CIS 678 Machine Learning

Genetic Algorithms

Plan

- Genetic Algorithms
- Travelling Salesman Problem (TSP)

Genetic Algorithms

- Mainly used to solve **search** and **optimization** problems

Genetic Algorithms

- Mainly used to solve **search** and **optimization** problems
- A subclass of **evolutionary algorithms**

Genetic Algorithms

- Mainly used to solve **search** and **optimization** problems
- A subclass of **evolutionary algorithms**
- Inspired by **Charles Darwin's** theory of natural evolution

Genetic Algorithms

- Mainly used to solve **search** and **optimization** problems
- A subclass of **evolutionary algorithms**
- Inspired by **Charles Darwin's** theory of natural evolution
- **Survival of the fittest!**

GA Algorithm (pseudocode)

```
START
Generate the initial population
Compute fitness
REPEAT
    Selection
    Crossover
    Mutation
    Compute fitness
UNTIL population has converged
STOP
```

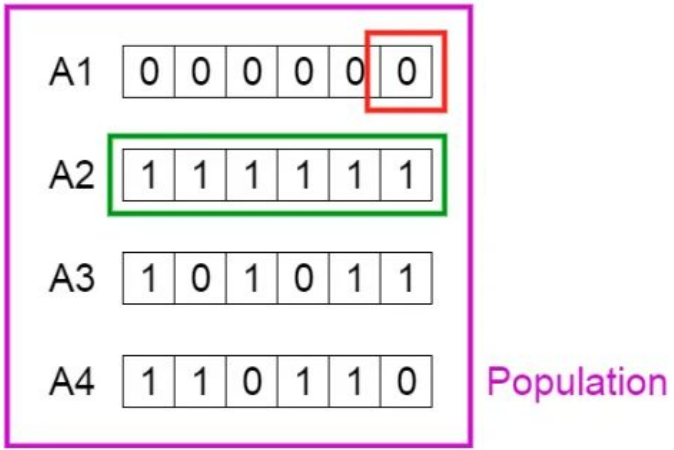
- The sequence of phases is repeated to produce individuals in each **new generation** which are better than the previous generation.

ref

Terminologies/concepts

Population

- Initial population

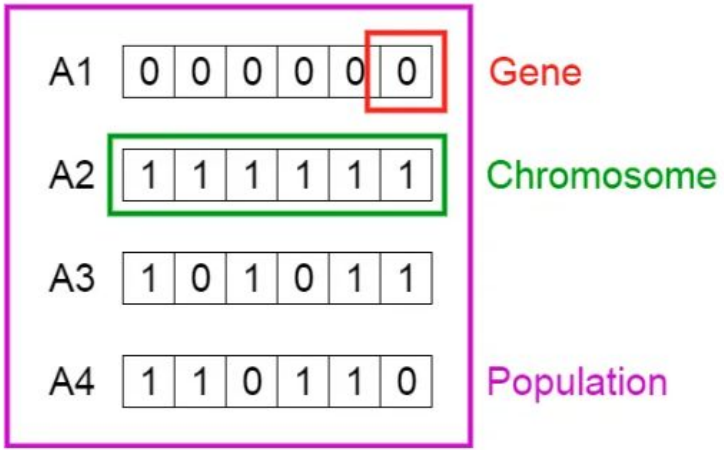


Population, Chromosomes and Genes

ref

Population

- Initial population
- **Gene & Chromosome**

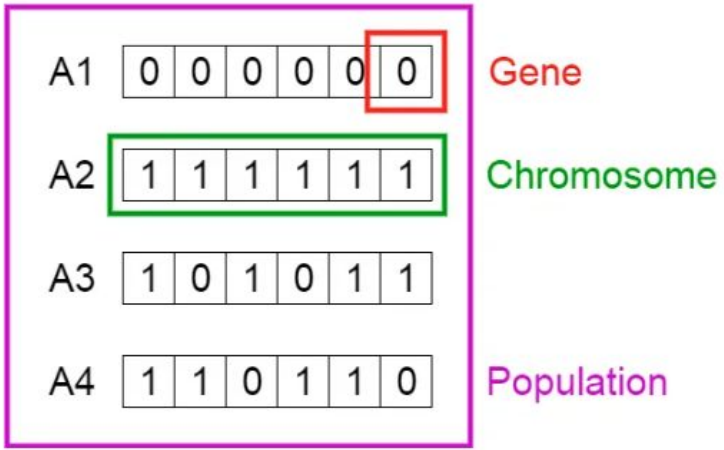


Population, Chromosomes and Genes

ref

Fitness function

- The fitness function determines how fit an individual is (the ability of an individual to compete with other individuals).
- It gives a fitness score to each individual.
- The probability that an individual will be selected for reproduction is based on its fitness score.



