Genetic Algorithm

Genetic Algorithm (GA)

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.

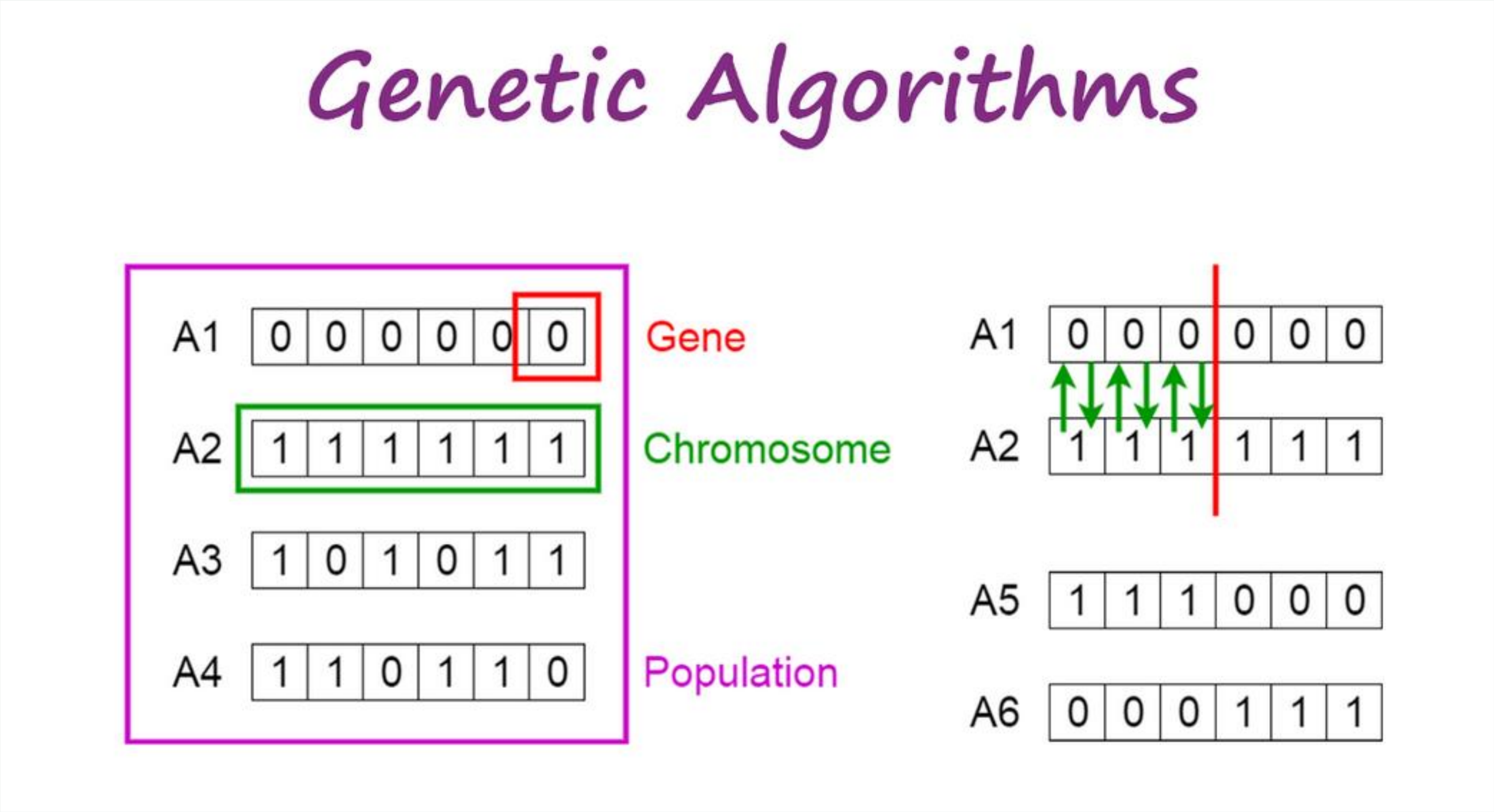
Genetic algorithms are commonly used to generate high-quality solutions to optimization and search problems by relying on biologically inspired operators such as mutation, crossover and selection. Some examples of GA applications include optimizing decision trees for better performance, solving sudoku puzzles, hyperparameter optimization, etc.

