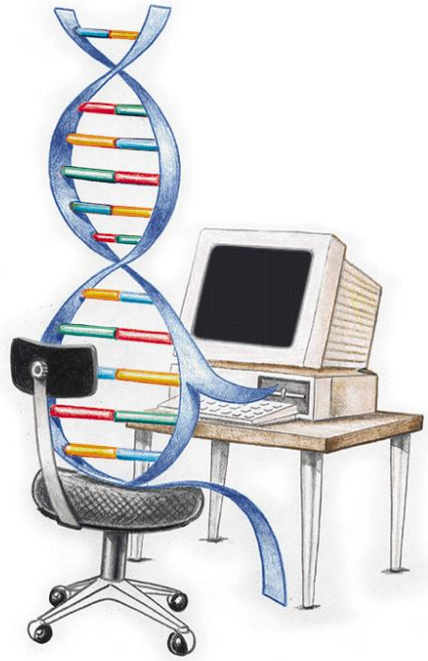
Advanced Search
Genetic algorithm

Yingyu Liang
yliang@cs.wisc.edu
Computer Sciences Department
University of Wisconsin, Madison

[Based on slides from Jerry Zhu, Andrew Moore http://www.cs.cmu.edu/~awm/tutorials]

GENETIC ALGORITHM



http://www.genetic-programming.org/

Evolution

- Survival of the fittest, a.k.a. natural selection
- Genes encoded as DNA (deoxyribonucleic acid), sequence of bases: A (Adenine), C (Cytosine), T (Thymine) and G (Guanine)

Evolution

- Survival of the fittest, a.k.a. natural selection
- Genes encoded as DNA (deoxyribonucleic acid), sequence of bases: A (Adenine), C (Cytosine), T (Thymine) and G (Guanine)
- The chromosomes from the parents exchange randomly by a process called **crossover**. Therefore, the offspring exhibit some traits of the father and some traits of the mother.
  - Requires genetic diversity among the parents to ensure sufficiently varied offspring

Evolution

- Survival of the fittest, a.k.a. natural selection
- Genes encoded as DNA (deoxyribonucleic acid), sequence of bases: A (Adenine), C (Cytosine), T (Thymine) and G (Guanine)
- The chromosomes from the parents exchange randomly by a process called **crossover**. Therefore, the offspring exhibit some traits of the father and some traits of the mother.
  - Requires genetic diversity among the parents to ensure sufficiently varied offspring
- A rarer process called **mutation** also changes the genes (e.g. from cosmic ray).
  - Nonsensical/deadly mutated organisms die.
  - Beneficial mutations produce “stronger” organisms
  - Neither: organisms aren’t improved.

Natural selection

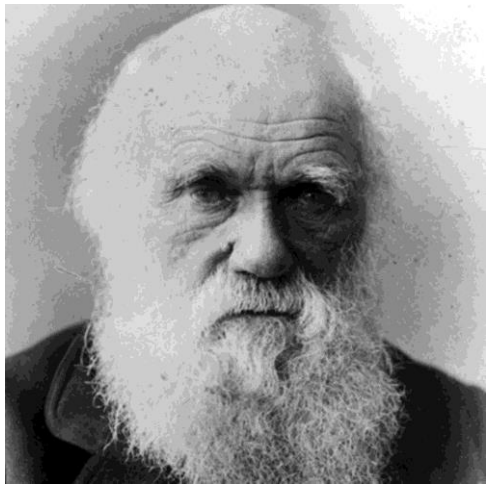
- Individuals compete for resources
- Individuals with better genes have a larger chance to produce offspring, and vice versa

Natural selection

- Individuals compete for resources
- Individuals with better genes have a larger chance to produce offspring, and vice versa
- After many generations, the population consists of lots of genes from the superior individuals, and less from the inferior individuals
- Superiority defined by fitness to the environment

Evolution and Natural Selection

- Popularized by Darwin
- Mistake of Lamarck: environment does not force an individual to change its genes

