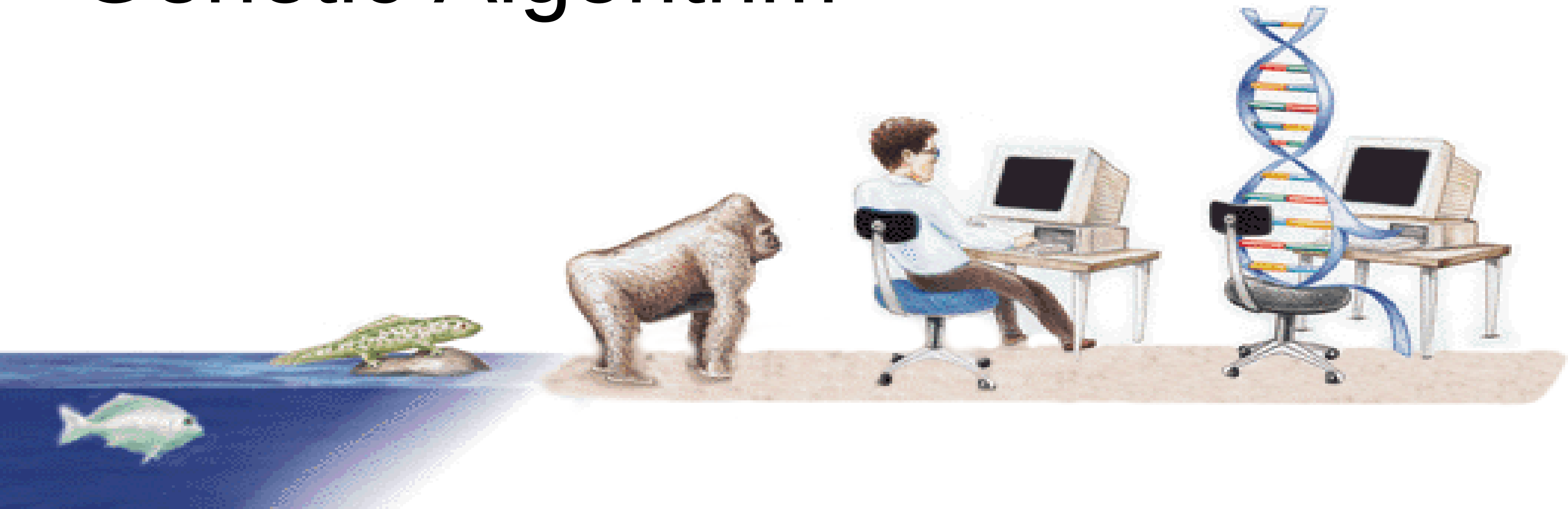
Intelligent systems

Lecture 9:

Dr. Bassam Kurdy

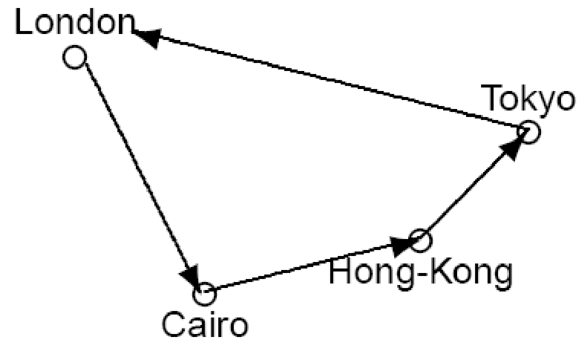
Genetic Algorithm



Combinatorial Problem

- No clear method to find the optimum solution
- Seem to have only one way of finding a solution: by random search
- They have many possible combinations and it is impossible to try all of them.
- “Neural networks are second best way to solve about anything… and GA are the third.”

The travel salesman problem



- A salesman goes on a tour around several cities.
- Each city have to be visited only once.
- What is the shortest tour ?
- N City => (n-1)! possible solution
- 50 City => $6*10^{62}$
  - each calculation 1 ηs => $2*10^{45}$ years to try all possible solution !!!

The Timetabling problem

- Balancing resource (Teachers, students, rooms)
  - Teachers at one place at one time
  - Rooms can only be used by one teacher
  - Each teacher can only teach certain subjects
  - …

Charles Darwin

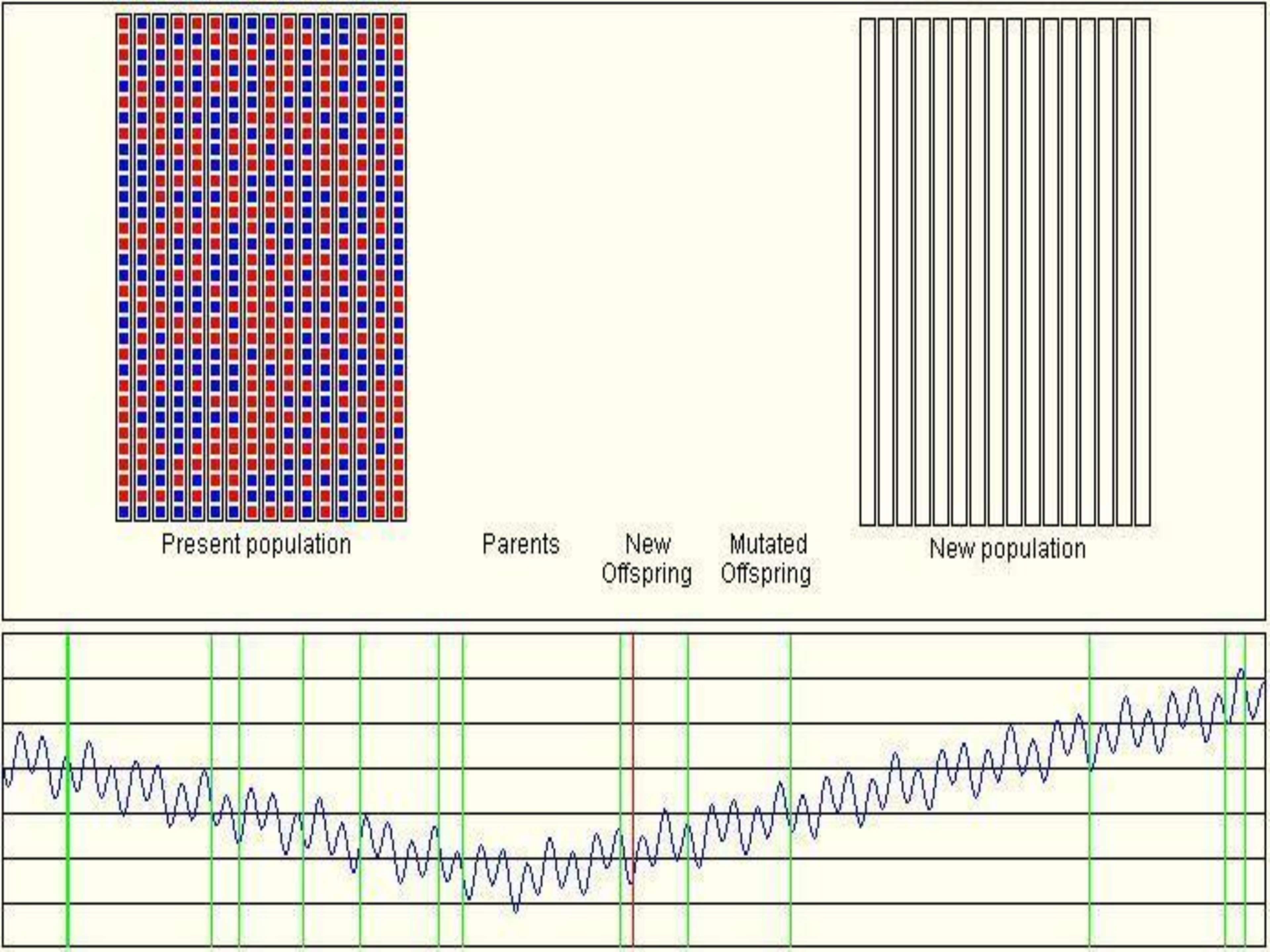
- Living organisms are fighting the forces of nature to survive. Those who are the fittest (strongest, fastest, biggest) are most likely to survive
- Those who survive mate and reproduce **(selection)**
- Children are similar **(inheritance)**, but not exactly like parents because of cross-fertilization and mutation, thus children can be more or less fitness than parents
- Children repeat the path of their parents, after several generations the organisms become much fitter.

