Introduction to Genetic Algorithm

Debasis Samanta

Indian Institute of Technology Kharagpur

dsamanta@sit.iitkgp.ernet.in

26.02.2016

Limitations of the traditional optimization approaches

Limitations:

- Computationally expensive.
- For a discontinuous objective function, methods may fail.
- Method may not be suitable for parallel computing.
- Discrete (integer) variables are difficult to handle.
- Methods may not necessarily adaptive.

Evolutionary algorithms have been evolved to address the above mentioned limitations of solving optimization problems with traditional approaches.

Evolutionary Algorithms

The algorithms, which follow some biological and physical behaviors:

Biologic behaviors:

- Genetics and Evolution –> Genetic Algorithms (GA)
- Behavior of ant colony –> Ant Colony Optimization (ACO)
- Human nervous system –> Artificial Neural Network (ANN)

In addition to that there are some algorithms inspired by some physical behaviors:

Physical behaviors:

- Annealing process –> Simulated Annealing (SA)
- Swarming of particle –> Particle Swarming Optimization (PSO)
- Learning –> Fuzzy Logic (FL)

Genetic Algorithm

It is a subset of evolutionary algorithm:

- Ant Colony optimization
- Swarm Particle Optimization

Models biological processes:

- Genetics
- Evolution

To optimize highly complex objective functions:

- Very difficult to model mathematically
- **NP-Hard** (also called combinatorial optimization) problems (which are computationally very expensive)
- Involves large number of parameters (discrete and/or continuous)

Background of Genetic Algorithm

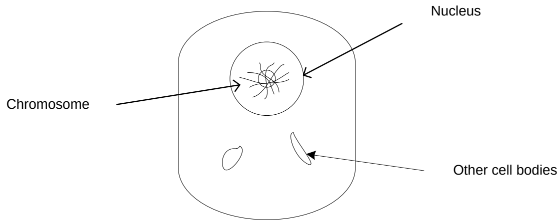
Firs time itriduced by Ptrof. John Holland (of Michigan University, USA, 1965).
But, the first article on GA was published in 1975.

Principles of GA based on two fundamental biological processes:

- **Genetics:** Gregor Johan Mendel (1865)
- **Evolution:** Charles Darwin (1875)

A brief account on genetics

The basic building blocks in living bodies are cells. Each cell carries the basic unit of heredity, called gene



- For a particular specie, number of chromosomes is fixed.

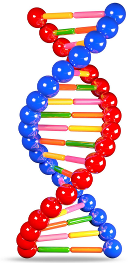
Examples

- Mosquito: 6
- Frogs: 26
- Human: 46
- Goldfish: 94

etc.

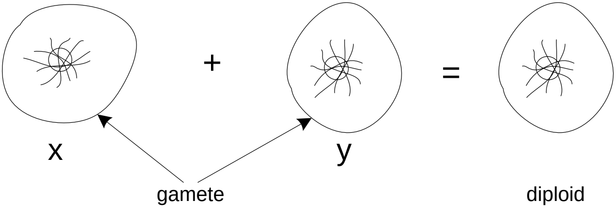
A brief account on genetics

Genetic code



- Spiral helix of protein substance is called DNA.
- For a specie, DNA code is unique, that is, vary uniquely from one to other.
- DNA code (inherits some characteristics from one generation to next generation) is used as biometric trait.

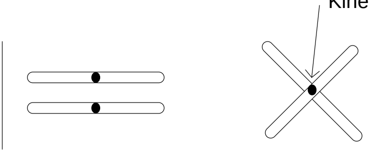
Reproduction



Organism's cell :
Cell division

haploid (Reproductive cell has half the number of chromosomes)

diploid Each chromosome from both haploids are combined to have full numbers


Information from two different organism's body cells
Kinetochore
Combined into so that diversity in information is possible
Random crossover points makes infinite diversities

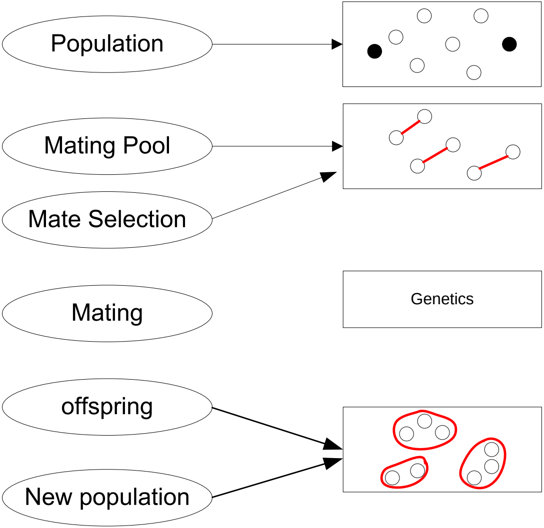
Four primary premises:

1. ***Information propagation:*** An offspring has many of its characteristics of its parents (i.e. information passes from parent to its offspring). **[Heredity]**
2. ***Population diversity:*** Variation in characteristics in the next generation. **[Diversity]**
3. ***Survival for exitence:*** Only a small percentage of the offspring produced survive to adulthood. **[Selection]**
4. ***Survival of the best:*** Offspring survived depends on their inherited characteristics. **[Ranking]**

Mutation:

To make the process forcefully dynamic when variations in population going to stable.

