



Artificial Intelligence

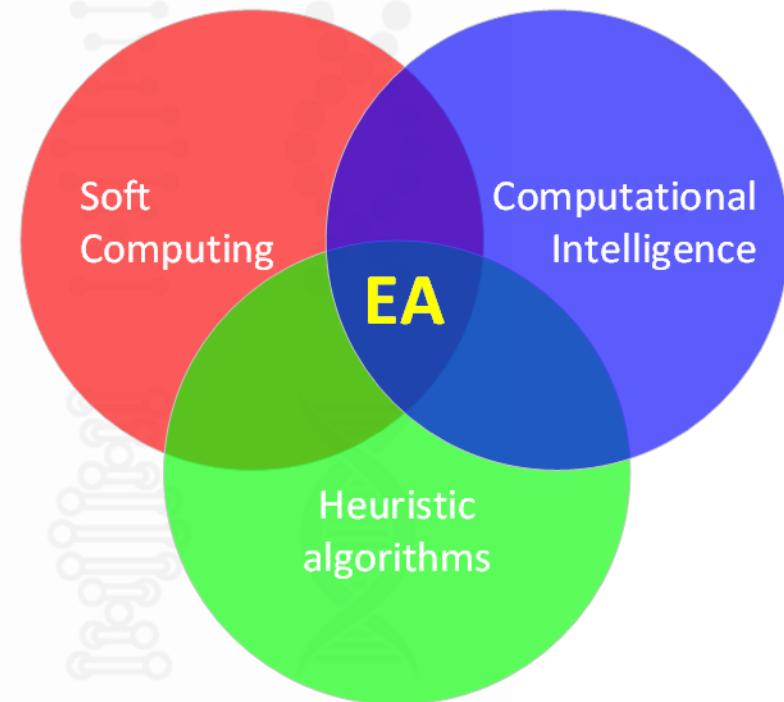
Topic 8: Evolutionary Computation

What are Evolutionary Algorithms (EAs)?

EAs are derivative-free stochastic approach based on biological evolutionary processes that mimics the mechanism of natural selection ("survival of the fittest") to solve optimization and learning problems.

Main characteristics:

- **population-based** - they maintain a group of solutions, called a population, to optimize or learn the problem in a parallel way,
- **fitness(cost)-oriented** - every solution in a population is called an individual. Every individual has its gene representation, called code, and performance evaluation, called its fitness (or cost) value,
- **variation-driven** - individuals undergo a number of variation operations to mimic genetic gene changes.



Historical background

- **Aristotle** (4th century BC) - grouping of living organisms.
Thesis: living organisms strive for achieving a state of higher perfection.
- **Leonardo da Vinci** (15th century) - fossils are the remains of species long extinct in the past.
- **Jean-Baptiste de Lamarck** (18th, 19th century) - the first scientifically-based theory of evolution.
Thesis: living organisms are endowed with a certain vital force, which leads them to ever-higher complexity.
- **James Hutton** (18th century) - theory of gradualism.
Thesis: the formation of the Earth arose as a result of the gradual culmination of slow processes happening over great amounts of time of that take place also today.
- **Charles Robert Darwin** (19th century) - the theory of evolution: *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.*
- **Alfred Russel Wallace** (19th, 20th century) - independently conceived the theory of evolution through natural selection.

- **Gregor Johann Mendel** (19th century) - the founder of the modern genetics.
Mendel's laws: law of segregation of genes, law of independent assortment, law of dominance.
- **Neo-Darwinism** (20th century) - the gene-centered view of evolution. Explains the variability described by Darwin theory among the offspring, using the concept of gene recombination
- **Godfrey Harold Hardy and Wilhelm Weinberg** (20th century) -
Hardy–Weinberg principle: allele and genotype frequencies in a population will remain constant from generation to generation in the absence of other evolutionary influences.
Conditions:
 - *a large populations,*
 - *no mutation,*
 - *isolation,*
 - *random crossing of pairs.*
- **Andrey Kolmogorov and Nikolai Vasilyevich Smirnov** (20th century) - *Kolmogorov-Smirnov theorem:* the smaller the random sample, the higher the probability of deviations from the actual distribution. We have a *genetic drift*.

Modern biology:

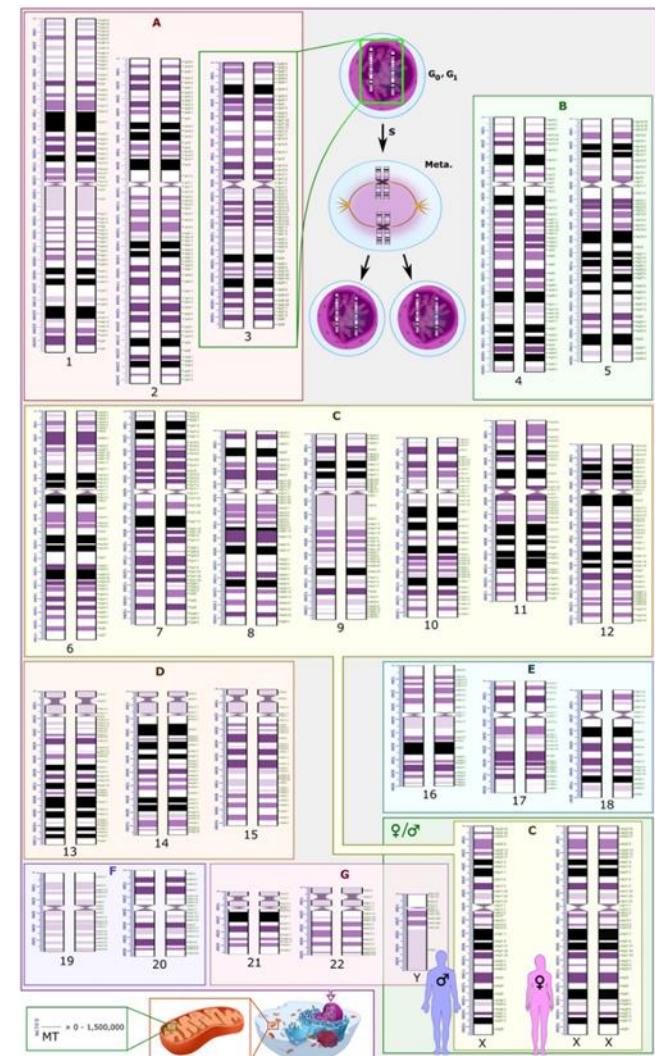
Chromosomes - carriers of genetic information.

Chromatin – a complex of protein, deoxyribonucleic (DNA) and ribonucleic (RNA) acids.

Karyotype - the number and appearance of chromosomes in the nucleus.