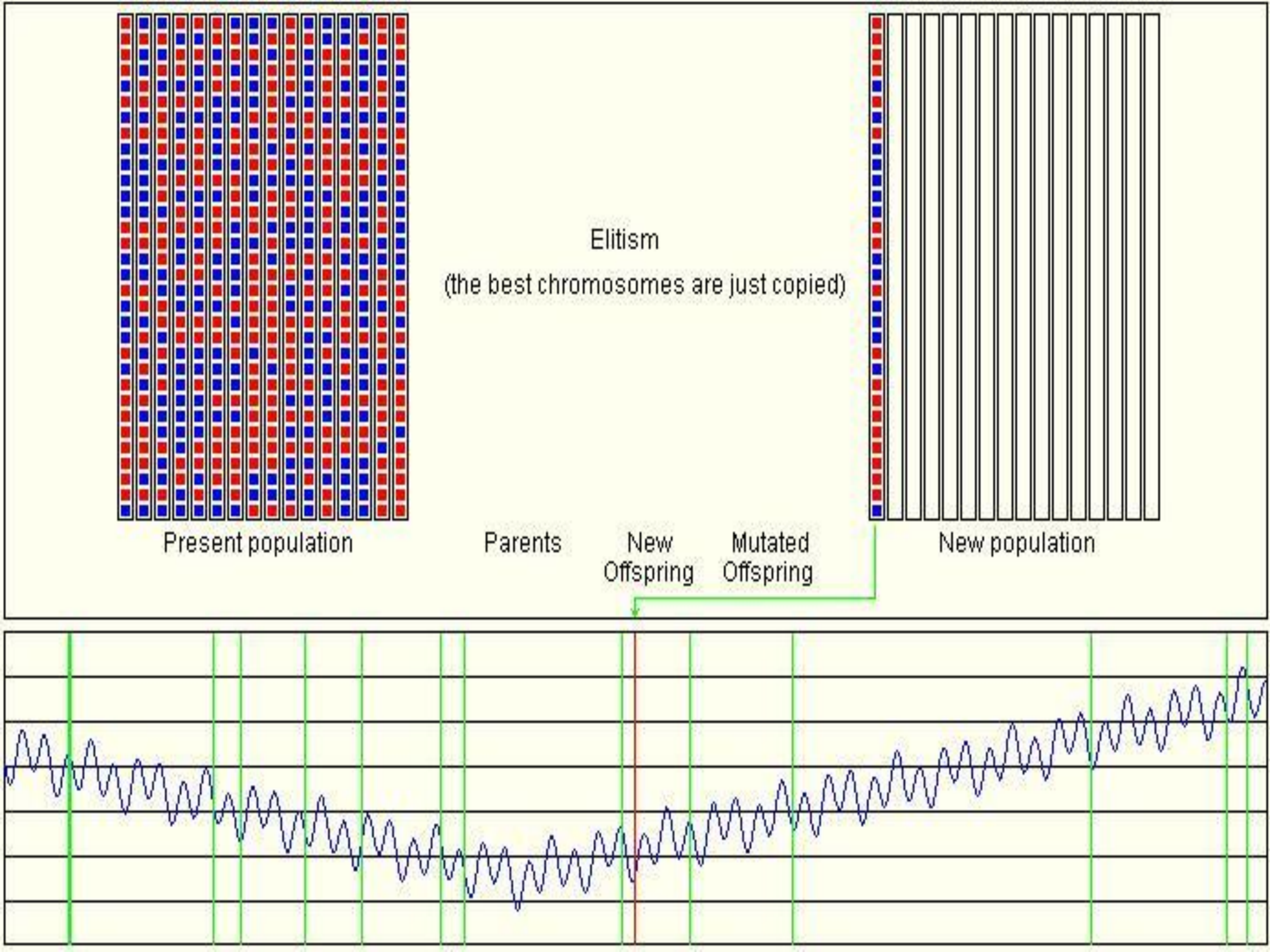
Genetic Algorithm

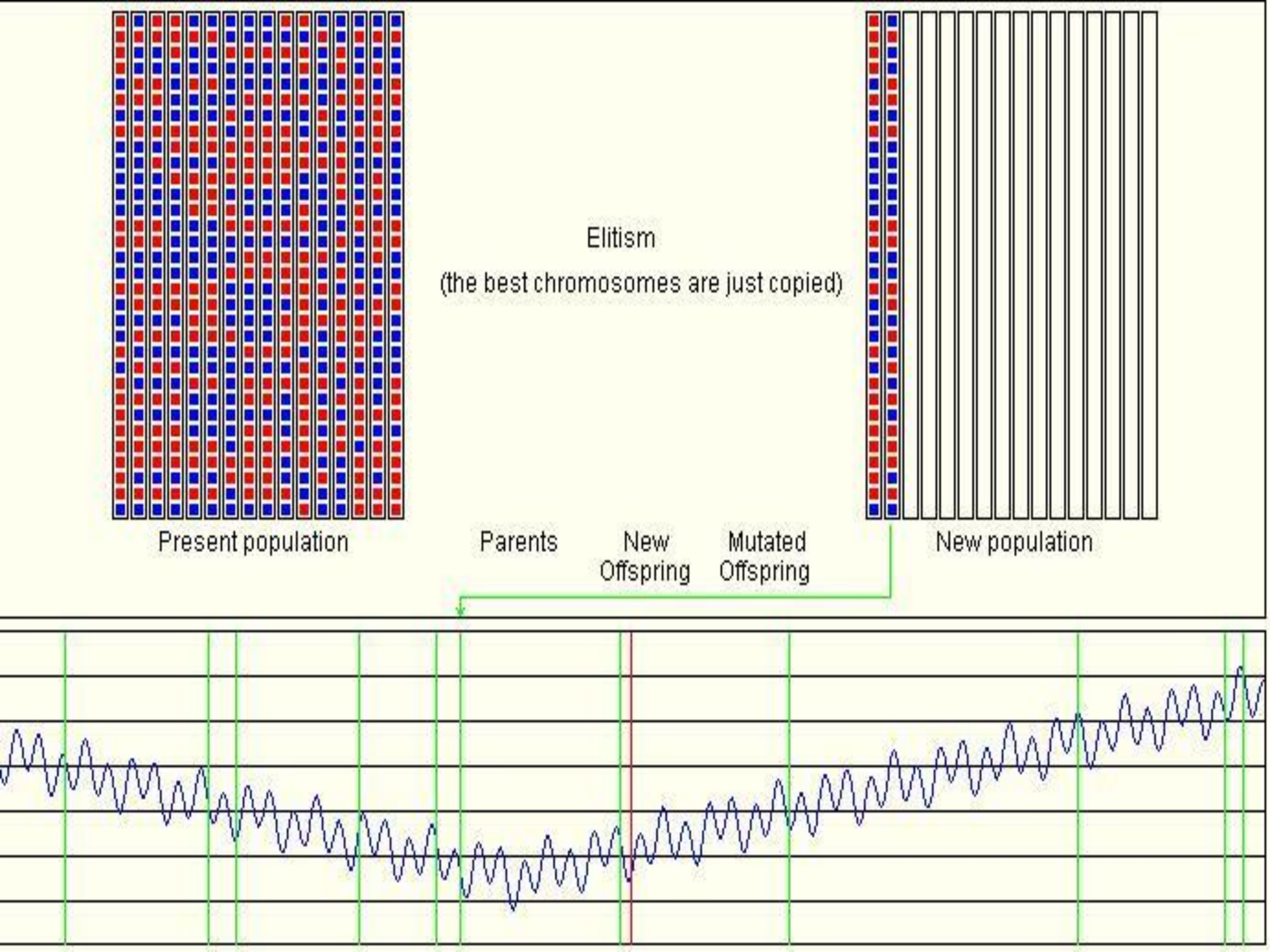
- Suppose that there is many possible solutions for the problem: $x_1,x_2,x_3,x_4,\ldots$. The main idea is to view each solution x_i of the problem as an individual living organism.
- the number of possible solution can be incredibly large $n \rightarrow \infty$, so we consider $m < n$ and chose a **Population:**
- $P(t) = \{x_1^t, x_2^t, \ldots , x_m^t\}$
- With time the organisms and the whole population will be evolving.

