__/__/____

Unit-5

Genetic Algorithm (Basic Terms)

Evolutionary Algorithms:
Evolutionary algorithms are those algorithms which follow some biological and physical behaviors.

Biological behaviors:

* Genetics & Evolution → Genetic Algorithm (GA)
* Human Nervous system → Artificial Neural Network (ANN)
* Behavior of ant colony → Ant colony optimization (ACO)

Physical behaviors:-

* Learning → Fuzzy Logic (FL)
* swarming of particle → Particle swarming optimization.

Genetic Algorithm:- It is a subset of Evolutionary algorithm that models biological process (Genetics & Evolution) to optimize high complex functions (very difficult to model mathematically or NP-hard problems which are computationally very expensive to solve or involve large number of parameters.

Background of Genetic Algorithm (GA)

* Introduced by Prof. John Holland (Michigan University, USA) in 1965.

* First article on GA was published in 1975.

* It is based on two fundamental biological processes.

(a) Genetics : G. J. Mendel (1865)

(b) ~~Evolution~~ Evolution :- C. Darwin (1875)

Genetics: branch of biology that deals with the study of genes, genes variation & heredity.

Evolution: Evolution is the process by which population of organisms change over generations

Genetic variations causes these changes

__/__/__

Genetic Algorithm

Genetic Algorithm (working principle)

Genetics

↓

Cells

(Basic building block in living bodies)

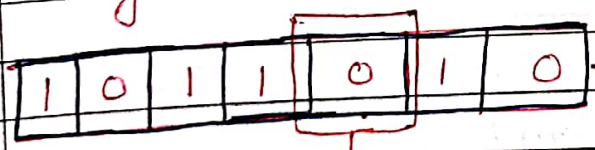
↓

Each Cell has some sets chromosomes

(man 46, Frogs : 26, mosquitos : 6)

↓

Chromosome are the strings of DNA (0's and 1's)



0 Gene → allele

(variant form of a gene)

Chromosome consist of Gene (block of DNA)

↓

Each Gene encodes a specific traits (eg hair color, age etc)

Evolution

Heredity Diversity Selection Ranking

① Heredity: An offspring has many of its characterstics of its parents.

② Diversity: variation in characterstics in the next generation.

③ Selection: only a small percentage of offspring produced to ad

④ Ranking offspring survival depends on their inherited characterstics (survival of the best)

__/__/____

Intuition based Genetic Algorithm

Consider a hypothetical situation

If in a country, we want only Good peoples, so we implement a policy like this:

1. Select only all the good people using a specific criteria and ask them to extend their generation by having their children.
2. Repeat this for a few generations.
3. So after some generations, we have an entire generation of good people.
4. The basic idea is to make some changes in the input (i.e population) such that we get better output (i.e better country)

Genetic Algorithm (GA)

* Genetic Algorithm is an optimization technique which tries to find out such values of input so that we get the best output values or results for highly complex problems by simulating biological processes (Genetics and Evolution)

Flow chart of GA (Next Lecture)

