

# Genetski algoritam

Marina Caran, Nela Todorović, Jelena Milivojević, Dara Milojković

Recenzija: Lazar Vasović

10. april 2020.

## 1 O čemu rad govori?

Rad govori o genetskom algoritmu kroz prizmu njegovih elemenata. Predstavljena je biološka evolucija, a zatim i reprezentacije jedinki, kratak istorijat, teorema o shemama kao teorijska podrška ideji, načini generisanja inicijalne populacije, funkcija prilagođenosti, genetski operatori, domen primene, uslovi zaustavljanja, kao i dodaci „za one koji žele da čuju više“.

## 2 Krupne primedbe i sugestije

1. Uvod, kako je navedeno u uputstvima za pisanje seminarskog rada, treba da bude svojevrsna proširena (u odnosu na apstrakt) najava rada i/ili opis teme odnosno nabranje tema koje se nadalje obrađuju. Ovde je to odlično započeto kreativnim poređenjem sa ljudskim promenama, ali je završeno tabelarnim predstavljanjem mogućih reprezentacija. Time je prvi od šest nabrojanih koraka stavljen u uvod, što ne odgovara ulozi podnaslova. Osim toga, kodiranje je često jedan od najbitnijih i najtežih koraka u dizajnu genetskog algoritma za konkretan problem, a svaki naredni korak je prilično vezan za njega (računanje prilagođenosti, inicijalizacija, reprodukcija izvršavaju se nad konkretnom reprezentacijom), tako da u svakom slučaju zaslužuje poseban podnaslov, i to tek posle istorijata, kako ne bi bilo prekida u opisu elemenata. Neka i ostane tabela – u tom slučaju pogledati četvrtu sitnu zamerku – ali ipak je treba izdvojiti iz uvoda.
2. Postoji greška u Primeru 2.1 [str. 4], koja bi mogla da učini definiciju sheme nerazumljivom čitaocu koji se dosad nije susretao sa tim pojmom. Fiksne su pozicije 1, 3, 5 i 6, a ne 1, 3, 4 i 5. Dodatno ceo podnaslov ugrožava vrlo nejasna rečenica koja sledi nakon uvođenja pojmova: „Sa novouvedenim pojmovima i genetskim operatorima, [Da li je ovaj zarez na pravom mestu?] selekcije i mutacije shema kaže [Možda teorema kaže?] da kratke sheme čiji je fitness [fitness] iznad proseka postaje eksponencijalno bolji [Ko postaje?] sa svakom sledećom generacijom.“ Pretpostavljam da je trebalo da piše nešto poput sledećeg: „Sa novouvedenim pojmovima i genetskim operatorima selekcije i mutacije, **teorema** kaže da kratke sheme čiji je fitness iznad proseka postaju eksponencijalno bolje sa svakom sledećom generacijom.“ Ipak, ni ovo nije najadekvatnije, pošto je interpretacija teoreme shema zapravo da kratke natprosečno prilagođene sheme tj. jedinke koje

im pripadaju postaju eksponencijalno **učestalije/češće/zastupljenije** u narednim generacijama.