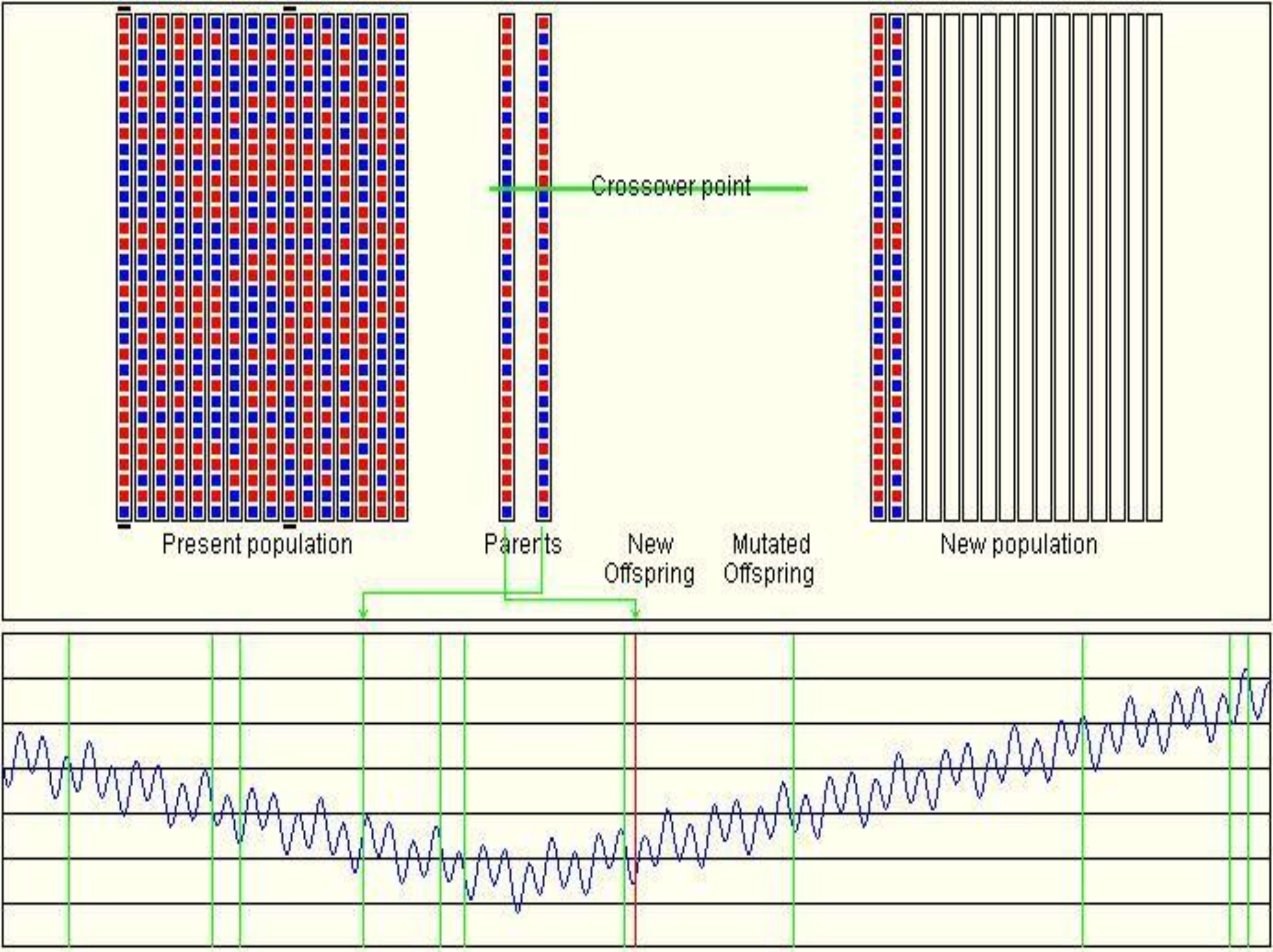
Genetic Algorithm

- The next slides show how Genetic Algorithm work. Like Charles Darwin theory

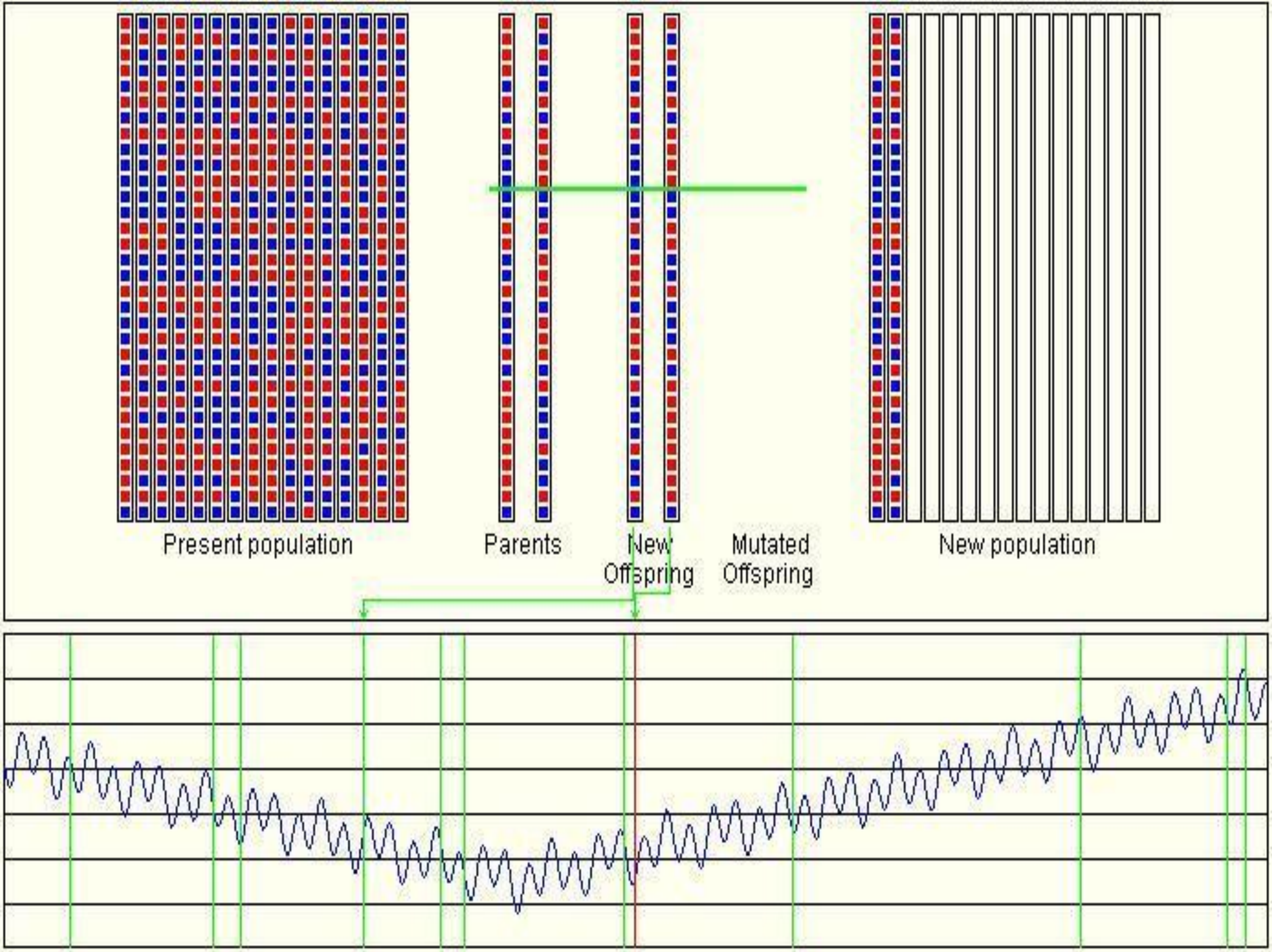

Present population
Parents
New
Offspring
Mutated
Offspring
New population

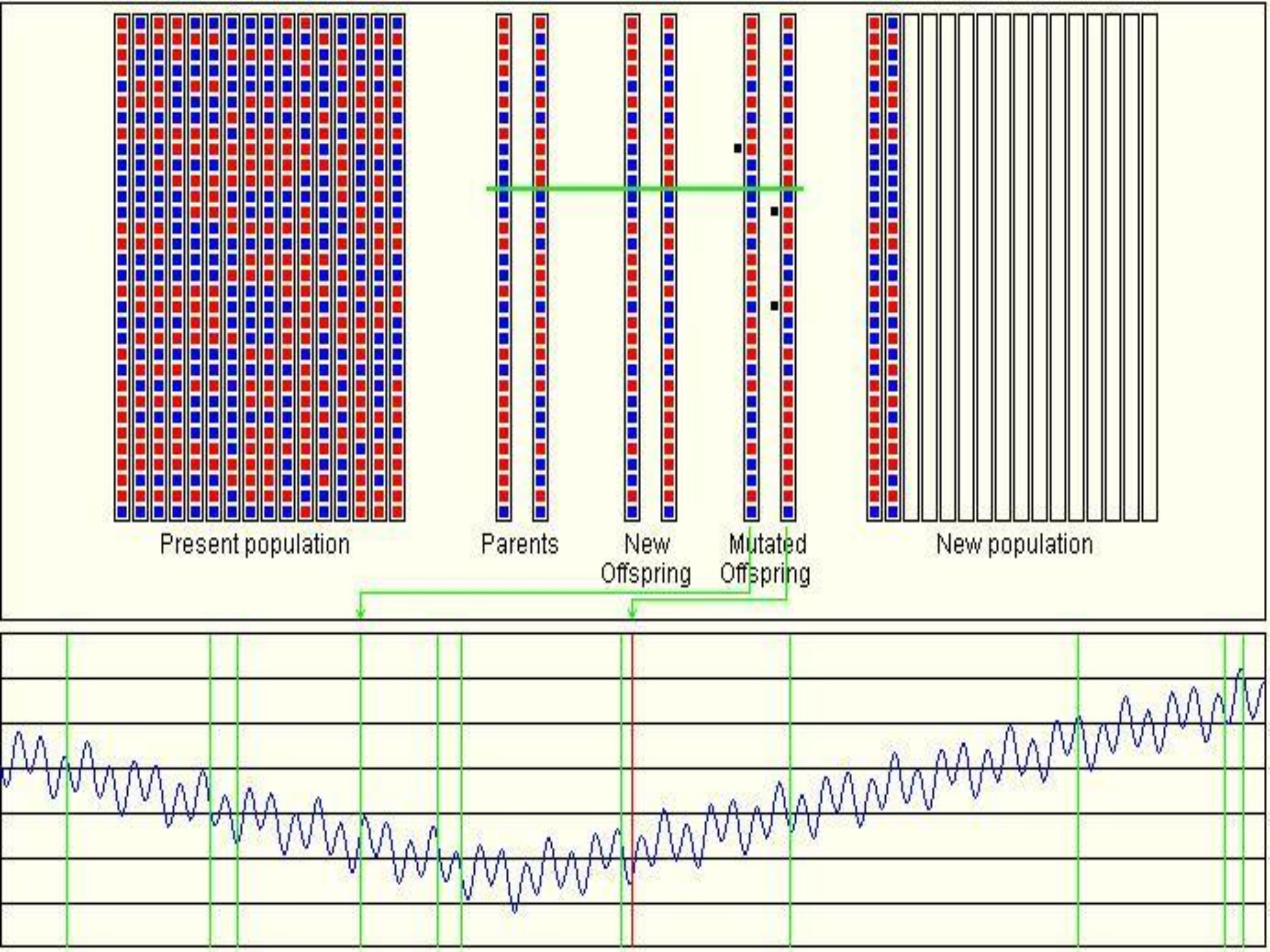

Elitism
(the best chromosomes are just copied)
Present population
Parents
New Offspring
Mutated Offspring
New population


Elitism
(the best chromosomes are just copied)
Present population
Parents
New Offspring
Mutated Offspring
New population

