Teda aby sa $\sigma_D = 1$ tak $b < 1$ a $d = 0$. Môže byť ale stratégia, ktorá má $b < 1$ a $d = 0$ schopná spolupráce so svojim klonom? Akákoľvek zrada vedie s nenulovou pravdepodobnosťou $1 - b$ k nekonečnému cyklu vzájomného zrádzania, čiže musíme zabezpečiť, aby k zradeniu nikdy nedošlo. A teda a aj e sa musí rovnať 1. Ale stratégia s hodnotami $a = 1$ a $e = 1$ zas naopak neustále spolupracuje so stratégiou $ALLC$, z čoho vyplýva, že žiadna 2- ťahová stratégia, kde $\sigma_D = \sigma_S = 1$ nemôže mať zároveň $\sigma_C > 0$. 2- ťahová stratégia môže byť maximálne slabo optimálna. Predstavme si $GRIM$ stratégiu $(1, 0, 0, 0)$ s pravdepodobnosťou počiatočnej spolupráce $e = \frac{1}{3}$. Vypočítané hodnoty σ sú potom: $\sigma_S = \frac{1}{9}$,$\sigma_C = \frac{2}{3}$ a $\sigma_D = 1$, čo spĺňa podmienky slabej optimality. Pre nájdenie stratégií, ktoré spĺňajú podmienky silnej optimality, sa budeme musieť pozrieť na stratégie s vyššou pamäťou. □

1.6 Kritéria optimality pre APD hru s minimálnym informačným šumom

V tejto kapitole sa budeme venovať APD hre s informačným šumom. Informačný šum je definovaný ako nenulová pravdepodobnosť ϵ, že sa stratégia v danom kole dopustí chyby. Teda stratégia zradí, keď má spolupracovať a naopak bude spolupracovať, keď má zradiť s pravdepodobnosťou ϵ. V APD hre s informačným šumom budeme funkciu výplat značiť $w_\epsilon(X|Y)$. V APD hre s minimálnym informačným šumom je možné, že hráčove rozhodnutie bude chápané nesprávne, ale pravdepodobnosť, že sa to stane, je blízka nule. Minimálny informačný šum môžeme považovať za hraničný prípad konečného šumu, kde pravdepodobnosť šumu vieme zapísať ako $\epsilon \to 0$. Teda funkcia výplat $W(X|Y)$ v APD hre s minimálnym informačným šumom sa rovná: $\lim_{\epsilon\to 0} W_\epsilon(X|Y)$ [16, 5]. Pomocou tejto limity vieme rozšíriť definície $C - exploiting$, $D - unexploitable$ a $Self - cooperating$ z predošlej kapitoly.

Definícia č.6 Stratégia X je $Self - cooperating$ v hre s minimálnym informačným šumom, ak $\lim_{\epsilon\to 0} w_\epsilon(X|X) > P$ a je úplne $Self - cooperating$, ak $\lim_{\epsilon\to 0} w_\epsilon(X|X) = R$.

Definícia č.7 Stratégia X je $C - exploiting$ v hre s minimálnym informačným šumom, ak $\lim_{\epsilon\to 0} w_\epsilon(X|ALLC) > R$ a je úplne $C - exploiting$,
ak $\lim_{\epsilon\to 0} w_\epsilon(X|ALLC) = T$.

Definícia č.8 Stratégia X je $D - unexploitable$ v hre s minimálnym informačným šumom, ak $\lim_{\epsilon\to 0} w_\epsilon(X|ALLD) > S$ a je úplne $D - unexploitable$,
ak $\lim_{\epsilon\to 0} w_\epsilon(X|ALLD) = P$.

Taktiež si rozšírime definície pre σ.

$$\sigma_S = \frac{\lim_{\epsilon\to 0} w_\epsilon(X|X) - P}{R - P} \tag{16}$$

$$\sigma_C = \frac{\lim_{\epsilon \to 0} w_\epsilon(X|ALLC) - R}{T - R} \tag{17}$$

$$\sigma_D = \frac{\lim_{\epsilon \to 0} w_\epsilon(X|ALLD) - S}{P - S} \tag{18}$$

Kritéria „silnej“ a „slabej“ optimality pre σ_S,σ_C a σ_D sú definované rovnako ako pri APD hre bez informačného šumu.

Je dôležité si uvedomiť, že kvalita stratégie môže byť signifikantne odlišná v prípade hry bez informačného šumu ($\epsilon = 0$) a v prípade hry s minimálnym informačným šumom ($\epsilon \to 0$). Ako príklad si predstavme stratégiu TFT v APD hre v súboji so svojím klonom. V hre bez informačného šumu budú tieto dve stratégie neustále spolupracovať, čím dosiahnu $\sigma_S = 1$. V hre s minimálnym informačným šumom jedna chyba bude viesť ku kolotoču vzájomného zrádzania s výplatou P, až kým druhá chyba neobnoví spoločnú spoluprácu. V nekonečne dlhej hre môžeme očakávať, že TFT stratégia proti svojmu klonu bude polovicu času zrádzať a polovicu času spolupracovať, čo vedie k priemernej výplate $\frac{R+P}{2}$. Teda pred TFT stratégiu v hre s informačným šumom sa $\sigma_S = 1/2$. Rovnakým spôsobom teraz dopočítame σ_S,σ_C a σ_D pre všetky stratégie ako v minulej kapitole.

Stratégia	σ_S	σ_C	σ_D
TFT	1/2	0	1
$GTFT$	1	0	1 -g
$ALLC$	1	0	0
$ALLD$	0	1	1
PAV	1/2	1/2	1/2
FBF	1	0	$\frac{1}{1+g}$

Tabuľka 2: σ hodnoty pre APD s minimálnym informačným šumom

Z tabuľky vidíme, že žiadna zo stratégií nie je ani "slabo" optimálna. Rovnako ako v prípade APD hry bez informačného šumu aj teraz vieme dokázať, že žiadna stratégia s menej ako 3- ťahovou históriou nie je "silno" optimálna. V APD hre s minimálnym informačným šumom vieme ale dokázať, že žiadna z týchto stratégií nie je ani “slabo” optimálna.

Veta č.3 Žiadna 0,1 ani 2- ťahová stratégia nie je optimálna v APD hre s minimálnym informačným šumom.

Dôkaz. Rovnako ako v hre bez informačného šumu predstavme si 2- ťahovú stratégiu (a, b, c, d). Ukážeme, že takáto stratégia nemôže mať hodnoty $\sigma_S > 0$,$\sigma_C > 0$ a $\sigma_D = 1$ a teda, že nie je ani slabo optimálna. Na to aby $\sigma_D = 1$, musí byť pravdepodobnosť vstupu stratégie proti stratégii ALLD do hrania akcii Dd vyššia ako pravdepodobnosť, že z hrania týchto akcií odstúpi aspoň o $O(\epsilon)$. Z toho dostávame $1-b = O(1)$ a $d = O(\epsilon)$ a teda $b < 1$ a $d = 0$. Na to aby $\sigma_C > 0$, musí byť pravdepodobnosť vstupu stratégie proti stratégii ALLC do hrania akcií Dc vyššia ako pravdepodobnosť, že z hrania týchto akcií odstúpi aspoň o $O(\epsilon)$. Z toho dostávame $1 - a = O(1)$ alebo $d = O(\epsilon)$ a teda $a < 1$ alebo $c = 0$. Vznikli tri možnosti ako dostať hodnoty $\sigma_C > 0$ a $\sigma_D = 1$:

1. $a < 1$, $b < 1$, $c > 0$, $d = 0$. V hre proti svojmu klonu, stratégia s pravdepodobnosťou $O(1)$ začne hrať akcie Dd a s pravdepodobnosťou $O(\epsilon)$ skončí hranie týchto akcií, čiže $\sigma_s = 0$.

2. $a < 1$, $b < 1$, $c = 0$, $d = 0$. V hre proti svojmu klonu, stratégia s pravdepodobnosťou $O(1)$ začne hrať akcie Dd a s pravdepodobnosťou $O(\epsilon^2)$ skončí hranie týchto akcií, čiže $\sigma_s = 0$.

3. $a = 1$, $b < 1$, $c = 0$, $d = 0$. V hre proti svojmu klonu, stratégia s pravdepodobnosťou $O(\epsilon)$ začne hrať akcie Dd a s pravdepodobnosťou $O(\epsilon^2)$ skončí hranie týchto akcií, čiže $\sigma_s = 0$.

Vidíme, že žiadna úplne $D - unexploitable$ 2- ťahová stratégia nemôže byť zároveň aj $Self - cooperating$ aj $C - exploiting$[3, 6]. Preto aj pri hľadaní optimálnych stratégii pre APD hru s minimálnym informačným šumom sa musíme pozrieť na stratégie s väčšou pamäťou. □

1.7 Kritéria optimality pre APD hru s konečným informačným šumom

V tejto časti budeme rozoberať APD hru s informačným šumom na úrovni $0 < \epsilon < 1/2$. Budeme sledovať ako sa mení hodnota výplat v závislosti od funkcie ϵ. Očakávaná hodnota výplat dvoch stratégií, ktoré majú tendenciu spolupracovať, je v hre s informačným šumom nižšia ako R. Ak dve stratégie majú tendenciu spolupracovať s pravdepodobnosťou chyby ϵ, v skutočnosti to znamená, že spolupracujú s pravdepodobnosťou $1-\epsilon$ a zrádzajú s pravdepodobnosťou ϵ [4, 21]. Teda očakávaný výnos pre každú stratégiu po vzájomnej spolupráci je:

$$E_{cc} = (1-\epsilon)^2 R + \epsilon^2 P + \epsilon(1-\epsilon)(T+S) \tag{19}$$

Rovnako očakávaná hodnota výplat dvoch stratégií, ktoré majú tendenciu zrádzať, je v hre s informačným šumom vyššia ako P. Ak dve stratégie majú tendenciu zrádzať s pravdepodobnosťou chyby ϵ, v skutočnosti to znamená, že zrádzajú s pravdepodobnosťou $1-\epsilon$ a spolupracujú s pravdepodobnosťou ϵ. Teda očakávaný výnos pre každú stratégiu po vzájomnej zrade je:

$$E_{dd} = (1-\epsilon)^2 P + \epsilon^2 R + \epsilon(1-\epsilon)(T+S) \tag{20}$$

Taktiež vieme zapísať E_{dc} - stratégiu, ktorá neustále "vykorisťuje" druhú stratégiu a E_{cd} - stratégiu, ktorá je neustále "vykorisťovaná" druhou stratégiou.

$$E_{dc} = (1-\epsilon)^2 T + \epsilon^2 S + \epsilon(1-\epsilon)(R+P) \tag{21}$$

$$E_{cd} = (1-\epsilon)^2 S + \epsilon^2 T + \epsilon(1-\epsilon)(R+P) \tag{22}$$

Teraz si vypočítame konkrétne hodnoty pre tabuľku výplat (T, R, P, S). Pre (5,3,1,0) dostávame $E_{cc} = 3 - \epsilon - \epsilon^2$,$E_{dd} = 1 + 3\epsilon - \epsilon^2$, $E_{dc} = 5 - 6\epsilon - \epsilon^2$ a $E_{cd} = 4\epsilon + \epsilon^2$. Pre (4,3,1,0) dostáveme $E_{cc} = 3 - 2\epsilon$,$E_{dd} = 1 + 2\epsilon$, $E_{dc} = 4 - 4\epsilon$ a $E_{cd} = 4\epsilon$.

Rozšírime definíciu σ_S,σ_C a σ_D pre hru s konečným informačným šumom nahradením R za E_{cc}, P za E_{dd}, T za E_{dc}, S za E_{cd}.

$$\sigma_S = \frac{w_\epsilon(X|X) - E_{dd}}{E_{cc} - E_{dd}} \tag{23}$$

$$\sigma_C = \frac{w_\epsilon(X|ALLC) - E_{cc}}{E_{dc} - E_{cc}} \tag{24}$$

$$\sigma_D = \frac{w_\epsilon(X|ALLD) - E_{cd}}{E_{dd} - E_{cd}} \tag{25}$$

Vieme, že $\lim_{\epsilon\to 0} \sigma_S(\epsilon)$ je rovná σ_s definovanej pre minimálny šum. Rovnako to platí aj pre σ_C a σ_D. Na základe týchto definícií vieme zapísať očakávané hodnoty σ pre stratégie $ALLC$ a $ALLD$. Pre $ALLC$ platí $\sigma_S = 1$, $\sigma_C = \sigma_C = 0$. Pre $ALLD$ platí $\sigma_S = 0$, $\sigma_C = \sigma_C = 1$. Avšak takmer pre všetky stratégie hodnoty σ závisia od funkcie ϵ. Vo väčšine prípadov σ_C rastie a σ_S a σ_D klesá s rastúcim ϵ. Ako príklad si vypočítame hodnoty σ pre stratégiu $FBF(1,0,1,1)$ ako funkciu ϵ. Hodnoty pre stratégiu v hre s minimálnym informačným šumom boli $\sigma_S = 1$, $\sigma_C = 0$ a $\sigma_D = 1/2$.

