Genetic Algorithm (GA)

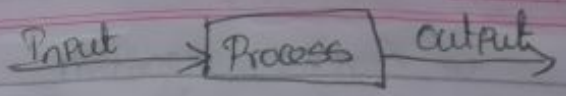
(*) Genetic Algorithm (GA) is a Search based optimization technique (heuristic search) based on the Principles of Genetic and Natural Selection.

(*) It is frequently used to solve optimization Problems, in ~~xxxx~~ research, and in machine Learning.

Optimization :

Optimization is the Process of making Something better.

In any Process, we have a set of Inputs and a set of outputs



Optimization refers to Finding the Values of Inputs in such a way that we get the "Best output".

Genetic Algorithms were developed by John Holland and His Students and Colleagues of University of Michigan.

Genetic Algorithms are a subset of a much larger branch of Computation known as Evolutionary Computation.

Natural Selection:

(x) Genetic Algorithm Simulate the Process of natural selection. which means, who can adapt to Changes in their environment are able to Survive and reProduce and Goto next Generation.

Adapting Changes → Survive → ReProduce

In Simple words "Survival of Fittest"

Genetic:

(x) Each Generation Consist of a Population of Individuals.

(x) Each Individual represents a Point in Search Space and possible solution.

(x) Each Individual is represented as a string of character is called Chromosome.

In humans, each cell Contains 23 Pairs of chromosomes, for a total of 46.

22 of these Pairs are called autosomes.

Autosomes are look the Same in Both Male and Female.

23rd Pair is Sex chromosomes. That is differ between male and Female.

Females have Two Copies of X-X (Chromosomes)

Males have one X-Y one (Chromosome)

Pairs of chromosomes called "Karyotype"

Foundation of Genetic Algorithms: (Coding)

(x) GA are based on an Analogy with Genetic Structure and Behaviour of chromosome of the Population.

(x) Individual in Population Compete for resources and mate.

(x) Those Individuals who are successful (fittest) then mate to create more offspring than others.

(x) Genes from "fittest Parent" throughout the Generation.

(x) Thus each successive generation is more suited for environment.

Gene:

A Gene is the basic physical and functional unit of Heredity (वंशानुगत).

(x) Genes are made up of DNA.
DNA => Deoxyribonucleic acid
RNA => Ribonucleic Acid.

(x) Some Genes act as Instructions to make Molecules called Proteins.

(x) Many Genes do not code for Proteins.

Human Genes are varying in size. from a few Hundred DNA bases to more than 2 million bases.
2 Million = 20,00,000
(Twenty lakhs)

An International research effort called Human Genome Project, to determine the human genome and Identify the genes it Contains. estimated. 20,000 and 25,000 Genes.

Each Person has 2 Copies of each Gene. One Inherited from each Parent.

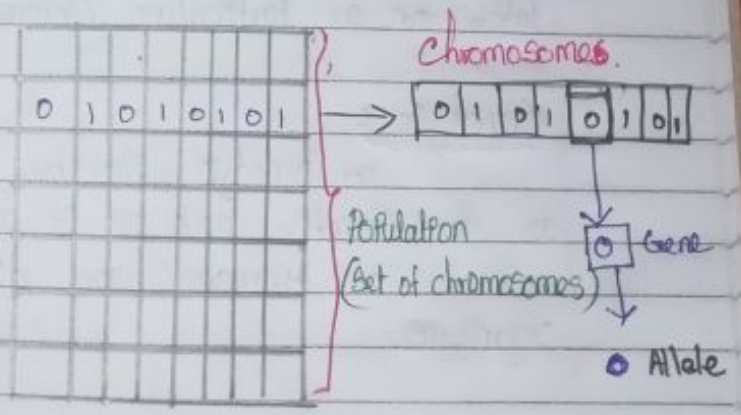
Most Genes are the Same In all People, but a Small number of genes are Slightly different between People.

Basic Terminology :

Population :

(*) It is a Subset of all possible (encoded) Solutions to the given Problem.

(*) The Population of GA is analogous to the Population for human beings except that Instead of human beings, we have Candidate Solutions representing human beings.



Chromosomes :

A chromosome is one such solution to the given Problem.

Gene :

A Gene is one element position of a Chromosome.

