

# Pravdepodobnostná väzenská dilema

Michal Vaľko, valko~sturak.sk, FMFI UK BA

*Tento text bol vypracovaný ako „case study“ na predmet Evolučné algoritmy prednášaný  
Doc. J. Pospíchalom v zimnom semestri 2003/2004 na Fakulte matematiky, fyziky a informatiky Univerzity Komenského v  
Bratislave.*

Väzenská dilema je klasickým a veľmi študovaným problémom na štúdium sociálnych javov v spoločnosti, ich modelovanie a hľadanie optimálnych stratégií v rôznych oblastiach ako biológia, sociálne vedy, filozofia, ekonómia, matematika, teória hier a informatika. Prvý krát uzrela svetlo sveta v roku 1950. Zverejnili ju Merrill Flood spolu s Melvinom Drescherom. Vernosť jej modelovania plynie z toho, že úspešnosť stratégie vo väzenskej dileme závisí na dobrom bodovom zisku v súťaži s rovnako úspešnými stratégiami ale aj s veľkou plejádou ďalších stratégií. Cieľom tejto práce je evolúciou získať optimálne stratégie jednak s deterministickým rozhodovaním ale aj stratégie so stochastickým vplyvom a vzájomne tieto stratégie porovnať.

## Teória hier a teória rozhodovania

Teória hier rozlišuje dva typy hier: hry s nulovým a hry s nenulovým súčtom. Hry s nulovým súčtom sú hry typu: buď vyhrám 100 € alebo 100 € prehrám, čiže „vyhrám“ -100 €, a teda súčet mojej možnej výhry je nulový. Väzenská dilema však patrí k tým druhým, hrám s nenulovým súčtom, teda hrám, kde takéto pravidlo neplatí. Takéto hry sa vyznačujú väčším počtom možností ako hrať a teda majú blízko k teórii rozhodovania, z ktorej teória hier vychádza. Teória rozhodovania sa snaží získať postupy ako sa rozhodnúť, ak nepoznáme reakciu prostredia, prípadne nevieme v akom prostredí sa nachádzame. Cieľ je motivovaný očividne reálnymi situáciami: získať aj v neznámom prostredí, čo najviac. Teória hier neuvažuje len vzťah jedinec-prostredie, ale oproti nám stojí protihráč alebo aj protihráči, ktorí majú (ako inak) rovnaký cieľ.

## Väzenská dilema

### Príbeh

Dvaja kriminálnici Bert a Ernie boli zatknutí za menší zločin: krádež auta. Uväznili ich v oddelených celách, teda nedokážu medzi sebou komunikovať. Polícia si však myslí, že Bert ako aj Ernie sú vinní v závažnejších zločinoch, chcú im našiť krádež banky a to, že zabili člena ochrannej služby pri krádeži mikročipu. Prešibaný a všetkými masťami mazaný detektív chce, aby obaja zradili toho druhého (teda nespolupracovali navzájom). Ak ani jeden z nich neprehovorí (teda spolupracujú, a predstierajú nevinu), obaja dostanú rok väzenia za krádež auta. Ak obidvaja zradia druhého, dostanú 10 rokov za vraždu. Ak jeden prehovorí a druhý ostane mlčať, zrádzajúceho (ktorý prehovoril) prepustia za pomoc pri vyšetrowaní a druhý (spolupracujúci s komplicom) dostane celý trest 20 rokov väzenia. Dilemou je, či prehovoriť, alebo mlčať. Ak budú spolupracovať obaja získajú viac ako keď obaja zradia, ale keď zradí práve jeden, tak ten získa ešte viac...

## Všeobecný prípad

Konkrétna inštancia väzenskej dilemy závisí len od distribúcií trestov a odmien pre jednotlivých väzňov. Nazýva sa *tabuľka platieb*<sup>1</sup>. Štandardne je symetrická, teda každý z väzňov je v rovnakej pozícii. Pre náš príbeh vyzerá tabuľka takto<sup>2</sup>:

	<i>Ernie spolupracuje</i>	<i>Ernie zradí</i>
<i>Bert spolupracuje</i>	[1,1]	[20,0]
<i>Bert zradí</i>	[0,20]	[10,10]

Tabuľka 1 Distribúcia trestov

Väčšinou však tabuľku uvádzame vo forme odmien (väčšie číslo znamená väčší zisk). V tejto práci budeme pracovať s nasledujúcou tabuľkou platieb:

	<i>B spolupracuje</i>	<i>B zradí</i>
<i>A spolupracuje</i>	[3,3]	[0,5]
<i>A zradí</i>	[5,0]	[1,1]

Tabuľka 2 Tabuľka platieb

Vo všeobecnosti, tabuľka ma takýto tvar ( $C^3$  – spolupracuje,  $D^4$  – zradí):

