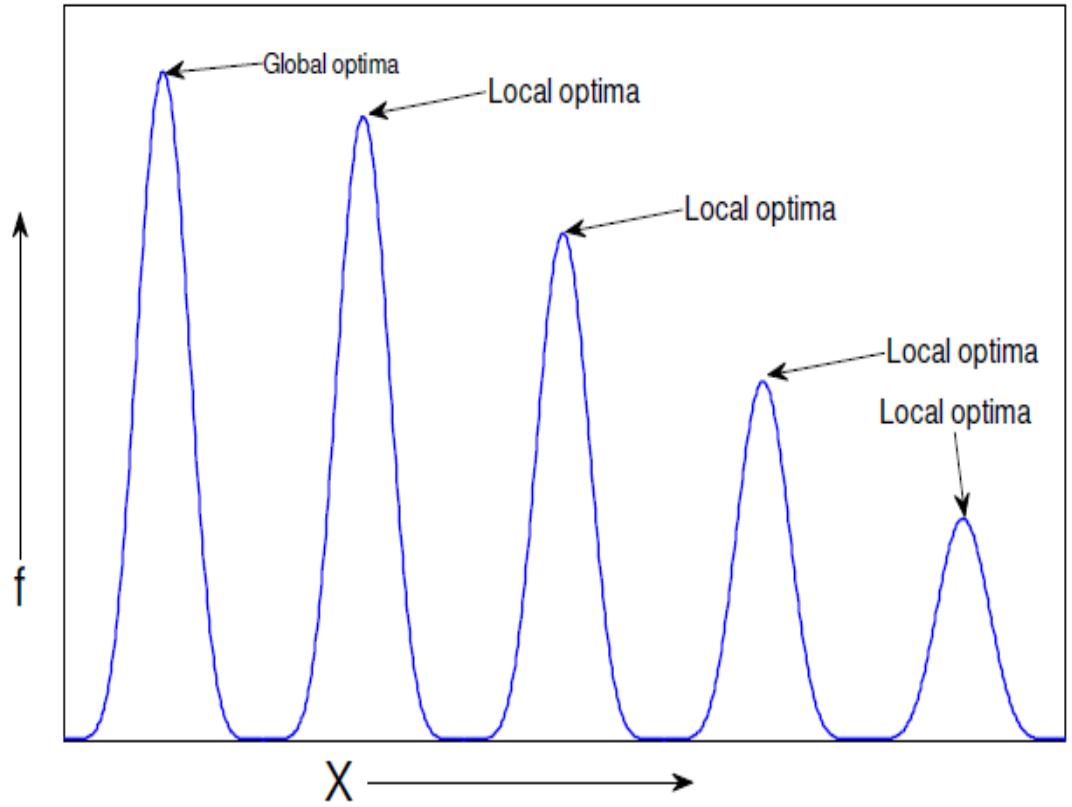
A Basic Introduction to Genetic Algorithms

