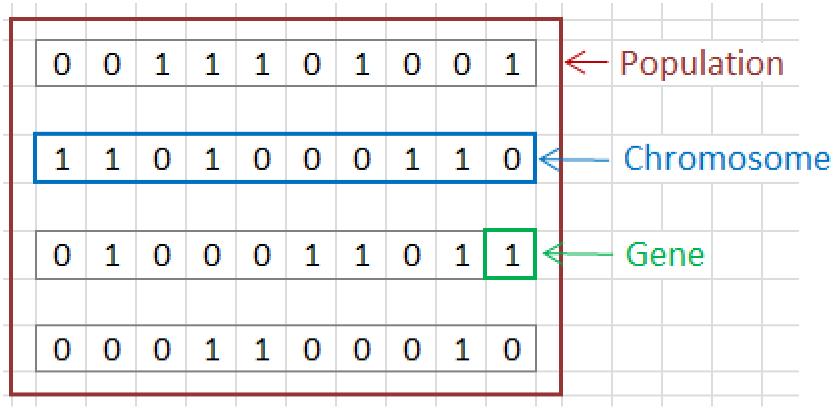
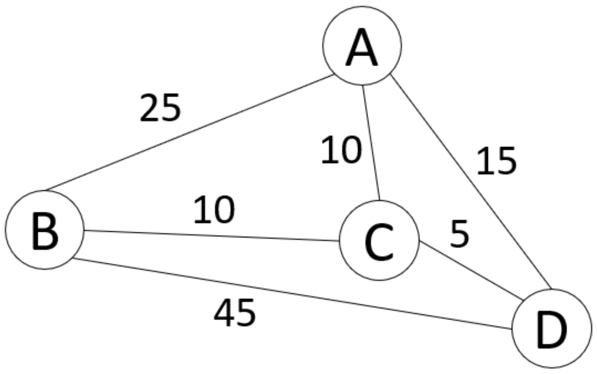
Genetic Algorithms (GA)

John Holland proposed genetic Algorithms in the 1970s. Initially, they were called “Reproductive Plans.” These algorithms are maybe the most famous of the evolutive algorithms family.

The inspiration comes from the DNA structure, which is people's genetic code. All the information is stored in chromosomes that have a lot of genes. Holland’s proposal consists of representing the solutions by binary arrays.



Example 1: Traveling salesman problem

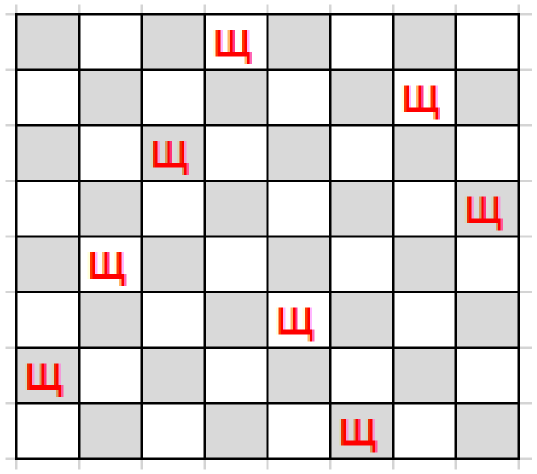


The problem consists of determining the path a salesperson must follow to minimize the distance. He/she starts and ends in the same node, and nodes cannot be repeated.

Example:

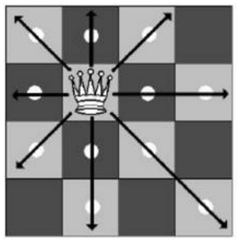
A – B – D – C – A = 25 + 45 + 5 + 10 = 85

Example 2: Traveling salesman problem



The problem consists of placing 8 queens on an 8x8 chessboard. The queens must not attack among them.

A queen attacks all the pieces that are in the same row, column, or diagonal.


wikipedia.org

Individuals' representation

The individuals' representation can be divided into genotype and phenotype:

- **Genotype**: it is the codified version of the solution
- **Phenotype**: it is the solution that represents an individual

We can use some of the following representations:

- **Binary representation**: It the Holland's original proposal. It consists of representing the solution using a binary array. It could be beneficial because a lot of problems can be represented using this technique.

- **Integer representation**: The representation of individuals as integer arrays is the best option for specific problems. For example: for evolving the trajectory of an agent using numbers for directions (left, right, up, and down).