	<i>B spolupracuje</i>	<i>B zradí</i>
<i>A spolupracuje</i>	CC	CD
<i>A zradí</i>	DC	DD

Tabuľka 3 Všeobecná tabuľka platieb

Platí  $DC > CC > DD > CD$  (z pohľadu prvého hráča). Teda najvýhodnejšie pre mňa (ako väzňa) je keď ja zradím a komplic spolupracuje, potom ak spolupracujeme obaja, menej výhodne ak spolu zrádzame a najmenej výhodné keď ma komplic zradí a pritom ja sa pokúšam spolupracovať. Ako bude asi uvažovať Bert? „Ak ma Ernie zradí, tak by som bol hlupák aby som zatĺkal. Ale ak ma podrží tak lepšie z toho vyjdem, ak ho zradím ja.“ A tak ho zradí. Ernie však bude uvažovať rovnako a tak dostaneme riešenie DD. Stav nazveme rovnovážnym. Pre oboch väzňov je však nepríjemné že je horší ako tiež symetrický stav CC.

Rôzne typy podobných študovaných hier uvádza nasledujúca tabuľka<sup>5</sup>:

<i>Prisoner dilemma</i>	<i>Chicken</i>	<i>Stag hunt</i>	<i>Deadlock</i>
DC	DC	CC	DC
CC	CC	DC	DD
DD	CD	DD	CC
CD	DD	CD	CD

Tabuľka 4 Rôzne typy hier

Napríklad hra Chicken vyzerá nasledovne. Podgurážení kamaráti Puf a Muf si idú dokázať, že sú tvrdí chlapi.

<sup>1</sup> payoff matrix

<sup>2</sup> prvé/druhé z dvoch čísel v zátvorke, uvádza, koľko rokov si odsedí Bert/Ernie

<sup>3</sup> cooperates

<sup>4</sup> defect

<sup>5</sup> CC, CD, DC a DD sú v tabuľke zoradené podľa výhodnosti z pohľadu prvého hráča

Nasadnú na motorky a vyrazia oproti sebe, kto prvý uhne je mäkký (slaboch, padavka, žena, kurča...) . Samozrejme, stav keď uhne aspoň jeden je lepší pre oboch, a z tých stavov je lepší ten ak uhne ten druhý.

## Iterujeme...

Klasický prípad väzenskej dilemy sa konal jednorazovo. Stratégia bola jasná: zradiť. Do inej situácie sa však dostaneme, ak hru budeme opakovať. Teda celým procesom (budeš zatĺkať alebo sa priznáš?) prejdú väzni napríklad 10-krát. Môžeme tiež uvažovať viac väzňov. Zvykneme pridať k podmienke  $DC > CC > DD > CD$  aj podmienku  $2CC > DC + CD$ . Jej praktický význam je taký že by sa malo viac oplatíť spolupracovať ako striedavo sa priznávať a zatĺkať.

Ak by si väzni nepamätali ťahy svojho protihráča a pravdepodobnosť komplicovej spolupráce bola rovnaká ako zrada, dostaneme sa k tomu istému ako v neiterovanej verzii problému. Totiž priemer zo zisku keď zrádzam je vyšší ako keď spolupracujem<sup>6</sup>. Aj v reálnom živote si však väzni „ľahy“ svojho protihráča pamätajú. Pre zjednodušenie uvažujme jednoduchú (jednobitovú) pamäť, teda väzeň si pamätá, ako sa k nemu zachoval komplic naposledy a podľa toho sa rozhoduje. K tomu ešte pridáme informáciu ako väzeň ťahá prvýkrát, teda či začína spolupracou alebo zradou. Práve sme si formalizovali pojem *stratégia*. Každú takúto stratégiu možno kódovať trojbitovým číslom a teda máme celkom 8 možných stratégií. Kód stratégie je trojica  $[a,b,c]$ ,  $[a,b,c] \in [0,1]^3$  a je prvý ťah, b je ťah ak komplic v prechádzajúcom ťahu spolupracoval, a c keď zrádzal. 0 znamená *spolupracuj* a 1 *zraď*. Pozrime sa ako dopadne týchto 8 stratégií vo vzájomných súbojoch. Necháme ich „ľahy“ 10 ťahov. Priemer za 1 ťah ukazuje tabuľka 5:

	[000]	[001]	[010]	[011]	[100]	[101]	[110]	[111]
[000]	3	3	0,3	0,3	2,7	2,7	0	0
[001]	3	3	2,1	1,1	2,9	2,5	1,9	0,9
[010]	4,8	2,6	2	0,4	4,3	2,1	0	0
[011]	4,8	1,6	4,4	1,2	4,5	1,3	4,1	0,9
[100]	3,2	2,9	0,8	0,5	2,8	2,5	0,4	0,1
[101]	3,2	2,5	2,6	1,3	3	1	2,4	1
[110]	5	2,4	5	0,6	4,4	1,9	2	0,1
[111]	5	1,4	5	1,4	4,6	1	4,6	1

