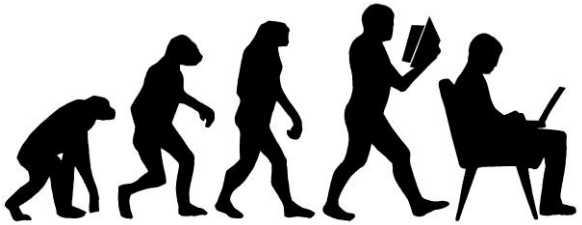
Evolutionary Algorithm: application to event selection in High Energy Physics

Seminari Dottorandi XXXII ciclo

21/09/17 Matteo Cencini

Darwinian Evolution

Four postulates



1. Individuals within species are variable

2. Some of the variations are passed on to offspring

3. In every generation, more offspring are produced than can survive

4. Survival and reproduction of individuals are not random: the individuals with the most favourable variations survive and go on to reproduce

   → they are **naturally selected**.

Concept and Classes

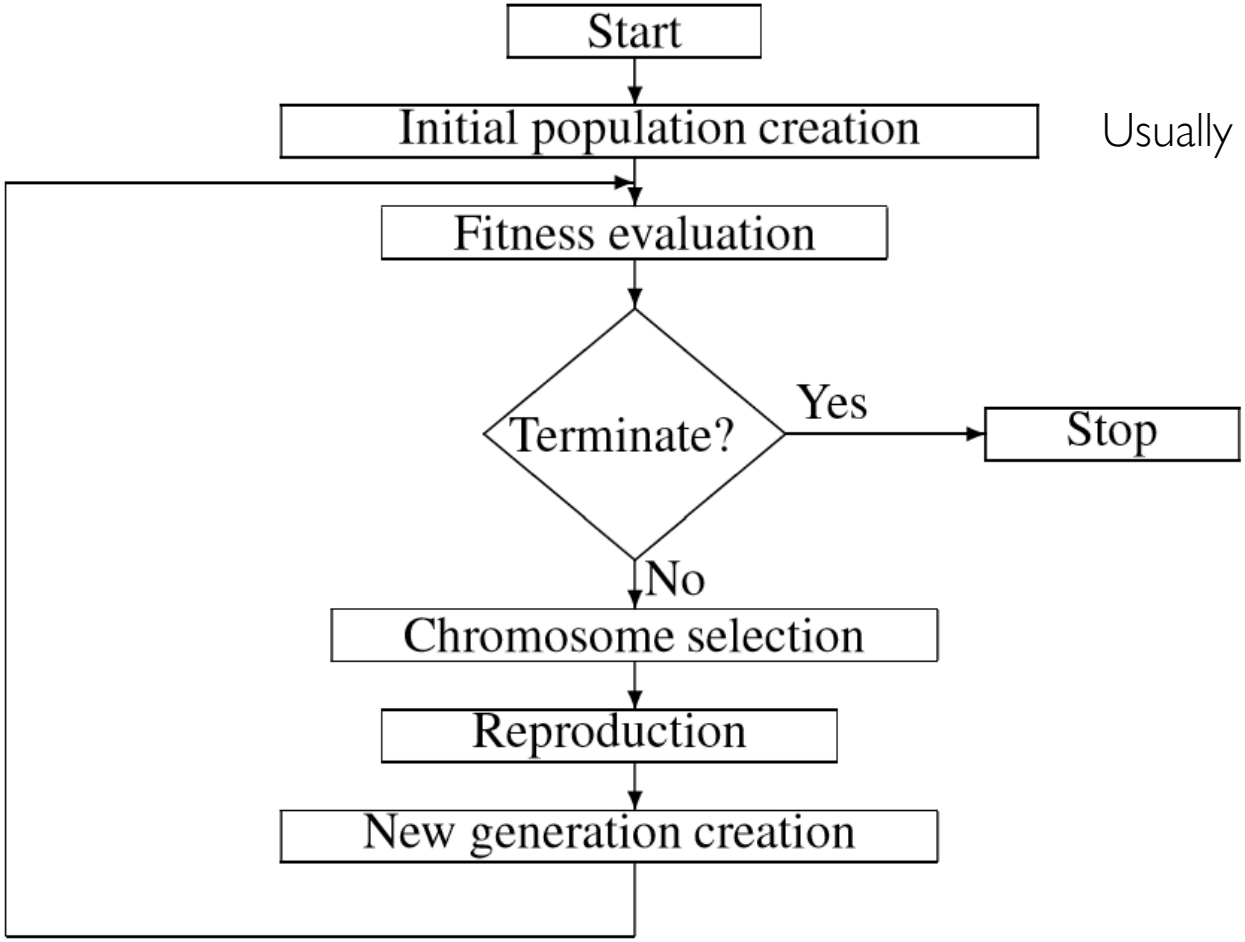
Solving complex problems by modelling the natural evolution

Some types of Evolutionary Algorithms (EA):

Genetic Algorithm (GA): probabilistic search algorithm to find the solution of a problem using the principle of Natural selection.