Elements

population

candidate solutions (called individuals, creatures, organisms, or phenotypes)

chromosomes or genotype


Genetic Algorithms
A1 0 0 0 0 0 0
Gene
A2 1 1 1 1 1 1
Chromosome
A3 1 0 1 0 1 1
A4 1 1 0 1 1 0
Population
A1 0 0 0 0 0 0
A2 1 1 1 1 1 1
A5 1 1 1 0 0 0
A6 0 0 0 1 1 1

Population

Population typically refers to the number of people in a single area, whether it be a city or town, region, country, continent, or the world.

In genetics, a population is often defined as a set of organisms. All members belong to the same species.

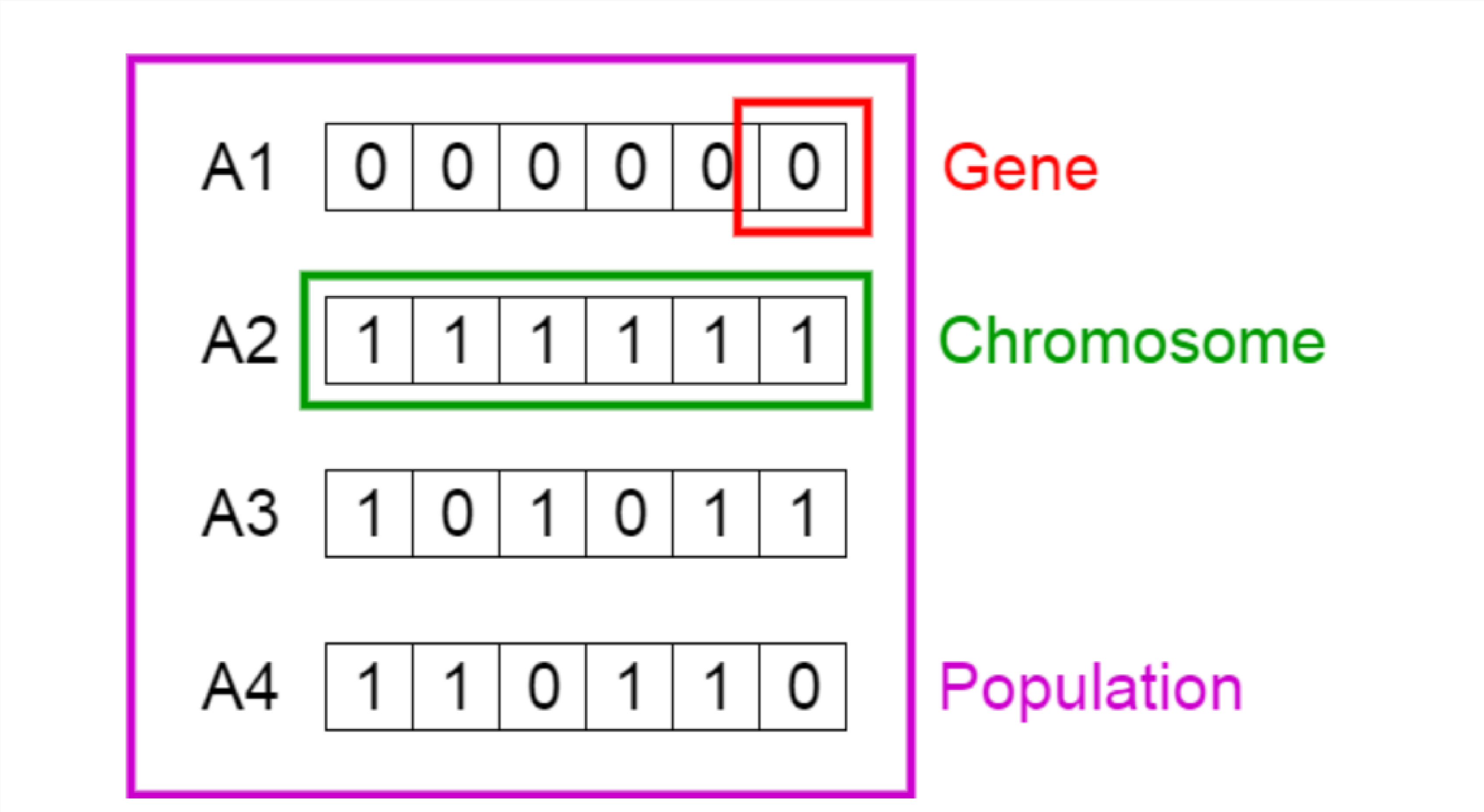
Initial Population

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 (STRING OF 1S AND 0S). WE SAY THAT WE ENCODE THE GENES IN A CHROMOSOME.

[redacted]


A1
0 0 0 0 0 0
A2
1 1 1 1 1 1
A3
1 0 1 0 1 1
A4
1 1 0 1 1 0
Gene
Chromosome
Population

Genetic Algorithm

Generation of Initial Population (random).

Evaluation (Generation of New Population).

Selection.

Mutation and Crossover.

Have we achieved the goal?

- If YES - end of algorithm
- If NOT – go to 2

Initialization

Initialization, i.e. the creation of an initial population, consists in the random selection of a certain number of chromosomes (population elements) represented by binary (or numerical) strings of a fixed length.