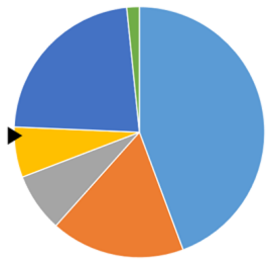
- **Real representation**: Many problems can be represented as real arrays, it is, $[x_1, x_2, \ldots, x_n]$ where $x_i \in \mathbb{R}$.

- **Permutation representation**: Some problems, such as sudoku or traveling agent, can be represented as a permutation of a set.

Selection of parents

There are some techniques for parents' selection:

Roulette selection, also called proportional selection, was the original proposal of Holland. The idea is simple. You can imagine a roulette where each section is assigned to an individual. If we have 10 individuals, the roulette is divided into 10 sections. The section size is proportioned to the individual's fitness.



The roulette's implementation can be as follows. Given an individual i, f_i represents its fitness, and, p_i is the individual's proportion, where it can be calculated as $p_i = \frac{f_i}{\sum_k f_k}$. The sum of all proportions must be 1. The segment between 0 and 1 is divided using those proportions for calculating ranges for each individual. Finally, to select an individual, a number is randomly calculated between 0 and 1, and the individual whose range corresponds to that number is selected as a parent.

Individual	**Fitness f_i**	**Proportion p_i**	**Range**
A	2	0.04	[0.00 – 0.04]
B	10	0.20	[0.04 – 0.24]
C	7	0.14	[0.24 – 0.38]
D	1	0.02	[0.38 – 0.40]
E	30	0.60	[0.40 – 1.00]

The roulette has some problems:

- If an individual has a fitness significantly bigger than others, the algorithm selects the same individual several times.
- If the individuals' fitness is similar among them, there is no pressure in the selection. In other words, the individuals will be selected as entirely random.

Tournament selection, may be the most known technique and easy to implement for parent selection. It consists of randomly choosing k individuals and selecting the fittest one. k represents the tournament size.

Advantages of the tournament:

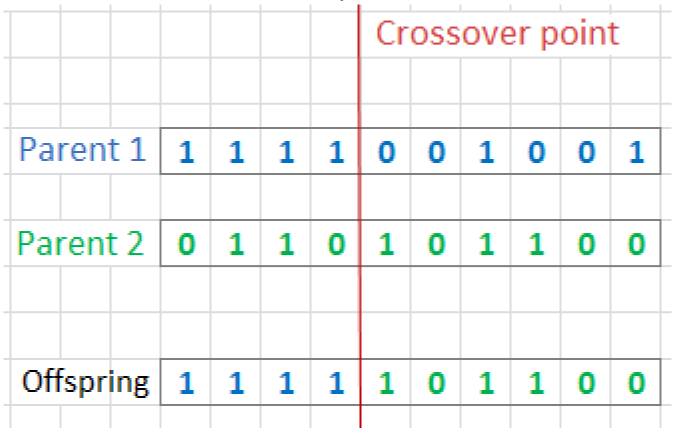
- It is too easy to implement.
- To change the selection's pressure, we can change the value of k. The bigger the value, the more the pressure.

Reproduction (crossover or recombination)

The goal of reproduction is to generate new individuals (offspring) combining the parents' properties. It can be implemented based on the individuals' representation.

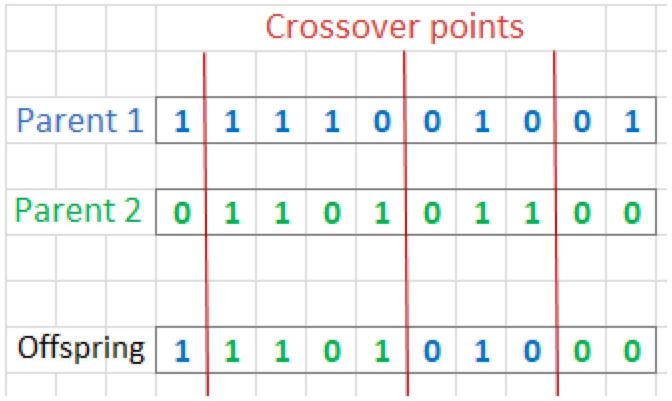
Binary and integer representation

- **1 point crossover**: This technique divides the parents into two sections randomly choosing a crossover point, represented in the color red in the following figure. The offspring is created using the first section of the first parent and the second section of the second parent.



