

Univerzitet Crne Gore
Elektrotehnički fakultet

Genetički algoritmi

- Sedmi termin -

Istorijat

- Darwinova teorija evolucije se oslanja na prirodni odabir, selekciju, koja podržava jedinke bolje prilagođenim uslovima sredine, dok one slabije prilagođene odumiru.
- Kompjuterska simulacija evolucije je počela 1954. godine radom N. A. Barićelija (*Nils Aall Barricelli*)
- Tokom 1960-ih i 1970-ih godina vještačka evolucija je postepeno razvijana, da bi početkom 1970-ih genetički algoritmi postali popularni zahvaljujući Džonu Holandu (*John Holland*).
- U knjizi „*Adaptation in Natural and Artificial Systems*“ je postavio ideje i principe današnjih genetičkih algoritama.
- Genetički algoritmi su bili prisutni samo u teoriji sve do sredine 1980-ih kada je održana prva internacionalna konferencija na temu genetičkih algoritama u Pitsburgu.
- Najveći doprinos genetički algoritmi daju pri traženju globalnog ekstremuma nepravilnih funkcija u velikom prostoru pretrage.

Genetički algoritmi vs genetike

- Genetički algoritmi prate principe na kojima se zasniva genetika/evolucija i na taj način, polazeći od nekog mogućeg rješenja, brzo pronalaze konačno rješenje.
- Genetički algoritmi stvaraju čitav niz (populaciju) potencijalnih rješenja (jedinki), koja primjenom raznih transformacija evoluiraju stvarajući u narednim populacijama jedinke koje su bolje prilagođene datoj okolini.
- Transformacije, koje se primjenjuju nad jedinkama (rješenjima), su ekvivalentne promjenama koje se dešavaju nad nekom vrstom u procesu evolucije (odabir roditelja, ukrštanje, mutacija).
- Određeni termini upotrijebljeni u genetici uvedeni su u terminologiju genetičkih algoritama, ali dosta pojednostavljeni:
 - Hromozomi predstavljaju jednu jedinku i prikazani su korišćenjem odgovarajućeg koda. Predstavljaju iskodirano rješenje.
 - Geni su elementi hromozoma i predstavljeni su pomoću pojedinačnih simbola ili grupe susjednih simbola koji čine jedan blok simbola i „opisuju“ pojedinačnu osobinu jedne jedinke (hromozoma).

Faze genetičkog algoritma

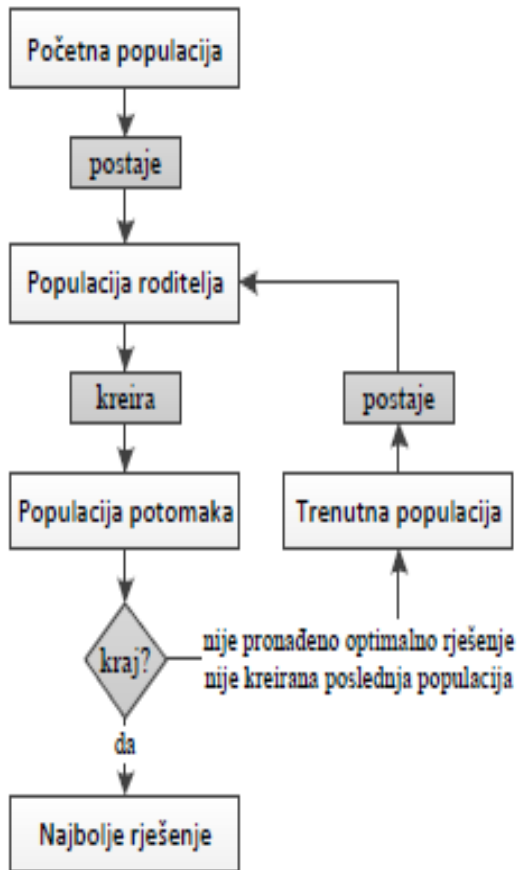
- Početna generacija predstavlja niz rješenja koja mogu imati neko *apriori* znanje ili mogu biti sasvim slučajno generisana.

Ta generacija zatim postaje populacija roditelja.