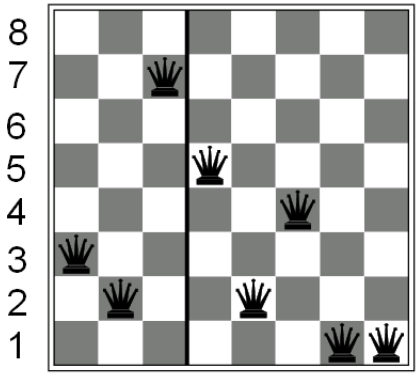


Genetic algorithm

- Yet another AI algorithm based on real-world analogy
- Yet another heuristic stochastic search algorithm

Genetic algorithm

- Yet another AI algorithm based on real-world analogy
- Yet another heuristic stochastic search algorithm
- Each state *s* is called an individual. Often (carefully) coded up as a string.



(3 2 7 5 2 4 1 1)

- The score *f*(*s*) is called the fitness of *s*. Our goal is to find the global optimum (fittest) state.
- At any time we keep a fixed number of states. They are called the population. Similar to beam search.

Individual encoding

- The “DNA”
- Satisfiability problem

What is the individual encoding scheme?

$A \vee \neg B \vee C$
$\neg A \vee C \vee D$
$B \vee D \vee \neg E$
$\neg C \vee \neg D \vee \neg E$
$\neg A \vee \neg C \vee E$

Individual encoding

- The “DNA”
- Satisfiability problem

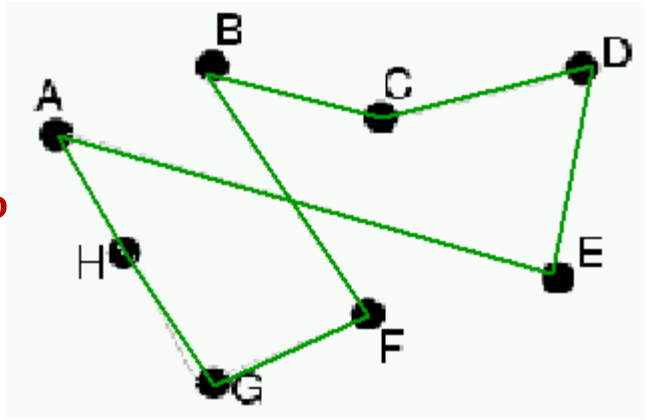
  (A B C D E) = (T F T T T)

$A \vee \neg B \vee C$
$\neg A \vee C \vee D$
$B \vee D \vee \neg E$
$\neg C \vee \neg D \vee \neg E$
$\neg A \vee \neg C \vee E$

Individual encoding

- The “DNA”
- TSP

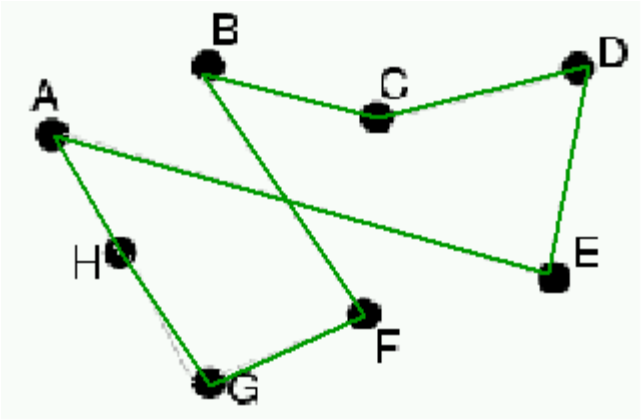
What is the individual encoding scheme?



