

Problem trgovačkog putnika i genetski algoritmi

Daniel Silađi

Zašto?

- »» Zato što Problem trgovačkog putnika spada u klasu NP problema, koji još uvek nisu rešeni u potpunosti!

P ili NP?

P problemi:

- Problemi koji se mogu rešiti u polinomnom vremenu, tj, čiji najbolji algoritmi rade u složenosti $n^{O(1)}$ (na determinističkim Tjuringovim mašinama)
- Ovakvi problemi se smatraju rešivima
- Podklasa: P kompletni problemi (svi problemi iz P klase se mogu svesti na njih u polinomnom vremenu)

NP problemi:

- Problemi za koje se ne zna da li postoji rešenje u polinomnom vremenu
- Primeri: TSP, bojenje grafova, hamiltonova putanja,...
- Podklase:
 - NP tvrdi (hard): teški bar koliko i najteži NP, ali možda i teži
 - NP kompletni: svi NP problemi se mogu svesti na njih u polinomnom vremenu

Pitanje (za \$7.000.000): da li je $P=NP$?

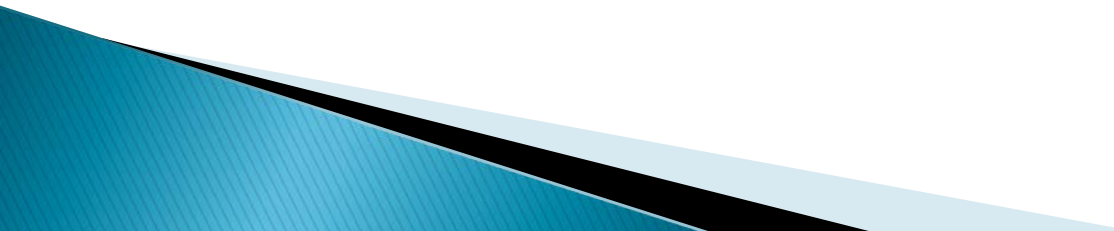


Ako jeste, onda to ima veoma značajne posledice...

Dok ne znamo odgovor, moramo se dovijati raznoraznim metodama (npr. genetskim algoritmima) da bi rešili NP probleme (TSP)