Tabuľka 5 Vzájomný súboj všetkých možných deterministických stratégií (priemer za 10 ťahov)

## Pravdepodobnostná verzia

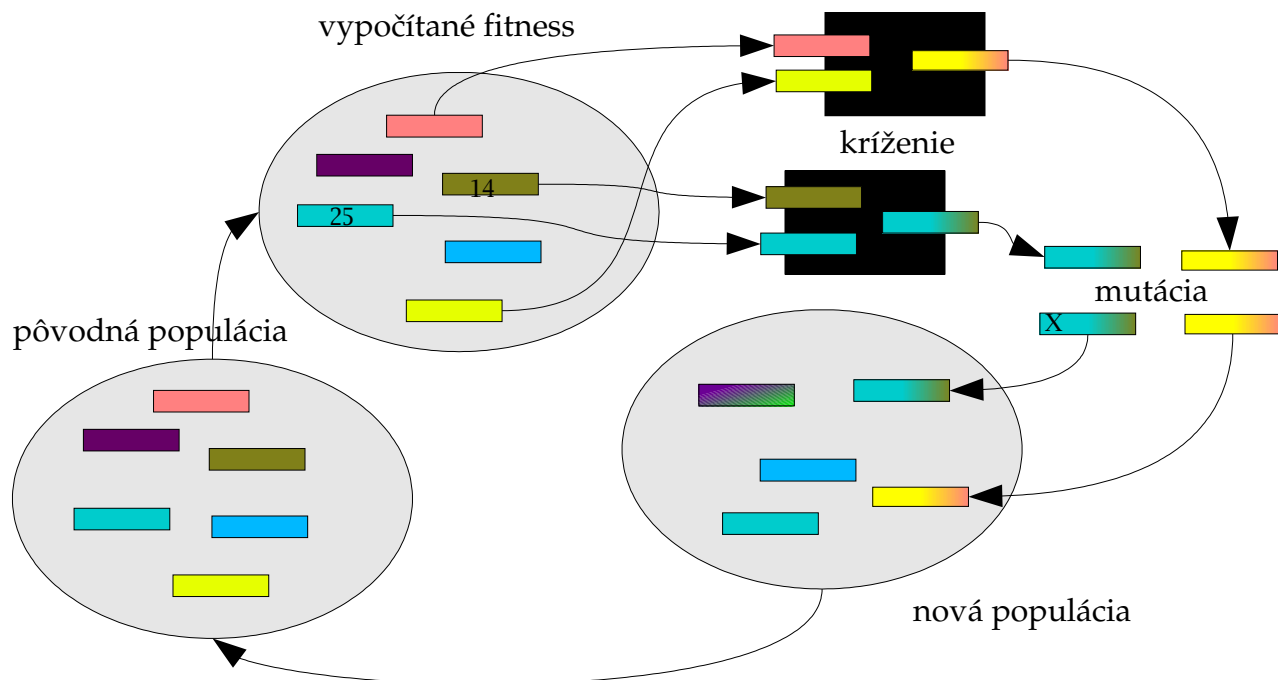
Skúsime upustiť od požiadavky, že rozhodnutie v nasledujúcom ťahu je striktné determinované predchádzajúcim a umožníme väzňovi sa rozhodnúť na základe akejsi pravdepodobnosti. Presnejšie, stratégia bude trojica  $[a,b,c]$  kde jednotlivé zložky nadobúdajú hodnoty medzi 0 a 1 pričom každá vyjadruje pravdepodobnosť zrady. Všimnime si, že takto definovaná (pravdepodobnostná) verzia je pre celé  $a,b,c$  spätne kompatibilná s pôvodnou nedeterministickou verzou, čo nám uľahčí ich vzájomné porovnávanie a súťaženie. Takýchto stratégií je samozrejme nekonečne veľa.

<sup>6</sup> priemer pri zrade je  $(5+1)/2=3$ , pri spolupráci  $(3+0)/2=0,5$  - v prípade rovnakých pravdepodobností, že sa stretneme: so spolupracujúcim ako so zrádzajúcim

