

Genetski algoritmi

Osnovni koncepti

Osnovni pojmovi vezani za genetske algoritme (GA)

- Spada u grupu stohastičkih **optimizacionih metoda** koje ne koriste derivacije – izvode funkcija.
- Bazirani su na konceptu **prirodne selekcije** i **evolucije**.
- Karakteristike:
 - GA predstavljaju proceduru **paralelnog pretraživanja** pogodnu za primenu u sistemima paralelnog procesiranja (obrade) podataka – značajno ubrzavanje rada;
 - GA su podjednako primenljivi u rešavanju kontinualnih i diskretnih (kombinatornih) **optimizacionih problema**;
 - GA predstavljaju **stohastičke metode**, manje podložne zapadanju u lokalni minimum (čime su, inače, opterećeni mnogi praktični optimizacioni problemi);
 - GA poseduju veoma **veliki stepen prilagodljivosti** čime se olakšava njihova primena u okviru problema strukturne i parametarske identifikacije složenih modela kao što su neuronske mreže i/ili fuzzy sistemi zaključivanja.

- GA svaku tačku parametarskog ili prostora rešenja predstavljaju preko **binarnog niza (stringa)** koji se zove **hromozom**.
- Binarni niz se vrlo često u modernijim primenama posmatra kodirano na višem nivou kao niz drugačijih elemenata npr. realnih brojeva, stringova (u suštini bilo čega)
- Svakom hromozomu je pridružena određena **“vrednost podobnosti (stepen prilagođenosti)”**.
- Umesto jedne izolovane tačke GA često koriste set (niz, skup) tačaka koji se naziva **populacija**, i koji evoluiraju u pravcu što boljeg zadovoljenja kriterijumske funkcije.
- U svakoj **generaciji** GA primenom **genetskih operatora** proizvodi novu populaciju.

- Osnovni genetski operatori su **selekcija**, **ukrštanje** i **mutacija**.
- Članovi populacije koji u **većoj meri** zadovoljavaju **kriterijumsku funkciju** imaju veću šansu da prežive i da učestvuju u budućim ukrštanjima.
- Nakon većeg broja generacija, populacija sadrži članove sa boljim (većim) **stepenom prilagođenja** (zadovoljenja kriterijumske funkcije) – što je analogno sa Darwinovim modelom evolucije preko slučajne mutacije i prirodne selekcije.
- GA se često navode kao primer **optimizacije bazirane na populaciji**, gde se popravljaju performanse cele populacije a ne samo određenih jedinki.

Osnovne komponente GA

- ◆ Osnovni elementi genetskih algoritama su:
 - ◆ Kodiranje;
 - ◆ Prilagođenost (procena pogodnosti);
 - ◆ Selekcija (odabir roditelja);
 - ◆ Ukrštanje;
 - ◆ Mutacija;
 - ◆ Elitizam.

Kodiranje

- ❑ Ovom transformacijom se tačke parametarskog prostora predstavljaju uglavnom nizovima (stringovima) **binarnih oznaka**.
- ❑ Uobičajeni postupci kodiranja binarnim nizovima su **binarno** i **sivo (Gray) kodiranje**.
- ❑ Moguće je **kodirati** pozitivne realne brojeve, ali i brojeve zapisane u formatu pomičnog zareza, negativne brojeve, diskretne vrednosti, i dr.
- ❑ Šema kodiranja predstavlja **način transformacije** specifičnog znanja vezanog za određeni problem u okvire genetskog algoritma, što ujedno predstavlja i korak koji će direktno (ključno) uticati na **uspešnost** genetskog algoritma.
- ❑ Generalna preporuka je da se i genetski operatori ukrštanja i mutacije projektuju **zajedno** sa šemom kodiranja kako bi što bolje odgovorili zahtevima pojedinačnih (specifičnih) aplikacija (problema).

Prikazivanje (opisivanje) hromozoma

■ Klasičan način predstavljanja hromozoma u okviru GA jeste **binarni vektor fiksne dužine**.

■ U slučaju ***N-dimenzionalnog*** prostora pretraživanja (parametarski prostor) svaka **jedinka (hromozom)** se sastoji od ***N* promenljivih vrednosti (koordinata)** od kojih je svaka kodirana binarnim stringom.