Allele :

It is the value a gene takes for a particular Chromosome.

Genotype :

(x) Genotype can be described as the Genetic make up of an organism.

(x) Humans are diploid organisms.

Diploid => A cell that contains 2 copies of each chromosome.

Haploid => A cell that contains single set of chromosome

Humans have Two alleles at Genetic position, with one Inherited from Parent.

Phenotype :

Phenotypes are Physical Properties of an Organisms, which can be observed with our eyes.

(x) Phenotype is determined by Individual Genotype and Expressed Genes or by visible.

Genotype	Phenotype
✍ The Hereditary Information of the Organisms in the form of Gene in DNA, and remains the same through out the life	✍ The characteristics of an organism, which are visible.
✍ The organism's heredity characters that may (or) may not be expressed in next Generation.	✍ Characters are not Inherited
Same Genotype Produce Same Phenotype.	Same Phenotype may (or) may not belong to same Genotype.
✍ Present Inside the human body. (Not visible)	✍ Expression of Genes as the External appearance.
✍ Genotype is Inherited from the Parent to offspring.	✍ The Phenotype is not Inherited from the Parent.
✍ It Can be determined by Scientific Method such as Polymerase chain reaction.	✍ It Can be determined by observing Method of organism.
✍ It is affected by Genes.	✍ It's affected by Genotype and Environment Conditions.
eg: Blood group, eye colour, height, genetic disease.	eg: weight, character, Behaviour.

Encoding:

Encoding is a Process of Transforming from Phenotype to Genotype.

Decoding:

Decoding is a Process of Transforming from Genotype to Phenotype.

Fitness Function:

A Fitness Function is a Function which takes the Solution as Input and Produce the Suitability of the Solution as output.

Genetic Operators:

These alter the genetic Composition of the offspring.

Blood Group:

Human blood type is determined by Codominant alleles.

(x) There are 3 different alleles

(i) I^A (ii) I^B (iii) i

I^A and I^B are Codominant.

i allele is recessive.

Human Phenotypes for Blood group are. type A, type B, type AB, Type O.

Type A and Type B => homozygous (I^AI^A (or) I^BI^B)

(or)

heterozygous (I^Ai (or) I^Bi)

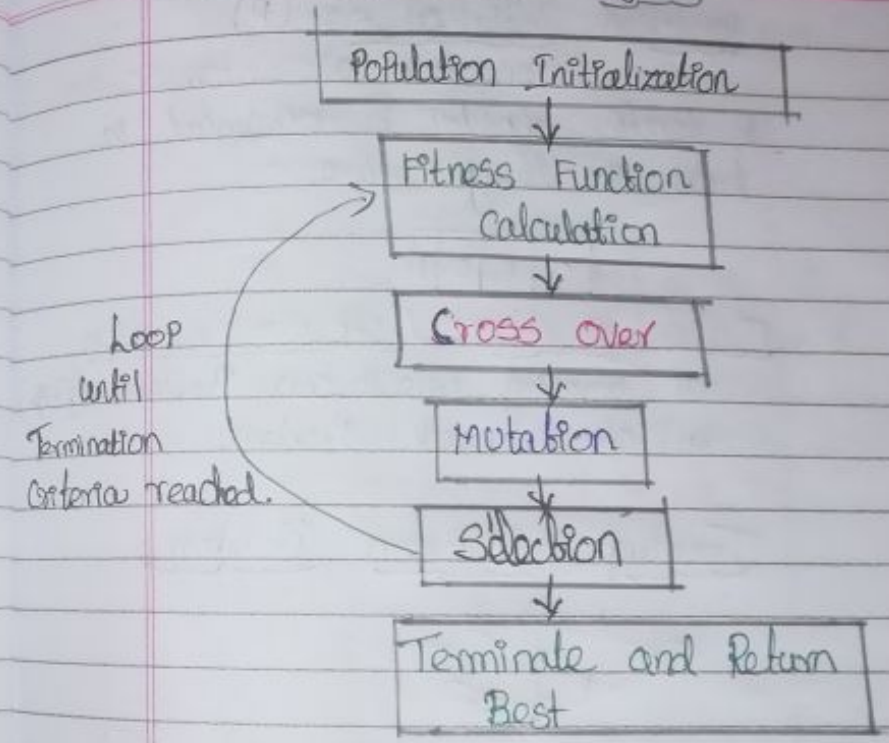
Blood type and Genotype :