Partha Pratim Sarangi

School of Computer Engineering

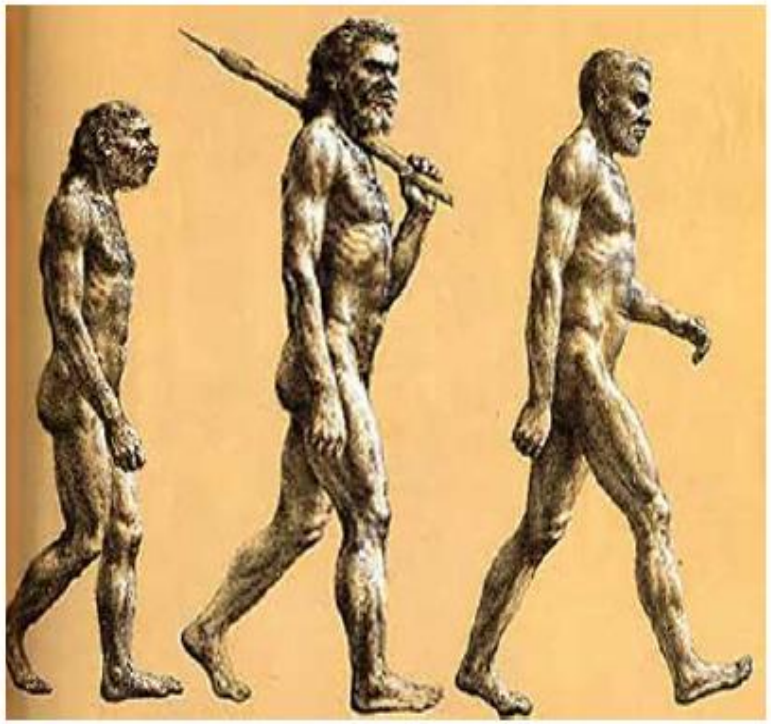
Introduction to optimization



Genetic Algorithms

Genetic Algorithms are the heuristic search and optimization techniques that mimic the process of natural evolution.