ϵ	σ_S	σ_C	σ_D
0.001	0.999	0.001	0.499
0.010	0.990	0.007	0.495
0.050	0.953	0.034	0.475
0.100	0.913	0.064	0.450
0.200	0.843	0.111	0.396

Tabuľka 3: σ hodnoty stratégie FBF v závislosti od ϵ

Využitím Markovho reťazca [12, 14], vieme vypočítať σ hodnoty stratégie ako funkciu v závislosti od ϵ. Napríklad pre tabuľku výplat (5,3,1,0) dostávame:

$$W(FBF|FBF) = \frac{3 + 3\epsilon}{1 + 2\epsilon} \tag{26}$$

$$\sigma_S = \frac{2 + 2\epsilon - \epsilon^2}{2 + 4\epsilon} \tag{27}$$

Teraz, keď vidíme, že hodnoty σ závisia od ϵ, musíme si zadefinovať nové pravidlá, kedy je stratégia $C-exploiting$, $D-unexploitable$ a $Self-cooperating$. Jedným zo spôsobov je určiť si takzvanú konštantu "noise resistance threshold" (hraničná hodnota voči informačnému šumu) NRT. Teda pre hocijaké ϵ platí:

Definícia č.7 Stratégia je $Self - cooperating$ v hre s informačným šumom ϵ ak $\sigma_S \geq NRT$.

Definícia č.8 Stratégia je $C - exploiting$ v hre s informačným šumom ϵ ak $\sigma_C \geq NRT$.

Definícia č.9 Stratégia je $D - unexploitable$ v hre s informačným šumom ϵ ak $\sigma_D \geq NRT$.

Ak položíme $NRT = 0.9$ a hodnotu $T = 5$, z tabuľky výplat dostanem, že stratégia FBF je $self - cooperating$ ak $\epsilon \leq 0.117$, ale nie je $D - unexploitable$. Stratégia TFT zase na opačnej strane je $D - unexploitable$ pre $\epsilon \leq 0.697$, ale nie je $self - cooperating$.

Okrem zadefinovania optimálneho kritéria pre hodnotu ϵ sa pozrieme ako sa bude správať funkcia výplat $w_\epsilon(X|Y)$ s meniacim sa ϵ. Zadefinujeme si koeficient odolnosti voči informačnému šumu ako NRC, ktorý bude predstavovať spodnú hranicu pre w_ϵ ako lineárnu funkciu od ϵ. To znamená, že $w_{\epsilon=e} \geq w \rightarrow 0 - (NRC)e$ pre všetky e [22, 3]. Vieme, že level informačného šumu sa môže pohybovať od 0 po nejakú maximálnu hodnotu, ktorú si označíme ako ϵ_{max}. Na teraz si túto hodnotu položíme $\epsilon_{max} = 0.4$. Dostávame vzorec:

$$NRC = \sup_{e=0 \ldots \epsilon_{max}} \frac{w_{\epsilon \rightarrow 0} - w_{\epsilon=e}}{e} \tag{28}$$

Zoberieme stratégie, ktoré sú úplne $self - cooperating$ v hre s minimálnym informačným šumom. Pre tieto stratégie platí, $w(X|X)_{\epsilon \rightarrow 0} = R$. Zadefinujeme si NRC_S ako koeficient voči informačnému šumu pre $self - cooperation$:

$$NRC_S = \sup_{e=0 \ldots \epsilon_{max}} \frac{R - w(X|X)_{\epsilon=e}}{e} \tag{29}$$

'

Akonáhle nájdeme hodnotu NRC_S, tento koeficient nám určí spodnú hranicu hodnoty výplat stratégie pre všetky $\epsilon \leq \epsilon_{max}$. Vieme, že $w(X|X) \geq R - (NRC_S)\epsilon$. Teda ak má stratégia $NRC_S = 5$ a $R = 3$, potom vieme, že $w(X|X) \geq 3 - 5\epsilon$. Pre stratégie, ktoré nie sú uplné $self - cooperating$ v hre s minimálnym informačným šumom, platí $NRC_S = \infty$.

Z vyššie uvedeného vyplýva, čím má stratégia nižšiu hodnotu koeficientu NRC_S, tým je stratégia odolnejšia voči informačnému šumu. Bližšie sa pozrieme, aká by musela byť hodnota koeficientu NRC_S, aby bola stratégia odolná voči konečnému informačnému šumu. Najprv si predstavme, k akej hodnote výplat vedie jedna jediná chyba stratégie. Pre stratégiu $ALLC$ proti svojmu klonu, táto chyba vedie k opakujúcemu sa cyklu hrania akcií $CdCcCc$. Ten, čo zradil, získa $(2T - 2R)$ a ten, čo spolupracoval, stratí $(2R - 2S)$. Čiže táto chyba v priemere na hráča predstavuje stratu $2R - (T + S)$. Pre klasickú tabuľku výplat (5,3,1,0) potom priemerne každý z hráčov stratí $2*3-(5+0) = 1$ bod za chybu a $w \approx 3 - \epsilon$ a $NRC_S \approx 1$. Pre stratégiu FBF proti svojmu klonu chyba vedie k opakujúcemu sa cyklu hrania akcií $CdDcCc$. Každý z hráčov v priemere stratí $3R - (T + S + P) = 3$ body za chybu, z čoho vyplýva, že $w \approx 3 - 3\epsilon$ a $NRC_S \approx 3$. Vo všeobecnosti jedna jediná chyba vedie k N početnému zrádzaniu, pri ktorom hráč priemerne stratí $(N + 1)R - (N - 1)P - (T + S)$ bodov za chybu. Určíme si hranicu, pri ktorej je stratégia odolná voči konečnému informačnému šumu, ak jedna chyba v priemere nevedie k viac ako dvom zradám. Z toho vyplýva:

Definícia č.10 Stratégia je odolná voči konečnému informačnému šumu ak $NRC_S \leq 3R - P - T - S$.

Pre tabuľku výplat (5,3,1,0) dostávame $NRC_S = 1+2(N-1)$ a stratégia je odolná voči konečnému informačnému šumu ak $NRC_S \leq 3$. Pre tabuľku výplat (4,3,1,0) dostávame $NRC_S = 2 + 2(N - 1)$ a stratégia je odolná voči konečnému informačnému šumu ak $NRC_S \leq 4$. Teraz zavedieme novú definíciu pre stratégie, ktoré sú optimálne v hre s konečným informačným šumom[3, 9].

Definícia č.11 Stratégia X je optimálna pri konečnom informačnom šume, ak je voči tomuto informačnému šumu odolná a je uplné $D - unexploitable$ a $C - exploiting$ v hre s minimálnym informačným šumom.

Na to, aby bola stratégia optimálna v hre s konečným informačným šumom, musí spĺňať všetky kritéria silnej optimality z predošlých kapitol a zároveň musí platiť, že $NRC_S \leq 3$ (pre T=5) a $NRC_S \leq 4$ (pre T=4). Bohužiaľ, žiadna z doteraz spomínaných stratégií tieto podmienky nespĺňa. Takéto stratégie sa pokúsime nájsť neskôr.

1.8 Kritéria optimality pre APD hru s ľubovoľným informačným šumom

Predstavme si stratégiu s 3-ťahovou históriou X=(1 0 1 1, 1 0 1 0) v hre s konečným náhodným informačným šumom ϵ. Ak nastanú dve chyby hneď po sebe, hra sa zacyklí v Dd cykle a stratégie nebudú schopné obnoviť spoluprácu, až kým nenastane ďalšia chyba: $CcCc\underline{D}d\underline{D}dDd...Dd\underline{C}cCc$. Vo všeobecnosti sa hra zacyklí v cykle Dd každých $1/\epsilon^2$ ťahov a v tomto cykle strávi $1/\epsilon$ ťahov. Keďže čas strávený v cykle Dd vieme približne vyjadriť ako ϵ, $\lim_{\epsilon\to 0} w_\epsilon(X|X) = R$, stratégia X je úplne $self-cooperating$ v hre s minimálnym informačným šumom[23, 3].

V skutočnosti, ale pravdepodobnosť, že nastane chyba ϵ je v každej hre iná. Predstavme si znova stratégiu X v hre proti svojmu klonu. V najhoršom prípade chyba nastane pri prvom a treťom ťahu a potom sa už nikdy nezopakuje. História výplat by vyzerala: $\underline{D}d\underline{D}dDd.....$ Spolupráca nebude obnovená a obaja hráči budú dostávať totožnú výplatu P. Stratégia, ktorá je naozaj "nepriestrelná" voči informačnému šumu, musí byť schopná obnoviť vzájomnú spoluprácu po ľubovoľne konečnej postupnosti chýb. Predtým, ako si sformulujeme túto definíciu do matematickej podoby, zadefinujeme si, čo znamená konečná postupnosť chýb. Konečnú postupnosť chýb budeme označovať ako ω. ω bude reprezentovať konečný reťazec 0 a 1 $\omega_1, \omega_2, \omega_3 ... \omega_n$, pričom $\omega_i = 1$ ak chyba nastala v i- tom ťahu a $w_i = 0$ ak chyba nenastala. Konečný počet chýb, ktoré nastali : $\sum_n^{i=1} w_i$, sú potom norma Ω postupnosti chýb. Pre danú postupnosť chýb ω si tiež zadefinujeme ako stratégia na tieto chyby reaguje. D_ω je očakávaný počet hrania akcie D (zrada) v hre proti svojmu klonu, bez toho aby nastali iné chyby:

$$D_\omega = \sum_{k=0}^{\infty} kProb(D = k) \tag{30}$$

Tiež si zadefinujeme "koeficient nepriestrelnosti voči informačnému šumu" NPC_S. NPC_S predstavuje maximálny počet hrania akcie D (zrada) spôsobených jednou chybou:

$$NPC_S = \max_{\forall\omega} \frac{D(\omega)}{\Omega} \tag{31}$$

Stratégia bude "nepriestrelná" voči informačnému šumu, ak žiadna postupnosť chýb nespôsobí viac ako 2 hrania akcie D v priemere na jednu chybu.

Definícia č.12 Stratégia X je "nepriestrelná" voči informačnému šumu ak $NPC_S \leq 2$.

Z tejto definície vyplýva, že takáto stratégia rovnako ako musí byť odolná voči konečnému informačnému šumu, musí byť zároveň aj úplne $self-coorperating$ voči minimálnemu informačnému šumu. Vypočítame si teraz NPC_S pre viaceré stratégie. Pre stratégiu $ALLC$ počet hraní akcie D; D_ω zapríčinené postupnosťou chýb ω je rovné číslu Ω. Teda $NPC_S = 1$ pre stratégiu ALLC a táto stratégia je "nepriestrelná" voči informačnému šumu. Pre stratégiu TFT jediná chyba ($\omega = 1$) vedie k $D(\omega) = \infty$. Potom aj $NPC_S = \infty$ a stratégia TFT nie je "nepriestrelná" voči informačnému šumu. Nakoniec si zoberieme stratégiu FBF, kde jedna chyba vedie k 2 ťahom hrania stratégie D a teda žiadna postupnosť chýb nebude viesť k viac ako dvom zradeniam v priemere na chybu. $NPC_S = 2$ pre FBF čiže táto stratégia je "nepriestrelná" voči informačnému šumu. Keď si ale zoberieme FBF stratégie s meniacou sa "štedrosťou" (1,0,1,g) pre $0 \leq g \leq 1$ dostávame $NPC_S = max(1 + 1/g, 2)$. Na to, aby aj tieto stratégie boli "nepriestrelné" voči informačnému šumu, musí platiť $g = 1$.

Veta č.4 Žiadna stratégia s konečnou pamäťou nie je aj "nepriestrelná" voči informačnému šumu aj úplne $D - unexploitable$.