Roditelji učestvuju u procesu ukrštanja i na taj način generišu populaciju potomaka:

- Selekcija roditelja u zavisnosti od kvaliteta
- Razmjena genetičkog materijala (ukrštanje)
- Mutacija

Algoritam se prekida prolaskom kroz unaprijed definisani broj generacija ili pronalaženjem optimalnog rješenja (rješenje se ne popravlja u par generacija).



Prikaz rješenja

- Da li je određeni problem tj. potencijalno rješenje problema moguće predstaviti pomoću hromozoma i gena?
- Jedinka posjeduje karakteristike koje je opisuju i koje su bitne u donošenju odluke da li je upravo ona odgovarajuće rješenje. Te karakteristike su predstavljene pomoću gena.
- U zavisnosti od preciznosti kojom želimo okarakterisati neku jedinku, potreban nam je manji ili veći broj gena.
- Jedinke se najčešće predstavljaju binarnim kodom, a to mogu biti i drugi karakteri, vektori itd.
- Binarnim kodom se mogu kodirati gotovo svi problemi koji se rješavaju genetičkim algoritmima.
- Binarni kod nije uvijek najbolji vid prikaza rješenja. 7-> 0111, 8->1000, 15->1111 Veliko Hamming-ovo rastojanje – promjena jednog bita ne dovodi uvijek do promjene za istu vrijednost. Gray-ev kod ima Hammingovo rastojanje jedan između susjednih vrijednosti.

2-bit	4-bit
00	0000
01	0001
11	0011
10	0010
	0110
	0111
3-bit	0101
000	0100
001	1100
011	1101
010	1111
110	1110
111	1010
101	1011
100	1001
	1000

Fitnes funkcija

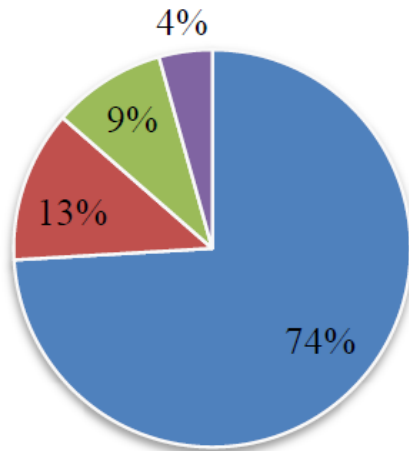
- Ocjena prilagođenosti jedinke okruženju se naziva fitnes funkcija.
- Svaka jedinka ima svoj fitnes – funkciju dobrote.
- Ova funkcija se izračunava na startu svake iteracije algoritma i to za svaku jedinku ponaosob.
- Od fitnes funkcije zavisi vjerovatnoća selekcije jedinke koja će biti uparena, jer ukazuje na kvalitet jedinke – potencijalnog rješenja.
- Veći fitnes znači bolju adaptaciju jedinke ka okolini, samim tim i češći odabir u procesu razmnožavanja.
- Način na koji će se izračunavati pomenuta funkcija je od ključnog značaja za funkcionisanje svakog genetičkog algoritma.
- Upravo pomenuta fitnes funkcija predstavlja nepravilnu funkciju sa velikim brojem lokalnih ekstremuma od kojih treba pronaći globalni ekstremum.

Selekcija

- Selekcija se obavlja unutar populacije roditelja. Biraju se parovi roditelja između kojih će se izvršiti razmjena genetičkog materijala.
- Vjerovatnoća selekcije roditelja direktno je povezana sa fitnessom.
- Hromozomi (roditelji) koji imaju veći fitness (bolje adaptirani životnoj sredini), biće češće birani i imaće više potomaka. Samim tim imaće i veću vjerovatnoću prenošenja svog genetičkog materijala (znanje roditelja) u narednu generaciju.
- Nakon određenog broja generacija sve jedinke će početi da liče jedna na drugu i na taj način sužavamo prostor za pretragu (broj mogućih rješenja). To može usloviti da genetički algoritam dosegne neki lokalni maksimum i tu ostane u nemogućnosti da dosegne i globalni maksimum, što mu je i cilj.
- Da bi se izbjegao ovaj nedostatak prilikom vršenja selekcije uvodi se faktor slučajnog odabira. On omogućava da se prilikom selekcije odaberu i jedinke sa manjom vrijednošću funkcije dobrote, što će kreirati generaciju kod koje sve jedinke ne liče jedna na drugu i na taj način se pokušava izbjeći lokalni maksimumi.