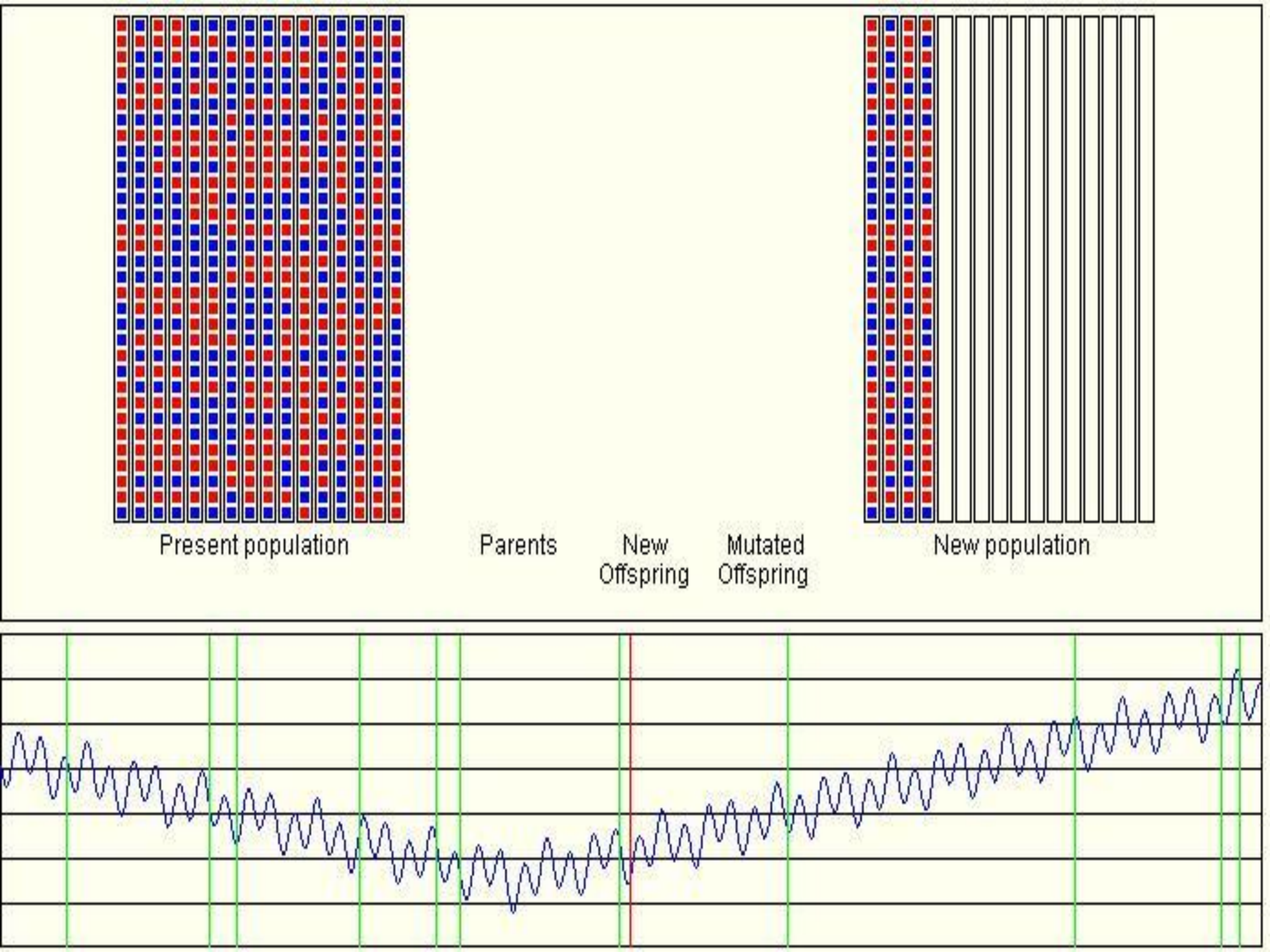


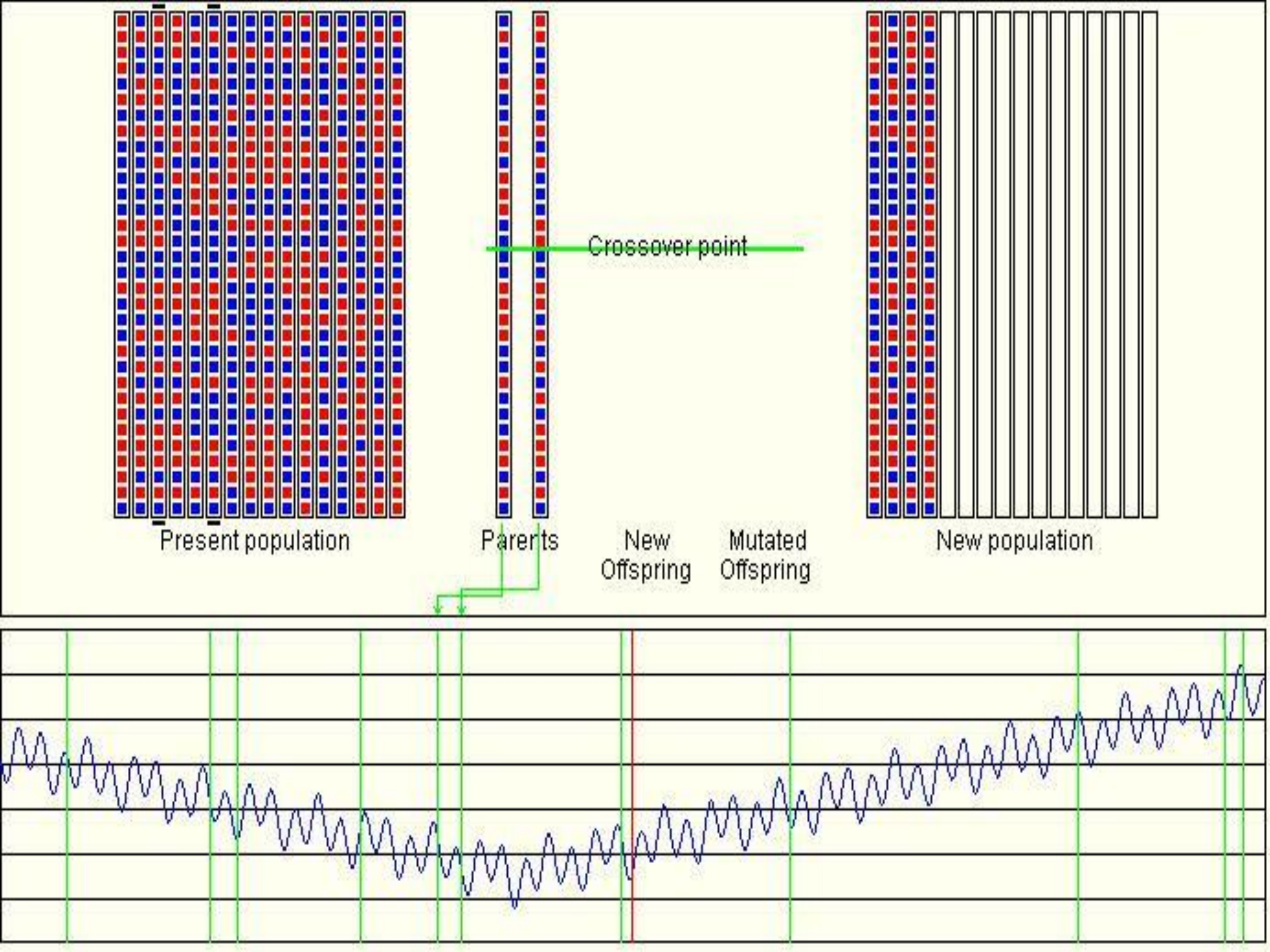
Crossover point
Present population
Parents
New Offspring
Mutated Offspring
New population

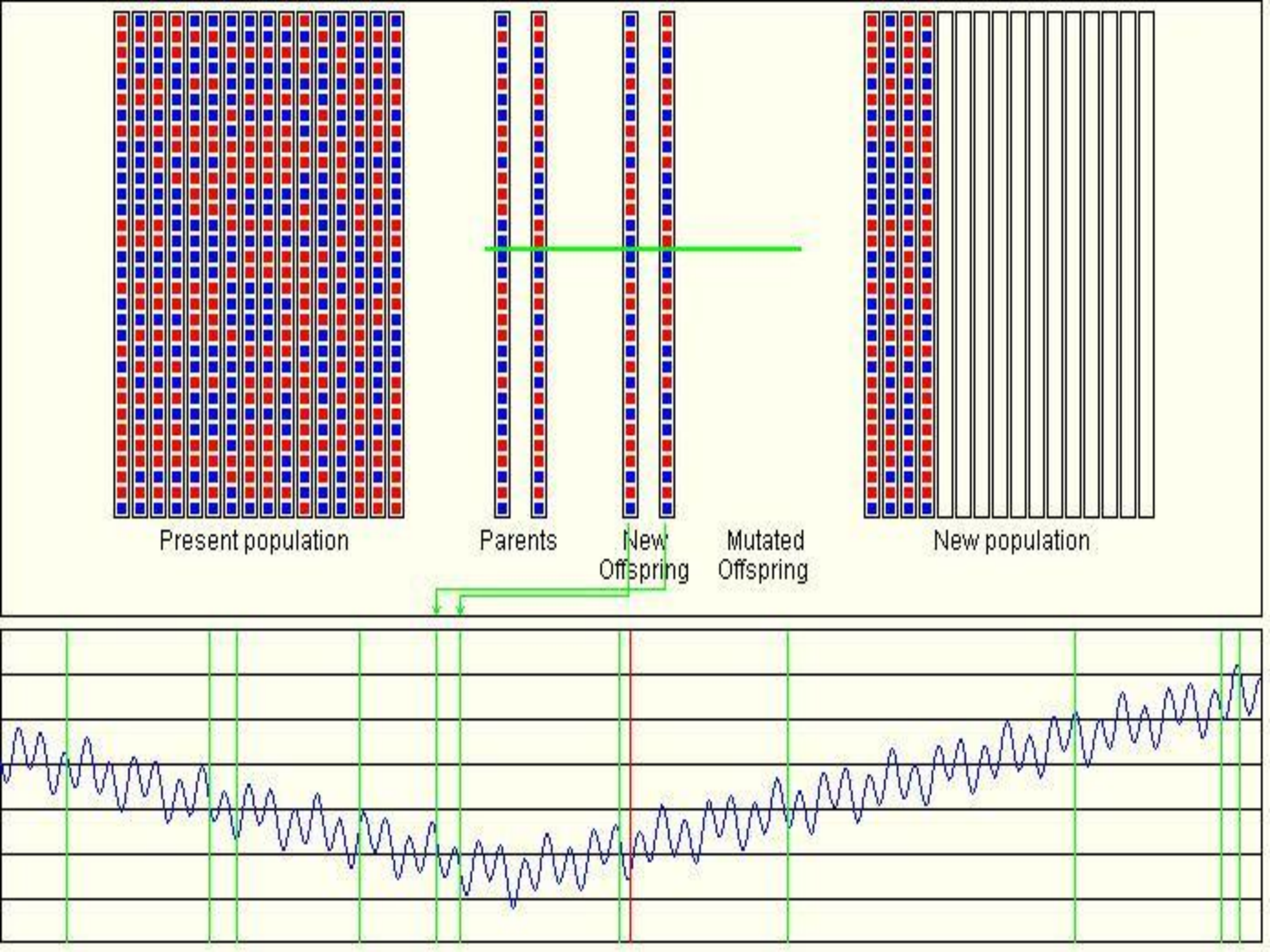

Present population
Parents
New
Offspring
Mutated
Offspring
New population



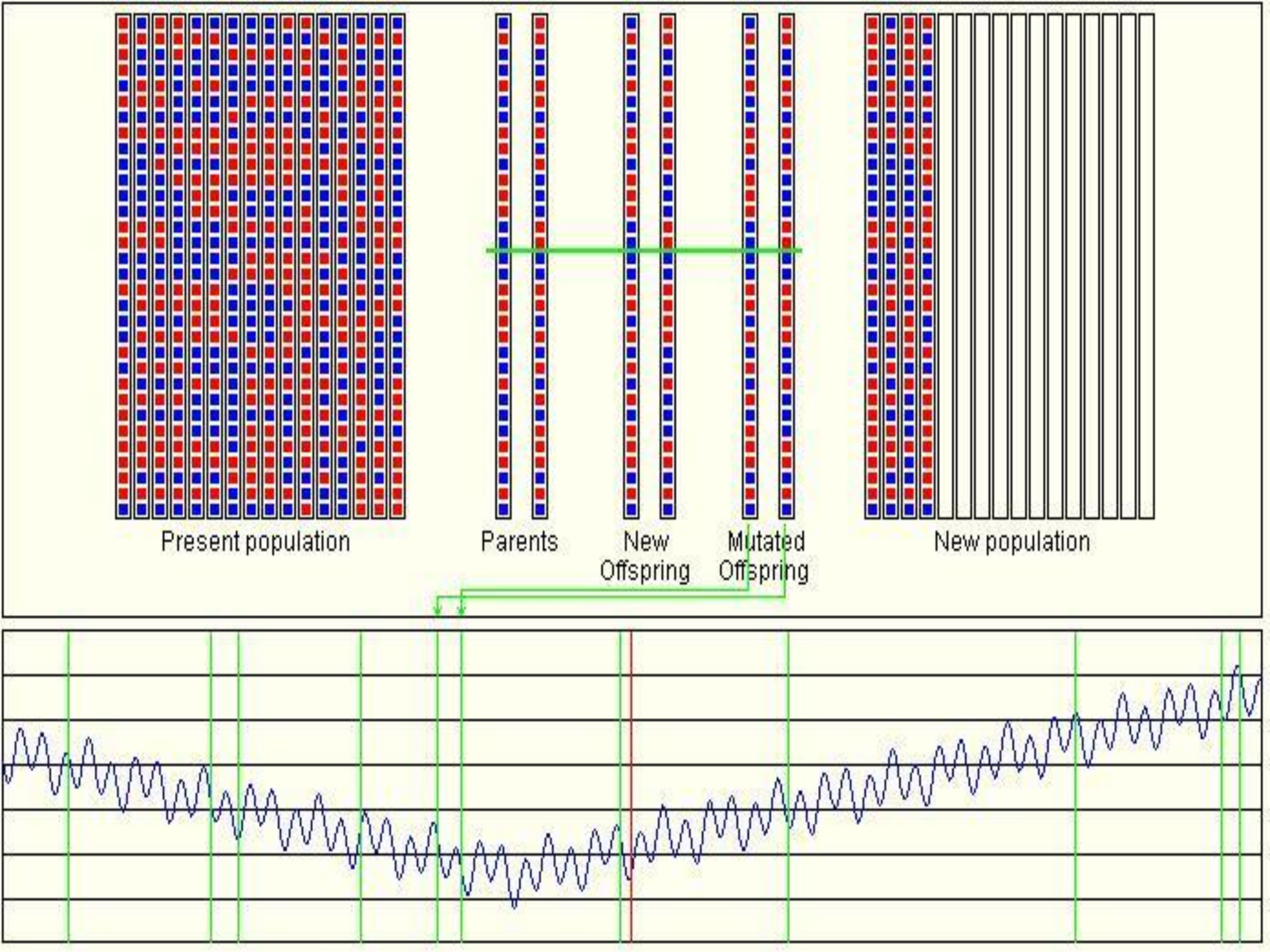
Present population
Parents
New Offspring
Mutated Offspring
New population

