Istraživanje pozitivne i negativne recipročnosti u evolucionarnim igrama

Petar Fumić, PMF-Fizički odsjek, Bijenička c. 32, 10 000 Zagreb

28. studenog 2016.

Sažetak

Rad je reprodukcija članka Attila Szolnokija i Matjaža Perca[1]. Cilj je bio ispitati evolucionarnu prednost strategije koja u sebi kombinira pozitivnu i negativnu recipročnost. Nalazimo da kombinacija može postojati u vrlo uskom spektru vrijednosti te da ima veće šanse za egzistenciju ako su uvjeti evolucije teži. Također smo našli prisutnost nekih zanimljivih ponašanja sustava poput ciklične dominacije.

1. Uvod

Kooperacija je proces u kojem grupa individua djeluje zajedno i iz toga dobiva uzajamnu korist. Ljudi su još od prapovijesti doveli kooperaciju na prilično visoku razinu te ona čini temelj svake civilizacije. Od prvih civilizacija plodnog polumjeseca i Hamurabijevog zakona pa do današnjeg dana temelj društva i društvenog života čini sankcioniranje nepoželjnog ponašanja (poput pljačke ili ubojstva). Također, u mnogim uspješnim društvima možemo naći i primjere nagrađivanja poželjnog ponašanja (promocije, dobivanje društvenog ugleda). Recipročnost se smatra bitnim mehanizmom propagacije kooperativnog ponašanja u društvu (nagrađivanje pozitivnog i kažnjavanje negativnog ponašanja). Prema teoriji jake recipročnosti pozitivna i negativna recipročnost su korelirane kako bi dale optimalnu strategiju [2]. Ispitujemo tu tvrdnju pomoću igre javnog dobra na mreži uz jedno defektorsko i tri kooperativna ponašanja (strategije) od kojih jedno uključuje koreliranu pozitivnu i negativnu recipročnost. Gledamo u koju će fazu sustav evoluirati nakon relaksacijskog vremena, ovisno o parametrima sustava, te razmatramo za koje je vrijednosti tih parametara korelirana strategija uspiješna.

2. Igra javnog dobra na mreži

Sustav u kojemu ćemo modelirati naše ideje je tzv. public goods game (igra javnog dobra)[3] . Ta igra je standardni alat eksperimentalne ekonomije. U svojoj osnovnoj formi svi igrači ulažu određenu svotu novca u javni fond. Vrijednost novca u fondu se množi s faktorom (većim od jedan a manjim od ukupnog broja igrača, u daljnjem tekstu faktor sinergije, r) te biva ravnomjerno podijeljena među svim igračima. Naša verzija igre će biti na kvadratnoj mreži dimenzija $L \times L$ i igrati će ju ju L^2 igrača. Igrači su raspoređeni u preklapajuće grupe veličine $G = 5$ tako da je svatko povezan sa svoja $G-1$ najbliža susjeda. Svaki igrači pripada u G grupa. Kvadratna rešetka je najjednostavnija rešetka na kojoj možemo uzeti u obzir činjenicu da su interakcije među ljudima strukturirane. Postoje mnoštvo dokaza da je za igre s grupnim interakcijama proučavanje kvadratne mreže dovoljno kako bi otkrili sve moguće evolucionarne ishode i ti ishodi su u pravilu neovisni o strukturi interakcije [4].

Svaki igrač će imati zadanu "strategiju" koja može biti jedna od četiri moguće, defektiranje ($s_x = D$, označava igrača koji ne doprinosi ništa u javni fond), kooperacija koja kažnjava defektiranje ($s_x = P$), kooperacija koja nagrađaju druge kooperatore ($s_x = R$) te kooperacija koji nagrađuje i kažnjava ($s_x = B$). Sve tri kooperativne strategije doprinose zadani iznos (u ovom slučaju 1) u javni fond. Suma svih doprinosa u svakoj grupi se množi s faktorom sinergije r i dijeli među igračima. Defektor ima dodatni faktor $-\beta(G-1)$ za svakog susjednog kooperanta koji kažnjava. Kooperator koji kažnjava ima faktor $-\gamma/(G-1)$ za svakog defektora u grupi što predstavlja cijenu koju on mora platiti kako bi mogao kazniti defektora. Svaki kooperator dobiva nagradu $\beta/(G-1)$ od R i B strategija u grupi a svaka od tih taktika mora platiti cijenu nagrađivanja $-\gamma/(G-1)$. Koristimo jednako jake faktore kazne i nagrade za istu cijenu, tj. isti par (β, γ) vrijednosti za nagradu i kaznu. Vrijednosti isplate za četiri sukobljene strategije za svaku grupu su dane sa:

$\Pi_D^g = r(N_P + N_R + N_B)/G - \beta(N_P + N_B)/(G-1)$,

$\Pi_P^g = r(N_P + N_R + N_B + 1)/G - 1 - \gamma(N_D)/(G-1) + \beta(N_R + N_B)/(G-1)$,

$\Pi_R^g = r(N_P + N_R + N_B + 1)/G - 1 - \gamma(N_P + N_R + N_B)/(G-1) + \beta(N_R + N_B)/(G-1)$,

$\Pi_B^g = r(N_P + N_R + N_B + 1)/G - 1 - \gamma + \beta(N_R + N_B)/(G-1)$,

gdje N_{sx} predstavlja broj drugih igrača sa strategijom s_x u grupi.

Radimo Monte Carlo simulacije igre javnog dobra s sljedećim elementarnim koracima. Nasumični igrač x igra igru javnog dobra kao član svih grupa kojima pripada i njegova ukupna isplata Π_{sx} je suma svih isplata Π_{sx}^g skupljenih u pojedinačnim grupama. U sljedećem koraku se nasumično bira jedan od igračevih prvih susjeda y i odabrani suigrač računa svoju isplatu Π_{sy} na isti način. Konačno, igrač x pokušava prenijeti svoju strategiju na igrača y sa vjerojatnoti danom Fermijevom funkcijom

$\omega(s_x \to s_y) = 1/\{1 + exp[(\Pi_{sy} - \Pi_{sx})/K]\}$, gdje je $K = 0.5$. Strategije uspješnijih igrača su spremnije prihvaćene premda nije nemoguće ni prenijeti strategiju manje uspješnog igrača. Svaki Monte Carlo korak se sastoji od L^2 nasumično izabranih igrača koji pokušaju prenijeti svoju strategiju na susjede.

Nakon dovoljne količine Monte Carlo koraka računaju se frakcije defektora (ρ_D), kooperatora koji kažnjavaju (ρ_P), kooperatora koji nagrađuju (ρ_P) i kooperatora koji nagrađuju i kažnjavaju (ρ_B). Veličinu sustava je $L = 200$ a broj Monte Carlo koraka je 10^4. Na kraju se radi fazni dijagram ovisno o parametrima (β, γ).

3. O numerici

Računanje faznih dijagrama ovakvih sustava može biti prilično numerički zahtijevno. Naime, za svaku točku faznog dijagrama treba provesti 10^4 MCS (Monte Carlo step). Vrijeme izračuna jednog MCS raste kvadratično s dimenzijama sustava ($t^2 \sim L$). Te već za $L = 200$ postoji $4 \cdot 10^4$ točaka za koje treba napraviti proračun u svakom MCS. Prvo sam radio kalkulacije na CPU no vidjevši da je za jednu točku potrebno$\approx 40min$ uvidio sam da bi za izračun cijelog faznog dijagrama bilo potrebno vrlo puno vremena. S obzirom da imam relativno pristojnu grafičku kartice odlučio sam se za metodu kalkulacija na GPU-u pomoću OpenCL-a. Inicijalno razvijan od strane Apple-a, OpenCL je postao kolaboracija između timova iz AMD-a, IBM-a, Intel-a i Nvidi-e koji su verziju 1.0 izdali 21. listopada 2009.-e.