Genetic Programming (GP): solution is a computer program to solve the specific problem, rather than the solution of the problem itself.

Algorithm Flowchart



Usually randomized!

Fig. 1 from L. Teodorescu, **arXiv:0804.0369v1**(2008)

Solution Representation

Chromosome: candidate solution of the problem. Can be subdivided into **genes**.

GA: Binary String

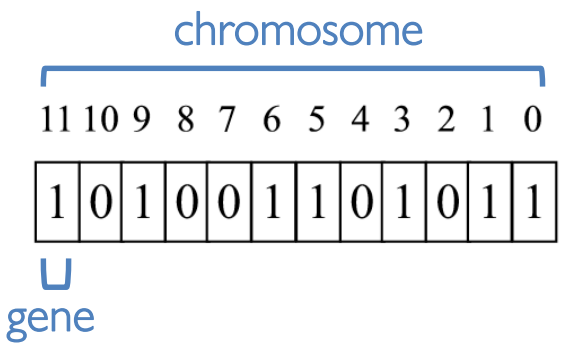


Fig. 2.3 from X. Yu, M. Gen, Introduction to evolutionary algorithms

GP: Binary Tree

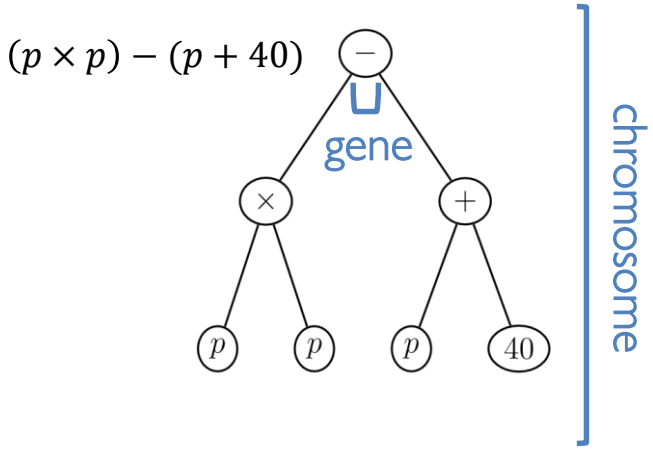
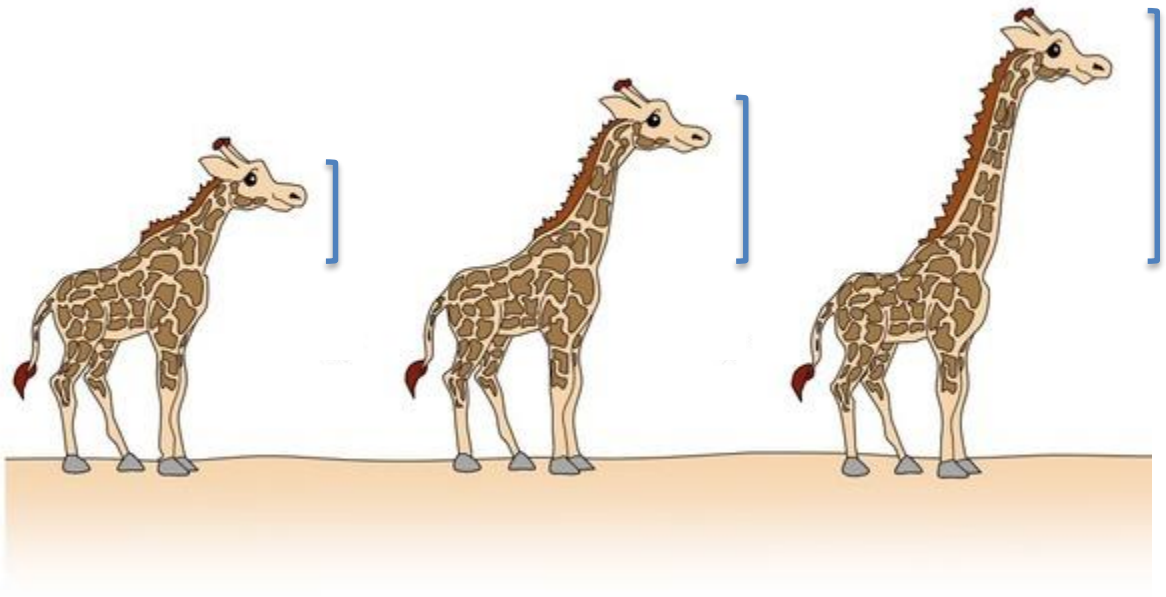


