Evolutionary Computing

Biological Background

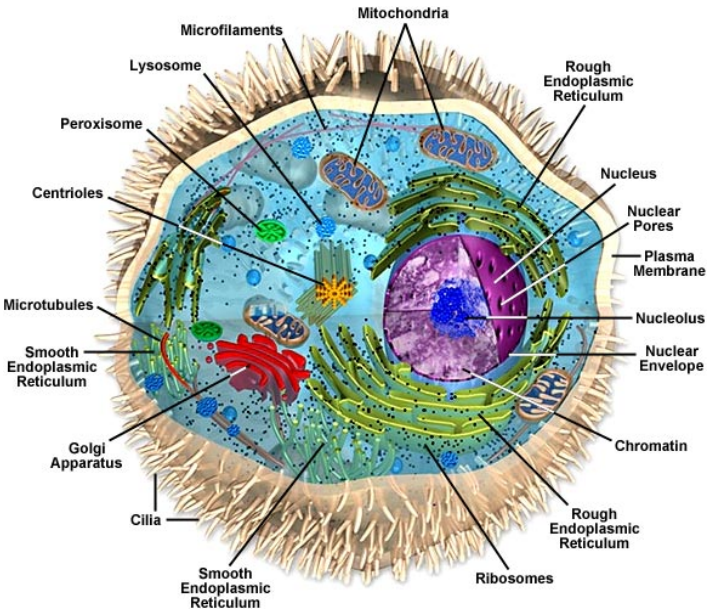
Dr. Petr Musilek

Department of Electrical and Computer Engineering
University of Alberta

Fall 2019

Cells, nuclei

- small factories working together
- the center of each cell is cell nucleus
- nucleus contains genetic information



