Machine Intelligence
Fitness, Selection and Genetic Operators

Dr. Arti Arya
Department of Computer Science
artiarya@pes.edu
+080-66186629 Extn 6629

MACHINE INTELLIGENCE

Fitness, Selection and Genetic Operators

Dr. Arti Arya

Department of Computer Science and Engineering

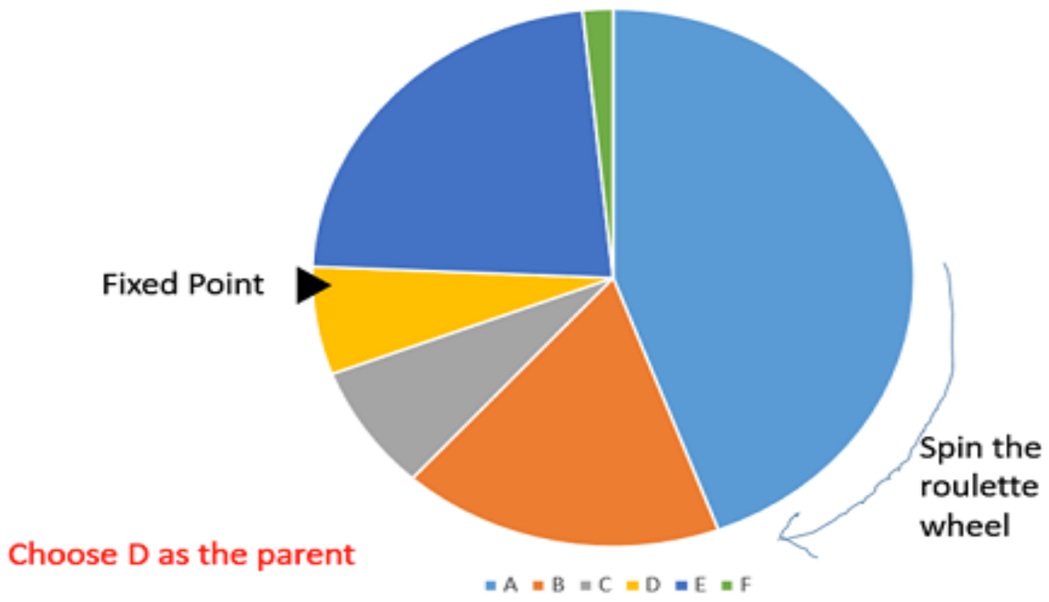
The **fitness function defines the criterion for ranking potential hypotheses** and for probabilistically selecting them for inclusion in the next generation population.

- **Fitness proportionate selection** (***Roulette wheel***) **:** Ratio of fitness to the fitness of other members of the current population.

Roulette Wheel Selection

- In a roulette wheel selection, the circular wheel is divided as described below. A fixed point is chosen on the wheel circumference as shown and the wheel is rotated.

- The region of the wheel which comes in front of the fixed point is chosen as the parent. For the second parent, the same process is repeated.