Dôkaz. Predstavme si stratégiu X s n- ťahovou pamäťou, kde n je nejaké konečné číslo. Majme "štedrosť" g stratégie, ktorá predstavuje spoluprácu po n zradách. Hra dvoch rovnakých stratégií s $g = 0$ v hre proti sebe, čo i len po jednej zrade vedie k nekonečnému cyklu zrád. Inými slovami, pre postupnosť chýb ω s $\Omega \leq n$, vieme, že $D(\omega) = \infty$. Potom tieto stratégie majú $NPC_S = \infty$ a nie sú "nepriestrelné" voči informačnému šumu. Ak $g > 0$, potom stratégia X vie byť zneužitá stratégiou $ALLD$ minimálne v nejakom úseku hry. Vypočítame, aký najmenší úsek hry (teda koľko ťahov) sa musí stratégia X vedieť brániť voči stratégií $ALLD$ s výnimkou cyklu n zradení. Každé zneužitie od stratégie $ALLD$ vedie k $n/2 - 1$ zradám od stratégie X plus k priemerne $1/g$ zradám predtým, než stratégia X začne znova spolupracovať. Stratégia X bude zrádzať proti stratégii $ALLD$ s pravdepodobnosťou:

$$\sigma_D = \frac{\lfloor \frac{n}{2} \rfloor - 1 + \frac{1}{g}}{\lfloor \frac{n}{2} \rfloor + \frac{1}{g}} = \frac{(\lfloor \frac{n}{2} \rfloor - 1)g + 1}{\lfloor \frac{n}{2} \rfloor g + 1} \tag{32}$$

Z toho vyplýva, že ak $g > 0$, stratégia má $\sigma_D < 1$ a stratégia nie je úplne $D-unexploitable$. Iba stratégie s nekonečnou pamäťou môžu byť zároveň "nepriestrelné" voči informačnému šumu aj úplne $D-unexploitable$ [16, 27, 24]. □

Definícia č.12 Stratégia s nekonečnou pamäťou, ktorá je "nepriestrelná" voči informačnému šumu, $D-unexploitable$ a zároveň aj $C-exploiting$ je optimálna v APD hre s ľubovoľným informačným šumom.

2 Nové stratégie

V tejto kapitole sa budeme sústrediť na dve hlavne úlohy. Prvou bude nájsť stratégiu s konečnou pamäťou ťahov, ktorá bude optimálna v APD hre s konečným informačným šumom a tou druhou bude nájsť stratégiu s nekonečnou pamäťou ťahov, ktorá bude optimálna v APD hre s ľubovoľným informačným šumom. Každá z týchto stratégii bude musieť byť úplne $D - unexploitable$, úplne $self - cooperation$ a aspoň slabo $C - exploiting$ v hre s minimálnym informačným šumom. Z matematického hľadiska to bude znamenať, že každá stratégia spĺňa $\sigma_S = \sigma_D = 1$ a $\sigma_C > 0$ pre $\epsilon \to 0$. Rovnako musia byť odolné voči konečnému informačnému šumu, čiže $NRC_S \leq 3$ pre tabuľku výplat (5,3,1,0) a $NRC_S \leq 4$ pre tabuľku výplat (4,3,1,0). Pre optimálnu stratégiu v hre s ľubovoľným informačným šumom musí platiť ešte jedno dodatočné kritérium a to stratégia musí byť "nepriestrelná"voči informačnému šumu $NPC_S \leq 2$.

2.1 Stratégie prvej zrady

Teraz si zadefinujeme "Stratégie prvej zrady"pre stratégie s nekonečnou pamäťou. Takáto n-ťahová stratégia sa v APD hre rozhoduje, či spolupracovať alebo zrádzať na základe n- ťahovej histórie H spolu s takzvaným "memory bit"b. Tento *bit* zaznamenáva, ktorý hráč naposledy bezdôvodne zradil. $b = 1$ vždy, keď je 2- ťahová história Cd (protihráč bezdôvodne zradil) a $b = 0$ vždy, keď je 3- ťahová história cDc alebo cDd (hráč bezdôvodne zradil). Stratégia je definovaná číslami $2^n + 1(a_1, ..., a_{2^n+1})$, kde $0 \leq a_i \leq 1$ pre všetky i. Pre $i = 1...2^n - 1$, a_i reprezentuje pravdepodobnosť spolupráce ak $H = H_i$. Ak $H = H_{2^n}$ (v n- ťahovej histórii boli všetko zrady), či už použijeme a_{2^n} alebo a_{2^n+1} pravdepodobnosť spolupráce je a_{2^n} ak $b = 0$ a a_{2^n+1} ak $b = 1$. Ako príklad si predstavme 2- ťahovú stratégiu (1 0 0 [1 0]). Táto stratégia bude zrádzať, ak 2- ťahová história hry bude Cd alebo Dc a bude spolupracovať, ak 2- ťahová história hry bude Cc. V prípade histórie Dd, bude spolupracovať, ak $b = 0$ (hráč bezdôvodne naposledy zradil) a zrádzať, ak $b = 1$ (protihráč bezdôvodne naposledy zradil). Tieto stratégie budeme označovať ako FD_∞, aby sme ich odlíšili od skupiny pre stratégie s konečnou pamäťou FD_n [12, 17].

2.2 Aproximácia pre stratégie s konečnou pamäťou

Predstavme si k- ťahovú stratégiu "prvej zrady"$FD_\infty = (a_1, ..., a_{2^k+1})$, potom n-ťahová aproximácia FD_n je definovaná nasledovne. Pre $n = k$, $FD_n = (a1, ...a_{2^k-1}, g)$, kde g predstavuje "štedrosť"stratégie (pravdepodobnosť spolupráce po n zradách). 2-ťahová aproximácia ku (1 0 0 [1 0]) by bola 2- ťahová stratégia (1 0 0 g).

Pre $n > k$ definujeme FD_n rekurzívne spojením dvoch kópií

FD_{n-1}: $FD_n = FD*_{n-1} \oplus FD_{n-1}$, kde $FD*_{n-1}$ je definované ako FD_{n-1} s:

$$g = \begin{cases} a_{2^k} \text{ ak n je nepárne} \\ a_{2^k+1} \text{ ak n je párne} \end{cases}$$

Z toho vyplýva, že 3- ťahová aproximácia ku (1 0 0 [1 0]) by bola (1 0 0 1)⊕(1 0 0 g)=(1 0 0 1, 1 0 0 g) a 4- ťahová aproximácia by bola (1 0 0 1, 1 0 0 0)⊕(1 0 0 1, 1 0 0 g)=(1 0 0 1, 1 0 0 0, 1 0 0 1, 1 0 0 g). Takýmto spôsobom vieme napísať akúkoľvek aproximáciu k- ťahovej stratégie FD_∞ pomocou n- ťahovej aproximácie FD_n pre hocijaké $n \geq k$ [29, 30].

2.3 Firm But Fair (FBF) stratégia

Túto stratégiu rozdelíme na dve skupiny. Prvou je "stratégia prvej zrady"FBF_∞ a druhou je jej n- ťahová aproximácia FBF_n. FBF_∞ je 2- ťahová stratégia (1 0 1 [1 0]), ktorá sa správa rovnako ako stratégia TFT. 2- ťahová aproximácia $FBF_2(g)$ je vlastne štandardná FBF stratégia (1 0 1 g). Troj a viac ťahová aproximácia tejto stratégie je definovaná nasledovne:

$$FBF_3(g) = FBF_2(g = 1) \oplus FBF_2(g) = (1\;0\;1\;1\;, 1\;0\;1\;g)$$

$$FBF_4(g) = FBF_3(g = 0) \oplus FBF_3(g) = (1\;0\;1\;1\;, 1\;0\;1\;0\;, 1\;0\;1\;1\;, 1\;0\;1\;g)$$

$$FBF_n(g) = \begin{cases} FBF_{n-1}(g = 1) \oplus FBF_{n-1} \text{ ak n je nepárne} \\ FBF_{n-1}(g = 0) \oplus FBF_{n-1} \text{ ak n je párne} \end{cases}$$

Tieto stratégie spolupracujú s protihráčom, ktorý taktiež spolupracuje a zrádzajú po tom, čo protihráč bezdôvodne zradil. Ak tieto stratégie prvé bezdôvodne zradili, v prípade šnúry viacerých zrád, táto skupina stratégií začne spolupracovať napr. $(cDd, cDdDd, ...)$. V prípade, že protihráč prvý bezdôvodne zradil, budú

zrádzať $(Cd, CdDd, ...)$. V prípade, že pamäť stratégií bude prekročená (zrada prišla pred viac ako n- ťahmi a stratégie nevedia určiť, kto zradil prvý), budú spolupracovať s pravdepodobnosťou g [29, 30].

2.4 Firm Pavlov Stratégia

Táto stratégia je istý variant Firm But Fair stratégie, ktorá vie tiež dosiahnuť bezpodmienečnú spoluprácu so svojim protihráčom. Túto stratégiu rozdelíme rovnako na dve skupiny FP_∞ a FP_n. FP_∞ je troj- ťahovou stratégiou "prvej zrady"(1 0 0 1, 1 0 1 [1 0]). 3- ťahová aproximacia $FP_3(g)$ je stratégia (1 0 0 1, 1 0 1 g). 4 a viac ťahová aproximácia tejto stratégie je definovaná nasledovne:

$$FP_4(g) = FP_3(g = 0) \oplus FP_3(g) = (1\ 0\ 0\ 1\ , 1\ 0\ 1\ 0\ , 1\ 0\ 0\ 1\ , 1\ 0\ 1\ g)$$

$$FP_n(g) = \begin{cases} FP_{n-1}(g = 1) \oplus FP_{n-1} \text{ ak n je nepárne} \\ FP_{n-1}(g = 0) \oplus FP_{n-1} \text{ ak n je párne} \end{cases}$$

Stratégia je podobná stratégií FBF_n, ale zrádza po úspešnom využití protihráča (cDc). Naopak ale nezrádza po 3- ťahovej histórií dDc, vďaka čomu sa vie dostať z cyklu vzájomného zrádzania v hre proti svojmu klonu[26, 19].

V ďalšej kapitole sa pokúsime nájsť variácie stratégií FBF a FP, ktoré budú spĺňať podmienky optimality pre hry s minimálnym, konečným a ľubovoľným informačným šumom.

3 Optimalita stratégii FBF a FP

3.1 Optimalita v hre bez informačného šumu

V tejto hre stratégia $FBF_n(g)$ dosahuje hodnoty $\sigma_S = 1$ a $\sigma_C = 0$, čo znamená, že bude neustále spolupracovať v hre proti svojmu klonu a rovnako aj proti stratégií $ALLC$. Hodnota σ_D závisí od oboch parametrov n aj g. $FBF_n(g)$ stratégia proti stratégii $ALLD$ bude spolupracovať s pravdepodobnosťou g, ak to spôsobí pri najmenšom $n/2$ ťahovú šnúru zrádzania. To znamená, že bude spolupracovať každý $n/2 + 1/g$ ťah. Z toho dostávame:

$$1 - \sigma_D = \frac{1}{\frac{n}{2} + \frac{1}{g}} \longrightarrow \sigma_D = \frac{(\frac{n}{2} - 1)g + 1}{\frac{n}{2}g + 1} \tag{33}$$

Vidíme že, stratégia $FBF_n(g)$ dosahuje hodnotu $\sigma_D = 1$ pre $g = 0$ a rovnako tak aj pre stratégiu s nekonečnou pamäťou FBF_∞. Firm Pavlova stratégia $FP_n(g)$ s počiatočnou spoluprácou má rovnaké σ hodnoty ako stratégia $FBF_n(g)$. Naopak táto stratégia s počiatočným zradením, vie naplno zneužiť stratégiu $ALLC$, vďaka čomu nadobúda hodnotu $\sigma_C = 1$. To znamená, že pre všetky n, $FP_n(0)$ s počiatočným zradením nadobúda hodnoty $\sigma_S = \sigma_D = \sigma_C = 1$ a je silno optimálna v tejto hre. Rovnako aj stratégia FP_∞ s počiatočným zradením nadobúda rovnaké hodnoty a je tiež silno optimálna v hre bez informačného šumu [15, 20, 18].

3.2 Optimalita v hre s minimálnym informačným šumom

V hre s minimálnym informačným šumom ostanú hodnoty σ_D pre obidve stratégie rovnaké ako v minulej kapitole. Stratégie FBF majú dokonca rovnakú aj hodnotu $\sigma_C = 0$, ale stratégie FP budú nadobúdať hodnoty $\sigma_C = 1/3$. Pri vzájomnej spolupráci stratégii FP a $ALLC$ iba chyba zo strany stratégie FP môže začať "vykorisťovanie"stratégie $ALLC$. Ale toto "vykorisťovanie"môže byť zastavané chybou či už zo strany stratégie FP alebo $ALLC$. Stratégia FP má dva krát väčšiu tendenciu opustiť šnúru, v ktorej "vykorisťuje"stratégiu $ALLC$, ako do nej vstúpiť. Z toho vyplýva, že "vykorisťovaním"$ALLC$ strávi 1/3 času.