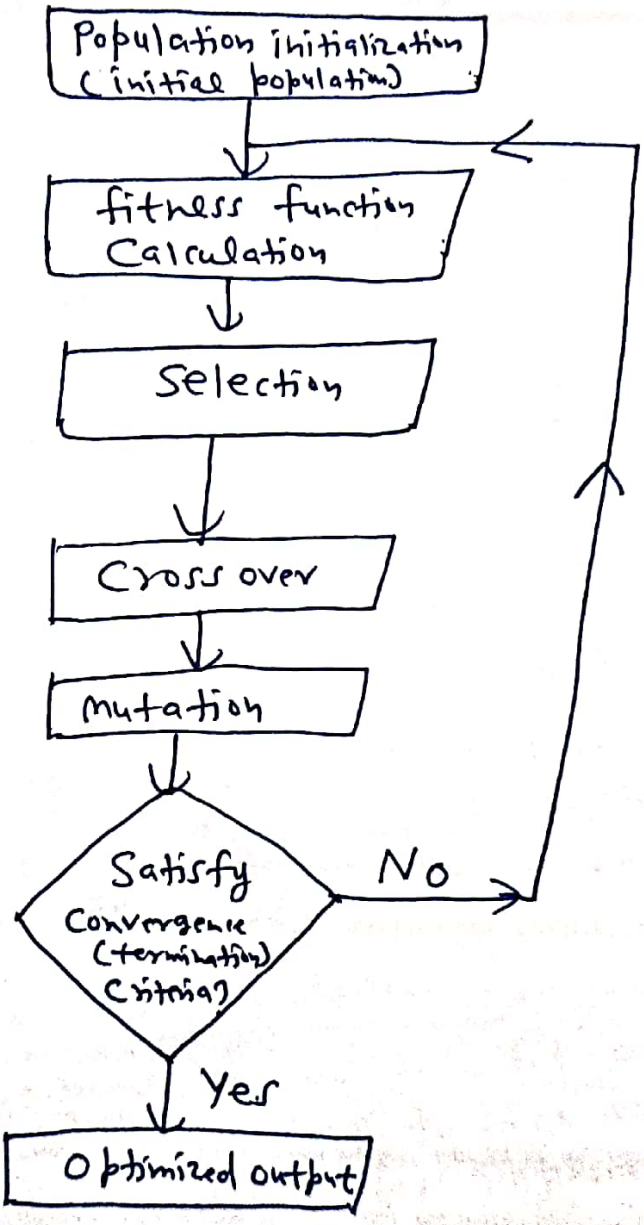
__/__/____

Difference between Genetic algorithm and Traditional Algorithm

Genetic Algorithm	Traditional Algorithm
(1) used to find the optimal solutions for difficult problems.	(1) It provides a step by step procedure to solve a problem.
(2) used in artificial Intelligence machine learning and research fields.	(2) used in programming, mathematics etc.
(3) More advanced	(3) Not advanced
(4) It works on a population of point to search best solution among them.	(4) It works on a single point to search the solution of a given problem.
(5) It makes use of probabilistic rules.	(5) It makes use of deterministic rules.

Basic Structure (Flow Chart) of Genetic Algorithm

The basic structure (flow chart) of genetic algorithm is as given:



Flow Chart of genetic algorithm

The various phases in a genetic algorithm can be understand in the following way:

① **Population Initialization:** Population is considered as a subset of all the possible solutions (to the given problem. So individual (Solution) is represented using chromosome. (i.e. each chromosome represent one Solution among a set of Solutions.

* The population is usually defined as a 2 dimensional array of – Population size, & Chromosome size.
* There are 2 primary methods to initialize a population in Genetic Algorithm (GA):
  (i) Random initialization: Populate the initial population with Completely random solutions.
  (ii) Heuristic initialization: Populate the initial population using a known heuristic for the problem.
* Population models:
  (i) Steady State: In steady state GA, we generate one or two offsprings (Children) solution in each iteration and they replace one or two individuals from the population.
  (ii) Generational: In a generational model, we generate 'n' offsprings, where 'n' is the population size, and the entire population is replaced by the new one at the end of the iteration.

(2) Fitness function Calculation

* Fitness function determines how fit an individual is. It gives a fitness score to each individual.
* individuals selected for reproductions is based on the fitness score.
* A fitness function should posses the following ~~data~~ Characteristics:
  (i) The fitness function should be sufficiently fast to Compute.
  (ii) It must quantitatively measure how fit a given solution is & or how fit individuals can be produced from the given solution.

(3) **Selection (Parent Selection)**: Parent selection is the process of selecting parents which mate and recombine to create off-springs for the next generation. In this two pairs of individuals based on their fitness score are selected and pass their genes to next generation.

The various selection methods are as given:

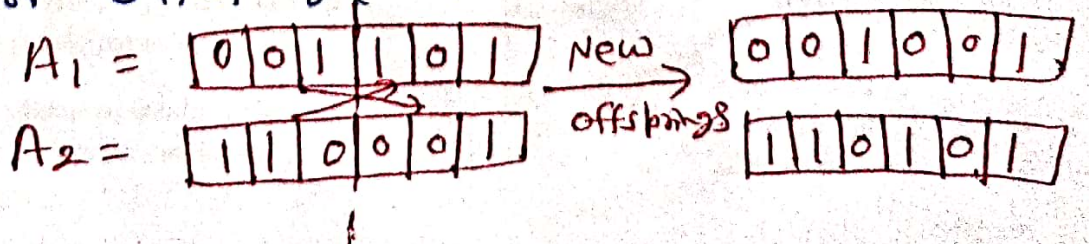
(a) **Fitness Proportionate Selection**: In this, every individual can become a parent with a probability which is proportional to its fitness. Two implementation of fitness proportionate selection are:
   (i) Roulette Wheel Selection
   (ii) Stochastic Universal Sampling

(b) **Tournament Selection**: In tournament selection, we select k-individuals from the population at random and select the best among them to become a parent. The same process is repeated to select next parent.

(c) **Rank Selection**: Rank selection also works with negative fitness values and mostly used when the individuals in the population have very close fitness value.

(d) **Random Selection**: In this method, we randomly select parents from the existing population.