Principle Of Natural Selection

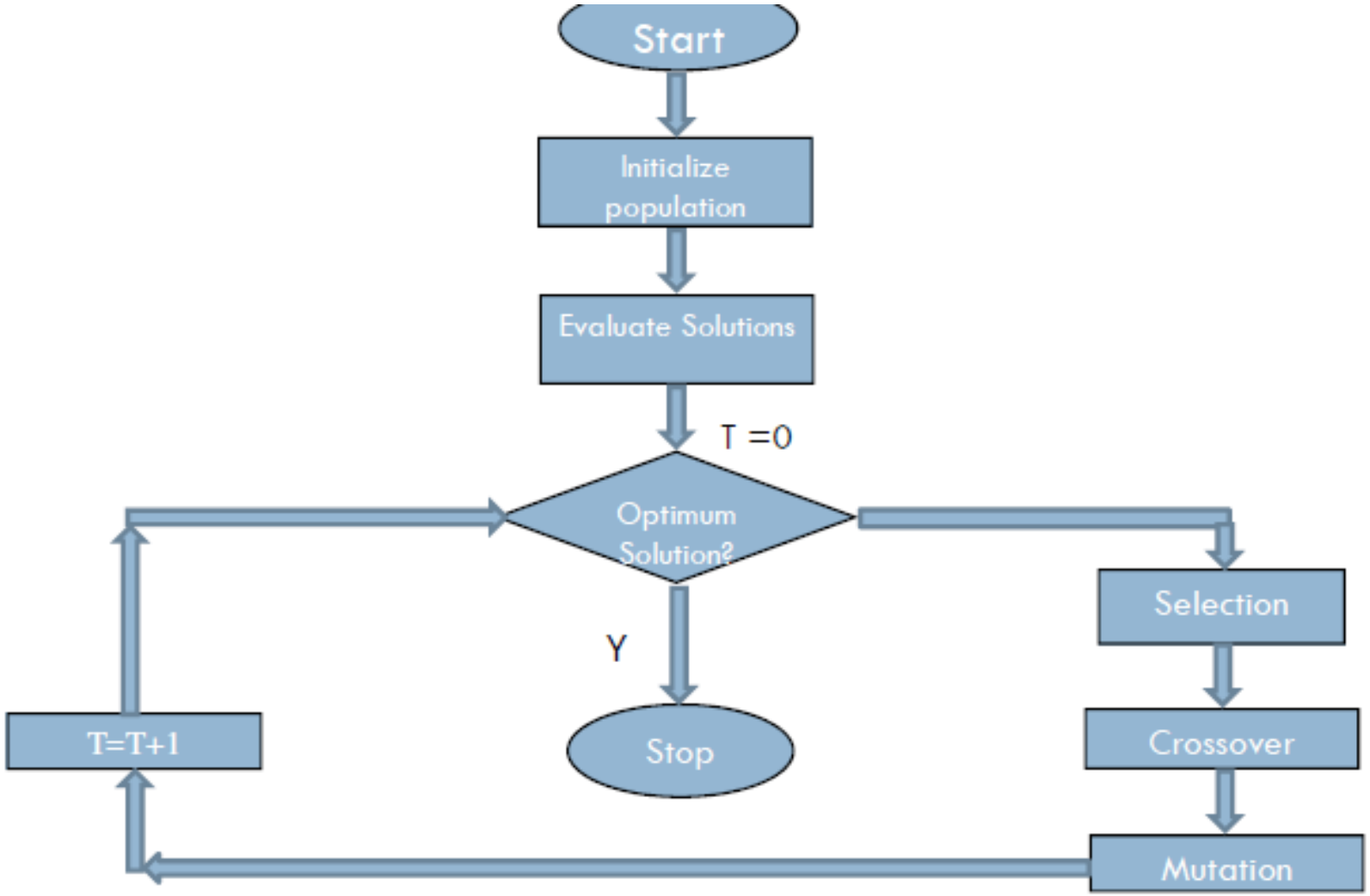


"Select The Best, Discard The Rest"

Evolution of species

Genetic algorithms implement the optimization strategies by simulating evolution of species through natural selection.

Simple Genetic Algorithms



Simple Genetic Algorithm

```
function sga()
        {
        Initialize population;
        Calculate fitness function;

        While(fitness value != termination criteria)
                {
                    Selection;

                    Crossover;

                    Mutation;

                Calculate fitness function;
                }
        }
```

Nature to Computer Mapping

Nature	Computer
Population	Set of solutions
Individual	Solution to a problem
Fitness	Quality of a solution
Chromosome	Encoding for a solution
Gene	Part of the encoding solution
Reproduction	Crossover

Terminologies

- **Key Elements**

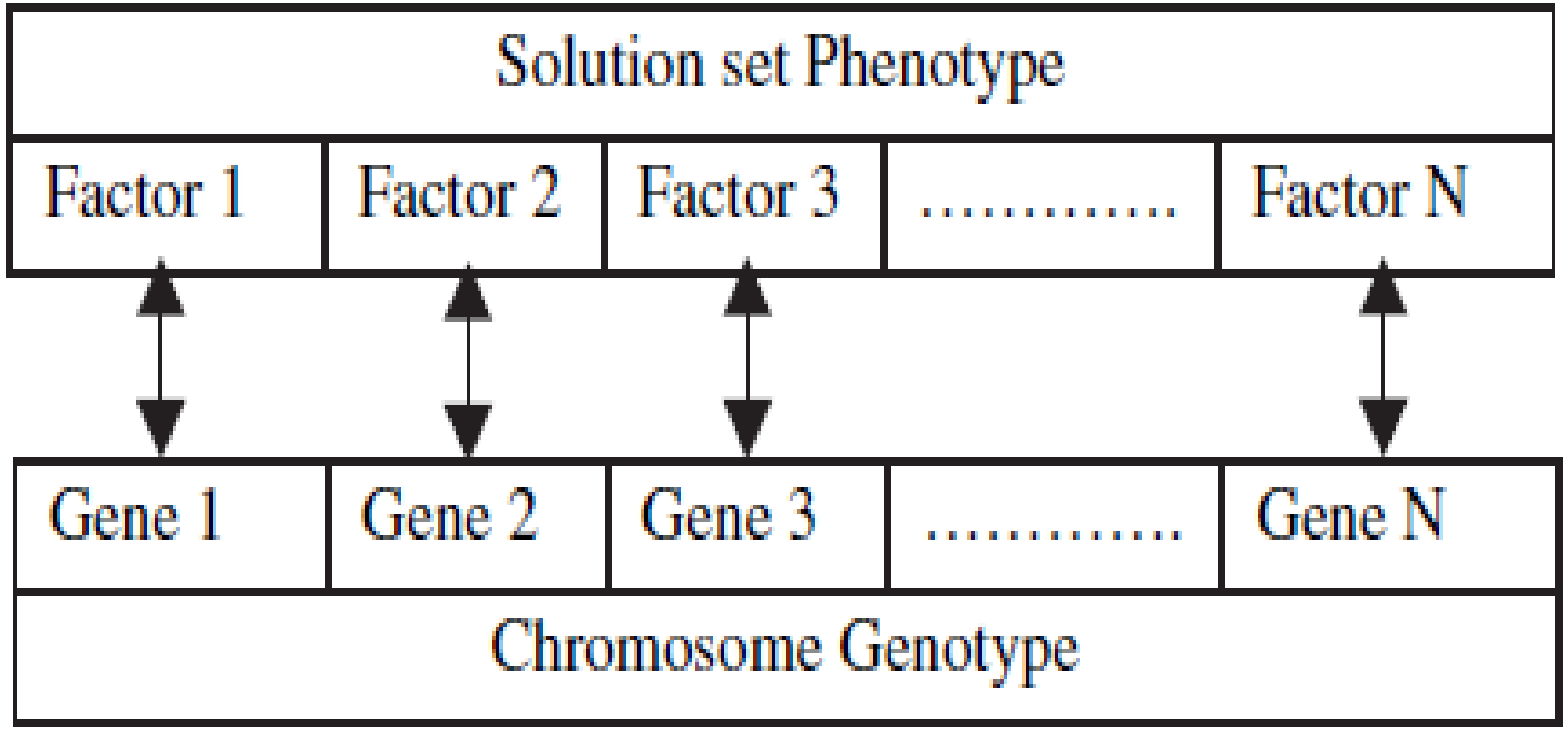
The two distinct key elements in the GA are individuals and populations. An individual is a single solution while the population is the set of individuals