Binarno kodiranje

- Svaka koordinata parametarskog prostora se predstavlja **binarnim stringom** dužine D (D -dimenzionalni binarni vektor).
- Ako se promenljiva z sa kontinualnog intervala $[z_{min}, z_{max}]$ želi **konvertovati** u D-bitni string može se primeniti sledeći obrazac:

$$\left(2^D - 1\right) \frac{z - z_{min}}{z_{max} - z_{min}}$$

Primer: tačka (2,3,7) trodimenzionalnog parametarskog prostora može biti predstavljena sledećim binarnim stringom:

01001111
 └─┘ └─┘ └─┘
 2 3 7

Svaka koordinata je kodirana kao **gen** sastavljen od tri binarna bita.

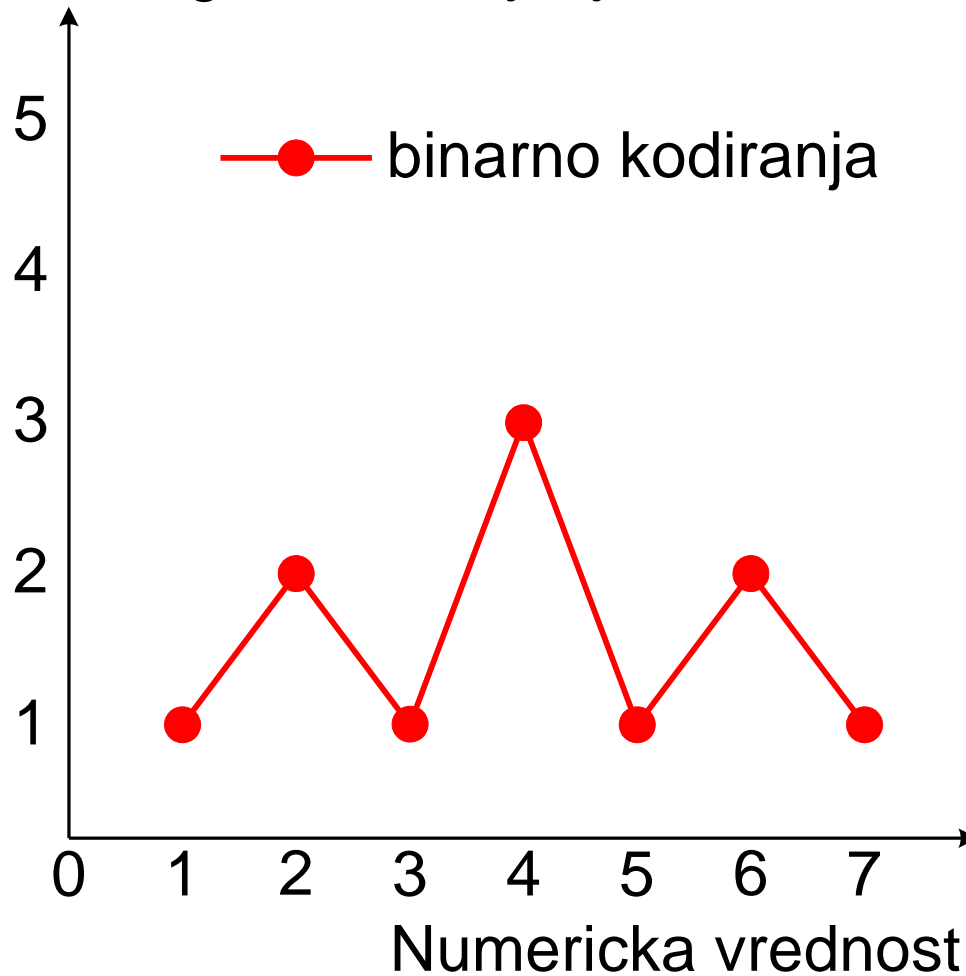
Posmatra se interval $[0, 7]$ i neka promenljiva z može poprimati celobrojne vrednosti sa tog intervala. Binarno kodiranje datih vrednosti promenljive z se može predstaviti sledećom tabelom:

Koordinata	0	1	2	3	4	5	6	7
Binarno kodiranje	000	001	010	011	100	101	110	111

Iako se binarno kodiranje veoma često koristi ono sa sobom nosi i značajan problem – **Hamming-ove litice**.