3. Vrlo neobičan problem je što nigde nije definisano šta je tačno genetski algoritam. Navedeni su njegovi delovi, istorijat, primeri, ali sam algoritam nigde nije izložen. Moguće je da se pretpostavlja da je čitaocu to već poznato, ali čak i u tom slučaju nije dobro pisati o nečemu bez njegove definicije. Takođe postoji poglavlje „Opšti genetski algoritam“ bez naznaka u čemu je razlika između njega i „neopštih“. Upućeni čitalac bi mogao da pretpostavi da se misli na jednostavni/prosti/kanonski genetski algoritam (SGA), ali ipak nije to u pitanju pošto se u nastavku govori o raznim vrstama – deo o prilagođenosti govori u nizovima bitova (jeste SGA), dok npr. deo o selekciji nabraja mnogo operatora (nije SGA, jer on podrazumeva samo proporcionalni model), a pominje se i dosta novijih ideja.
4. Postoji nekoliko problema sa literaturom. Prvi se tiče propusta u samoj literaturi. Od toga je prvo pomešano izdanje između stavki 6 i 8, što je detaljnije opisano u manjoj zamerci 16. Sledeće je link u stavci 3, za koji Tutorialspoint vraća kod 404, jer se loše formatirao, sa crticom. Situacija sa upotrebom pretprinta opisana je u manjoj zamerci 16. Postoje i određeni problemi sa upotrebom BibTeX-a u slučaju više autora. Verovatno je do toga što autorima seminarskog nije bilo jasno da je spoljašnji separator autora citata „and“, a unutrašnji separator prezimena i imena zapeta, npr. „Janičić, Predrag and Nikolić, Mladen“. To je dovelo do situacije da ni za jedan rad sa više od jednog autora osim donekle „Veštačke inteligencije“ autori nisu korektno navedeni, već stoje random pomešana prezimena i zapete. Drugi deo problema je samo citiranje, koje nije u skladu sa primerima iz uputstva i šablona, mada nije jasno da li je tako neophodno ili je samo predlog. Naime, na prezentaciji se citati nalaze razmaknuti pre interpunkcijskog znaka (npr. „Broj je 42 [1].“), dok je u recenziranom radu razmaknuto posle znaka (npr. „Broj je 42. [1]“), mada ima i pre u nabrananju selekcija [str. 5]. Bilo kako bilo, treba ujednačiti.

### 3 Sitne primedbe

1. Autori i njihovi kontakt podaci nisu dobro centrirani, pa Darino prezime i mejl osetno izlaze van margine [str. 1]. U najboljem slučaju, samo nije lepo za oko; u najgorem, došlo bi do odsecanja pri štampanju. Ne znam gde je nastao problem, pa ne mogu da pomognem konkretnim predlogom.
2. U apstraktu se pominju operativna istraživanja [str. 1]. Ovde se nesumnjivo misli na matematičku disciplinu koja se uglavnom naziva operacionim istraživanjima. Nisam stručan u toj oblasti, ali na našem fakultetu sa koristi drugi naziv, pa bih ga koristio i u seminarskom. Primećujem upotrebu prvog naziva na internetu, ali uglavnom u drugim naukama (npr. pedagogiji), što možda i ne odgovara u potpunosti ovoj matematičkoj, ili u hrvatskim izvorima (npr. Ekonomski fakultet Osijek i prevod strane knjige izdat u Zagrebu). Slično, u razradi se pominje selektivni pritisak [str. 5-6], za koji je najverovatnije bolje da bude promenjen u selekcioni.

