Unit 8

U 10 1

Akanksha

3

GA is non conventional (non traditional) optimization technique. Basically GA is search algorithm which is based on mechanism of natural selection.

Biological Background -

The Cell-

Every animal / human cell is a complex of many "small" factories that work together.

The Center of all this is the cell nucleus. The genetic information is contained in the cell nucleus.

Chromosomes-

All the genetic information gets stored in the chromosomes. In human there are 23 pairs exists in 1 chromosome. Each chromosome is build of DNA.

The Chromosomes are divided into several parts called genes.

Gene code the properties of species i.e. the characteristics of an individual.

The combination of the genes for 1 property are called alleles.

A gene can take different alleles -

For Ex- there is a gene for eye color and all the different possible alleles are black, brown, blue etc.

The set of all possible alleles present in a particular population forms a gene pool.

This gene pool can determine all the different possible variations for the future generations.

The size of the gene pool helps in determining the diversity of the individuals in the population.

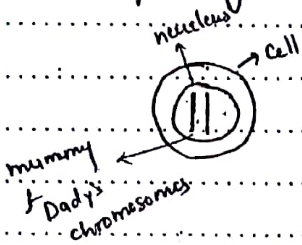
Akanksha

§ Reproduction →

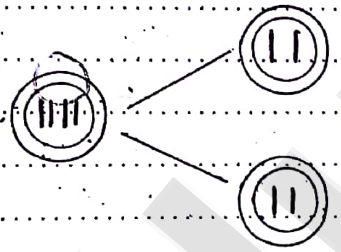
Is carried out by the following

Mitosis →

- The same genetic Informatis is copied to new offspring.

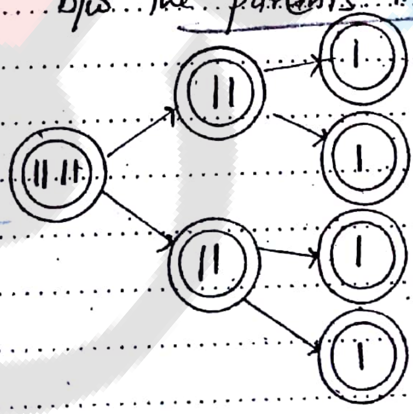


Would many to saw a vision to 2 get 2 cells, and the process of producing them involves the duplication of chromosomes. And 1 pair ends up in each of the child cells. That is Mitosis.



There is no exchange of information.

Meiosis → In this case, the genetic information is shared b/w the parents in order to create new offspring.



§ Difference b/w genetic algorithm and traditional methods

GA	Traditional Method
1. It works with coding of solution set not with solution itself	It works with parameters.
2. It improves the chance of reaching the global optimum and also helps in avoiding local stationary point.	
3. GA uses fitness func^n for finding optimum sol^n.	3. It uses derivatives for finding solution.
4. It narrow down the search space as the search progresses.	4.

§ Simple Genetic Algorithms –

GA handles a population of possible solutions. Each solution is represented through a chromosome.

Then reproduction operators are applied on the chromosomes.

The simple form of GA is given by the following –

handles
ulation of
...lu sol^n

pg 401
Deepa (5 steps)

Now, we will discuss each iteration of this process-

Generation-
Selection is supposed to be able to compare each Individual in the population. Selection is done by using a fitness funcⁿ.

Each chromosome has an associated value corresponding to the fitness of the solution. It represents.

The fitness should correspond to an evaluation of how good the candidate solution is.

The optimal solⁿ is the one which maximizes the fitness funcⁿ.

G.A deals with the problems that maximize the fitness function.

Generally the initial population is generated randomly.

The GA performs iterations to get desired population.

Each iteration consists of following steps:

1. Selection-
Here individuals are selected for reproduction. The individuals whose fitness is best are chosen for reproduction.

2. Reproduction-
Here offsprings are bred for generating new chromosomes. Here also uses both recombination and mutation.