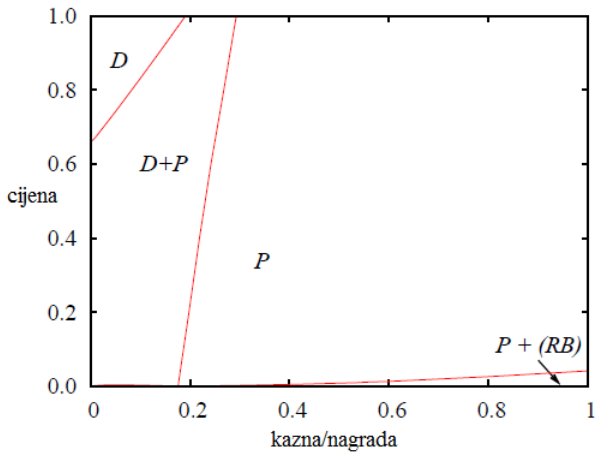
GPU ima arhitekturu koja je optimalna za izračun paralelnih problema (probleme koje možemo podijeliti na više neovisnih dijelova, poput procesiranja slika, shadere, etc.). GPU-i rade na nižim frekvencijama no imaju mnogo veći broj jezgri od klasičnih CPU-ova. Dobivanje faznog dijagrama možemo lako paralelizirati tako da svakoj jezgri GPU-a dodijelimo jednu točku faznog dijagrama. GPU treba $\approx 2h30min$ da izračuna 255 točaka faznog dijagrama što je velika vremenska ušteda u usporedbi s CPU-om.

Općenito govoreći, ako je problem prikladan za GPU vrijeme izračuna može bit drastično kraće nego da isti problem vrtimo na CPU-u istog cijenovnog ranga. GPU Computing je obećavajuće područje i može imati široku primjenu u fizici kao potencijalno vrlo snažan numerički alat. Jedina mana ovakvog pristupa je što nije dobar za sekvencijalne probleme (probleme u kojima svaki sljedeći korak ovisi o prethodnima).

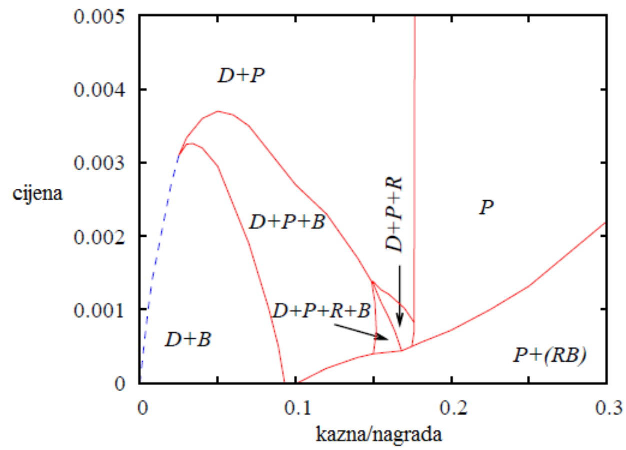
OpenCL standard ne dolazi s implementiranim generatorom nasumičnih brojeva. U ovom projektu koristim Linear Congruential Generator (linearni sukladni generator) koji se bazira na Java Random funkciji. U osnovi generator je definiran s rekurzivnom funkcijom:

$X_{n+1} = (aX_n + c)\ mod\ m$

gdje su



Slika 1: $\beta - \gamma$ fazni dijagram za vrijednosti $r = 4.5$



Slika 2: $\beta - \gamma$ fazni dijagram za vrijednosti $r = 4.5$, uvećano, plavo su diskontinuirani fazni prijelazi

$0 < a < m$,$0 \leq c < m$,$0 \leq X_0 < m$. Eksponencijalna funkcija $exp(x)$ je Taylorov razvoj do desetog člana s cutoffom na 0 za vrijednosti x manje od -4.

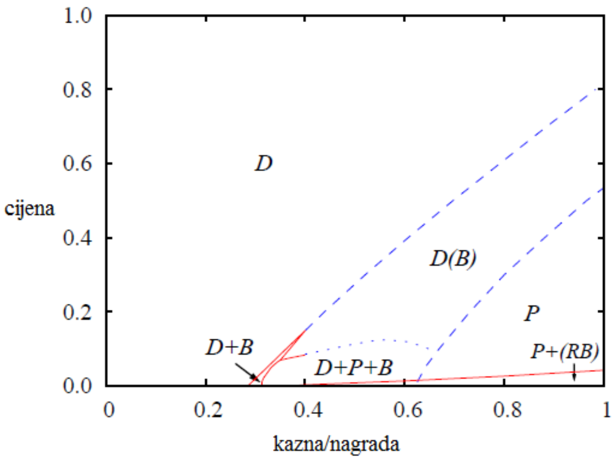
Izračuni su rađeni na kartici Radeon HD 7870, s veličinom globalne memorije od 2GB i 255 jezgara koje rade na 860 MHz.

