

# EVOLUCIJA SURADNJE U NOWAK-SIGMUNDOVOM MODELU ITERIRANE ZATVORENIKOVE DILEME

## EVOLUTION OF COOPERATION IN NOWEG-SIGMUNDS ITERATED PRISONER'S DILEMMA

*Jurica Hižak*

Stručni članak

**Sažetak:** Ovaj rad opisuje metodu i rezultate istraživanja koje su Martin Nowak i Karl Sigmund objavili 1992.g. u časopisu *Nature* u članku pod naslovom „Tit for tat in heterogeneous populations“. Autori su kroz računalno simulirani turnir Iterirane zatvorenikove dileme pokazali da je strategija „Milo za drago“ (eng. „Tit for Tat“, TFT) ključna za pojavu kooperativnosti u heterogenoj skupini igrača. Njihova simulacija pokazuje da prvobitno miješana populacija igrača evoluirala do istrebljenja gotovo svih strategija osim odmetničke, a zatim dolazi do naglog porasta TFT igrača. Uz to, njihova simulacija pokazuje da TFT zapravo i nije kraj evolucije, nego početak prevlasti velikodušnih strategija. Prema riječima autora, TFT je pivot (osovina ili os zaokreta), a ne završna točka u evoluciji prema suradnji.

**Ključne riječi:** evolucija suradnje, iterirana zatvorenikova dilema, strategija Tit for Tat

Professional paper

**Abstract:** This paper describes the method and the results of a research that Martin Nowak and Karl Sigmund published 1992 in *Nature* in an article entitled "Tit for tat in heterogeneous populations." Through computer-simulated tournament of iterated prisoner's dilemma, the authors showed that the strategy "Tit for Tat" (TFT) is essential for the emergence of cooperation in a heterogeneous group of players. Their simulation shows that initially mixed population of players evolves to the extinction of almost all strategies except renegades, and then there is a sudden increase in TFT players. Moreover, the simulation shows that TFT is not really the end of evolution, but the beginning of the predominance of a more generous strategy. According to the authors, TFT is a pivot, rather than the end point of an evolution towards cooperation.

**Key words:** evolution of cooperation, iterated prisoner's dilemma, Tit for tat strategy

### 1. UVOD

Zatvorenikova dilema (eng. Prisoner's dilemma, PD) poznata je igra u kojoj dva igrača neovisno jedan o drugom moraju donijeti odluku: surađivati (C) ili izdati suradnika (D). Ukoliko donesu različite odluke, jedan od njih (izdajnik) maksimalno profitira uz isplatu T, dok drugi igrač (suradnik) završava s najmanjom isplatom S. Obostrana suradnja donosi korist obojici, dakle svaki igrač dobiva isplatu R. Međutim, kako nemaju informacije o ponašanju svog suigrača, oni moraju odvagati sve opcije koje se slikovito mogu prikazati tabelom 1.

**Tabela 1.** Matrica isplate u igri Zatvorenikova dilema. Kako bi se stvorila dilema nužno je namjestiti relativne odnose između isplata, Sucker (S) < Punishment (P) < Reward (R) < Temptation (T).

	C	D
Prvi igrač	C R, R	D S, T
	D T, S	P, P

Tabela nam pokazuje: ako je prvi igrač spreman za suradnju, te izabere prvi redak (C kao „cooperate“), dovodi se u opasnost da umjesto u (R,R) završi u (S,T). Stoga on nužno dolazi do zaključka da je najsigurnije izdati suigrača, odnosno izabrati drugi redak (D kao „defect“) jer je  $T > R$ , a  $P > S$ . Potpuno isti „rezon“ drugog igrača navodi da on također izabere izdaju (D), što znači da igra završava obostranom izdajom te isplatom (P,P). Zanimljivost igre, dakle, leži u činjenici da je pametnije igrati „izdaju“ umjesto „suradnje“, iako suradnja zapravo donosi povoljniji ishod za obojicu.

Kad se igra stavi u kontekst biologije, kako bi se razmatralo ponašanje životinja, tada veća isplata (bez obzira da li je riječ o hrani, vodi ili životnom prostoru) predstavlja veću podobnost (eng. fitness) tj. veću vjerojatnost ostavljanja potomstva. Na prvi pogled, evolucija se protivi suradnji, jer suradnja umanjuje isplatu, a time i podobnost, međutim, kad se igračima dozvoli da opetovano igraju, pomalo neočekivano otvaraju se vrata prema suradnji [1]. Takva opetovana igra naziva se Iterirana zatvorenikova dilema (eng. Iterated prisoner's dilemma, IPD) i, prema Nowakovim

riječima, danas se smatra paradigmatom za evoluciju suradnje između sebičnih jedinki [2].