Pozrieme sa ešte na vlastnosť $self-cooperating$ v tejto hre, čím zistíme koľko času strávia stratégie vo vzájomnej spolupráci v hre proti svojim klonom. Najprv sa pozrieme na stratégiu $FP_3(0)$ (1 0 0 1, 1 0 1 0) a ukážeme, že $\sigma_S = 1$. Ako môže byť táto stratégia úplné $self-cooperating$ ak jej "štedrosť"(pravdepodobnosť spolupráce po histórií dDd) je rovná 0? Vieme, že stratégia $FP_3(0)$ sa môže ocitnúť v cykle vzájomného zrádzania so svojim klonom, ale sú na to potrebné dve chyby, zatiaľ čo stačí len jedna chyba, aby naopak zase nadviazali vzájomnú spoluprácu. Potom do takého cyklu vstúpia každý $1/\epsilon^2$ ťah a opustia ho každý $1/\epsilon$ ťah. V nekonečne dlhej hre potom môžeme očakávať, že stratégia $FP_3(0)$ bude v cykle vzájomného zamietania s pravdepodobnosťou:

$$\frac{1/\epsilon}{1/\epsilon^2 + 1/\epsilon} = \frac{\epsilon}{1+\epsilon} \approx 0. \tag{34}$$

Z toho vyplýva, že stratégia $FP_3(0)$ nadobúda hodnoty $\sigma_D = \sigma_S = 1$ a $\sigma_C = 1/3$ a táto stratégia spĺňa podmienky silnej optimality v hre s minimálnym informačným šumom.

Teraz spravíme rovnaký výpočet s tým, že sa pozrieme koľko času strávia stratégie vzájomným zrádzaním v hre proti svojim klonom. Vo všeobecnosti $n/2$ chýb vedie k cyklu vzájomného zrádzania a stratégie tento cyklus opustia s pravdepodobnosťou g ($g \geq \epsilon$). Čas strávený v tomto cykle je vypočítaný nasledovne:

$$\frac{\epsilon^{(n/2)}}{g + \epsilon^{(n/2)}} \tag{35}$$

Pre $\epsilon \approx 0$ je tento čas zanedbateľný, pokiaľ $n \leq 2$ a g je $\Omega(\epsilon)$. Teda stratégie FBF a FP s 3 a viac ťahovou históriou nadobúdajú hodnotu $\sigma_S = 1$. Rovnako aj stratégie FBF_∞ a FP_∞. Pre 2- ťahovú stratégiu FBF platí $\sigma_S = 1/2$ pre $g = \epsilon$ a $\sigma_S = 1$ pre $g \gg \epsilon$.

Stratégie, ktoré sú silno optimálne v APD hre s minimálnym informačným šumom musia spĺňať $\sigma_D = \sigma_S = 1$ a $\sigma_C > 0$. Našli sme teda stratégie $FP_n(0)$ pre všetky n a stratégie FP_∞, ktoré tieto kritéria v APD hre s minimálnym informačným šumom spĺňajú[15, 28].

3.3 Optimalita v hre s konečným informačným šumom

V tejto hre si zadefinujeme koeficient odolnosti voči informačnému šumu NRC_S pre rozličné hodnoty n a g stratégií FBF a FP. Pripomeňme si, že tento koeficient pre stratégiu X má tvar $w(X|X) \geq R-(NRC_S)\epsilon$ pre všetky $\epsilon \leq 0.2$. Z tabuľky č.4 môžeme vidieť, že všetky stratégie s $n \geq 5$ alebo $g \gg \epsilon$ sú odolné voči informačnému šumu pre T=4. Pre T=5, FP stratégie s $n \geq 7$ alebo $g \gg \epsilon$ sú odolné voči informačnému šumu. Pre T=4, stratégia $FP_5(0)$ spĺňa všetky podmienky optimality v hre s konečným informačným šumom. Úspešne sme tak našli stratégiu s konečnou pamäťou, ktorá spĺňa podmienky optimality v tejto hre. Rovnako aj všetky FP_n stratégie spĺňajú podmienky optimality pre $n \geq 5$.

Pre rôzne stratégie FP a FBP vypočítame maximálny stupeň informačného šumu $\epsilon = e$ tak, aby tieto stratégie boli $self-cooperating$ ($\sigma_S \geq 0.9$) a $D-unexploitable$ ($\sigma_D \geq 0.9$). Z tabuľky č.4 vidíme, že s rastúcou "štedrosťou" g stratégie vedia dosiahnuť spoluprácu samé so sebou aj pri vyššom stupni informačného šumu, ale veľmi rýchlo klesá ich schopnosť $D-unexploitability$. S rastúcim n vidíme, že rastie aj schopnosť $self-cooperating$, ale tento parameter len máličko ovplyvňuje schopnosť $D-unexploitability$. Môžeme vidieť, že stratégie $FP_5(0)$ $FBF_5(0)$ sú aj $self-cooperating$ aj $D-unexploitable$ pre úroveň informačného šumu približne do 0.09.

	NRC_S		e pre $\sigma_S = 0.9$		e pre $\sigma_D = 0.9$	
Stratégia	$T = 4$	$T = 5$	$T = 4$	$T = 5$	$T = 4$	$T = 5$
$FBF_2(0)$	∞	∞	0	0	0.100	0.0697
$FBF_2(0.5)$	6	5	0.0605	0.0585	0	0
$FBF_2(1)$	4	3	0.125	0.117	0	0
$FBF_3(0)$	6	5	0.0605	0.0585	0.0922	0.0668
$FBF_3(0.5)$	4	3	0.112	0.106	0	0
$FBF_3(1)$	4	3	0.125	0.117	0	0
$FP_3(0)$	6	5	0.0556	0.0547	0.0925	0.0669
$FP_3(0.5)$	4	3	0.0949	0.0942	0	0
$FP_3(1)$	4	3	0.102	0.101	0	0
$FBF_5(0)$	4	3	0.104	0.0988	0.0917	0.0666
$FBF_5(0.5)$	4	3	0.111	0.105	0	0
$FBF_5(1)$	4	3	0.112	0.106	0	0
$FP_5(0)$	4	3.011	0.0898	0.0890	0.0919	0.0667
$FP_5(0.5)$	4	3	0.0943	0.0937	0	0
$FP_5(1)$	4	3	0.0949	0.0942	0	0
FBF_∞	4	3	0.111	0.105	0.0917	0.0666
FP_∞	4	3	0.0943	0.0936	0.0918	0.0667

Tabuľka 4: Hodnoty stratégii v APD hre s konečným informačným šumom

Vypočítame minimálnu hodnotu g pre každú stratégiu s konečnou pamäťou tak, aby bola odolná voči konečnému informačnému šumu (tabuľka č.5). Zase sa ukázali stratégie $FBF_5(g)$ a $FP_5(g)$ ako odolné voči informačnému šumu pre všetky hodnoty g pre tabuľku výplat s T=4.

Stratégia	$g(T=4)$	$g(T=5)$
FBF_2	1	1
FBF_3	0.20	0.25
FP_3	0.33	0.40
FBF_5	0	0
FP_5	0	0.047

Tabuľka 5: Minimálna hodnota g pre odolnosť voči konečnému informačnému šumu

Nakoniec vypočítame akú maximálnu "štedrosť" môžu mať stratégie na to, aby vedeli odolať "vykorisťovaniu"zo strany $ALLD$. Z tabuľky č.6 vidíme, že všetky stratégie vedia zabrániť „vykorisťovaniu“ pre $\epsilon \leq 1$ pre T=4. Pre T=5 a $\epsilon \geq 0.1$ niektoré veľmi štedré stratégie nie sú voči 'vykorisťovaniu' imúnne. Taktiež vidíme, že FBF_∞ a FP_∞ sú imúnne voči ALLD. Pri predpoklade informačného šumu do výšky 20%, stratégie FBF_n sú imúnne voči ALDD pre všetky $n \geq 5$ a stratégie FP_n sú imúnne voči ALDD pre všetky $n \geq 7$.

	T=4				T=5			
Stratégia	$\epsilon = 0.01$	$\epsilon = 0.05$	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.01$	$\epsilon = 0.05$	$\epsilon = 0.1$	$\epsilon = 0.2$
FBF_2	1	1	1	0.814	0.968	0.846	0.708	0.456
FBF_3	1	1	1	0.811	0.968	0.855	0.734	0.530
FP_3	1	1	1	0.744	0.968	0.850	0.720	0.502
FBF_5	1	1	1	1	1	1	1	1
FBF_5	1	1	1	1	1	1	1	0.938

Tabuľka 6: Maximálna hodnota g pre odolnosť voči vykorisťovaniu od $ALLD$

3.4 Koeficient odolnosti voči informačnému šumu NRC

V tejto časti sa pozrieme na koeficient odolnosti voči informačnému šumu NRC_S podrobnejšie s tým, že sa budeme sústrediť na 3 a 5 ťahové stratégie FP a FBF pri tabuľke výplat (5 3 1 0). Prečo sa tento koeficient pre 3- ťahové stratégie pohybuje v rozmedzí od 3 do 5 a prečo pre $FP_5(0)$ je tento koeficient o máličko väčší ako 3? Najprv si vypočítame očakávanú hodnotu výplat $w(X|X)$ pre každú stratégiu ako funkciu jej "štedrosti" g a konečného informačného šumu na úrovni $\epsilon = e$:

$$w(FP_3|FP_3) = \frac{5ge^2 + 3g + e^2 - e^3}{ge^2 + g + ge + e^2 - e^3}$$

$$w(FBF_3|FBF_3) = \frac{-ge^2 + 3g + 3ge + e^2 - e^3}{-ge^2 + g + 2ge + e^2 - e^3}$$

$$w(FP_5|FP_5) = \frac{7ge^2 + 3g - 3ge^3 + ge^4 + e^3 - 2e^4 + e^5}{3ge^2 + g + ge - 3ge^3 + ge^4 + e^3 - 2e^4 + e^5}$$

$$w(FBF_5|FBF_5) = \frac{ge^2 + 3g + 3ge - 3ge^3 + ge^4 + e^3 - 2e^4 + e^5}{ge^2 + g + 2ge - 3ge^3 + ge^4 + e^3 - 2e^4 + e^5}$$

Najprv budeme uvažovať prípad $g \gg e$ a každú z týchto funkcií vyjadríme pomocou Taylorovho polynómu:

$$w(FP_3|FP_3) = 3 - 3e + \left(5 - \frac{2}{g}\right)e^2 + \left(\frac{7}{g} - 2\right)e^3 + O(e^4)$$

$$w(FBF_3|FBF_3) = 3 - 3e + \left(8 - \frac{2}{g}\right)e^2 + \left(\frac{9}{g} - 19\right)e^3 + O(e^4)$$

$$w(FP_5|FP_5) = 3 - 3e + e^2 + \left(14 - \frac{2}{g}\right)e^3 + O(e^4)$$

$$w(FBF_5|FBF_5) = 3 - 3e + 4e^2 + \left(1 - \frac{2}{g}\right)e^3 + O(e^4)$$

V každom z týchto prípadov linearizáciou $w(X|X)$ v okolí bodu $e = 0$ dostávame $3 - 3e$ a z toho vyplýva:

$$NRC_S = \sup_{(e=0\ldots\epsilon_{max})} \frac{R - w(X|X)_{\epsilon=e}}{e} \approx \lim_{\epsilon\to 0} \frac{R - w(X|X)}{\epsilon} = \frac{3 - (3 - 3\epsilon)}{\epsilon} = 3$$

Pre FBF_5 a FP_5, $w(X|X) > 3 - 3e$, čiže tieto stratégie majú $NRC_S = 3$ pre $g \gg \epsilon$. Pre FP_3, $w(X|X) \approx 3 - 3e + (5 - 2/g)e^2$ z čoho vyplýva, že $NRC_S = 3$ ak $(5 - 2/g) > 0$ respektíve $g > 0.4$. Pre $g < 0.4$ bude výraz pred kvadratickým členom záporný, čiže NRC_S pre FP_3 bude väčší ako 3. Rovnako to platí aj pre stratégiu

FBF_3, $w(X|X) \approx 3 - 3e + (8 - 2/g)e^2$ z čoho vyplýva, že $NRC_S = 3$ ak $(8 - 2/g) > 0$ respektívne $g > 0.25$.

Vypočítame koeficient odolnosti voči informačnému šumu pre stratégie FBF_3 a FP_3 ako funkciu g za predpokladu, že úroveň informačného šumu sa pohybuje od 0 po g (tabuľka č.7).

g	$NRC_S(FP_3)$	$NRC_S(FBF_3)$
0.01	4.863	4.816
0.02	4.731	4.645
0.05	4.364	4.192
0.10	3.833	3.594
0.20	3.206	3.043
0.40	3	3

Tabuľka 7: NRC_S pre stratégie $FP_3(g)$ a $FBF_3(g)$

Teraz budeme uvažovať prípad $g = O(\epsilon)$. Položíme $k = g/e$ a nájdeme mocninový rad pre $w(X|X)$ v bode k na okolí bodu e :

$$w(FP_3|FP_3) = 3 - \Big(3 + \frac{2}{k}\Big)e + O(e^2)$$

$$w(FBF_3|FBF_3) = 3 - \Big(3 + \frac{2}{k}\Big)e + O(e^2)$$

$$w(FP_5|FP_5) = 3 - 3e + \Big(1 - \frac{2}{k}\Big)e^2 + O(e^3)$$
$$w(FP_5|FP_5) = 3 - 3e + \Big(4 - \frac{2}{k}\Big)e^2 + O(e^3)$$

Dostávame, že NRC_S pre stratégie FBF_3 a pre FP_3 je približne rovný $3 + 2/k$. Pre $k = 1(g = e)$ dostávame $NRC_S = 3 + 2 = 5$ a pre $k \to \infty$ $NRC_S = 3$. Toto vysvetľuje, prečo sa koeficient odolnosti voči informačnému šumu pre tieto stratégie pohybuje medzi 3 až 5. Pre FP_5 dostávame:

$$\frac{R - w(X|X)}{\epsilon} \approx \frac{3 - (3 - 3e + (1 - 2/k)e^2)}{e} = 3 + \Big(\frac{2}{k} - 1\Big)e$$

Z toho vyplýva, že $NRC_S = 3$ pre FP_5 ak $(2/k - 1) < 0$ respektívne ak $k > 2$. Pre $k < 2$, vieme že NRC_S bude máličko väčšie ako 3. V najhoršom prípade keď je $k = 1$, sme zistili, že supremum nadobúda svoju hodnotu v bode $e \approx 0.23$ a teda $NRC_S = 3.011$ (tabuľka č.4).