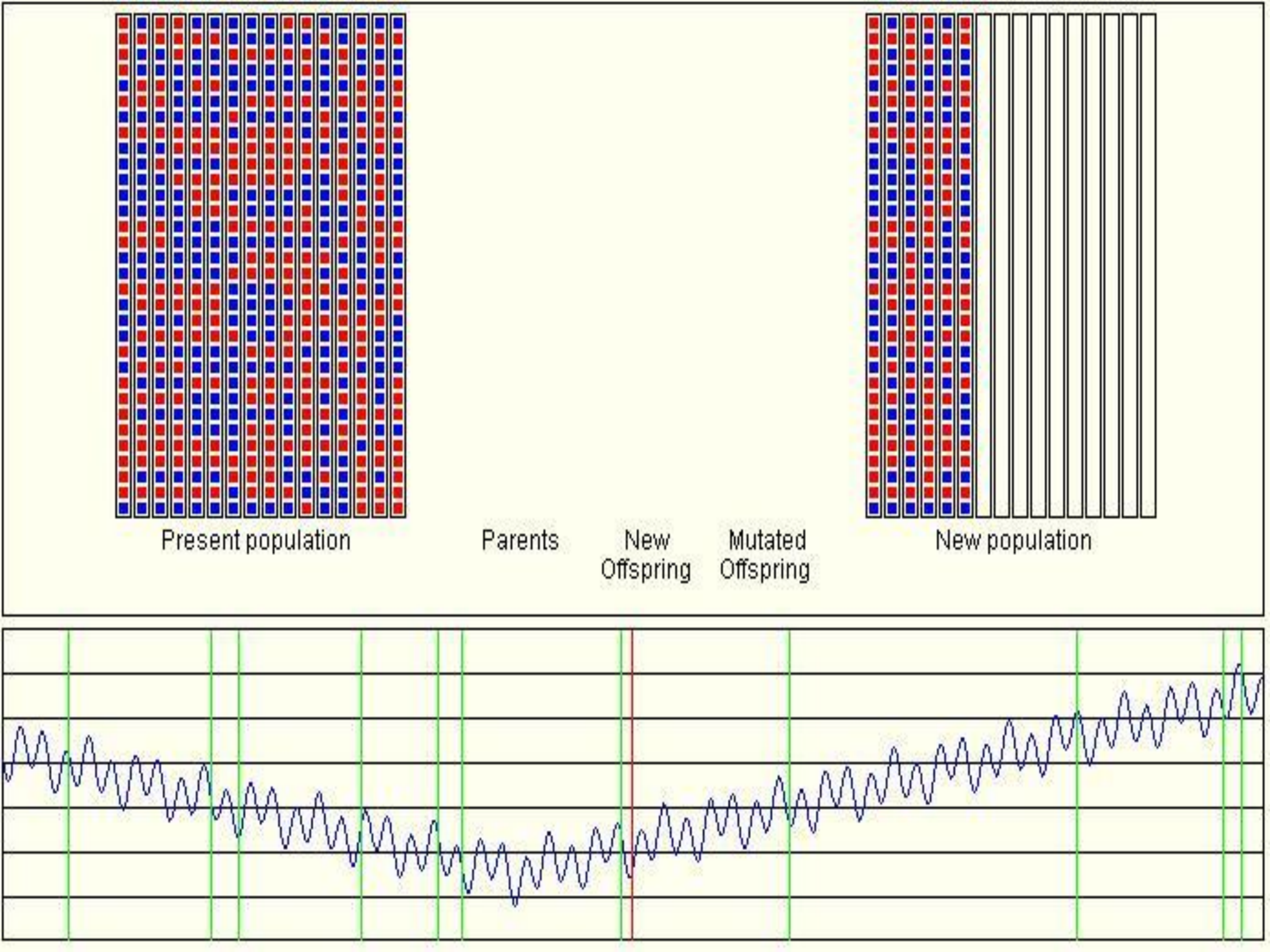

Present population
Parents
New Offspring
Mutated Offspring
New population

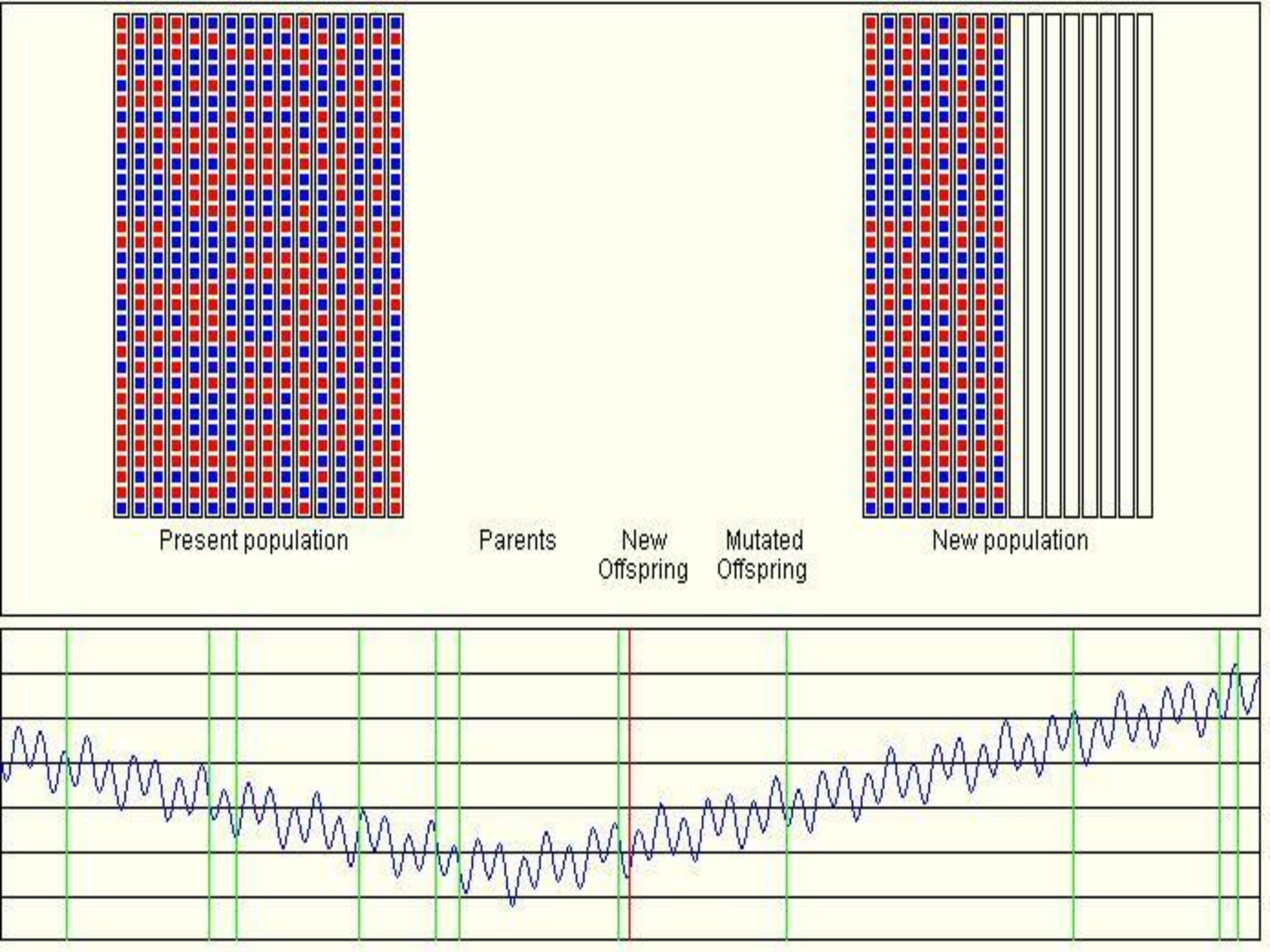

Crossover point
Present population
Parents
New Offspring
Mutated Offspring
New population

