

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

Genetski algoritmi

Nikolina Mišljenčević

Branko Spasojević

Voditelj: *Marin Golub*

Zagreb, svibanj, 2008

Sadržaj

1. Uvod.....	1
2. Evolucija u prirodi.....	2
2.1 Genetski materijal – geni, kromosomi i molekula DNA.....	2
2.2 Križanje i mutacija.....	3
3. Genetski algoritmi	4
3.1 Populacija	4
3.2 Funkcija dobrote.....	4
3.3 Genetski operatori.....	5
3.3.1 Križanje	5
3.3.2 Mutacija	6
3.3.3 Selekcija.....	7
3.4 Selekcija jedinki	7
3.4.1 Jednostavna	7
3.4.2 Turnirska selekcija.....	7
3.4.3 Eliminacijska selekcija	8
3.4.4 Elitizam.....	8
4. Primjer genetskog algoritma.....	9
5. Zaključak.....	10
6. Literatura.....	11
7. Sažetak	12

1. Uvod

2. Evolucija u prirodi

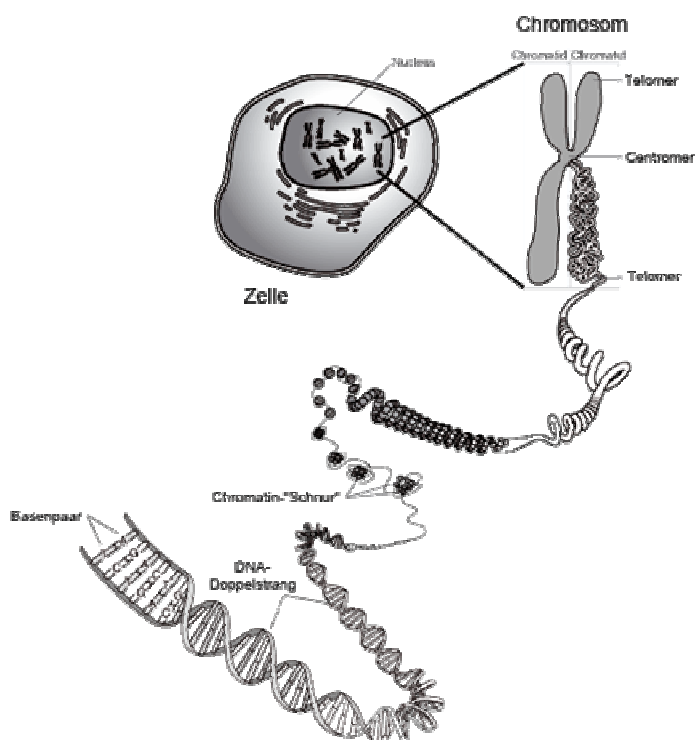
Charles Darwin je ustanovio da je glavni pokretač evolucije prirodna selekcija. Priroda određuje uvjete za život - one jedinke koje imaju bolji genetski materijal prilagodit će se tim uvjetima i opstati, dok će slabije jedinke s vremenom izumrijeti. Možda najpoznatiji primjer vrste koja se nikako nije uspjela prilagoditi novim okolišnim uvjetima upravo su dinosauri koji su, kao što nam je poznato, izumrli. Kako jedna vrsta izumire, druga vrsta stvara nove potomke i na taj način broj živih bića uvijek ostaje barem približno konstantan.

2.1 Genetski materijal – geni, kromosomi i molekula DNA

Naše se tijelo sastoji od 6 bilijuna stanica, od kojih svaka u sebi nosi 23 para homolognih kromosoma. Unutar svakog para homolognih kromosoma jedan je kromosom naslijeđen od oca, a drugi od majke. Homologni kromosomi sadrže gene za ista svojstva na istom mjestu, tj. lokusu. Npr. na jednom od tih parova imamo gen koji određuje boju očiju. Razlika je u tome što gen naslijeđen od oca može određivati plavu boju očiju a od majke smeđu boju očiju. Od tih svojstava samo će se jedno ispoljiti ovisno o tome koje svojstvo je dominantno a koje recesivno. U konkretnom slučaju potomak će imati smeđe oči jer je to dominantno svojstvo.