Individuals

An individual is a single solution. Individual solution is represented in two forms as given below:

1. The chromosome, which is the raw ‘genetic’ information (genotype) that the GA deals.
2. The phenotype, which is the expressive of the chromosome in the terms of the model.

Representation of Genotype and phenotype



Chromosome

- A chromosome is a binary string.
- A chromosome is subdivided into genes.
- A gene is the GA's representation of a single factor value of a range.
- The range can be divided into the number of intervals that can be expressed by the gene's bit string.
- The size of the interval would be (range)/(2n-1).

Populations

A population is a collection of individuals. A population consists of a number of individuals being tested, the phenotype parameters defining the individuals and some information about search space. The two important aspects of population used in Genetic Algorithms are:

1. The initial population generation.
2. The population size.

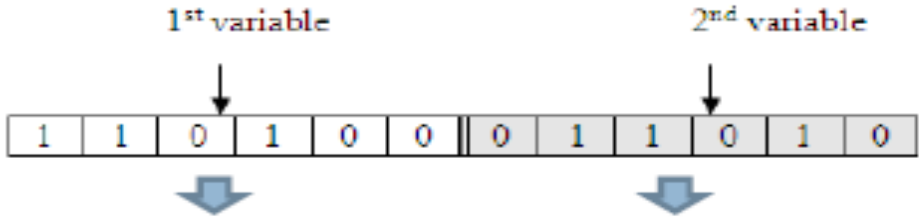
Population	Chromosome 1	1 1 1 0 0 0 1 0
	Chromosome 2	0 1 1 1 1 0 1 1
	Chromosome 3	1 0 1 0 1 0 1 0
	Chromosome 4	1 1 0 0 1 1 0 0

Encoding

- The process of representing a solution in the form of a string that conveys the necessary information. The process can be performed using bits, numbers, trees, arrays, lists or any other objects.
- The encoding depends mainly on solving the problem.

Encoding Methods

- Most common method of encoding is binary coded. Chromosomes are strings of 1 and 0 and each position in the chromosome represents a particular characteristic of the problem.



Decoded value 52 26

Mapping between decimal and binary value

$$x_i = x_i^{\min} + \frac{x_i^{\max} - x_i^{\min}}{2^{l_i} - 1} \mathrm{DV}(s_i)$$

Fitness

A fitness function value quantifies the optimality of a solution. The value is used to rank a particular solution against all the other solutions.