Parent 1	1	1	1	1	0	0	1	0	0	1
Parent 2	0	1	1	0	1	0	1	1	0	0
Offspring	1	1	1	1	1	0	1	1	0	0

- **N point crossover**: It is a generalization of 1-point crossover. In this case, the parents are divided into several sections. The offspring are created using different sections of parents, as is shown in the following figure.



Parent 1	1	1	1	1	0	0	1	0	0	1
Parent 2	0	1	1	0	1	0	1	1	0	0
Offspring	1	1	1	0	1	0	1	0	0	0

- **Uniform crossover**: This technique chooses one by one the elements of the new individual. For each gene, it copies the gene of the first or the second parent randomly.

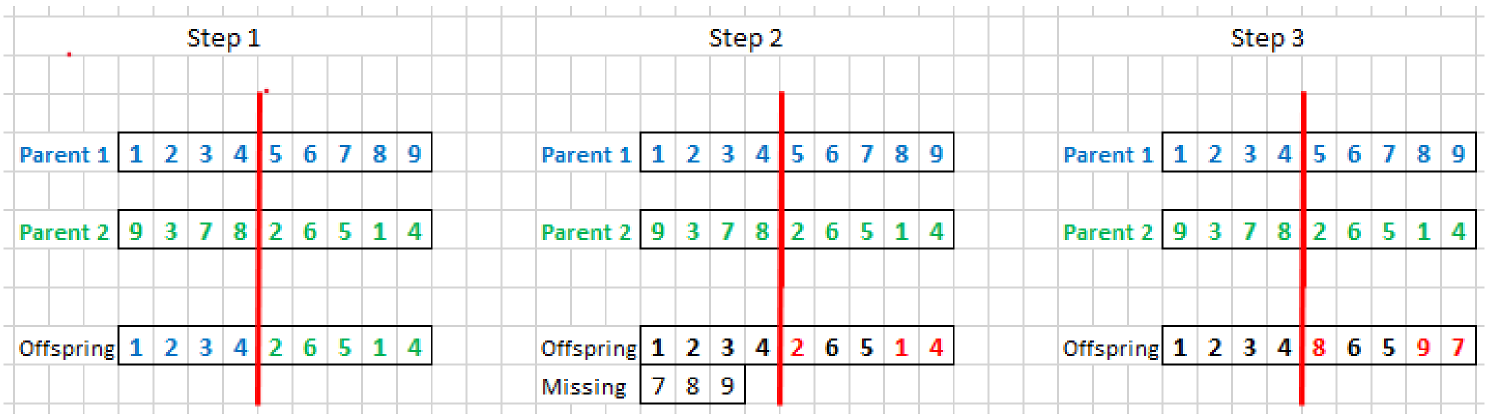
Real-valued representation

- **Discrete reproduction:** It is the same idea of a uniform crossover. For each element of the children, we randomly copy the value of parent 1 or parent 2.

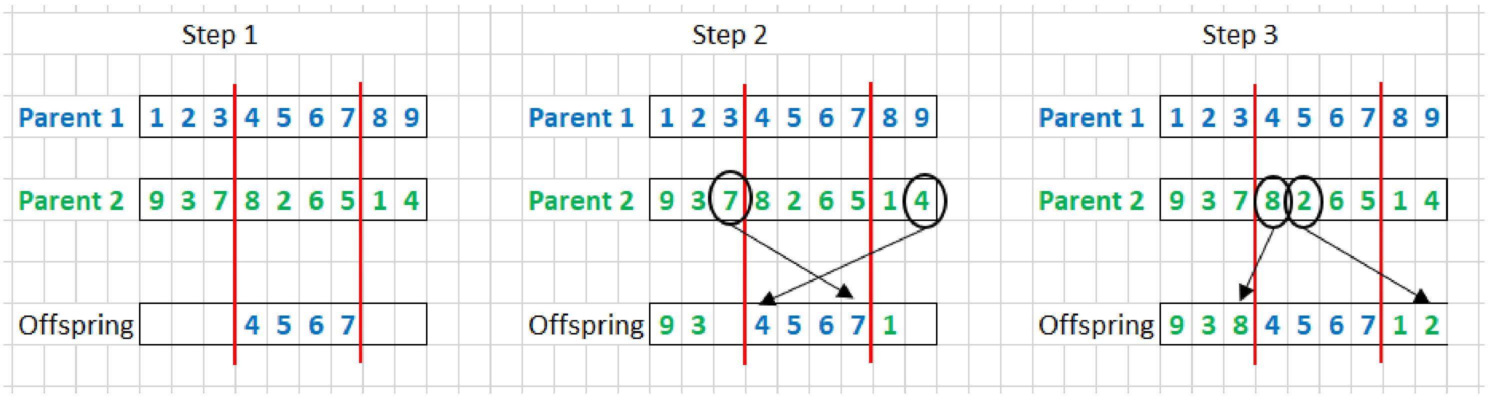
- **Asymmetric reproduction:** This technique consists of calculating the offspring o as the average of their parents p_1 and p_2. Formally, $o = \alpha p_1 + (1 - \alpha p_2)$, where α is a number between 0 and 1, generally, its value is 0.5.

