Artificial Intelligence

Dr. Mubashir Ahmad (Ph.D.)

Genetic Algorithm

- A **genetic algorithm** is a search heuristic that is inspired by Charles Darwin's theory of natural evolution. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce offspring of the next generation.

- Genetic Algorithm? What is it about? It is a heuristic search that was inspired by the great Charles Darwin's theory of evolution. Once again the computer science field has drawn its inspiration from the biological field.

Genetic Algorithm

- Genetic Algorithms are used to provide quality solutions for optimization problems and search problems. In this article, we would demonstrate the use of the genetic algorithm in searching for some strings.

- The genetic algorithms simulate the survival of the fittest amongst individuals of consecutive generations to solving a problem. So before we delve in too deep, let us remind ourselves of some key Charles Darwin principles in the theory of natural selection.

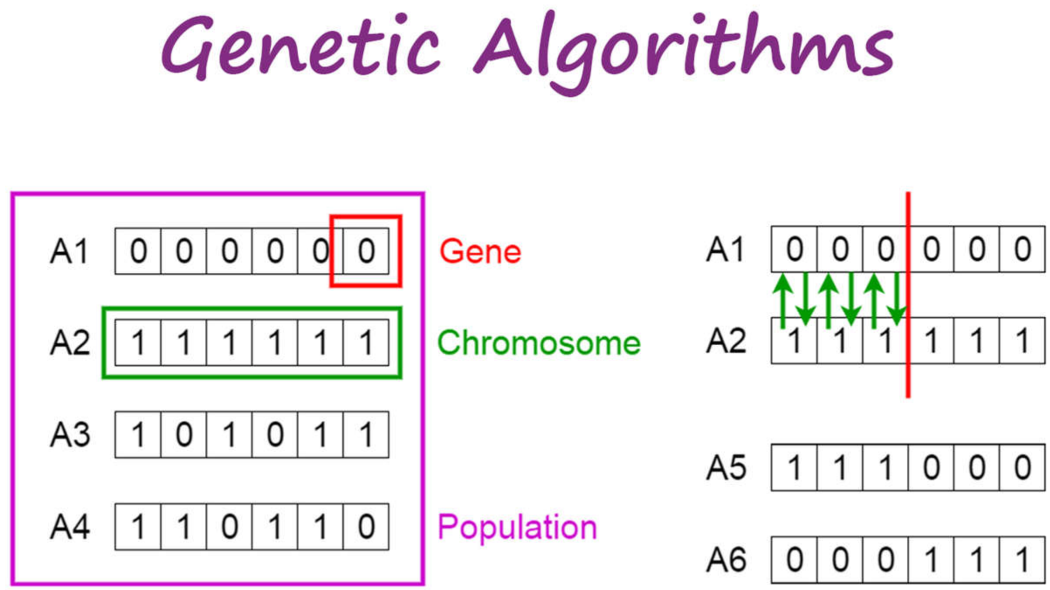
Genetic Algorithm

- ***Heredity*** – There is a process that must take place whereby the children inherit some characteristics from their parents.

- ***Variation*** – There must be diversity in the traits of the population or ways variation or diversity can be introduced into the population.

- ***Selection*** – There has to be a way in the population where parents pass down their genetic characteristics to their children and a way where some do not pass down these genetic characteristics.

Genetic Algorithm



The notion of Natural Selection

- The process of natural selection starts with the selection of fittest individuals from a population. They produce offspring which inherit the characteristics of the parents and will be added to the next generation. If parents have better fitness, their offspring will be better than parents and have a better chance at surviving. This process keeps on iterating and at the end, a generation with the fittest individuals will be found.
- This notion can be applied for a search problem. We consider a set of solutions for a problem and select the set of best ones out of them.

Phases of GA

- Five phases are considered in a genetic algorithm.
- Initial population
- Fitness function
- Selection
- Crossover
- Mutation