## 2. ITERIRANA ZATVORENIKOVA DILEMA

Budući da se igrači u IPD više puta sreću, omogućeno im je da korigiraju svoje odluke zavisno od prethodno odigrane runde. Primjerice, igrač može kazniti partnera koji ga je u prethodnoj rundi izdao. Da bi otkrio koja je strategija igranja najbolja, poznati politolog Robert Axelrod je, krajem 70-tih, organizirao turnir po principu svaki sa svakim (round-robin) i tada je, na sveopće iznenađenje, pobijedila strategija "Milo-za-drago" koja u prvom koraku bira suradnju, a zatim kopira suparničke poteze. Iako sama po sebi nije najjača, pokazala se najisplativijom jer je uzela mnogo bodova od suradničkih strategija, a nije izgubila mnogo od izdajničkih strategija (budući da ne dopušta izrabljivanje). Spoznaja da suradnja može evoluirati u populaciji sebičnih jedinki, izazvala je veliko uzbuđenje u znanstvenoj zajednici, što je rezultiralo mnoštvom IPD turnira i matematičkih analiza.

**Tabela 2.** Tok igre za deset rundi između četiri najpoznatije strategije ALLD, ALLC, TFT i GRIM.

													A
ALLD	D	D	D	D	D	D	D	D	D	D	D	10T	50
ALLC	C	C	C	C	C	C	C	C	C	C	C	10S	0
ALLD	D	D	D	D	D	D	D	D	D	D	D	1·T+9P	14
TFT	C	D	D	D	D	D	D	D	D	D	D	1·S+9P	9
ALLD	D	D	D	D	D	D	D	D	D	D	D	1·T+9P	14
GRIM	C	D	D	D	D	D	D	D	D	D	D	1·S+9P	9
ALLC	C	C	C	C	C	C	C	C	C	C	C	10R	30
TFT	C	C	C	C	C	C	C	C	C	C	C	10R	30
ALLC	C	C	C	C	C	C	C	C	C	C	C	10R	30
GRIM	C	C	C	C	C	C	C	C	C	C	C	10R	30
TFT	C	C	C	C	C	C	C	C	C	C	C	10R	30
GRIM	C	C	C	C	C	C	C	C	C	C	C	10R	30

Treba nešto reći i o problemu konačnog odnosno beskonačnog broja iteracija. U prvom Axelrodovom turniru igrao se fiksni broj rundi ( $n=200$ ) [1]. U kasnijoj analizi pokazalo se da bi TFT izgubila turnir od nekih podmuklijih strategija. Primjerice, čista strategija TFT može biti poražena od strategije koja igra TFT do predzadnje runde ( $n-1$ ), a u  $n$ -toj rundi igra izdaju. Takva strategija bi pobijedila na turniru. Da bi izbjegao takve smalice, na sljedećem turniru Axelrod je umjesto fiksnog broja partija, uveo vjerojatnost odigravanja sljedeće partije  $\omega < 1$ . U tom slučaju prosječna isplata TFT između ALLC i TFT nije  $A(\text{ALLC}, \text{TFT}) = nR$  kao u tabeli 2, nego je dana kao suma geometrijskog niza, pa uz beskonačni broj iznosi:

$$A(\text{ALLC}, \text{TFT}) = R + \omega R + \omega^2 R + \dots = \frac{R}{1-\omega} \quad (1)$$

Nowak i Sigmund ističu da evolucijsku ulogu TFT-a, u suštini, možemo razmatrati na dva načina [2]. Prvo, možemo se pitati hoće li se homogena TFT skupina oduprijeti napadu *mutantskih* strategija, te zatim analitički ispitivati uvjete u kojima TFT jest ili nije evolucijski stabilna strategija. (Npr. usporedbom prosječnih isplata TFT u duelu s ALLD i dr. može se zaključiti da je TFT stabilna strategija ako je  $\omega > 2/3$ ). Naime, TFT je evolucijski stabilna strategija samo ako postoji dovoljno velika vjerojatnost ponovne interakcije kako bi se osiguralo da isplata  $A(\text{TFT}, \text{TFT}) = \frac{R}{1-\omega}$  sigurno bude veća od isplate  $A(\text{ALLD}, \text{TFT}) = T + \frac{P\omega}{1-\omega}$  i zarade tzv. alternirajuće strategije u duelu s "milo-za-drago":  $A(\text{ALT}, \text{TFT}) = \frac{T+S\omega}{1-\omega^2}$  [3]. Iz navedenog uvjeta slijedi:

$$\frac{R}{1-\omega} > \max\left(T + \frac{P\omega}{1-\omega}, \frac{T+S\omega}{1-\omega^2}\right) \quad (2)$$