Individual encoding

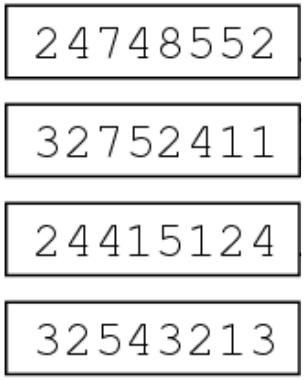
- The “DNA”
- TSP

A-E-D-C-B-F-G-H-A



Genetic algorithm

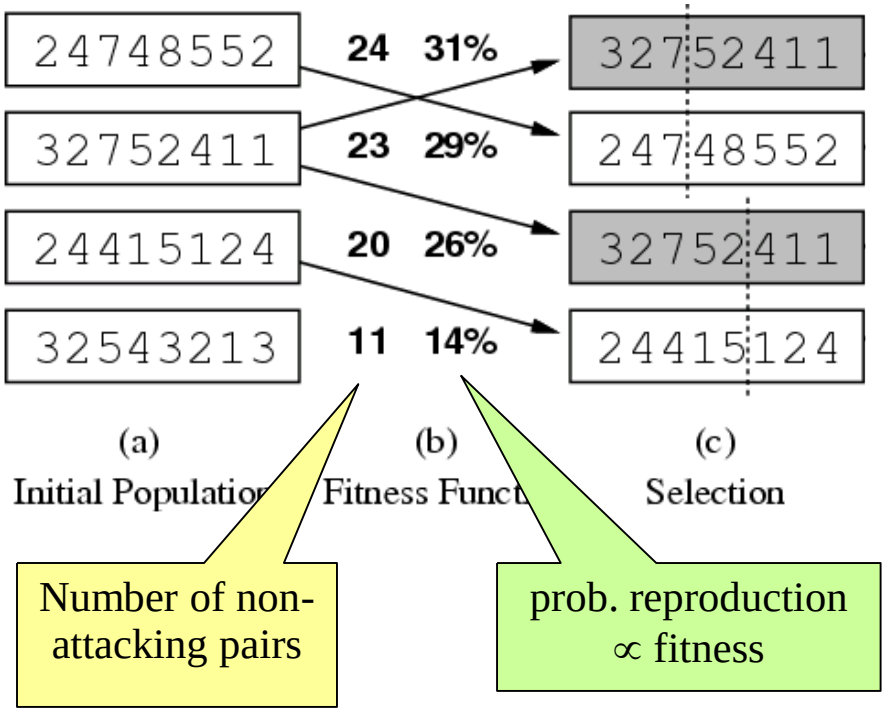
- **Genetic algorithm**: a special way to generate neighbors, using the analogy of cross-over, mutation, and natural selection.