Permutation representation

- **Simple permutation crossover:** the steps are described as follows:
  Step 1: Divide the individuals into two sections and copy the elements.
  Step 2: Calculate the repeated and the missing elements.
  Step 3: Randomly assign the missing values in the positions of the repeated elements.


Step 1
Parent 1 1 2 3 4 5 6 7 8 9
Parent 2 9 3 7 8 2 6 5 1 4
Offspring 1 2 3 4 2 6 5 1 4
Step 2
Parent 1 1 2 3 4 5 6 7 8 9
Parent 2 9 3 7 8 2 6 5 1 4
Offspring 1 2 3 4 2 6 5 1 4
Missing 7 8 9
Step 3
Parent 1 1 2 3 4 5 6 7 8 9
Parent 2 9 3 7 8 2 6 5 1 4
Offspring 1 2 3 4 8 6 5 9 7

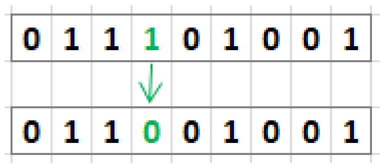
- **Partially mapped crossover:** the steps are described as follows:
  Step 1: Divide the individuals into 3 sections and copy the elements of the intermediate section of the first parent.
  Step 2: Copy the elements of the second parent (first and third sections) but ignore the elements that appear in the offspring.
  Step 3: Assign the missing values using the elements of the second part of the second parent that do not appear in the offspring.


Step 1
Parent 1 1 2 3 4 5 6 7 8 9
Parent 2 9 3 7 8 2 6 5 1 4
Offspring 4 5 6 7
Step 2
Parent 1 1 2 3 4 5 6 7 8 9
Parent 2 9 3 7 8 2 6 5 1 4
Offspring 9 3 4 5 6 7 1
Step 3
Parent 1 1 2 3 4 5 6 7 8 9
Parent 2 9 3 7 8 2 6 5 1 4
Offspring 9 3 8 4 5 6 7 1 2

Mutation

Mutation's goal is to modify individuals to explore the search space. Some of the techniques are the following:

- **Bitwise mutation** is used in binary representation and consists of randomly selecting one or several genes and changing their values.



- **Random resetting** is used in integer presentation and consists of randomly selecting one or several genes and resets its values.

- **Uniform mutation** is used in real-valued representation. It randomly selects one or several genes and chooses a random value between the minimum and maximum values.

- **Swap mutation** is used in permutation representation and consists of randomly selecting two elements and swapping their values.

Holland's original proposal of Genetic Algorithms was:

- Representation: binary
- Parents' selection: roulette
- Crossover: 1 point
- Mutation: bitwise
- Population model: generational

Implementation of the Genetic Algorithm phases

Crossover

```
Parameters:
    • Population of N parents
    • Pr: probability of reproduction
Return: new population of N individuals

Begin:
  For each individual in the new population:
        If a random number between 0 and 1 is less than Pr:
             Select two parents (roulette or tournament) and combine them
             Add the offspring to the new population
        Else
             Select one parent (roulette or tournament) and clone it
             Add the clone to the new population
        End if
  Return the new population
End
```