Na opačnej strane pre FBF_5 dostávame:

$$\frac{R - w(X|X)}{\epsilon} \approx \frac{3 - (3 - 3e + (4 - 2/k)e^2)}{e} = 3 + \Big(\frac{2}{k} - 4\Big)e$$

Aj v najhoršom prípade, keď $k = 1$, FBF_5 bude mať $NRC_S = 3$.

Môžeme povedať, že sme získali niekoľko dôležitých výsledkov pre 3 a 5 ťahové stratégie. Po prvé hodnota suprema NRC_S sa nemusí nutne nachádzať na kraji intervalu, čiže musíme vyšetriť $(R - w)/\epsilon$ na celkom intervale. Taktiež sme zistili, že hodnota výplat pre každú stratégiu je $w \geq 3 - 3e$, ak ma dostatočne veľkú "štedrosť" g. Pre FP_3 musí platiť $g > 0.4$, pre FBF_3 musí platiť $g > 0.25$ a pre FP_5 musí platiť $g > 2\epsilon$. FBF_5 má $w \geq 3 - 3e$ pre všetky $g \geq \epsilon$.

3.5 Optimalita v hre s ľubovoľným informačným šumom

Najprv vypočítame koeficient „nepriestrelnosti“ voči informačnému šumu NPC_S pre stratégie FP a FBF pre rôzne hodnoty n a g. Pripomeňme si, že NPC_S predstavuje maximálny počet hrania akcie D (zrada) spôsobených jednou chybou. Na výpočet tohto koeficientu budeme uvažovať dva najhoršie možné cykly chýb: náprava jednej chyby dvoma zradami a cyklus vzájomného zrádzania vychádzajúci z $n/2$ zradení. Tento cyklus vedie k priemernému výsledku $n + 1/g - 1$ zradení, z čoho dostávame:

$$NPC_S = max\left(\frac{n + \frac{1}{g} - 1}{\frac{n}{2}}, 2\right)$$

Z toho vyplýva, že stratégie FP a FBF sú "nepriestrelné" voči informačnému šumu ak $1/2 < g < 1$ pre n nepárne alebo ak $g = 1$ pre n párne. Taktiež sme zistili, že $NPC_S = 2$ pre FBF_∞ a FP_∞. Dostávame, že stratégia FP_∞ spĺňa všetky kritéria optimality v hre s ľubovoľným informačným šumom a rovnako tak spĺňa aj všetky kritéria optimality v hre s konečným informačným šumom, preberaných vyššie. Úspešne sme tak našli stratégiu s nekonečnou pamäťou, ktorá je optimálna v APD hre s ľubovoľným informačným šumom[23, 7].

3.6 Výkonnosť stratégií FBF a FP

Ukázali sme, že stratégie FBF a FP s vyššou pamäťou spĺňajú všetky podmienky optimality. Jedná sa o stratégie $FP_n(0)$ v hre s konečným informačným šumom pre $n > 5$ (T=4) a pre $n > 7$ (T=5). Stratégia FP_∞ zase spĺňa všetky podmienky optimality v hre s ľubovoľným informačným šumom. Príbuzné FBF stratégie taktiež spĺňajú skoro všetky podmienky optimality, až na vlastnosť $C-exploiting$. Teraz musíme zistiť, či si naozaj tieto stratégie počínajú dobre v simuláciách APD hier. Nasimulujeme niekoľko round-robin turnajov. Vypočítali sme si priemerné skóre každej stratégie v každom kole proti všetkým ostatným stratégiám v turnaji. Následne na základe celkového skóre sme ich porovnali. Prvé štyri turnaje pozostávali zo 60 stratégií: $Pavlov$, $RAND$, $ALLC$, $ALLD$, 11 FBF_2 stratégií (g od 0 po 1), 10 FBF_3 stratégií (g od 0 po 0.9), 11 FP_3 stratégií (g od 0 po 1), 11 FP_5 stratégií (g od 0 po 1), 11 FBF_5 stratégií (g od 0 po 1), FBF_∞ a FP_∞. Tieto turnaje boli simulované s informačným šumom na úrovni $\epsilon = 0.01$ a $\epsilon = 0.05$ pre tabuľky výplat (4,3,1,0) a (5,3,1,0).

Stratégia	$(\epsilon = 0.01, T = 5)$	$(\epsilon = 0.01, T = 4)$	$(\epsilon = 0.05, T = 5)$	$(\epsilon = 0.05, T = 4)$
$PAVLOV$	2.922 (15)	2.891 (52)	2.766 (54)	2.707 (53)
$RAND$	2.411 (59)	2.214 (59)	2.385 (58)	2.176 (59)
$ALLD$	2.162 (60)	1.870 (60)	2.21 (60)	1.911 (60)
$ALLC$	2.484 (58)	2.622 (58)	2.379 (59)	2.36 (58)
$FBF2(0.0)$	2.885 (57)	2.879 (57)	2.723 (57)	2.661 (57)
$FBF2(0.1)$	2.9 (56)	2.883 (56)	2.749 (56)	2.689 (56)
$FBF2(0.2)$	2.908 (45)	2.893 (51)	2.752 (55)	2.705 (54)
$FBF2(0.3)$	2.902 (54)	2.886 (55)	2.771 (51)	2.701 (55)
$FBF2(0.4)$	2.907 (48)	2.897 (42)	2.779 (50)	2.724 (51)
$FBF2(0.5)$	2.906 (49)	2.900 (35)	2.77 (53)	2.708 (52)
$FBF2(0.6)$	2.909 (43)	2.894 (48)	2.785 (43)	2.727 (49)

$FBF2(0.7)$	2.914 (31)	2.902 (30)	2.789 (38)	2.741 (39)
$FBF2(0.8)$	2.911 (36)	2.910 (6)	2.771 (52)	2.743 (38)
$FBF2(0.9)$	2.91 (39)	2.902 (28)	2.79 (37)	2.75 (29)
$FBF2(1.0)$	2.904 (52)	2.895 (47)	2.782 (48)	2.737 (45)
$FBF3(0.0)$	2.902 (55)	2.89 (53)	2.785 (44)	2.745 (33)
$FBF3(0.1)$	2.918 (20)	2.898 (39)	2.799 (25)	2.736 (46)
$FBF3(0.2)$	2.905 (51)	2.904 (19)	2.781 (49)	2.764 (4)
$FBF3(0.3)$	2.92 (16)	2.896 (45)	2.785 (45)	2.76 (9)
$FBF3(0.4)$	2.913 (33)	2.906 (15)	2.806 (15)	2.741 (40)
$FBF3(0.5)$	2.916 (24)	2.893 (50)	2.793 (32)	2.754 (19)
$FBF3(0.6)$	2.918 (19)	2.898 (41)	2.801 (20)	2.743 (36)
$FBF3(0.7)$	2.907 (47)	2.901 (31)	2.789 (39)	2.756 (18)
$FBF3(0.8)$	2.909 (44)	2.89 (54)	2.796 (29)	2.767 (3)
$FBF3(0.9)$	2.915 (28)	2.902 (29)	2.798 (26)	2.741 (42)
$FBF5(0.0)$	2.915 (29)	2.895 (46)	2.797 (28)	2.727 (50)
$FBF5(0.1)$	2.916 (25)	2.899 (36)	2.793 (34)	2.741 (41)
$FBF5(0.2)$	2.916 (23)	2.903 (27)	2.795 (31)	2.74 (43)
$FBF5(0.3)$	2.926 (13)	2.898 (40)	2.807 (14)	2.759 (11)
$FBF5(0.4)$	2.915 (26)	2.905 (17)	2.804 (17)	2.756 (16)
$FBF5(0.5)$	2.911 (35)	2.899 (37)	2.808 (9)	2.748 (32)
$FBF5(0.6)$	2.91 (42)	2.908 (13)	2.789 (40)	2.751 (25)
$FBF5(0.7)$	2.917 (22)	2.905 (16)	2.792 (36)	2.732 (47)
$FBF5(0.8)$	2.911 (37)	2.901 (33)	2.786 (42)	2.756 (15)
$FBF5(0.9)$	2.91 (40)	2.909 (9)	2.786 (41)	2.763 (6)
$FBF5(1.0)$	2.913 (32)	2.904 (23)	2.807 (13)	2.75 (28)
$FP3(0.0)$	2.93 (8)	2.913 (3)	2.785 (46)	2.728 (48)
$FP3(0.1)$	2.915 (30)	2.909 (8)	2.807 (10)	2.749 (31)
$FP3(0.2)$	2.936 (3)	2.896 (44)	2.814 (3)	2.751 (24)
$FP3(0.3)$	2.933 (5)	2.897 (43)	2.817 (2)	2.751 (26)
$FP3(0.4)$	2.91 (41)	2.904 (24)	2.799 (24)	2.759 (12)

$FP3(0.5)$	2.918 (18)	2.908 (12)	2.793 (35)	2.76 (8)
$FP3(0.6)$	2.912 (34)	2.901 (32)	2.801 (21)	2.762 (7)
$FP3(0.7)$	2.915 (27)	2.903 (25)	2.813 (4)	2.758 (13)
$FP3(0.8)$	2.903 (53)	2.899 (38)	2.803 (19)	2.76 (10)
$FP3(0.9)$	2.941 (2)	2.904 (22)	2.808 (8)	2.757 (14)
$FP3(1.0)$	2.905 (50)	2.910 (5)	2.784 (47)	2.752 (23)
$FP5(0.0)$	2.928 (10)	2.904 (21)	2.807 (12)	2.756 (17)
$FP5(0.1)$	2.933 (4)	2.908 (10)	2.796 (30)	2.74 (44)
$FP5(0.2)$	2.925 (14)	2.900 (34)	2.801 (22)	2.743 (37)
$FP5(0.3)$	2.908 (46)	2.922 (1)	2.811 (6)	2.752 (22)
$FP5(0.4)$	2.91 (38)	2.893 (49)	2.806 (16)	2.751 (27)
$FP5(0.5)$	2.92 (17)	2.913 (2)	2.823 (1)	2.75 (30)
$FP5(0.6)$	2.928 (12)	2.911 (4)	2.811 (7)	2.764 (5)
$FP5(0.7)$	2.932 (6)	2.908 (11)	2.797 (27)	2.744 (35)
$FP5(0.8)$	2.949 (1)	2.905 (18)	2.801 (23)	2.768 (2)
$FP5(0.9)$	2.928 (11)	2.904 (20)	2.793 (33)	2.754 (20)
$FP5(1.0)$	2.929 (9)	2.906 (14)	2.803 (18)	2.771 (1)
FBF_∞	2.918 (21)	2.909 (7)	2.812 (5)	2.753 (21)
FPN_∞	2.931 (7)	2.903 (26)	2.807 (11)	2.744 (34)

Tabuľka 8: Round robin turnaj výsledky

Stratégie FP_5 vyhrali všetky štyri turnaje. Všeobecne môžeme povedať, že FP stratégie porazili FBF stratégie. Vidíme, že pre stratégie s konečnou pamäťou stúpala ich výkonnosť s rastúcim n. Avšak stratégie s nekonečnou pamäťou si v týchto turnajoch nepočínali až tak dobre. Domnievame sa, že toto slabé skóre môže byť zapríčinené hraním proti stratégiám s malou pamäťou a malou štedrosťou. Ako príklad môžeme porovnať stratégiu FBF_∞ so stratégiou Tit for Tat a so stratégiou $FBF_2(0)$. Chyba zo strany FBF_∞ bude rýchlo napravená, ale chyba zo strany stratégie Tit for Tat bude viesť k cyklu vzájomného zrádzania. Na otestovanie tohto predpokladu sme nasimulovali ďalší turnaj, z ktorého sme vynechali stratégie s nulovou "štedrosťou". Turnaj bol nasimulovaný s $\epsilon = 0.01$ pre tabuľku výplat (5,3,1,0).