3. Evaluation- Here the fitness of new chromosomes is evaluated.

4. Replacement-
Here, Individuals from the old population are killed and replaced by the new ones.

U10 5

The algo is stopped when the population converges toward the optimal solution.

```
BEGIN /* genetic algorithm*/
  Generate initial population;
  Compute fitness of each individual;
  WHILE NOT finished DO LOOP
    BEGIN
      Select individuals from old generations
        For mating;
      Create offspring by applying
        recombination and/or mutation
        to the selected individuals;
      Compute fitness of the new individuals;
      Kill old individuals to make room for
        new chromosomes and insert
        offspring in the new generalization;
      IF Population has converged
        THEN finishes: = TRUE;
    END
END
```

Flow Chart for Genetic Algorithm →

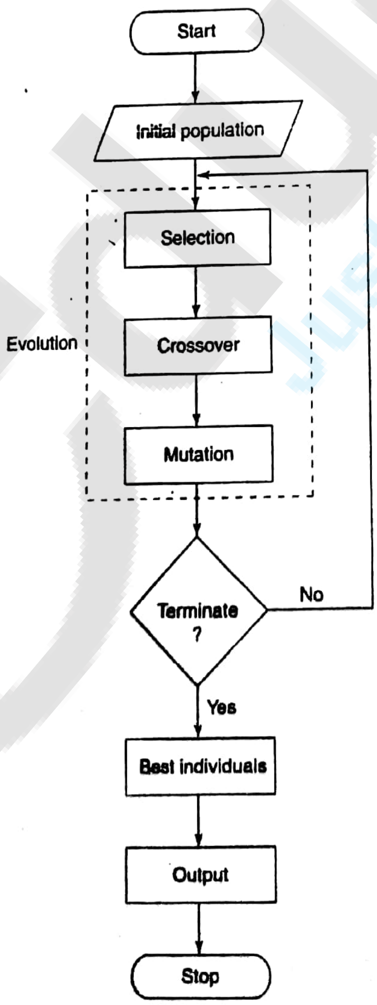


Figure 15-13 Flowchart for genetic algorithm.

§ Operators in Genetic Algorithm →

The 4 basic operators are -

1. Encoding
2. Selection
3. Crossover
4. Mutation

1. **Encoding**: It is a process of representation of individual genes.

This process can be performed by using-
- bits
- numbers
- trees
- arrays
- lists or any other objects.

a. **Binary Encoding** → It is the most common way of encoding. According to this, each chromosome encodes a binary string.

Binary coded string has 1's and 0's

Every chromosome is a string of bits 0 or 1.

Ex –

Chromosome 1   101101110

Chromosome 2.   100110101.

b. Octal Encoding → Here strings are made up of Octal numbers. [0 - 7]

Ex. Chromosome A : 03467216

Chromosome B : 15733214

c. Hexadecimal Encoding →

Here encoding strings are made up of Hexadecimal numbers [0 - 9 , A - F]

Ex: Chromosome A: 9CE7

Chromosome B : D6FA

d. Permutation Encoding → It can be used in ordering problem, such as travelling salesman problem or Task ordering problem.

In permutation encoding, every chromosome is a string of numbers, which represents number in a sequence.

Ex- Chromosome A: 1 5 3 2 6 4 7 9 8

Chromosome B: 8 5 6 3 2 7 9 4 1

Ex. of problem : Travelling salesman problem

The problem : There are cities and given distances b/w them. Travelling salesman has to visit all of them, but he does not to travel very much. Find the sequence of cities to minimize travelled distance.

Chromosome: It says order of cities, in which salesman will visit them.

e. Value Encoding →

Here, every chromosome is a string of some values. Values can be anything connected to problem, from numbe real numbers or chars to some complicated objects

Ex:

Chromosome A 1.2324 5.3243 0.4556 2.2329

Chromosome B ABDJEIFDCATDJPFCAD

Chromosome C (back), (right), (forward), (left)

Example of problem: Finding weights for neural network

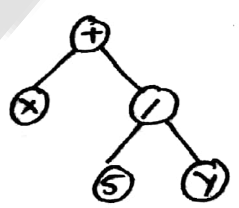
The problem: There is some neural network with given architecture. Find weights for inputs of neurons to train the network for wanted output.

Encoding: Real values in chromosomes represent corresponding weights for inputs.

f. Tree Encoding → It is used mainly for evolving programs or expressions for genetic programming.

Here, every chromosome is a tree of some object such as functions or commands in programming language.

Chromosome A



(+ x(/ 5 y))

Chromosome B

(do-until step wall)

Three Important Operators

1. Selection
2. Crossover
3. Mutation

Selection →

Selection is the process of choosing two parents from the population for crossing.

The purpose of selection is to emphasize fitter individual in the population so that their offsprings have higher fitness.