Fig. 1 from J.M. Link et al, Nuclear Instruments and Methods in Physics Research A551(2005) 504-527

Initial Population



Neck length:
gene value

Fitness Function

$$f(C) = x$$

C: Chromosome

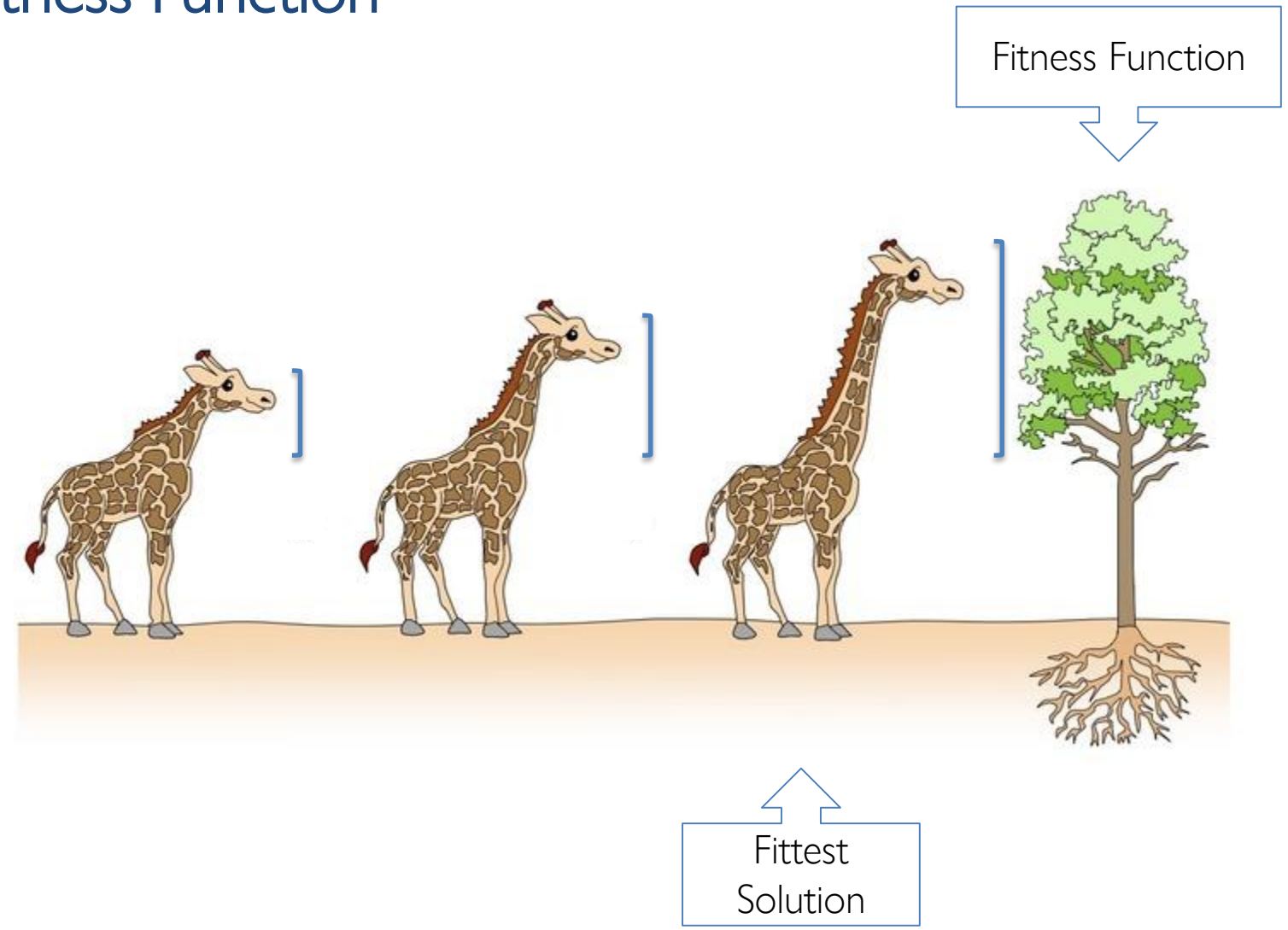
X: fitness value

Measure of how good the solution is for the given problem

Depends on the particular problem