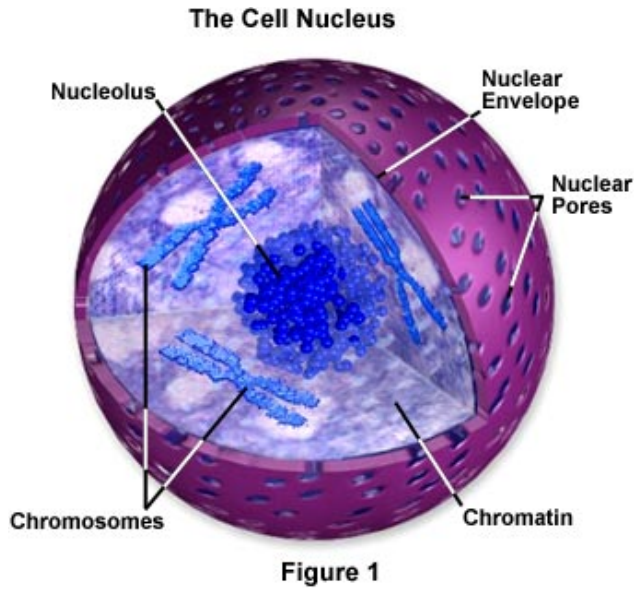


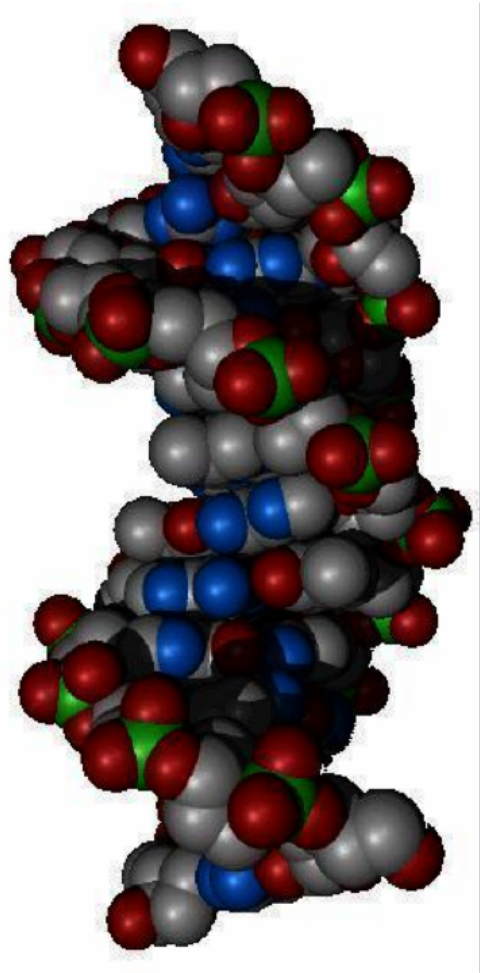
Figure 1

Chromosomes

- structures found in the nucleus of a cell, which contain the genes
- come in pairs
- store information
- built of DNA
- divided into *genes*

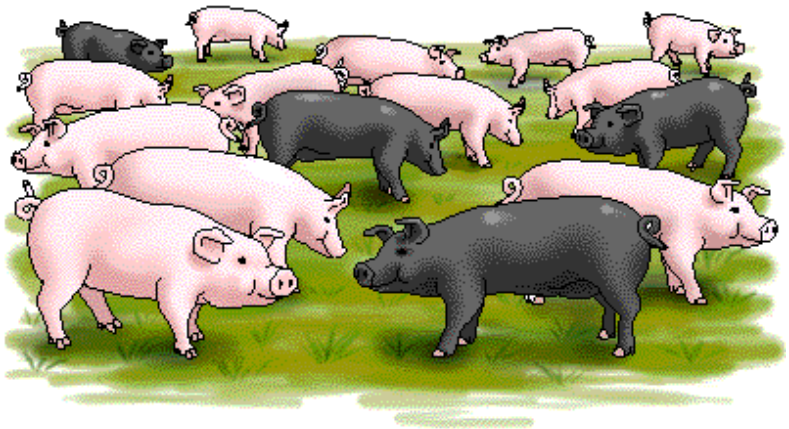
Gene

- a unit of inheritance
- a working subunit of DNA
- each of the body's 50,000 to 100,000 genes contains code for a specific product (typically, a protein such as an enzyme)



Genes - alleles

- any of the alternative forms of a gene that may occur at a given position in a chromosome (locus) are called allele
- different alleles produce variations in inherited characteristics such as eye color or blood type



Genetics

Genotype

- the entire combination of genes is called genotype, in other words it is a genetic constitution of an individual

Phenotype

- the visible properties of an organism that are produced by the interaction of the genotype and the environment

Expression

- dominant alleles are expressed from genotype to phenotype
- recessive alleles can survive in the population for many generations without being expressed