The basic selection process

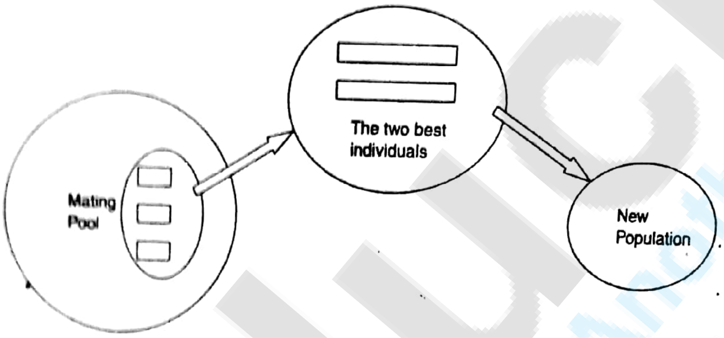


Figure 15-20 Selection.

Selection is a method that randomly picks chromosomes out of the population according to their evaluation function.

If an individual has higher fitness then it is better chance to get selected.

The degree to which the better individuals are favored is called selection pressure. So higher the selection pressured, the more the better individuals are favored.

The selection pressure help the GA to improve the population fitness over successive generations.

The convergence rate of GA is largely determined by the ie: If selection pressure is high it implies that the convergence rate is high.

If the selection pressure is low then GA will take time to find optimal solution.

There are 2 types of selection scheme-

① proportionate - based selection
It picks out individuals based upon their fitness values as compare to the fitness of other individuals in the population.

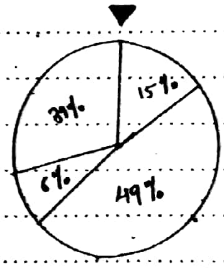
② Ordinal - based selection -
It selects individuals not upon their raw fitness, but upon their rank within the population.

There are 6 methods for selecting chromosomes -

1) Roulette - Wheel selection →
for Ex -
Here we are having population of 4 strings and their fitness values.

No	String	x values	Fitness $f(x) = x^2$	Probability of selecting string	% f
1	01101	13	169	0.15	15
2	11000	24	576	0.49	49
3	01000	8	64	0.06	6
4	10011	19	361	0.31	31
			$\Sigma f = 1170$		

- The total fitness score of the population is represented by roulette wheel.



- Now, we assign a slice of the wheel to each member of the population.

The size of the slice is proportional to the chromosomes fitness score.

i.e. the fitter member will have bigger slice.

Now, to chose a chromosome we have to spin the wheel and select the chromosome at the point it stops.

2) Random selection →

This technique randomly selects a parent from the population.

3) Rank selection →

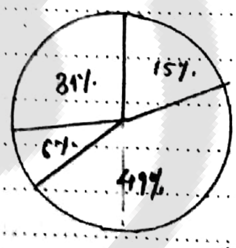
This method really focuses on fit members.

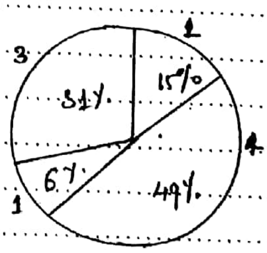
There is one problem with Roulette-wheel i.e. when the fitness values differ very much, i.e. If the best chromosome fitness is 90% then it's circumference occupies 90% of Roulette-wheel and the other chromosomes will have very few chances to be selected.

This method first rank the population and receives fitness from the ranking.

i.e. The worst will have fitness 1, the next 2 - 2 and the best will have fitness N.

where N = no of chromosomes





Roulette-wheel according to fitness

Roulette-wheel according to rank.

4) Tournament Selection →
This method extracts k individuals from the population with uniform probability and makes them play a "tournament"

The winner of the tournament is one whose fitness value is high.

The winner is inserted into the mating pool.

Ex –

Individuals	1	2	3	4	5	6
Fitness	1	2.1	3.11	4.01	4.66	1.19

Step 1 – First select individuals 2 and 4 at random

ϕ_2 ϕ_4

2.10 4.01

4 is the winner and hence select the string as 1000

Step 2 – Select individuals 1 and 6

ϕ_1 ϕ_6

1 1.19

6 is the winner and hence select the string as 0110

Step 3 – Select individuals 3 and 5

ϕ_3 ϕ_5

3.11 4.66

here 5 is the winner

Repeatedly this process is performed and those individuals are selected to place into the mating pool which are selected more times.

5) Boltzmann Selection →
This method follows Simulated Annealing. According to this

This method simulates the process of slow cooling of molten metal to achieve minimum function value in a minimization problem.