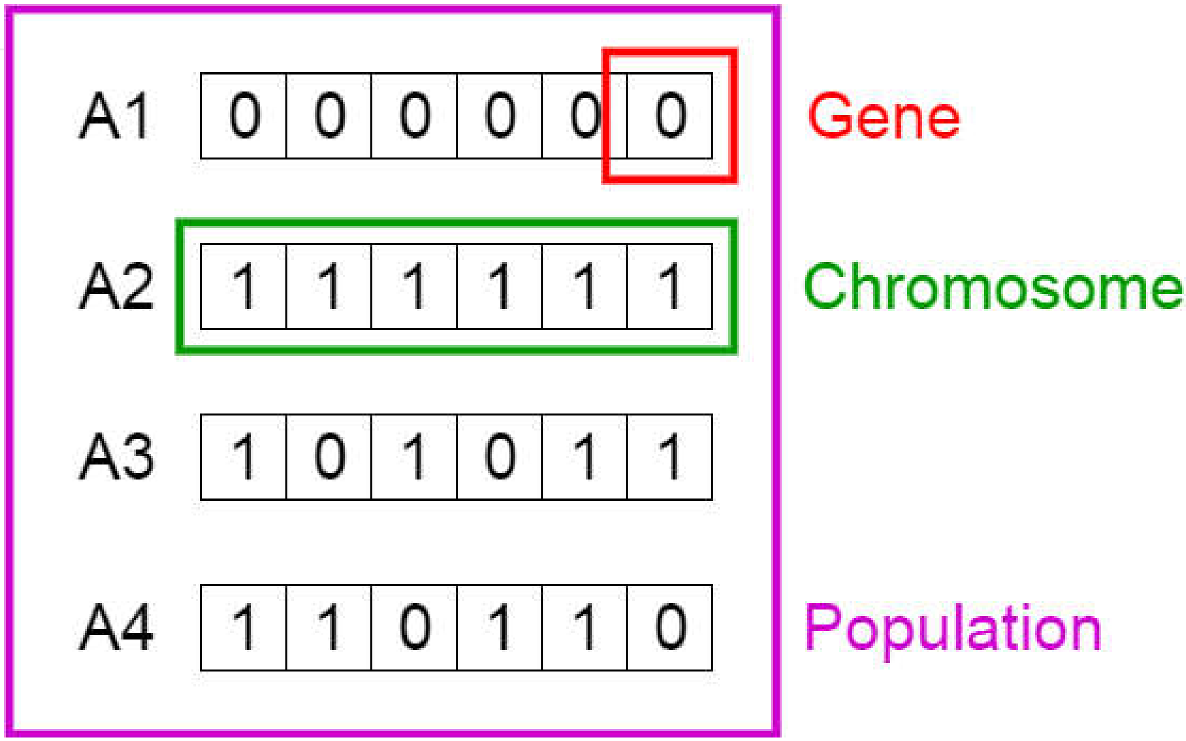
Initial Population

- The process begins with a set of individuals which is called a **Population**. Each individual is a solution to the problem you want to solve.

- An individual is characterized by a set of parameters (variables) known as **Genes**. Genes are joined into a string to form a **Chromosome** (solution).

- In a genetic algorithm, the set of genes of an individual is represented using a string, in terms of an alphabet. Usually, binary values are used (a string of 1s and 0s). We say that we encode the

Initial Population

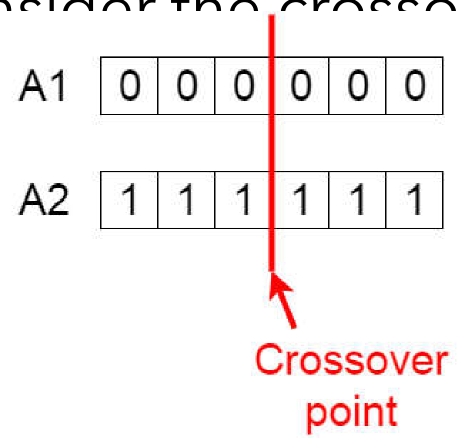


Fitness Function and selection

- The **fitness function** determines how fit an individual is (the ability of an individual to compete with other individuals). It gives a **fitness score** to each individual. The probability that an individual will be selected for reproduction is based on its fitness score.
- The idea of the selection phase is to select the fittest individuals and let them pass their genes to the next generation.
- Two pairs of individuals (parents) are selected based on their fitness scores. Individuals with high fitness have more chance to be selected for reproduction.

