

# Implementacija igre iterativna zapornikova dilema

Tomaž Piko  
Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko  
Koroška cesta 46  
2000 Maribor, Slovenija  
tomaz.piko@student.um.si

## POVZETEK

V tem delu smo se preizkusili v implementaciji igre iterativne zapornikove dileme. To je strateška igra v kateri igrata dva igralca. Cilj igre je premagati nasprotnika, in sicer v smislu, da pridobimo manj točk oziroma let zavora kot nasprotnik. Za igranje igre smo razvili tudi lastno strategijo, v treh različicah. Te različice smo preizkusili proti nekaterimi znanimi strategijami v iterativni zapornikovi dilemi in glede na pridobljene rezultate izmed njih izbrali najboljšo. Iste strategije vključno z našo lastno smo preizkusili v dveh turnirjih. Oba turnirja sta delovala po formatu krožnega dodeljevanja, kjer vsak igralec igra z vsakim. V drugem turnirju smo dodali možnost predhodne zaključitve igre z določeno verjetnostjo. Ugotovili smo, da naša strategija vrača relativno dobre rezultate proti ostalimi implementiranimi strategijami, vendar se rezultati razlikujejo od vrste turnirja. Prav tako rezultati zaradi naključnih elementov strategije niso konstantni, vendar so ti elementi ključni del naše strategije.

## Ključne besede

Igra, zapornikova dilema, strategija, turnir

## 1. UVOD

V strateški igri imamo končno množico igralcev, vsak od njih pa ima svojo strategijo, ki diktira njegove akcije. Igralčeve akcije so nagrajene oziroma kaznovane s funkcijo koristi, ki jo v večini primerov prikažemo z matriko izplačil. Zapornikova dilema je igra z neničelno vsoto, ki je večkrat analizirana v teoriji iger. Predstavlja nam zakaj se lahko dva posameznika odločita, da ne bosta sodelovala, čeprav se sprva zdi, da je to v njunem najboljšem interesu. V igri nastopata igralca A in B, oba sta zapornika, brez možnosti medsebojne komunikacije. Pravosodstvo ima dovolj dokazov le za lažjo obsodbo obeh, zato vsakemu ponudijo možnost dogovora, da izda drugega in se izogne zaporniški kazni. V kolikor ponudbe ne sprejme in nasprotnika ne izda pravimo, da igralec sodeluje, v nasprotnem pa izdaja. [5]

		Igralec B	
		Sodeluj	Izdaj
Igralec A	Sodeluj	(1, 1)	(5, 0)
	Izdaj	(0, 5)	(3, 3)

Tabela 1: Primer matrike izplačil

Tabela 1 prikazuje primer matrike izplačil in sicer prikazuje koliko let zavora bo moral odslužiti določeni igralec glede na njegovo in nasprotnikovo odločitev oziroma izvedeno akcijo. Tukaj lahko opazimo dilemo, ki se pojavi v igri. Po eni strani je za oba najbolje, da sodelujeta in odslužita krajšo kazen ampak nihče od njiju ne ve kaj naklepa drugi, po drugi strani pa z izdajo veliko pridobita. [1]

## 2. ITERATIVNA ZAPORNIKOVA DILEMA

Iterativna zapornikova dilema predstavlja razširjeno različico igre zapornikove dileme, kjer se klasična igra večkrat ponovi. Zapornika še vedno ne poznata nasprotnikove trenutne akcije ampak se lahko drug drugemu maščujeta za pretekle odločitve. Takšna igra omogoča več medsebojnega sodelovanja ter izkoriščanja nasprotnika, pri tem pa igrajo pomembno vlogo osebnosti igralcev. [3]