(4) **Crossover**: Crossover is a genetic operator used to vary the programming of a chromosome or chromosomes from one generation to the next. In this, two parents are selected from the mating pool at random to crossover in order to produce superior offspring (children).

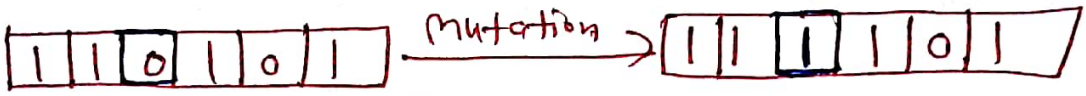


The different types of crossover are

(i) Single point Crossover (ii) Two point Crossover (iii) Uniform Crossover (iv) Half Uniform Crossover (v) Three parent Crossover (vi) Suffle Crossover

(5) **Mutation:** Mutation is defined as a small random tweak (change) in the chromosome of the offspring (children) solution during its growth to give new solution.

* It is used to maintain and introduce diversity in the genetic population



Different types of mutations are

(i) Flip-bit method (ii) Boundary Method (iii) Interchange method (iv) Reversing method

(6) **Convergence (Termination) Condition:**

The termination or convergence condition is important in determination when a Genetic Algorithm will end.

* Usually we keep one of the following termination conditions

(i) when there has been no improvement in the population for X-iterations.

(ii) when we reach an absolute no. of generations

(iii) when the objective function value has reached a certain pre-defined value.

Genetic operators and its types:

* **Genetic operators:** Genetic operators are the operators used in genetic algorithm to guide the algorithm towards a solution to a given problem.
* These operators alter (change) the genetic composition of the offspring (children).
* These operators work in conjunction with one another in order for the algorithm to be successful.

Types of genetic operators:

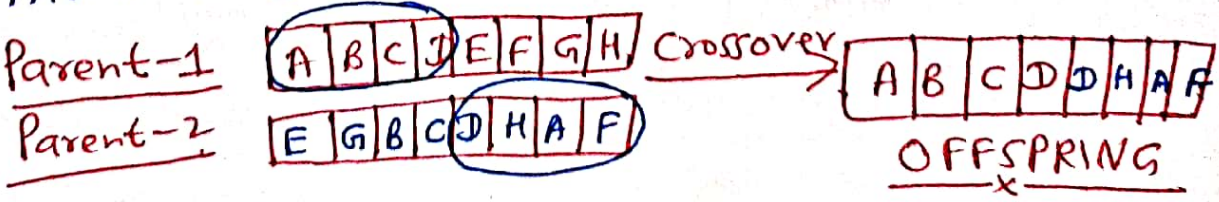
(1) Selection operator (Survival of the fittest)

The primary objective of the selection operator is to emphasize the good solution and eliminate the bad solutions in a population while keeping the population size constant. Different methods (fitness function, tournament selection etc) exist for choosing the best solution.

(2) Cross-over (Reproduction using recombination)

~~Operator~~

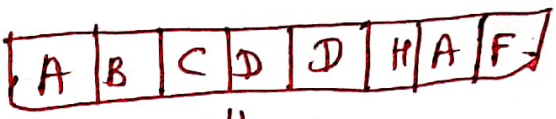
Cross over operators are used in the reproduction process. They take more than one parent solutions (Chromosomes) and produces a child solution from them. By re-combining portions of good solutions, the Genetic Algorithm (GA) is more likely to create a better solution.



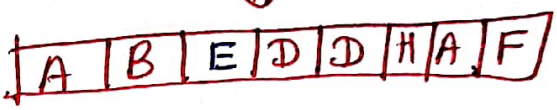
(3) **Mutation Operator:** The mutation operator encourages genetic diversity amongest solutions and attempts to prevent the genetic algorithm to give solutions that are very close to one another.

* In mutation, a given solution may change entirely from the precious solution to give a much better solution.

Before Mutation



After Mutation



Encoding in Genetic Algorithm

Encoding: Encoding is a process of representing individual genes. The encoding depends mainly on solving the problem.

Types of encoding:

① **Binary encoding:** The most common way of encoding is a binary string:

Chromosome-1	1	0	1	1	0	1	0	0
Chromosome-2	1	1	0	0	1	1	0	1

each chromosome consists of a binary (bit) string. Each bit in the string can represent some characteristics of the solution. Every string is therefore a solution, but not necessarily the best solution.

② **Octal encoding:** In this encoding, each string uses a set of octal number (0 to 7).