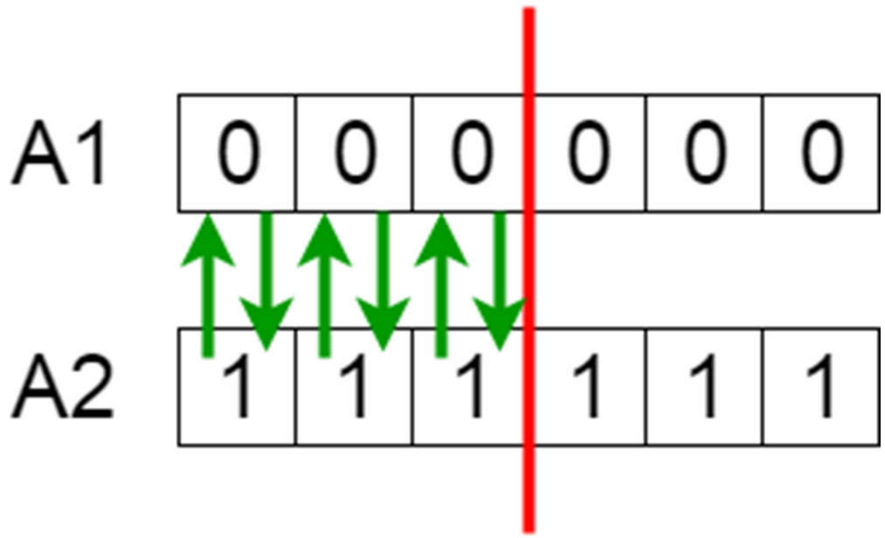
Crossover

- **Crossover** is the most significant phase in a genetic algorithm. For each pair of parents to be mated, a **crossover point** is chosen at random from within the genes.
- For example, consider the crossover point to be 3 as shown below.



Crossover

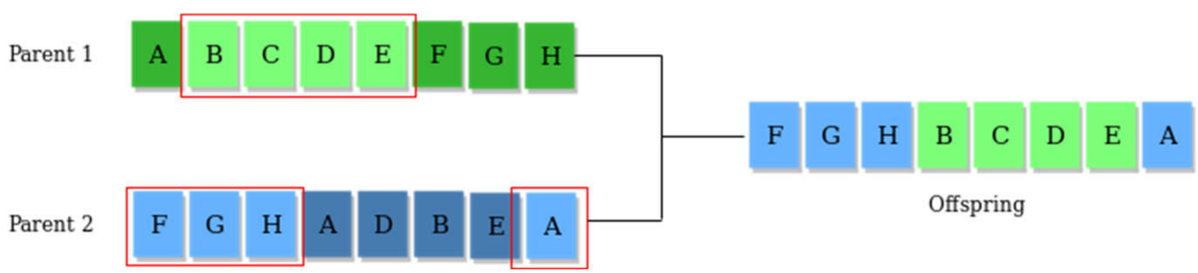
- **Offspring** are created by exchanging the genes of parents among themselves until the crossover point is reached.



Crossover

The new offspring are added to the population.

A5	1	1	1	0	0	0
A6	0	0	0	1	1	1



Mutation

- In certain new offspring formed, some of their genes can be subjected to a **mutation** with a low random probability. This implies that some of the bits in the bit string can be flipped.

Before Mutation

A5	1	1	1	0	0	0

After Mutation

A5	1	1	0	1	1	0

Mutation

Mutation occurs to maintain diversity within the population and prevent premature convergence.

Before Mutation

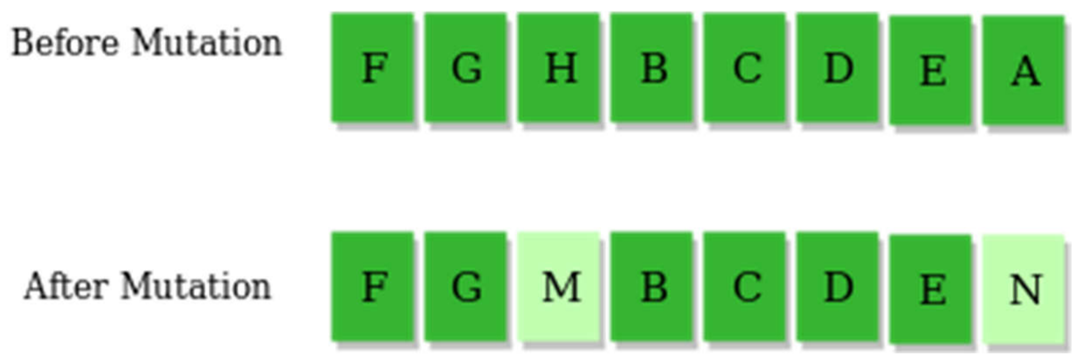
A5	1	1	1	0	0	0

After Mutation

A5	1	1	0	1	1	0

Mutation operator

- The key idea is to insert random genes in offspring to maintain the diversity in population to avoid the premature convergence. For example –