3. Zašto je kod operatora selekcije i reprodukcije korišćeno veliko početno slovo, kao da su im to vlastita imena [str. 2]? U daljem tekstu je sve malo, kako bi i trebalo, sa izuzenom podnaslova „Operator Reprodukcijske“.
4. Malo je nekonzistentan format teksta u tabeli za tipove reprezentacija [str. 2]. U redu je da prve rečenice budu teze, a ostale pune, ali poslednja rečenica za binarnu je čudna kao teza. Valjalo bi ubaciti negde npr. glagol „koristiti“. Isto važi za tezu „Primer 5.2“ koja izgleda zalutalo; najbolje je staviti je u zagradu uz problem trgovačkog putnika ili možda podbelešku (fusnotu). Primer za binarnu završava se tačkom, a ostali ne, pa je treba izbrisati. Primeri za celobrojnu i realnu imaju razmake između zapeta, a za permutacije ne. Realni brojevi nisu napisani u skladu sa srpskom normom (decimale razdvaja zapeta, ne tačka – npr. treba 1,01 umesto 1.01). Realna nije vlastito ime, pa bi trebalo „Celobrojna, realna“ ili, još bolje, razdvajanje na dve vrste. I samo referisanje na tabelu u tekstu je neobično – koriste se crtica i veliko početno slovo umesto crte i malog.
5. Nije preporučljivo da rečenice počinju brojem, tako da predlažem: „Dana 24. novembra 1859. . .“ [str. 3]
6. Tjuringov citat je već naglašen kosim slovima, suvišni su navodnici [str. 3]. Isto važi i za rad sa iste strane, kao i citat u zaključku [str. 12].
7. Prilagođavanje imena iz stranih jezika je prilično nezgodna stvar u srpskom jeziku. Iz tog razloga, preporučujem navođenje originalnih imena koja se transkribuju [str. 3]. To čitaocima olakšava razumevanje o kome se radi, pogotovu ako postoje greške. U vezi sa tim, „Hans-Hoakim Brememan“ zapravo je Hans-**Joahim** Bremerman (nem. *Hans-Joachim Bremermann*), dok je „Frejžer“ zapravo Frejzer (engl. *Fraser*).
8. Iz primera 3.1 i 3.2 vidi se da je uobičajeno da se funkcija prilagođenosti [str. 5] minimizuje odnosno da za bolje jedinice ona ima manju vrednost, ali to u tekstu nigde nije naznačeno. Štaviše, tvrdnja da „[v]erovatnoća da će se jedinica koristiti za generisanje sledeće generacije raste sa vrednošću funkcije prilagođenosti te jedinice“ sugerise suprotno – maksimizaciju. Traži li se minimim ili maksimum, naravno, zavisi od prirode funkcije prilagođenosti, ali nije dobro mešati dva suprotna pristupa bez ikakvog objašnjenja. Konkretni predlozi: nadalje je sve podređeno maksimizaciji (npr. priča u ruletskoj selekciji), tako da je najbolje obrnuti primere, zameniti ih ili izbaciti i umesto toga dati dodatna pojašnjenja.
9. Sve vreme počev od [str. 5] nadalje zamašćeni (boldirani) pojmovi objašnjavaju se iza crtice (-) umesto crte (—). Za slučaj da je autorima to nepoznanica, za crtu se koristi -- (dve crtice jedna do druge), što se u generisanom pdf-u prikazuje kao željeni karakter. Izuzetak je poglavlje „Primeri genetskog algoritma“ [str. 9], gde se koristi dvotačka, verovatno zato što su različite osobe pisale, pa bi trebalo ispraviti i tu nekonzistentnost. Treći način je u dodatku A.2 [str. 13], gde je separator dupla belina.
10. Postoji nekonzistentnost u upotrebi navodnika. Uglavnom se koriste klasični „. . .“ (otvoreni donji 99, zatvoreni gornji 99), ali na [str. 5] i [str. 13] nalaze se i engleski “. . .“ (otvoreni gornji 66, zatvoreni gornji 99). Engleski

nisu karakteristični za srpski, pa ih treba zameniti klasičnim, a prihvatljivi su i nemački („...“; otvoreni donji 99, zatvoreni gornji 66, njih koristim u recenziji), pa čak i neke egzotičnije varijante sa strelicama.