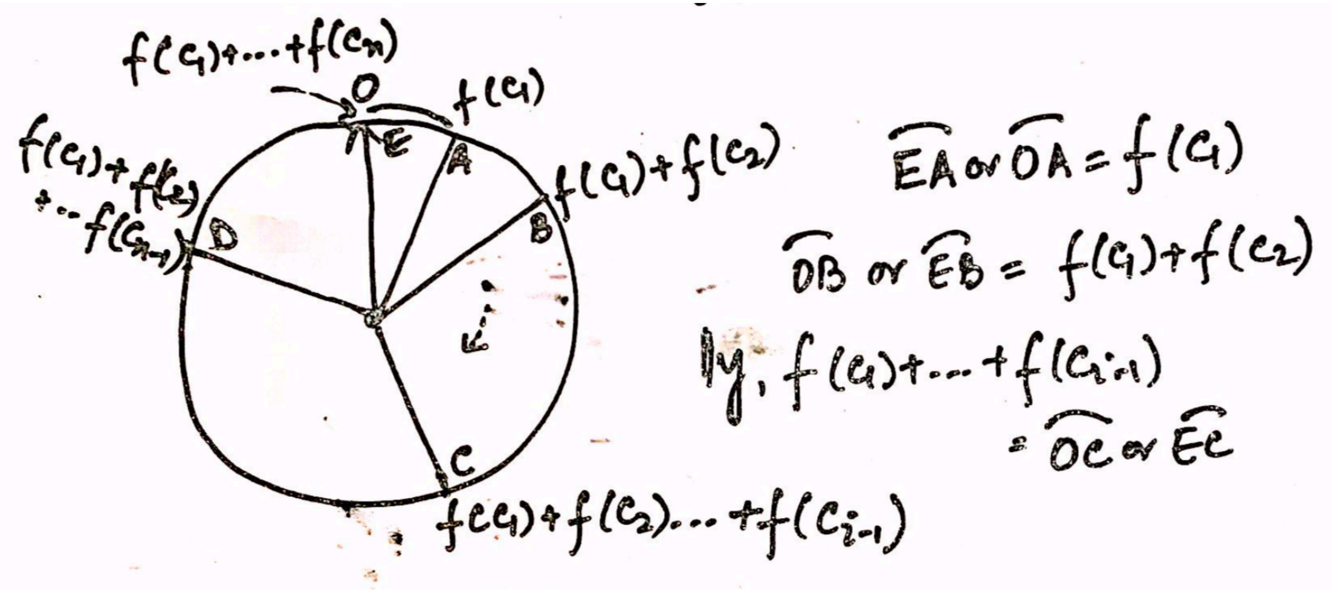


Chromosome	Fitness Value
A	8.2
B	3.2
C	1.4
D	1.2
E	4.2
F	0.3

Roulette Wheel Selection-Example

Consider a population of size n.

Let c_1, c_2,…..c_n be n chromosomes with fitness values $f(c_1)$, $f(c_2)$,….$f(c_n)$.
Suppose there is a disc of circumference= $\sum_{i=1}^{n} f(c_i)$



Roulette Wheel Selection-Example

In order to select a chromosome to be included in the mating pool from the current population,

a random no. is generated betn
0 & $\sum_{i=1}^{m} f(c_i)$

Let one such random no. be 'x'.
Now locate a point Y on the circumference of the wheel which is at a distance 'x' from the starting point O or E.

The chromosome within whose limits Y is located will be selected for the mating pool.
This process is repeated 'n' times as population size=n.

Roulette Wheel Selection-Example

Keeping a TSP for 5 cities in mind. Consider the following set of chromosomes(possible solutions) in an initial population:

S.No	Chromosome	Fitness
1	d-e-a-b-c	193
2	c-e-b-d-a	172
3	e-a-b-c-d	193
4	d-e-a-b-c	193
5	a-d-e-b-c	141
6	c-e-d-a-b	197
7	e-a-d-b-c	222
8	c-d-e-a-b	193
9	e-a-b-d-c	224
10	a-d-e-b-c	141

1089

1311

$\sum f(c_i)=1869$, i=1,…10

Let a random no. is generated between (0,1869), say 1279

(it can be any number between (0,1869)

Pls refer [2] in References for a simple mathematical problem using GA

Since 1279 lies between $\sum f(c_i)=1089$, i=1…6. and $\sum f(c_i)=1311$, i=1….7.

As 1089 < 1279 < 1311

So, 7th Chromosome is selected.

- **Tournament Selection:**
  - Two hypothesis (chromosomes) chosen at random, with some predefined probability, p, the more fit is selected.

- **Rank Selection:**
  - Sorted by fitness. The probability that a hypothesis will be selected is then proportional to its rank.

- **Elitist Selection (Elitism):**
  - Ensures that the best chromosomes are retained in the next generation, if not selected by any other method.

THANK YOU

Dr. Arti Arya
Department of Computer Science
artiarya@pes.edu
+91 9972032451 Extn 029

Machine Intelligence
Genetic Operators

Dr. Arti Arya
Department of Computer Science
artiarya@pes.edu
+080-66186629 Extn 6629

MACHINE INTELLIGENCE

Genetic Operators

Dr. Arti Arya
Department of Computer Science and Engineering

Broadly 2 categories of Operators

Crossover (recombination)

Exchanging genetic material between the individuals of the chromosomal population

Mutation

A genetic operator used to maintain genetic diversity from one generation of a population of chromosomes to the next

Note: A genetic operator which works well for one problem may not work well for another problem. [1]

MACHINE INTELLIGENCE
GA Crossover- Single Point Crossover

- Crossover point is randomly selected.