Stratégia	$(\epsilon = 0.01, T = 5, g \neq 0)$
$PAVLOV$	2.91(20)
$RAND$	2.42(54)
$ALLD$	2.233(55)
$ALLC$	2.697(53)
$FBF2(0.1)$	2.895(51)
$FBF2(0.2)$	2.907(27)
$FBF2(0.3)$	2.901(45)
$FBF2(0.4)$	2.898(49)
$FBF2(0.5)$	2.905(35)
$FBF2(0.6)$	2.909(22)
$FBF2(0.7)$	2.904(38)
$FBF2(0.8)$	2.895(52)
$FBF2(0.9)$	2.904(39)
$FBF2(1.0)$	2.901(46)
$FBF3(0.1)$	2.916(7)
$FBF3(0.2)$	2.911(18)
$FBF3(0.3)$	2.916(8)
$FBF3(0.4)$	2.907(28)
$FBF3(0.5)$	2.907(29)
$FBF3(0.6)$	2.913(14)
$FBF3(0.7)$	2.901(47)
$FBF3(0.8)$	2.907(30)
$FBF3(0.9)$	2.903(41)
$FBF5(0.1)$	2.907(31)
$FBF5(0.2)$	2.903(42)
$FBF5(0.3)$	2.914(12)
$FBF5(0.4)$	2.911(19)

$FBF5(0.5)$	2.897(50)
$FBF5(0.6)$	2.912(15)
$FBF5(0.7)$	2.91(21)
$FBF5(0.8)$	2.903(43)
$FBF5(0.9)$	2.902(44)
$FBF5(1.0)$	2.908(25)
$FP3(0.1)$	2.919(5)
$FP3(0.2)$	2.933(2)
$FP3(0.3)$	2.915(10)
$FP3(0.4)$	2.912(16)
$FP3(0.5)$	2.908(26)
$FP3(0.6)$	2.9(48)
$FP3(0.7)$	2.93(3)
$FP3(0.8)$	2.916(9)
$FP3(0.9)$	2.918(6)
$FP3(1.0)$	2.907(32)
$FP5(0.1)$	2.928(4)
$FP5(0.2)$	2.909(23)
$FP5(0.3)$	2.904(40)
$FP5(0.4)$	2.906(34)
$FP5(0.5)$	2.907(33)
$FP5(0.6)$	2.905(36)
$FP5(0.7)$	2.905(37)
$FP5(0.8)$	2.909(24)
$FP5(0.9)$	2.912(17)
$FP5(1.0)$	2.915(11)
FBF_∞	2.914(13)
FP_∞	2.94(1)

Tabuľka 9: Round robin turnaj výsledky bez stratégii s $g = 0$

Ako sa očakávalo, stratégia FP_∞ sa stala víťaznou. Taktiež sme si mohli všimnúť, že optimálna "štedrosť"stratégie závisí od pamäte n. Ďalej vyberieme 29 stratégií FBF a FP: 4 FBF_2 (g od 0.2 po 0.8), 5 FBF_3 (g od 0 po 0.8), 6 FP_3 (g od 0 po 1), 6 FP_5 (g od 0 po 1), 6 FBF_5 (g od 0 po 1), FBF_∞ a FP_∞. Každá z týchto stratégií bude hrať individuálne round- robin turnaj proti 9 stratégiám: $g1$,$g2$,$g3$,$Pavlov$,$RAND$, $ALLD$, $ALLC$, TFT. Tieto turnaje sme simulovali pre tabuľku výplat (4,3,1,0) s $\epsilon = 0.01$. 28 z týchto 29 stratégií svoj turnaj vyhrali, výnimkou bola stratégia $FBF_3(0)$, ktorá skončila druhá. Najlepší výsledok zaznamenali stratégie $FP_5(0.2)$ a $FP_3(0.2)$, ktoré vyhrali s náskokom 0.0711 respektíve 0.0677. Znova sa ukázalo, že FP stratégie porazili FBF stratégie.

Z výsledkov je jasné, že oboje stratégie boli úspešné v robin round turnajoch a stratégie FP si vedú lepšie ako stratégie FBF. Schopnosť dosiahnuť bezpodmienečnú spoluprácu je kľúčom k úspechu v tomto turnaji. Stratégia $FP_5(0.2)$ bola najlepšia v rámci simulácií, ktoré sme zrealizovali. Je úplne $self - cooperating$, a takmer úplne $D - unexploitable$ a $C - exploiting$. Nenulová "štedrosť"jej dovoľuje spolupracovať aj so stratégiami s malou pamäťou a malou štedrosťou ako je napríklad TFT.

Záver

Cieľom diplomovej práce bolo zadefinovanie a oboznámenia sa s problémom alternatívnej väzňovej dilemy a vyriešenie problému optimálnej stratégie v alternatívnej väzňovej dileme bez informačného šumu, s minimálnym informačným šumom, s konečným informačným šumom a nakoniec s ľubovoľným informačným šumom.

Kedy je v takejto hre stratégia optimálna? Napriek tomu že sme ukázali, že žiadna zo stratégii nie je schopná dosiahnuť maximálne možné skóre, zadefinovali sme kritéria optimality, ktoré úzko korelujú s maximálnou priemernou hodnotou výplat. Na to, aby stratégia bola optimálna musí byť schopná nadviazať bezpodmienečnú spoluprácu s inou stratégiou, musí vedieť dosiahnuť spoluprácu so svojim klonom a musí vedieť odolávať príliš častému zrádzaniu. Tieto 3 kritériá sme zadefinovali a vyšetrili pre všetky vyššie spomínané prípady informačného šumu. Následne sme rozobrali najčastejšie spomínane 2- ťahové stratégie alternatívnej väzňovej dilemy. Tieto stratégie robia rozhodnutie len na základe posledných 2 ťahov a ukázali sme, že žiadna z týchto stratégii nespĺňa všetky podmienky optimality. 2- ťahové stratégie, ktoré sú odolné voči častému zrádzaniu a vedia nadviazať bezpodmienečnú spoluprácu s inou stratégiou, nie sú schopné dosiahnuť potrebný úroveň spolupráce v hre proti svojmu klonu.

V ďalšej časti sme sa venovali stratégiám, ktoré si pamätajú viac ako 2 ťahy dozadu a ukázalo sa, že tieto stratégie sú oveľa efektívnejšie, nakoľko spĺňajú všetky podmienky optimality. Napríklad Firm Pavlov stratégia $FP_3(0)$ (1 0 0 1, 1 0 1 0) spĺňa všetky 3 podmienky optimality v alternatívnej väzňovej dileme s minimálnym informačným šumom. Firm Pavlov a Firm But Fair stratégie s meniacou sa „štedrosťou“ a vyšším n spĺňajú ešte prísnejšie kritéria. Firm Pavlov stratégie s 5 a viac ťahovou pamäťou dosahujú v APD hre vysoké skóre pri ľubovoľnom konečnom informačnom šume a preto môžeme povedať, že sú voči nemu odolné. Firm Pavlov stratégia s nekonečnou pamäťou navyše dokáže rýchlo napraviť vzájomnú spoluprácu po sérii chýb v hre proti svojmu klonu. Vďaka spojeniu tejto vlastnosti s 3 kritériami optimality je FP_∞ stratégia optimálna v APD hre s ľubovoľným informačným šumom.

V poslednej časti sme testovali či stratégie, ktoré spĺňajú všetky podmienky optimality sú naozaj úspešné v hre proti ostatným stratégiám. Silu týchto stratégií sme otestovali rovnakým spôsobom ako Robert Axelrod. Nasimulovali sme round robin turnaje s rozličným informačným šumom a rozličnou tabuľkou výplat pre 60 stratégií. Každá dvojica stratégií sa stretla presne 200-krát. Víťaznými sa stali Firm Pavlov a Firm But Fair triedy stratégií, ktoré dominovali vo všetkých turnajoch, čím sme potvrdili výsledky z teoretickej časti. Zo simulácii sa ukázalo, že Firm Pavlov stratégie si viedli lepšie ako Firm But Fair stratégie. Najoptimálnejšia stratégia s konečnou pamäťou medzi všetkými testovanými sa ukázala $FP_5(0.2)$ a s nekonečnou pamäťou stratégia FP_∞.

Zoznam použitej literatúry

[1] John von Neumann, Oskar Morgenstern : *Theory of Games and Economic Behavior*, Princeton University Press, 1944

[2] Martin J. Osborne : *An Introduction to Game Theory*, 2003

[3] Jianzhong Wu, Robert Axelrod :*How to Cope with noise in the Iterated Prisoner's Dilema*, Journal of Conflict Resolution 39(1), March 1995

[4] The Alternating Prisoner's Dilema, dostupné na internete (2.01.2018): http://homepage.univie.ac.at/Karl.Sigmund/JTB94a.pdf

[5] Extending the Iterated Prisoner's Dilemma without Synchrony, dostupné na internete (2.01.2018): $https://www.math.ubc.ca/ hauert/publications/reprints/hauert_jtb98.pdf$

[6] Prisoner's Dilemma, dostupné na internete (24.04.2018) https://plato.stanford.edu/entries/prisoner-dilemma/

[7] Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent, , dostupné na internete (24.04.2018) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3387070/

[8] Iterated Prisoner's Dilemma, dostupné na internete (24.04.2018) https://www.psychestudy.com/behavioral/learning-memory/iterated-prisoners-dilemma

[9] Marcus Frean: *The Prisoner's Dilemma without Synchrony*, DOI: 10.1098/rspb.1994.0096

[10] Strategies for IPD, dostupné na internete (24.04.2018) http://prisoners-dilemma.com/

[11] Iterated Prisoner's Dilemma with Choice and Refusal of Partners, dostupné na internete (24.04.2018) https://lib.dr.iastate.edu/cgi/viewcontent.cgi?referer=https://scholar.google.sk/&httpsredir=1&article=1028&context=econ_las_economicreports

[12] The Iterated Prisoner's Dilemma: Good Strategies and Their Dynamics, dostupné na internete (24.04.2018)
https://arxiv.org/pdf/1211.0969/

[13] J. Maynard Smith : *Optimization Theory in Evolution*, 1978

[14] J. Samuel Karlin, Howard E. Taylor : *A first course in stochastic processes*, ISBN: 978-0-08-057041-9

[15] Mohammed Abdellaoui : *Theory and Decision*. ISSN 0040-5833, Volume 1 / 1971 - Volume 84 / 2018

[16] Rafik A Aliev : *Decision Theory with Imperfect Information*, 2014

[17] On some winning strategies for the Iterated Prisoner's Dilemma, dostupné na internete (24.04.2018)
https://arxiv.org/ftp/cs/papers/0609/0609017.pdf

[18] The Iterated Prisoner's Dilemma on a Cycle, dostupné na internete (24.04.2018) //https://arxiv.org/pdf/1102.3822.pdf

[19] C. Wedekind and M. Milinski : *Human cooperation in the simultaneous and the alternating prisoner's dilemma: Pavlov versus generous tit-for-tat*, Proceedings of the National Academy of Sciences of the United States of America, 1996

[20] Daniel W. Dyer : *Opponent Modelling and Strategy Evolution in the Iterated Prisoner's Dilemma*, dostupné na internete (24.04.2018)
https://pdfs.semanticscholar.org/833d/64359e90abb5f0b38cd5191a25b52c89d3b7.pdf

[21] Björn Brembs :*Chaos, Cheating and Cooperation: Potential Solutions to the Prisoner's Dilemma*, Vol. 76, No. 1 (May, 1996), pp. 14-24

[22] The Continuous Prisoner's Dilemma and the Evolution of Cooperation through Reciprocal Altruism with Variable Investment, dostupné na internete (24.04.2018)
http://www.math.ubc.ca/ doebeli/reprints/Doe44.pdf

[23] Martin Nowak, Karl Sigmund : *The evolution of stochastic strategies in the Prisoner's Dilemma* , September 1990, Volume 20, Issue 3, pp 247–265

[24] Imhof LA, Fudenberg D, Nowak MA :*Tit-for-tat or win-stay, lose-shift?*, J Theor Biol. 2007 Aug 7;247(3):574-80. Epub 2007 Mar 24.