Reflects optimization criteria and problem constraints

Fitness Function

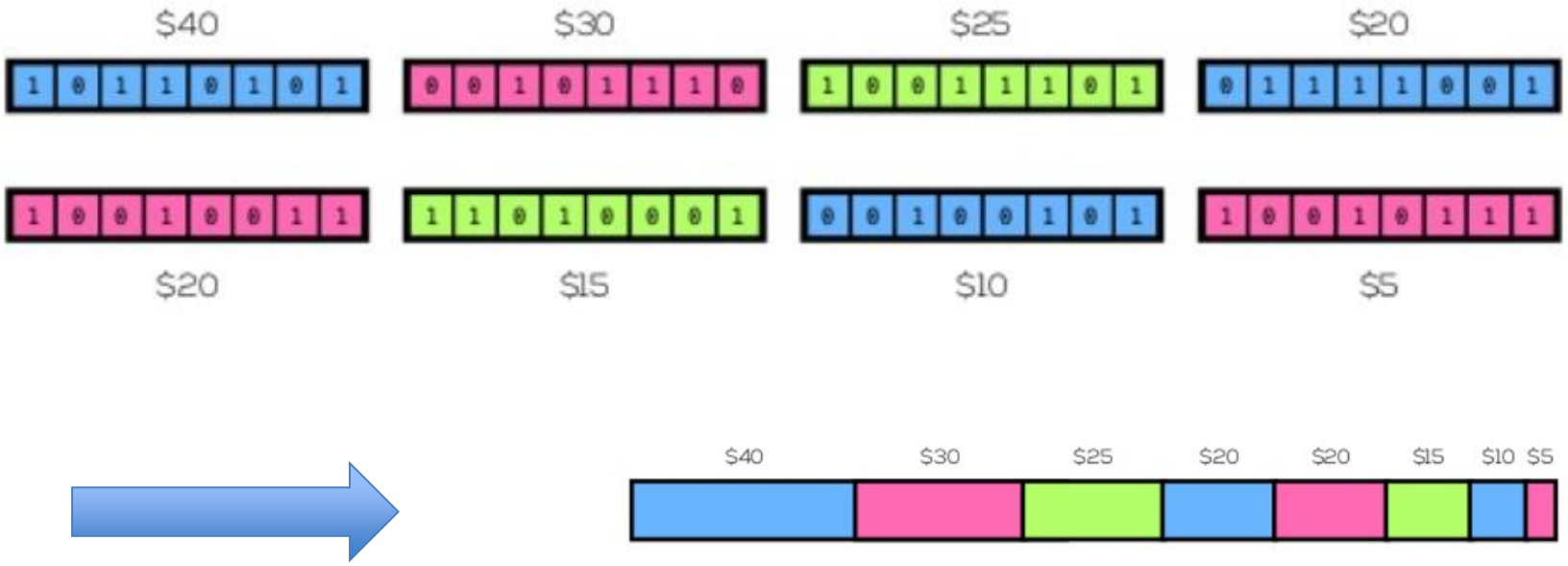


Selection Operators (1)

Select individuals for applying genetic operators and for creating the new generation.

Some examples:

Proportional Selection: the probability to select an individual is proportional to the fitness value