Generation of New Population

Application of genetic operators to create a new population

Fitness Function

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.

Selection

The idea of **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.

Selection

There are many selection methods. The most popular is the so-called the roulette method, which is named after the analogy of drawing with a roulette wheel. Each chromosome can be assigned a part of the roulette wheel with a size proportional to the value of the fitness function for that chromosome. Thus, the greater the value of the fitness function F, the larger part on the roulette wheel.

The whole roulette wheel corresponds to the sum of the values of the fitness function of all chromosomes of the population under consideration.

Each chromosome denoted by ch_i for i = 1, 2, ..., N, where N is the size of the population, corresponds to part of the roulette wheel v (ch_i), expressed as a percentage of whole wheel according:

$$v(ch_i) = p_s ch_i \cdot 100\%,$$

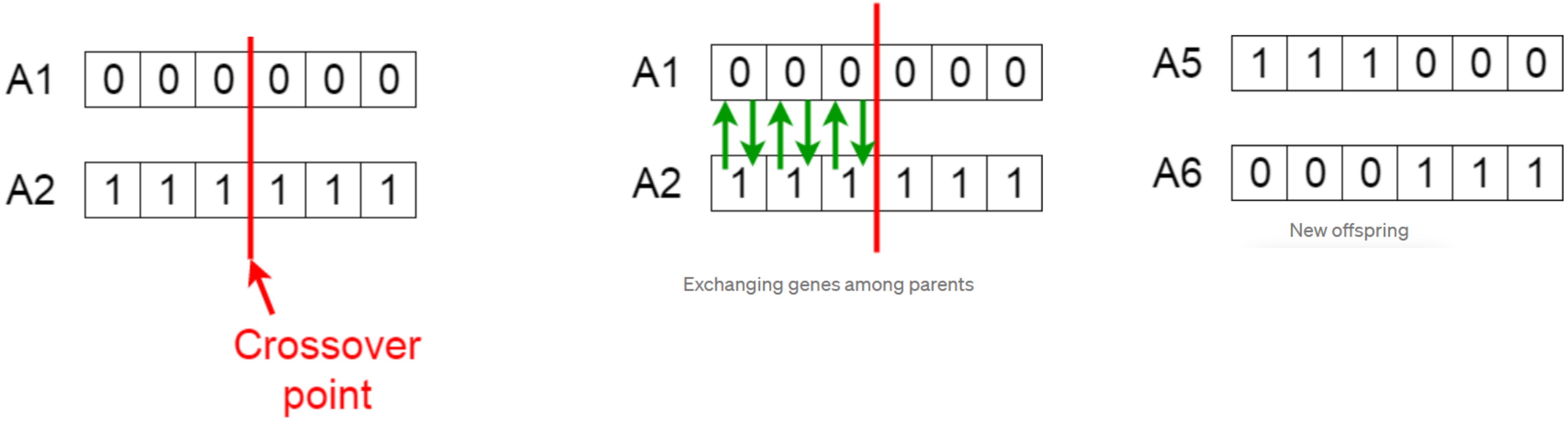
$$p_s(ch_i) = \frac{F(ch_i)}{\sum_{i=1}^{N} F(ch_i)},$$