[25] Hisashi Ohtsuki, Martin A. Nowak :*Evolutionary games on cycles* , Proceedings of the Royal Society B: Biological Sciences 273(1598):2249-56,October 2006

[26] Fingerprint Analysis of the Noisy Prisoner's Dilemma Using a Finite State Representation, dostupné na internete (24.04.2018) http://eldar.mathstat.uoguelph.ca/dashlock/eprints/FingerprintsIII.pdf

[27] Stephens D.W., Nishimura K., Toyer K.B. :*Error and Discounting in the Iterated Prisoner's Dilemma*, Journal of Theoretical Biology,Volume 176, Issue 4, 21 October 1995, Pages 457-469

[28] Jonathan Bendor, Roderick M. Kramer, Suzanne Stout :*When in Doubt...Cooperation in a Noisy Prisoner's Dilemma*, Journal of Conflict Resolution 35:691-719

[29] Anatolii F.Kleimenov, Alexei A.Semenishchev :*Repeated Prisoner's Dilemma: Stackelberg Solution with Finite Memory*, IFAC Proceedings Volumes, Volume 33, Issue 16, July 2000, Pages 567-572

[30] K-Memory Strategies in Repeated Games, dostupné na internete (24.04.2018) https://pdfs.semanticscholar.org/0693/ac626ab01459b0fa364dbaca17d7f982906c.pdf