# Evolujeme stratégie

Kvalita, úspešnosť alebo fitness konkrétnej stratégie priamo súvisí s bodmi ktoré získa konfrontáciou s ostatnými stratégiami. Preto použijeme evolučné algoritmy na zistenie úspešnosti jednotlivých stratégií a na nájdenie tej najperspektívnejšej. Na začiatku náhodne vyberieme populáciu. Pre každú stratégiu spočítame fitness, ako súčet bodov získaných zo vzájomných súbojoch s ostatnými stratégiami. Ďalšiu generáciu vytvoríme kvázináhodným výberom dvoch stratégií z populácie, ich potenciálnym krížením a mutáciou a zaradením potomkov do nasledujúcej generácie. Treba podotknúť, že v tomto algoritme nejde o to, aby sme získali maximálnu fitness. Tu je odpoveď známa, má ju stratégia 111 alebo 110 v populácii so samými 010 a 000 (tab. 5). Našou snahou je evolúciou získať stratégiu, ktorá bude úspešná v kontexte so širokým spektrom stratégií. Pseudokód nášho algoritmu:

```
P := {náhodne vybratá populácia stratégií};
while (t < tMAX) and (not stop_criterion) do
begin
  t := t + 1;  Q ← ∅ ;
  while |Q| < |P| do begin
    pre každú stratégiu v P spočítaj fitness v súboji s ostatnými;
    vyber ruletou 2 stratégie s1 a s2 z populácie P;
    if (random < PREPRO) [q1, q2] := Reproduction(s1, s2);
    else [q1, q2] := [s1, s2];
    Q ← Q ∪ {q1, q2};
  end;
  P := Q;
  if (convergent_criterion) stop_criterion := true;
end;
```



Obrázok 1 Schéma použitého genetického algoritmu

## Popis algoritmu

Súboj znamená daný počet ťahov medzi dvoma stratégiami (napr. 10). Ruletu realizujeme rozdelením intervalu dĺžky súčtu všetkých fitness na podintervaly (podľa fitness jednotlivých stratégií) náhodným výberom prvku z intervalu a potom binárnym prehľadávaním zistíme, ktorej stratégii patrí. Reprodukcia pozostáva z jednobodového kríženia (kríženie v takýchto chromozómoch môže nastať len v 2 bodoch) a eventuálnou mutáciou s voliteľnou pravdepodobnosťou. Konvergenčné kritéria budú napríklad homogenita populácie.

## Klasický prípad

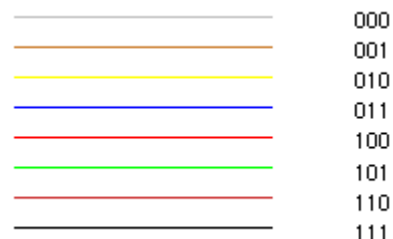
Použijeme navrhnutý genetický algoritmus na štúdium klasického deterministického iterovaného prípadu. Nech je veľkosť populácie 100 (300 500), pravdepodobnosť reprodukcie 0,5 (0,2 0,7) pravdepodobnosť mutácie 0,01 (0,05 0,005). Počet generácií obmedzme na 500.

## Výsledky

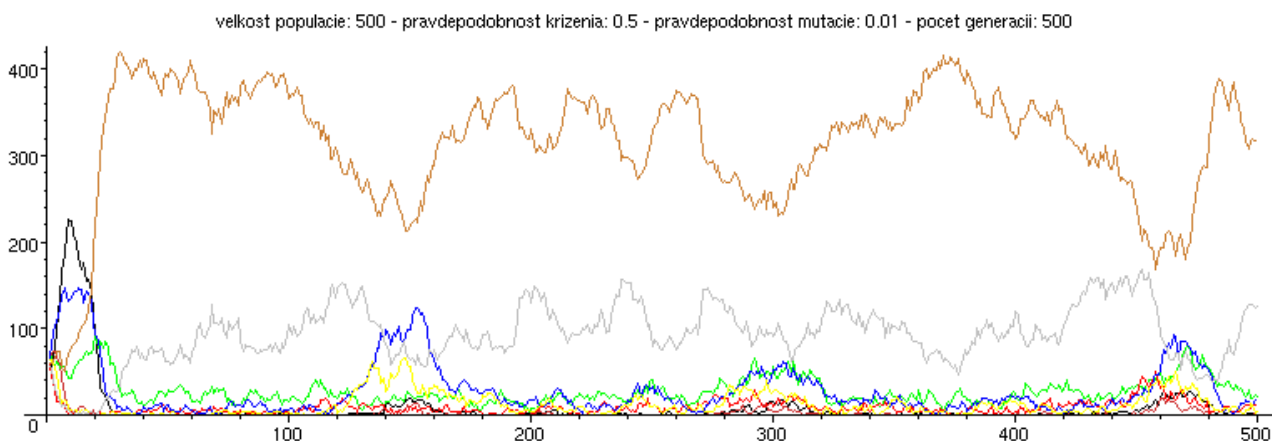
Urobili sme desiatky pokusov s rôznymi veľkosťami populácie a rôznymi pravdepodobnosťami. Väčšina pokus dopadla ako ukazuje graf na obrázku 3. Zo začiatku má najväčšie zastúpenie stratégia 111, takzvaný *zloduch* (vždy zrádza), neskôr je zastúpenie klesá na úkor známej stratégie 101 alebo TFT<sup>7</sup> (začína spolupracou a potom opakuje „ľahý“ svojho protihráča v predchádzajúcom ťahu, recipročná stratégia). Potom sa ešte v menšej miere objaví stratégia 000, tzv. *dobrák* (vždy spolupracuje). Ako to vysvetlíme?