Phenotype	Genotype
Type A	$I^A I^A$ / $I^A i$
Type B	$I^B I^B$ / $I^B i$
Type AB	$I^A I^B$
Type O	ii

Type A and Type B Cross :

Mother (type A)		Father (type B)
$I^A I^A$	or	$I^B i$
$I^A I^A$	or	$I^B I^B$
$I^A i$	or	$I^B I^B$
$I^A i$	or	$I^B i$

Genetic Algorithm Life cycle :



Population Initialization: (P)

Population of chromosomes in Genetic algorithm is represented in terms of Binary numbers.

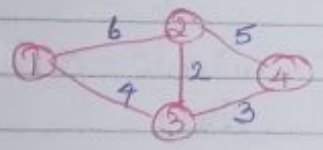
0	1	0	1

Fitness Function: / objective Function.

(1) Traveling Salesperson Problem (TSP)

(2) 0/1 Knapsack Problem.

Traveling Salesperson Problem:



Path 1: 1 – 2 – 4 – 3 – 1 = (18)
6 5 3 4

Path 2: 1 – 3 – 4 – 2 – 1 = (18)
4 3 5 6

You can visit every node at once and return back to source node.

Always select the minimum.

0/1 Knapsack Problem.

(A) Neuron has weight of 5 and Profit is 100, (B) Neuron has weight of 10 and Profit is 150, (C) Neuron has a weight of 15 and Profit is 200. and maximum weight is 25.

Find the Fitness?

Solution:

	W	P
(A)	5	100
(B)	10	150
(C)	15	200.

	A B C	W	P.
(A)	100	5	100
(B)	010	10	150
(C)	001	15	200

	w	p	
A	100	5	100
AB	110	15	250
AC	101	20	300
BC	011	25	350 Maximum
ABC	111	30	

Q1. The Population of chromosomes in Genetic algorithm is represented in terms of binary number. The Strength of fitness of a chromosome in decimal form x, is given by

$$S f(x) = \frac{f(x)}{\leq f(x)} \quad \text{where } f(x) = x^2$$

Population (P) = {(01101), (11000), (01000), (10011)}

Find the Strength of chromosome (11000) is?

32 16 8 4 2 1

0 1 1 0 1 = 13

1 1 0 0 0 = 24

0 1 0 0 0 = 8

1 0 0 1 1 = 19

Solution:

Population	Decimal (f(x))	$f(x) = x^2$
01101	13	169
11000	24	576
01000	8	64
10011	19	361
		1170

$$S f(x) = \frac{f(x)}{\leq f(x)} = \frac{\quad}{1170} = \quad$$

but Question is.

The Strength of chromosome (11000) is?

11000 = 24 $f(x) = x^2$ $f(x) = 24$ $f(x) = 576$

$\leq f(x) = 1170$

So $= \frac{576}{1170} = 0.492$

Genetic Algorithms are
Adaptive heuristic Search Algorithm
that belong to the larger Part of
Evolutionary Algorithms.

Each Generation Chromosome (candidate solution) Undergo Mutation, Cross over, Selection to Produce better Population as offspring.

Mutation:

Mutation is a Unary Operator, So One Parent is enough.

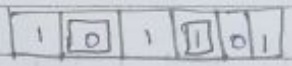
There are 5 types of Mutation Algorithms.

(1) Bit Flip Mutation
(2) Random Resetting Mutation
(3) Swap Mutation
(4) Scramble Mutation
(5) Inversion Mutation.

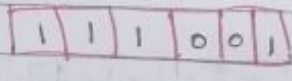
Bit Flip Mutation:

In Population of Chromosome you can select one (or) more gene and Flip their values.

Ex: Change 0 to 1, 1 to 0.



← Before Mutation



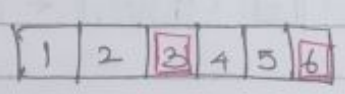
← After Bit Flip Mutation.

Random Resetting mutation:

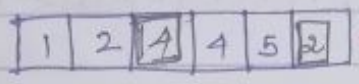
In Population of chromosome you can select one or more gene and replace their values with another random value from their range.

e.g:

Population of chromosome of gene range from [1-6], the random resetting mutation range also [1-6], you can select one value and replace it.



