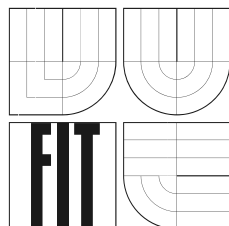


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ



**Analýza dynamiky evoluce
kooperace v opakovaných hrách
s nenulovým součtem**

2010

Jaroslav Bendl

Obsah

Obsah	3
1 Úvod a formulace zadání	4
2 Analýza řešeného problému	5
2.1 Vězňovo dilemma	5
2.2 Typologie strategií	6
2.3 Ekologická evoluce	8
3 Abstraktní model	9
4 Simulační model	10
5 Experimenty	11
5.1 Experimenty zkoumající úspěšnost strategií	11
5.2 Experimenty zkoumající dynamiku vývoje populace	14
6 Závěr	18

Kapitola 1

Úvod a formulace zadání

Evoluční teorie her je založena na myšlence, že úspěšnější strategie chování se při evolučním běhu šíří populací a mají tendenci potlačovat strategie slabší. Postupně dochází k úplné eliminaci těch nejslabších a vzniku dominance jedné strategie, která je označována jako *evolučně stabilní* [3]. Na základě tohoto předpokladu lze evoluční teorii her využít při zkoumání vzniku a evoluce kooperace v populaci hráčů hrajících různé strategie a díky tomu objasnit principy chování živočichů i rostlin při vzájemných interakcích. V rámci této práce bude cílem analyzovat vybrané strategie chování a využít evoluční teorii her jako analytický nástroj problémů typu opakovaného věžňova dilematu. Cílem projektu je tedy implementovat program simulující interakce jednotlivých forem chování na třídě opakovaných her s nenulovým součtem a pokusit se formulovat obecné závěry ohledně úspěšnosti zkoumaných typů strategií. Pozornost bude věnována rovněž samotné dynamice evoluce kooperace strategií.

Kapitola 2

Analýza řešeného problému

V rámci této kapitoly bude popsán problém věžňova dilematu a několik jeho obvyklých příkladů. Dále budou představeny analyzované herní strategie a jejich rozdělení do tříd.

2.1 Věžňovo dilema

Věžňovo dilema představuje třídu her s nenulovým součtem sloužící pro analýzu vzájemného chování jedinců. Úloha je obvykle popisována tak, že hráči zastupují vězně v káznici, kteří mají možnost spolupracovat či nespolutpracovat, přičemž kombinace jejich rozhodnutí jim určuje výplatu, tj. dobu, kterou si odpykají ve vězení. Přestože nejvyššího společného součtu by dosáhli při situaci “spolupráce x zrada”, nejvyššího individuálního zisku by dosáhli, pakliže by zradili, zatímco soupeř by spolupracoval. Právě vidina většího zisku a také obava ze zrady spoluhráče simuluje konflikt mezi sobeckým přáním jedince hrát strategii “vítěz bere vše” a oboustrannou výhodností kompromisu pro uspokojení potřeb ostatních hráčů. Úlohu lze tedy považovat za optimalizační problém s existencí vzájemně si odporujících podmínek.

Obecně se jako věžňovo dilema chápe každá situace následujícího typu:

		Hráč 2	
		spolupráce	zrada
Hráč 1	spolupráce	<i>(odměna, odměna)</i>	<i>(oškubání, pokušení)</i>
	zrada	<i>(pokušení, oškubání)</i>	<i>(trest, trest)</i>

kde zisky jsou uspořádány následovně: $oškubání < trest < odměna < pokušení$. Původní a dodnes nejčastěji používané hodnoty parametrů úlohy jsou uvedeny v tabulce 2.1.

V případě jednokolové hry je zřejmé, že se jedinec bude snažit maximalizovat svůj prospěch bez ohledu na své okolí a potom je dominantní strategií nespolutpracovat. Druhý hráč uvažuje obdobným způsobem, tudíž pro hru existuje jediná rovnováha, a to situace “zrada

		Hráč 2	
		spolupráce	zrada
Hráč 1	spolupráce	$(3, 3)$	$(0, 5)$
	zrada	$(5, 0)$	$(1, 1)$

Tabulka 2.1: Nejčastěji používané hodnoty parametrů problému opakovaného věžňova dilematu [1].

x zrada”. Již v předchozím odstavci však bylo řečeno, že nejvyššího součtu lze dosáhnout při spolupráci obou hráčů, proto tato rovnováha nepředstavuje Paretovsky optimální řešení.

V případě vícekolové hry se však strategie hráčů mění. Jelikož mají možnost v dalším kole oplatit soupeři předchozí zradu, racionální strategií hráče je proto spolupráce. Je přitom dokázáno, že se zvětšujícím se počtem opakování se Nashova rovnováha více přibližuje Paretovu optimu [7].