4. Rezultati

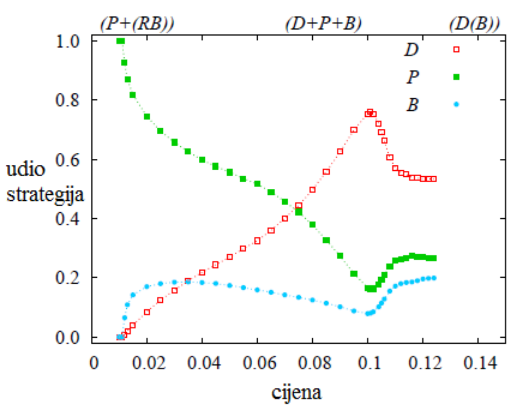
Fazne dijagrame smo reproducirali za vrijednosti faktora sinergjie $r = 4.5$ i $r = 2.5$. Ako promatramo sustav čistih defektora i kooperatora bez mehanizama kazne i nagrade kooperatori preživljavaju za $r > 3.74$ a u potpunosti pobjeđuju defektore za $r > 5.49$. Kada je vrijednost faktora sinergije $r = 4.5$ imamo dobre uvjete za evoluciju kooperacije dok je za $r = 2.5$ sustav nemože preživjeti bez kazna i nagrada što daje izazov preživljavanju kooperativnih strategija.

A. Faktor sinergije $r = 4.5$

Iz prikazanog faznog dijagrama možemo vidjeti da je dominatna faza kooperatora. Najefikasnija evolucionarna strategija je za većinu parova (β, γ) kažnjavanje defektora ili P. Kombinacija tri kooperativne taktike (P, R, B) isprva nadigra defektorsku taktiku (D) no taktike koje nagrađuju druge kooperatore imaju manju isplatu od čiste P taktike i ta taktika nadigra ostale. Čista defektorska faza postoji samo za niske vrijednosti udijeljene kazne (fine) i velike cijene nametanja kazne (cost). Povećavajući kaznu i spuštajući cijenu nametanja dolazimo do domene u kojoj D iP strategije mogu koegzistirati. Kako variramo (β, γ) prema čistoj P fazi možemo vidjeti da udio defektora opada prema nuli a udio kooperatora raste. Ako je cijena nagrađivanja vrlo niska a nagrada visoka strategije R i B koegzistiraju uz strategiju P u vrlo malenoj domeni faznog dijagrama. Za vrlo malene vrijednosti γ nalazimo zanimljivi dio faznog dijagrama. Za konstantnu vrijednost γ i povećavajući β možemo vidjeti prijelaz iz $D + P$ faze u $D + B$ fazu koja prelazi u $D+P+B$ fazu koja konačno prelazi u $D+P+R+B$ fazu. Granica faza $D+P \to D+B$ je diskontinuirana zbog indirektnog teritorijalnog nadmetanja. Naime, strategije P i B se indirektno nadmeću protiv defektora i pobjednik je ona strategija koja je efektivnija. Iz ovoga možemo vidjeti da kombinirana strategija B može preživjeti samo za vrlo specifični set vrijednosti (β, γ).



Slika 3: $\beta - \gamma$ fazni dijagram za vrijednosti $r = 2.5$



Slika 4: udio strategija za konstantni $\beta = 0.55$, diskontinuirani fazni prijelaz zbog divergencije amplitude oscilacije

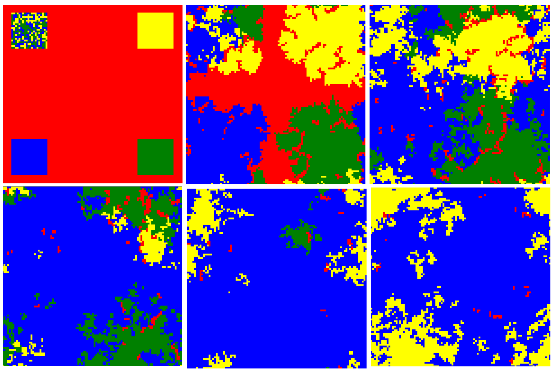
B. Faktor sinergije $r = 2.5$

Za teže uvjete kooperacije domena D strategije postaje puno šira, B također preživljava za puno veći set parametara (β, γ). Zanimljiva je pojava $D + P + B$ faze koja se bazira na cikličnoj dominaciji. U toj fazi strategija D pobijeđuje strategiju P, strategija P pobijeđuje strategiju B a strategija B pobijeđuje strategiju D. Ni jedna od faza u $r = 4.5$ slučaju nije nastala zbog ciklične dominacije već zbog stabilne koegzistencije. Zanimljiva je pojava diskontinuiranog faznog prijelaza iz $D+P+B$ faze u $D(B)$ fazu. Do nje dolazi uslijed amplitudnih oscilacija pojedinih strategija koje se povećavaju kako povećavamo vrijednosti γ sve dok oscilirajuća faza ne dosegne veličinu cijelog sustava. Te amplitudne oscilacije nisu efekt konačnih veličina sustava, štoviše neovisne su o veličini sustava.