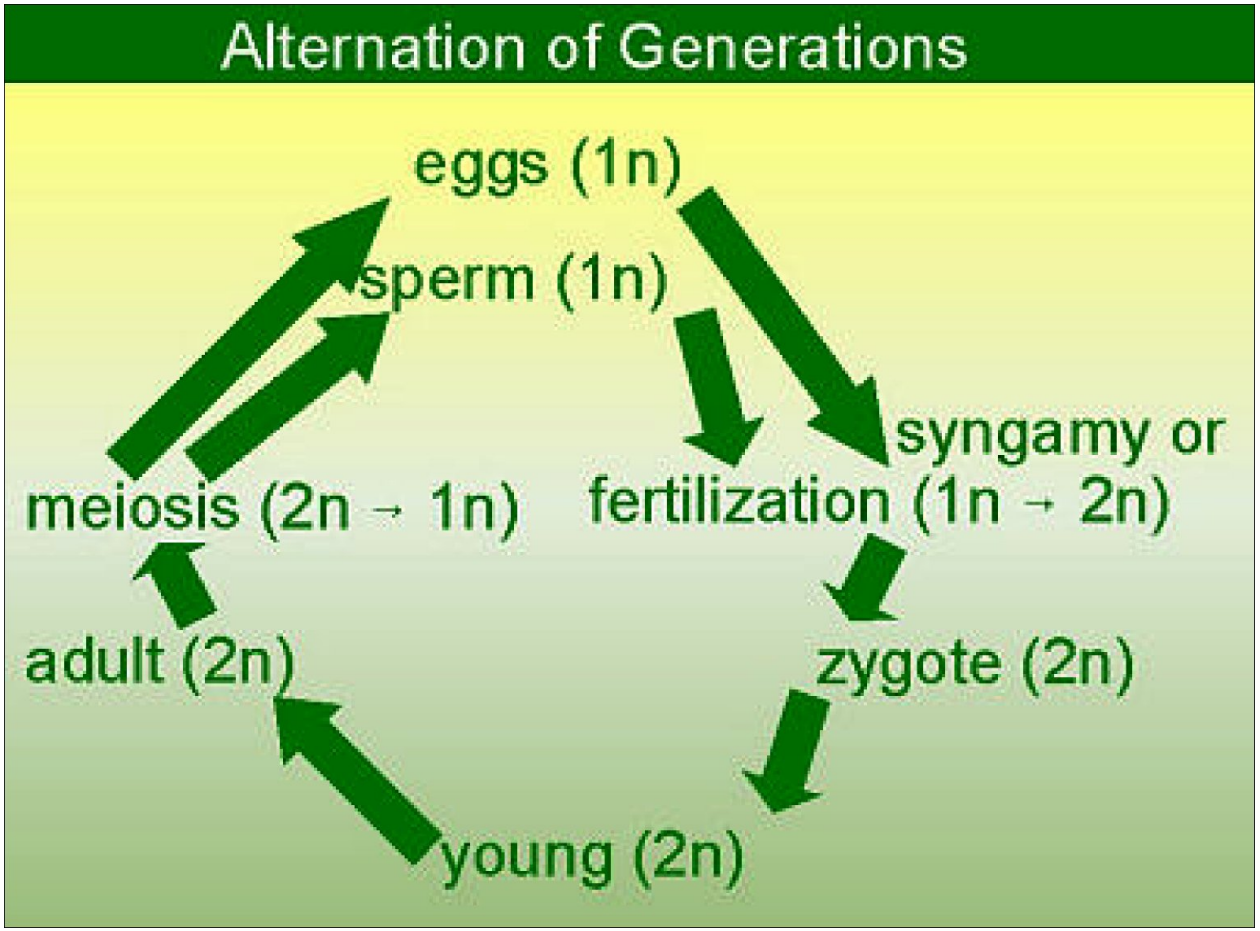
Reproduction

Mitosis: the process that takes place in the nucleus of a dividing cell, and results in the formation of two new nuclei each having the same number of chromosomes as the parent nucleus (no exchange of information)

Meiosis: the process of cell division that reduces the number of chromosomes in reproductive cells from diploid to haploid, leading to the production of reproductive cells – gametes (the nucleus divides into four nuclei each containing half the chromosome number)

Reproduction mechanism: syngamy or fertilization is the union of the egg and sperm to restore the 2n; this results in a zygote, the first cell formed by fertilization, a completely new and different organism with unique genetic information different from either parent

Reproduction cycle



Reproduction “errors”

Recombination (cross-over): a phenomenon that sometimes occurs during the formation of sex cells (meiosis); a pair of chromosomes (one from the mother and the other from the father) break and trade segments with one another

Mutation: a change in the number, arrangement, or molecular sequence of a gene

Natural selection

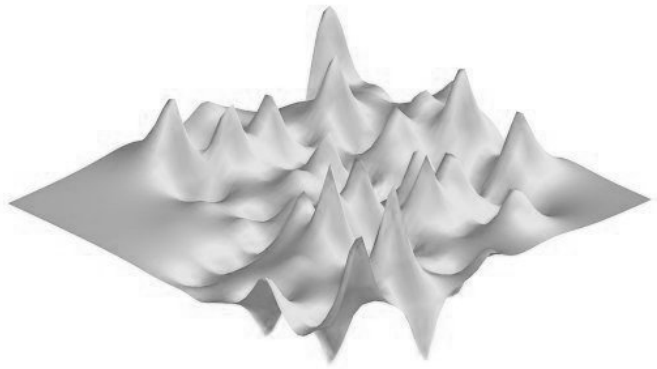
The origin of species: “preservation of favorable variations and rejection of unfavorable variations”