Věžňovo dilemma je názorným příkladem celé řady problémů reálného světa, ve kterých osobní ambice jednotlivců vedou ke kolektivnímu selhání. Ukázkovou situací může být placení koncesionářských poplatků (spolupráce = platit, zrada = neplatit) – odměnou za placení je možnost sledování programu, pokusem je program sledovat, ale neplatit, trestem je zrušení vysílání z důvodů nedostatečného výběru poplatků [4]. Podobných situací lze pochopitelně nalézt v každodenní realitě mnohem více.

Ochota ke zradě druhého hráče však nezávisí pouze na skutečnosti, zda se hra bude opakovat, nýbrž i na konkrétních hodnotách parametrů ve výplatní matici (payoff matrix). Čím vyšší obdrží hráč zisk v případě zrady, tím větší je pokušení ji provést. Uvedenou situaci demonstuje problém typu “lift dilemma” [2], ve kterém má parametr pokušení namísto 5 výplatu 8. Konkrétním příkladem pak budiž aukce uměleckých předmětů, kde je spolupráce vyjádřena dohodou mezi jedinými dvěma dražiteli, že v jednom kole nebude přihazovat jeden z nich a v druhém kole druhý z nich.

2.2 Typologie strategií

Interakci mezi dvěma strategiemi lze chápat jako sérii tahů hráčů na konkrétní dvojmatici výplat. Strategie určuje chování hráče a reakce na tahy soupeře, přičemž základní klasifikace strategií spočívá v rozdělení na statické a dynamické. Zatímco *statické strategie* své chování během interakce nemění, *dynamické strategie* obvykle určité změny během hry vykazuje – takové změny mohou být buď náhodné, nebo v nich může být nějaký řád. V dalším textu i v provedených experimentech budou zkoumány výhradně statické strategie.

Statické strategie lze rozčlenit na podmíněné a nepodmíněné. *Nepodmíněné statické strategie* se při svém jednání neohlíží na protihráče a jednají vždy podle pevně určeného

schématu. Příkladem budiž strategie *Gandhi*, která za všech okolností spolupracuje s protihráčem. *Podmíněné statické strategie* naopak při volbě tahu zohledňují chování protivníka a jejich rozhodnutí vychází z analýzy určitého počtu posledních svých kroků a kroků protivníka.

Statické strategie je možné dále klasifikovat i podle charakteru jejich chování. Zatímco *milé strategie* nikdy nevolí zradu jako první, *zákeřné strategie* alespoň v některých případech volí f, ačkoliv protihráč spolupracuje.

ALL C	Vždy spolupracuje - (C)*.
ALL D	Vždy zradí - (D)*.
PER CD	Střídá periodicky strategie (CD)*.
PER CCD	Střídá periodicky strategie (CCD)*.
PER DDC	Střídá periodicky strategie (DDC)*.
TFT	Půjčka na oplátku - Začíná spoluprací, potom kopíruje kroky soupeře z předchozího kola.
TF2T	Půjčka na dvě oplátky - Spolupracuje, pokud však soupeř dvakrát po sobě zradí, tak zradí i on.
SLOW TFT	Pomalá půjčka na oplátku - Začíná spoluprací, pokud však soupeř hrál dvakrát po sobě stejný tah, hraje opačný.
HARD TFT	Nelítostná půjčka na oplátku - Spolupracuje, pokud však soupeř alespoň v jednu z dvou předchozích kroků zradí, tak zradí i on.
SOFT MAJO	Něžná většinová - Začíná spoluprací, potom vybírá nejčastější volbu soupeře. V případě vyrovnaného poměru spolupráce a zrady volí spolupráci.
HARD MAJO	Krutá většinová - Začíná zradou, potom vybírá nejčastější volbu soupeře. V případě vyrovnaného poměru spolupráce a zrady volí zradu.
MISTRUST	Podezřívavá půjčka na oplátku - Začíná zradou, potom kopíruje kroky soupeře.
SPITEFUL	Nevraživec - Spolupracuje, pokud však soupeř zradí, bude navždy pouze zrazovat.
PAVLOV	Spolupracuje v případě, že v předchozím kroku oba hráči zvolili stejnou strategii, jinak zradí.
RANDOM	Spolupracuje s pravděpodobností $p = \frac{1}{2}$.

Tabulka 2.2: Charakteristika strategií uvažovaných v rámci této práce.