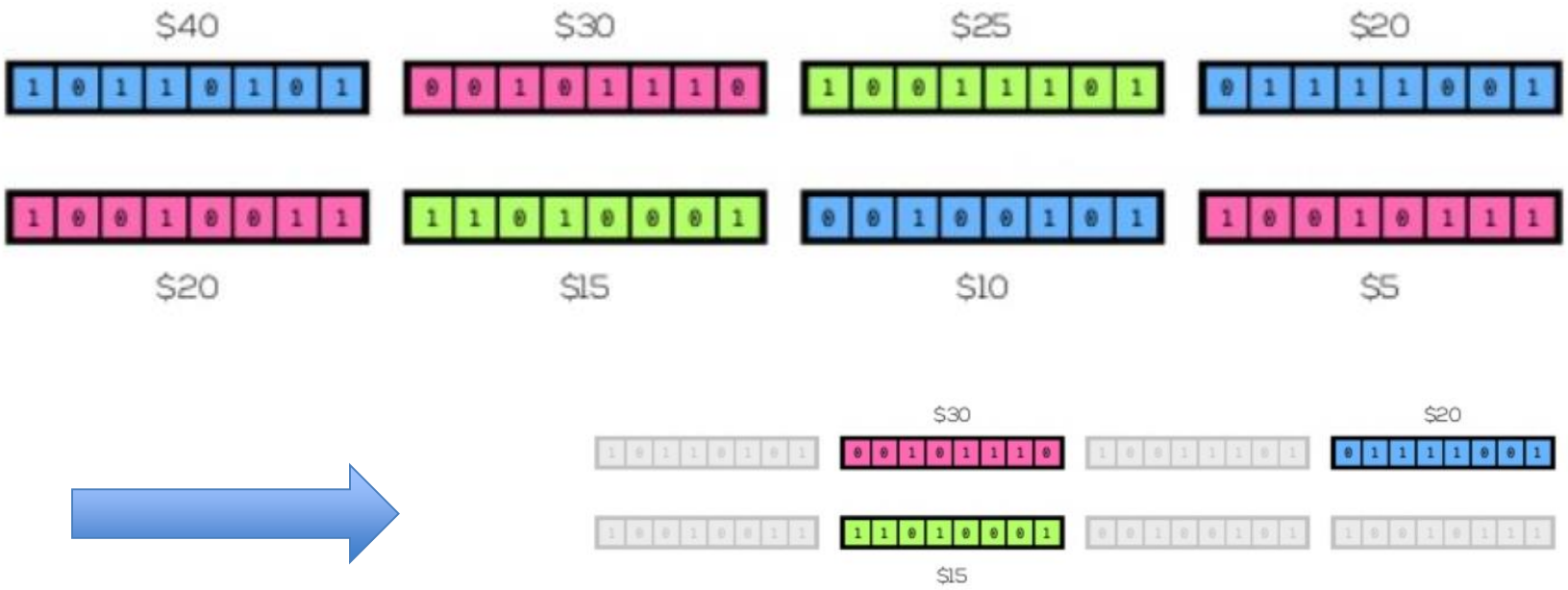


Selection Operators (2)

Select individuals for applying genetic operators and for creating the new generation.

Some examples:

Tournament Selection: the selection is performed within a randomly chosen sub-group

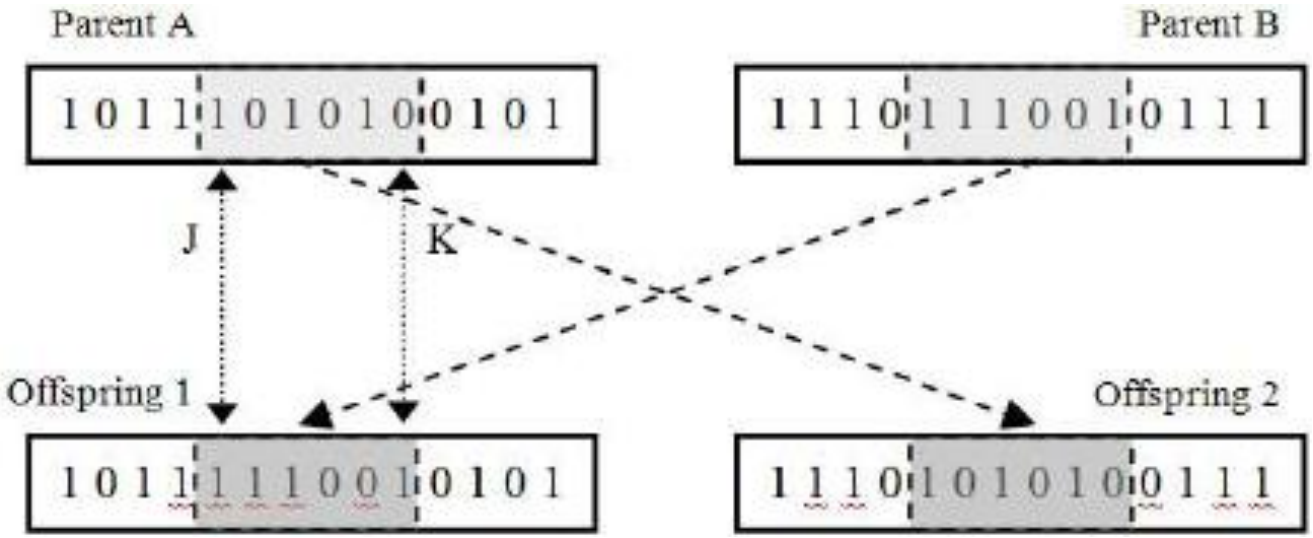


Reproduction Operators (1)

Reproduction operators are applied on the selected individuals to create offspring which will constitute the next generation.

Some examples:

Cross-over: combines genes of two parents, producing two new individuals

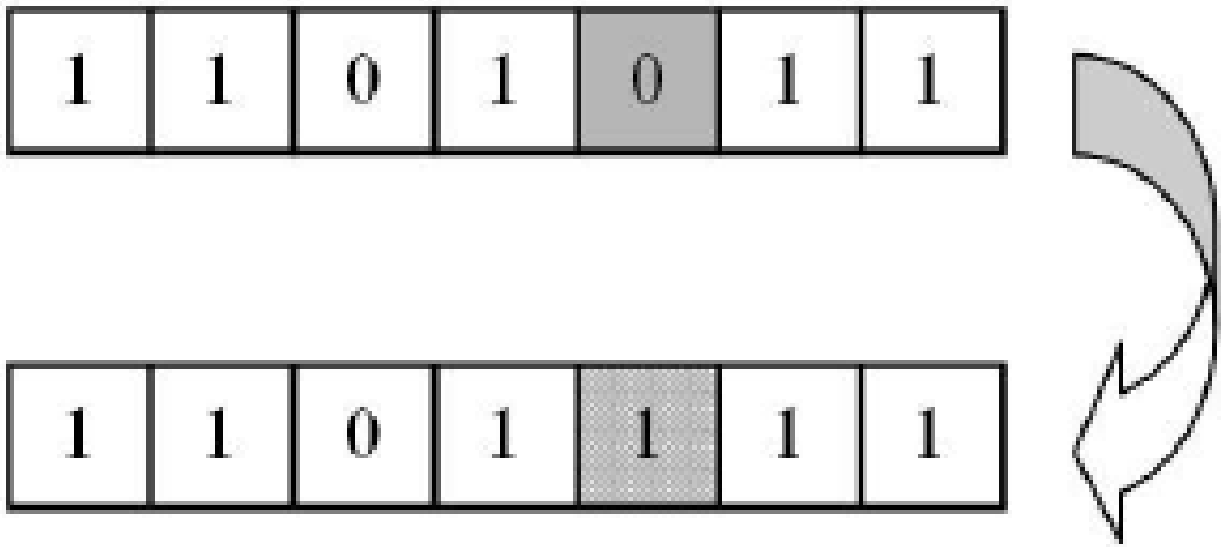


Reproduction Operators (2)

Reproduction operators are applied on the selected individuals to create offspring which will constitute the next generation.

Some examples:

Mutation: randomly changes the values of genes in the chromosome, introducing new genetic material



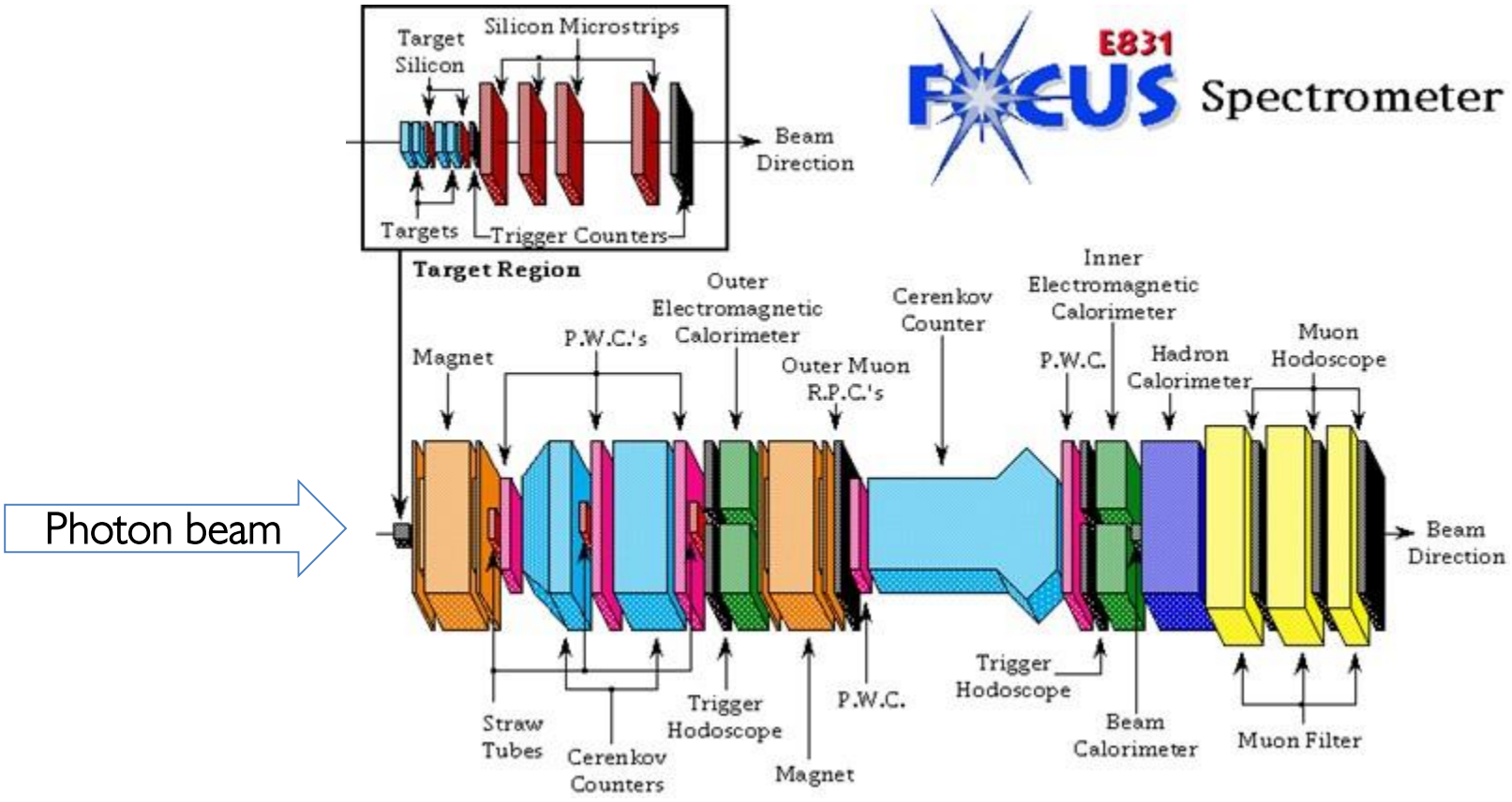
Reproduction Operators (3)