11. Postoji nekonzistentnost u uvlačenju redova. U većini poglavlja, uvučen je samo prvi red prvog pasusa. Ipak, u drugom poglavlju je uvučen prvi red svakog pasusa. Oba pristupa su ispravna, samo treba ujednačiti. Šablon koji smo dobili propagira uvlačenje svakog pasusa, tako da pretpostavljam da su neuvučeni paragrafi dobijeni upotrebom komande za novi red (//) umesto uobičajenim pravljjenjem novog pasusa sa dva entera.
12. „Što su manje razlike u prilagođenosti u generaciji, to je veći selektivni pritisak.“ [str. 6] Ovo je verovatno greška u brzini, pošto je obrnuto – mala razlika u prilagođenostima povlači za sobom loš, manji selektivni pritisak. Naime, sve jedinke su vrlo slične, tako da nema neke oštre dominacije jedne ili nekog dela jedinki nad ostalima.
13. „[...] jedinki koje međusobno odigravaju turnire.“ [str. 6] Odigravaju jedan turnir, zar ne, a i ovo „međusobno“ je višak. U istom pasusu, „[p]ri tome važi  $n_s < n_{uk}$ “, a zapravo je  $n_s \leq n_{uk}$  (zapravo i to tek pri uslovu  $n_s \in N$ , ali, naravno, ne treba preterivati sa formalizmima). Ne postoji poseban razlog da budu zabranjeni turniri u kojoj učestvuje celokupna populacija.
14. Uniformnosti radi, poglavlje „Operator Reprodukcijske“ [str. 7], nakon već diskutovanog preimenovanja u „Operator[i] reprodukcije“, treba priključiti poglavlju „Opšti genetski algoritam“, verovatno isto nakon preimenovanja (diskutovano kod veće zamerke 3). Inače ne postoji nijedan razlog zašto bi bio izdvojen, osim što ga je verovatno pisala druga osoba, pa je došlo do problema u spajanju. Alternativno, a možda bi čak to bilo preporučljivije, od svakog koraka je moguće napraviti zaseban odeljak. Ukoliko se odluči za spajanje, isto važi i za bezrazložno izdvojeno poglavlje „Uslovi zaustavljanja“ [str. 11], koje bi najverovatnije trebalo da se nalazi pre primera, kao deo samog algoritma koji se redom opisuje.
15. „Primenjuje se formula 4 na nasumično odabranu poziciju i iz izraza 3.“ [str. 7-8] Izraz 3 bira indeks jedinke u populaciji za selekciju (zapravo ni ovo nije tačno, samo govori o njegovoj raspodeli, ali uprostio sam priču), te nema veze sa odabirom indeksa promenljive unutar vektorom predstavljenih jedinki. Ne znam kako je ta referenca („iz izraza 3“) nastala, ali najbezbolnije je izbrisati je. U okolini se koristi i decimalna tačka umesto decimalnog zareza, što je situacija već pomenuta u sitnoj zamerci 4.
16. „[...] virusu COVID-19.“ [str. 9] U pitanju je aktuelna tema i pohvalno je pomenuti je. Iskoristio bih, međutim, priliku da razjasnim čestu grešku u terminologiji vezanoj za tekuću pandemiju. Bolest koja trenutno osetno utiče na normalan tok života naziva se COVID-19 (od *coronavirus disease 2019*). S druge strane, virus koji je izaziva skraćeno se naziva koronavirus (neki kažu virus korona), dok mu je puno ime SARS-CoV-2 (od *severe acute respiratory syndrome coronavirus 2*). U skladu sa tim, treba napisati da su istraživanja vršena na virusu SARS-CoV-2 odnosno novom koronavirusu. Dobra stvar je što je tvrdnja potkrepljena referencom, tako

da je jasno da nije greška do autora seminarskog, već do autora citiranog rada, koji su (pročitao sam im rad, dostupan je na ResearchGate-u) krajnje pogrešno citirali Svetsku zdravstvenu organizaciju u objašnjenju zašto umesto 2019-nCoV (od **2019 novel coronavirus**, ime korišćeno pre određivanja aktuelnog SARS-CoV-2) pišu COVID-19. Dodatan problem koji se ovde provukao jeste sam članak u literaturi. U pitanju je takozvani preprint (preprint), što je nešto slično ovim našim seminarskim, bez ozbiljne recenzije i neobjavljeno ni u jednom časopisu. Razlika je, doduše, što su autori mahom doktori nauka, pa je, po mom mišljenju, ipak u redu citirati ga, ali svakako ne kao „*Proceedings of the World Congress on Engineering 2011 Vol II*, 2:1–6, 7 2011“, što je zamenjeno sa referencom 8.

17. „Funkcija prilagođenosti je predstavljena kao par dužine puta i svih gradova u putu. Što je manja dužina puta bolje se uklapa gen.“ [str. 10] Zašto par? Valjda je samo dužina puta. U šta se uklapa i koji gen? Verovatno je želelo da se kaže nešto poput: „Što je manja dužina puta, to je jedinka bliža optimalnom rešenju.“ Takođe, optika je kritična kao primer, pošto problem nije precizno definisan – tek je sa slike jasno da jedinke karakteriše ništa drugo do ugao odnosno nagib poluge, a ne npr. dužina.
18. U dodatku A.5 [str. 14] koristi se crtica umesto crte, otvorene zagrade nisu odvojene od reči koja im prethodi, postoji razmak u „gospodar-sluga“ i dvaput je ponovljeno da se koristi za TSP.
19. Greške u kucanju i brzini: zamena slova u „ogranizma“ [str. 2], fali „je“ u „opstalo do dan danas“ [str. 3], ponavljanje u „da imam obavezu da moram uzeti u obzir“ [str. 3] (podsetnik, ako zatreba pri ispravci, rekao je „that one is obliged to regard the machine as showing intelligence“), „Popularnosti genetskim algoritmima“ [str. 3] umesto „genetskih algoritama“, „reprezentacije gena koja su uvedena“ [str. 4], ošišana latinica u „vrednošću“ [str. 5], suviše „da“ u „pitanje da koji je“ [str. 5], verovatno treba „u tabeli“ u tekstu „na slici 1 i tabeli 2“ [str. 6], zamena slova u „varijatni“ [str. 7], treba „povećavanjem veličine populacije“ ili „povećavanjem broja jedinki populacije“ umesto „povećavanjem broja populacije“ [str. 7], treba „najprilagodljiviji genom“ umesto „najprilagodljiviji gen“ [str. 7] (u nekim slučajevima bi moglo i gen, ali najmanje kad je u pitanju permutacioni problem poput TSP), treba „lozinki“ umesto „lozinka“ [str. 10], verovatno treba „došlo je do mogućnosti“ umesto „doveo je do potrebe“ [str. 14], treba „vrsta koje“ umesto „vrsti koji“ [str. 14], treba „granularni“ umesto „gradirani“ [str. 14], duplirano slovo u „selekcijje“ [str. 14], višak „se“ u „da se na“ [str. 14], verovatno „stavnovnici“ umesto „članovi“ [str. 14].