Rulet točak selekcija

- Disk izdijeljen na djelove koji su oblika kružnog isječka. Kružnih isječaka imamo onoliko koliko ima jedinki roditeljske populacije. Površina svakog isječka je proporcionalna fitnessu jedinke koju predstavlja.
- Zamislimo da "ispalimo" pucanj direktno na jedan ovako generisani disk koji se okreće. Jedinica koju predstavlja pogođeni isječek je izabrana za proces ukrštanja.
- Na ovaj način vjerovatnoća odabira jedinke u velikoj mjeri zavisi od vrijednosti njene dobrote, ali i od slučajnog odabira, čime je omogućeno da i jedinke sa manjom vrijednošću dobrote budu birane.



■ Hromozom 1 - Dobrota = 120

■ Hromozom 2 - Dobrota = 20

■ Hromozom 3 - Dobrota = 15

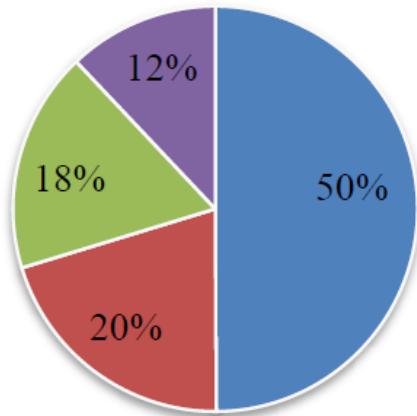
■ Hromozom 4 - Dobrota = 7

Rulet točak selekcija - skaliranje

- Može favorizovati neke "super individue" koje imaju veliki fitness, a samim tim zauzimaju veliku površinu na točku, pa su gotovo uvijek birane. Time se dolazi do prerane konvergencije.
- Da bi imali efikasnu selekciju, potrebno je skalirati fitness funkciju.
- Vjerovatnoća selekcije jedne individue x se tada izračunava kao:

$$P(x) = \frac{f'(x)}{\sum_{x \in \text{Populacija}} f'(x)}$$

- $f'(x)$ – skalirana vrijednost fitness funkcije



■ Hromozom 1 - Dobrota = 120

■ Hromozom 2 - Dobrota = 20

■ Hromozom 3 - Dobrota = 15

■ Hromozom 4 - Dobrota = 7

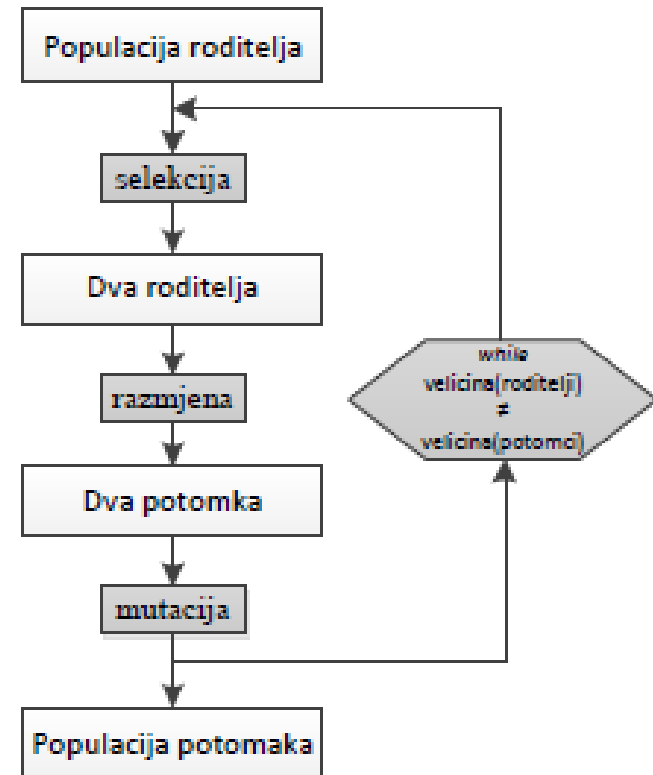
$$P(x) = \frac{\sqrt{f(x)}}{\sum_{x \in \text{Populacija}} \sqrt{f(x)}}$$