(a)
Initial Population

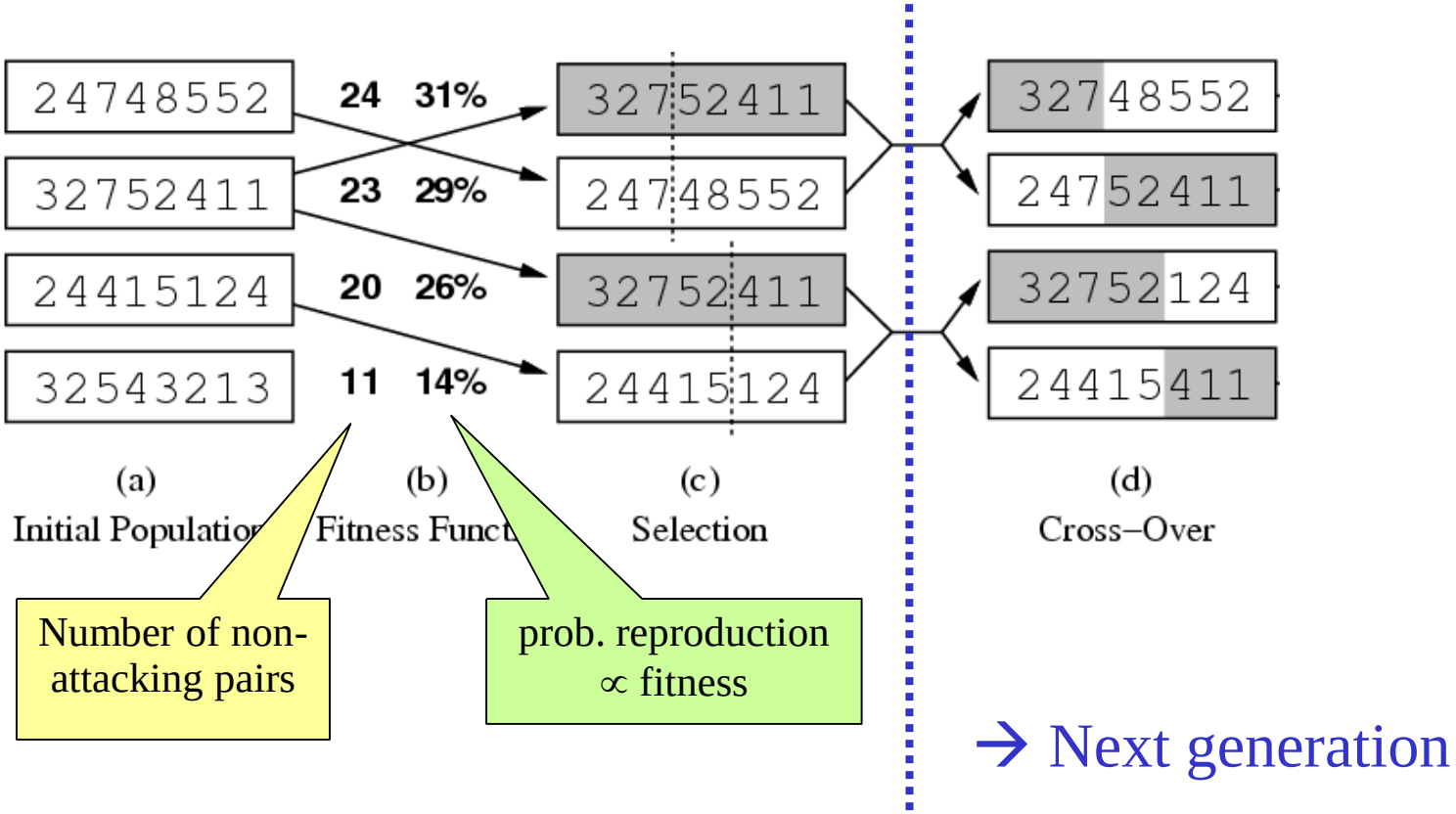
Genetic algorithm

- **Genetic algorithm**: a special way to generate neighbors, using the analogy of cross-over, mutation, and natural selection.