Strukturu molekule DNA otkrili su sredinom 20. st. Watson i Creek. DNA je molekula koja se sastoji od određenog slijeda nukleotida. Nukleotid je dušična baza na koju je vezan šećer deoksiriboza i fosfatna skupina. Dva lanca DNA se spajaju po principu komplementarnosti dušičnih baza (adenin se spaja sa timinom, a gvanin s citozinom). Određeni slijed parova dušičnih baza čini gen, odnosno prenosioca nasljednih svojstava.

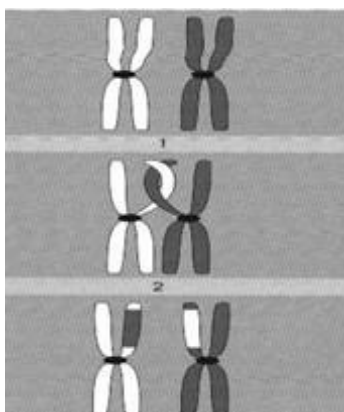
Kromosomi su transportni oblik molekule DNA koja se nalazi u svakoj našoj tjelesnoj stanici. Svaka molekula DNA dugačka je otprilike 2 metra. Prilikom diobe molekula DNA se mora replicirati kako bi svaka od novonastalih stanica mogla dobiti po jedan njen primjerak. Odmah je jasno da neće biti jednostavno odvojiti dvije molekule dugačke 2 metra na različite polove sićušne stanice. Zbog toga se DNA mora sažeti u svoj transportni oblik, već spomenuti kromosom.



Slika 1. Kromosom kao transportni oblik molekule DNA

2.2 Križanje i mutacija

Križanjem se nasljedna svojstva prenose sa parentalne generacije na potomstvo. Pritom roditelji na svoju djecu prenose genetski materijal, tj. svoje dobre i loše osobine. Slučajna promjena nekih svojstava jedinke naziva se mutacija. U genetskom algoritmu operacija mutacije je unarna operacija jer se promjene gena vrše nad jednom jedinom jedinkom. Za razliku od mutacije, operacija križanja je operacija višeg reda jer se nove jedinke stvaraju kombinacijom gena više različitih jedinki iz parentalne skupine.



Slika 2. Križanje

3. Genetski algoritmi

Genetski algoritmi su jedan od četiri tipa evolucijskih algoritama. Svoju ideju crpe iz procesa evolucije u prirodi koju koriste za rješavanje problema optimiranja. Simulacija evolucijskog procesa odvija se korištenjem analogija iz prirode. Prisutni su svi važni čimbenici procesa evolucije koji uključuju: populaciju (jedinke), reproduksijske metode, funkciju dobrote koja predstavlja sposobnost preživljavanja jedinke u prirodi, te selekcijski mehanizam ili opstanak jedinke. Promoviranjem željenih osobina jedinke moguće je usmjeravati evoluciju u željenom smjeru i na taj način pronaći optimalno rješenje problema čak i u slučajevima kada ne postoji eksplicitno rješenje.

Primjena genetskih algoritama je moguća u slučajevima kada problem možemo opisati kao pretraživanje ili optimizaciju proizvoljnih podataka, te poznamo način mogućeg mjerenja uspješnost svakog pojedinog rješenja. Iako genetski algoritam pretražuje i optimizira zadani skup rješenja ne predstavlja uvijek optimum procesa, te je potrebno pažljivo konstruirati okruženje evolucije kako bi se zaista dobili rezultati bliski rješenju. Prvi korak u konstrukciji ispravnog evolucijskog okruženja je razumijevanje svih njegovih dijelova (populacija, dobrota, selekcija, operatori) te njihove interakcija.