Hamming-ova litica se formira u slučaju kada dve numerički **bliske vrednosti** imaju veoma **udaljene kodne oznake**, kao što su npr. 3 i 4 iz gornje tabele. Neka je 3 optimalno rešenje problema, a 4 tekuće rešenje. Da bi se kodna oznaka promenila i postiglo optimalno rešenje potrebno je promeniti vrednost svih bitova kodne oznake (u ovom slučaju tri bita). Problem Hamming-ovih litica se može ilustrovati sledećim grafikom:

Hamming-ovo rastojanje



Problem Hamming-ovih litica se može rešiti uvođenjem **sivog (Grey) kodiranja**.

Grejevo (Sivo) kodiranje (Frank Gray, 1947)

- Nakon primene sivog kodiranja **Hamming-ovo rastojanje** između numerički susednih podataka poprima **vrednost 1**.
- **Primena** sivog kodiranja se može ilustrovati na prethodnom primeru sledećom tabelom:

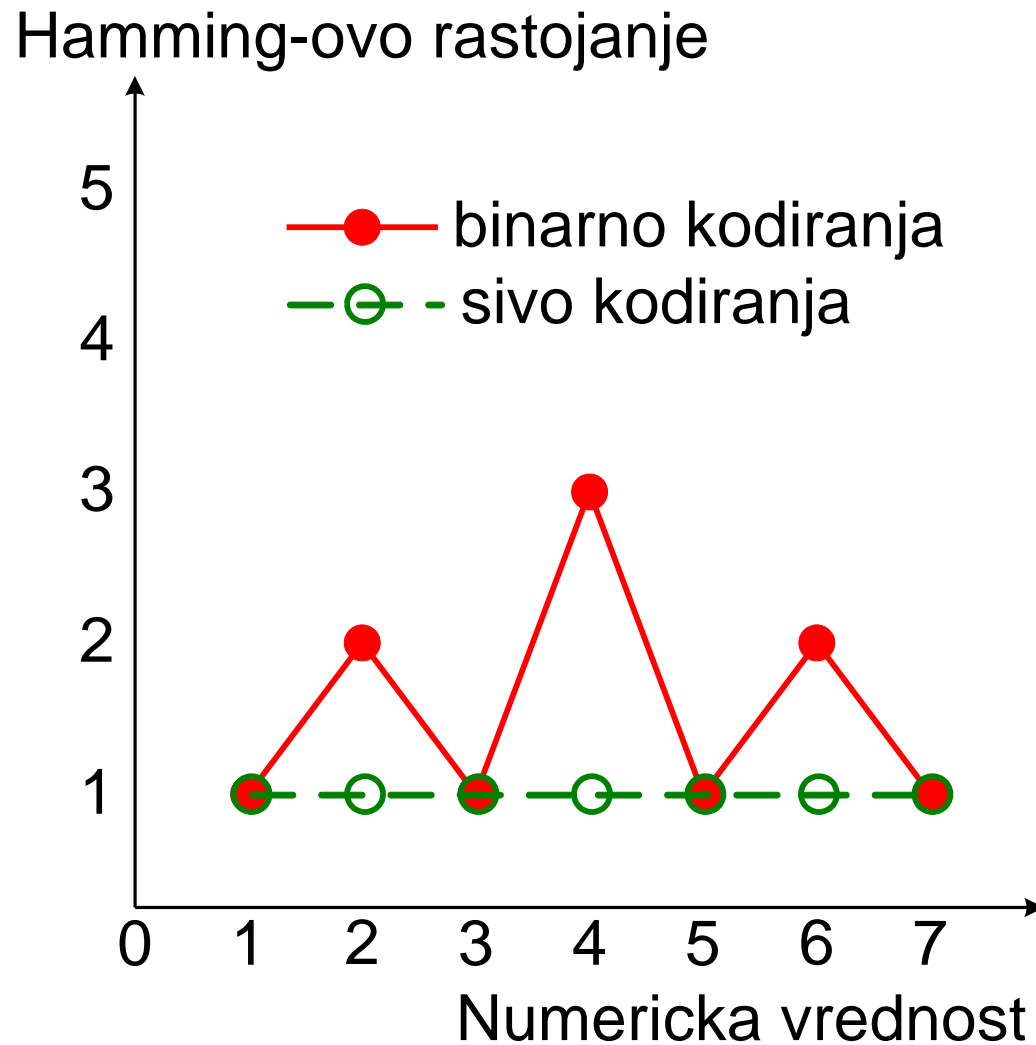
Koordinata	0	1	2	3	4	5	6	7
Binarno kodiranje	000	001	010	011	100	101	110	111
Sivo kodiranje	000	001	011	010	110	111	101	100