- there are more individuals born than can survive, so there is a continuous struggle for life
- individuals with an advantage have a greater chance for survive – survival of the fittest

Search and Optimization

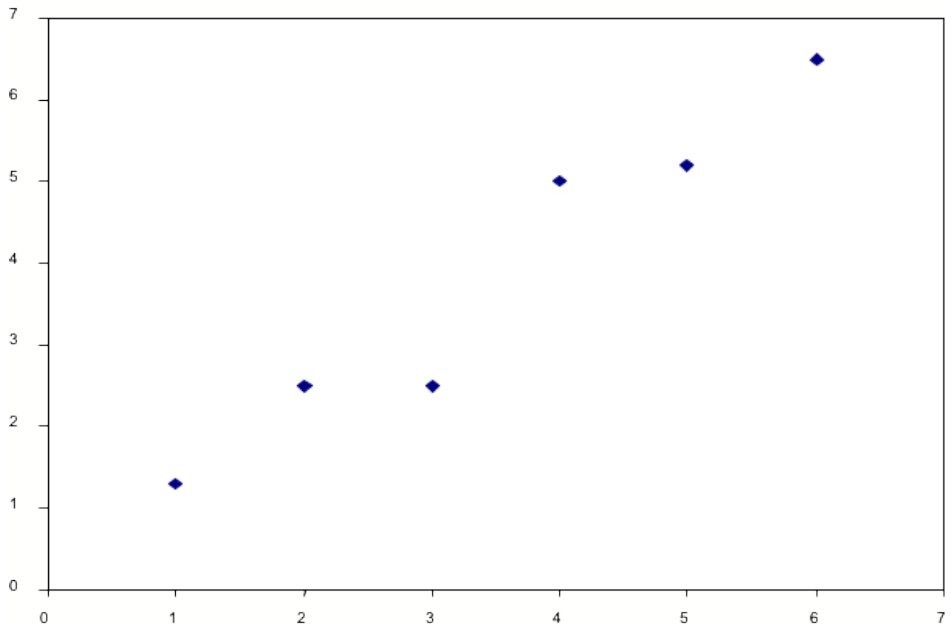
Optimization: search for the best set of values of adjustable parameters

Search space: representation of relationship between a measure of success of the search and the values of the parameters – e.g. error landscape