### 2.1 Turnirji Axelroda

V poznih sedemdesetih letih prejšnjega stoletja je ameriški politični znanstvenik in matematik Robert Axelrod organiziral turnir iterativne zapornike dileme. Povabil je več poznanih teoretikov igre, da prispevajo svoje strategije v obliki računalniških agentov. Posamezna igra je trajala 200 iteracij, matrika izplačil je bila enaka kot je prikazana v tabeli 1. Končni rezultat je predstavljal seštevek vseh let zavora, ki si jih je prislužil posamezni igralec tekom igre. V 200 iteracijah je navišji možni rezultat 1000, zmaga pa tisti, ki si je prislužil krajšo kazen. Turnir je potekal v obliki krožnega dodeljevanja, kjer vsak igralec igra z vsakim. Končni zmagovalec je bil tisti, ki je dosegel najboljši povprečni rezultat. [3] [4]

Na prvem turnirju je sodelovalo 15 strategij, zmagovalna strategija pa je bila TitForTat, ki jo bomo podrobneje spoznali kasneje. Axelrod je objavil rezultate prvega turnirja in organiziral še enega z enakimi pravili. Edina razlika je

bila, da se lahko vsaka igra zaključi na katerikoli iteraciji z verjetnostjo 0.00346%. Tokrat je na turnirju sodelovalo 63 strategij, zmagovalna pa je bila ponovno TitForTat. Axelrod je nekaj let pozneje organiziral še tretji turnir, ki je imel bolj evolucijski pristop in je imela vsaka strategija določeno populacijo igralcev. Zmagovalec pa je bil tudi takrat TitForTat. [3] [4]

Za uspešnost v njegovih turnirjih je Axelrod predlagal, da se strategije držijo naslednjih načel [1] [2] :

- Ne bodi preveč tekmovalen. Igro lahko zmagata oba tekmovalca.
- Začni prijazno in vzajemno sodeluj.
- Odpusti nasprotniku, v kolikor on začne sodelovati sodeluj tudi ti.
- Ustrezno se maščuj a ne pretiravaj.

### 3. IMPLEMENTACIJA IGRE

Lastno igro iterativne zapornikove dileme smo implementirali v programskem jeziku Python. Omogoča igro med dvema agentoma. Prav tako lahko spreminjamo pravila igre in sicer število iteracij, matriko izplačil ter verjetnost predhodne prekinitve igre po vsaki iteraciji. Prav tako nam omogoča, da odigramo turnir v katerega lahko vključimo poljubno število agentov in določimo skupna pravila za vse igre turnirja. Turnir se odigra po formatu krožnega dodeljevanja. Agentom v posameznih igrah ali turnirjih lahko določimo izbrano strategijo.

#### 3.1 Vključene strategije

Implementirali smo šest že uveljavljenih strategij v iterativni zaporniški dilemi in sicer:

- "RANDOM": Na vsaki iteraciji izbere naključno akcijo. Strategija je ekvivalentna metu kovanca.
- "ALL-C": Najbolj "prijazna" strategija, vedno sodeluje in nikoli ne izda nasprotnika.
- "ALL-D": Vedno izda nasprotnika. Strategija izkorišča prijaznejše strategije kot je "ALL-C".
- "TIT-FOR-TAT": Ime lahko prevedemo v "milo za drago". Na prvi iteraciji, ko še ne poznamo nasprotnikove predhodne poteze, začne s sodelovanjem. V kasnejših iteracijah pa ponovi nasprotnikovo prejšnjo potezo.
- "JOSS": Podobno kot "TIT-FOR-TAT", vendar v prvi iteraciji izdamo nasprotnika.
- "TESTER": V prvi iteraciji izdamo nasprotnika. Če se izkaže, da je nasprotnik maščevalen in nas v naslednji iteraciji izda, nadaljujemo s strategijo "TIT-FOR-TAT". V nasprotnem primeru ponavljamo zaporedje dvakrat sodelovanje in enkrat izdaja.