Chromosome-1	1	2	3	5	6	7	1
Chromosome-2	0	7	1	4	5	3	2

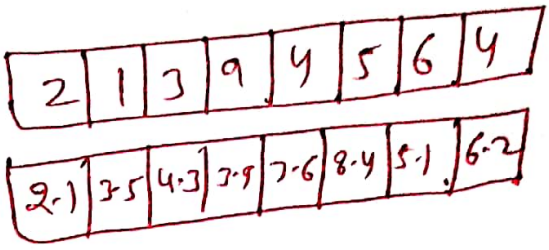
③ **Hexadecimal encoding:** This encoding uses strings made up of hexadecimal numbers (0-9, A-F).

Chromosome-1	2	A	D	4	B	F	5
Chromosome-2	A	0	1	E	F	0	C

(4) **Permutation encoding:** In permutation encoding, every chromosome is a string of integer/real values.

Chromosome-1

Chromosome-2



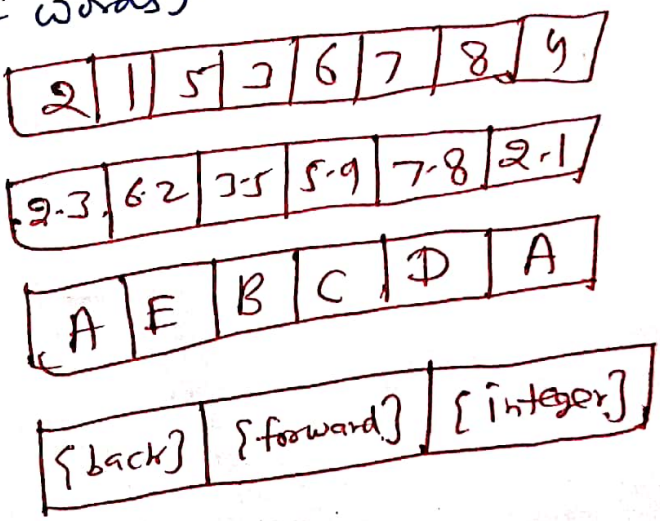
(5) **Value encoding:** In value encoding, every chromosome is a string of some values. Value can be anything connected to the problem (integer, real, characters, combination of words).

Chromosome-1

Chromosome-2

chromosome-3

Chromosome-4



Parent Selection in GA
(Roulette Wheel Selection)

Parent Selection Parent selection is the process of selecting parents which mate and recombine to create off-spring for the next generation.

Different technique for parent selection

1. Fitness Proportionate Selection
2. Tournament Selection
3. Rank selection
4. Random Selection

Fitness Proportionate Selection:

In this method, every individual can become a parent with a probability which is proportional to its fitness. A fitter individual have a higher chance of mating and propagate their features to the next generation.

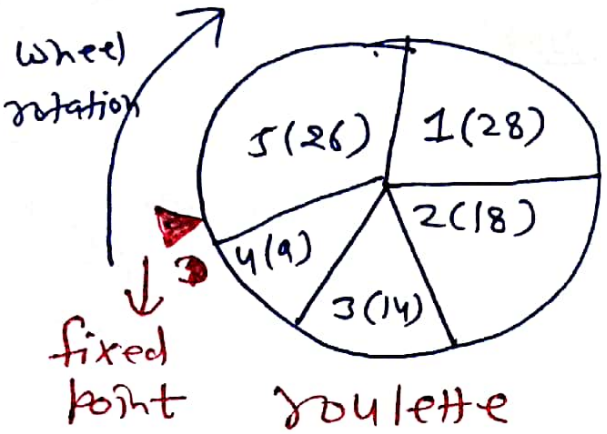
Types of fitness Proportionate Selection

(i) Roulette wheel Selection (ii) Stochastic Universal Sampling

(i) Roulette wheel Selection:

* In a Roulette wheel selection, the circular wheel is divided into n-parts, where n is the no-of individuals in the population.
* Each individual gets a ~~partition~~ portion of the circle which is proportional to the fitness value.
* A fixed point is chosen on the wheel circumference and the wheel is rotated.
* The region of the wheel which comes in front of the fixed point is chosen

as parent for the second parent, same process is followed.



Roulette wheel

Chromosome	1	2	3	4	5
Fitness (f)	28	18	14	9	26
Probability (P_i)	$\frac{28}{95}$ = 0.295	0.189	0.147	0.09	0.274
Commulative Probability (q_i)	0.295	0.484	0.631	0.721	$\frac{95}{100}$ = 1.00

* It is clear that a fitter individual has a greater pie on the wheel and therefore a greater chance of landing in front of the fixed point, where the wheel is rotated.

Algorithm ① Find the fitness of each individual

② Compute the probability of each individual (P_i)

③ Compute the commulative probability (q_i)

④ Generate a random No. $r \in [0,1]$.