- Binarno kodirane oznake se mogu jednostavno transformisati u sivo kodirane primenom sledeće transformacije

$$g_1=b_1$$
$$g_k=b_{k-1}b_k^*+b_{k-1}^*b_k$$

gde je: b_k k-ti bit binarnog broja $b_1b_2...b_K$; b_1 najstariji bit; b_k^* je neb_k , $+$ znači logičko “ili”; **množenje** logičko “i”.

Na sledećem grafikonu su prikazane uporedne vrednosti Hamming-ovih rastojanja za binarno i sivo kodiranje.



Prilagođenost (procena pogodnosti)

- Prvi korak nakon formiranja generacije jeste **izračunavanje stepena prilagođenosti** svakog člana populacije.
- **Za probleme maksimizacije** stepen pogodnosti f_i , i -tog člana populacije je najčešće **vrednost kriterijumske funkcije** izračunata za tog člana (u toj tački).
- **Uobičajeno je da stepen prilagođenosti bude pozitivna vrednost**, tako da bi u slučaju kada kriterijumska funkcija nema samo pozitivne vrednosti valjalo primeniti postupak **skaliranja** i/ili **translacije**.
- Drugi postupak je upotreba **ranga** članova populacije kao stepena prilagođenosti. U ovom slučaju nije potrebno precizno izračunavanje vrednosti kriterijumske funkcije, sve dok postoji korektna informacija o rangovima članova populacije.

Selekcija

- Nakon procene stepena pogodnosti (prilagođenosti) potrebno je formirati **novu populaciju** na osnovu postojeće (trenutne) generacije.
- Selekcijom se određuju **roditelji** koji će učestvovati u produkciji naredne generacije, i taj postupak je analogan prirodnom “opstanku najprilagođenijih”.
- Uobičajeno je da se članovi biraju sa **verovatnoćom** koja je proporcionalna vrednosti njihovog stepena prilagođenosti. Jedan od načina određivanja navedene verovatnoće je dat sledećim izrazom (n je veličina populacije)

$$f_i / \sum_{k=1}^n f_k$$

- Efekat ovakvog načina selekcije jeste da se dopusti članovima sa **natprosečnim** stepenom pogodnosti da se **reprodukuju** i zauzmu mesto članova sa ispodprosečnim stepenom prilagođenosti.