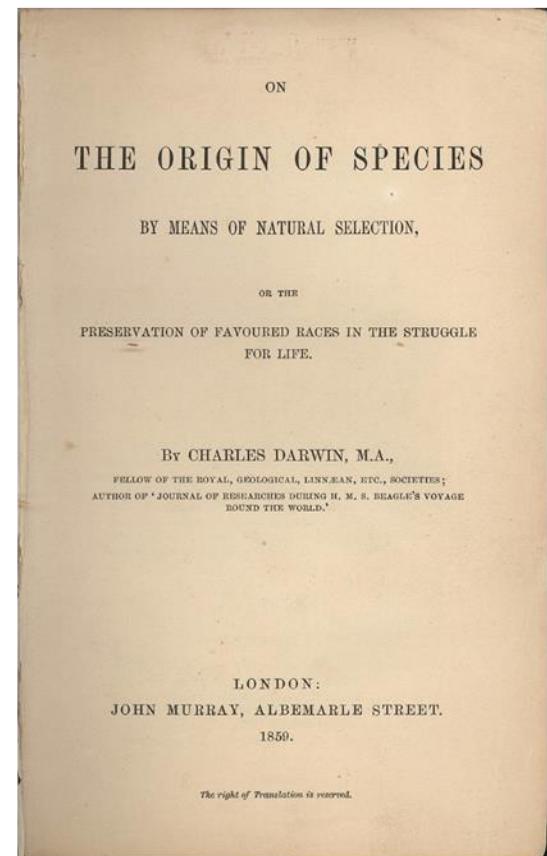
Basic concepts:

- Chromosomes store genetic information.
- Each chromosome is composed of DNA.
- Human chromosomes form pairs (23 pairs).
- Chromosomes are divided into parts (*genes*).
- Genes code for characteristics.
- Possible values of genes: *alleles*.
- Position of the gene in the chromosome: *locus*.



Darwin's theory (key concepts)

- In any population of organisms, due to differences in their genetic constitution, individuals have different characteristics, which can be inherited by offspring (**variability**).
- Organisms produce more offspring than they can survive (**overproduction**).
- Due to overproduction and limited resources, there is competition among organisms for survival (**struggle for existence, competition**).
- Individuals better adapted to their environment are more likely to survive and reproduce (**survival of the fittest**).
- Over generations, features that increase survival and reproduction become more common in the population.
Natural selection is the mechanism by which the environment “selects” these favorable characteristics (**natural selection**).
- Over time, populations evolve, becoming better adapted to their environment (**adaptation**).
- Over long periods of time, accumulated changes can lead to the formation of new species. When populations of a species become geographically isolated or are otherwise prevented from interbreeding, they can evolve independently, eventually becoming distinct species (**speciation**).



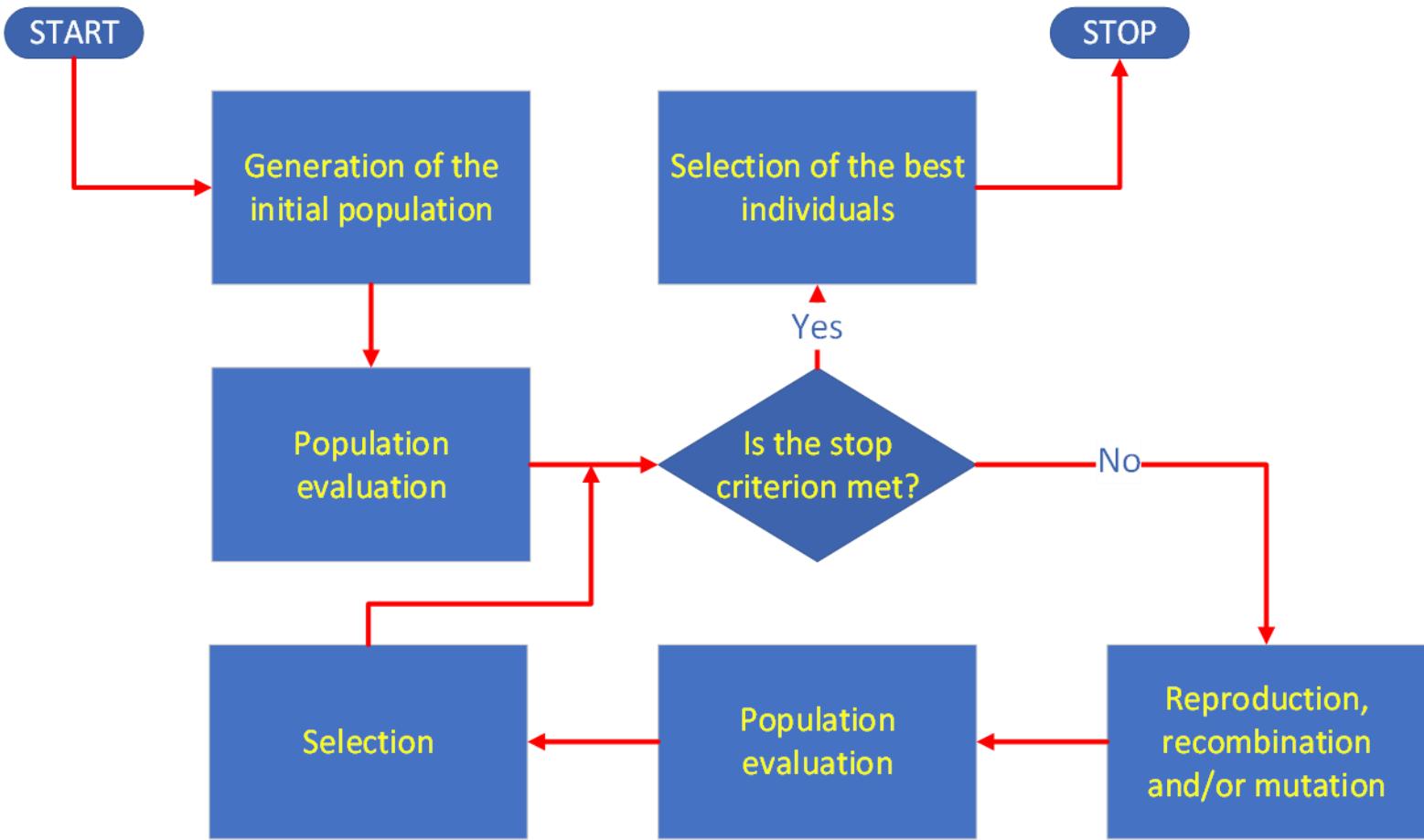
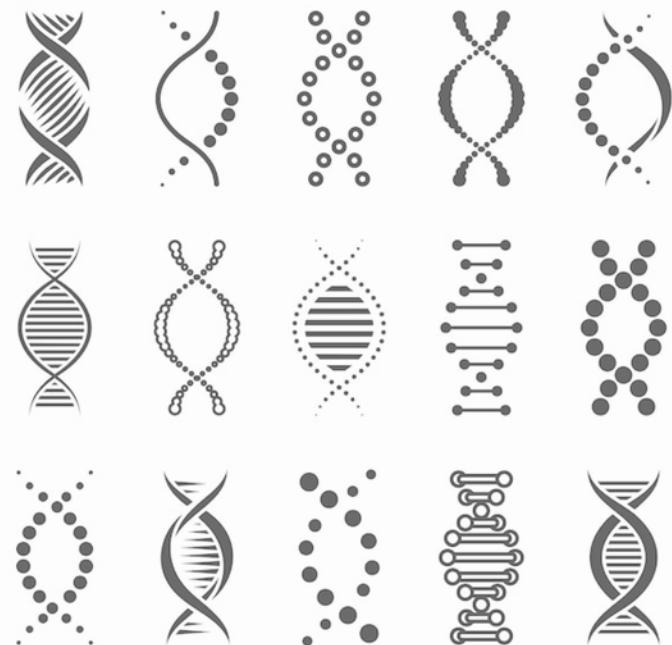


Fig. A typical structure of the EA algorithm.