A fitness value is assigned to each solution depending on how close it is actually to the optimal solution of the problem.

Operators of GA

- Selection
- Crossover
- Mutation
- Elitism

Selection

The process that determines which solutions are to be allowed to reproduce and which ones deserve to die out.

The primary objective of the selection operator is to emphasize the good solutions and eliminate the bad solutions in a population while keeping the population size constant.

“Selects the best, discards the rest”

Selection

- Identify the good solutions in a population.
- Make multiple copies of the good solutions.
- Eliminate bad solutions from the population so that multiple copies of good solutions can be placed in the population.

Selection

- There are different techniques to implement selection in Genetic Algorithms.
- They are:

Roulette wheel selection

Rank selection

Tournament selection, etc

Roulette wheel selection

- Parents are selected according to their fitness values.
- The better chromosomes have more chances to be selected.

Chromosome #	Fitness	% of Roulette wheel
Chromosome 1	127.50	51
Chromosome 2	50.00	20
Chromosome 3	25.00	10
Chromosome 4	20.00	8
Chromosome 5	15.00	6
Chromosome 6	12.50	5

Rank selection

- The Roulette wheel selection will have problems when the fitness values differ very much.
- In this case, the chromosomes with low fitness values will have a very few chance to be selected.
- This problem can be avoided using ranking selection.

Chromosome #	Fitness	% of Roulette wheel	Rank	% of Roulette wheel
Chromosome 1	127.50	51	6	28
Chromosome 2	50.00	20	5	24
Chromosome 3	25.00	10	4	19
Chromosome 4	20.00	8	3	14
Chromosome 5	15.00	6	2	10
Chromosome 6	12.50	5	1	5

Tournament selection

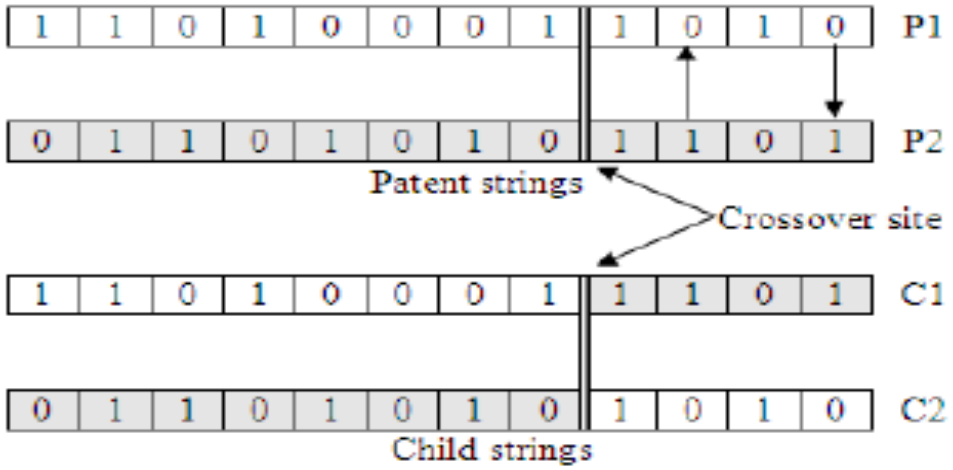
- In tournament selection several tournaments are played among a few individuals. The individuals are chosen at random from the population.
- The winner of each tournament is selected for next generation.
- Selection pressure can be adjusted easily by changing the tournament size.
- Weak individuals have a smaller chance to be selected if tournament size is large.