## Interpretácia výsledkov

Zo začiatku je distribúcia stratégií náhodná. Vyhráva stratégia 111 pretože sa najviac prižiuje na „hlúpych“ stratégiách ako 010 110 100 000 (všimnime si že žiadna z týchto zradu netrešťa). Takto sa 111 stane významnou zložkou populácie. Jej fitness je však stále menšia. Závisí, tak ako je to u všetkých stratégií vo väzenskej dileme, od kontextu stratégií v ktorom sa nachádza. Na zloduchoch tak zloduch veľa nezarobí, tým že zloduchovia pribúdajú, nemajú na kom profitovať. Spolu s ňou sa objaví stratégia 101 (podobná TFT, začína však zradou) ktorá ma zo zloduchmi rovnakú bilanciu (viď tab. 5). Následne však emerguje TFT. Zloduchovia na nej veľa nezískajú. Sama so sebou však zarába viac ako zloduchovia a vie získať aj v súťaži s hlúpychmi stratégiami. Nakoniec pribudne aj dobráčka 000, pretože TFT jej dobráctvo „nezneužíje“.



Obrázok 2 Farby stratégií



Obrázok 3 Pribeh evolúcie (počet stratégií v závislosti od epochy)

7 Tit For Tat – „a ko ty mne tak aj ja tebe“

Axelrod pripisuje úspech TFT v kontexte rôznych stratégií (aj iných ako sme uvažovali) jej štyrom vlastnostiam:

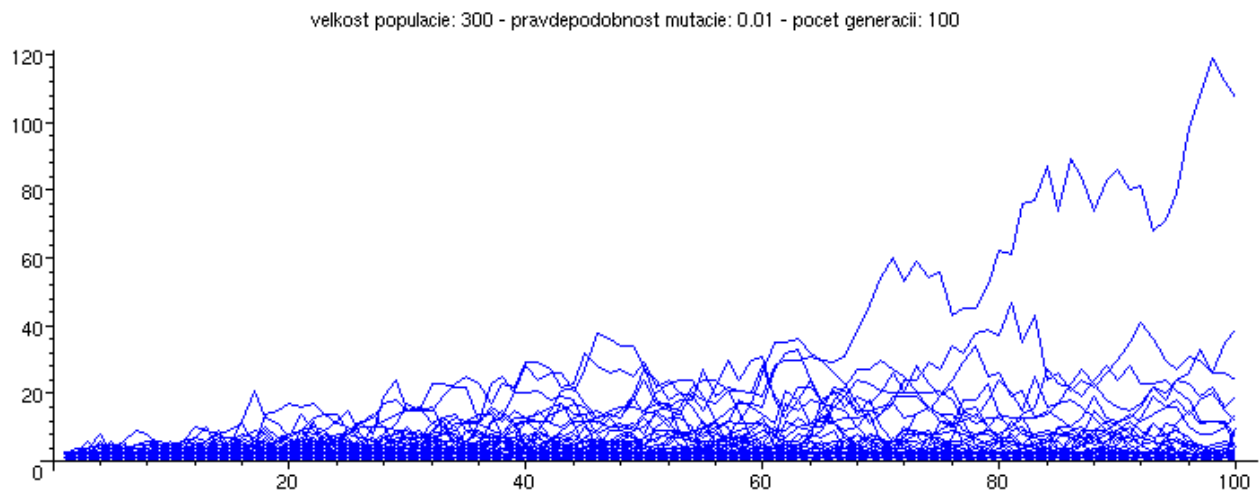
1. TFT je dobrá<sup>8</sup> – nikdy nie je prvá, čo zradí
2. TFT odvracia útok<sup>9</sup> – to ju robí odolnou voči zneužitiu stratégiami, ktoré nie sú dobré
3. TFT je odpúšťajúca<sup>10</sup> – má vôľu spolupracovať aj so stratégiami, ktoré ju zradili (okrem posledného kola)
4. TFT je prehľadná<sup>11</sup> – je jednoduché pre iné stratégie predvídať jej správanie

## Pravdepodobnostné rozhodovanie

TFT bola síce v deterministickom prostredí výhodnou stratégiou, ale uvažujme o prípade, kde máme 2 TFT stratégie a jedna je nútená v istom kroku zradiť (vo všeobecnosti je však ochotná spolupracovať). Ešte jasnejším príkladom je dvojica stratégií 001 a 101. Stratégií sa budú striedavo zrádzať a pokúšať sa spolupracovať, čím získavajú menej (spomeňme si na podmienku  $2CC > DC + CD$ ). Lepšie (pre obe stratégie) by bolo keby jedna z nich dokázala druhej odpustiť je chvíľkovú nelojálnosť a tak ich spoločne prinavrátila k dlhodobej spolupráci.