Turnirska selekcija

- Uključuje organizovanje "turnira" između nekoliko slučajno izabranih jedinki iz cjelokupne populacije.
- Pobjednik svakog turnira (pobjednik je onaj sa najvećom vrijednošću fitnesa) je izabran za proces ukrštanja.
- Najbitnije je dobro odrediti veličinu turnira, kako bi se izvršila što efikasnija selekcija.
- Pravljenjem velikih turnira jedinke sa niskim fitnesom će imati veoma malu šansu da budu izabrane.
- Moguće je pobjednike turnira odstranjivati iz populacije i tako omogućiti da u budućnosti ne mogu više biti birani, ili ih zadržavati i na taj način omogućiti da mogu stvarati potomke više puta.
- Efikasnost selekcije se vrlo jednostavno podešava promjenom samo jednog parametra, a to je veličina turnira.

Koraci nastanka potomaka

- Nakon izvršenog odabira roditelja nastupa proces ukrštanja u kom dolazi do razmjene genetičkog materijala - **crossover**.
- Razmjena se vrši sa određenom vjerovatnoćom - između 0.5 i 0.9.
- Hromozomi (potomci) formirani razmjenom se zatim podvrgavaju procesu koji se naziva mutacija (**mutation**).
- Mutacija se primjenjuje sa veoma malom vjerovatnoćom koja često opada kako dublje ulazimo u proces evolucije.



Ukrštanje i mutacija

- Operatori ukrštanja i mutacije vrše koncentraciju jedinki odnosno rješenja, svaki na svoj način.
- Ukrštanje koncentriše nove populacije oko optimalnog rješenja. Može se desiti da dovede algoritam do prerane konvergencije, tj. do koncentracije populacije oko rješenja približnog željenom. Ovo se dešava zato što operator ukrštanje vremenom smanjuje prostor pretraživanja grupisanjem populacije oko jednog regiona.
- Tu nastupa operator mutacije, koji za cilj ima povećavanje oblasti pretraživanja, tako što vrši slučajnu izmjenu genetičkog materijala novo-kreiranim jedinkama.
- Na taj način se stvaraju jedinke koje se udaljavaju od, u tom trenutku, optimalnog rješenja, pokušavajući slučajno pronaći novu oblast u kojoj bi se populacija trebala skoncentrisati.
- Ukoliko je populacija trenutno skoncentrisana oko pravog rješenja, onda će operator mutacije usporiti konvergenciju, ali ne i onemogućiti pronalazak pravog rješenja.

Modeli populacije

- **Generacijski model** - kada se svi roditelji iz N -te generacije zamijene potomcima i tako se kreira nova $(N+1)$ -va generacija. Stvaranjem nove populacije roditelji se zamijenjuju potomcima, nakon čega roditelji odumiru.
- Mogu se i spojiti roditeljska i generacija potomaka i **za potomke birati N najboljih**, bez obzira da li pripadaju roditeljima ili potpomcima.
- **Steady-state** model populacije kada određeni broj potomaka zamijeni isti broj roditelja iz N -te generacije i tako se kreira $(N+1)$ -va generacija.
 - Prilikom selekcije se odabira određeni broj jedinki nad kojima se vrše genetičke operacije da bi se kreirao određeni broj potomaka koji će zamijeniti isti broj već postojećih jedinki, najčešće onih sa najmanjom vrijednošću dobrote.