### 3.2 Lastna strategija

Razvili smo tudi lastno strategijo, ki smo jo potem preizkusili proti strategijam opisanimi v poglavju "Vključene strategije", strategijo smo poimenovali "Bipolar TIT-FOR-TAT". Je posebna različica zloglasne strategije "TIT-FOR-TAT", in sicer smo inspiracijo črpali iz bipolarne motnje človeških oseb, pri katerih prihaja do tako imenovanih nepričakovanih nihanj razpoloženja. Ideja je bila, da strategija taktično igra igro, vendar je sposobna nasprotnika zvesti z dokaj naključnimi nizi sodelovanj ali izdajanj. Delovanje strategije je sledeče:

- Strategiji predhodno določimo dve spremenljivki oziroma pragovni vrednosti  $G = \text{good}$  in  $B = \text{bad}$ , tako da velja:  $0 < B < G < 1$ .
- Na vsaki iteraciji igre tvorimo naključno realno število med 0 in 1, ki predstavlja temperament igralca, označimo ga s spremenljivko  $T$ .
- V kolikor velja:  $T < B$ , ima agent nihanje slabega razpoloženja in izda nasprotnika.
- V kolikor velja:  $T > G$ , ima agent nihanje dobrega razpoloženja in sodeluje.
- V kolikor  $T$  zavzema vrednost na intervalu  $[B, G]$ , je agent nevtralne volje in igra po principu milo za drago.

S spreminjanjem pragovnih vrednosti, lahko določimo ali bo naša strategija prijaznejša ali zahrbtnejša. Algoritem 1 prikazuje psevdokod naše strategije.

---

#### Algorithm 1 Bipolar TIT-FOR-TAT

---

```

1:  $B \leftarrow 0.3$ 
2:  $G \leftarrow 0.9$ 
3:  $T \leftarrow \text{Random}(0, 1)$ 
4: if  $T < B$  then
5:   return IZDAJ
6: else if  $T > G$  then
7:   return SODELUIJ
8: else
9:    $Iter \leftarrow \text{Trenutna\_Iteracija}$ 
10:  if  $Iter = 0$  then
11:    return SODELUIJ
12:  end if
13:   $X \leftarrow \text{Nasprotnikova\_Predhodna\_Poteza}$ 
14:  return  $X$ 
15: end if
```

---

### 4. REZULTATI

V nadaljevanju si bomo ogledali tri različice naše strategije "Bipolar TIT-FOR-TAT". V prvi različici prevladujejo nihanja slabega razpoloženja, v drugi nihanja dobrega razpoloženja, zadnja pa predstavlja uravnoteženo razmerje med nihanji. Najboljšo od treh različic bomo nato preizkusili še v različnih formatih Axelrodovega turnirja.

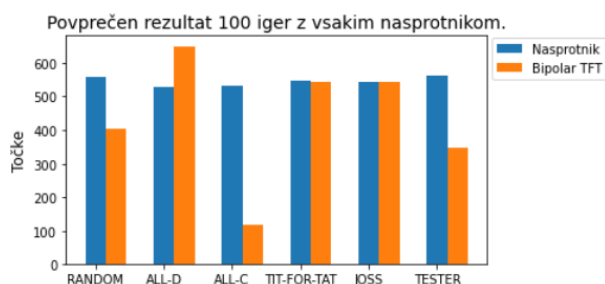
## 4.1 Analiza različic implementirane strategije

Tabela 2 prikazuje vrednosti, ki smo jih uporabili za naši pragovni spremenljivki B in G, v posamezni različici strategije. Vrednosti so izbrane tako, da v 60% iteracij strategija še vedno igra racionalno, preostalih 40% odločitev pa diktirajo nihanja razporeženja.

prevladujoča nihanja	B	G
slaba	0.3	0.9
dobra	0.1	0.7
uravnotežena	0.2	0.2

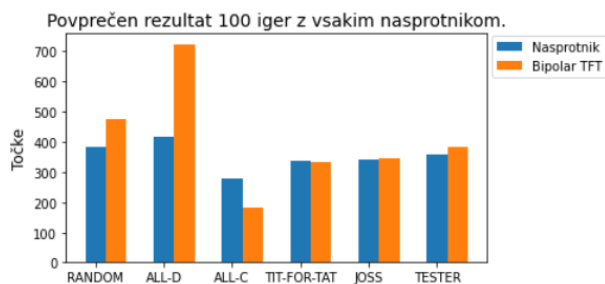
Tabela 2: Pragovi v posameznih različicah strategije

Vsako različico strategije smo preizkusili proti ostalimi strategijami, opisanimi v poglavju "Vključene strategije". Proti vsaki strategiji smo odigrali 100 iger, vsaka igra je trajala 200 iteracij brez možnosti predhodne zaključitve igre.