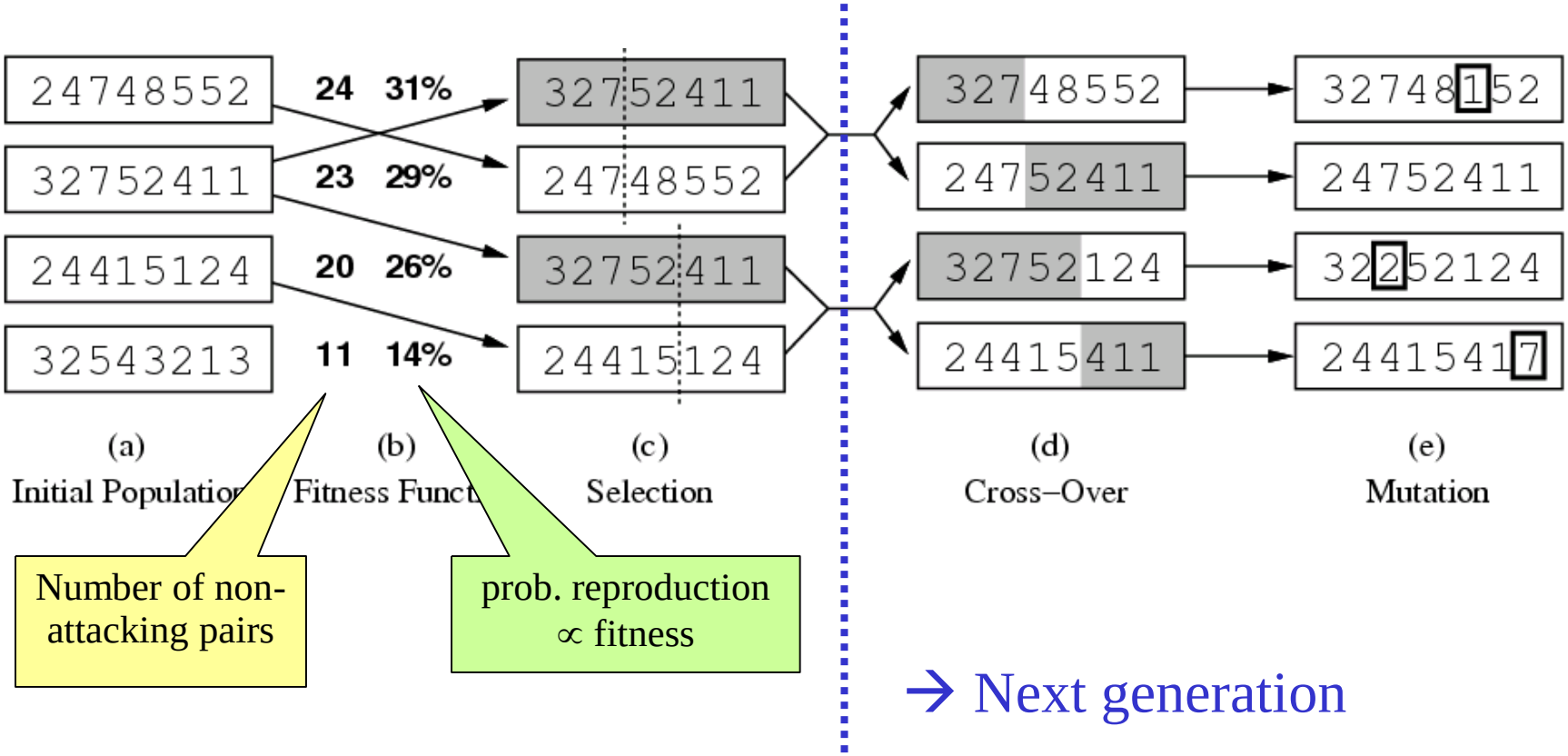
Genetic algorithm

- **Genetic algorithm**: a special way to generate neighbors, using the analogy of cross-over, mutation, and natural selection.