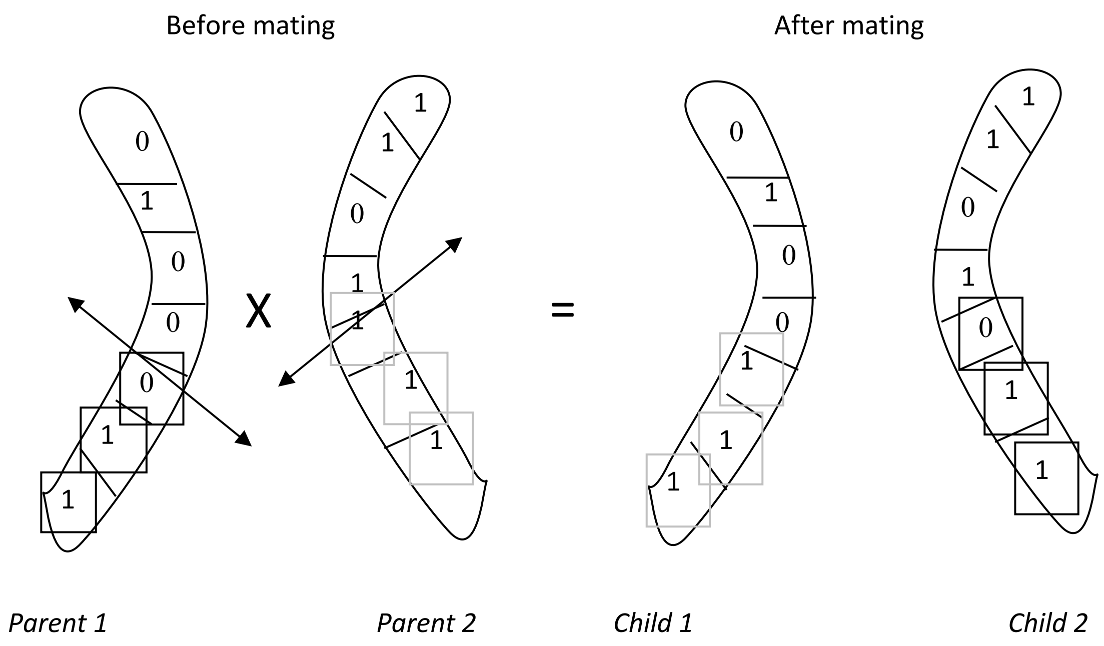


Figure Courtesy: N.P Padhy, "Soft Computing"

GA- Double Point Crossover

- Two random sites are chosen.
- The bits in between the two points are swapped between the parent chromosomes.