⑤ ~~If $r < q_1$~~ If $r < q_1$, select first chromosome else select chromosome i such that

$$q_{i-1} < r \leq q_i$$

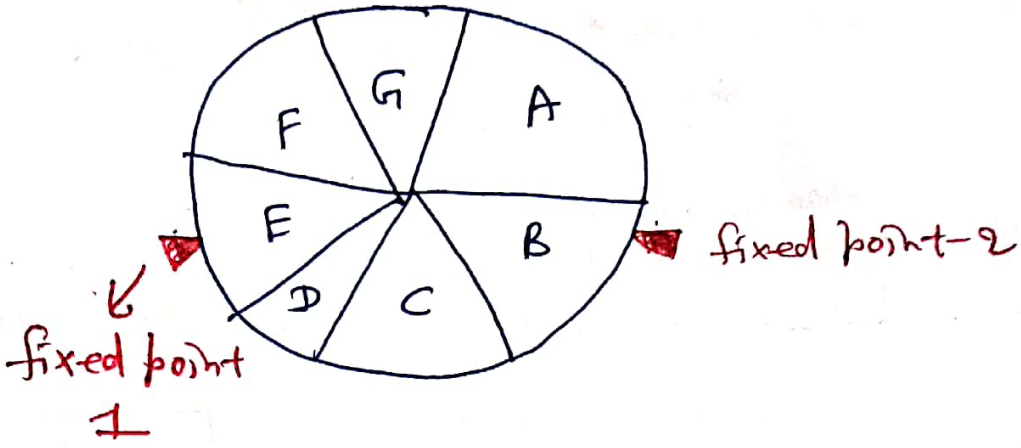
⑥ Repeat step ④ and ⑤ till the required No. of parents are selected.

(ii) Stochastic Universal Sampling

* It is quite similar to Roulette wheel selection
* However in this, instead of just one fixed point, we have multiple fixed points as shown

in figure.

* In this, all the parents are chosen in just one spin of the wheel.



(2) Tournament selection method:

* In this, we select k-individuals from the population at random and select the best out of these to become a parent.
* The same process is repeated to select the next parent.

(3) Tournament Selection method:

* In this method, we select k-individuals from the population at random and select the

(3) Rank Selection method

* Rank selection method also works with negative fitness values and is mostly used when the individual in the population have very close fitness value.
* It leads to each individual having an almost equal share on the pie.

(4) Random Selection method

In this method, we randomly select parents from the existing population. There is no selection pressure towards fitter individuals

Crossover and its types

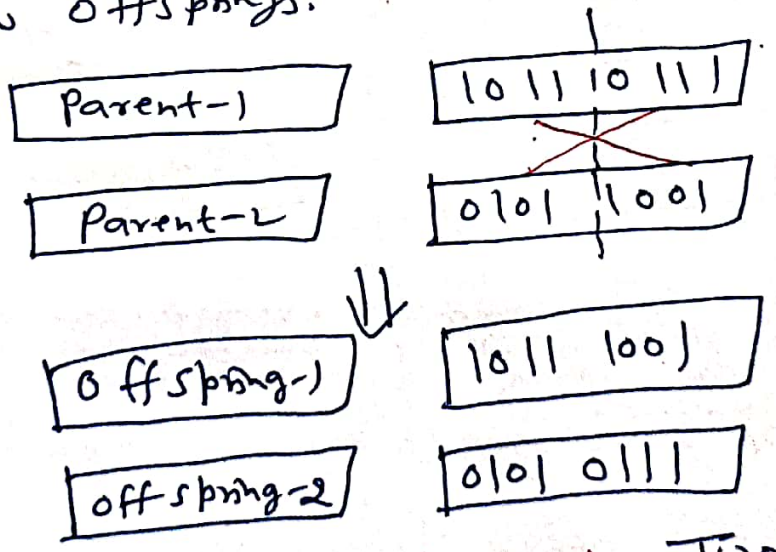
Crossover: Crossover is a genetic operator that combines (mates) two chromosomes (parents) to produce a new chromosome (offspring).

* The idea behind crossover is that the new chromosome may be better than both of the parents if it takes the best characteristics from each of the parents.