Problem trgovačkog putnika

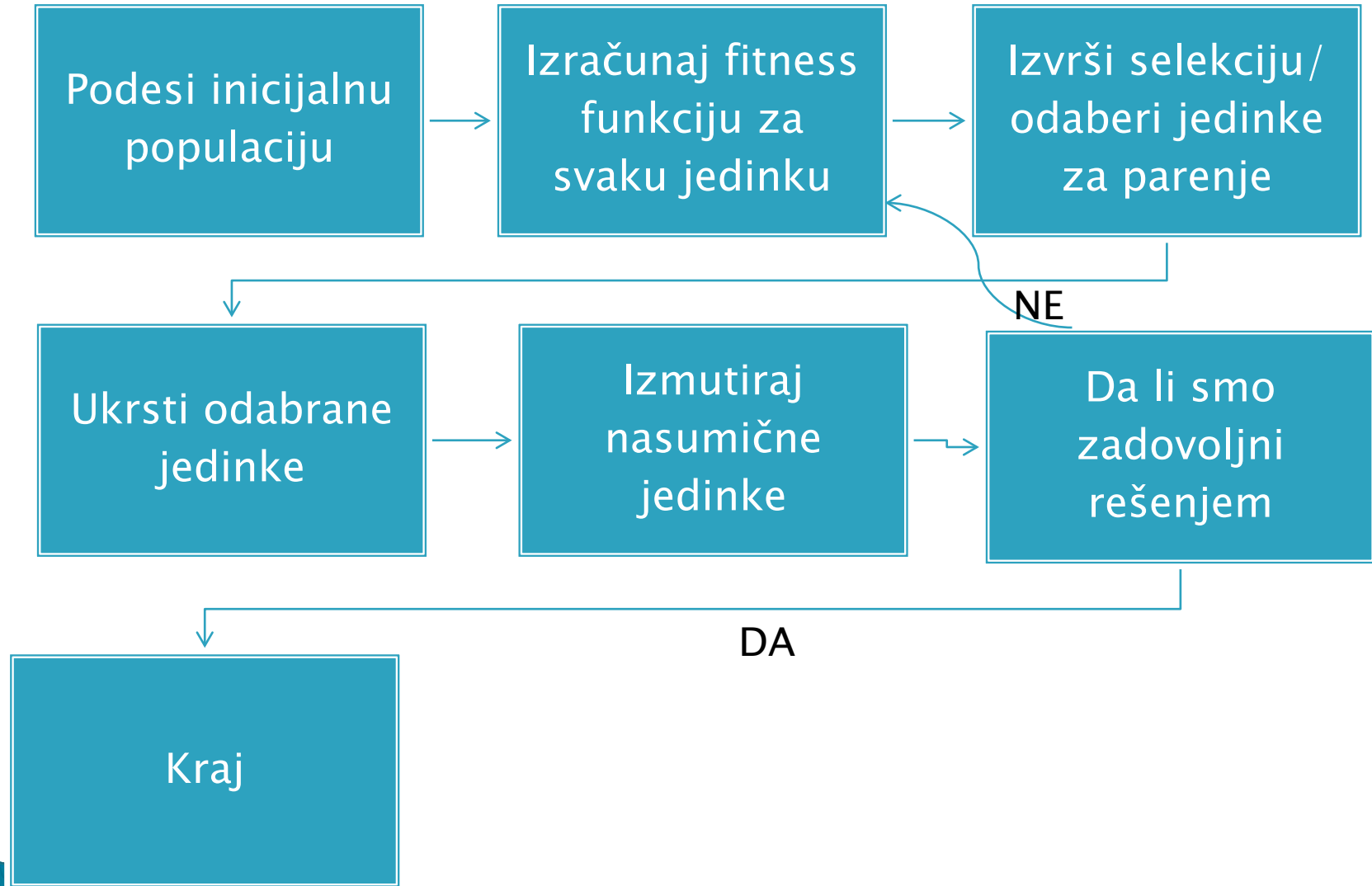
- Engleski: *Travelling salesman problem (TSP)*
- Definicija: ako je dat kompletan težinski graf, naći najkraću putanju koja prolazi kroz sve čvorove grafa (Hamiltonovu putanju)
- Tačnije, postoje 2 verzije problema:
 - “Decision problem”: Da li postoji tura kraća od L (tada je problem NP kompletan)
 - “Constructive problem”: Koja je najkraća tura (tada je problem NP hard) (ja ću se fokusirati na ovaj problem)
- U literaturi se prvi put pojavljuje 1832, a W. R. Hamilton ga opisuje matematički sredinom XIX veka.

- Pošto je potrebno poređati gradove, brute force algoritmu bi trebalo $n!$ koraka da nadje optimalan put
 - Zato su razvijene razne heuristike sa prilično malim odstupanjem od optimalnog rešenja
 - Jedan primer takvih algoritama su genetski algoritmi
 - Pošto je termin “genetski algoritam” prilično širok, a ni samo polje nije puno istraženo, verujem da tu ima dosta prostora za napredak
- 

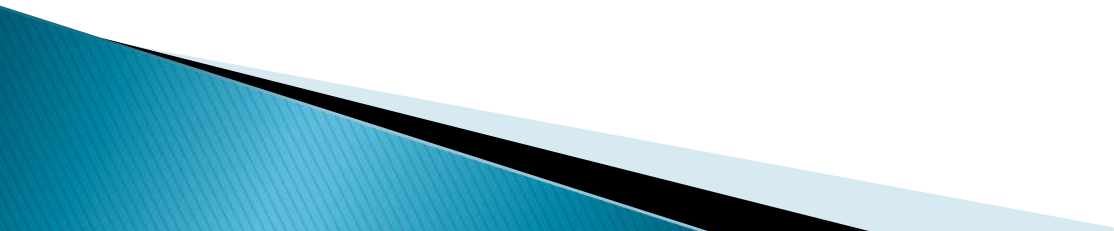
Genetski algoritmi

- Porodica algoritama inspirisana Darwinovom teorijom evolucije
- Imaju za cilj rešavanje problema kombinatorne optimizacije, tj problema u kojima se traži maksimum neke funkcije
- Koriste se kad je prostor pretraživanja prevelik, a nama nije potrebno optimalno rešenje, već smo zadovoljnim i nekim približnim => idealno (je delovalo) za TSP