Reproduction operators are applied on the selected individuals to create offspring which will constitute the next generation.

Some examples:

Elitism or Cloning: copies the best individuals in the next generation, without any modifications

Event selection for FOCUS: setup



Event selection for FOCUS Experiment: aim

Purpose: Analysis of experimental data to measure the following branching ratio in the charm photoproduction experiment

$$\frac{\mathrm{BR}(D^+ \to K^+\pi^+\pi^-)}{\mathrm{BR}(D^+ \to K^-\pi^+\pi^+)}$$

Doubly Cabibbo suppressed

Cabibbo Favoured

J.M. Link et al, Nuclear Instruments and Methods in Physics Research A551(2005) 504-527

Event selection for FOCUS: solution form (1)

Class of EA: Genetic Programming

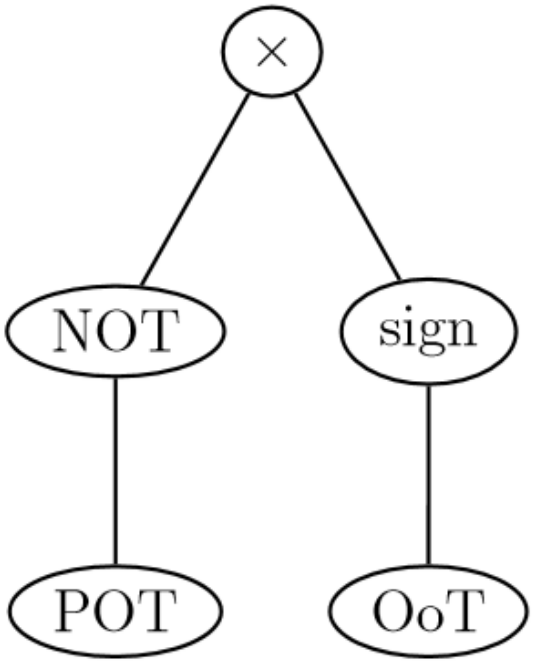
Solution Representation: Binary tree

- Variables: vertexing variables, kinematic variables ($p, p_T, \tau, Oo\mathrm{T}, \mathrm{POT}\ldots$)

- Operators: Boolean operators, algebraic and trigonometric functions ($+,\ \times, \cos(), AND, OR\ \ldots$)

Event selection for FOCUS: solution form (2)

Example of solution:



Fittest tree from the first gen:

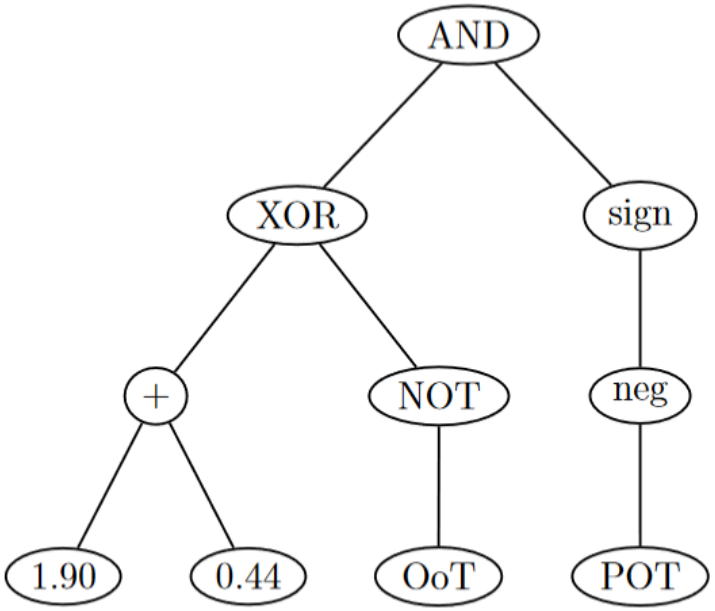
- Requires (NOT) **Production** vertex **Outside** the **Target** **(POT)** and vertex decay **Outside of Target (OoT)**