Types of Crossover:

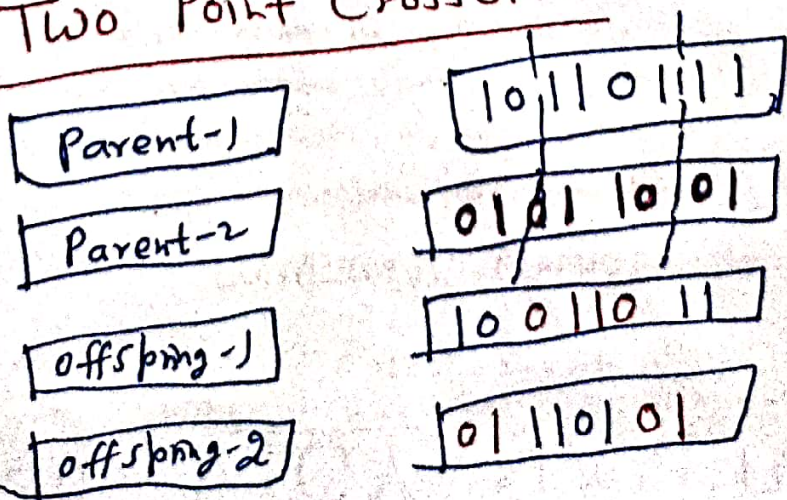
(1) One point Crossover:

In one point crossover, a random crossover point is selected and the tails of its two parents are swapped to get new offsprings:



(2) Two Point Crossover:

Two random crossover point is selected and swapped to get new offsprings



③ **Uniform Cross-over:** In a uniform crossover, we don't divide chromosome into segments, rather we treat each gene separately, we can bias to one parent to have more genetic material in the children (offspring) than the other. We can also use binary mask in uniform crossover.

example-1

Parent-1 0 0 0 | | | 0 0

Parent-2 | | 0 0 | 0 | |

⇓

Child-1 0 | 0 | | | 0

(Biased to Parent-1)

Child-2 | | 0 | | 0 | |

(Biased to Parent-2)

ex: 1

---

example-2

Binary Mask | 0 0 0 | | | 0

Parent-1 0 0 0 | | | 0 0

Parent-2 | | 0 0 | 0 | |

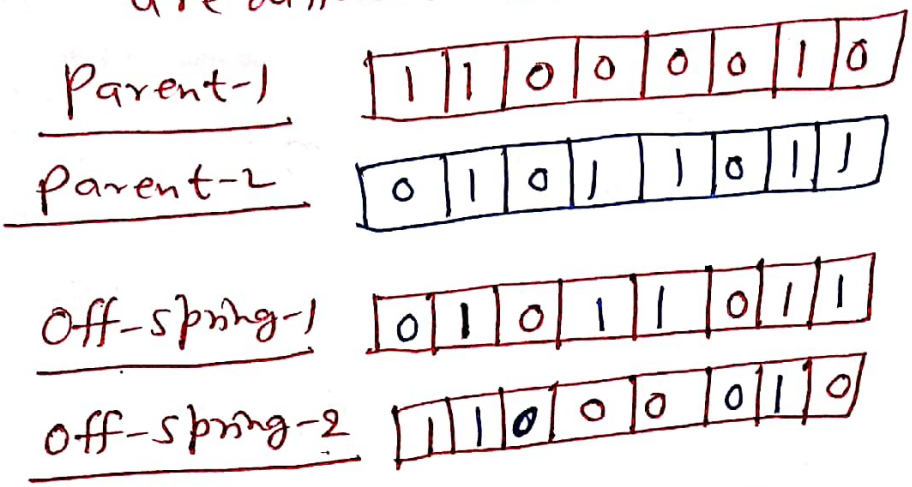
⇓ Crossover (Uniform)