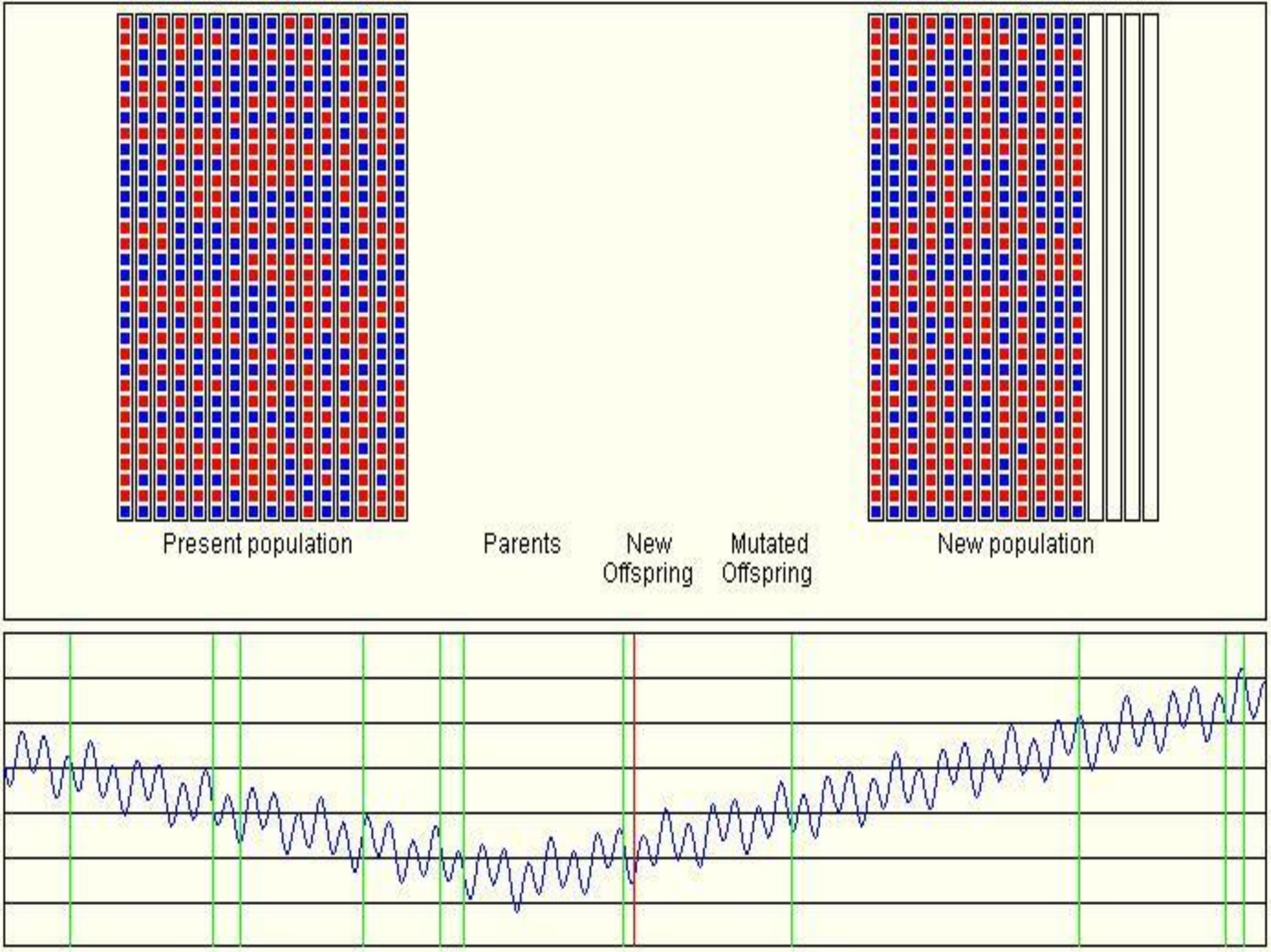


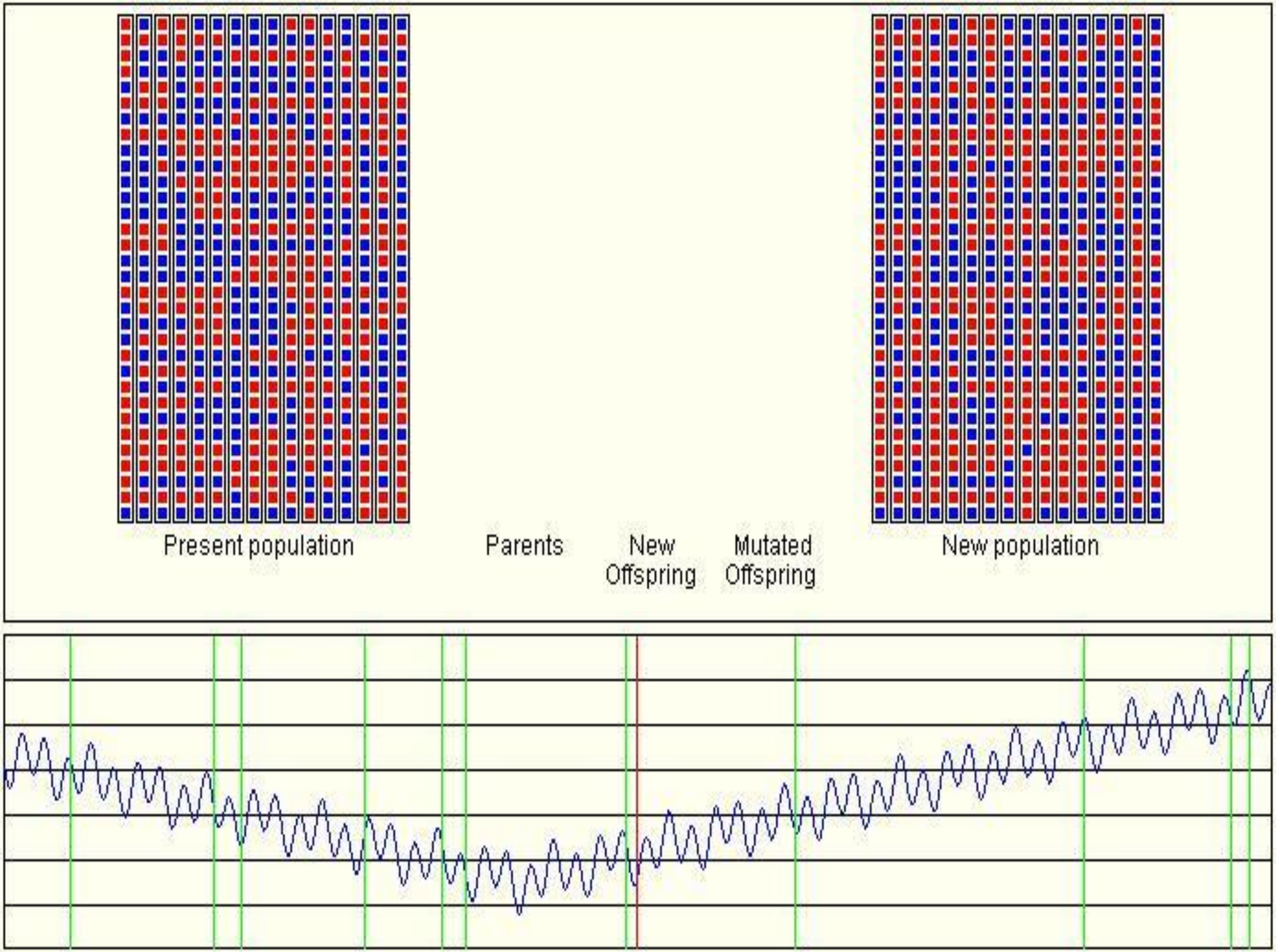
Present population
Parents
New Offspring
Mutated Offspring
New population