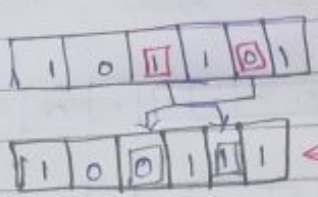
← Before Mutation



← After Random Mutation

Swap mutation:

Select Two genes from the chromosome and Interchange it.



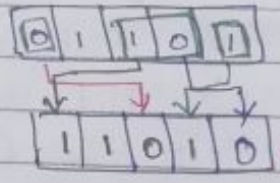
← Before Mutation

← After Swap mutation.

Scramble mutation:

You can select a subset of Gene from chromosome and scramble their value.

The selected Genes may not be Contingucus.

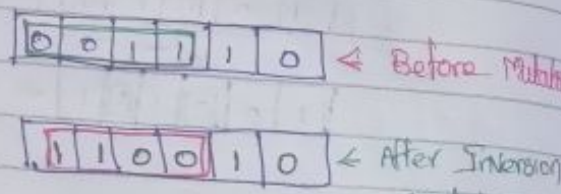


← Before Mutation

← After Scramble Mutation.

Inversion Mutation:

You can select a subset of Gene from Chromosome and reverse their order.



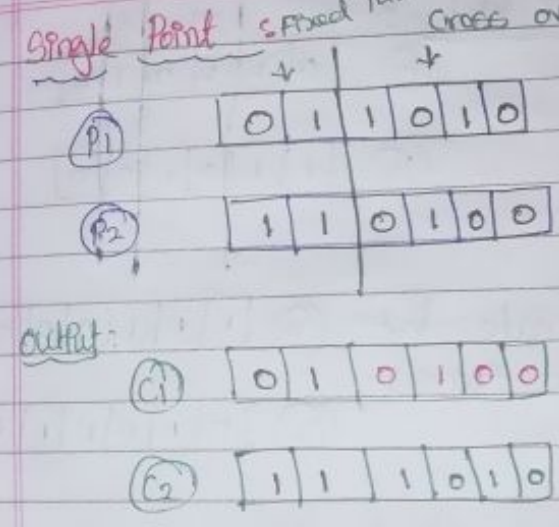
Cross Over:

Cross over need Two Parent Chromosomes to Produce offsprings by using Cross over.

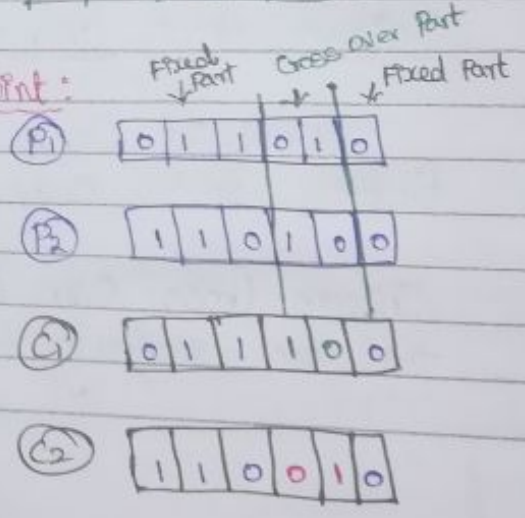
There are 3 types of Cross over.

(i) Single Point Cross over
(ii) Two Point Cross over
(iii) Multi Point Cross over.

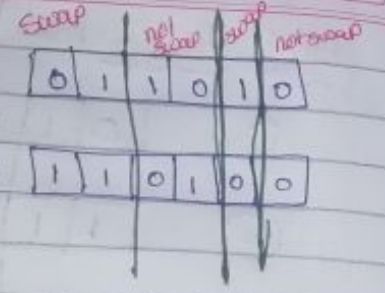
Single Point:



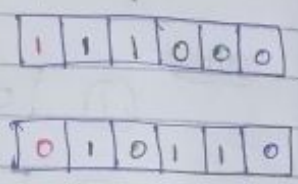
Two Point:



Multi Point :



Alternate Part to be swapped :



Uniform Cross over :

Multipoint Cross over is the Uniform Cross over.