što je ispunjeno ako i samo ako:

$$\omega > \max\left(\frac{T-R}{T-P}, \frac{T-R}{R-S}\right) \quad (3)$$

Uz parametre korištene na Axelrodovom turniru ( $S=0, P=1, R=3, T=5$ ) slijedi:

$$\omega > \max\left(\frac{2}{4}, \frac{2}{3}\right) \quad (4)$$

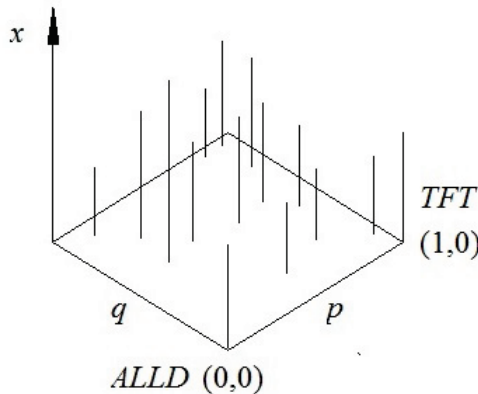
Drugi pristup je ono što autori članka nazivaju „the heteromorphic approach“; kreće se od široko raspršene distribucije strategija i zatim ih se prepušta „prirodnoj“ selekciji. To je bilo učinjeno na Axelrodovom „ekološkom“ turniru gdje su se 63 strategije borile, generaciju za generacijom, svaki put s frekvencijom proporcionalno zaradi prethodne generacije. Međutim, Nowak ističe: „Sudionici tog turnira dobrim su dijelom bili teoretičari igara, koji teško da mogu biti reprezentativni primjeri strategija u eko-sustavu“ [2].

## 3. NOWAK-SIGMUNDOW MODEL

Kako bi dobili na realističnosti i kako bi se što više približili stvarnim situacijama iz prirode, Nowak i Sigmund su uveli određenu nesigurnost u povlačenju poteza svojih igrača. Taj pristup Nowak komentira: „To bi trebalo oponašati biološke interakcije. Igrači mogu pogrešno interpretirati tuđi potez (ili identitet) ili nespretno sprovesti vlastite intencije“. Dakle, za razliku od autora stvarnih i simuliranih IPD koje su provedene prije ove studije, Nowak i Sigmund svoje igrače nisu determinirali nepromjenjivim strateškim algoritmima, nego su svakoj strategiji pridružili vjerojatnosti suradnje nakon protivničke suradnje, odnosno izdaje. Reaktivna strategija pojedinog igrača zadana je uređenom trojkom  $(y, p, q)$ , gdje je  $y$  vjerojatnost suradnje u prvoj rundi, a  $p$  i  $q$  su uvjetne vjerojatnosti za suradnju nakon protivničke suradnje, odnosno nakon izdaje. Dakle, TFT je

predstavljena s (1,1,0), ALLD (0,0,0), ALLC (1,1,1), sumnjičavi tit-for-tat STFT (0,1,0) itd.

Radi jednostavnosti, autori su razmatrali beskonačnu IPD (krajnji slučaj,  $\omega=1$ ). Ako zanemarimo prvi element uređene trojke (koji je ionako nevažan kad se igra na beskonačan broj iteracija), svaka strategija može se predstaviti naprosto kao točka (p,q) u jediničnom kvadratu. Treća dimenzija, autorima članka kasnije će poslužiti za grafički prikaz broja igrača (odnosno frekvencije). Izabравši uzorak reaktivnih strategija  $E_1$  do  $E_{100}$  uniformno distribuiranih po površini jediničnog kvadrata, s jednakom početnom frekvencijom, Nowak i Sigmund pustili su sustav da evoluiru.



Slika 1. Frekvencija igrača u nekom trenutku igre.

Na jediničnom kvadratu nalaze se strategije zadane vjerojatnošću suradnje p nakon suradnje protivničkog igrača, odnosno vjerojatnošću suradnje q nakon izdaje. Visina linije prikazuje frekvenciju igrača koji koriste strategiju (p,q). Strategije koje su prvobitno bile uniformno raspoređene, postupno odumiru ili napreduju zavisno od ostvarenog rezultata.

Frekvencija svake sljedeće generacije bila je zadana jednadžbom:

$$x'_i = x_i \frac{f_i}{\bar{f}} \quad (5)$$

gdje je  $f_i$  prosječna isplata (fitness) strategije  $E_i$  u populaciji, a  $\bar{f} = \sum x_k f_k$  prosječni fitness cijele populacije. Treba primijetiti da se ovdje radi o tzv. diskretnoj replikatorskoj dinamici izvedenoj iz pretpostavke da su generacije nepreklapajuće (svi odrasli umiru prije nove generacije, slično insektima) i da je reprodukcija aseksualna. Pretpostavlja se da svaka jedinka koja koristi i-tu strategiju, generira  $f_i$  vlastitih kopija, tako da je u sljedećoj generaciji  $N_i(t+1) = N_i(t)f_i$  [4]. Iz danih pretpostavki slijedi:

$$x'_i \equiv x_i(t+1) = \frac{N_i(t+1)}{N(t+1)} \quad (6)$$

$$x'_i = \frac{N_i(t)f_i}{\sum_k N_k(t)f_k} = \frac{N_i(t)}{N} \frac{f_i}{\sum_k x_k(t)f_k} \quad (7)$$

$$x'_i = x_i(t) \frac{f_i}{\bar{f}} \quad (8)$$

Isplata  $A(E_i, E_j)$  u duelu i-te i j-te strategije dana je relacijom[5]:

$$A(E_i, E_j) = 1 + 4c' - c - cc' \quad (9)$$

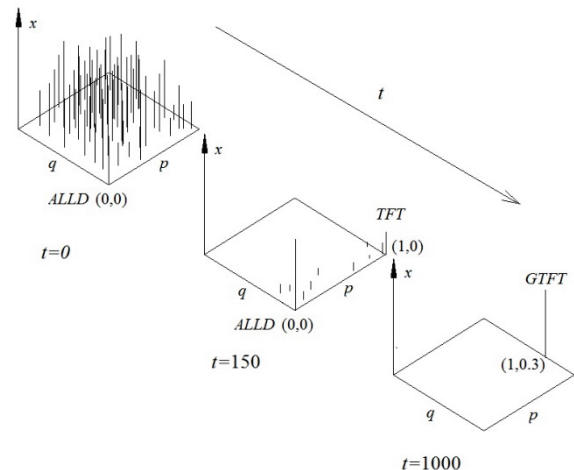
gdje su koeficijenti

$$c \equiv c_{ij} = \frac{q_i + (p_i - q_i)q_j}{1 - (p_i - q_i)(p_j - q_j)} \quad (10)$$

$$c'_{ij} = c_{ji}. \quad (11)$$

## 4. REZULTATI

Kad se simulacija turnira pusti u „pogon“, strategijama koje su blizu (0,0) frekvencija raste, dok ostale strategije vrlo brzo nestaju. Drugim riječima, preostaje praktički samo ALLD. Naime, daleko najveći udio strategija u uzorku ima visoke q-vrijednosti i ne osvećuje se izrabljivačima. Nowak slikovito kaže: „Kraj bogate zalihe naivaca, isplati se biti izdajnik.“



Slika 2. Evolucija suradnje u heterogenoj populaciji.

Prvih 100 generacija, strategije bliske ALLD dominiraju populacijom. Nakon 150 generacija, TFT se revitalizira i uspostavlja suradnju u populaciji, što omogućava oprastajućim strategijama da rastu. Konačni pobjednik je GTFT (0.99,0.3).

Ipak, ishod se dramatično mijenja negdje oko 150-te generacije, ako je jedna od početnih strategija (dodanih ručno ili slučajno) bliska strategiji TFT. Autori kažu: „Preokret se događa u momentu kad su naivčine toliko desetkovane da se izrabljivači više nemaju čime hraniti. Isprva pomalo, a zatim uzimajući sve jači zamah, osvetnici se vraćaju i sada izrabljivači iščezavaju“ [2]. Interesantno, strategija TFT i njoj slične strategije (koje su izazvale ovaj preokret) same neće profitirati: uklonivši izrabljivače, bivaju smijenjene od strane strategije koja je najbližnja GTFT (tzv. velikodušna strategija koja oprašta oko 30% izdajničkih poteza). Evolucija tada prestaje.

Čini se da u evoluciji suradnje strategija TFT ima ulogu policije, koja beskompromisno kažnjava odmetnike i na taj način „čisti“ sustav od izrabljivača. Međutim, kad je uspostavljena suradnja u sustavu, tada je

bolje imati određeni postotak opraštajućih poteza jer svaka i najmanja pogreška (npr. komunikacijski šum) kod igrača TFT uzrokuje spiralu osвете.

## 5. LITERATURA

- [1] Axelrod, R.: The Evolution of Cooperation, 1984.
- [2] Nowak, M.; Sigmund, K.: Tit for tat in heterogeneous populations, Nature, 1992.
- [3] Broom, M.; Riechert, J.: Game-Theoretical Models in Biology, 2013.
- [4] Bishop, D.; Cannings, C.: A generalized war of attrition, Journal of Theoretical Biology, 1976.
- [5] Nowak, M.: Stochastic strategies in the Prisoner's Dilemma, Theoretical Population Biology 38, (1990) 93-112

### Kontakt autora:

**Mr.sc. Jurica Hižak**  
Sveučilište Sjever  
104. brigade 3, 42000 Varaždin  
jurica.hizak@unin.hr