This cooling phenomenon is simulated by controlling a temperature like parameter introduced with the concept of Boltzmann probability distribution.

$$P(E) = \exp\left(\frac{-E}{kT}\right)$$

where "k" = Boltzmann constant

This expression suggests that a system at a high temperature has almost uniform probability of being at any energy state, but at a low temperature it has a small probability of being at a high energy state,

Therefore, by controlling the temperature T and assuming search process follows Boltzmann Probability distribution, the convergence of the algorithm is controlled.

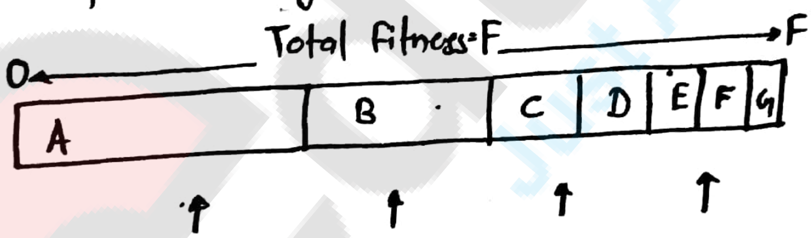
6 Stochastic Universal Sampling →

6. Stochastic Universal Sampling →

Here, the individuals are mapped to contiguous segment of a line, such that the size of the each individual is exactly same as it is in Roulette wheel.

And equally spaced pointers are placed over the line

i.e. if I want 4 individual for mating then I will have 4 pointers, if I want 6 individuals for mating then I will have 6 pointers and so on..



So, After selection the mating population consists of the individuals

A, B, C, F

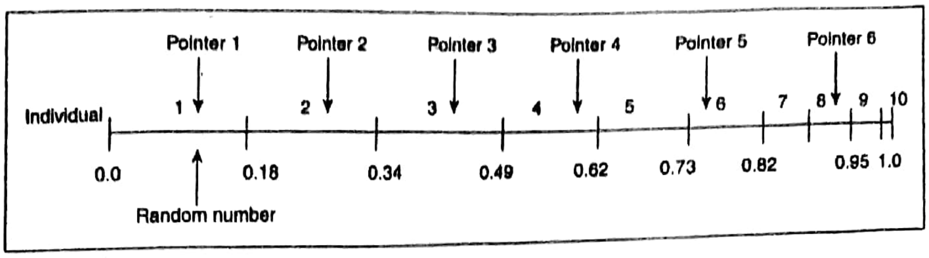


Figure 15-21 Stochastic universal sampling.

§ Crossover (Recombination)

Reproduction makes clones of good strings, but does not create new ones. But cross over operator is applied to mating pool to create a better string.

So, Crossover is a swapping of chromosomes which produces better offspring.

There are 10 types of cross-over →

1. Single - Point Crossover -
Here a cross site or cross point is selected randomly along the length of the mated string. And the bits which are next to the cross site are exchanged.

If an appropriate site is chosen then better children can be obtained otherwise it may give less fit string.

If good strings are not created by cross over they will not survive beyond next generation because reproduction will not select those strings for the next mating pool.

10 (U10) (12)

Parent 1 : 111010 | 0101

Parent 2 : 101011 | 1011

String before mating

Child 1 : 111010 1011

Child 2 : 101011 0101

String after mating.

2. Two - Point Crossover —

Here 2 random sites are chosen and the content bracketted by these sites are exchanged b/w two mated parents.

If the cross-site 1 is THREE and cross-site 2 is SIX then the string b/w THREE & SIX are exchanged.

	cross-site 1		cross-site 2
Parent 1	1 1 0	0 1 0	1 1 1 0 1
Parent 2	0 1 1	0 1 1	0 1 0 0 1