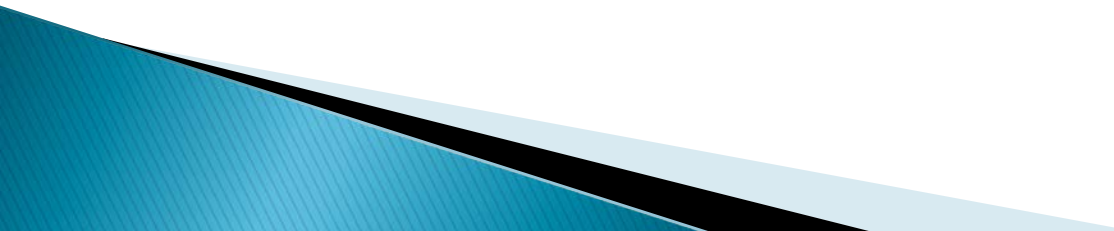
Princip rada genetskog algoritma



Terminologija

- Jedinka – jedno rešenje u pretraživačkom prostoru
 - Populacija – skup jedinki
 - Fitness–funkcija – funkcija kojom se određuje kvalitet neke jedinke/rešenja
 - Selekcija – odabir jedinki koje će opstati, razmnožavati se, ili umreti
 - Ukrštanje – kombinovanje najboljeg genetskog materijala iz 2 jedinke
 - Mutacija – menjanje nekog gena u rešenju
- 

Primena na TSP

- Genom – neka permutacija gradova (tačaka)
 - Fitness – funkcija – dužina puta
 - Konstrukcija početnog rešenja: greedy
 - Biram grad koji je najbliži prethodno odabranom gradu i nije još odabran
 - Algoritam nije optimalan, ali daje zadovoljavajuća rešenja (do 2x duži putevi)
 - Radi brzo, što je neophodno jer se na početku pravi 1000 jedinki
- 

□ Selekcija – 2 vrste:

- Rulet selekcija

- Najjednostavnija
- Verovatnoća da neka jedinka bude odabrana je proporcionalna njenom fitness-u, tj