V roce 1981 uspořádal Robert Axelrod počítačový turnaj, v němž čtrnáct různých strategií, navržených jeho kolegy, hrálo opakované vězňovo dilema ve snaze objevit ideální strategii [1]. Výsledky ukázaly, že milé strategie obsadili překvapivě všech prvních osm

míst. Konečným vítězem se stala *půjčka na oplátku* vytvořená psychologem a herním teoretikem Anatolem Roportem, která sice v prvním kole volila spolupráci, ovšem v dalších kolech kopírovala kroky svého soupeře z přechozího kola. O její univerzalitě svědčí i fakt, že vyhrála i v druhém Axelrodově turnaji při celkovém počtu dvašedesáti účastníků ze šesti zemí. Axelrod výsledky této strategie zdůvodňoval takto: “Za ohromný úspěch půjčky na oplátku může kombinace její vstřícnosti, oplácení, schopnosti odpouštět a čitelnosti. Chová se vstřícně, tak, že se nedostává do zbytečných konfliktů. Podrazý dokáže oplácet stejnou kartou, a tím protihráče odradí od trvání na tricích, které by na ni chtěli zkoušet. Dokáže odpouštět, takže může obnovit vzájemnou spolupráci. A konečně je dostatečně čitelná, což protihráče motivuje k dlouhodobé spolupráci.” [6].

Půjčka na oplátku je úspěšnou strategií i v oblasti přírodních ekosystémů, což poprvé zdokumentoval Michael Lombardo v roce 1985 v článku “Vzájemná interakce vlaštovek stromových: Testování reciprocity v modelu půjčka na oplátku” [5]. Ačkoliv půjčka na oplátku není evolučně stabilní strategií, přesto experimenty prokázaly, že na rozdíl od hodnotnějších strategií (ve smyslu jejich větší tolerance k soupeřově zradě) nemůže být z populace zcela vytlačena zákeřnými strategiemi. Strategie uvažované v rámci této práce jsou uvedeny v tabulce 2.2.

2.3 Ekologická evoluce

Ekologická evoluce ¹ (ecologic evolution) je simulační metoda napodobující Darwinův přirozený výběr s cílem zjistit, jaké strategie chování se v populaci rozšíří a jaké postupně zaniknou. Evoluční běh obvykle začíná s rovnoměrným zastoupením analyzovaných strategií v populaci. Poté je v diskretních časových krocích přepočítáván počet jedinců s danou strategií podle výplatní matice zahrnující všechny uvažované strategie. Tento postup se přitom opakuje tak dlouho, dokud se populace nestane stabilní. Úspěšnou strategií je pak taková, která má na konci evolučního běhu největší zastoupení v populaci.

¹zde je termín *ekologie* chápán ve svém původním významu, tj. jako biologická věda zabývající se vztahem organismů a jejich prostředí a vztahem organismů navzájem

Kapitola 3

Abstraktní model

V rámci tohoto projektu je analýza vybraných forem chování studována na problému opakovaného věžňova dilematu popsaného v kapitole 2.1. Vývoj zastoupení jednotlivých strategií populace v čase je simulován prostřednictvím ekologické evoluce představené v kapitole 2.3, přičemž výčet zkoumaných strategií včetně jejich stručné charakteristiky se nachází v kapitole 2.2.

Vzájemná interakce jednotlivých strategií v každém evolučním kroku je vyhodnocována pomocí výplatní matice zkonstruované před samotným spuštěním evoluce. Tato výplatní matice přiřazuje zisky pro střetnutí libovolné dvojice strategií. Zisky jsou přitom vypočítány na základě dříve simulovaných interakcí každé dvojice strategií o stanoveném počtu kroků (tento počet kroků však jedinci neznají, tudíž nemohou blížící se konec doby spolupráce zahrnout do svého rozhodování).

Výsledek evolučního běhu založeného na ekologické evoluci ukazuje podobu konečné stabilizované populace. Dynamika tohoto vývoje však může být poměrně složitá a lze ji významně ovlivnit rozdílným počtem jedinců hrajících různé strategie. Zkoumání dynamiky vývoje je jedním z hlavních cílů této práce, přičemž konkrétní experimenty se nachází v kapitole 5.

Kapitola 4

Simulační model

Implementace projektu je provedena v jazyce C++. Po spuštění programu dojde k rozpoznání předaných argumentů, přičemž povinné je uvedení cesty ke vstupnímu souboru se zadáním úlohy i k výstupnímu souboru pro uložení výsledků. Detekce těchto hodnot je realizována v rámci funkce `main()`.

Bezprostředně po získání obou argumentů příkazového řádku je řízení předáno instancí třídy `Manager`. V ní nejprve provede funkce `LoadData()` analýzu souboru k získání výplatní matice řešeného problému, množiny vyšetřovaných strategií a jejich počátečního zastoupení v populaci. Pro jednotlivé strategie se vytvoří instance tříd `ProbabilisticStrategy`, `BufferStrategy` či `PeriodicalStrategy`, které jsou odvozeny od abstraktní třídy `Strategy`. Ze vstupního souboru je dále získána délka interakce `matchLength` mezi strategiemi ovlivňující podobu výplatní matice a parametr maximálního počtu generací evolučního běhu `generationCount` včetně prahu `endTreshhold` představujícího hodnotu změny populace mezi dvěma posledními generacemi, při které již považujeme populaci za stabilní.