Classical EAs include genetic algorithms (**GAs**), evolution strategy (**ES**), evolutionary programming (**EP**), and genetic programming (**GP**). They are **random-based solution space searching metaheuristic algorithms**.

Application areas

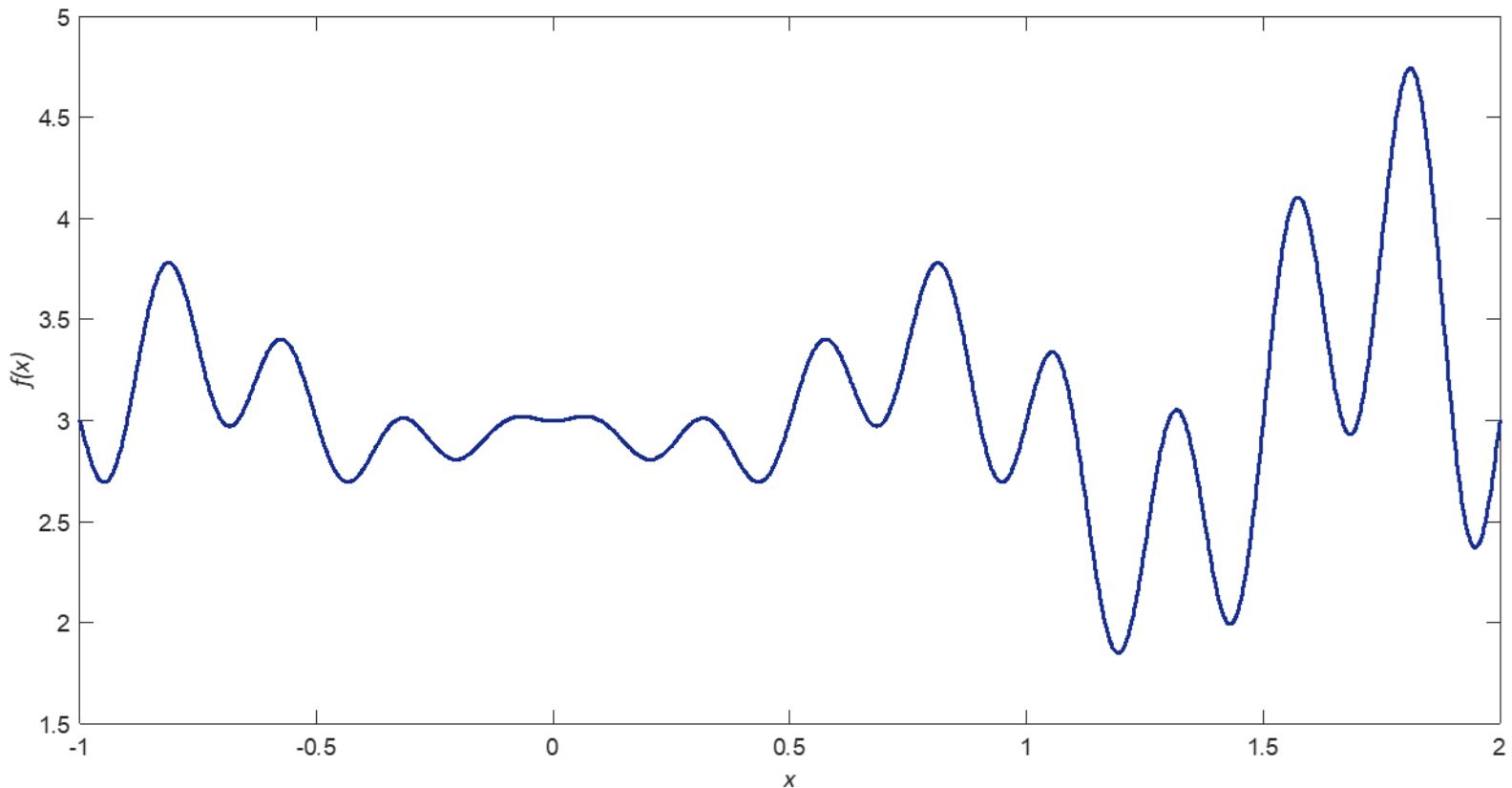
- function optimization,
 - multi-objective optimization,
 - combinatorial optimization,
 - engineering, structural optimization and design,
 - constraint satisfaction problems,
 - economics and finance,
 - biology,
 - data mining and data analysis,
 - mathematical problems,
 - electrical
- engineering and circuit design,
 - chemistry, chemical engineering,
 - scheduling,
 - robotics,
 - image processing,
 - networking and communication,
 - medicine,
 - resource minimization, environment surveillance and protection,
 - military and defense,
 - evolving behaviors, e.g., for agents or game players.



Simple Genetic Algorithm (SGA)

Let us consider a maximization problem:

$$\begin{aligned} & \underset{x}{\text{max}}, f(x) = x \cos(5\pi x) \sin(3\pi x) + 3, \\ & \text{s.t. } -1 \leq x \leq 2, \end{aligned}$$



In a typical SGA

- a **population** of P **individuals** is considered,
- an individual is characterized by its location (**chromosome** composed of **genes**) and its quality (**fitness value**),
- the quality determines the **selection** of **mating pool (parents)**, higher quality means higher probability of being selected,
- two parents may be selected randomly (without replacement) from the mating pool to generate two offspring (**crossover, recombination**),
- every offspring may undergo small changes to become a new individual (**mutation**),
- the newly generated population replaces the old one and another **generation** starts.

References to Darwinian natural selection theory:

- **selection** \leftrightarrow survival of the fittest,
- **crossover, recombination** \leftrightarrow mating,
- **mutation** \leftrightarrow small changes of the offspring.

Representation and Evaluation

Representation - the process of finding a way to represent a solution to the problem.

We use the binary code (binary representation) of a real number using ℓ binary codes. This makes the **representation error** equal to:

$$e = \frac{x_{\max} - x_{\min}}{2^\ell},$$

where x_{\min} and x_{\max} are the lower and upper bounds of the considered range of real numbers.

Assuming $\ell = 16$ we get a chromosome to be evolved:

0	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0		

Each part of a chromosome is called a gene.

A gene has two properties: its value (allele), and its location (locus).

So, the chromosome:

0	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1	1
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	

is the result of binary encoding.