## Prvé výsledky a vylepšenia

Použili sme ten istý algoritmus ako v deterministickom prípade pre rôzne veľkosti populácie (100, 300, 500, 1000), pravdepodobnosti reprodukcie (0,2 0,3 0,5 0,7) a pravdepodobnosti mutácie (0,01 0,05 0,001 0,005). Tu naštudujeme na menší problém: ako vlastne sledovať homogenitu a ako interpretovať výsledky. Problém je v tom, že pravdepodobnosti sú reálne čísla a tým pádom skoro nenájdeme v populácii identických jedincov. Tento problém navrhujeme riešiť takto: V každej generácii zaokrúhlime aktuálne pravdepodobnosti v chromozóme, na nejaký počet desatinných miest (10, 100) a zisťujeme, či nejaká skupina stratégií po zaokrúhlení neprevahuje nad ostatnými.



Obrázok 4 Emergencia „pravdepodobnostného zloducha“ (každá modrá krivka reprezentuje skupinu stratégií s rovnakými chromozómami na 1 desatinné miesto – vyhrávajúca krivka reprezentuje stratégie ([0,8-0,9],[0,9-1,0],[0,9-1,0]))

Populácie v drvivej väčšine prípadov skonvergovali ku „pravdepodobnostným zloduchom“, teda k stratégiám z intervalov (0,80-0,99 0,95-0,99 0,97-0,99). To v niektorých simuláciách umožnilo vznik „pravdepodobnostných TFT“ (0-0,2 0-0,05 0,7-0,9). Teda populácie sa stali v podstate deterministickými a výsledky nám nepriniesli, žiadne nové poznatky. Skúsme sa však zamyslieť ako zabrániť takémuto javu a prinútiť evolúciu k emergencii nejakej menej deterministickej stratégie. Jednou z príčin spomenutého javu môže byť to, že len 10 „ťahov“ pre

8 nice

9 retaliatory

10 forgiving

11 clear

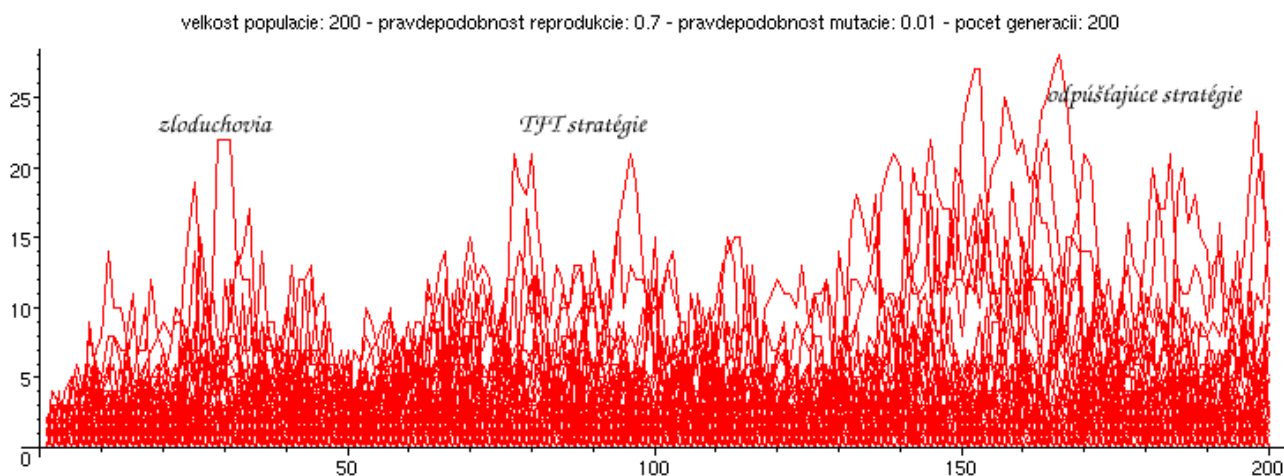
každý „súboj“ nedovolilo stochastičnosti tej-ktorej stratégii naplno sa prejavíť. Totiž v deterministickom prípade kvôli malému počtu stavov, v ktorých sa súperiace stratégie mohli nachádzať 10 ťahov dobre aproximovalo priemer z veľkého počtu ťahov. Úspešnosť stochastickej stratégie sa však môže prejavíť až po veľkom počte ťahov (súvisí to so zákonom veľkých čísel, známym z teórie pravdepodobnosti). Zvýšme teda v nasledujúcich simuláciách počet ťahov na 30 (50, 100). Mohli by sme však zároveň skúsiť ďalšie vylepšenie: ak sa nám nebude zdať stratégia dosť stochastická, jednoducho ju nezaradíme do nasledujúcej generácie. Určíme si niečo ako parameter stochastičnosti (napr. 0,03) a ak ruletou vyberieme stratégiu, ktorá bude mať aspoň jeden parameter v neprijateľnom rozmedzí (napr. menej 0,03 alebo viac ako 0,97) tak ju do reprodukcie nezaradíme a budeme vyberať znova.