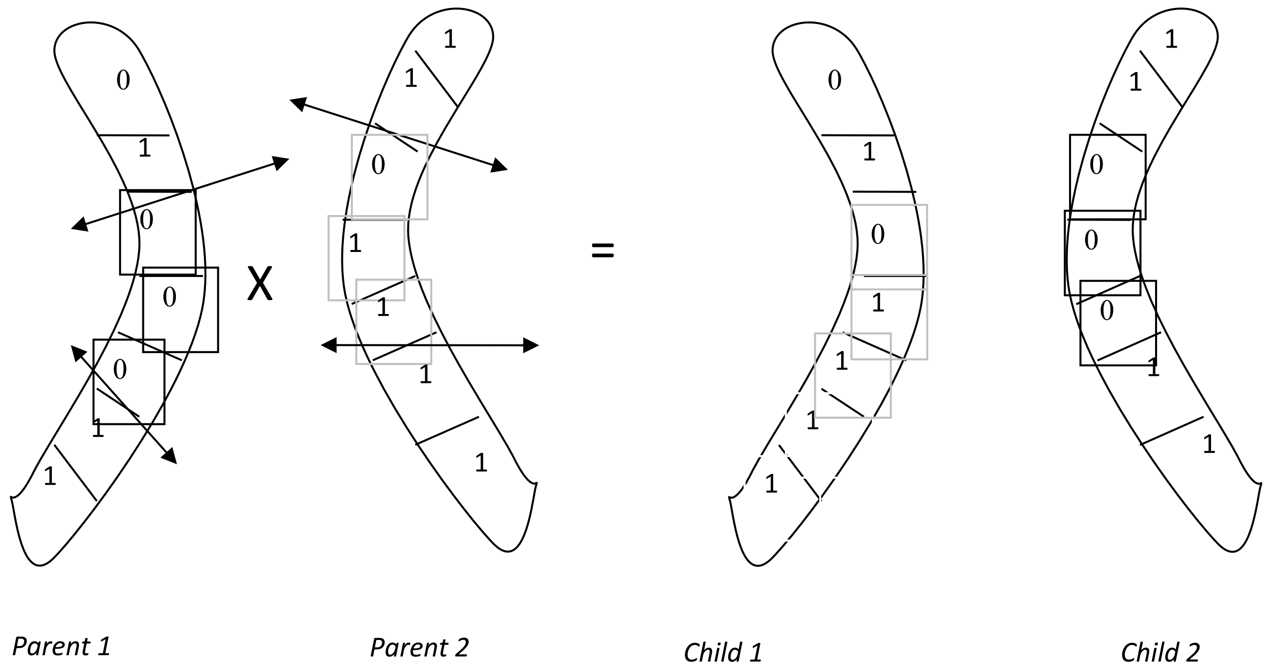
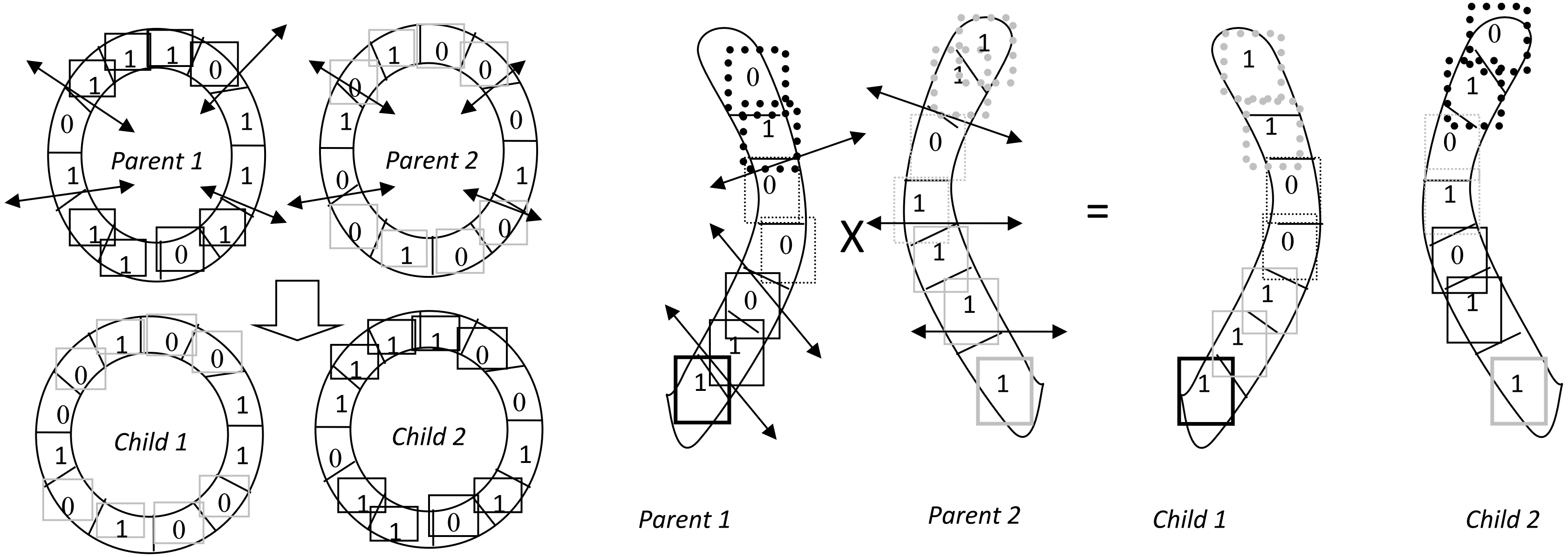


Figure Courtesy: N.P Padhy, “Soft Computing”