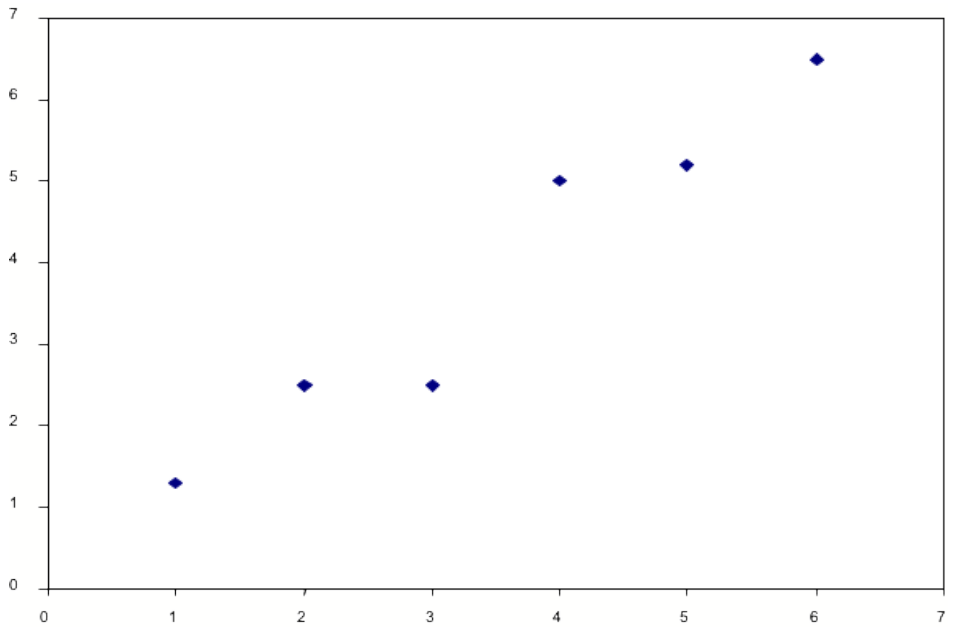
Fitting the experimental data

Graphical search: use ruler and visually estimate the best line through the points



Fitting the experimental data

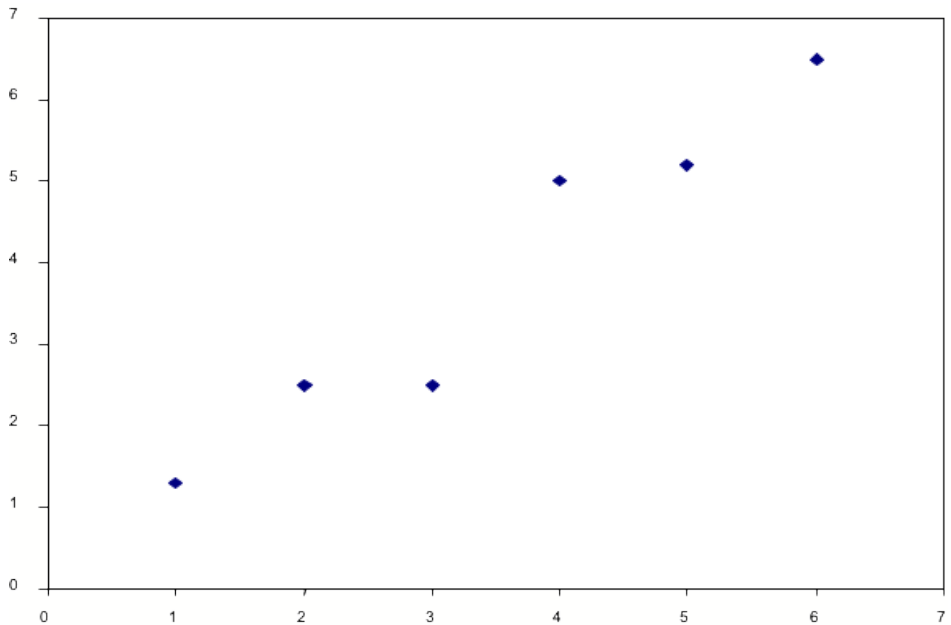
Enumerative search: use estimation (e.g. least-squares) at many points and select value of parameter(s) that yields minimal error



What if we have 10 variables and want 1% accuracy?

Fitting the experimental data

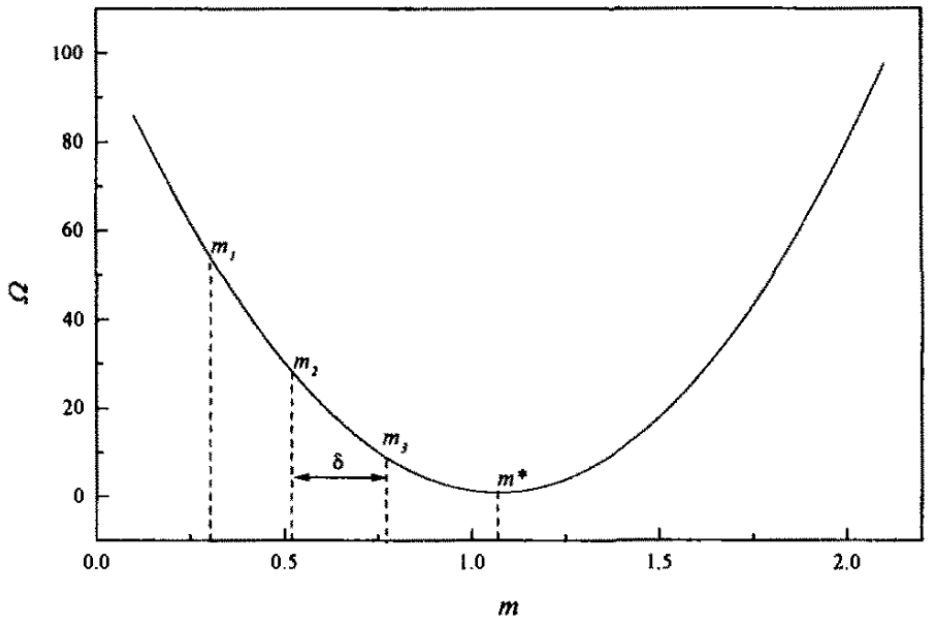
Enumerative search: use estimation (e.g. least-squares) at many points and select value of parameter(s) that yields minimal error