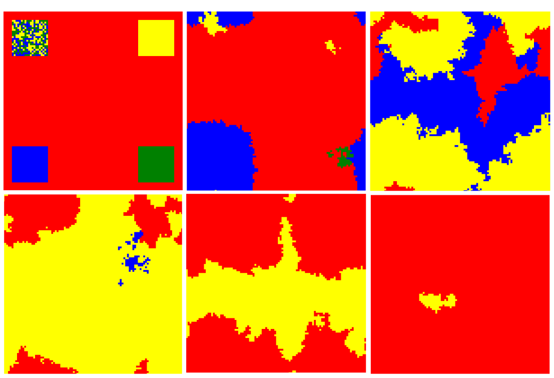
Fazni dijagram također ima prijespomenutu $D(B)$ fazu koja je zapravo ili čista D ili čista B faza. Nasumični efekti diktiraju u koju će konačnu strategiju sustav ući. Postoji i $P + (RB)$ faza koja se pojavljuje za iste vrijednosti (β, γ) kao u prethodnom slučaju te vrlo mala domena $D + B$ faze. To je jedino stabilno riješenje u kojem postoji strategija B i uspiješnija je od strategija P i R. Čini se da uspješnost B strategije jako ovisi o faktoru sinergije. Što su uvjeti za evoluciju kooperacije teži (manji r) veći je uspjeh kombinirane strategije.

5. Zaključak

Glavni cilj ovog seminara bio je provjera uspješnosti kombinirane strategije koja nagrađuje kooperante i kažnjava defektore. Iz dobivenih faznih dijagrama možemo vidjeti da su strategije čistoga surađivanja i kažnjavanja puno efikasnije od kombinirane strategije. To je u skladu s do sada izdanim eksperimentalnim podatcima koji nisu uspijeli potvrditi pretpostavku korelacije negativne i pozitivne recipročnosti [5]. Indirektno teritorijalno nadmetanje koje prelazi u cikličku dominaciju je vrlo zanimljiva pojava no teško ju je reproducirati u eksperimentu. Možemo vidjeti da metode statističke fizike imaju široku primjenu i mogu biti korisne u testiranju teorija evolucije ljudske kooperacije.



Slika 5: Slike kvadratne rešetke, prikaz evolucije sustava za vrijednosti parametara $\beta = 0.1$ i $\gamma = 0.01$ i faktora sinergije $r = 4.5$ nakon $0, 100, 400, 1400, 4400, 6400$ Monte Carlo koraka, možemo vidjeti da sustav u konačnici završava u $D + P + B$ fazi



Slika 6: Slike kvadratne rešetke, prikaz evolucije sustava za vrijednosti parametara $\beta = 0.55$ i $\gamma = 0.1$ i faktora sinergije $r = 2.5$ nakon $0, 100, 300, 400, 700, 900$ Monte Carlo koraka

Literatura

1. A. Szolnoki, M. Perc, Correlation of positive and negative reciprocity fails to confer an evolutionary advantage: Phase transitions to elementary strategies, Phys. Rev. X 3 (2013) 041021

2. H. Gintis, “Strong reciprocity and human sociality,” J. Theor. Biol. 206, 169–179 (2000).

3. https://en.wikipedia.org/wiki/Public_goods_game

4. A. Szolnoki, M. Perc, and G. Szab ´o, “Topology-independent impact of noise on cooperation in spatial public goods games,” Phys. Rev. E 80, 056109 (2009).

5. T. Yamagishi, Y. Horita, Nobuhiro Mifune, H. Hashimoto, Y. Li, M. Shinada, A. Miura, K. Inukai, H. Takagishi, and D. Simunovic, “Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity,” Proc. Natl. Acad. Sci. USA 109, 20364–20368 (2012).