The binary decoding:

$$x = \frac{\sum_{i=0}^{\ell-1} a_i 2^i}{2^\ell} (x_{\max} - x_{\min}) + x_{\min},$$

where a_i is the allele of the i -th locus.

In the considered example:

$$x = \frac{3}{2^{16}} \sum_{i=0}^{15} a_i 2^i - 1 = 0.30000.$$

encoding

0 1 1 0 1 1 1 0 1 1 1 1

genotype representation

chromosome
evaluation

$f(x) = 3.00000$

fitness value

(a function value in our example)

decoding

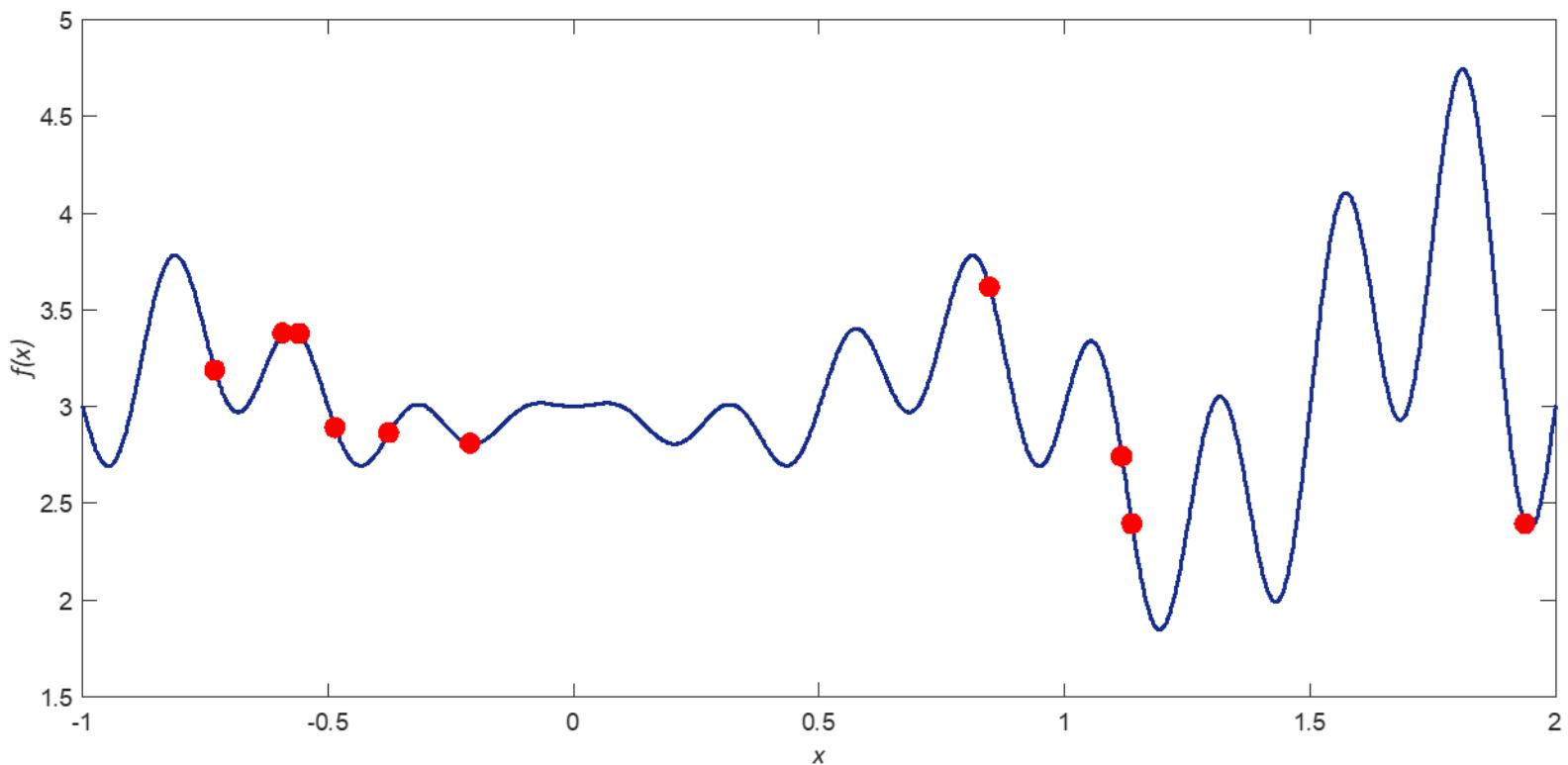
$x = 0.30000$

phenotype representation

Initialization

SGA starts with a randomly generated group of P individuals (**population**):

- with uniform distribution,
- evenly distributed individuals in the definition domain (grids),
- ...



Selection

SGA imitates natural selection by granting fitter individuals higher opportunity to breed.

Roulette wheel selection (RWS):

1. Calculate the **relative fitness value**:

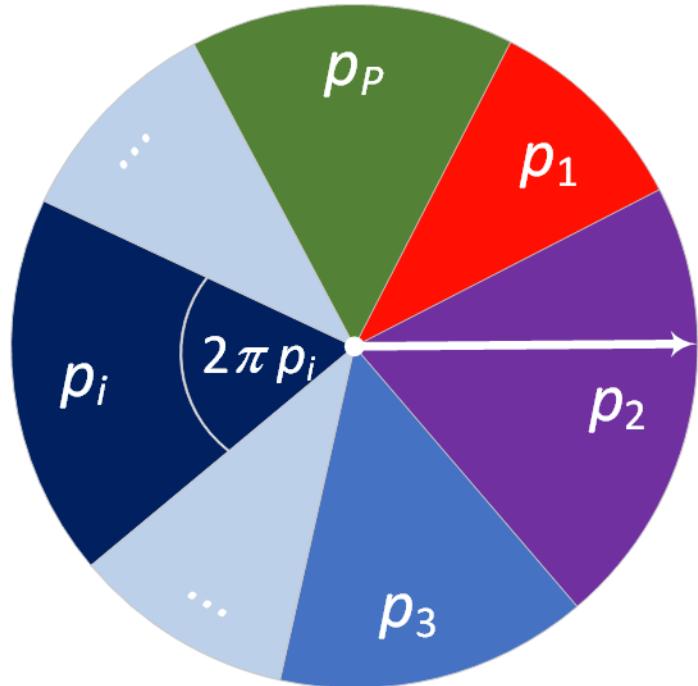
$$p_i = \frac{f_i}{\sum_{i=1}^P f_i}, \quad \sum_{i=1}^P p_i = 1,$$

which can be interpreted as **the probability of being selected as a candidate in the mating pool**.

2. Select P individuals by applying an accumulation process

$$\sum_{i=1}^k p_i < r < \sum_{i=1}^{k+1} p_i,$$

where $r \sim U(0,1)$ is a random number.



Some individuals in the population will be selected more than once and some will never be selected, even the most fitted individuals (**selection bias**).

Variation operators

Gene exchange between two or more individuals:
crossover (recombination).