Ukrštanje

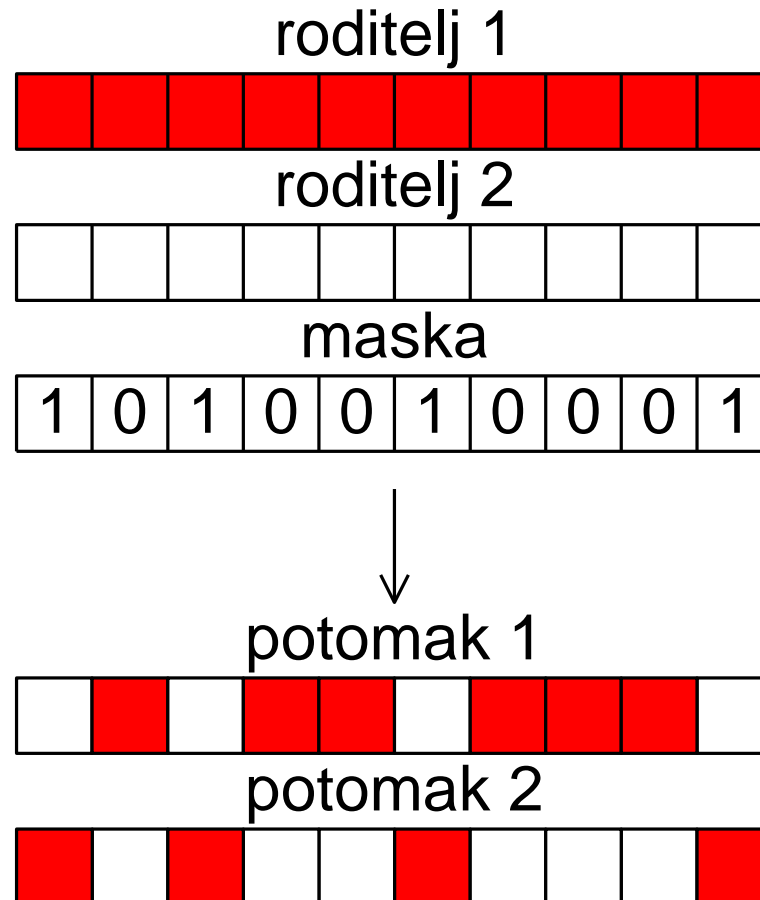
- Da bi se iskoristio **genetski potencijal** postojeće populacije koristi se **operator ukrštanja** u cilju generisanja novih hromozoma (nove generacije). Nova generacija bi trebalo da očuva dobre osobine prethodne generacije.
- Svaka grupa roditelja neće obavezno proizvesti potomstvo. Ukrštanje se primenjuje na odabrane parove roditelja čija je verovatnoća jednaka **stepenu ukrštanja** $p_c \in [0, 1]$.
- **Algoritam ukrštanja** između hromozoma \vec{C}_{n1} i \vec{C}_{n2} se može prikazati sledećim pseudokodom:
 1. Na slučajan način odrediti broj $\xi \in [0, 1]$;
 2. Ako je $\xi > p_c$, tada nema ukrštanja; u suprotnom se prelazi na korak 3;
 3. $\vec{\alpha} = \vec{C}_{n1}$; $\vec{\beta} = \vec{C}_{n2}$
 4. Izračunati masku \vec{m}
 5. Za $i = 1, \dots, N$ ako je $m_i = 1$ izvršiti zamenu genetskog materijala:
 - I. $\alpha_i = C_{n2,i}$
 - II. $\beta_i = C_{n1,i}$
 6. Vratiti potomke α i β .

- U prethodno opisanoj proceduri ukrštanja **maska** \vec{m} određuje **koji biti** roditelja će zameniti mesta i generisati potomke.
- Postoji **više operatora** ukrštanja koji na različite načine generišu masku:
 - **Uniformno** ukrštanje;
 - Ukrštanje u **jednoj tački**;
 - Ukrštanje u **dve tačke**.
- Generalno moguće je generisati ukrštanje u **N tačaka**, ukrštanje u jednoj i dve tačke su samo najčešće susretani slučajevi.

Uniformno ukrštanje

- Na **slučajan** način se kreira **maska dužine N** za svaki par jedinki izabranih za reprodukciju.
- Bit maske sa **vrednošću 1** znači da se na tom mestu vrši **zamena** bitova roditelja.
- **Algoritam** generisanja maske je opisan sledećim pseudokodom:
 1. $m_i=0$ za svako $i=1,\dots,N$;
 2. za svako $i=1,\dots,N$:
 - a) odrediti slučajnu vrednost $\xi \in [0,1]$;
 - b) ako je $\xi \leq p_x$, tada je $m_i=1$;
 3. rezultat je vektor maske \vec{m} .
- p_x je **verovatnoća ukrštanja** na svakoj poziciji u hromozomu. Ako je $p_x=0.5$, svaki bit hromozoma ima podjednaku šansu da uzme ili ne uzme učešće u ukrštanju.