What if we have 10 variables and want 1% accuracy?

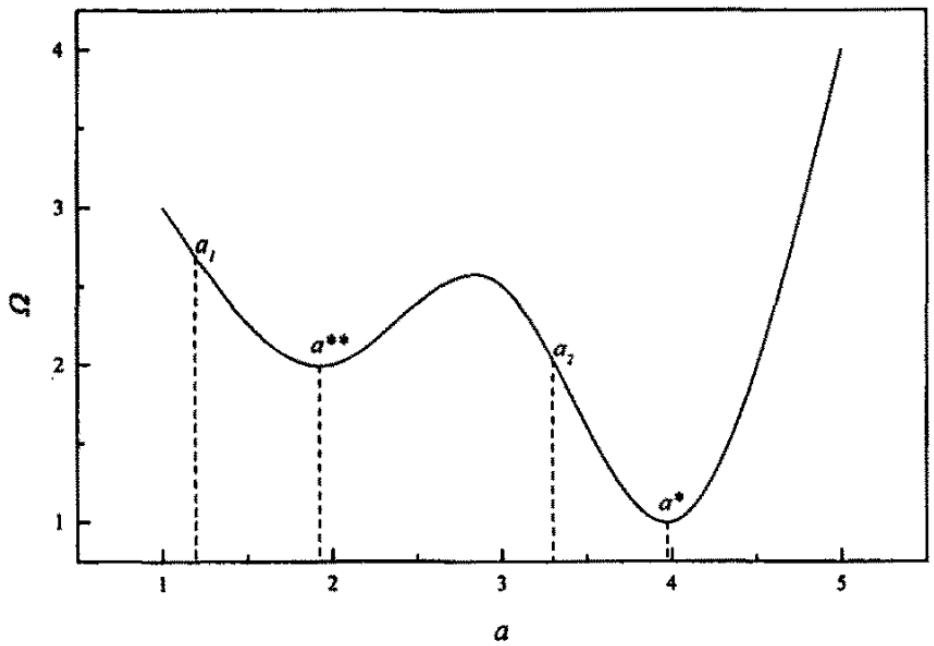
Fitting the experimental data

Direct search: estimate at two points and depending on the relation between their quality, go to third point at distance d left or right of the second/first point