Population, Chromosomes and Genes

ref

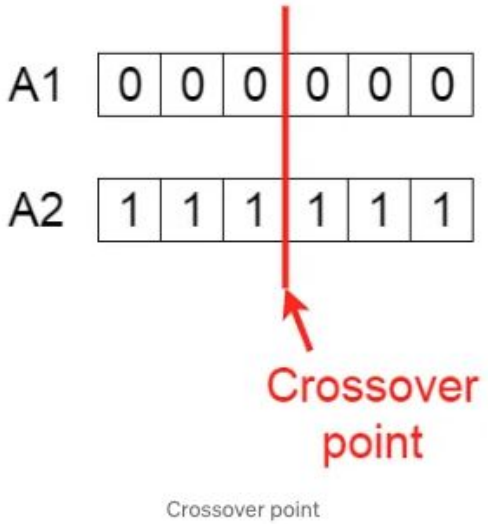
Selection

- Two individuals (parents) are selected based on their fitness scores.
- Individuals with high fitness have more chance to be selected for reproduction.

ref

Crossover

- For each pair of parents to be mated, a crossover point is chosen at random from within the genes.
- For example, consider the crossover point to be 3 as shown below.

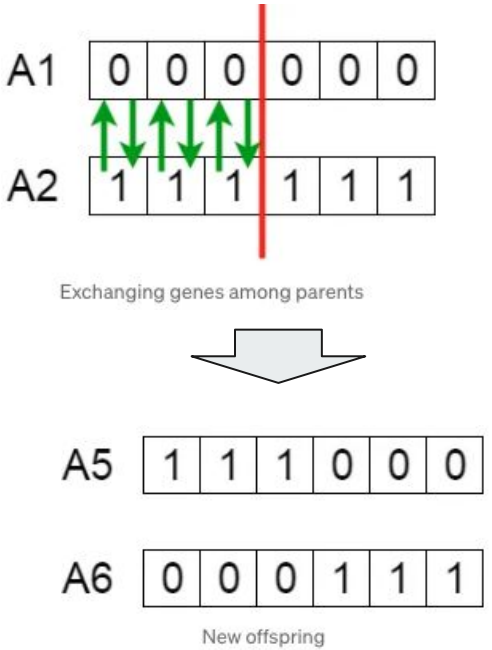


Crossover point

ref

Crossover

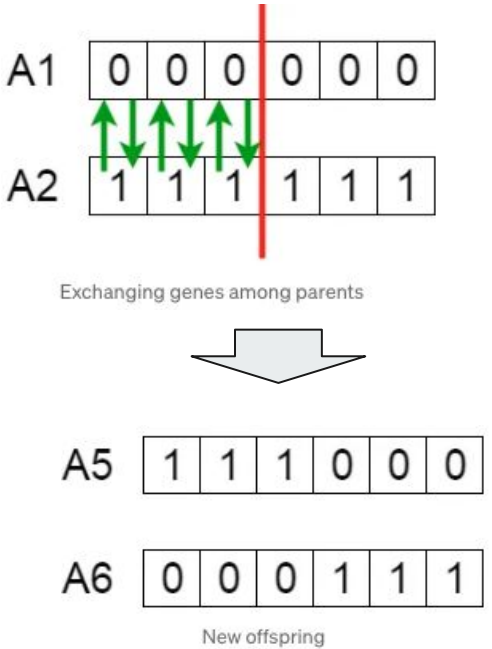
- For each pair of parents to be mated, a crossover point is chosen at random from within the genes.
- Offspring are created by exchanging the genes of parents among themselves until the crossover point is reached.



ref

Crossover

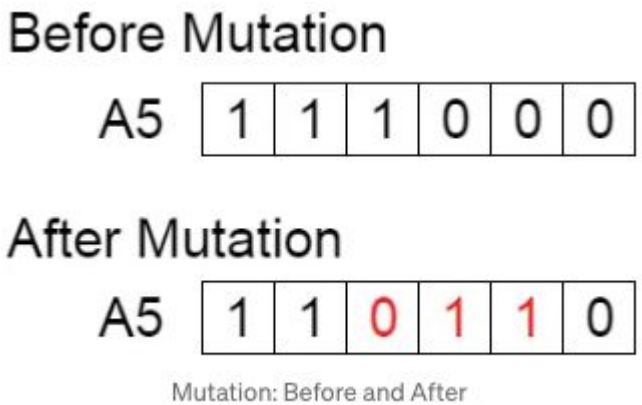
- For each pair of parents to be mated, a crossover point is chosen at random from within the genes.
- Offspring are created by exchanging the genes of parents among themselves until the crossover point is reached.
- **The new offspring are added to the population.**



ref

Mutation

- In certain new offspring formed, some of their genes can be subjected to a mutation with a low random probability.
- This implies that some of the bits in the bit string can be flipped.