Uniform Cross over Generate Mask that are randomly Generated.

Important ! For 1st offspring :

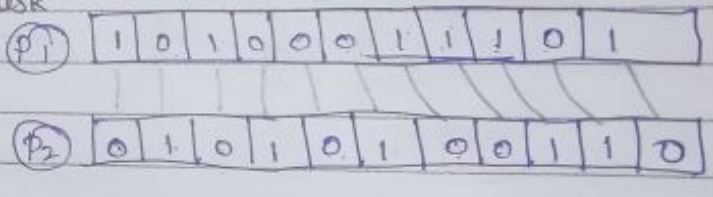
- when 1 in Mask = Parent 1 Gene
- when 0 in Mask = Parent 2 Gene.

For 2nd offspring :

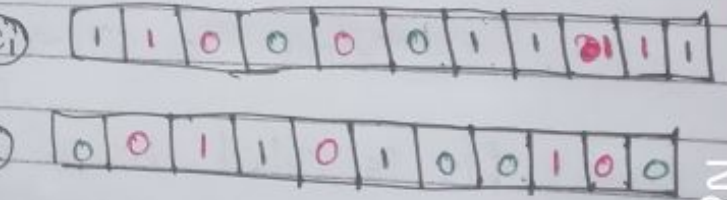
Change the Parent 2 at Top & Parent 1 at Bottom.

eg :

Cross over mask 1 0 0 1 0 1 1 1 0 0 1



output :



Selection:

Selection is a process of selecting individual from population based on fitness.

Reproduction is also known as selection.

There are 5 types

1. Roulette wheel selection
2. Boltzmann selection
3. Rank selection
4. Tournament selection
5. Steady state selection

Roulette wheel selection:

Select the fitness

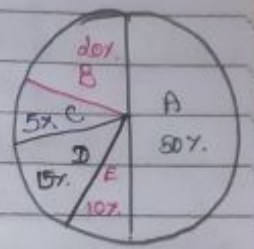
Selection ∝ fitness

This is called fitness probability selection.

eg:

Chromosome	Fitness	Percentage
A	5	50%
B	2	20%
C	0.5	5%
D	1.5	15%
E	1	10%
		100%

Roulette wheel



Output:

More the area occupied by the particular individual can be selected.

So A chromosome is selected.

Rank selection:

Disadvantage of Roulette wheel selection is most area occupied individual can be selected, others are ignored.

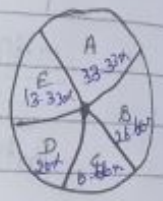
Rank based selection providing the opportunities to all.

eg:

Chromosome	Fitness	Rank	Percentage
(A)	5	5	33.33%
(B)	2	4	26.66%
(C)	0.5	1	6.66%
(D)	1.5	3	20%
(E)	1	2	13.33%
		15	

Calculate the Percentage:

$$\text{Percentage} = \frac{\text{Rank}}{\sum \text{Rank}} \times 100$$



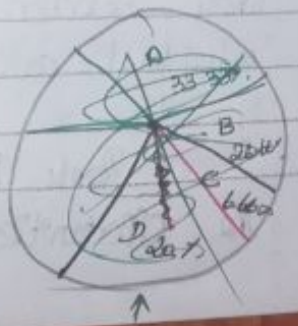
(i) $\frac{5}{15} \times 100 = 33.33$

(ii) $\frac{4}{15} \times 100 = 26.66$

(iii) $\frac{1}{15} \times 100 = 6.66$

(iv) $\frac{3}{15} \times 100 = 20\%$

(v) $\frac{2}{15} \times 100 = 13.33$

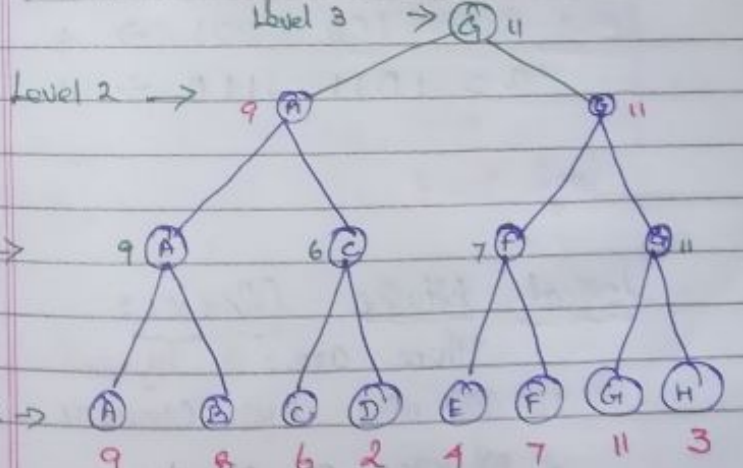


Tournament Selection:

chromosome	Fitness
A	9
B	8
C	6
D	2
E	4
F	7
G	11
H	3

Tournmant Selection is based on Tree Model

G is selected



Bit wise Operators:

There are 3 types.

1. The One's Complement Operator.
2. Logical Bitwise Operator.
3. Shift Operator.

1. One's Complement Operator:

One's Complement Operator is ~
Unary Operator.

Inverse 0 to 1, 1 to 0.

a = 0100 0001 → 4 1
~a = 1011 1110 → 11 14

length is = 8

Logical Bitwise Operator:

There are 3 types.

1. Bit-wise AND Operator
2. Bit-wise OR Operator
3. Bit-wise Exclusive-OR Operator

Bit-wise AND Operator:

Bit-wise AND Operator is &
It returns 1 if both bits otherwise 0

eg:

P1 = 1010 1010 → 10 10
P2 = 1100 0011 → 12 3
output C1&2 = 1000 0010 → 8 2
C2

Bit-wise OR Operator:

Bit-wise OR Operator is |
It returns 1 if any bits have 1.
otherwise 0.

eg:

P1 = 10 10 1010 → 10 10
P2 = 1100 0011 → 12 3
output
C1|C2 = 1110 1011 → 14 11

Bitwise Exclusive - OR :

Bitwise Exclusive-OR Operator is ^
It returns 1 if any one of bits have 1 but not in both

eg:

P_1 = 1010 1010 → 10 10

P_2 = 1100 0011 → 12 3

output

a^2c_2 = 0110 1001 → 6 9

Shift Operator :

There are 2 shift Operator.

1. Shift left Operator <<
2. Shift right Operator. >>

Shift left Operator (<<) :

All the bits in First Operand to be shifted to the left, no. of positions Indicated by Second Operand.

- left most bit in original bit Pattern is lost.
- Right most bit are Vacant and filled with Zero's.

eg:

a = 1010 0110 → 10 6

a << 2.

Solution:

a = 1010 0110 → 10 6.

a << 2 = 1001 1000 = right most bit vacant

(9) (8)

Shift - right Operator (>>)

First Operand shifted to right by the no of positions indicated by second Operand.

(x) right most bits of the original bit's are lost.

(x) Left most bit position be vacant and filled with zero.

eg:

a = 1010 0110 → 10 6

a >>2

Solution:

00 10 1001 → 2 9

vacant

Genetic schema:

Schema is a Template, in Genetic Algorithm that Identify a subset of strings with similarities at certain string positions.

A Template is made up of 0's and 1's and *'s (* is a dont care symbol) that matches either 0 or 1.

eg:

Schema *111***01.

Consider schema you can find

(i) Order (ii) length

Order: o(h).

The order of schema is the number of fixed bits in the string.

fixed bit

*111***01 o(h) = 5.

don't care

Length :

Length of the schema is Consider the Distance between first & last fixed bits.

eg:

1 2 3 4 5 6 7 8 9
* 1 1 1 * * * 0 1

first fixed bit Positions = 2
Last fixed bit Position = 9

Length :
9 − 2 = 7

A Schema of Order O represent 2^{N-O} different strings of length N.

	Schema	Order "O"	2^{N-O}	Represented strings
$2^3=8$	* * *	0	2^{3-0}	000, 001, 010, 011, 100, 101, 110, 111.
$2^2=4$	* 1 *	1	2^{3-1}	010, 011, 110, 111.
$2^1=2$	* 1 0	2	2^{3-1}	010, 110
$2^0=1$	* 1 0	3	2^{3-3}	110.

Q The Order of Schema ?10?101? and ???0??1 are ___ and ___ respectively.

Solution:

?10?101? = 5

???0??1 = 2

So Answer 5 and 2 respectively