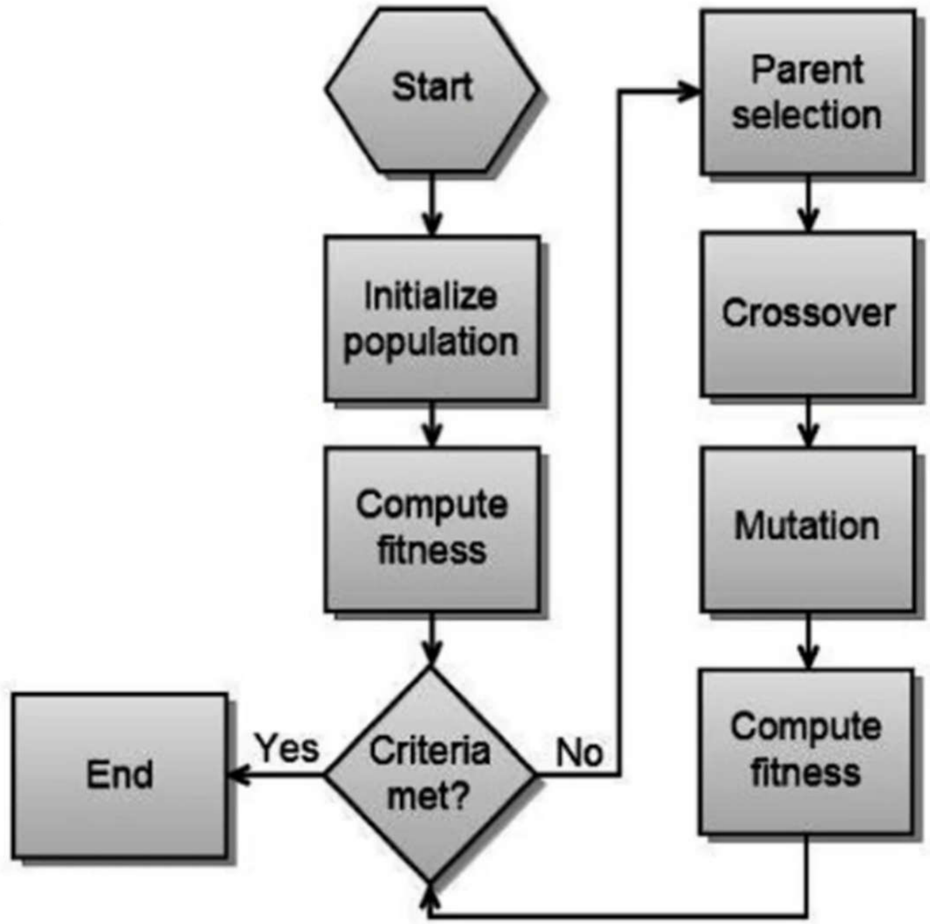


Termination

- The algorithm terminates if the population has converged (does not produce offspring which are significantly different from the previous generation). Then it is said that the genetic algorithm has provided a set of solutions to our problem.

START
    Generate the initial population
    Compute fitness
    REPEAT
    Selection
    Crossover
    Mutation
    Compute fitness
    UNTIL population has converged
STOP

Flow Diagram of GA



GA Solved Example

- Maximize the value of the function

$$f(x) = -x^2 + 2x$$

- over the range of real number from 0 to 2 with initial population 11010, 00111, 10110, 00101, with random number 0.4 ,0.15 ,0.7 ,0.9.
- select the crossover between the first and fifth digits?

GA Solved Example

- **Select Encoding Technique**
- The minimum value is 0 and maximum value is 2
- As a part of problem definition encoding technique is already given

  11010, 00111, 10110, 00101

GA Solved Example

Select Initial Population

- To start with, select initial population are random.
- Individual 1: 11010, = 26
- Individual 2: 00111, = 7
- Individual 3: 10110, = 22
- Individual 4: 00101 = 5

The binary number 11010 represents:

1 * 2^4 + 1 * 2^3 + 0 * 2^2 + 1 * 2^1 + 0 * 2^0

Calculating each term:

1 * 2^4 = 1 * 16 = 16

1 * 2^3 = 1 * 8 = 8

0 * 2^2 = 0 * 4 = 0

1 * 2^1 = 1 * 2 = 2

0 * 2^0 = 0 * 1 = 0

Adding these together: 16 + 8 + 0 + 2 + 0 = 26

Decode the value between 0 and 2

The formula for scaling a value between two ranges can be written as:

$$\text{scaled value} = \text{min\_range2} + \left(\frac{\text{value} - \text{min\_range1}}{\text{max\_range1} - \text{min\_range1}}\right) \times (\text{max\_range2} - \text{min\_range2})$$

In this case, you want to map the value 26 from a range of 0 to 31 to a range of 0 to 2:

$$\text{scaled value} = 0 + \left(\frac{26-0}{31-0}\right) \times (2 - 0)$$
$$\text{scaled value} = \frac{26}{31} \times 2 \approx 1.6774$$