Child-1 0 | 0 0 | | 0 |

Child-2 | 0 0 | | 0 | 0

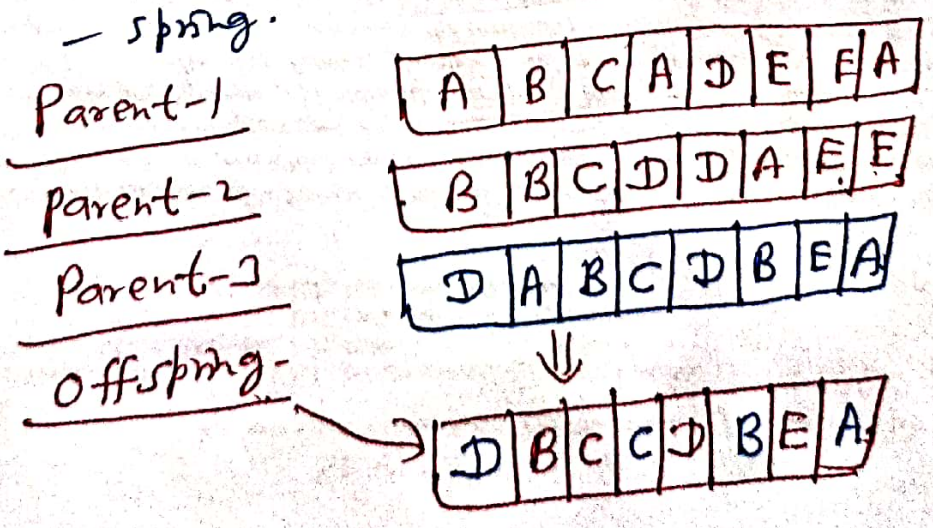
(4) Half-Uniform Crossover

In this method, first the hamming distance (No. of different bits in parent-1 and parent-2) is calculated. Swap the bits in offspring-1 and 2, if bits are different else remain the same.



(5) Three parents Crossover

* In this method, three parents are taken.
* each bit of parent-1 is compared with each bit of parent-2
* If bit is same in both parents, it is taken in offspring; if bit is different in both parents then bit from parent-3 is taken for the off-spring.



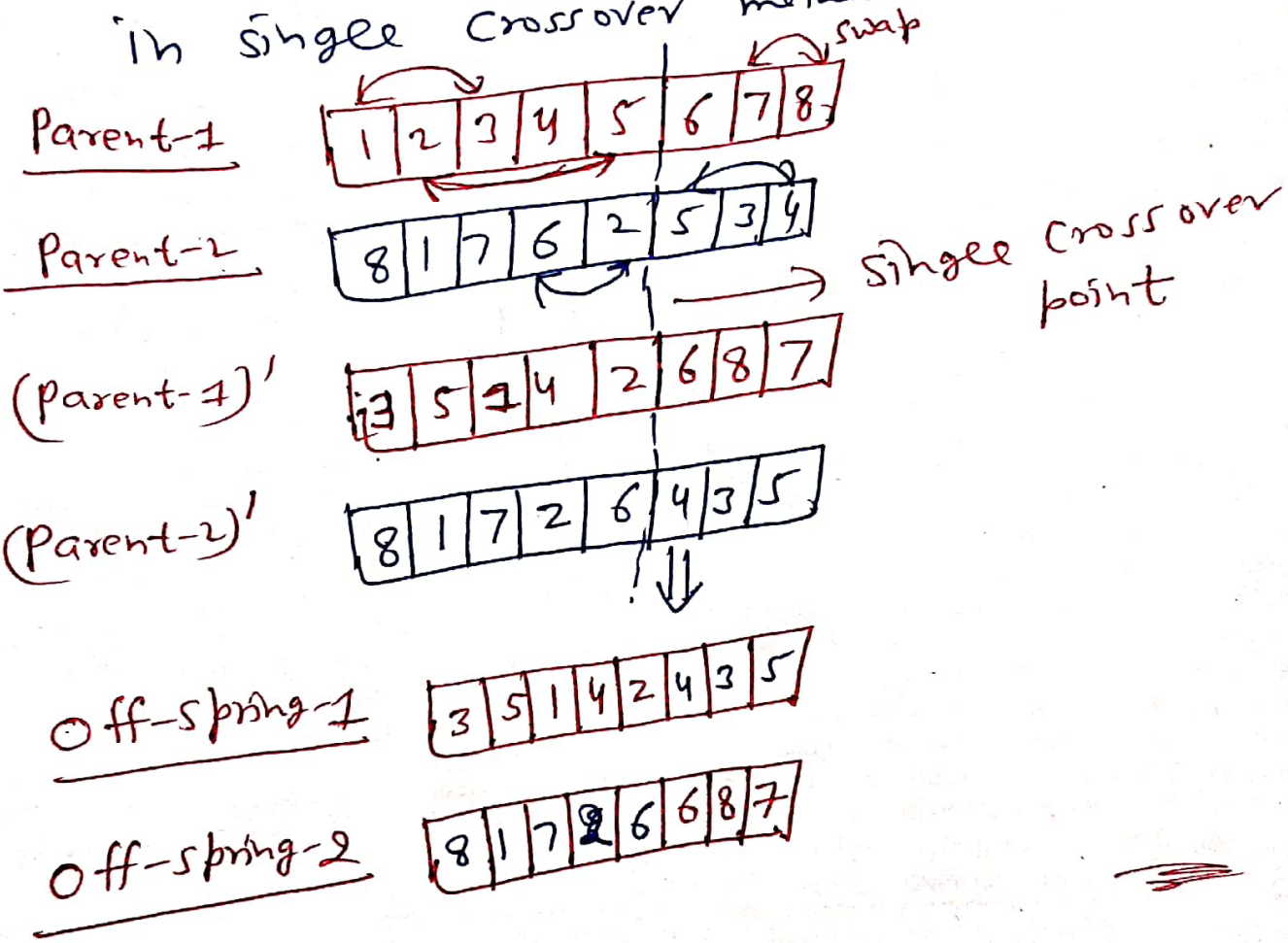
(6) Suffle Cross-over:

In this, a single crossover point is selected. This

point divides the chromosomes in two parts.