Fitting the experimental data

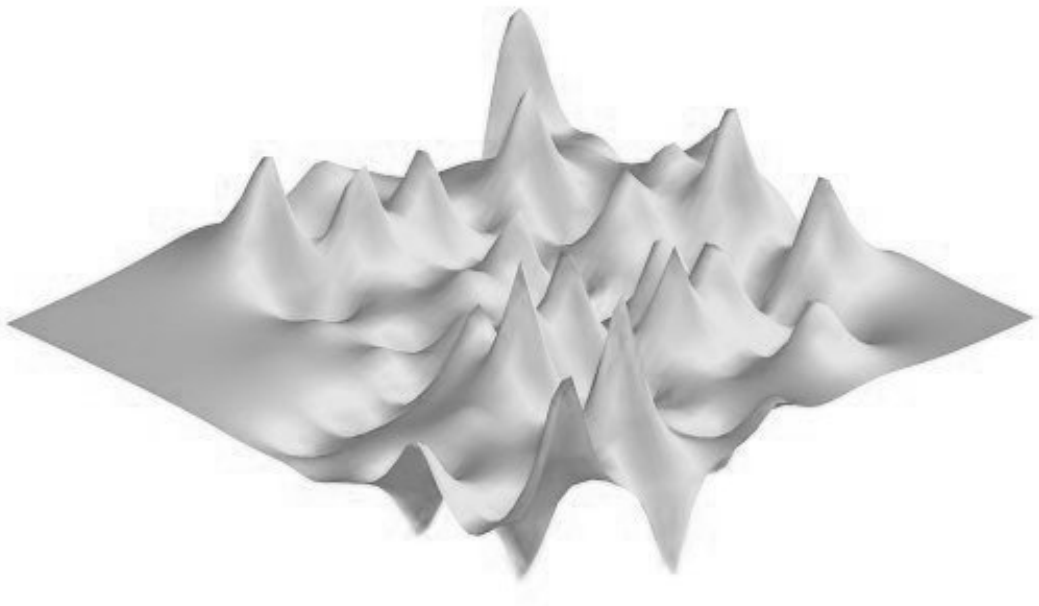
What if we have local minima? Start direct or derivative based search from multiple points and assume that the lowest result is the global minimum



Fitting the experimental data

What if we have a complex search space?

Evolutionary computing may help: it's is a form of heuristic search driven by the principles of natural evolution.



Evolutionary computing

Evolutionary Algorithms

Stochastic search algorithms based on abstraction of the processes of Darvinian evolution.

There are several types of EAs, all sharing the following basic components:

- they work with population of individuals (candidate solutions), rather than a single candidate solution at a time
- they use selection method biased by fitness (measure of quality of the candidate solutions)
- they generate new individuals via mechanisms of inheritance from existing individuals, mainly through the use of probabilistic operators of crossover and mutation.

Darwin's hypotheses (Origin of Species, 1859)

1. Earth is very old, and organisms have been changing steadily through the history of life.
2. All organisms are descendants of a single common ancestor.
3. Species multiply by splitting into daughter species; such speciation has resulted in the great diversity of life found on Earth.
4. Evolution proceeds via gradual changes in populations, not by the sudden production of individuals of dramatically different types.
5. The major agent of evolutionary change is natural selection.

Types of EAs

Evolution Strategies (ES) individuals: real value vectors; mutation and recombination are used; individuals represent also self–adaptive parameters controlling mutation distribution.

Evolutionary Programming (EP) individuals: originally FSMs; recently also real–valued vectors; do not use crossover; mutation also self-addaptive.

Genetic Algorithms (GA) individuals: binary or real–valued vectors; crossover is the main search operator, mutation minor; follow more the natural genetic mechanisms (ES and EP) are more concerned with behavioral aspects of individuals).

Genetic programming (GP) individuals: computer programs or functions; otherwise quite similar to GAs.

Pseudocode of EAs

GA and GP

```
create initial population
compute fitness of individuals
REPEAT
  select ind. based on fitness
  apply gen. operators to
    selected individuals,
    creating offspring
  compute fitness of offspring
  update the current population
UNTIL (stopping criteria)
```

ES and EP

```
create initial population
compute fitness of individuals
REPEAT
  apply gen. operators to
    individuals, creating
    offspring
  compute fitness of offspring
  select ind. based on fitness
  update the current population
UNTIL (stopping criteria)
```

Key design issues of EAs

Fitness function: should measure quality of individuals as precisely as possible; fine grained.

Individual representation: suitable for given problem (binary, integer, real, tree, ...)

Genetic operators: must be appropriate and balanced with other components of the system

Exploration vs. exploitation: appropriate tradeoff must be chosen

Selective pressure: various selection methods put different pressure on the population (how long does it take until the population is taken over by strong candidates)