GA- Multipoint Crossover



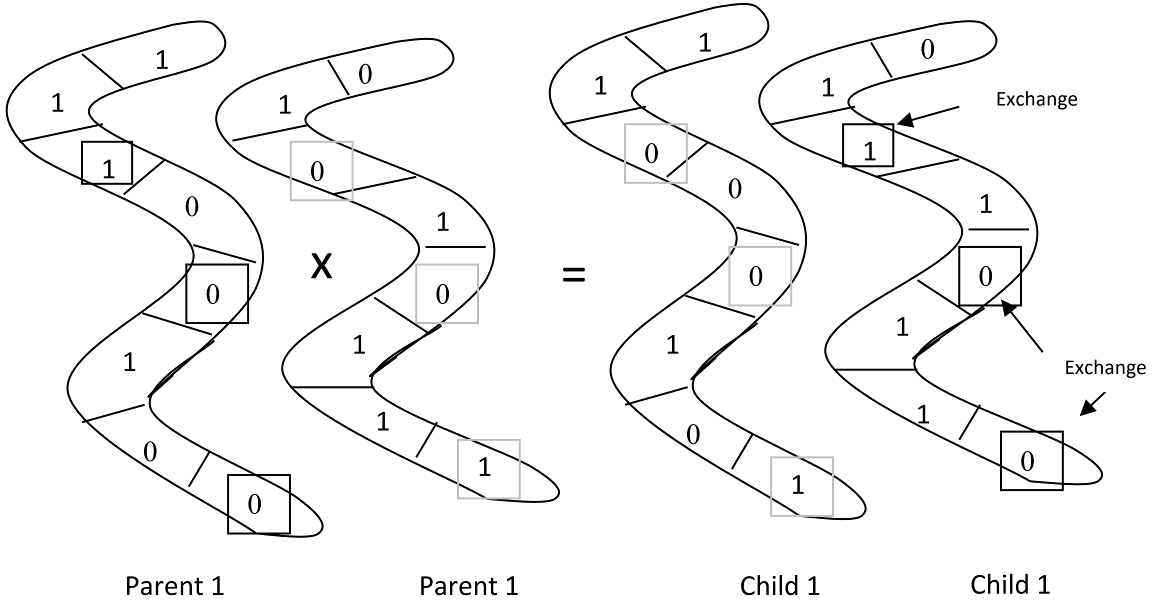
Even cross-site multi - point cross over Figure Courtesy: N.P Padhy, “Soft Computing”

- Multi point crossover is a generalization of the one-point crossover wherein alternating segments are swapped to get new off-springs.

- One-point and N-point crossover have a tendency to exhibit positional bias and inherent bias which have proved to be experimentally evident.

MACHINE INTELLIGENCE

GA- Uniform Crossover



Exchanging bits from two individual parental chromosomes by maintaining a probability of 0.5 to produce offsprings.

Figure Courtesy: N.P Padhy, “Soft Computing”

GA- Masking Based Uniform Crossover

Uniform crossover does not exhibit positional bias but do exhibit distributional bias due to which uniform crossover has a strong tendency towards transmitting 50% of the genes from each parent.

P1	1	1	0	0	0	1	0	1	
Mask	1	0	1	1	0	0	0	1	
P2	1	0	1	0	1	0	0	1	
C1	1	0	0	0	1	0	0	1	P1=1, P2=0
C2	1	1	1	0	0	1	0	1	P1=0, P2=1

GA- Matrix Crossover

P1 X P2

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad X \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

C1 C2

Cross sites of row and column are chosen randomly.

Select any region between in each layer, either vertically or horizontally and then exchange the information in the region between the mated populations.

GA- Order Crossover

P1	4	5	2	1	3	7	8	6	9
C1	5	2	4	1	3	7	8	6	9
P2	5	7	3	2	1	4	6	8	9
C2	5	3	7	2	1	4	6	8	9

- Set of randomly selected string of genes from one parent is considered (1-3-7-8 from parent 1)
- Transfer the selected genes of parent 1 from child 1 at the same locations.

- Delete 1-3-7-8 from *parent 2* and remaining genes are 9,6,4,2,5 in P2.
- Starting from the bottom to top of C1, fill the empty strings of C1 with the remaining genes of Parent 2 ie 9,6,4,2,5.

GA- Position Based Crossover

P1	4	5	2	1	3	7	8	6	9
C1	7	5	3	1	2	4	8	6	9

P2	5	7	3	2	1	4	6	8	9
C2	4	5	2	1	3	7	8	6	9

- Set of randomly selected positions from P1 (5-1-8-6) are copied to C1.
- Delete the selected nodes of P1 from P2.