$$p_i = \frac{f_i}{\sum_{j=1}^n f_j}$$

- Ipak, neke jedinke dobijaju preveliku prednost u selekciji

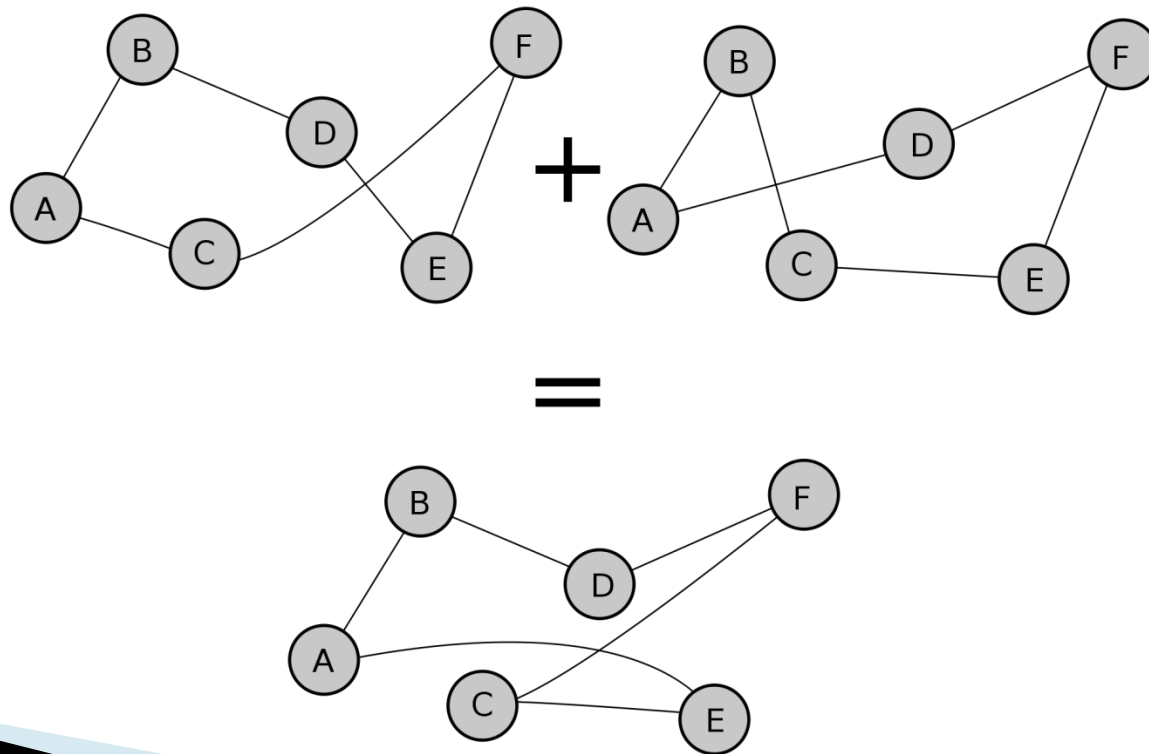
- Selekcija zasnovana na rangu

- Verovatnoća da neka jedinka bude odabrana je proporcionalna njenoj poziciji u listi sortiranoj po fitness-ima odgovarajućih jedinki

$$p_i = \frac{r_i}{\sum_{j=1}^n r_j}$$

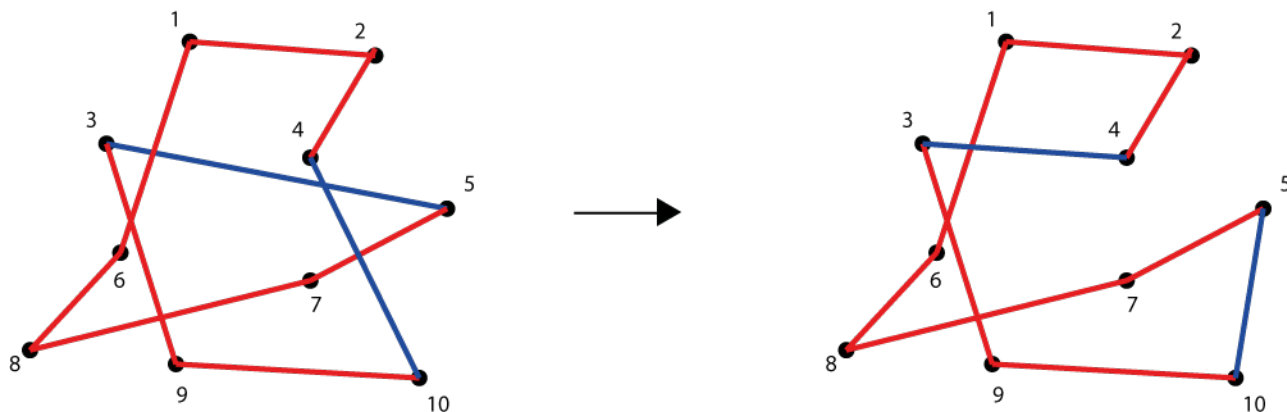
□ Ukrštanje

- Jedan od najtežih i najspecifičnijih delova problema
- Pošto su pojedinačna rešenja u stvari permutacije, ne mogu da koristim uobičajene metode jednopozicionog ili dvopozicionog ukrštanja
- Zato sam koristio „Edge recombination crossover“