GA Solved Example

Decode the individual into a real number

- Individual 1: 11010 $\rightarrow = 0 + \frac{(2-0)}{(2^5-1)} x(26) = \frac{52}{31} = 1.677$
- Individual 2: 00111 $\rightarrow = 0 + \frac{(2-0)}{(2^5-1)} x(7) = \frac{14}{31} = 0.451$
- Individual 3: 10110 $\rightarrow = 0 + \frac{(2-0)}{(2^5-1)} x(22) = \frac{44}{31} = 1.419$
- Individual 4: 00101 $\rightarrow = 0 + \frac{(2-0)}{(2^5-1)} x(5) = \frac{10}{31} = 0.322$

GA Solved Example

String No.	Initial Population	X Value	Fitness $f(x) = -x^2 + 2x$	Prob	Cumulative	Intervals of R N
1	11010	1.677	0.541			
2	00111	0.451	0.699			
3	10110	1.419	0.824			
4	00101	0.322	0.541			
Sum						
Average						
Maximum						

GA Solved Example

String No.	Initial Population	X Value	Fitness $f(x) = -x^2 + 2x$	Prob	Cumulative	Intervals of R N
1	11010	1.677	0.541			
2	00111	0.451	0.699			
3	10110	1.419	0.824			
4	00101	0.322	0.541			
Sum			2.6056			
Average			0.6514			
Maximum			0.824			

$$Prob = \frac{f(x)}{\sum f(x)}$$

GA Solved Example

String No.	Initial Population	X Value	Fitness $f(x) = -x^2 + 2x$	Prob	Cumulative	Intervals of R N
1	11010	1.677	0.541	0.21		
2	00111	0.451	0.699	0.27		
3	10110	1.419	0.824	0.32		
4	00101	0.322	0.541	0.2		
Sum			2.6056			
Average			0.6514			
Maximum			0.824			

GA Solved Example- Crossover step

String No.	Initial Population	X Value	Fitness $f(x) = -x^2 + 2x$	Prob	Cumulative	Intervals of R N
1	11010	1.677	0.541	0.21	0.21	
2	00111	0.451	0.699	0.27	0.48	
3	10110	1.419	0.824	0.32	0.8	
4	00101	0.322	0.541	0.2	1	
Sum			2.6056			
Average			0.6514			
Maximum			0.824			

GA Solved Example

String No.	Initial Population	X Value	Fitness $f(x) = -x^2 + 2x$	Prob	Cumulative	Intervals of R N
1	11010	1.677	0.541	0.21	0.21	0 to 0.21
2	00111	0.451	0.699	0.27	0.48	0.22 to 0.48
3	10110	1.419	0.824	0.32	0.8	0.49 to 0.8
4	00101	0.322	0.541	0.2	1	0.81 to 1

Random Number	Region	Chosen string
0.4	0.22 to 0.48	00111 ✓
0.15	0 to 0.21	11010
0.7	0.49 to 0.8	10110
0.9	0.81 to 1	00101

GA Solved Example- Crossover step

Crossover:

- Within each pair swap parts of the members solutions to create offspring which are a mixture of the parents:

Crossover Points
first and fifth digits

- For the first pair of strings: **00111 , 11010**
- Crossing these two strings at that point yields:
- 00111 → 0|011|1 → 01011
- 11010 → 1|101|0 → 10110
- For the second pair of strings: **10110, 00101**
- Crossing these two strings at that point yields:
- 10110 → 1|011|0 → 10100
- 00101 → 0|010|1 → 00111

GA Solved Example

String No.	New Population	X Value	Fitness $f(x) = -x^2 + 2x$	Prob	Cumulative	Intervals of R N
1	01011					
2	10110					
3	10100					
4	00111					
(sum)						
Average						
Max						

GA Solved Example

- Individual 1: 01011 $\rightarrow = 0 + \frac{(2-0)}{(2^5-1)} x(11) = \frac{22}{31} = 0.709$
- Individual 2: 10110 $\rightarrow = 0 + \frac{(2-0)}{(2^5-1)} x(22) = \frac{44}{31} = 1.419$
- Individual 3: 10100 $\rightarrow = 0 + \frac{(2-0)}{(2^5-1)} x(20) = \frac{40}{31} = 1.29$
- Individual 4: 00111 $\rightarrow = 0 + \frac{(2-0)}{(2^5-1)} x(7) = \frac{14}{31} = 0.451$

GA Solved Example

String No.	New Population	X Value	Fitness $f(x) = -x^2 + 2x$	Prob	Cumulative	Intervals of R N
1	01011	0.709	0.915			
2	10110	1.419	0.824			
3	10100	1.29	0.915			
4	00111	0.451	0.699			
(sum)			3.3548			
Average			0.8387			
Max			0.915			

Previous Maximum value was **0.824**, after first generation we got **0.915**