Jakmile jsou získány výše uvedené údaje o úloze, lze přistoupit ke konstrukci výplatní matice pro všechny uvažované strategie. Postupně se vyšetřuje zisk ze střetnutí všech možných dvojic, přičemž toto střetnutí strategií probíhá po dobu `matchLength` kroků a využívá základní výplatní matice získanou ze vstupního souboru v kombinaci s tahy jednotlivých hráčů odpovídajících jejich strategií. Tento výpočet má na starost metoda `GetPayoffMatrix()` třídy `Payoff`.

Znalost výplatní matice pro jednotlivé strategie umožňuje provedení samotné ekologické evoluce, která je implementována ve třídě `Evolution`. Pro provedení evoluce trvající maximálně `generationCount` kroků je určeno konečné zastoupení vyšetřovaných strategií v poslední generaci a zároveň pomocí programu `gnuplot` vygenerován graf vývoje.

Kapitola 5

Experimenty

Experimenty s vytvořeným modelem jsou rozděleny do dvou skupin. V první z nich je zkoumána úspěšnost jednotlivých strategií s cílem zjistit takovou formu chování, která se stane dominantní v řadě různých populací s odlišným obsazením. V druhé z nich je pozornost věnována různým podobám dynamiky evoluce kooperace.

5.1 Experimenty zkoumající úspěšnost strategií

Jak již bylo řečeno v úvodu kapitoly, tato sada experimentů je zaměřena na vzájemné porovnání vyšetřovaných strategií s cílem zjistit co možná nejuniverzálnější úspěšnou formu chování. Ve všech experimentech se shodně používá 100 jedinců v počáteční populaci pro každou vyšetřovanou strategii, 100 vzájemných interakcí při konstrukci výplatní matice, práh pro ukončení evoluce 0.005 a standardní výplatní matice pro problém věžňova dilematu (viz kapitola 2.1).

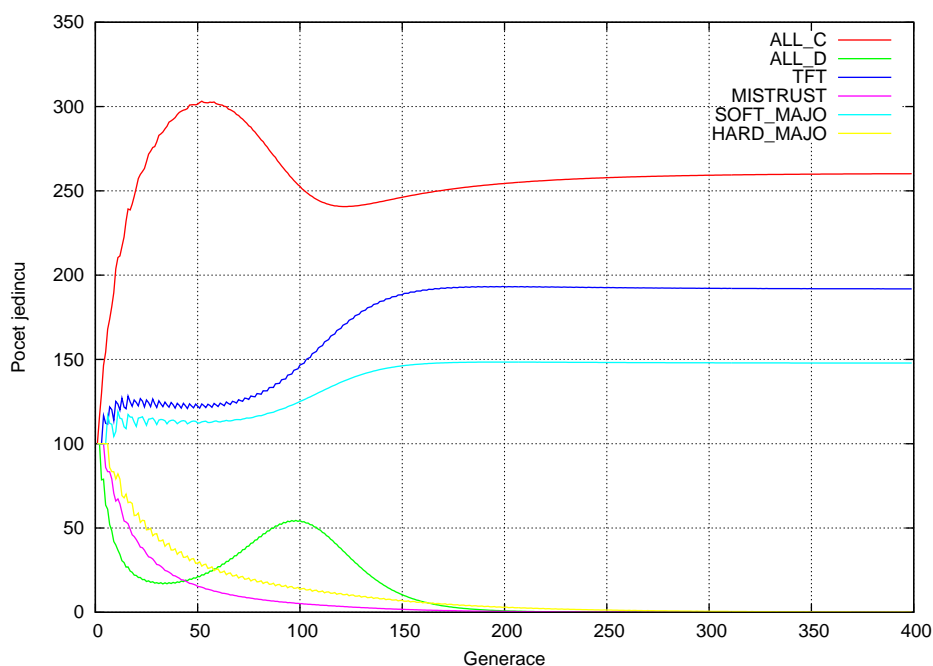
V prvním experimentu (viz obrázek 5.1) jsou porovnáváni tři zástupci “milých” strategií (ALL C, TFT, SOFT MAJO) a tři zástupci “zákeřných” strategií (ALL D, MISTRUST, HARD MAJO). Z výsledného grafu je patrné, že zástupci milých strategií jsou všeobecně úspěšnější.

V druhém experimentu (viz obrázek 5.2) je pozornost opět věnována porovnání milých a zákeřných strategií. Ačkoliv v předcházejícím pokusu v populaci jednoznačně převládly milé strategie, zde je ukázáno, že bez milé strategie disponující schopností oplácet získají zákeřné strategie rychle převahu a stanou se v populaci dominantní. Graf ukazuje, že všechny milé (ALL C, PAVLOV) i neutrální (RANDOM, PER CD) strategie jsou v tomto pokusu záhy vytěsněny těmi zákeřnými (ALL D, HARD MAJO). Zatímco v prvním případě plnila roli jakéhosi ochránce ostatních milých strategií “půjčka na oplátku” (TFT), zde žádná taková milá strategie neexistuje.