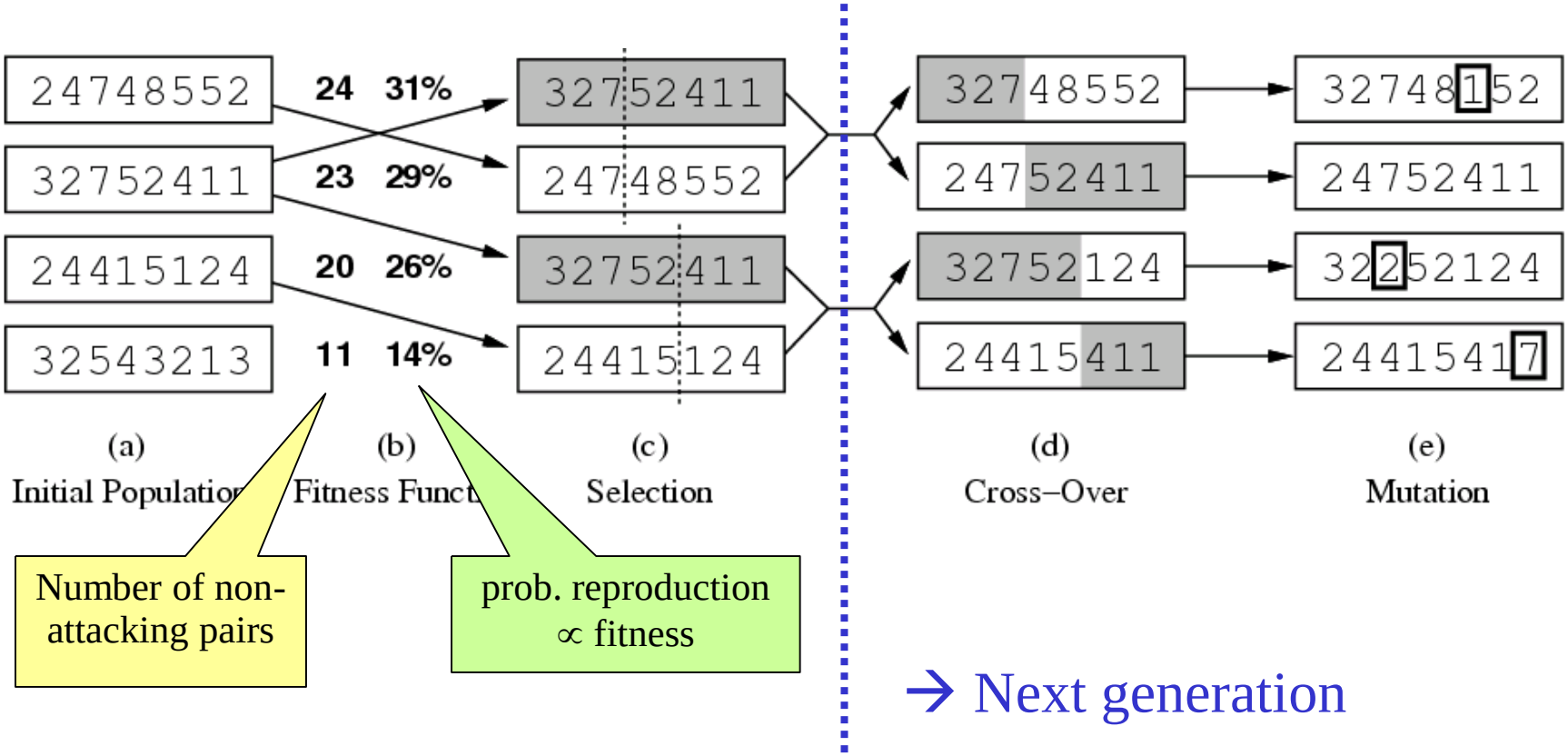
Genetic algorithm

- **Genetic algorithm**: a special way to generate neighbors, using the analogy of cross-over, mutation, and natural selection.



Genetic algorithm

- **Genetic algorithm**: a special way to generate neighbors, using the analogy of cross-over, mutation, and natural selection.



How about for the SAT and TSP problems?

Genetic algorithm (one variety)

1. Let s_1, …, s_N be the current population
2. Let $p_i = f(s_i) / \Sigma_j f(s_j)$ be the reproduction probability
3. FOR $k = 1$; $k<N$; $k+=2$
   - parent1 = randomly pick according to p
   - parent2 = randomly pick another
   - randomly select a crossover point, swap strings of parents 1, 2 to generate children $t[k]$, $t[k+1]$
4. FOR $k = 1$; $k<=N$; $k++$
   - Randomly mutate each position in $t[k]$ with a small probability (mutation rate)
5. The new generation replaces the old: { s }←{ t }. Repeat.

Proportional selection

- $p_i = f(s_i) / \Sigma_j f(s_j)$
- $\Sigma_j f(s_j) = 5+20+11+8+6=50$
- $p_1=5/50=10\%$

Individual	Fitness	Prob.
A	5	10%
B	20	40%
C	11	22%
D	8	16%
E	6	12%

Variations of genetic algorithm

- Parents may survive into the next generation
- Use ranking instead of $f(s)$ in computing the reproduction probabilities.
- Cross over random bits instead of chunks.
- Optimize over sentences from a programming language. Genetic programming.
- …

Genetic algorithm issues

- State encoding is the real ingenuity, not the decision to use genetic algorithm.
- Lack of diversity can lead to premature convergence and non-optimal solution
- Not much to say theoretically
  - Cross over (sexual reproduction) much more efficient than mutation (asexual reproduction).
- Easy to implement.
- Try hill-climbing with random restarts first!