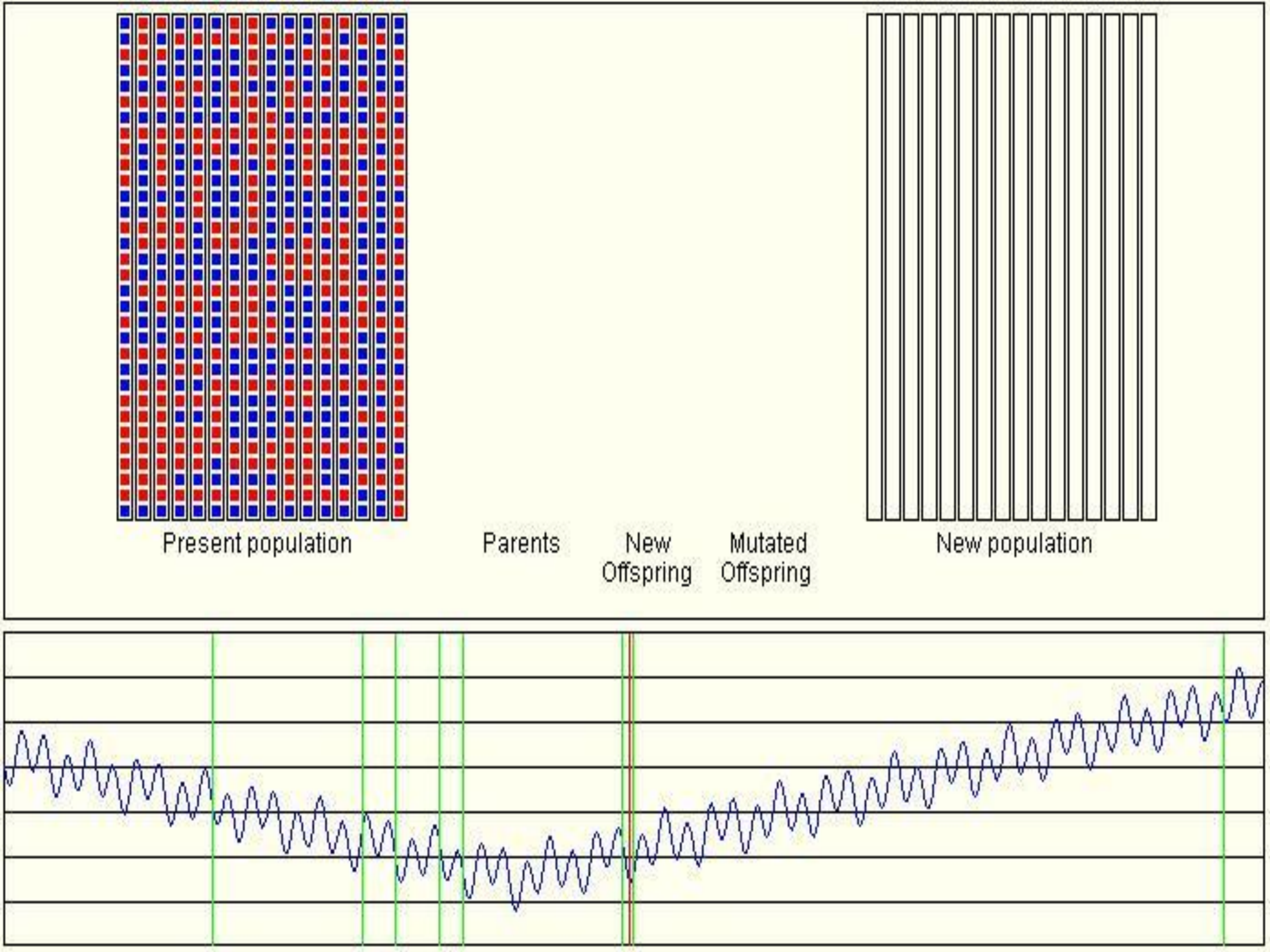

Present population
Parents
New
Offspring
Mutated
Offspring
New population

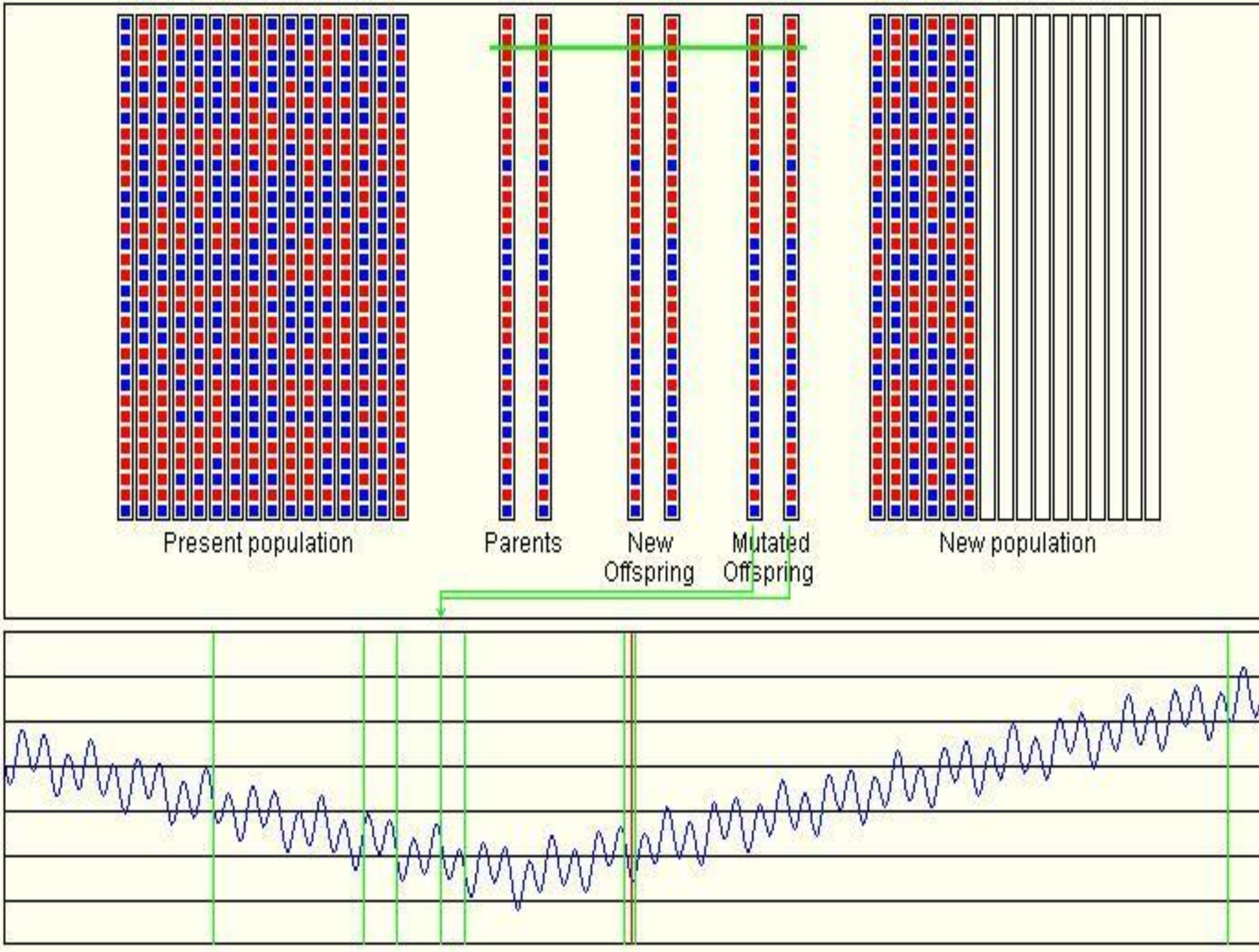

Present population
Parents
New
Offspring
Mutated
Offspring
New population

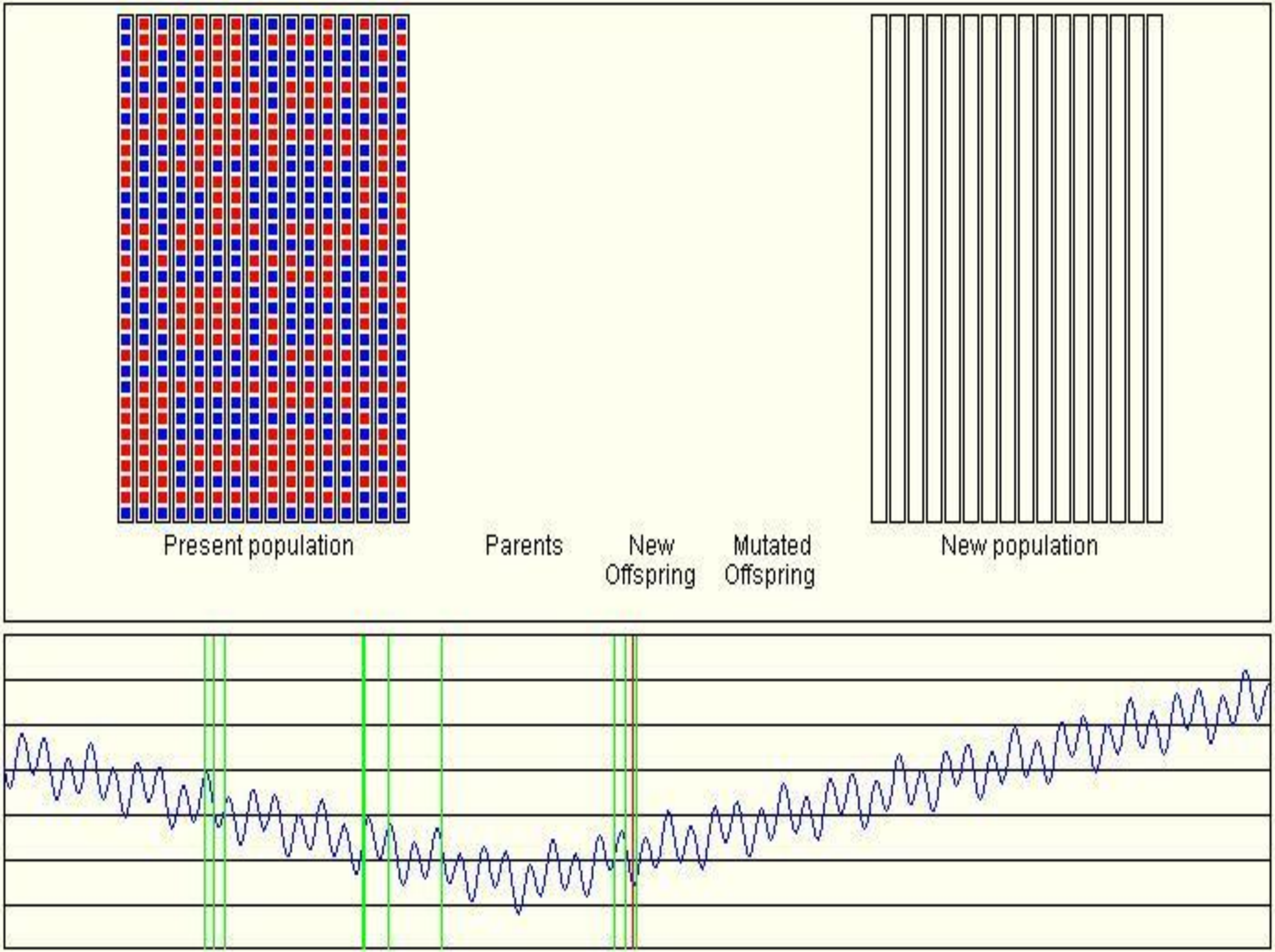

Present population
Parents
New Offspring
Mutated Offspring
New population


Present population
Parents
New Offspring
Mutated Offspring
New population


Present population
Parents
New Offspring
Mutated Offspring
New population


Present population
Parents
New
Offspring
Mutated
Offspring
New population


Present population
Parents
New
Offspring
Mutated
Offspring
New population


Present population
Parents
New Offspring
Mutated Offspring
New population

Representation of genes (chromosome)

- Identify the simplest elements (Building block) found in a solution
- The number of these elements should be fixed and not very large
- X_i will be represented with a string of gene

We call it a string of chromosome.

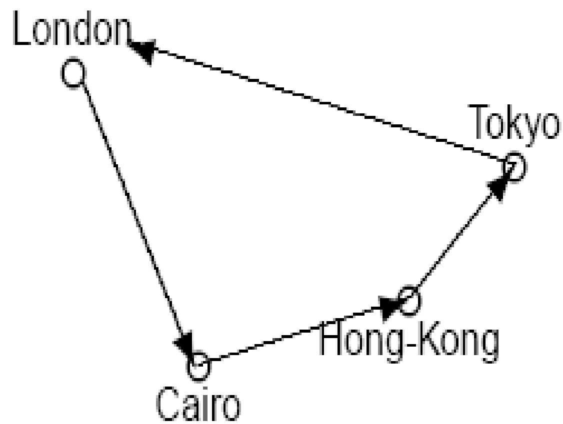
Example TSP

L London

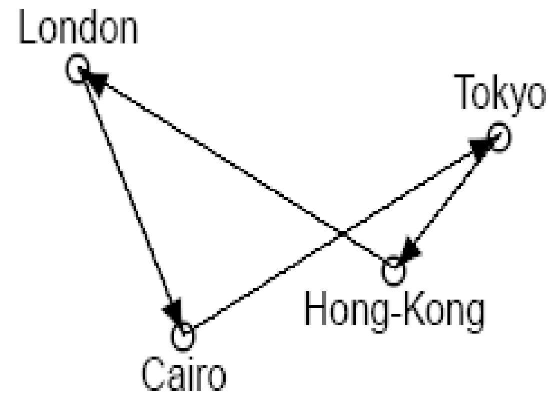
C Cairo

H Hong–Kong

T Tokyo



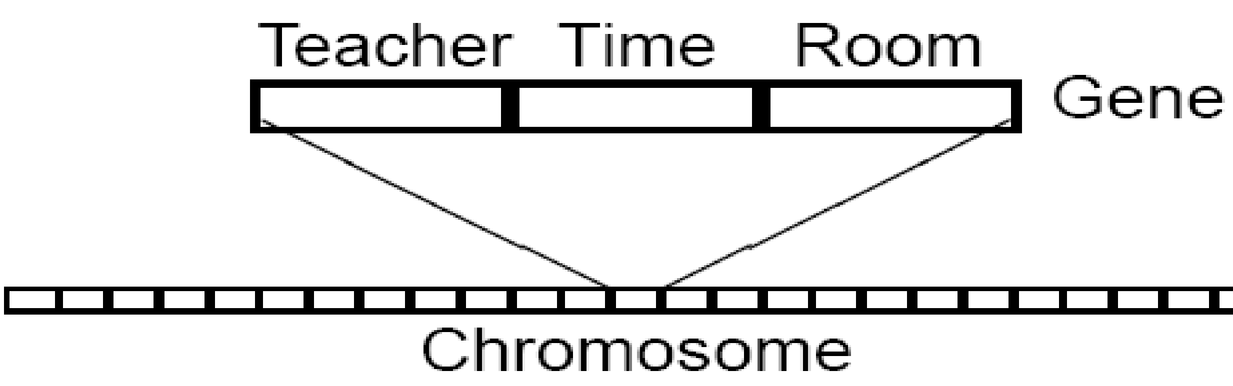
$$x_1 = LCHTL$$



$$x_2 = LCTHL$$

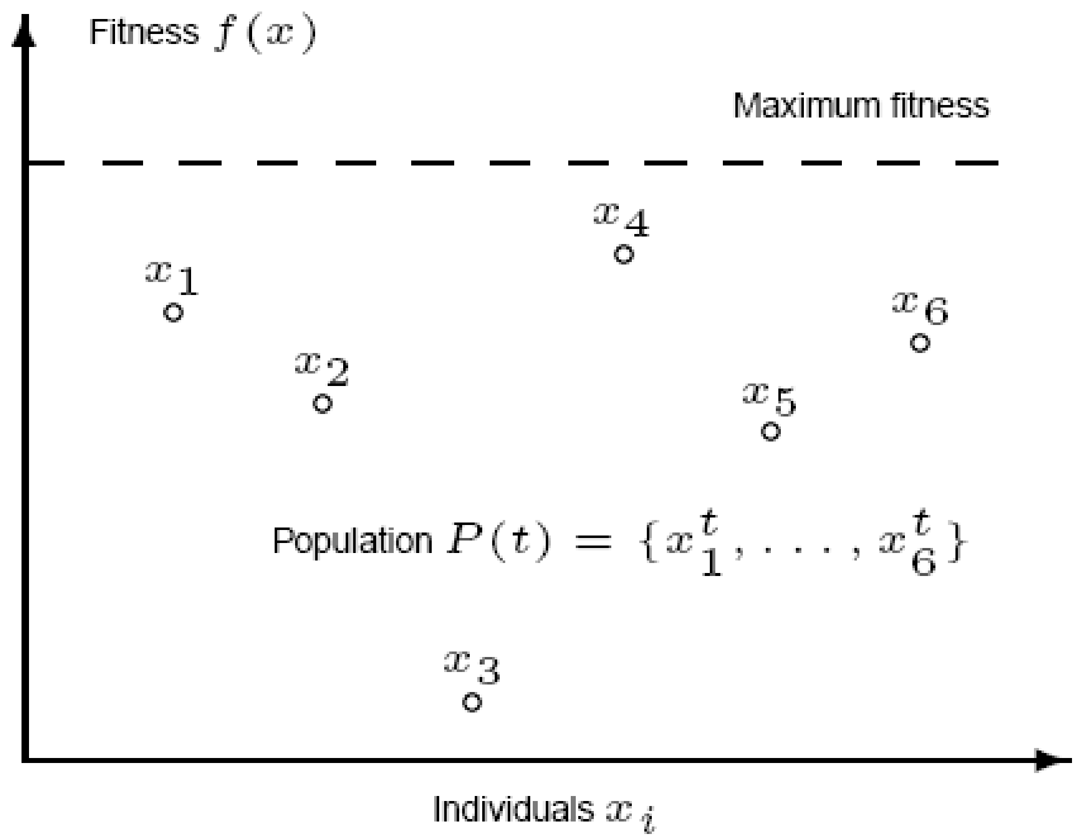
Example Time Table

Teacher	Time	Room
Prof. Newell	9:00am	Room B48
Dr. Ritter	11:00am	Room C52
Prof. Wood	1:00pm	
⋮	⋮	



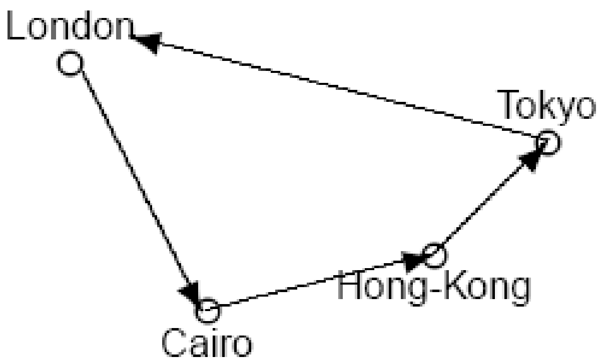
Fitness function

- A function that tell you how good the individual is.