Ukrštanje

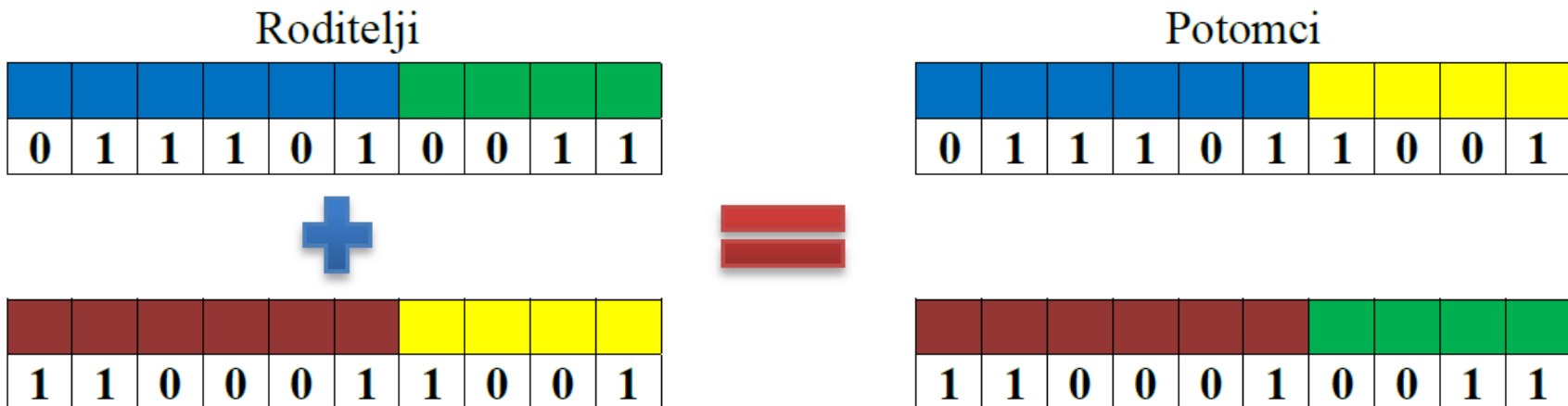
- Nakon što se selektuju odgovarajuće jedinke (roditelji), vrši se njihovo ukrštanje po određenom pravilu i stvaranje dvije nove jedinke.
- Prilikom ukrštanja dolazi do razmjene genetičkog materijala i potomci naslijeđuju svojstva svojih roditelja.
- U slučaju dobrih roditelja povećava se vjerovatnoća da i djeca budu dobra, ako ne i bolja od svojih roditelja.
- Razmjena genetičkog materijala će omogućiti konvergenciju potomaka oko jedinki koje u tom trenutku imaju najbolju fitnes funkciju.
- Postupak ukrštanja se razlikuje za različite načine prikazivanja jedinki (binarni prikaz, prikaz pomoću cijelih ili realnih brojeva, permutacija karaktera, itd.).

Ukrštanje

- Tri su standardna tipa ukrštanja za binarni prikaz rješenja. Primjenljiva su kod drugih načina prikaza. Od dva roditelja dobija se jedan ili dva potomka sačinjeni od kombinacije gena roditelja.

Ukrštanje u jednoj tački

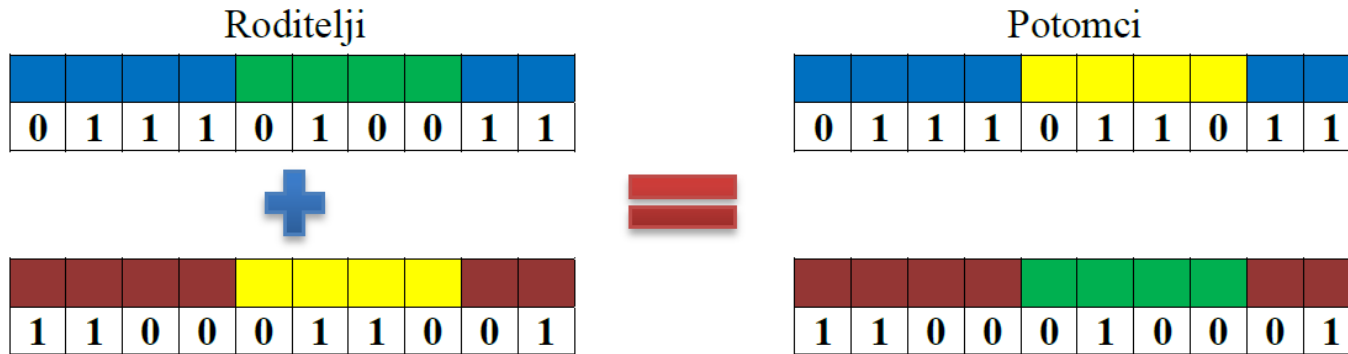
- Slučajno se odabere jedna tačka unutar hromozoma i do odabrane tačke se uzimaju geni prvog roditelja, a od odabrane tačke se uzimaju geni drugog roditelja. Na ovaj način je moguće kreirati jednog ili dva potomka.