3.1 Populacija

Populacija je skup jedinki iste vrste smještenih na nekom području. Kako dio populacije stari i umire, tako se razmnožavanjem stvaraju novi potomci i veličina populacije u svakoj generaciji ostaje konstantna. U genetskom algoritmu populacija predstavlja skup jedinki od kojih je svaka jedinka potencijalno rješenje zadanog problema. Početna populacija može biti odabrana slučajnim odabirom ili nekim drugim optimizacijskim postupkom. Odumiranje slabijih jedinki koje se nisu uspjele prilagoditi novim životnim uvjetima i opstanak jedinki koje su to uspjele u genetskim se algoritmima provodi uporabom takozvane funkcije dobrote (detaljnije objašnjeno u idućem poglavlju). Razmnožavanje jedinki koje su se svojim dobrim svojstvima uspjele probiti provodi se operacijom križanja. Iz iteracije u iteraciju jedinke u populaciji poprimaju sve poželjnija svojstva. Algoritam obično završava dosežanjem zadanog broja iteracija. Kada je uvjet završetka ispunjen, iz dobivene populacije odabire se najbolja jedinka i ona predstavlja rješenje optimizacijskog problema.

3.2 Funkcija dobrote

Kao što je već spomenuto u prethodnom poglavlju, funkcija dobrote predstavlja prirodnu okolinu koja vrši selekciju nad jedinkama. Što je jedinka bolje prilagođena okolini u kojoj živi, tj. što joj je veća dobrota, veće su joj i šanse za preživljavanje. Odmah treba napomenuti da nije ključ problema odmah u prvoj iteraciji odabrati jedinke s najvećom dobrotom jer jedinke koje su po tom pitanju prosječne nose neka bitna svojstva koja nije poželjno izgubiti. Dakle, izbacuju se one jedinke koje imaju ekstremno nisku razinu dobrote unutar promatrane populacije.

Odabir adekvatne funkcije dobrote obično je ključan problem kod implementacije genetskog algoritma. Također, obzirom da se radi o funkciji koja se u algoritmu najviše koristi, funkcija dobrote trebala bi biti što je moguće jednostavnija i brža.

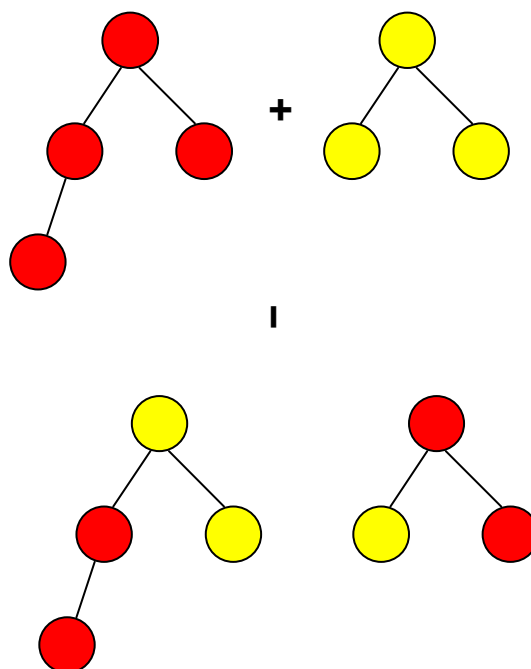
3.3 Genetski operatori

Kako bi se osigurao napredak populacije potrebno je osmisliti načine stvaranja novih jedinki. Cilj je osiguranje opstanka genetskog materijala boljih jedinki, tj. onih koje posjeduju veći koeficijent dobrote i na taj način promovirati što više željenih osobina. Promatrajući evoluciju živih bića nailazimo na nekoliko načina stvaranja novog genetskog materijala: križanje, mutacija i reprodukcija. S obzirom na strukturu prikaza jedinice promatrat ćemo tri slučaja: stablo, graf i linearan prikaz.