□ Mutacija – 2 vrste

- 2-opt – zamenim svake 2 grane koje se ukrštaju

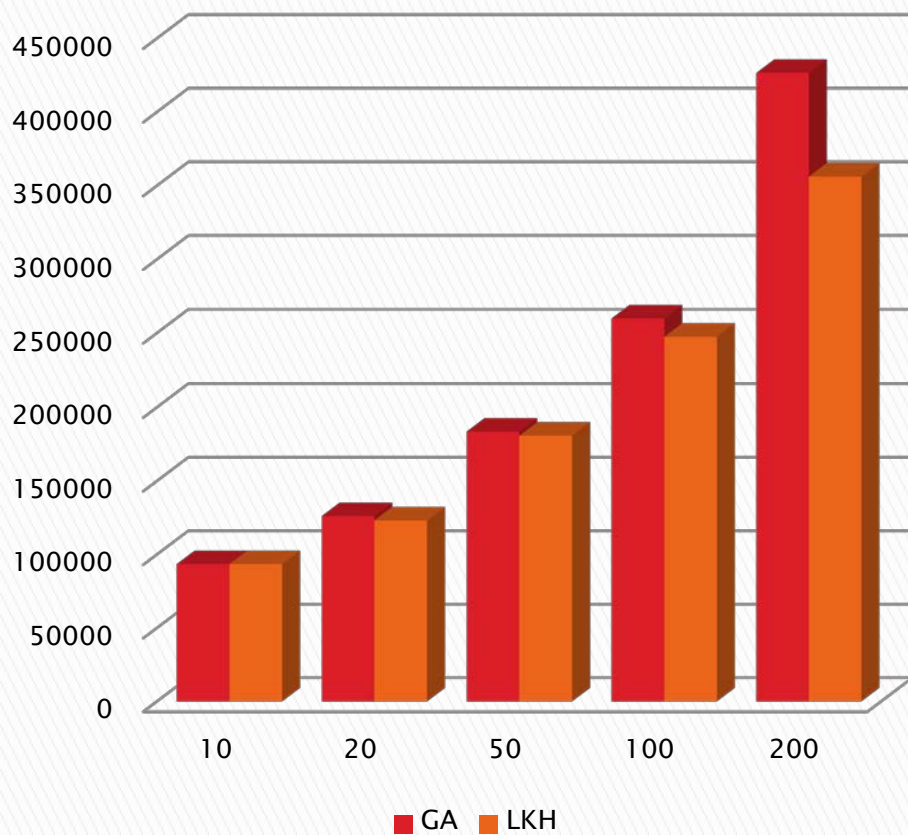


- Greedy – zamenim 2 random grada i gledam da li je to bolje

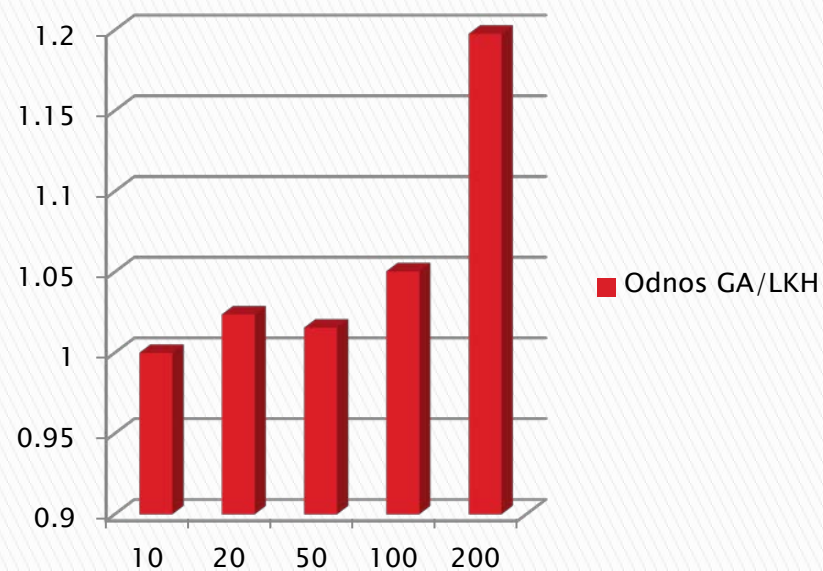
□ Kriterijum završetka: kada razlika fitnessa padne ispod neke vrednosti (to znači da smo dostigli optimum, bilo lokalni, bilo globalni)

Rezultati

Dužine puteva



Odnos GA/LKH



HVALA NA PAŽNJI!