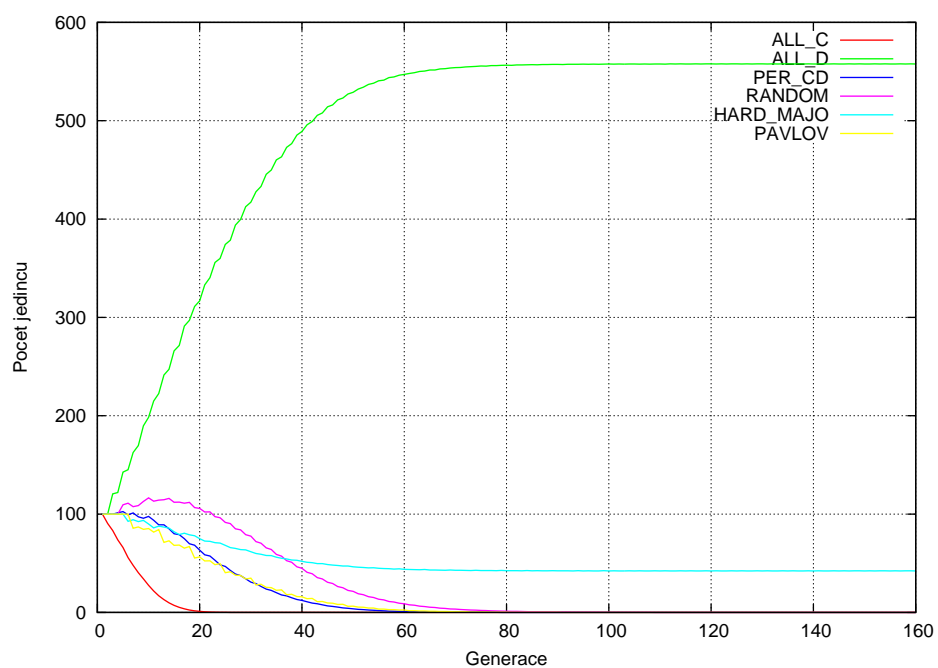
Třetí experiment (viz obrázek 5.3) ukazuje, jak si vedou různé varianty strategie půjčky na oplátku (TFT, TF2T, SLOW TFT, HARD TFT) proti zákeřným strategiím (ALL D, PER DDC, HARD MAJO). V tomto případě zvítězila strategie nelítostné půjčky na oplátku (HARD TFT),

která je méně tolerantnější ke zradě než čistá TFT.

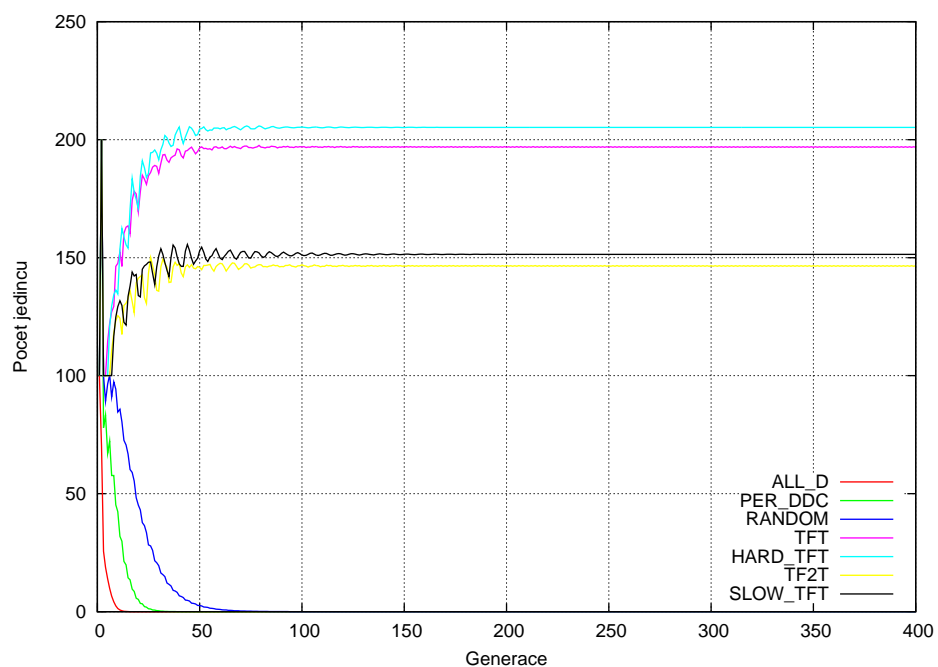
Pokud bychom připustili model, v němž funguje reprodukce a mutace strategií, dostali bychom zřejmě jiné výsledky. Po eliminaci zákeřných forem chování by pravděpodobně v populaci začalo přibývat hodnějších strategií než TFT. Takový stav by nebyl evolučně stabilní, neboť s výskytem mutovaného jedince se zákeřným chováním by půjčka na více oplátek nebyla efektivní. Záleží tedy vždy na konkrétním prostředí a jeho modelu.