## 4 Provera sadržajnosti i forme seminarskog rada

1. Da li rad dobro odgovara na zadatu temu?  
Rad u najvećoj meri odgovara na zadatu temu. Iako je propušteno navođenje samog algoritma, pristojno je obrađen svaki njegov pojedinačan deo.
2. Da li je nešto važno propušteno?  
Jeste, kako je i navedeno u odeljku „Krupne primedbe i sugestije“. Ukrat-

ko – nedostatak precizne definicije, neadekvatan uvod, loša interpretacija teoreme o shemama, loše navedena literatura.

3. Da li ima suštinskih grešaka i propusta?  
Nejasna je razlika između ovog i prethodnog pitanja, ali, ukoliko se pod suštinskim greškama podrazumevaju neke koje bi učinile da rad bez mogućnosti popravke bude odbijen u nekom kontekstu, takvih nema. Sve navedene greške mogu se ispraviti u jednom danu, sigurno za mnogo manje vremena nego što je trebalo za pisanje ove recenzije.
4. Da li je naslov rada dobro izabran?  
Naslov je dobar, govori se stvarno o genetskim algoritmima.
5. Da li sažetak sadrži prave podatke o radu?  
Sažetak sadrži prave podatke, odnosno dve rečenice o samom algoritmu i jednu kratku o sadržaju rada, ali koja u potpunosti pogađa suštinu.
6. Da li je rad lak-težak za čitanje?  
Sa izuzetkom izdvojenih nejasnih delova, rad nije težak za čitanje. Sve rečenice su jednostavne i prikladne dužine, bez prenatrpane upotrebe stranih ili stručnih termina. Da je čitljivost na zadovoljavajućem nivou, slažu se i tri algoritma provere čitljivosti kroz koje sam propustio tekst.
7. Da li je za razumevanje teksta potrebno predznanje i u kolikoj meri?  
Kako nije naveden sam genetski algoritam, neophodno je znati ga kako bi bilo jasno na koji način se tačno opisani elementi algoritma uklapaju u širu sliku. Neophodno je znati i šta su Grejov kod i Monte Karlo metode odnosno posedovati osnovna znanja iz računarstva i statistike. Pohvalno je što nije neophodno znati šta je evolucija, jer je ona slikovito objašnjena. Postoji i nekoliko iznenadnih neobjašnjenih termina, kao što je „alel“.
8. Da li je u radu navedena odgovarajuća literatura?  
Postoji nekoliko problema koji se ne mogu zanemariti i koji su navedeni u odeljku „Krupne primedbe i sugestije“. Ukratko – loše formatirani i pomešani podaci o izdanju, problemi sa BibTeX-om. Ipak, napominjem da ne smatram da je ijedna stavka takva da bi se morala izbaciti. Takođe, nijedan od časopisa ne nalazi se na crnim listama. Ispunjen je uslov da je citirano barem sedam članaka iz časopisa, knjiga i internet stranica.
9. Da li su u radu reference korektno navedene?  
Reference su korektno navedene, što kao potvrda potencijalno spornih tvrdjenja, što kao upućivanje čitaoca na dodatne informacije, s tim što se u podnaslovu „Operator Reprodukcijske“ ne nalazi nijedna, a valjalo bi da je svaki deo pokriven makar uputima na dalju literaturu. Važno je napomenuti da najverovatnije nema plagijata – proverena je navedena literatura i sav napisan tekst od nekoliko uzetih uzoraka nije doslovno navođen odnosno preveden niti sa tek malim izmenama. Jediniti potencijalni problem, sa stilom citiranja, naveden je u odeljku „Krupne primedbe i sugestije“.
10. Da li je struktura rada adekvatna?  
Struktura prati preporučenu shemu uvod–razrada–zaključak, a u razradi postoji logička veza prethodnog sa sledećim, s tim što postoje primedbe za organizaciju podnaslova koje su iznete u odeljku „Sitne primedbe“.