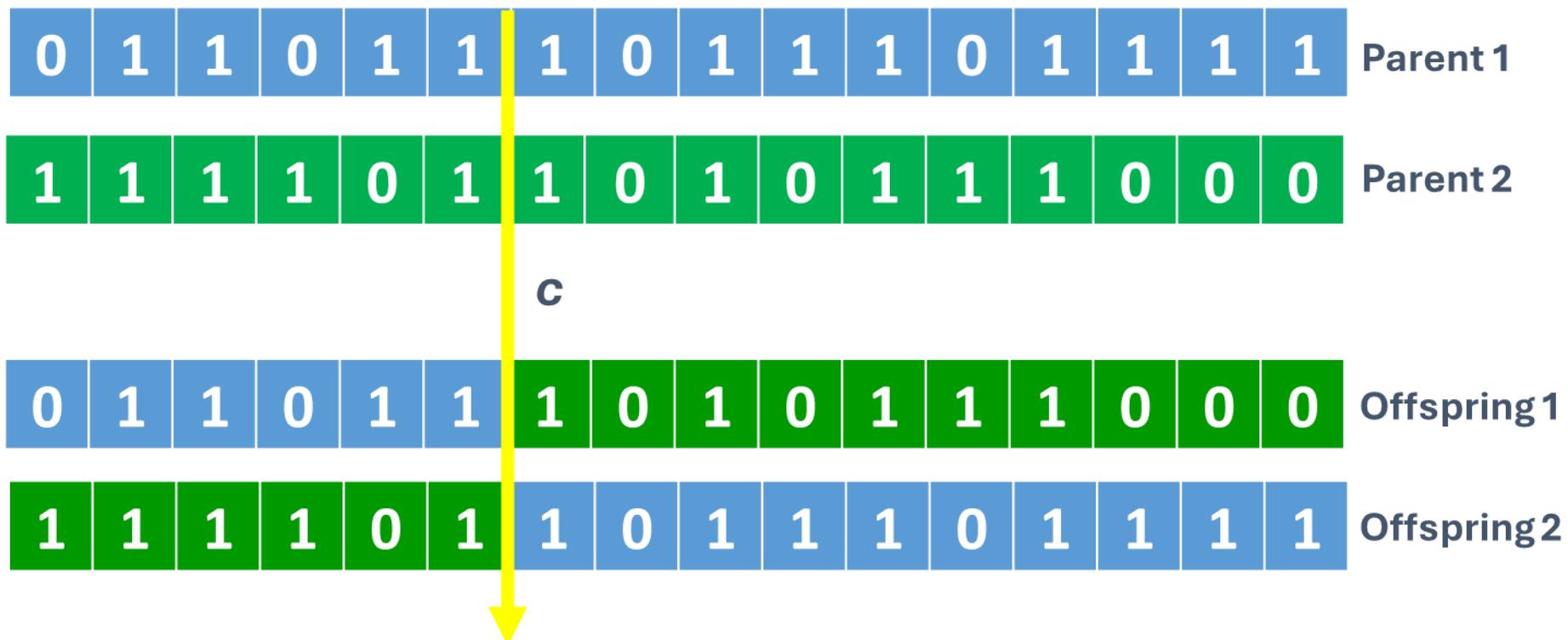
Genes of one individual changes on its own:
mutation.

Crossover

There are many ways to select two individuals from the mating pool (random shuffle, random permutation, random selection, ...).

Crossover rate - the probability of crossover p_c .

Single point crossover: if $r < p_c$ then generate random point $c \in [1, \ell - 1]$. The genes after c are exchanged between parents and the resulting chromosomes are offspring.



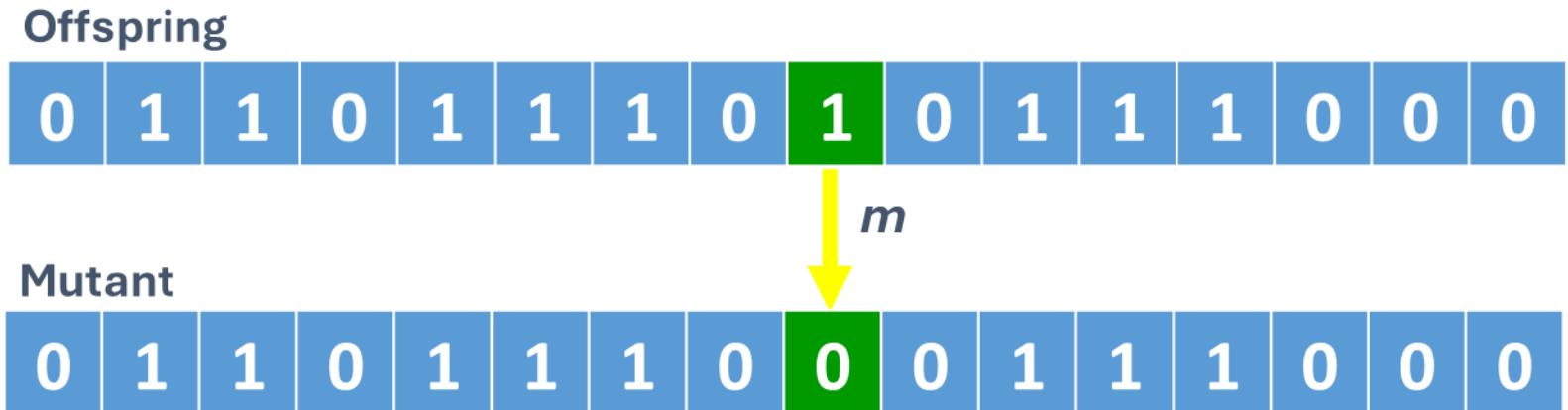
Mutation

There are many ways to mutate the resulting offspring.

Mutation rate - the probability of mutation p_m .

Bit-flip mutation: if $r < p_m$ then generate a random point $m \in [1, \ell]$.

Flip the gene m .