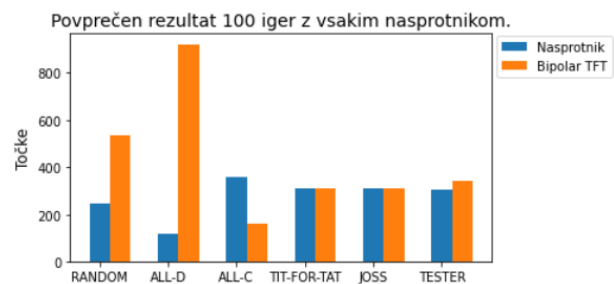
Slika 1: Prevladujoča nihanja slabega razporeženja

Na grafu na sliki 1 lahko vidimo rezultate prve različice naše strategije. Opazimo lahko, da je bila strategija izenačena s strategije "JOSS", sicer pa premagala vse razen "ALL-D", ki pridobi točke vsakokrat, kadar ima naša strategija nihanje dobrega razporeženja in se odloči za sodelovanje.



Slika 2: Prevladujoča nihanja dobrega razporeženja

Na grafu na sliki 2 opazimo, da različica, kjer prevladujejo nihanja dobrega razporeženja, vrača precej slabše rezultate. Uspelo nam je premagati le strategiji "ALL-C" in "TIT-FOR-TAT". Ostane nam še le zadnja, in sicer uravnotežena različica strategije, katere rezultate prikazuje graf na sliki 3. Opazimo lahko, da zadnja različica vrača najslabše rezultate, saj je bil uspešna proti eni sami strategiji.



Slika 3: Uravnotežena nihanja razporeženja

Čeprav imajo vse različice precej izenačene rezultate proti strategijama "JOSS" in "TIT-FOR-TAT" in je večkrat za zmago odločilna zadnja iteracija igre ter naključni element dodan v naši strategiji, iščemo vsesplošno učinkovito strategijo. Glede na pridobljene rezultate smo ugotovili, da je to različica naklonjena k nihanjam slabega razporeženja.

## 4.2 Rezultati prvotnega turnirja

V turnir s krožnim dodeljevanjem brez izločanja smo vključili vse strategije opisane v poglavju "Vključene strategije" in strategijo "Bipolar TIT-FOR-TAT". Vsaka igra turnirja se je igrala po pravilih prvotnega Axelrodovega turnirja. Uporabljena je bila matrika izplačil prikazana v tabeli 1. Zmagovalec turnirja je tista strategija, ki proti vsem ostalimi strategijami skupno zbere najmanj točk. Zaradi naključnih elementov v določenih strategijah smo turnir odigrali petkrat. Rezultate smo prikazali v obliki povprečnega števila točk, ki ga je zbrala posamezna strategija in povprečno uvrstitev tekom vseh petih odigranih turnirjev.

strategija	povprečno št. točk	povprečna uvrstitev
ALL-D	2253.0	1.0
Bipolar TFT	2620.6	2.4
ALL-C	2651.2	2.6
JOSS	2671.0	4.0
Tit-For-Tat	2688.2	5.0
Random	2726.8	5.8
Tester	2836.0	7.0

Tabela 3: Rezultati prvega turnirja

Rezultati prvotnega turnirja so prikazani v tabeli 3. V vseh petih iteracijah turnirja je prvo mesto dosegla strategija "ALL-D" in zbrala občutno manj točk kot ostale. Razlog za dominanco te strategije je izkoriščanje strategije "ALL-C", ki se nikoli ne maščuje. Prav tako je strategiji naklonjen tudi format turnirja, saj ji vedno zagotavlja igro s strategijo "ALL-C". Strategija "Bipolar TIT-FOR-TAT" je dosegla skupno drugo mesto po povprečnem številu doseženih točk vendar po povprečni uvrstitvi lahko vidimo, da drugo mesto ni bilo konstantno, kar je posledica naključnih elementov strategije.