Obrázek 5.1: Vývoj strategií v populaci. Graf ukazuje, že milé strategie jsou úspěšnější než zákeřné strategie.



Obrázek 5.2: Vývoj strategií v populaci. Graf ukazuje, že milé strategie bez zástupce schopného odvety neobstojí oproti zákeřným strategiím.



Obrázek 5.3: Vývoj strategií v populaci. Graf porovnává varianty strategie půjčky na oplátku.

5.2 Experimenty zkoumající dynamiku vývoje populace

Křivky v grafu ekologické evoluce znázorňující vývoj počtu jedinců s danou strategií v čase mají ve většině případů monotónní průběh, přičemž typicky dojde během relativně malého počtu kroků k ustálení stavu celé populace. V některých případech je však možné pozorovat určité oscilace, podle nichž lze dynamiku vývoje klasifikovat do několika skupin, které zde budou postupně prezentovány včetně ukázkových příkladů s následujícím nastavením parametrů:

- **Uvažované strategie:** PER CCD, PER DDC, SOFT MAJO,
- **Počáteční populace:** 100 výskytů jedinců každé strategie,
- **Výplatní matice:** $spolupráce = 3$, $trest = 1$, $pokušení = 5$, $oškebání = 0$,
- **Délka střetnutí:** 100,
- **Maximální počet generací:** 1000,
- **Práh pro ukončení:** 0.005.

Monotónní konvergence

Monotónní konvergenci lze pozorovat u většiny úloh (například u již diskutovaného experimentu z obrázku 5.2). Při tomto druhu chování mají křivky monotónní průběh a rychle konvergují, takže se populace brzy ustálí v určitém stavu, který se dále nemění.

Tlumená oscilace

Tlumená oscilace značí chování, při kterém postupně dochází k tlumení kmitavého pohybu křivek, tedy zmenšování amplitudy kmitu. Populace postupně konverguje k ustálenému stavu, což ovšem trvá déle, než v předchozím případě. Ukázka takového chování je uvedena na obrázku 5.4. Tato ukázka vychází ze základních parametrů uvedených na začátku této podkapitoly.

Buzená oscilace

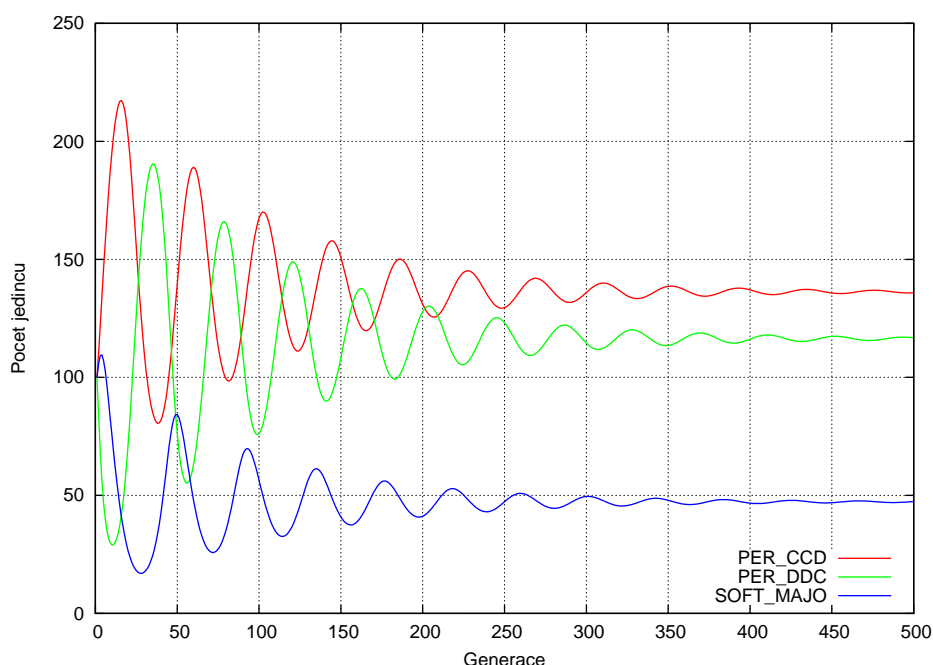
Buzená oscilace značí chování, při kterém postupně dochází k buzení kmitavého pohybu, tedy zvětšování amplitudy kmitu. Populace tedy nedojde k ustálenému stavu. Ukázka takového chování je uvedena na obrázku 5.5. Tato ukázka vychází ze základních parametrů uvedených na začátku této podkapitoly, ovšem počty jedinců v počáteční populaci jsou tyto: PER CCD=400, PER DDC=300, SOFT MAJO=200 a délka střetnutí je nastavena na 8.