The remaining nodes 9-4-2-3-7 of *parent 2* is transferred to its child 1 from bottom to top.

GA- Mutation

- Mutation is a genetic operator used to maintain *genetic diversity* from one generation of a population of genetic algorithm chromosomes to the next.

- It is analogous to biological mutation. Mutation alters *one or more gene values* in a chromosome from its initial state.

- Mutation helps preventing the population of chromosomes from becoming too similar to each other, thus slowing or even stopping evolution. So, it helps the algorithm to *avoid local extreme*

- Mutation allows the *development of un-inherited characteristics* --- it promotes diversity by allowing an offspring to also evolve in ways not solely ***determined by inherited traits***.

GA- Mutation

- After crossover, the strings are subjected to mutation.

- This process enables GAs to jump out of local or suboptimal solutions to avoid premature convergence.

- Mutation plays the *role of recovering the lost genetic material* as well as *for randomly distributing genetic info.*

GA- Mutation

a) **Uniform mutation**

$X^t=[X_{1,} X_{2\ldots} X_m]$ - Chromosome

A random number is selected such that k Є [1,m]

$X^{t+1} =[X_{1,} \ldots X'_k \ldots X_m]$ – Off- Spring

X'_k - A random value generated according to uniform probability distribution from the range $[X_k^L, X_k^U]$

b) **Boundary mutation**

$X'_k \rightarrow X_k^L$ or X_k^U

Replacement takes place with equal probability

GA- Mutation

- Mutation can be implemented as *One's complement Operator.*
- Let C= [01101011] be a chromosome, then one's complement operator gives C=[10010100].

- **Inversion**
  - Two positions are randomly selected within a chromosome. Then the substring between these positions is inverted.

$$P_X=[\ 2\ 5\ \underline{4\ 9\ 6\ 8}\ 1\ 3\ 7]$$

$$C_X=[2\ 5\ \underline{8\ 6\ 9\ 4}\ 1\ 3\ 7]$$

- **Insertion**

➢ A node is randomly selected and inserted in a random position.

- P_X=[2 5 4 9 6 8 1 3 7]

- C_X=[2 6 5 8 9 4 1 3 7]

- **Heuristic Mutation**

➢ It is based on neighborhood technique

P_X=[1 2 3 4 5 6 7 8 9]

Neighbors are formed with these nodes.

N_X=[1 2 3 4 5 8 7 6 9]

N_X=[1 2 8 4 5 3 7 6 9]

N_X=[1 2 8 4 5 6 7 3 9]

N_X=[1 2 6 4 5 8 7 3 9]

N_X=[1 2 6 4 5 3 7 8 9]

After evaluating all the neighbors the best neighbor is selected.

GA- Crossover and Mutation related terms

Crossover Probability:
Measure of likeliness that the selected chromosomes will be undergo crossover operation.

Crossover rate:
The total no. of crossovers that will be performed on the selected chromosomes.

Mutation Probability:
Measure of likeliness that random elements of the chromosomes will be changed with something else.

Mutation rate:
The no. of chromosomes that will undergo the mutation out of a total population.

Genetic Algorithm vs Genetic Programming (FYI)

Genetic Algorithm

1.Often produces invalid states.

2. GA’s are probabilistic search algorithms which mimic the process of natural evolution.

3. In GA’s, each individual is a candidate solution.

4. GA has an o/p which is a quantity

Genetic Programming

1. Special case of GAs, where each individual is a computer program

2. GP explores the algorithmic search space and evolve computer program to perform a defined task.

3. GP is an application of GA. O/p of GP is another program.

References

[1] Xin Yao, An empirical study of genetic operators in genetic algorithms, Microprocessing and Microprogramming, Volume 38, Issues 1–5, 1993, Pages 707-714, ISSN 0165-6074.

[2] Hermawanto D. Genetic algorithm for solving simple mathematical equality problem. arXiv preprint arXiv:1308.4675. 2013 Aug 16.

[3] An empirical study of genetic operators in genetic algorithms X Yao Microprocessing and Microprogramming, 1993 - Elsevier

THANK YOU

Dr. Arti Arya
Department of Computer Science
artiarya@pes.edu
+91 9972032451 Extn 029