* In both parents, suffle bits (genes) in each parts using any logic.

* After Suffle, modified parents are mixed as in singee crossover method.



Mutation and its types

Mutation: Mutation is performed after crossover.

* Mutation is a genetic operator used to maintain genetic diversity from one generation of a population of chromosomes to the next.
* Mutation alters one or more gene values in a chromosome from its initial value.
* With the new gene value, the genetic algorithm may be able to arrive at better solution.
* Mutation is an important part of the genetic search.

Types of Mutation:

① Flip-bit Method (Flipping)

In this, a mutation chromosome is created. for a 1 in mutation chromosome, the corresponding bit in offspring chromosome is flipped (0 to 1 or 1 to 0) and mutated chromosome is produced.

Offspring Chromosome | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |

Mutation Chromosome | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |

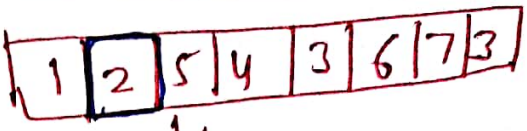
⇓

Mutated Chromosome | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 |

② Boundary Method

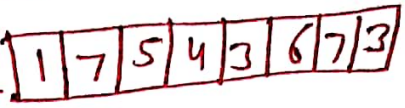
The mutation operator replaces the value of the chosen gene with either the upper or lower bound for that gene (chosen randomly).

Offspring Chromosome



Mutation (Upper bound)

Mutated Chromosome



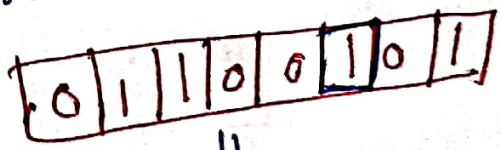
(3) **Interchange Method:** Two position of child's Chromosome are chosen randomly and the bits corresponding to those position are interchanged.

Offspring Chromosome

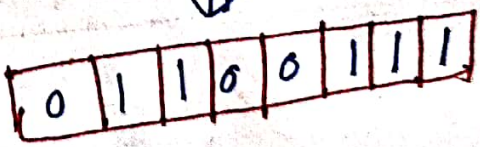
Mutated Chromosome

(4) **Reversing Method:** A position is chosen at random and the bit next to that position is reversed to produced mutated child.

Offspring Chromosome:



Mutated Chromosome



Convergence & Convergence criteria in genetic algorithm
(Termination Condition in GA)

Convergence in GA

for any problem, If GA is correctly implemented the population (set of solutions) evolves over successive generations with fitness value increasing towards the global optimum.

So Convergence is the progression towards increasing uniformity

" A population (set of solutions) is said to have converged when 95% of the individuals (solutions) constituting the population share the same fitness value.

Convergence (Termination) Criteria :

Each iteration should be tested with some convergence test.

Various Convergence (Termination) Criteria / Conditions are as given:

1. A solution is found that satisfy the objective criteria.
2. fixed no. of generation is executed.
3. Allocated budget (computation time) reached.
4. The highest ranking solution fitness is reaching or has reached a point such that successive iterations no longer produce better result.
5. Manual Inspection.
6. Combination of the above.

Application of Genetic algorithms (GAs)

Genetic Algorithms are primarily used in optimization problem of various kinds, however they are also used in other application areas as well.

(1) Optimization GAs are most commonly used in optimization problems where we have to maximize/minimize a given objective function value under a given set of contraints.

(2) Neural Networks : GAs are also used to train neural networks specially recurrent neural network.

(3) Image Processing :- GAs are used for various digital image processing tasks.

(4) DNA analysis :- GAs have been used to determine the structure of DNA.

(5) Travelling salesman problem :- GAs are also used to solve the Travelling salesman problem (TSP).

(6) Scheduling Application :- GAs are used in solving various scheduling problem.

(7) Machine Learning :- GAs are also used in GBML (Genetic Based Machine Liscencing).

(8) Parametric design of Aircrafts
GAs are also used to design aircrafts by varying the parameters and evolving better solutions.

(9) Robotics :- GAs are used to create, learning robots which will behave as a human.

(10) Economics :- GAs are used to characterize various economic models like cobweb model, game theory equilibrium resolution and asset pricing etc.