Periodická oscilace

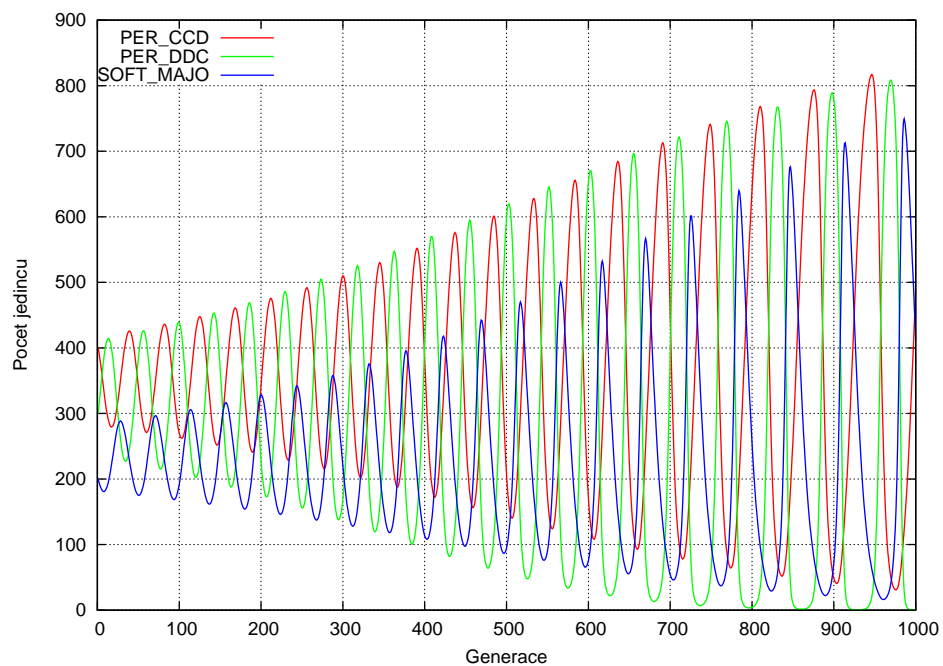
Periodická oscilace značí chování, při kterém není kmitání křivek ani buzeno, ani tlumeno. Populace tedy nedojde k ustálenému stavu. Ukázka takového chování je uvedena na obrázku 5.6. Tato ukázka vychází ze základních parametrů uvedených na začátku této podkapitoly, ovšem ke třem základním strategiím je přidána ještě strategie **HARD MAJO** a délka střetnutí je nastavena na 8.

Anharmonická oscilace

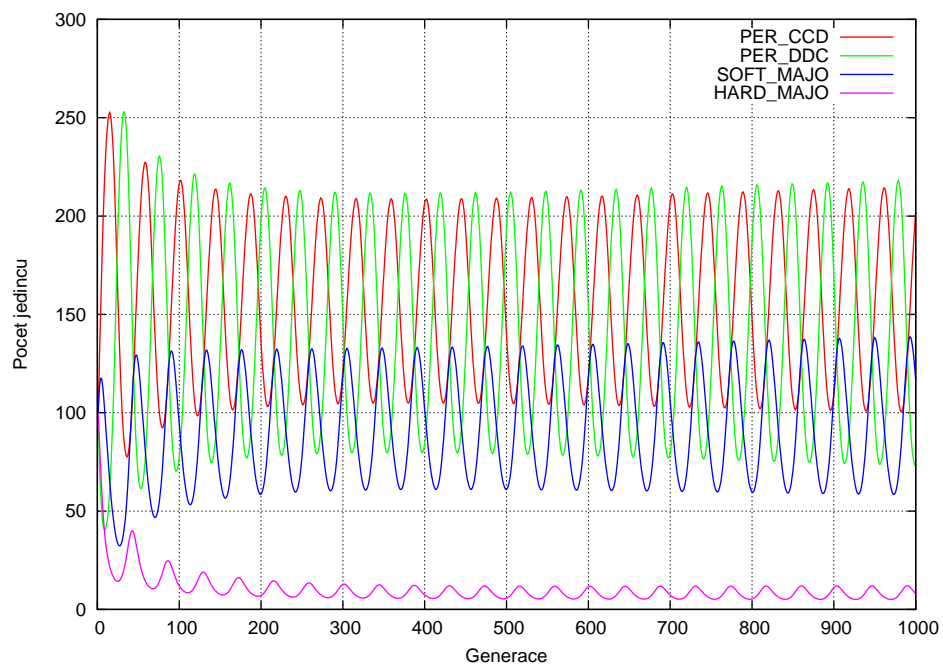
Anharmonická oscilace značí chování, při kterém nelze pohyb křivek vnímat jako harmonický. Ukázka takového chování je uvedena na obrázku 5.7. Tato ukázka vychází ze základních parametrů uvedených na začátku této podkapitoly, ovšem parametr *pokušení* ve výplatní matici se snížil na hodnotu 4 – zrada tedy není tak výhodná.