Ukrštanje

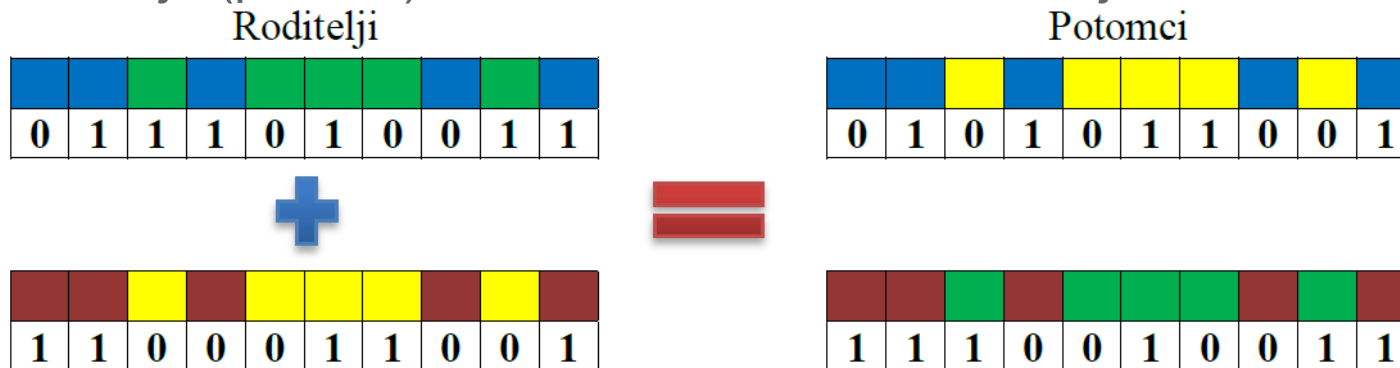
Ukrštanje u više tačaka

- Ukrštanje se može izvesti i sa više tačaka prekida od jedne. Ukoliko je broj bita u hromozomu $N+1$ tada se ukrštanje može izvesti u N tačaka.



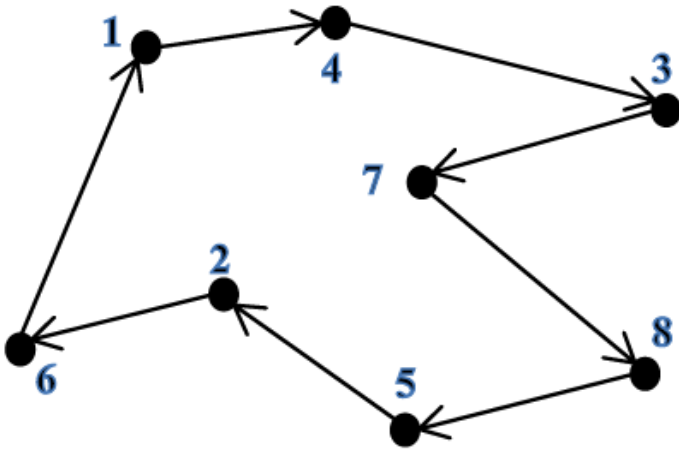
Uniformno ukrštanje

- Ukrštanje kod koga je podjednaka vjerovatnoća naslijeđivanja gena od oba roditelja ($p=50\%$) naziva se uniformno ukrštanje