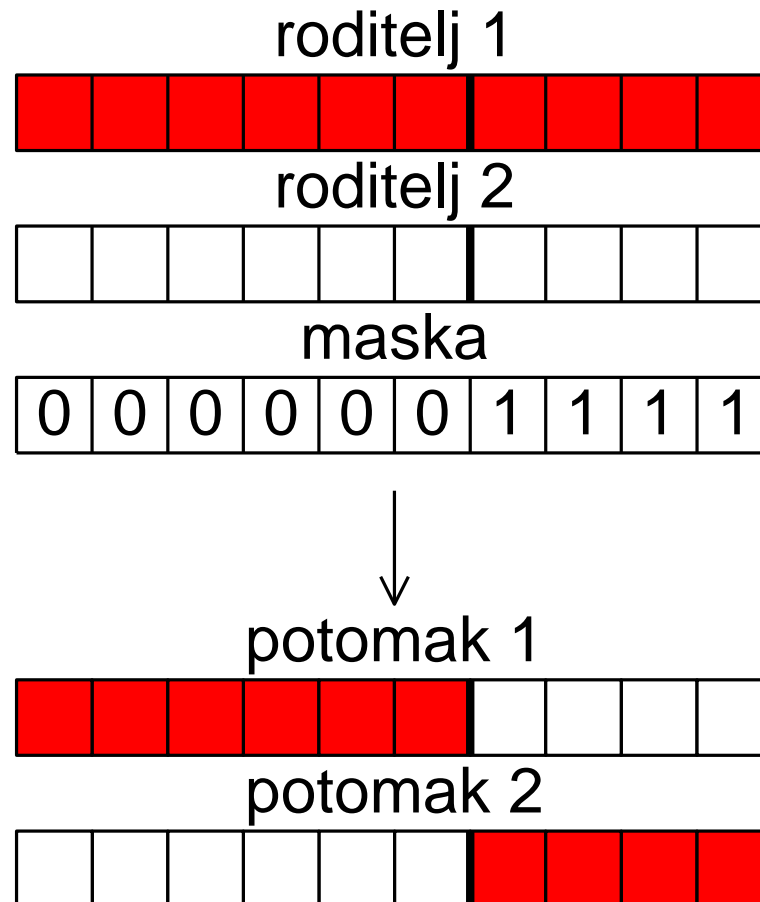
Na slici je grafički prikazan princip **uniformnog ukrštanja**.



Ukrštanje u jednoj tački

- Na **slučajan** način se bira **pozicija jednog bita**.
- Substring nakon** izabranog bita se razmenjuje između hromozoma.
- Algoritam** generisanja maske je opisan sledećim pseudokodom:
 1. određuje se slučajna veličina $\xi \in (1, N-1)$
 2. $m_i = 0$ za svako $i = 1, \dots, N$;
 3. za svako $i = \xi + 1, \dots, N$ je $m_i = 1$;
 4. rezultat je vektor maske \vec{m} .

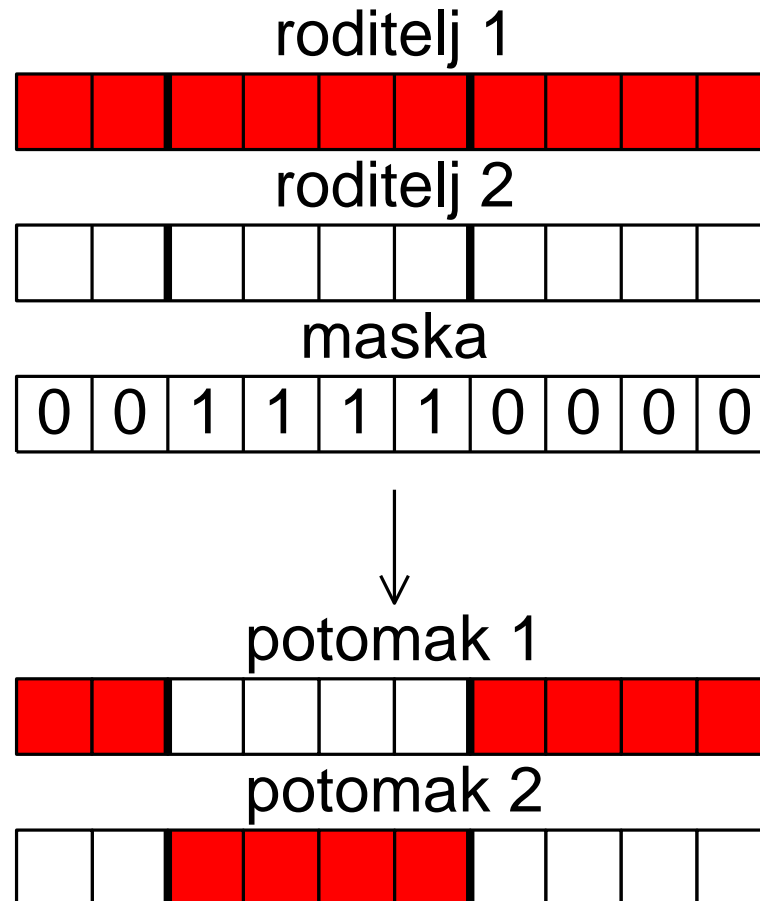
Na slici je grafički prikazan princip **ukrštanja u jednoj tački.**



Ukrštanje u dve tačke

- Na **slučajan** način se bira **pozicija dva bita**.
- Substring između** dva izabrana bita se razmenjuje između hromozoma.
- Algoritam** generisanja maske je opisan sledećim pseudokodom:
 - određuju se dve slučajne veličine $\xi_1, \xi_2 \in (1, N)$
 - $m_i = 0$ za svako $i = 1, \dots, N$;
 - za svako $i = \xi_1, \dots, \xi_2$ je $m_i = 1$;
 - rezultat je vektor maske \vec{m} .