11. Da li rad sadrži sve elemente propisane uslovom seminarskog rada (slike, tabele, broj strana...)?

Rad sadrži originalne slike i tabele i na sve postoji prikladna referenca. I broj strana je adekvatan – dvanaest glavnog teksta sa naslovnom stranom i literaturom i dve dodatka. Zapravo je primetno da je ograničenje maksimalnog broja strana pravilo problem autorima, jer su očigledno mogli i želeli još više da napišu. U skladu sa tim, dodaci nisu samo kodovi, kako se zahteva u šablonu, ali smatram da su ipak prikladni. Postoji sitan problem kod podnaslova „Opšti genetski algoritam“. Naime, verovatno u svrhe ispunjavanja pravila da mora da postoji pasus između naslova i podnaslova, stavljena je jedna rečenica: „Slede osnovne osobine genetskog algoritma.“ Međutim, na ovaj način je nastao novi problem, pošto takođe postoji pravilo po kom paragraf ne sme da ima samo jednu rečenicu. Van toga, pasusi su prikladno strukturirani i koherentni u iznesenim idejama.

12. Da li su slike i tabele funkcionalne i adekvatne?

Tehnički gledano, u uputstvu stoji da ne treba komplikovati prikaz sa 3D elementima, a isto je rekao i profesor na predavanjima, ali meni je ova trodimenziona pitica baš simpatična, ne bih joj smanjivao broj dimenzija. Doduše, kad malo bolje razmislim, možda bih je drugačije obojio, pošto je malo nezgodno što su sve nijanse plave, pri čemu su dve jedinke sa jednakim verovatnoćama izbora iste boje, pa možda može doći do nekog zbunjivanja u interpretaciji. Osim toga, slike su dobre, ali postoji problem sa tabelama – opis se nalazi ispod njih umesto iznad, kako bi trebalo.

## 5 Ocenite sebe

b) Jasno je da kao student nisam ekspert (zar je iko?), ali smatram da sam veoma upućen u temu. Naime, o genetskim algoritmima sam slušao na kursu Računarska inteligencija, pri čemu sam pročitao sav tekst o njima iz knjige „Computational Intelligence: An Introduction“, koja je i citirana u radu. Pritom sam izučavao problematiku i iz druge literature. Većina informacija iznetih u radu bila mi je poznata, a prilikom recenziranja sam pročitao i sve citirane stavke koje dosad nisam, tako da se usput zahvaljujem autorima na novom znanju, pogotovu za deo o upotrebi genetskih algoritama pri stvaranju muzike.

## 6 Poverljivi komentari

Komentar nije direktno vezan za recenzirani rad, ali je nimenjen profesoru, tako da ga ovde ostavljam. Predlažem da ubuduće recenzentima budu dostupne i tex i bib datoteke uz generisane pdf-ove. Na taj način bi se i uz tehničke zamerke moglo predložiti rešenje, a ne samo zamerka koja je bez propratne sugestije donekle beskorisna. U mom slučaju, problem je ispadanje imena i kontakt podataka iz margine, što iskreno sumnjam da sami autori nisu primetili, pošto je na naslovnoj strani i baš zapada za oko. Najverovatnije samo nisu umeli da reše, pa bi korisno bilo da sam im uz zamerku mogao reći i kako to da popra-ve. Druga stvar je kod neadekvatno prikazane literature, za šta sam morao da pretpostavim da je problem sa nepoznavanjem separatora u BibTeX-u.