3.3.1 Križanje

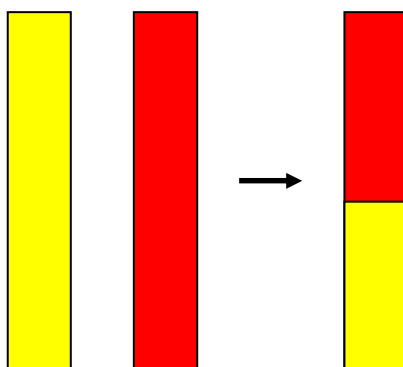
Postupak križanja imitira spolnu reprodukciju dviju jedinki (roditelji), prilikom čega dolazi do zamjene jednog dijela roditeljskog gena onime drugoga. Novonastala jedinka (dijete) posjeduje genetski kod obaju roditelja, te kombinaciju njihovih osobina. Križanjem dviju jedinki koje posjeduju pozitivne osobine očekujemo dijete koje će također posjedovati pozitivne osobine, tim više što je moguće nadomještanje mogućih negativnih osobina nekog od roditelja genima onog drugog čije su osobine bolje na nepovoljnom segmentu.

Križanje kod stablastog prikaza prikazujemo grafički slikom 1. Prilikom križanja odabere se čvor ispod kojega se odvija križanje, te se kod djeteta zamjene roditeljska pod stabla.



Slika 3. Križanje stabala

Kod linearnog križanja potrebno je odrediti proizvoljan broj točaka prekida na kojima se odvija zamjena segmenata roditeljskih gena. Grafički prikaz linearnog križanja predočen je na slici 2. Broj točaka prekida je proizvoljan i može varirati između jedne i broja gena u jedinki. Dodatno je moguće definirati i vjerojatnost križanja u određenim točkama i na taj način poboljšati učinkovitost rada algoritma ako su poznata mjesta grupiranja gena zaduženih za isto obilježje.



Slika 4. Linearno križanje

Prilikom križanja roditelja opisanih strukturom grafa koristi se sljedeći algoritam:

- Podijeli svaki od grafova u skupove od po dva čvora
 - Dodijeliti ime svim bridovima „unutarnji“ ukoliko povezuju čvorove unutar fragmenta, inače „vanjski“
 - Dodijeliti ime svakom čvoru unutar fragmenta „izlaz“ ukoliko je izvor „vanjskom“ bridu, a „ulaz“ ukoliko je odredište vanjskom bridu
- Zamijeniti fragmente roditelja
- Rasporediti bridove tako da svi vanjski bridovi fragmenta koji sada pripadaju zajedno pokazuju na nasumično odabrane ulazne čvorove drugih fragmenata

3.3.2 Mutacija

Mutacija je trajna promjena genetskog materijala, najčešće uzrokovana vanjskim čimbenicima. Smatraju se jednim od preduvjeta evolucije jer se procesom prirodnog odabira u populaciji nakupljaju mutacije koje omogućuju bolju prilagodbu uvjetima okoliša i time poboljšavaju vjerojatnost preživljavanja jedinke koja ih nosi te prijenose na sljedeće generacije.

Mutacija se odvija nad pojedinom jedinkom, obično nakon postupka križanja svako dijete prolazi kroz proces mutacije s proizvoljnim parametrom (najčešće malim) vjerojatnosti. Ukoliko je jedinka odabrana za mutaciju bira se proizvoljan dio

genetskog koda jedinke koji će biti zamijenjen nasumičnim nizom (ili drugom strukturom u ovisnosti o prikazu jedinke).

3.3.3 Selekcija

Operator reprodukcije je najjednostavniji. Jedinica je odabrana, te je njena kopija dodana populaciji. Sada se nalaze dvije identične jedinke u populaciji.