Na slici je grafički prikazan princip **ukrštanja u dve tačke**.



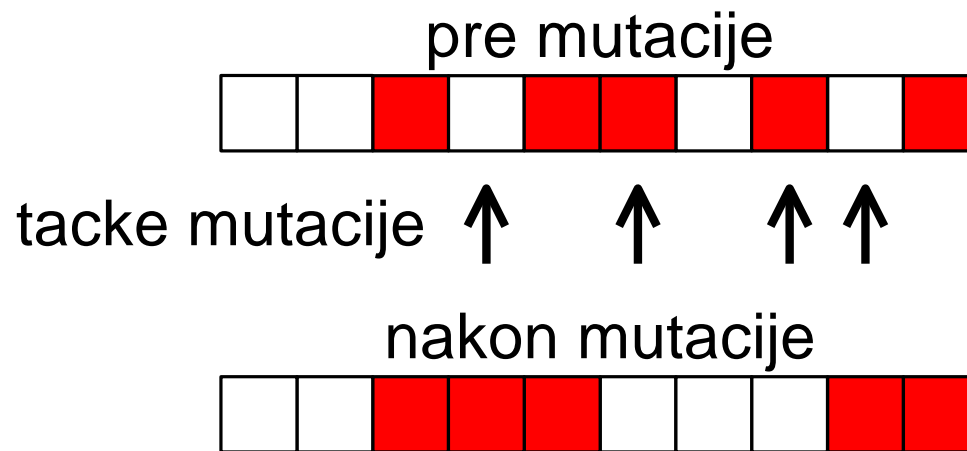
Mutacija

- Vrší se u cilju unošenja **novog** genetskog materijala u već postojeće jedinke (hromosome).
- Primenjuje se kada populacija **ne sadrži** kodirane sve informacije neophodne za rešavanje konkretnog problema.
- Mutacija se vrši sa određenom verovatnoćom **p_m** koja se naziva **stepen mutacije**.
- Operator mutacije može da **spreči** populaciju da konvergira ka i završi u **lokalnom minimumu**.
- Stepen mutacije se obično drži na **niskom nivou** da se dobri hromozomi dobijeni ukrštanjem ne bi izgubili.
- Ako se stepen mutacije drži na **visokom nivou**, karakteristike GA se približavaju algoritmu **slučajnog pretraživanja**.
- Često primenjivani algoritmi mutacije su:
 - **slučajna** mutacija;
 - **uređena** mutacija.

Slučajna mutacija

- **pozicija bita** koji se mutira se bira na **slučajni način**, nakon čega se vrši operacija negacije nad vrednošću tog bita.
- **Algoritam** slučajne mutacije je prikazan sledećim pseudokodom:
 1. za svako $i=1, \dots, N$:
 - a) odrediti slučajnu veličinu $\xi \in (0, 1)$
 - b) ako je $\xi \leq p_m$ tada je $C_{n,i} = \overline{C_{n,i}}$, gde je $\overline{C_{n,i}}$ komplement od $C_{n,i}$

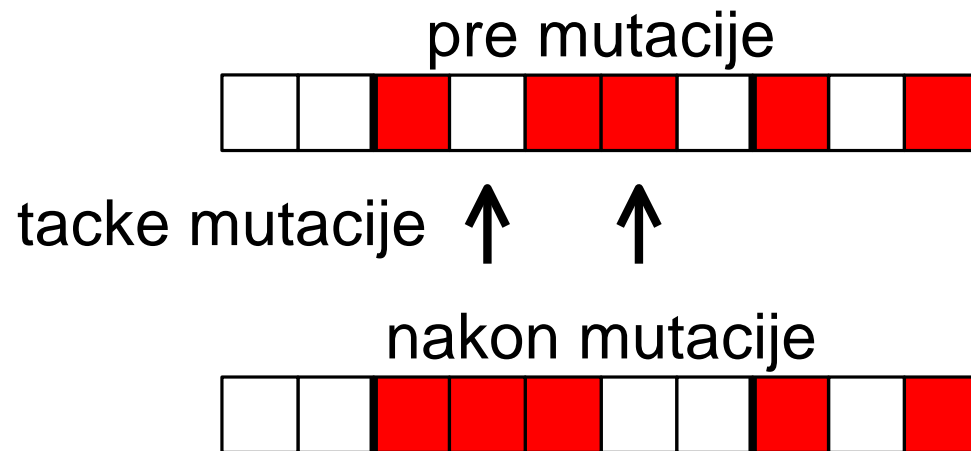
Na slici je grafički prikazan princip **slučajne mutacije**.