Ukrštanje

- Kod problema putujućeg trgovca - prikaz pomoću permutacije karaktera

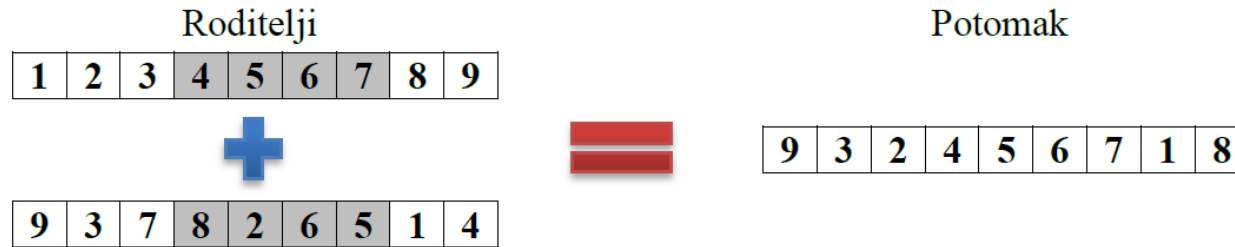


1	4	3	7	8	5	2	6
---	---	---	---	---	---	---	---

- Nije moguće uzeti jedan dio gena od prvog roditelja, drugi dio gena od drugog i nakon toga gene spojiti stvarajući potomke.
- Problem nastaje ukoliko se ista vrijednost gena ponavlja više puta na različitim pozicijama u jednom hromozomu, što ga u tom slučaju ne čini jednim od mogućih rješenja problema.
- Poželjno je ne mijenjati u potpunosti redoslijed pojavljivanja vrijednosti gena u hromozomu, već djelimično sačuvati red pojavljivanja kombinujući oba roditelja.

Ukrštanje

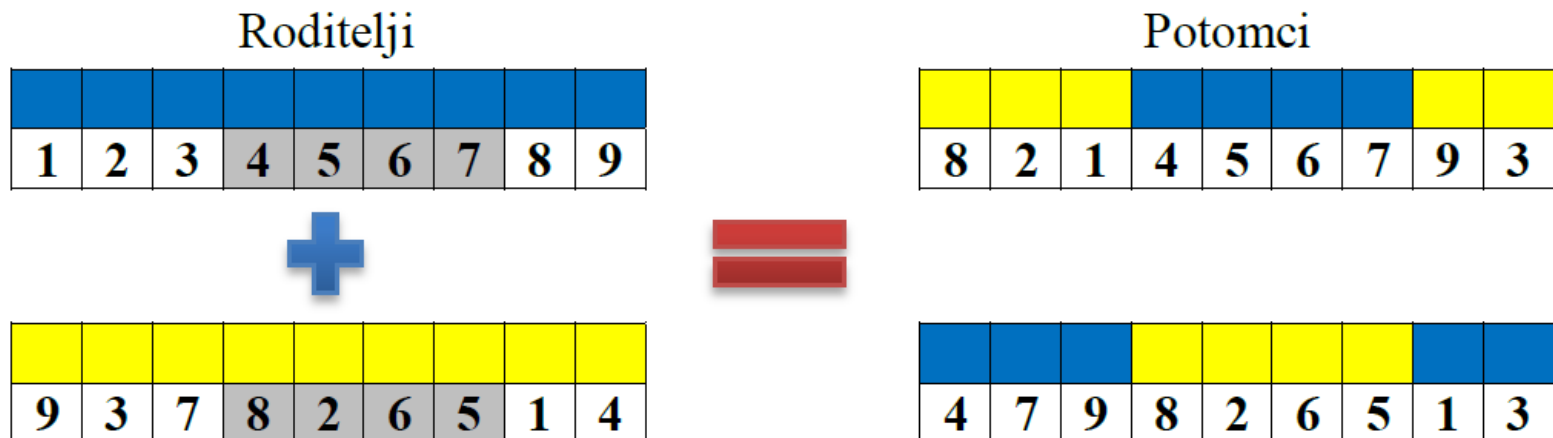
PMX ukrštanje (eng. *Partially Mapped Crossover*)



- *Korak 1* Slučajnim odabirom se odrede dvije tačke ukrštanja.
- *Korak 2:* U opsegu između te dvije tačke kopiraju se geni prvog roditelja.
- *Korak 3:* Sada se posmatraju geni drugog roditelja koji se nalaze na odabranim pozicijama prvog roditelja.
 - Na poziciji 4 drugog roditelja se nalazi vrijednost 8. Njenu poziciju u potomku je zauzela vrijednost 4 iz prvog roditelja, te se vrijednost 8 upisuje na poziciju 8, gdje se nalazi vrijednost 4 drugog roditelja. Na poziciji 5 drugog roditelja se nalazi vrijednost 2 koju treba upisati na poziciju 7, ali je tu već upisana vrijednost 7 u potomku. Tražimo gdje je upisana vrijednost 7 drugog roditelja i ukoliko je slobodna, tu upisujemo vrijednost 2.
- *Korak 4:* Kada se u potomku upišu sve vrijednosti iz određenog opsega oba roditelja preostale vrijednosti iz drugog roditelja se redom upisuju na slobodnim pozicijama u potomku.