String before mating

Child 1	1 1 0	0 1 0	1 1 1 0 1
Child 2	0 1 1	0 1 1	0 1 0 0 1

String after mating

3. Multipoint Crossover [N-Point Crossover]

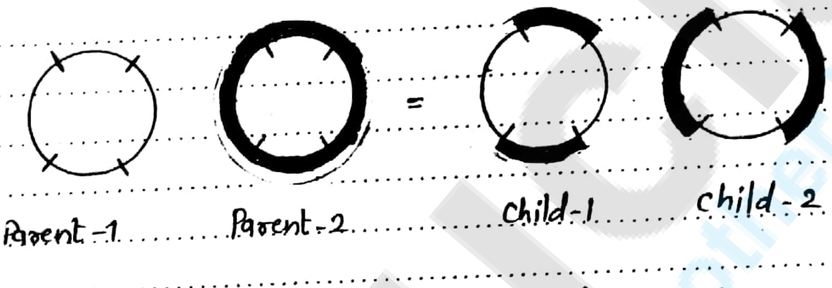
There are 2 Cases in this crossover-

1. Even number of cross-sites
2. Odd number of cross-sites

1. In case of Even numbered cross site, the string is treated as a ring with no beginning or end.

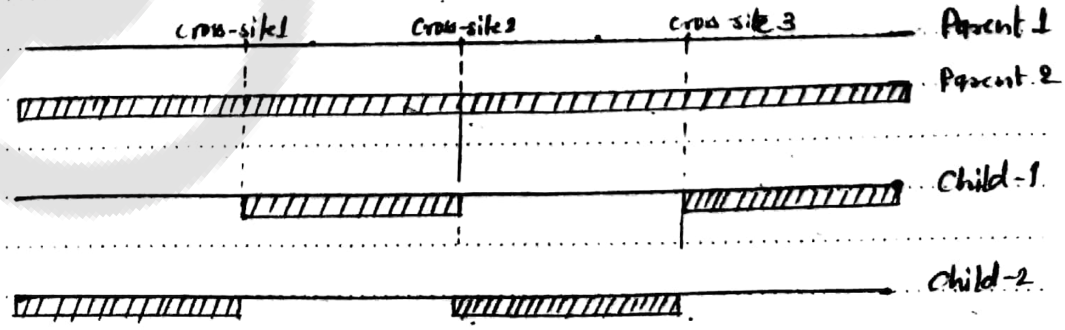
The cross sites are selected around the circle uniformly at random.

Then the Information b/w alternate pairs of site is interchange



fig(a) Multi-point cross over with even number of cross-sites.

2. If the number of cross-sites are odd, then string is treated as a straight line. Here also the info^n b/w alternate pair of site is interchanged.



fig(b) Multi point cross over with odd number of cross-sites.

4. Uniform Crossover-

Here, each bit from either parent is selected with a probability of 0.5 and then interchanged.

Here, when there is 1 in the mask, the gene is copied from the first parent and when there is 0, the gene is copied from the second parent.

Parent 1: 1 0 1 1 0 0 1 1

Parent 2: 0 0 0 1 1 0 1 0

Mask : 1 1 0 1 0 1 1 0

Child 1: 1 0 0 1 1 0 1 0

Child 2: 0 0 1 1 0 0 1 1

Uniform Crossover

5. Three - Parent Cross Over →

In this crossover technique, 3 parents are randomly chosen.

Each bit of the first parent is compared with the bit of the second parent.

If both are the same, the bit is taken for the offspring, otherwise the bit from the third parent is taken for the offspring.

Parent 1 1 1 0 1 0 0 1

Parent 2 0 1 1 0 1 0 1

Parent 3 0 1 1 0 1 1 0

Child 0 1 1 0 1 0 1

Three parent Cross-over

(U10) (15)

5

§ Mutation →

After cross over mutation is performed on strings by this mechanism we get variety of population.

Mutation prevents the algorithm to be trapped in a local minima. Mutation has been traditionally considered as a simple search operator.

If crossover is failed to find optimum solution then mutation is supposed to help for the exploration of the whole search space.

Mutation is viewed as a background operator to maintain genetic diversity in the population.

A simple mutation can consist in inverting the value of each gene with a small probability.

The probability is usually taken about 1/L, where L is the length of chromosome.

Mutation ~~of~~ a bit means flipping a bit i.e. changing 0 to 1 and vice-versa.

a) Flipping →

A mutation chromosome is generated on the basis of it flipping of bit is performed.

For a 1 in mutation chromosome, the corresponding bit in parent chromosome is flipped and child chromosome is produced.

Parent	1 0 1 1 0 1 0 1
Mutation chromosome	1 0 0 0 1 0 0 1
Child	0 0 1 1 1 1 0 0

Mutation Flipping

U10 16

b) Interchanging →

Here 2 random positions of the string are chosen and the bits corresponding to those positions are interchanged.

Parent 1 0 1 1 0 1 0 1

child 1 1 1 1 0 0 0 1

Interchanging

c) Reversing →

Here a random position is chosen and the bits next to that positions are reversed and child chromosome is produced.

Parent 1 0 1 1 0 | 1 0 1

Child 1 0 1 1 0 | 1 1 0

Reversing