Uređena mutacija

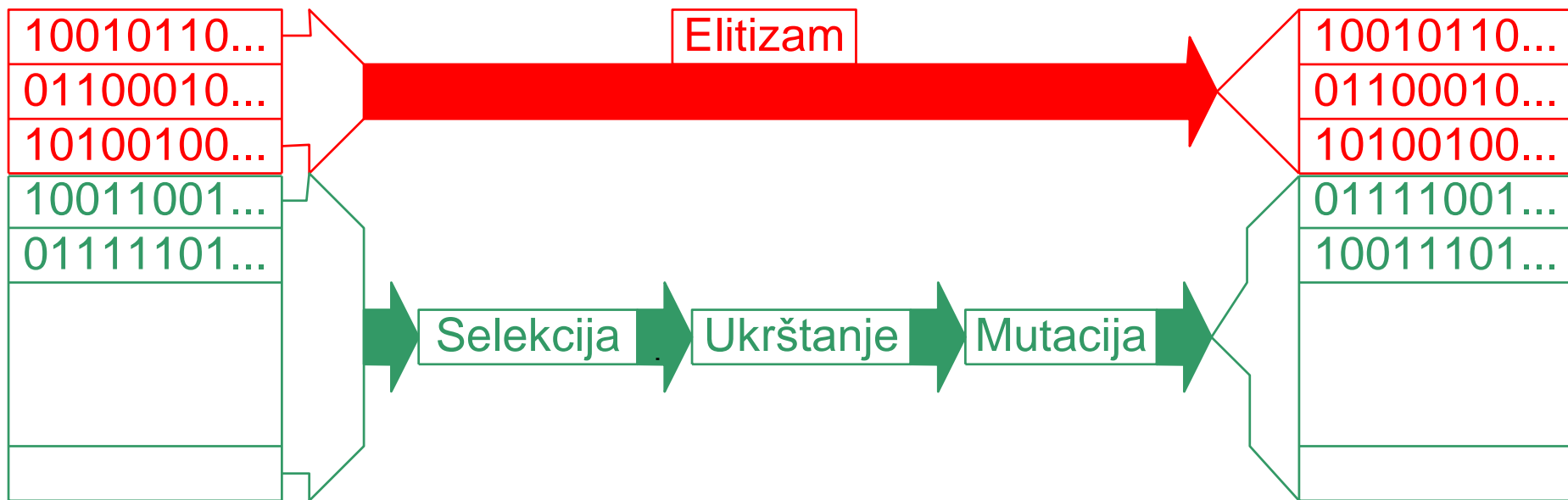
- ✱ Na **slučajni način** se bira **pozicija dva bita**.
- ✱ Nakon toga se vrši mutacija samo onih **bitova** koji se nalaze **između** prethodno određenih granica.
- ✱ **Algoritam** uređene mutacije je prikazan sledećim pseudokodom:
 1. odrediti dve slučajne veličine $\xi_1, \xi_2 \in (1, N)$
 2. za svako $i = \xi_1, \dots, \xi_2$:
 - a) odrediti slučajnu veličinu $\xi \in (0, 1)$
 - b) ako je $\xi \leq p_m$ tada je $C_{n,i} = \overline{C}_{n,i}$

Na slici je grafički prikazan princip **uređene mutacije**.



Elitizam

- Pošto je GA ipak veštački i komandovani sistem “evolucije” jedne populacije, u njegovom okviru je moguće izvršiti **izbor** određenog broja “**najboljih**” jedinki, i direktno ih preneti u narednu generaciju.
- Ovakav postupak se naziva **elitizam**, i grafički je prikazan na sledećoj slici:



KRAJ