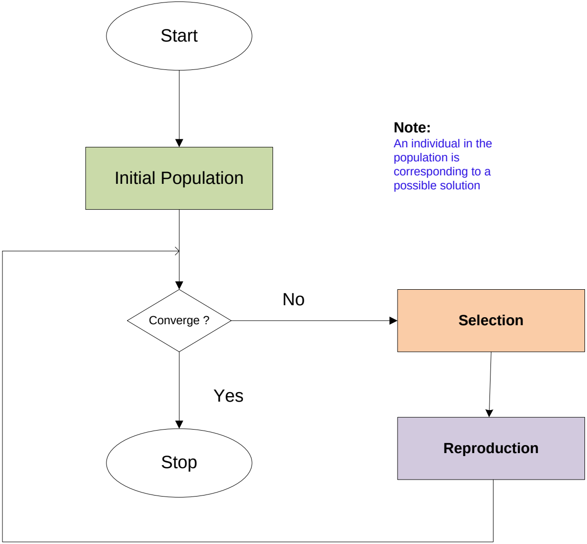
Biological process : A quick overview



Definition of GA:

Genetic algorithm is a population-based probabilistic search and optimization techniques, which works based on the mechanisms of *natural genetics* and *natural evaluation*.

Framework of GA



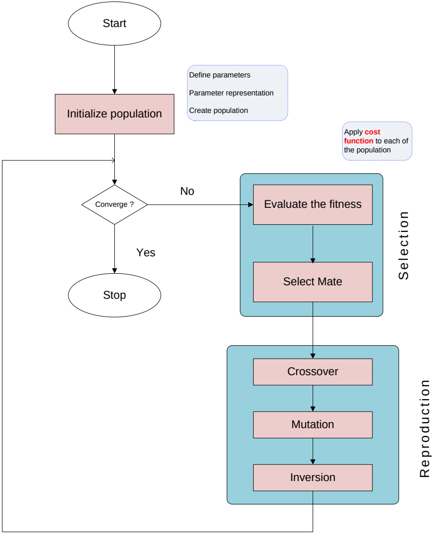
Note:
An individual in the population is corresponding to a possible solution

Working of Genetic Algorithm

Note:

1. GA is an iterative process.
2. It is a searching technique.
3. Working cycle with / without convergence.
4. Solution is not necessarily guranteed. Usually, terminated with a local optima.

Framework of GA: A detail view



Optimization problem solving with GA

For the optimization problem, identify the following:

- Objective function(s)
- Constraint(s)
- Input parameters
- Fitness evaluation (it may be algorithm or mathematical formula)
- Encoding
- Decoding

GA Operators

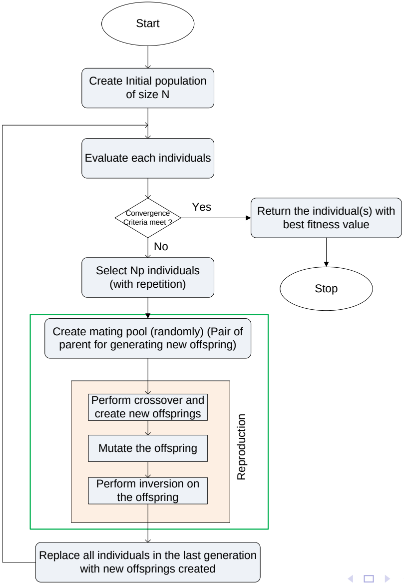
In fact, a GA implementation involved with the realization of the following operations.

1. **Encoding:** How to represent a solution to fit with GA framework.
2. **Convergence:** How to decide the termination criterion.
3. **Mating pool:** How to generate next solutions.
4. **Fitness Evaluation:** How to evaluate a solution.
5. **Crossover:** How to make the diverse set of next solutions.
6. **Mutation:** To explore other solution(s).
7. **Inversion:** To move from one optima to other.

Different GA Strategies

- Simple Genetic Algorithm (SGA)
- Steady State Genetic Algorithm (SSGA)
- Messy Genetic Algorithm (MGA)

Simple GA



Start
Create Initial population of size N
Evaluate each individuals
Convergence Criteria meet ?
Yes
Return the individual(s) with best fitness value
Stop
No
Select Np individuals (with repetition)
Create mating pool (randomly) (Pair of parent for generating new offspring)
Perform crossover and create new offsprings
Mutate the offspring
Perform inversion on the offspring
Reproduction
Replace all individuals in the last generation with new offsprings created

Important parameters involved in Simple GA

SGA Parameters

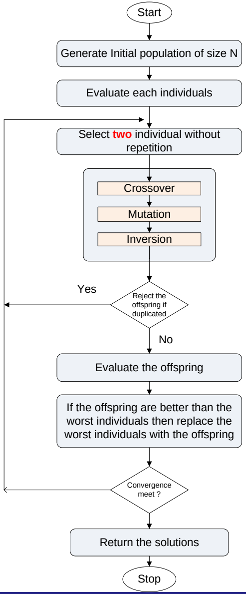
- Initial population size : N
- Size of mating pool, N_p : $N_p = p\% of N$
- Convergence threshold δ
- Mutation μ
- Inversion η
- Crossover ρ

Salient features in SGA

Simple GA features:

- Have overlapping generation (Only fraction of individuals are replaced).
- Computationally expensive.
- Good when initial population size is large.
- In general, gives better results.
- Selection is biased toward more highly fit individuals; Hence, the average fitness (of overall population) is expected to increase in succession.
- The best individual may appear in any iteration.

Steady State Genetic Algorithm (SSGA)



Salient features in Steady-state GA

SGA Features:

- Generation gap is small.
  Only two offspring are produced in one generation.
- It is applicable when
  - Population size is small
  - Chromosomes are of longer length
  - Evaluation operation is less computationally expensive (compare to duplicate checking)

Limitations in SSGA:

- There is a chance of stuck at local optima, if crossover/mutation/inversion is not strong enough to diversify the population).
- Premature convergence may result.
- It is susceptible to stagnation. Inferiors are neglected or removed and keeps making more trials for very long period of time without any gain (i.e. long period of localized search).

Any Questions??