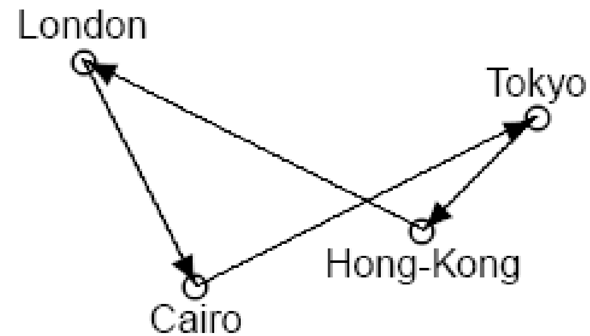


Example TSP

- X_1 better than X_2



x_1



x_2

$$\text{length}(x_1) \quad < \quad \text{length}(x_2)$$

$$f(x_1) \quad > \quad f(x_2)$$

SUMMARY OF GA

1. Choose the initial population
2. Select parent chromosomes
3. Perform crossover
4. Perform mutation
5. Evaluate fitness of the new population
6. Repeat 2 until satisfied

Pseudo-code algorithm

- Choose initial population
- Evaluate the fitness of each individual in the population
- Repeat
  - Select best-ranking individuals to reproduce
  - Breed new generation through crossover and mutation (genetic operations) and give birth to children
  - Evaluate the individual fitnesses of the children
  - Replace worst ranked part of population with offspring
- Until termination

Selection

- Select some of the population for reproduction:
  - Roulette Wheel selection: Probability selection from all population with probability proportional of their fitness
  - Ranked selection: few fittest individuals

Selection Methods

- Proportional (roulette) selection:
  - Probability of selection is proportional to the individual's fitness.

Fitness proportionate selection:

$$\Pr(h_i) = \frac{Fitness(h_i)}{\Sigma_{j=1}^{p} Fitness(h_j)}$$

- Ranking method:
  - All Individuals are sorted, and probabilities of their selection are according to their ranking rather than their fitness.
- Tournament selection:
  - Some number, e.g., 2, of individuals compete for selection
  - The competition step is repeated *popsize* times for each generation.
  - More diverse

CROSSOVER

- Replacing some genes in the parent by the corresponding genes of the other.

$$\begin{array}{rr|l} P_1 = & \mathbf{1\,0\,1} & \mathbf{0\,0\,1\,0} \\ P_2 = & 0\,1\,1 & 1\,0\,0\,1 \end{array} \Rightarrow \begin{array}{rr|l} O_1 = & \mathbf{1\,0\,1} & 1\,0\,0\,1 \\ O_2 = & 0\,1\,1 & \mathbf{0\,0\,1\,0} \end{array}$$

MUTATION

- Randomly chose gene and replace it with other gene
- Mutation help to add diversity to the population
- Help avoiding local maximum

$$O_1 = 1\,0\,\underline{1}\,1\,\underline{0}\,0\,1 \quad \Rightarrow \quad O_1 = 1\,0\,\underline{0}\,1\,\underline{1}\,0\,1$$

Why GA Work

- Why should one believe that this is going to result in an effective form of search or optimization?
  - Parallel search, since we check several solution at once
  - Fitness out bad solution from good one

GA: other applications

- Nurse Rostering with Genetic Algorithms
- Solving of the uncapacitated warehouse location problem
  - Shipping goods from warehouse to clients
  - Best solution : minimum cost (the sum of storage cost and cost of shipment)
- Volumetric Segmentation of Brain Images Using Parallel Genetic Algorithms (convert 2D pictures into 3D pictures to see abnormality in brain image)

GA: other applications (2)

- GA-based exam timetabling
- A Genetic Algorithm for Resource-Constrained Scheduling
- Using Genetic Programming To Evolve an Algorithm For Factoring Numbers
  - Modern cryptography is based on the assumption that factoring numbers takes a long time.

Research trends

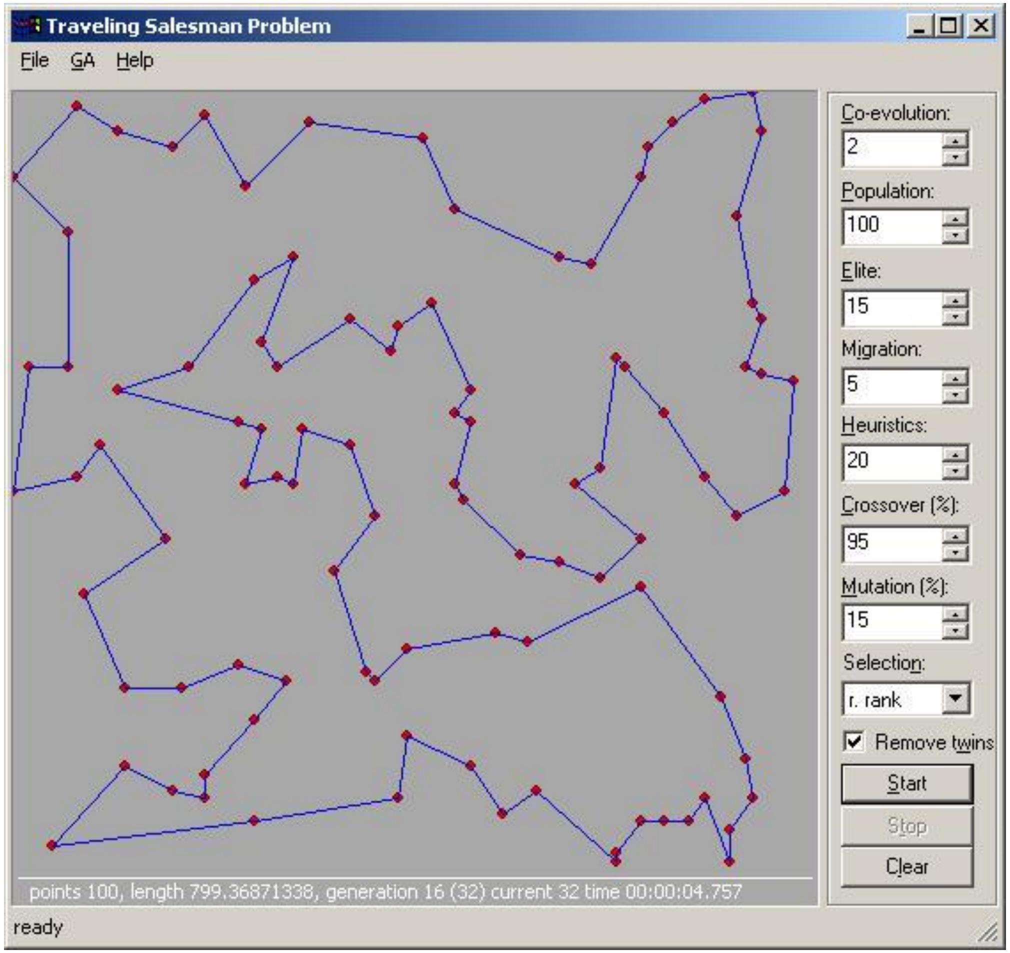
- Accelerating GA processing
  - using GRADIENT Search
  - Parallel GA
- Improve finding more optimized solutions

Other new technique

- **Ant colony optimization**
  - Advantage : the graph may change dynamically
- **Harmony search**
  - Does not need initial vector, and no hard calculations
- **swarm optimization**

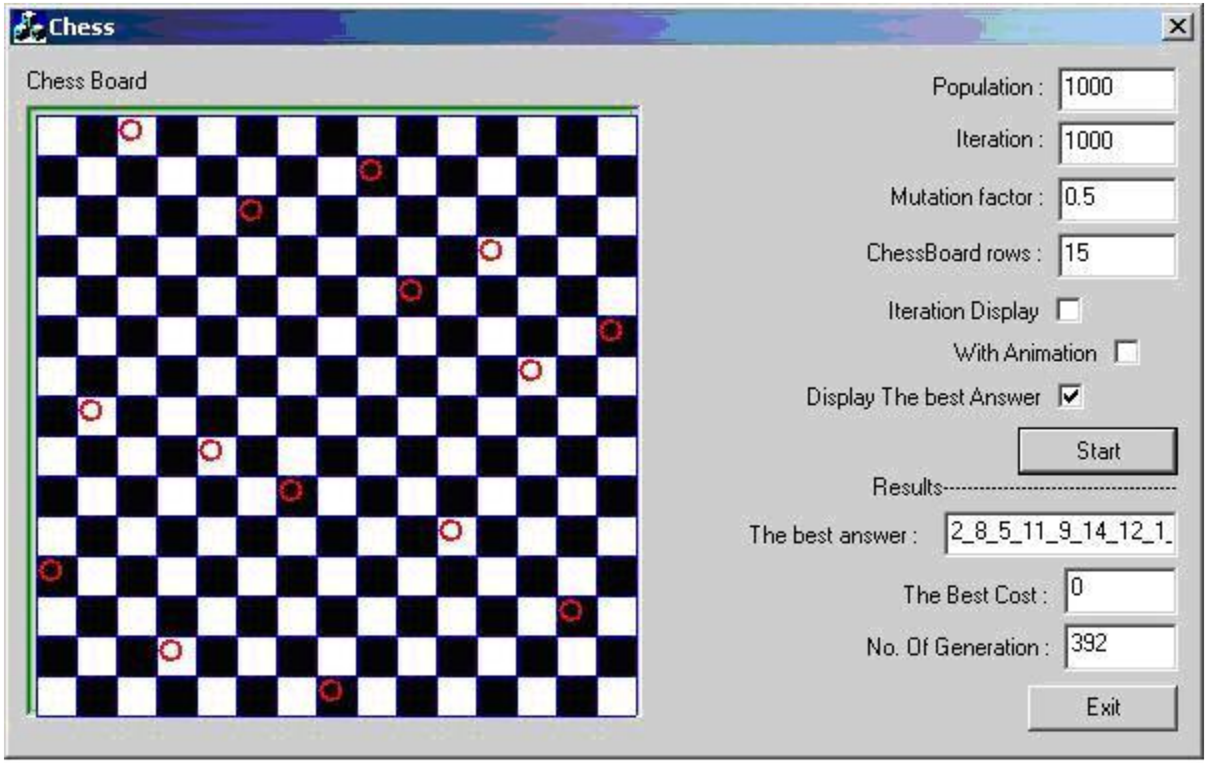
Examples

- VC++ Example on TSP



Examples

- how to place 8 queens on an ordinary chess board so that none of them can hit any other in one move



Networking & Security

- Immune System Model for Detecting Web Server Attacks using genetic algorithm

- Using Genetic Algorithm in Cryptanalysis

Cryptanalysis

- the practice of **codebreaking**
- this involves finding the secret key
  - *Cipher text-only*: the cryptanalyst has access only to a collection of cipher texts
  - *Known-plaintext*: the attacker has a set of ciphertexts to which he knows the corresponding plaintext
  - *Chosen-plaintext*: the attacker can obtain the ciphertexts (plaintexts) corresponding to an arbitrary set of plaintexts (ciphertexts) of his own choosing.

Methods of cryptanalysis

- **Brute force attack**：try all possible keys
- **Linear cryptanalysis**：finding affine approximations to the action of a cipher
- **Differential cryptanalysis**：study of how differences in an input can affect the resultant difference at the output
- **Timing attack**：analyzing the time taken to execute cryptographic algorithms
- …….