Príloha A

A. Zdrojové kódy v jazyku PHP

1, Kódy pre simuláciu round robin turnaja

```
<?php

set_time_limit(60*5);
error_reporting(E_ALL ^E_DEPRECATED ^E_NOTICE ^E_WARNING);
if (function_exists("date_default_timezone_set"))
date_default_timezone_set("Europe/Bratislava");

// Inicializacia nahodneho generatora
srand((double) microtime() * 1000000);

$show = 0;

// ================================
// Vztah väzòov C-cooperuju
// Obaja kooperuju spoloène CC - skrati sa kazdemu trest o 3 roky
// Obaja nespolupracujú  (defect) DD - skrati sa im obom trest o 1 rok
// Ak spolupracuje s policiou iba jeden èize zradza kolegu (D-Deffect)
//tak sa mu skati trest o 5 rokov a druhemu o 0
//
$param_hra = array('CC'=>3, 'DD'=>1, 'CD'=>0, 'DC'=>5);
// Vzdy prve pismenko je prvy hrac
$epsilon = 0.01;
// Informaèny sum 0.05 znamena ze s 5% pravdepodobnostou nastane chyba v tahu
$pocet_kol = 200;
$nekonecno = 10;
$nekonecno_g = 0.2;

$zoznam_strategii = array();
```

```
/*
$zoznam_strategii[] = "ALLD";
$zoznam_strategii[] = "FBF3(0.5)";
$zoznam_strategii[] = "FPN()";
*/

/*
$zoznam_strategii[] = "ALLD";
$zoznam_strategii[] = "ALLC";
$zoznam_strategii[] = "RAND";
*/


$zoznam_strategii[] = "PAVLOV";
$zoznam_strategii[] = "RAND";
$zoznam_strategii[] = "ALLD";
$zoznam_strategii[] = "ALLC";
//$zoznam_strategii[] = "FBF2(0.0)";
$zoznam_strategii[] = "FBF2(0.1)";
$zoznam_strategii[] = "FBF2(0.2)";
$zoznam_strategii[] = "FBF2(0.3)";
$zoznam_strategii[] = "FBF2(0.4)";
$zoznam_strategii[] = "FBF2(0.5)";
$zoznam_strategii[] = "FBF2(0.6)";
$zoznam_strategii[] = "FBF2(0.7)";
$zoznam_strategii[] = "FBF2(0.8)";
$zoznam_strategii[] = "FBF2(0.9)";
$zoznam_strategii[] = "FBF2(1.0)";
//$zoznam_strategii[] = "FBF3(0.0)";
$zoznam_strategii[] = "FBF3(0.1)";
```

```
$zoznam_strategii[] = "FBF3(0.2)";
$zoznam_strategii[] = "FBF3(0.3)";
$zoznam_strategii[] = "FBF3(0.4)";
$zoznam_strategii[] = "FBF3(0.5)";
$zoznam_strategii[] = "FBF3(0.6)";
$zoznam_strategii[] = "FBF3(0.7)";
$zoznam_strategii[] = "FBF3(0.8)";
$zoznam_strategii[] = "FBF3(0.9)";
//$zoznam_strategii[] = "FBF5(0.0)";
$zoznam_strategii[] = "FBF5(0.1)";
$zoznam_strategii[] = "FBF5(0.2)";
$zoznam_strategii[] = "FBF5(0.3)";
$zoznam_strategii[] = "FBF5(0.4)";
$zoznam_strategii[] = "FBF5(0.5)";
$zoznam_strategii[] = "FBF5(0.6)";
$zoznam_strategii[] = "FBF5(0.7)";
$zoznam_strategii[] = "FBF5(0.8)";
$zoznam_strategii[] = "FBF5(0.9)";
$zoznam_strategii[] = "FBF5(1.0)";
//$zoznam_strategii[] = "FP3(0.0)";
$zoznam_strategii[] = "FP3(0.1)";
$zoznam_strategii[] = "FP3(0.2)";
$zoznam_strategii[] = "FP3(0.3)";
$zoznam_strategii[] = "FP3(0.4)";
$zoznam_strategii[] = "FP3(0.5)";
$zoznam_strategii[] = "FP3(0.6)";
$zoznam_strategii[] = "FP3(0.7)";
$zoznam_strategii[] = "FP3(0.8)";
$zoznam_strategii[] = "FP3(0.9)";
$zoznam_strategii[] = "FP3(1.0)";
//$zoznam_strategii[] = "FP5(0.0)";
```

```
$zoznam_strategii[] = "FP5(0.1)";
$zoznam_strategii[] = "FP5(0.2)";
$zoznam_strategii[] = "FP5(0.3)";
$zoznam_strategii[] = "FP5(0.4)";
$zoznam_strategii[] = "FP5(0.5)";
$zoznam_strategii[] = "FP5(0.6)";
$zoznam_strategii[] = "FP5(0.7)";
$zoznam_strategii[] = "FP5(0.8)";
$zoznam_strategii[] = "FP5(0.9)";
$zoznam_strategii[] = "FP5(1.0)";
$zoznam_strategii[] = "FBFN()";
$zoznam_strategii[] = "FPN()";
// $zoznam_strategii[] = "TFT";        // 60
// $zoznam_strategii[] = "FBF3(1.0)";  // 61


if ($GLOBALS['show'] >= 1) echor($zoznam_strategii);

// ---- Tu si budem napocitavat po kazdej hre vysledne roky
$strategia_rokov = array();
// Kluc v array èislo 0-60 korespondujúce so zoznamom strategii
//a hodnota bude sumar z vysledkov hier
// Vynulujem vysledne roky
for ($i=0; $i<count($zoznam_strategii); $i++) $strategia_rokov[$i] = 0;

$pocet_strategii = count($zoznam_strategii);

$pocet = 0;
for ($s1 = 0; $s1 <= $pocet_strategii-1; $s1++) {
for ($s2 = $s1+1; $s2 <= $pocet_strategii-1; $s2++) {
$pocet++;
```

```
// echo("<br>$pocet   $s1,$s2  
//".$zoznam_strategii[$s1]."     ".$zoznam_strategii[$s2]."   ");
zahraj_hru($s1, $s2);
}
}


if ($GLOBALS['show'] >= 1) echo("<br>Pocet zahranych = $pocet  
 Count strategii: ".count($zoznam_strategii)."<br>");
// Hry su odohrane. V poli $strategia_rokov mam nasumovane vysledky
//pre kazdu strategiu. Predelim poctom strategii-1

for ($i=0; $i<count($zoznam_strategii); $i++)
$strategia_rokov[$i] = $strategia_rokov[$i] / (count($zoznam_strategii)-1);

$pole_vysledok = array();
foreach($zoznam_strategii as $key=>$nazov)
  $pole_vysledok[$nazov] = $strategia_rokov[$key];

$pole_vysledok_sort = $pole_vysledok;
asort($pole_vysledok_sort);

echo("<TABLE BORDER=0><TR><TD>Nezotriedene");

echo("<table border=1 cellspacing=0 cellpadding=5>");
foreach($pole_vysledok as $key=>$val)
  echo("<tr><td>$key</td><td>$val</td></tr>");
echo("</table>");

echo("</TD><TD>      </TD><TD>Zotriedene vysledky");
```

```
echo("<table border=1 cellspacing=0 cellpadding=5>");
foreach($pole_vysledok_sort as $key=>$val)
  echo("<tr><td>$key</td><td>$val</td></tr>");
echo("</table>");

echo("</TD></TR></TABLE>");
die("<br>Koniec");



// ----------------------------------------------------------
// Vstupujuce strategie su idèka do $zoznam_strategii
// Funkcia prevedie 100 kôl (kazdy hraè potiahne 100x)
//
function zahraj_hru($strategia1, $strategia2) {
  global $zoznam_tahov;
 // Globalny zoznam jednotlivych tahov hracov.
//Na zaciatku si ho vynulujem (resp. naplnim poèiatoènym C resp. D pre ALLD)
global $sum_rokov1;
 // Sumacia odmeny hraca1 a hraca2
global $sum_rokov2;
global $strategia_rokov;
// Globalny celkovy zoznam rokov pre kazdú hranú strategiu.
//Tato premenna sa nuluje iba na zaèiatku programu
global $zoznam_strategii;
// Pole s textovymi nazvami strategii (definovane na zaèiatku programu)

// echo("<br>Hra: $strategia1 (".$zoznam_strategii[$strategia1].")
//  vs.   $strategia2 (".$zoznam_strategii[$strategia2].") ");

  // Zoznam tahov zaèina prvym hraèom a predvyplnenym C
  // (resp. pre strategiu ALLD hodnotou D)
```

```
  // Z toho vyplyva, ze prvy hraè uz potiahol a je na tahu druhy hraè !!!

if ($zoznam_strategii[$strategia1] == 'ALLD') $zoznam_tahov = "D";
  else                                          $zoznam_tahov = "C";

  $sum_rokov1 = 0;                    // Reset odmien
  $sum_rokov2 = 0;

priprav_strategiu(1, $strategia1);
// Pre FP a FBF strategie si predpoèitam kombinacie
//tahov a odpovedi oboch hraèov
// do globalnych poli ($p_strategia1, $p_strategia2)
priprav_strategiu(2, $strategia2);

// Inicializacia nahodneho generatora
srand((double) microtime() * 1000000);

$kto_je_na_tahu = 2;
   // Kedze hrac 1 potiahol inicializaciou zoznam_tahov tak je na rade hrac 2
for ($kolo = 1; $kolo <= $GLOBALS['pocet_kol']; $kolo++) {
if ($kto_je_na_tahu == 1) {
$kto_je_na_tahu = 2;
 // Posunieme tah na dalsieho hraca
  $volba = get_strategia_tah(1, $strategia1);
// Zisti aky tah zahra hraè èislo 1
  // echo("<br>Hrame kolo:$kolo Hrac1 voli tah: $volba ");
    }
else {
$kto_je_na_tahu = 1;
$volba = get_strategia_tah(2, $strategia2);
 // Rovnako druhy hraè
```

```
// echo("<br>Hrame kolo:$kolo Hrac2 voli tah: $volba ");
  }
hra($kolo, $volba);
  // Zahraj tento tah = prilozi $zoznam_tahov + napoèita odmenu obom
//hraèom za tento tah (premenne $sum_rokov1, $sum_rokov2)
    }

// Vypisanie vysledkov tejto hry po 100 kolach
if ($GLOBALS['show'] >= 1) {
  echo("<hr>Hra: $strategia1 (".$zoznam_strategii[$strategia1].")
  vs.   $strategia2 (".$zoznam_strategii[$strategia2].") ");
  echo("<br><br>Vysledok<br>Hrac1 rokov: $sum_rokov1<br>Hrac2 rokov: $sum_rokov2")
  }

// Celkove sumovanie strategii - ich úspesnost
$strategia_rokov[$strategia1] = $strategia_rokov[$strategia1] +
    $sum_rokov1 / $GLOBALS['pocet_kol'];
$strategia_rokov[$strategia2] = $strategia_rokov[$strategia2] +
    $sum_rokov2 / $GLOBALS['pocet_kol'];
// echor($strategia_rokov); echo("<hr>");
  }



// ----------------------------
// tah jedneho z hraèov
//
// Vstupujuce $pc je èislo kola 1 .. 100  (potrebne iba pre vypisovanie)
//             $tah = ako zahral hraè: C-cooperative, D-defect (zrada)
//
// Rutina pripoji tah hraèa do globalneho $zoznam_tahov +
//napoèita odmenu obom hraèom
```

```
//
function hra($pc, $tah) {
global $zoznam_tahov, $param_hra, $sum_rokov1, $sum_rokov2;

$zoznam_tahov = $zoznam_tahov . $tah;

$poc_tahov = strlen($zoznam_tahov);
if ($poc_tahov <= 1) {
echo("<br>Vyskakujem lebo dlzka je 1");
return;
// Pri prvom tahu sa este nerobi vyhodnotenie vyplaty
  }

$predosly = substr($zoznam_tahov,-2, 1);        // Preèitam predosly tah
if (mod2($poc_tahov,2) == 0) {
 // Parne cislo => teraz hral hraè èislo 2  a teda hraè 1 mal predosly tah
  $hrac1 = $predosly;
$hrac2 = $tah;
  }
else {                      // Neparne èislo => teraz hral hraè èislo 1
  $hrac2 = $predosly;
$hrac1 = $tah;
  }
$vysledok1 = $param_hra[$hrac1.$hrac2];
// Ohodnotenie z pohladu prveho hraèa   CC,DD,CD,DC => 3,1,0,5
$vysledok2 = $param_hra[$hrac2.$hrac1];
 // Ohodnotenie druheho hraèa

$sum_rokov1 = $sum_rokov1 + $vysledok1;
// Sumovanie odmeny prveho a druheho hraèa
$sum_rokov2 = $sum_rokov2 + $vysledok2;
```

```
// echo("<br>Kolo:$pc Tah Hrac1/Hrac2:  
$hrac1/$hrac2   =   $vysledok1/$vysledok2
    Historia posl.20 tahov: ".substr($zoznam_tahov,-20)."");
  }

// -------------------------------------
// Rutina vrati tah hraèa $pc = 1 alebo 2 pre vstupujúce $id_strategia
//
function get_strategia_tah($pc, $id_strategia) {
global $zoznam_tahov, $zoznam_strategii;

$nazov = $zoznam_strategii[$id_strategia];

if (substr($nazov,0,3) == 'FBF') {      // FBF2(0.1)
$zoznam = $zoznam_tahov;
// Lokalna kopia zoznamu tahov (kedze ho budem mozno zlava doplnat C znakmi)
  $poc_pamatam = substr($nazov,3,1);
  $pravdepodobnost_c = substr($nazov,5,3);
if ($poc_pamatam == 'N') $poc_pamatam = $GLOBALS['nekonecno'];
while (strlen($zoznam) < $poc_pamatam)
// Zoznam si musim zlava vyplnit C az na dlzku $poc_pamatam
$zoznam = "C".$zoznam;

global $p_strategia1, $p_strategia2;
$tahy = substr($zoznam, -$poc_pamatam);
// Tu mam poslednych x tahov
if ($pc == 1) $vysledok = $p_strategia1[$tahy];
if ($pc == 2) $vysledok = $p_strategia2[$tahy];
// echo("<br>TAH pc=$pc id=$id_strategia vysledok=$vysledok ");
return nahoda($vysledok);
```

```
  }

if (substr($nazov,0,2) == 'FP') {      // FBF2(0.1)
$zoznam = $zoznam_tahov;
  // Lokalna kopia zoznamu tahov (kedze ho budem mozno zlava doplnat C znakmi)
  $poc_pamatam = substr($nazov,2,1);
  $pravdepodobnost_c = substr($nazov,4,3);
if ($poc_pamatam == 'N') $poc_pamatam = $GLOBALS['nekonecno'];
while (strlen($zoznam) < $poc_pamatam)
 // Zoznam si musim zlava vyplnit C az na dlzku $poc_pamatam
$zoznam = "C".$zoznam;

// echo("<br>Zoznam: ".substr($zoznam,-20)." ");

global $p_strategia1, $p_strategia2;
$tahy = substr($zoznam, -$poc_pamatam);
 // Tu mam poslednych x tahov ktore potrebujem
if ($pc == 1) $vysledok = $p_strategia1[$tahy];
if ($pc == 2) $vysledok = $p_strategia2[$tahy];
// echo("<br>TAH pc=$pc id=$id_strategia vysledok=$vysledok ");
return nahoda($vysledok);
  }

  // ------------
if ($nazov == 'ALLD') return nahoda(0.0);     // 'D';
if ($nazov == 'ALLC') return nahoda(1.0);     // 'C';

// ------------
if ($nazov == 'TFT') {
$zoznam = $zoznam_tahov;
$poc_pamatam = 2;
```

```
while (strlen($zoznam) < $poc_pamatam)
 // Zoznam si musim zlava vyplnit C az na dlzku $poc_pamatam
$zoznam = "C".$zoznam;

$last = substr($zoznam, -1, 1);
$prev = substr($zoznam, -2, 1);
if ($last == 'C' and $prev == 'C') return nahoda(1.0);   // 'C';
 // TFT (1,0,1,0)  = (CD,CD,DC,DD)            1=C 0=D
if ($last == 'C' and $prev == 'D') return nahoda(0.0);   // 'D';
if ($last == 'D' and $prev == 'C') return nahoda(1.0);   // 'C';
if ($last == 'D' and $prev == 'D') return nahoda(0.0);   // 'D';
die("Error Last:$last Prev:$prev");
  }

// ------------
if ($nazov == 'RAND') {
$zoznam = $zoznam_tahov;
$poc_pamatam = 2;
while (strlen($zoznam) < $poc_pamatam)
  // Zoznam si musim zlava vyplnit C az na dlzku $poc_pamatam
$zoznam = "C".$zoznam;

$last = substr($zoznam, -1, 1);
$prev = substr($zoznam, -2, 1);
if ($last == 'C' and $prev == 'C') return nahoda(0.5);
// nahoda 1 => C   nahoda 0 => D   nahoda  0.1 èastejsie vrati D
if ($last == 'C' and $prev == 'D') return nahoda(0.5);
if ($last == 'D' and $prev == 'C') return nahoda(0.5);
if ($last == 'D' and $prev == 'D') return nahoda(0.5);
die("Error Last:$last Prev:$prev");
  }
```

```
// ------------
if ($nazov == 'PAVLOV') {
$zoznam = $zoznam_tahov;
$poc_pamatam = 2;
while (strlen($zoznam) < $poc_pamatam)
  // Zoznam si musim zlava vyplnit C az na dlzku $poc_pamatam
$zoznam = "C".$zoznam;

$last = substr($zoznam, -1, 1);
$prev = substr($zoznam, -2, 1);
if ($last == 'C' and $prev == 'C') return nahoda(1.0);   // 'C';
if ($last == 'C' and $prev == 'D') return nahoda(0.0);   // 'D';
if ($last == 'D' and $prev == 'C') return nahoda(0.0);   // 'D';
if ($last == 'D' and $prev == 'D') return nahoda(1.0);   // 'C';
die("Error Last:$last Prev:$prev");
  }

die("CHYBA - neznama naprogramuj - $nazov !!!!");
}

// ------------------------------------------
// Vstup  0 => D
//        0.001 => vysoko pravdepodobne D
//        1 => C
function nahoda($pravdepodobnost_c) {
global $epsilon;

if ($pravdepodobnost_c == 1) $out = 'C';
elseif ($pravdepodobnost_c == 0) $out = 'D';
else {
```

```
$x = rand(1, 10000);                                   // vratane
    if ($x <= $pravdepodobnost_c * 10000) $out = 'C'; else $out = 'D';
  }
// V premennej $out mame spravnu odpoved. Teraz zavedieme epsilon
if ($epsilon == 0) return $out;


$x = rand(1, 10000);                                   // vratane
  if ($x > $epsilon * 10000) return $out;               // sum sa neuplatnil

// Zamenime vysledok
if ($out == 'C') return 'D';
if ($out == 'D') return 'C';
  }


// -----------------------------------------------
// Potrebujem zlozit vsetky kombinacie $n miestneho èisla
// (a mat ich spravne zoradene)
// Vyuzijem na to binarne (dvojkove) èislo,
//ktore budem zväèsovat dovtedy kym nedosiahne
// poèet znakov $n. Na zaver prekonvertujem
// dvojkove èislo do pismen 0->C  1->D
// s ulozenim do pola ktore vratim
//
// 0 = 00000 => CCCCC
// 1 = 00001 => CCCCD
// 2 = 00010 => CCCDC
// 3 = 00011
// 4 = 00100
//
```

```
function zloz_kombinacie($n) {
$pole = array();
$i = 0;
while(1==1) {
  $bin = base_convert($i, 10, 2);
// Skonvertuje èislo $i z desiatkovej do dvojkovej (binarnej) sústavy
if (strlen($bin) > $n) break;
// Ak uz dostavam dlhsie èislo ako potrebujem => odchod
while (strlen($bin) < $n)
 // Dokial je èislo kratsie ako potrebujem tak ho zlava doplnim nulami
$bin = "0".$bin;
$bin = str_replace("0", "C", $bin);
// V retazci predstavujúcom binarne èislo zamenim 0=>C a 1=>D
$bin = str_replace("1", "D", $bin);
$pole[] = $bin;
 // Prilozim vysledok do pola
$i++;
    }
return $pole;
  }

// ----------------------------------------------
// K zoznamu kombinacii mam naplnene pole s vysledkami (0,1,g)
//
function zloz_vysledky_fbf($n, $in_g) {
$pole = array();
$pocet_opakovani = pow(2, $n-2);
  // Retazec 101g sa bude opakovat: dva na n-tú minus 2 krat
$g = 0;
  // prièom g sa bude striedat hodnota 0 a 1 (parne, neparne opakovanie)
for ($i=1; $i<=$pocet_opakovani; $i++) {
```

```
if ($g == 1) $g = 0; else $g = 1;
 // Posledne cislo sa iba strieda
if ($i == $pocet_opakovani) $g = $in_g;
  // Úplne posledne èislo bude nasa vstupujúca pravdepodobnost_c
$pole[] = 1;
$pole[] = 0;
$pole[] = 1;
$pole[] = $g;
  }
return $pole;
  }

// -----------------------------------------------
// K zoznamu kombinacii mam naplnene pole s vysledkami (0,1,g)
//
function zloz_vysledky_fp($n, $in_g) {
$pole = array();
$pocet_opakovani = pow(2, $n-3);                // Oproti fbf $n-3
$g = 1;                                          // Oproti fbf zaèina $g=1
for ($i=1; $i<=$pocet_opakovani; $i++) {
if ($g == 1) $g = 0; else $g = 1;
if ($i == $pocet_opakovani) $g = $in_g;
$pole[] = 1;
$pole[] = 0;
$pole[] = 0;
$pole[] = 1;
$pole[] = 1;
$pole[] = 0;
$pole[] = 1;
$pole[] = $g;
  }
```

```
return $pole;
  }

// ------------------------------------------------
// Pre strategie FBF a FP
// Predpoèitanie vsetkych kombinacii tahov a odpovedi hraèa
// vysledkom je naplnene globalne pole $p_strategia1 resp. $p_strategia2
//
function priprav_strategiu($pc, $id_strategia) {
  global $zoznam_strategii, $p_strategia1, $p_strategia2;

  $nazov = $zoznam_strategii[$id_strategia];
  if (substr($nazov,0,3) != "FBF" and substr($nazov,0,2) != "FP") return;
 // Pre ine ako FBF a FP nepripravujem niè => odchod

  if (substr($nazov,0,3) == "FBF") {
 // Z nazvu strategie vyberiem poèet posledne
// pamätanych èlenov a pravdepodobnost_c
  $poc_pamatam = substr($nazov,3,1);
  $pravdepodobnost_c = substr($nazov,5,3);
if ($poc_pamatam == 'N') { $poc_pamatam = $GLOBALS['nekonecno'];
$pravdepodobnost_c = $GLOBALS['nekonecno']; }
// Pre nekoneèno si pamätam poslednych x-tahov
    $pole_vysledky   = zloz_vysledky_fbf($poc_pamatam, $pravdepodobnost_c);
 // pole s odpoveïami hraèa 0,1,g  pre kazdú kombinaciu tahov
}

  if (substr($nazov,0,2) == "FP") {
   // Z nazvu strategie vyberiem poèet posledne pamätanych
//èlenov a pravdepodobnost_c
  $poc_pamatam = substr($nazov,2,1);
```

```
  $pravdepodobnost_c = substr($nazov,4,3);
if ($poc_pamatam == 'N') { $poc_pamatam =
$GLOBALS['nekonecno']; $pravdepodobnost_c = $GLOBALS['nekonecno_g']; }
    $pole_vysledky   = zloz_vysledky_fp($poc_pamatam, $pravdepodobnost_c);
}

// Spoloène pokraèovanie pre obe strategie FBF aj FP
  $pole_kombinacie = zloz_kombinacie($poc_pamatam);
// pole s kombinaciami tahov. Napr. CCCCD  (binarne=1)
$pole_strategia = array();
 // vysledne pole kde kluè je kombinacia tahov a
//hodnota je odpoveï:  CCCCD = èislo
foreach($pole_kombinacie as $key=>$val)
  $pole_strategia[$val] = $pole_vysledky[$key];



// echor($pole_strategia);
if ($pc == 1) { $p_strategia1 = $pole_strategia; }
 // Naplnim prislusne globalne polia podla vstupujúceho
//prveho alebo druheho hraèa
if ($pc == 2) { $p_strategia2 = $pole_strategia; }
}

// -----
// return => modulo po deleni (urcovanie parne / neparne)
//
function mod2($x,$y) {
  $t = $x - (floor($x/$y)*$y);
  return $t;
  }
```

```
?>
```

```
?>
```