3.4 Selekcija jedinki

Selekcija je posljedica generacije velikog broja jedinki. Kako nije moguće sve jedinke očuvati u populaciji mora se odrediti način odabira jedinki koje će opstati, te onih koje ćemo ukloniti, tj. koje će uginuti. Iako je cilj da populacija sadrži samo jedinke pozitivnih svojstava, prerana konvergencija kao posljedica selekcije samo najboljih jedinki odražava se negativno na uspješnost algoritma. Postoji nekoliko različitih vrsta selekcija koje pokušavaju riješiti taj problem.

Neke od popularnijih tehnika selekcije su jednostavna, turnirska, eliminacijska i elitizam.

3.4.1 Jednostavna

Jednostavna selekcija generira populaciju istog broja jedinki korištenjem dobrote svake jedinke kao vjerojatnosti odabira. Što je veća dobrota jedinke to je veća vjerojatnost da će ta jedinka biti odabrana, ali ujedno pruža mogućnost da se odaberu i jedinke manje dobrote.

Ukoliko neka jedinka posjeduje velik iznos dobrote proporcionalna je vjerojatnost njene prisutnosti veći broj puta u novoj populaciji. Pojavljivanje duplikata u populaciji pokazalo se kao nepoželjan fenomen koji usporava algoritam te su osmišljene tehnike koje ih nastoje ublažiti. Jedan od načina je da se jedinke poredaju uzlazno po vrijednosti dobrote, te se vjerojatnost odabira jedinke računa položajem u nizu i visini dobrote. Rezultat ove korekcije je ujednačavanje vjerojatnosti odabira jedinki, dok se u isto vrijeme zadržava veća vjerojatnost pojave jedinki s većom dobrotom.

3.4.2 Turnirska selekcija

Turnirska selekcija simulira turnire u kojima se pobjednici dodaju u skup jedinki za razmnožavanje. Broj sudionika turnira predstavlja selekcijski pritisak koji kontrolira brzinu konvergencije. Ukoliko je broj sudionika veći manja je vjerojatnost odabira jedinki koje posjeduju manji koeficijent dobrote, te je brža konvergencija prema rješenju. Brza konvergencija dvosjekli je mač koji često dovodi do suboptimalnog rješenja, stoga je važno pažljivo odrediti selekcijski pritisak.

Eliminacijska turnirska selekcija nakon nasumičnog odabira proizvoljnog broja jedinki ukloni najlošiju jedinku nadomještajući je djetetom dviju slučajno odabranih preživjelih jedinki.

3.4.3 Eliminacijska selekcija

Eliminacijska selekcija temelji se na izbacivanju loših jedinki radije nego odabiru boljih pri generiranju nove populacije. Kako bi odabrala lošu jedinku izračunava se kazna svake jedinke koja predstavlja vjerojatnost odabira jedinke za izbacivanje. Kazna jedinke je razlika maksimalne dobrote i dobrote jedinke. Ovim načinom odabira jedinke za izbacivanje iz postupka selekcije u potpunosti smo uklonili jedinku s najvećom dobrotom, pošto ona ima kaznu nula. Zbog očuvanja broja jedinki u populaciji nakon izbacivanja stvara se nova jedinka nekim od genetskih operatora te nastavlja proces selekcije.

3.4.4 Elitizam

Kako se dobra rješenja ne bi izgubila uporabom genetskih operatora ili eliminacijom tijekom selekcije javlja se potreba za zaštitom najbolje jedinke od izmjena ili eliminacije. Zaštita najbolje jedinke naziva se elitizam i osigurava da se svaka nova generacija kreće prema globalnom optimumu. Ukoliko je populacija velika, prilikom pretrage za najboljom jedinkom uvodi se nepotreban overhead koji može dodatno usporiti rad algoritma.

4. Primjer genetskog algoritma

5. Zaključak

Još nemamo.

6. Literatura

- [1] AI Application Programming, M. Tim Jones, Charles River Media, 2003
- [2] Genetski algoritam: Prvi Dio, Marin Golub, 2004

7. Sažetak