### 4.3 Rezultati turnirja z verjetnosti zgodnje zaključitve igre

Enak eksperiment kot v prejšnjem podpoglavju smo ponovili, z izjemo, da smo turnirju dodali možnost, da se vsaka igra predhodno zaključi na katerikoli iteraciji z verjetnostjo 0.00346%.

strategija	povprečno št. točk	povprečna uvrstitev
Bipolar TFT	1506.8	3.2
All-D	1515.0	2.6
Tit-For-Tat	1606.8	2.8
Tester	1911.6	5.0
JOSS	1911.8	4.6
Random	1945.8	4.8
All-C	2067.6	4.8

Tabela 4: Rezultati drugega turnirja

Rezultati drugega turnirja so prikazani v tabeli 4. Opaziti je občutno nižje povprečno število doseženih točk, kot v tabeli 3, saj je nekaj iger trajalo manj iteracij. Prav tako lahko opazimo, da glede na povprečno uvrstitev nimamo točno določenega zmagovalca ampak so v ospredju tri strategije "Bipolar TIT-FOR-TAT", "ALL-D" in "TIT-FOR-TAT". Najnižje povprečno število točk je dosegla strategija "Bipolar TIT-FOR-TAT" zaradi njene sposobnosti, da hkrati sodeluje, se ustrezno maščuje in občasno izdati nasprotnika, ki je pripravljen sodelovati.

## 5. ZAKLJUČEK

V raziskavi smo implementirali igro iterativne zapornikove dileme. Razvili smo lastno strategijo in jo primerjali s šestimi drugimi strategijami, vključno s čisto naključno strategijo. Našo strategijo smo proti ostalimi sprva pomerili v množici medsebojnih iger nato še v dveh turnirjih s krožnim dodeljevanjem, kjer je vsaka strategija igrala proti vsemi ostalimi. Naša strategija, ki smo jo poimenovali "Bipolar TIT-FOR-TAT" je zasnovana tako, da vsebuje dva parametra s katerima spreminjamo njeno obnašanje. Izkazalo se je, da najboljše rezultate dosega s takšnimi vrednostmi parametrov, da strategija večkrat nepričakovano izda nasprotnika. V igrah proti posameznimi strategijami se je izkazala za zelo prilagodljivo strategijo, ki premaga večino nasprotnikov ali pa je razlika v točkah zelo majhna. Zaradi te lastnosti se je dobro izkazala tudi v turnirski selekciji, kjer je zaradi pomanjkanja dominantnega nasprotnika pridobila zelo majhno število točk.

V kolikor bi dovolili sprotno prilagajanje parametrov strategije menimo, da bi lahko naša strategija dosegala še boljše ter seveda bolj konstantne rezultate. Potrebno je tudi dodati, da je v naših preizkusih uporabljenih relativno nizko število strategij in bi se z dodajanjem strategij rezultati lahko spremenili. Prav tako se je izkazalo, da je format naših turnirjev zelo naklonjen naši strategiji. Za igro iterativne zapornikove dileme je bilo razvitih je bilo že mnogo različnih strategij ampak univerzalna strategija, ki bi nadvladovala vsem ostalim ne obstaja. Rezultati turnirjev pa so odvisni od strategij, ki v njem sodelujejo in načina izvajanja turnirja.

## 6. REFERENCES

- [1] R. Axelrod and W. D. Hamilton. Science: The evolution of cooperation. pages 1390–1396, 03. 1981.
- [2] M. Q. Azhar. The success of tit for tat in computer tournaments, 2005.
- [3] A. L. C. Carvalho, H. P. Rocka, F. T. Amaral, and F. G. Guimaraes. Iterated prisoner's dilemma - an extended analysis. 03. 2016.
- [4] A. Kaznatcheev. Short history of iterated prisoner's dilemma tournaments, 03. 2015.
- [5] Wikipedia. Prisoner's dilemma.