Replacement

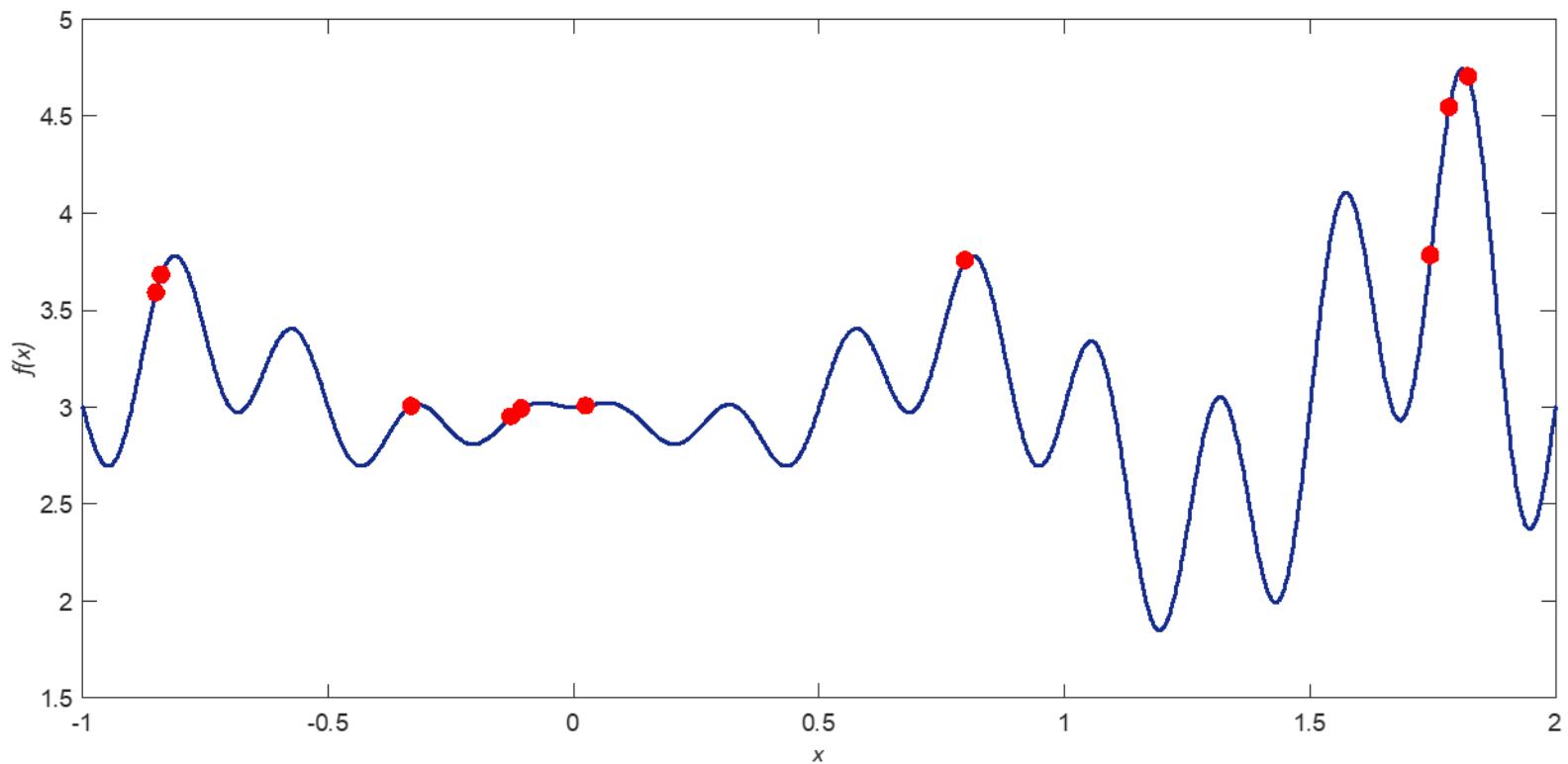
After variation operators (crossover, mutation) P individuals are generated, which replace the old population.

Stop criterion

Maximum number of generations N_{\max} .

Solution process of SGA:

1. Assign the parameters for SGA, such as p_c , p_m , P , N_{\max} , etc.
2. Generate P uniformly distributed individuals randomly to form the initial population. Evaluate their fitness values.
3. Select P individuals from current population using RWS to generate the mating pool.
4. Select two individuals from the mating pool randomly without replacement to perform single-point crossover with probability p_c , and perform bit-flip mutation for every gene of the offspring with probability p_m .
5. Evaluate the fitness value of individuals in the new population.
6. Replace the current population with the new population.
7. Increase the generation index $t = t + 1$, if $t < N_{\max}$ then go to the Step 3.
8. Submit the final P individuals as the results of the SGA.



$p_c = 0.8, p_m = 0.01, P = 10, N_{\max} = 10.$

Best individual: $x = 1.8224, f(x) = 4.7018, f_{\max} = 4.7447$

Crossover OR mutation?

(a decade long debate: which one is better or necessary).

It depends on the problem. In general, it is good to use both.

- Crossover is explorative, it jumps to an area “in between” two (parent).
- Mutation is exploitative, it creates random (small) diversions, staying near (or in the area of) the individual.
- Only crossover can combine information from two parents.
- Only mutation can introduce new information.
- Crossover does not change the allele frequencies of the population.
- To hit the optimum, often a ‘lucky’ mutation is needed.
- Mutation-only-EA is possible, crossover-only-EA would not work well.

Evolution Strategy (ES)

In ESs a chromosome representing an individual consists of two components being two real-valued vectors

$$\mathbf{w} = [\mathbf{x}^\top, \boldsymbol{\sigma}^\top]^\top, \quad \mathbf{x} \in \mathbb{R}^n, \quad \boldsymbol{\sigma} \in \mathbb{R}_+^n.$$

In ES mutation (not crossover) is the most important exploration technique over a solution landscape.

The mutation operator works by adding a normally distributed random vector

$$\boldsymbol{\zeta} = \mathbf{N}(\mathbf{0}, \sum_\zeta) \in \mathbb{R}^n,$$

with a mean of zero and the covariance matrix

$$\sum_\zeta = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} = \text{diag}(\sigma^2),$$

to the objective variable

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \zeta,$$

where k denotes the generation (iteration) number.

The strategy parameter is modified as follows

$$\boldsymbol{\sigma}^{(k+1)} = \boldsymbol{\sigma}^{(k)} \cdot \exp(\zeta_{\sigma 1}) \cdot \exp(\zeta_{\sigma 2}),$$

where

$$\zeta_{\sigma 1} = N(0, \tau_1) \in \mathbb{R},$$

is a single realization of zero-mean normally distributed variable with variance

$$\tau_1 = \frac{1}{\sqrt{2n}},$$

and

$$\zeta_{\sigma 2} = N(\mathbf{0}, \Sigma_{\zeta}) \in \mathbb{R}^n, \quad \Sigma_{\zeta} = \tau_2^2 \mathbb{I},$$

is a normally distributed random zero-mean vector with covariance matrix,
where

$$\tau_2 = \frac{1}{\sqrt{2\sqrt{n}}}.$$

Population sizes: μ parents, λ offspring.

Replacement procedures

- $(\mu + \lambda)$ -ES - μ current individuals and λ new individuals are combined to get the μ best ones (new population) according to their fitness values,
- (μ, λ) -ES - the best μ individuals from λ offspring form the new population,
- $(1 + 1)$ -ES, $(\mu, 1)$ -ES...



Solution process of $(\mu + \lambda)$ -ES

1. Assign the parameters for ES, such as λ , μ , and N_{\max} .
2. Generate μ uniformly distributed individuals randomly to form the initial population and evaluate their fitness values.
3. Generate λ individuals:
 - randomly select two individuals with replacement to perform crossover (optional),
 - perform a mutation for every gene of the offspring,
 - insert the mutant into the new population.
4. Calculate the fitness value for every new individual in the new population.
5. Combine μ current and λ new individuals and pick the μ best ones to form a new population.
6. Increase the generation index $t = t + 1$, if $t < N_{\max}$ then go to the Step 3.
7. Submit the final μ individuals as the results of ES.