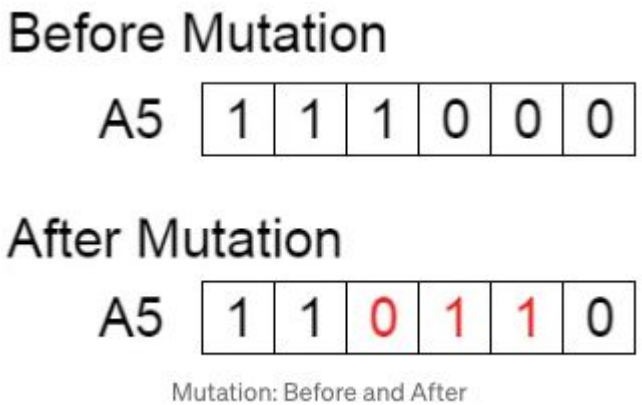


Mutation: Before and After

ref

Mutation

- In certain new offspring formed, some of their genes can be subjected to a mutation with a low random probability.
- This implies that some of the bits in the bit string can be flipped.
- **Mutation occurs to maintain diversity within the population and prevent premature convergence.**



Mutation: Before and After

ref

Termination (Convergence point)

- The algorithm terminates if the population has converged (does not produce offspring which are significantly different from the previous generation).
- Then it is said that the genetic algorithm has provided a set of solutions to our problem.

ref

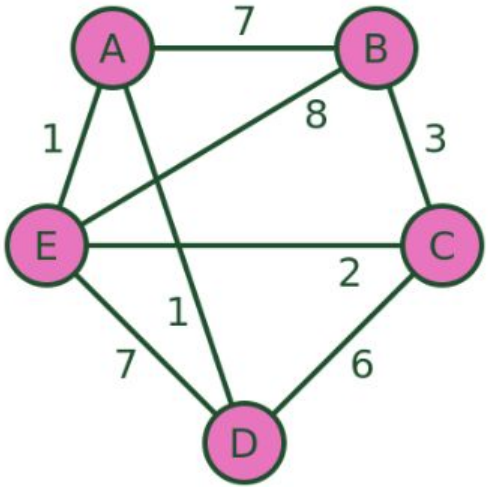
GA Algorithm (pseudocode)

```
START
Generate the initial population
Compute fitness
REPEAT
    Selection
    Crossover
    Mutation
    Compute fitness
UNTIL population has converged
STOP
```

- The sequence of phases is repeated to produce individuals in each **new generation** which are better than the previous generation.

ref

TSP



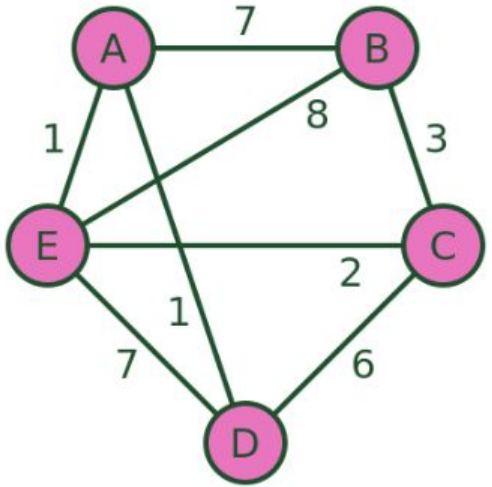
TSP

BRUTE-FORCE SOLUTION:

$O(n!)$

DYNAMIC PROGRAMMING ALGORITHMS:

$O(n^2 2^n)$



TSP (Notebook presentation)

QA

Course Material Review (part 1)

Week of	Topic (higher level)	Topic	Activity
Jan 08	Course introduction & Math and Probability Basics	Course introduction, regulations, and policies Math, Probability Basics (very brief; mainly directives)	
Jan 15	General idea of ML (connection to Math and Probability)	Polynomial curve fitting, connection between method of least squares and maximum likelihood learning.	Martin Luther King Jr. Day recess (Jan 15)
Jan 22	Supervised learning	Parametric and non parametric regression models	Class test 1 (30 min)
Jan 29		Parametric and nonparametric classification models	
Feb 05		Ensemble methods	Class test 2 (30 min)
Feb 12		Model selection, HP optimization	