Mutation

```
Parameters:
    • Population of N parents
    • Pm: probability of mutation
Return: new population of N individuals

Begin:
  For each individual in the population:
        If a random number between 0 and 1 is less than Pm:
             Mutate the individual
             Add the mutated individual to the new population
        Else
             Clone the individual and put it in the new population
        End if
  Return the new population
End
```

Elite individual

```
If in the new population, we find an individual better than the elite:
    Replace the elite with the fittest individual
Else:
    Select an individual from the population with negative tournament
    Replace the selected individual with the elite
```

Genetic Algorithm

```
Parameters:
    N, population size
    G, Maximum number of generations
    Pr, Reproduction's probability
    Pm, Mutation's probability
Return: the elite individual

Begin
  Create the initial population
  Calculate the population fitness
  Get the elite
  While the number of generations is less than G or we haven't found a good solution
      Apply crossover selecting the parents
      Apply mutation
      Calculate the population fitness
      Get the elite or include the elite in the population
  End while
End
```

Recommended values:
N = 30, G = 100, Pr = 0.8, Pm = 0.3

Population models

We can distinguish between two evolution models:

- **Steady-state model**: In each iteration, a few number of individuals (mostly only one) are created using crossover and mutation, and those individuals replace other individuals from the population. For selecting the individual that will be replaced, it can be used a negative tournament, it is, randomly choosing several individuals and remove the one with worse fitness.

- **Generational model**: Mostly used in Genetic Algorithms. It randomly creates an initial population of N individuals. Each generation, the whole population is modified applying the following process. First, the parents are selected based on their fitness. Those parents are using for creating the offspring using the crossover and mutation algorithms. To warranty the algorithm convergence, we need to store the elite individual, the fittest individual. If in the new generation we found an individual better than the elite, we replace the elite. If not, we include the elite in the generation.

Evolution Strategies (ES)

Evolution Strategies were proposed by Rechenberg y Schwefel (Technical University of Berlin) in 1960's. The goal is to solve continuous multidimensional optimization problems. The main characteristic is the **self-adaptation** of parameters. It means that some evolutive algorithm parameters change during the execution. Those parameters are included in the individual representation and evolve at the same time that the solution.

Individuals' representation

Given the goal is solving continuous multidimensional optimization problems, the individuals' solutions are represented as vectors whose inputs are the values of the variables, the individual i's solution is represented as the vector $\vec{x_i} \in \mathbb{R}^d$, where d represents the number of features. One of the main characteristics of Evolution Strategies is self-adaptation of parameters, where each individual contains the mutation parameters σ_i, in addition to the values of the variables that are stored in the vector $\vec{x_i}$. The individual's representation is as follows:

$$< \vec{x_i}, \sigma_i >$$

Mutation

The individuals can be seen as points in a d –multidimensional space, where the mutation's goal is to move those points so that the position of the mutated individual is close to the position of the individual before mutation.

The individual's position $\vec{x_i}$ is modified by adding a random number, noise, to each entry. The noise follows a normal distribution zero-centered and with standard deviation σ_i. The value σ_i is known as *mutation step size*, where the bigger the value, the bigger the individual modification.

Rechenberg proposed the famous **1/5 success rule**, it states that the ratio of successful mutations (those in which the child is fitter than the parent) to all mutations should be 1/5. Hence if the ratio is greater than 1/5 the step size should be increased to make a wider search of the space, and if the ratio is less than 1/5 then it should be decreased to concentrate the search more around the current solution. The rule is executed at periodic intervals, for instance, after k iterations, each σ is reset by

$$\sigma = \begin{cases} \sigma/c & si\ p > 1/5 \quad Increase \\ \sigma * c & si\ p < 1/5 \quad Decrease \\ \sigma & si\ p = 1/5 \end{cases}$$

Where p is the relative frequency of successful mutations measured over a number of trials, and the parameter c is in the range $0.817 \leq c \leq 1$.