- Short tree

Fig. 13a from J.M. Link et al, Nuclear Instruments and Methods in Physics Research A551(2005) 504-527

Event selection for FOCUS: solution form (3)

Example of solution:



Second fittest tree from the first gen:

- Same requirements of best fitting solution, but longer tree

Fig. 13b from J.M. Link et al, Nuclear Instruments and Methods in Physics Research A551(2005) 504-527

Event selection for FOCUS: fitness function

Fitness Function: $\frac{S+B}{S^2} \times 10000 \times (1 + 0.05 \times n)$

- S, B: number of signal and background events
- n: number of nodes in the tree (privileges shorter trees)

Event selection for FOCUS: fitness evaluation

Fitness Evaluation:

1. Each tree is tested on each physical event

2. For surviving events, signal **S** and background **B** are fitted

3. **S** and **B** are used to compute fitness

Event selection for FOCUS: operators

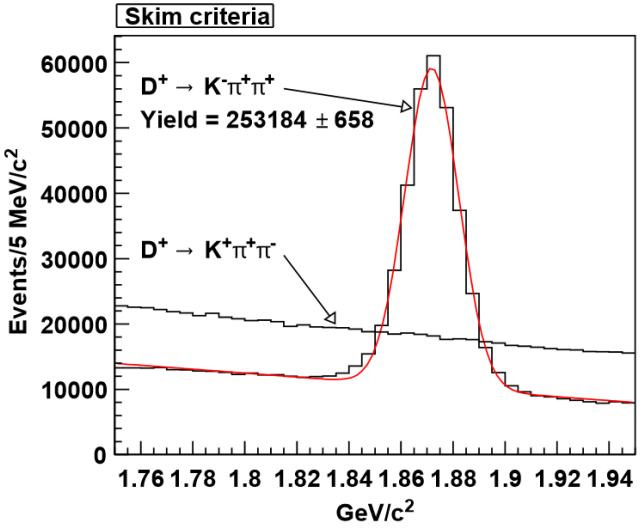
Selection: Proportional selection

- 80% probability of performing selection within best individuals subgroup (subgroup size: 320/number of individuals)
- 20% probability of performing selection within the rest of the population

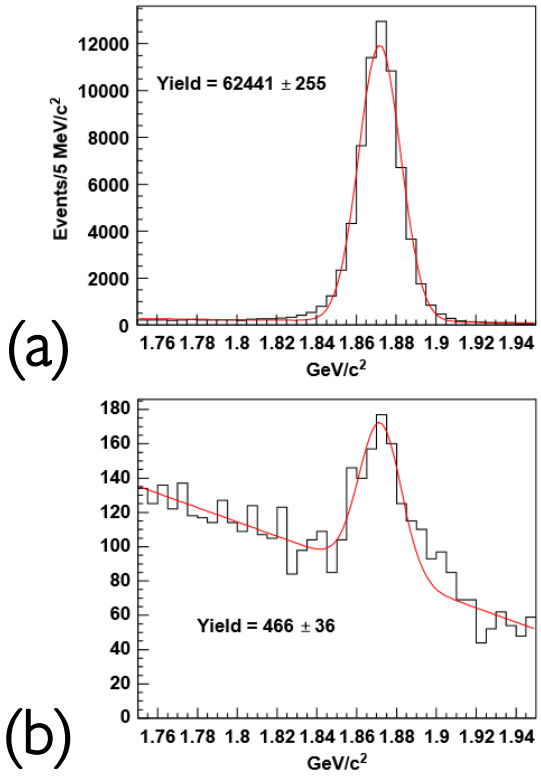
Reproduction:

- Cross-over
- Mutation
- Cloning

Event selection for FOCUS Experiment: results



The initial $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K^+\pi^+\pi^-$ candidate distributions. Fig. 9 from J.M. Link et al, Nuclear Instruments and Methods in Physics Research A551(2005) 504-527



$D^+ \to K^-\pi^+\pi^+$ (a) and $D^+ \to K^+\pi^+\pi^-$ (b) signals selected by genetic programming. Fig. 10 from J.M. Link et al, Nuclear Instruments and Methods in Physics Research A551(2005) 504-527

Predicted BR (PDG): $(0.75 \pm 0.16)\%$
Traditional cut BR: $(0.65 \pm 0.08 \pm 0.04)\%$
GP cut BR: $(0.76 \pm 0.06)\%$

Other applications

1. Optimization of cuts for event selection

2. Optimization of parameters of theoretical models

3. Dalitz plot analysis

Thank you for your attention!

REFERENCES:

L. Teodorescu, Evolutionary Computation in High Energy Physics, **arXiv:0804.0369v1**(2008)

J.M. Link et al, Application of genetic programming to high energy physics event selection, Nuclear Instruments and Methods in Physics Research A551(2005) 504-527