Redno ukrštanje

- Odrede se dvije tačke ukrštanja i između njih se elementi iz prvog roditelja upisuju na identičnim pozicijama u potomku.
- Od drugog roditelja se uzimaju elementi koji nisu upisani u potomku i redom se upisuju u potomka, startujući od druge tačke ukrštanja.
- Ukoliko se dosegne zadnji element u hromozomu potomka nastavlja se od prvog elementa. Drugog potomka je moguće kreirati zamjenom mjesta roditelja.

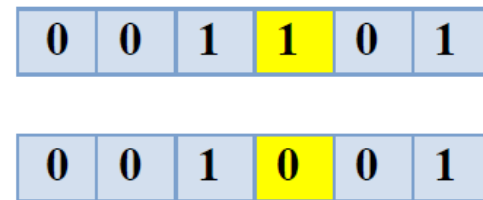
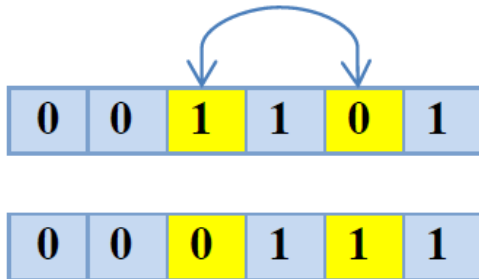


Mutacija

- Mutacija se obično vrši nakon operatora ukrštanja, sa unaprijed definisanom malom vjerovatnoćom.
- Mutacija je operator koji za stvaranje potomka koristi isključivo gene jedne jedinke. Vršiti se mutiranje gena na pojedinim pozicijama unutar jedinke, čime ona dobija nove karakteristike.
- Mutacijom se sprječava da sva potencijalna rješenja kojima algoritam raspolaže kroz generacije upadnu u neki lokalni ekstremum i time se ograniči prostor pretrage genetičkog algoritma.
- Principi realizacije zavise od načina prikazivanja jedinki i variraju od slučaja do slučaja. Najčešće se izvršava zamjenom pozicija gena ili zamjenom pojedinih gena nekom drugom vrijednošću iz dozvoljenog intervala i dodatno se može modifikovati tako da daje bolje rezultate.
- Jedan od načina modifikacije je da se hromozomi sa manjom vrijednošću dobrote izlože većoj vjerovatnoći mutacije od onih sa većom vrijednošću dobrote. Na taj način bolje jedinke čuvamo, sprječavajući da eventualno postanu lošije, dok one sa manjom vrijednošću dobrote pokušavamo učiniti boljim.

Mutacija

- Za binarni prikaz rješenja dva su najčešća principa realizacije mutacije: **zamjena pozicija gena** (eng. *swap*) i **invertovanje gena** (eng. *flip*).



- Pri realizaciji mutacije za prikaz pomoću permutacije karaktera nije više moguće posmatrati pojedinačne gene hromozoma nezavisno od ostalih.
- Nakon izvršene operacije mutacije hromozom i dalje mora predstavljati jedno od mogućih rješenja problema.
- Dakle, kada se govori o mutaciji za ovaj prikaz rješenja u razmatranje se uzimaju čitavi hromozomi, a ne pojedinačni geni što prouzrokuje da se procenat mutacije, kao i svi drugi parametri mutacije odnose na čitave hromozome.

Mutacija

- Slučajnim odabirom se biraju dva gena iz hromozoma i mijenjaju im se pozicije.



- Operator mutacije ubacivanjem takođe slučajno bira dva gena iz hromozoma i postavlja ih jedan pored drugoga, pomjerajući ostale gene u odgovarajućem smjeru.



- *Scramble* mutacija odabira dvije pozicije gena u hromozomu i svim genima između tih pozicija nasumično mijenja mjesta



- Invertujuća mutacija odabira dvije pozicije gena i pravi presjeke na tim pozicijama stvarajući tako tri odvojena niza karaktera i središnji niz invertuje tako da prvi i posljednji element mijenjaju mjesta, i tako redom.