We describe two special cases of mutation in evolution strategies:

- **Uncorrelated Mutation with One Step Size**. In this case, the individual is represented by its position $\vec{x}$ and one scalar that represents the mutation step size σ.

$$\underbrace{< x_1, x_2, \dots, x_d,}_{\vec{x}} \ \underbrace{\sigma >}_{\sigma}$$

In the mutation, we add to all the vector entries a random number obtained from a normal distribution zero-centered with standard deviation equals to standard σ. The mutation step size must also be modified for self-adaptation. The mutation is performed as follows:

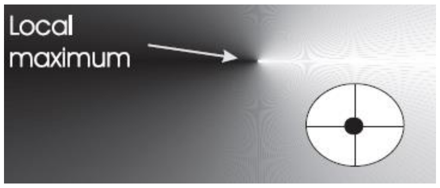
$$\sigma' = \sigma * e^{N(0,\tau)}$$
$$x_i' = x_i + N(0,\sigma)$$

the recommended value for τ is $\tau = 1/\sqrt{n}$. For avoiding step sizes equals to 0, it is recommended

$$If\ \sigma' < \epsilon \ \Rightarrow \sigma' = \epsilon$$

where ϵ is an small scalar, it is recommended $\epsilon = 1e - 3$.

The mutation can be seen as modification in a hypersphere, where the modifications near to the original position have more probability than the ones far from the original position. This mutation is recommended for problems whose features have the same values range.



En este caso, la mutación de un individuo se puede ver como una nueva posición en una hiperesfera cuyo centro es la posición actual del individuo, además de que posiciones cercanas al individuo son más probables (por la distribución Gaussiana).

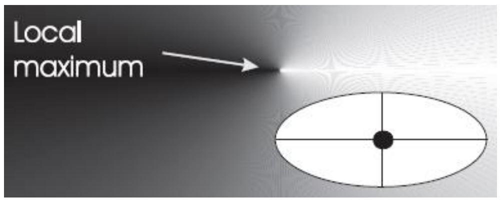
- **Uncorrelated Mutation with d Step Sizes**. The main idea is to treat each feature independently. It is useful where the features have different ranges values. In this case, the mutation parameter is a vector $\vec{\sigma} \in \mathbb{R}^d$. The individuals' representation is:

$$\underbrace{< x_1, x_2, \dots, x_d,}_{\vec{x}} \ \underbrace{\sigma_1, \sigma_2, \dots, \sigma_d >}_{\vec{\sigma}}$$

The mutation is performed as follows:

$$\sigma_i' = \sigma_i * e^{N(0,\tau_1) + N(0,\tau_2)}$$
$$x_i' = x_i + N(0,\sigma_i')$$

where $\tau_1 = 1/\sqrt{2n}$ y $\tau_2 = 1/\sqrt{2\sqrt{n}}$. It is important to avoid values of σ_i near to 0.



Recombination

The recombination consists of create one child from two parents. In Evolution Strategies there are two recombination variants. In *intermediate recombination* the values of the parents are averaged. Using *discrete recombination* one of the parent's values is randomly chosen with equal chance for either parents. The parents can be represented as the vectors $\overrightarrow{p_1}$ and $\overrightarrow{p_2}$, and the child is the resultant vector $\vec{c}$. Using this notation, the recombination techniques are defined as:

$$c_i = \begin{cases} (p_{1i} + p_{2i})/2 & intermediate\ recombination \\ random\ selection: p_{1i}\ or\ p_{2i} & discrete\ recombination \end{cases}$$

Selection of Parents

Parent selection in ES is completely random, it is because here the whole population is seen as parent.

Survivor Selection

After creating λ offspring and calculating their fitness, the best μ of them are chosen deterministically. There are two schemes of survivor selection:

- (μ, λ) selection, where only the best μ offspring are selected
- $(\mu + \lambda)$ selection, where the best μ individuals (from the union of parents and offspring) are selected

To sum up, the book Introduction to Evolutionary Algorithms (Eiben) presents the following summary of ES:

- Representation: real-valued vectors
- Recombination: discrete or intermediary
- Mutation: Gaussian perturbation
- Parent selection: Uniform random
- Survivor selection: (μ, λ) or $(\mu + \lambda)$
- Specialty: self-adaptation of mutation step sizes

Evolution Strategies Algorithm

```
Parameters:
    N, population size
    $\lambda$, offspring size
    G, Maximum number of generations

Return: the elite individual

Begin
  Create the initial population
  Calculate the population fitness
  Get the elite
  While the number of generations is less than G or we haven't found a good solution
      Recombination, generate $\lambda$ offspring
      Mutation of parents and offspring
      Calculate the population fitness
      Survivor selection (the best individuals)
      Get the elite or include the elite in the population
  End while
End
```