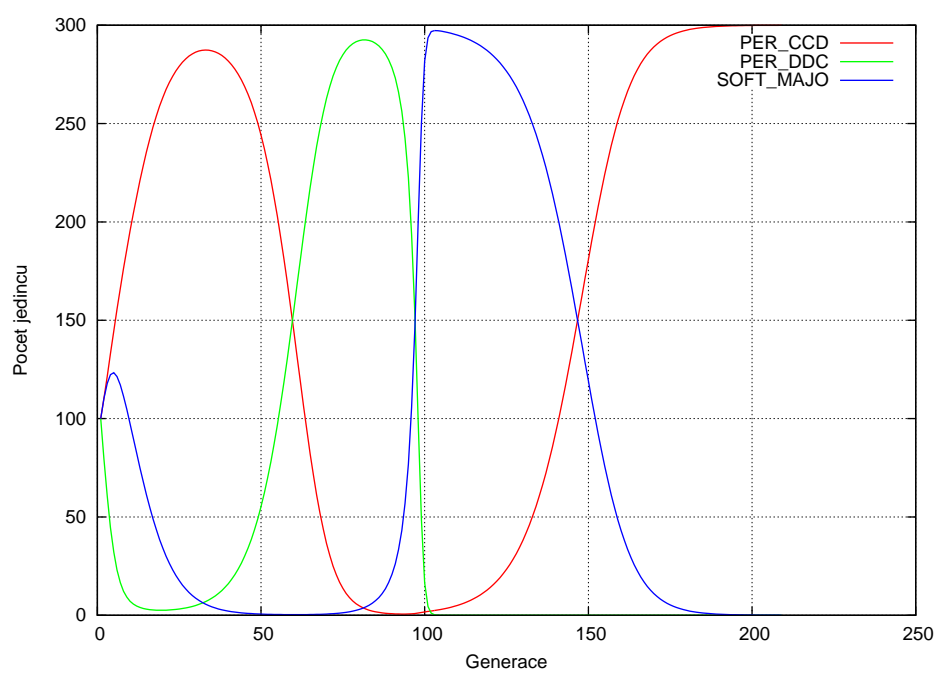
Obrázek 5.4: Vývoj strategií v populaci. Ukázka tlumené oscilace.



Obrázek 5.5: Vývoj strategií v populaci. Ukázka buzené oscilace.



Obrázek 5.6: Vývoj strategií v populaci. Ukázka periodické oscilace.



Obrázek 5.7: Vývoj strategií v populaci. Ukázka anharmonické oscilace.

Kapitola 6

Závěr

Při vypracování projektu byly dodrženy cíle formulované v kapitole 1. Program je schopen provádět ekologickou evoluci na opakované hře typu vězňova dilematu s libovolnou výplatní maticí. Celkem je možné experimentovat až s patnácti odlišnými strategiemi chování, nastavovat jejich počáteční výskyty a délky střetnutí.

V rámci této práce byly studovány strategie chování z hlediska jejich úspěšnosti ve vzájemných interakcích, přičemž jako nejuniverzálnější úspěšná strategie byla v souladu s tvrzením citované literatury zjištěna tzv. půjčka na oplátku. Práce se dále zabývala analýzou dynamiky vývoje kooperace a na základě četných pokusů se došlo k rozdělení na pět různých kategorií (monotónní konvergence, tlumená oscilace, periodická oscilace, buzená oscilace a anharmonická oscilace).

Program by jistě bylo možné dále rozvíjet. Ačkoliv je uvažováno patnáct různých strategií, dnes jich lze nalézt v literatuře daleko více a jejich implementace by jistě přispěla k větší důvěryhodnosti získaných výsledků. Rovněž by bylo vhodné experimentovat s modelem umožňující reprodukcí a mutaci strategií chování.

Literatura

- [1] Axelrod, R.: *The Evolution of Cooperation*. Basic Books, 1985.
- [2] Delahaye, J. P.; Mathieu, P.: The Iterated Lift Dilemma or How to Establish Meta-cooperation with your Opponent. In *Complexity, Society and Liberty*, editace P. Lemiux, Trois-Rivieres, Canada, 1996.
- [3] Flégr, J.: *Zamrzlá evoluce*. Academica, první vydání, 2001.
- [4] Hykšová, M.: Teorie her. [online].
URL <http://euler.fd.cvut.cz/predmety/teorie_her/hry.pdf>
- [5] Lombardo, M. P.: Mutual Restraint in Tree Swallows: A Test of the TIT FOR TAT Model of Reciprocity. In *Science* 227, 1985, s. 1363–1365.
- [6] Ridley, M.: *The Origins of Virtue*. Penguin Books Ltd, 1997.
- [7] Vincent, T. L.; Brown, J. S.: *Evolutionary Game Theory, Natural Selection and Darwinian Dynamics*. Cambridge University Press, první vydání, 2005.