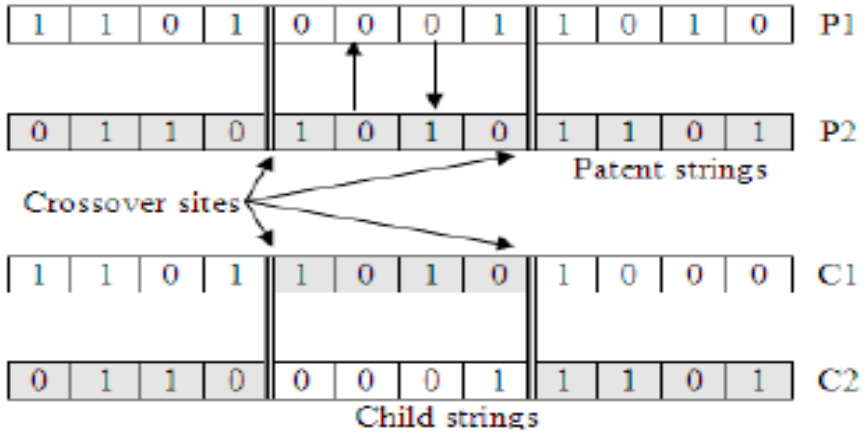
Binary Crossover

- The crossover operator is used to create new solutions from the existing solutions available in the mating pool after applying selection operator.
- This operator exchanges the gene information between the solutions in the mating pool.
- The most popular crossover selects any two solutions strings randomly from the mating pool and some portion of the strings is exchanged between the strings.

Binary Crossover

- A probability of crossover is also introduced in order to give freedom to an individual solution string to determine whether the solution would go for crossover or not.

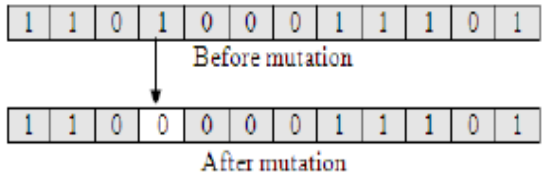




Binary Mutation

- Mutation is the occasional introduction of new features in to the solution strings of the population pool to maintain diversity in the population.
- Though crossover has the main responsibility to search for the optimal solution, mutation is also used for this purpose.
- Mutation operator changes a 1 to 0 or vise versa, with a mutation probability of .
- The mutation probability is generally kept low for steady convergence.
- A high value of mutation probability would search here and there like a random search technique.

Binary Mutation



Elitism

- Crossover and mutation may destroy the best solution of the population pool.
- Elitism is the preservation of few best solution of the population pool.
- Elitism is defined in percentage or in number.

An example problem

$$Maximize\ f(x) = \sin(x)$$

$$0 \leq x \leq \pi$$

Consider 5 bit string to represent the solution, then
00000 = 0 and 11111 = π

Let assume a population size of 4

One generation of a GA simulation

Table 1. One generation of a GA simulation on function $\sin(x)$.

	Initial population							New population			
String	DV[a]	x	$f(x)$	$f_i/\bar{f}$	AC[b]	Mating pool	CS[c]	String	DV	x	$f(x)$
01001	9	0.912	0.791	1.39	1	01001	3	01000	8	0.811	0.725
10100	20	2.027	0.898	1.58	2	10100	3	10101	21	2.128	0.849
00001	1	0.101	0.101	0.18	0	10100	2	11100	28	2.838	0.299
11010	26	2.635	0.485	0.85	1	11010	2	10010	18	1.824	0.968
	Average, $\bar{f}$		0.569						Average, $\bar{f}$		0.711

[a] DV decoded value of the string; [b] AC actual count of strings in the population; [c] CS stands for cross site

Source: Deb 1999

GA Stopping Criteria

- X number of generations completed-typically O(100)
- Mean deviation in performance of individuals in the population falls below a threshold J<x (genetic diversity has become small)
- Stagnation-no or marginal improvement from one generation to the next:*(Jn+1-Jn)<X*

Disadvantage of binary coded GA

- more computation
- lower accuracy
- longer computing time
- solution space discontinuity
- hamming cliff

References

- D. E. Goldberg, ‘Genetic Algorithm In Search, Optimization And Machine Learning’, New York: Addison –Wesley (1989)
- Holland J., “Adaptation in Natural and Artificial Systems”, University of Michigan Press, 1975
- Kalyanmoy Deb, ‘An Introduction To Genetic Algorithms’, Sadhana, Vol. 24 Parts 4 And 5.