Obrázok 5 Emergencia pravdepodobnostných TFT (od 130 generácie)

## Nové výsledky a diskusia

Po implementovaní nápadov do algoritmu sme vykonali ďalšiu sériu simulácií s rovnakými parametrami ako predtým, ale ešte aj s rôznymi počtami kôl (20, 30, 50, 100) a rôznymi parametrami stochastičnosti (0,01 0,03 0,05) sme dostali zaujímavejšie výsledky. Vo väčšine prípadov sme sa dostali k podobnému zloženiu populácie ako v prípadoch bez vylepšenia. Teda dominovali „111“<sup>12</sup> stratégie – alebo pravdepodobnostní zloduchovia, ktorí zdola aproximovali horné ohraničenia pre deterministickú zradu. Avšak našli sa aj zaujímavejšie simulácie v prípade, že sa objavili (na začiatku alebo v priebehu evolúcie krížením a mutáciou) nejaké „TFT“ stratégie. Zo začiatku tieto tvorili v populácii menšinu popri tom ako sa agresívne stratégie typu „11 1“ a „0 11“ nabaľovali na dobráckych, či hlúpych stratégiách ako „000“, „0 10“, „100“ a „110“. Avšak len dovtedy kým ich úplne nevykosili. V tejto chvíli emergovali recipročné „TFT“ stratégie ktoré sa rozmnožili a vytlačili tie agresívne. Doteraz sa situácia veľmi podobala deterministickému prípadu. Avšak teraz nastúpili odpúšťajúce stratégie,



Obrázok 6 Emergencia odpúšťajúcich stratégií.

Víťaziaci skupina stratégií na v 150-200 generácii má hodnoty [0.3-0.4 0.0-0.1 0.6-0.7]

12 deterministické stratégie v úvodzovkách v tomto odstavci znamenajú, pravdepodobnostné stratégie s pravdepodobnosťami blízky ich deterministickému profajskom

ktoré postupne nahradili „TFT“ stratégie. Tie postupne nahradili väčšinu stratégií a dlhodobo sa udržali v populácii. Vo väčšine prípadov potom vytesnili každú minoritnú stratégiu. (obr. 6).

Príklady ďalších evolúciou získaných stratégií uvádzame v tabuľke 6.

počet ťahov	populácia	p. reprodukcie	p. mutácie	emergovala <sup>13</sup>
30	200	0,7	0,01	[0,706-0,093-0,789]
30	500	0,3	0,01	[0,198-0,051-0,763]
20	100	0,5	0.005	[0.706-0.381-0.761]

Tabuľka 6 Príklady emergencie „odpúšťajúcich stratégií“

## Porovnanie deterministických a stochastických stratégií

Na porovnanie si vezmeme stratégie získané evolúciou v čisto deterministických prípadoch a v prípadoch s pravdepodobnosťami. Urobíme porovnanie pre rôzny počet ťahov 10 a 100. Keďže porovnáваме rôzne typy stratégií, z ktorých sa niektoré rozhodujú na základe pravdepodobnosti. Urobíme pre každú dvojicu 10000 simulácií a spočítame aritmetický priemer. Z deterministických stratégií budeme porovnávať dobráka (000), zloducha (111), TFT (001) a „ľakmer TFT“ (101). Zo stochastických si na porovnanie vezmeme pravdepodobnostného zloducha (0,944-0,938-0,935) a dobráka<sup>14</sup> (0,186-0,057-0,299) a jednu z emergovaných odpúšťajúcich stratégií (0,198-0,051-0,763).

	[000]	[001]	[101]	[111]
[0,186-0,057-0,299]	3,14 – 2,79 <sup>15</sup>	2,93 – 2,89	2,57 – 3,03	0,29 – 3,85
[0,198-0,051-0,763]	3,13 – 2,8	2,88 – 2,8	2,34 – 2,68	0,71 – 2,17
[0,944-0,938-0,935]	4,88 – 0,17	1,55 – 1,08	1,15 – 1,18	0,94 – 1,23

Tabuľka 7 Porovnanie pre 10 ťahov (10 000 opakovaní)

	[000]	[001]	[101]	[111]
[0,186-0,057-0,299]	3,17 – 2,82	2,92 – 2,92	2,89 – 2,93	0,3 – 3,81
[0,198-0,051-0,763]	3,1 – 2,84	2,84 – 2,8	2,73 – 2,77	0,76 – 1,97
[0,944-0,938-0,935]	4,88 – 0,19	1,23 – 1,18	1,19 – 1,19	0,93 – 1,26

Tabuľka 8 Porovnanie pre 100 ťahov (10 000 opakovaní)

Všimnime si napríklad vyznačené políčko v tabuľke 8, pre 100 ťahov dáva v priemere o 0,4 vyššiu platbu pre odpúšťajúcu stratégiu ako pre 10 ťahov (tab. 7, výhodne vlastnosti sa tu prejaví až po väčšom počte ťahov). V porovnaní s deterministickou TFT (tab. 5) je tu nárast o 0,61.