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


A1 0 0 0 0 0 0
A2 1 1 1 1 1 1
Crossover point
A1 0 0 0 0 0 0
A2 1 1 1 1 1 1
Exchanging genes among parents
A5 1 1 1 0 0 0
A6 0 0 0 1 1 1
New offspring

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.

Mutation

Before Mutation

A5	1	1	1	0	0	0

After Mutation

A5	1	1	0	1	1	0

Example 1.

Consider the functions f (x) = 2 x ^ 2 + 1 and assume that x takes integer values between 0 and 15. The task of optimizing this function consists in searching a space composed of 16 points and finding such among the values 0, 1, ..., 15, for which the function takes the maximum (minimum) value. In this case, the task parameter is x. The set {0, 1, ..., 15} is the search space. It is also a set of potential solutions to the task.

Each of the 16 numbers belonging to this set is called a search space point, solution, parameter value, phenotype.

It is worth noting that the solution optimizing functions is called the best or optimal solution. Successive values of the parameter x from 0 to 15 can be encoded as follows:

0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111.

Example 2.

Consider a simplified, somewhat artificial example of finding the chromosome with as many ones as possible.

Suppose the chromosomes are made up of 12 genes and the population is 8 chromosomes.

It is known that the best chromosome will be composed of 12 ones.

Initial population

ch1 = [111001100101]

ch2 = [001100111010]

ch3 = [011101110011]

ch4 = [001000101000]

ch5 = [010001100100]

ch6 = [010011000101]

ch7 = [101011011011]

ch8 = [000010111100]

Fitness function

$F(ch1) = 7$

$F(ch2) = 6$

$F(ch3) = 8$

$F(ch4) = 3$

$F(ch5) = 4$

$F(ch6) = 5$

$F(ch7) = 8$

$F(ch8) = 5$

Selection

For each of the 8 chromosomes in the current population, we get the percentage of the roulette wheel:

Selection

$v(ch1) = 15,22$

$v(ch2) = 13,04$

$v(ch3) = 17,39$

$v(ch4) = 6,52$

$v(ch5) = 8,70$

$v(ch6) = 10,87$

$v(ch7) = 17,39$

$v(ch8) = 10,87$

Selection

because the sum of the values of the function F for 8 chromosomes is 46 and for ch1 we have

$$v(ch1) = 7/46 * 100 = 15.22\%$$

Selection

Random generation with a roulette wheel is as easy as randomly selecting a number from [0, 100] as a given chromosome.

Let's assume 8 numbers are randomly generated:

79, 44, 9, 74, 44, 85, 48, 23.

This means selecting the following chromosomes:

ch7, ch3, ch1, ch7, ch3, ch7, ch4, ch2

Example 3

The problem of magic squares

The problem of magic squares is to generate a square n × n, in which the elements are the numbers 1, 2, ..., n ^ 2 arranged in such a way that the sum of the values of the numbers in each column, row and diagonal is the same

Magic square 3x3

$$\begin{matrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{matrix}$$