Evolutionary Programming

In EP a finite state machine can be used to represent the chromosome.

Solution process of EP:

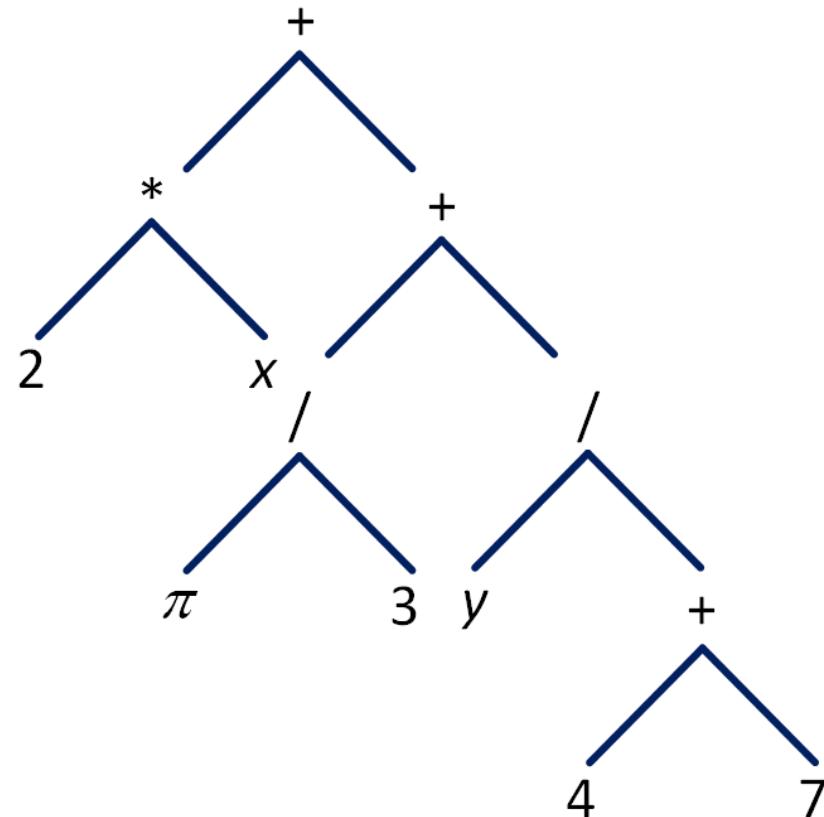
1. Assign the parameters for EP, such as λ , μ , and N_{\max} .
2. Generate μ uniformly distributed individuals randomly to form the initial population and evaluate their fitness values.
3. Perform a mutation for every gene of the μ individuals to get a new population of μ individuals (usually $\lambda = \mu$).
4. Calculate the fitness value for every new individual in the new population.
5. Combine μ current and μ new individuals and pick the μ of them using a probabilistic function based on the fitness.
6. Increase the generation index $t = t + 1$, if $t < N_{\max}$ then go to the Step 3.
7. Submit the final μ individuals as the results of EP.

Genetic Programming

Trees can represent different types of expressions:

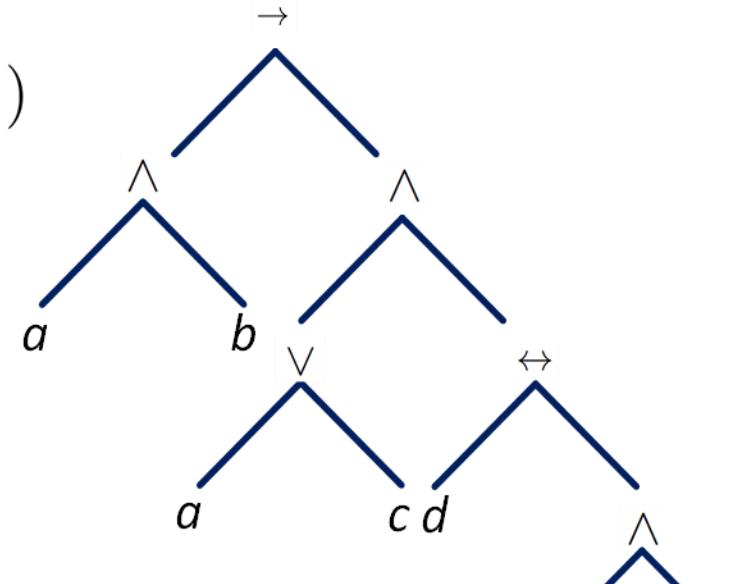
- arithmetic formulas:

$$\frac{\pi}{3} + 2x + \frac{y}{4+7}$$



- logical expressions:

$$(a \wedge b) \rightarrow ((a \vee c) \wedge (d \leftrightarrow (b \wedge c)))$$



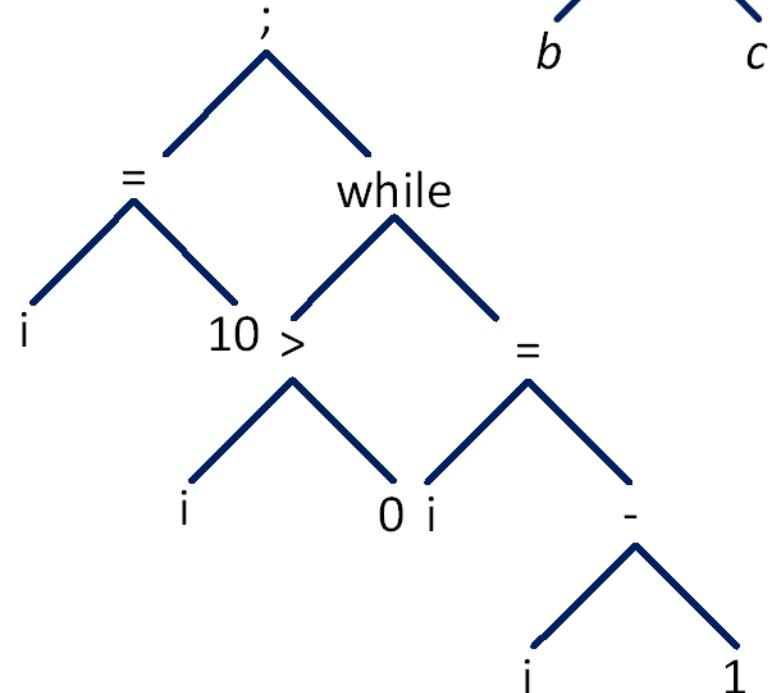
- computer programs:

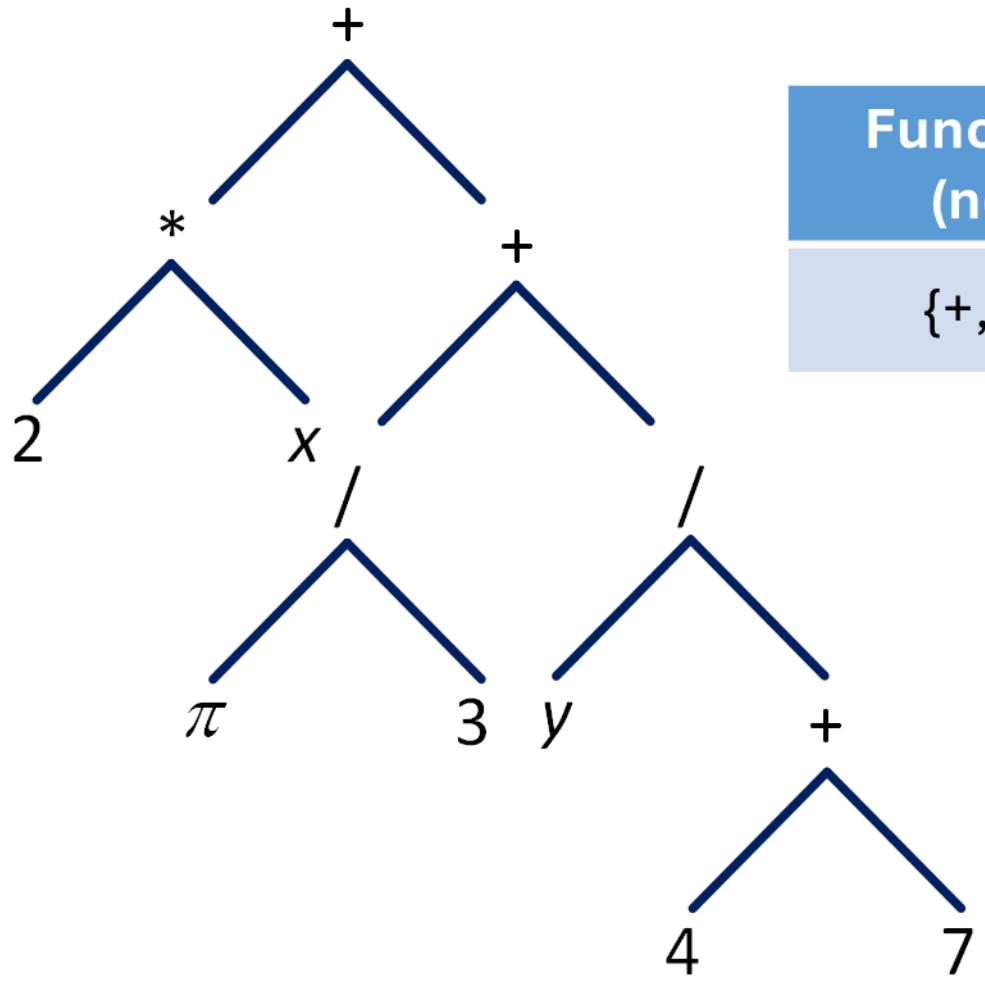
```
i = 10;
```

```
while (i > 0) {
```

```
    i = i - 1;
```

```
}
```

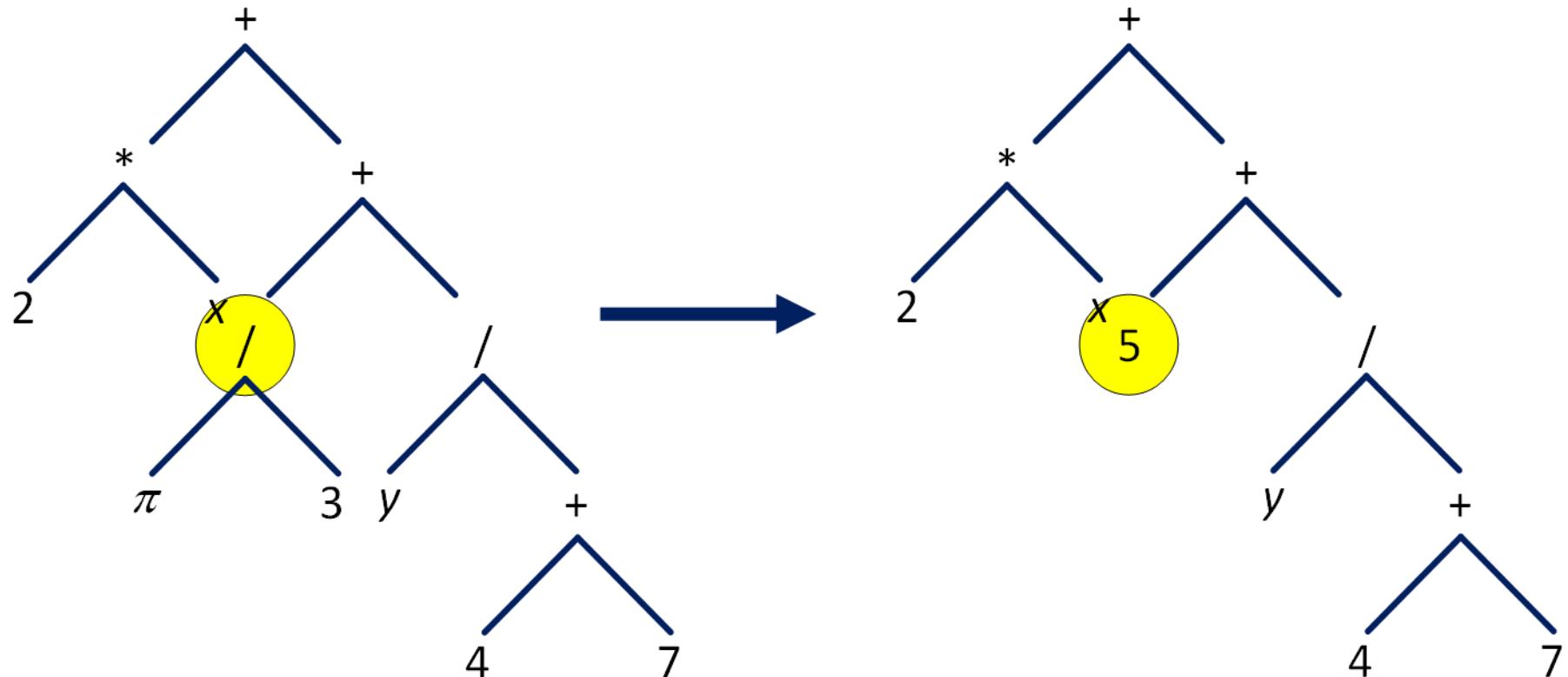




Function set (nodes)	Terminal set (leaves)
$\{+, -, *, /\}$	$\text{IR} \cup \{x, y\}$

Mutation for tree representation

Selects a node at random from the tree, and replaces the subtree starting there with a randomly generated tree:



Parameters:

- the probability of choosing mutation (mutation rate),
- the probability of choosing a point within the parent tree as the root of the subtree to be replaced.

Recombination for tree representation

Subtree crossover - interchanges the subtrees starting at two randomly selected nodes in the given parents.

Parameters:

- the probability of choosing recombination,
- the probability of choosing nodes as crossover points.