## Ďalšie možnosti

Nowak a Sigmund uvažovali dva typy turnajov, jeden veľmi podobný tomuto nášmu. Podobne ako Molander ukázali, že ak je pravdepodobnosť kooperácie, ak komplic v predchádzajúcom ťahu zradil, rovná  $q = \min\{1 - (T - R)/(R - S), (R - P)/(T - P)\}$  dosahuje stratégia najvyššie platby. Pre náš prípad je to  $\min\{1 - (5 - 3)/(3 - 0), (3 - 1)/(5 - 1)\} = 1/3$ . Teda v našom kódovaní je to chromozóm s poslednou zložkou rovnou 0,667. Toto číslo je veľmi blízke tým, ktoré sme získali evolúciou (napr. 0,761, v jednej zo simulácií vyšlo dokonca presne 0,667). Takáto stratégia sa

<sup>13</sup> v tabuľke je stratégia s najväčšou fitness v populácii, vyhlásenie, že emergovala sa zakladá na tom, že cca 25-30% populácie tvorili stratégie z tohto intervalu a ďalších cca 20% sa vyskytovalo v susedných intervaloch

<sup>14</sup> stratégie, ktoré emergovali v náhodne vybratej simulácii

<sup>15</sup> hodnota vľavo znamená platbu pre stochastickú stratégiu, vpravo pre deterministickú



nazýva GTFT (Generous TFT) . Nowak so Sigmund predpokladajú, že pokiaľ je TFT nevyhnutnou pre emergenciu kooperácie, stratégia, ktorá je biologicky plauzibilná je skôr GTFT.

Existuje však podobný typ turnaja, so širším spektrom stratégií, ktorý umožňuje rozhodovanie v ďalšom kroku, nie len na základe súperovho, ale aj vlastného ťahu. Stratégia môže byť teraz reprezentovaná ako  $S(p_1, p_2, p_3, p_4)$  kde  $p_1, p_2, p_3, p_4$  sú pravdepodobnosti po CC, CD, DC, a DD. Tu tiež emerguje GTFT stratégia, ale ešte viac (90%) takzvaná Pavlovovská stratégia, ktorá je blízka  $S(1, 0, 0, 1)$ , teda (väčšinou) spolupracuje po CC a DD a zrádza po CD a DC.

My sme ukázali, že síce pravdepodobnostné stratégie nedosahujú vyššie hodnoty v kontexte evolúcie s ich deterministickými proťajškami s malým počtom ťahov, vyhrávajú však v prostredí, kde nie je možné uplatniť determinizmus absolútne. V takomto prostredí však emergujú stratégie, ktoré dokážu odpúšťať.

## Prílohy

Prílohou k programu sú programy *inter* (porovnáva stratégie) a program *dilemma* (simuluje evolúcie pre deterministické aj stochastické stratégie s nastaviteľnou pravdepodobnosťou reprodukcie, mutácie a veľkosťou populácie a počtom generácií), oba v C++ (ANSI kompatibilné, využívajú triedu *prisoner*, súčasťou je aj Makefile). Ďalej skript pre *Maple*, pomocou ktorého sa dajú získať grafy (uvedené v práci) z výstupných súborov z programu dilemma. Parametre každého programu sa vypíšu po jeho spustení.

## Literatúra

- Martin Nowak, Karl Sigmund, "Tit for Tat in Heterogeneous Populations," Nature, 355 (16 January 1992): 250-253.
- Robert Axelrod, The Emergence of Cooperation Among Egoists, The American Political Science Review, 75 (1981): 306-318.
- Steven Kuhn: Prisoner's Dilemma, Stanford Encyclopedia of Philosophy 1997-2000
- Robert Axelrod: The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration, Princeton, New Jersey, Princeton University Press, 1997.
- Vladimír Kvasnička, Jíří Pospíchal, Peter Tiño: Evolučné algoritmy, Vydavateľstvo STU, Bratislava, 2000
- Melanie Mitchell: An introduction to genetic algorithms, The MIT Press, 1996
- C Wedekind and M Milinski: Human Cooperation in the Simultaneous and the Alternating Prisoner's Dilemma: Pavlov Versus Generous Tit-For-Tat, Proc Natl Acad Sci U S A. 1996 April 2; 93 (7): 2686-2689