Genetic Algorithms and Function Minimization

University of Geneva

Metaheuristics for Optimization
November 28 2022

Evolutionary Algorithms and Genetic Algorithm

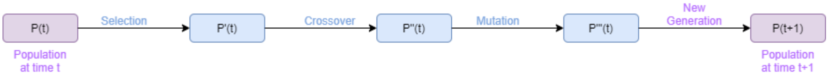
Evolutionary Algorithms:

- Family of algorithms based on Darwin's "survival of the fittest" evolution law
- Organisms adapt to more complex tasks
- Computer Model: individuals evolve through crossover during reproduction and mutation

Genetic Algorithms:

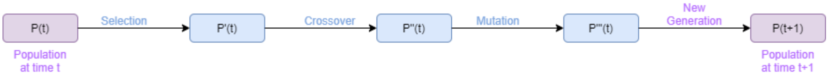
- Population Metaheuristics
- Individual = genome/chromosome
- Fitness = degree of adaptation of an individual to the environment
- Problem evolves by generations (iterations)
  - Selection of best individuals
  - Crossover + Mutation

Selection, Crossover, and Mutation



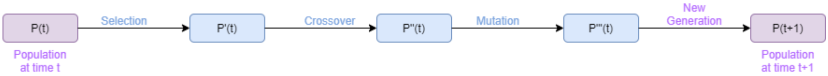
- Selection
  - Randomly draw N individuals from P(t) according to fitness values, to increase the good individuals in each generation

Selection, Crossover, and Mutation



- <u>Crossover</u>
  - P''(t) is formed by hybridization:
    Parents → Offsprings

Selection, Crossover, and Mutation



- Mutation
  - Individuals in P''(t) undergo mutation with small $p_{mutation}$
  - Errors in genome allow the emergence of new individuals/solutions
  - P'''(t) becomes P(t+1)

Selection, Crossover, and Mutation

The three GA operations: Selection, Crossover, and Mutation do not guarantee that the best individual/selection is present in the next generation/iteration.

We can perform Elitism:

- Best individual at Generation P(t) replaces worst individual at P(t+1) if fitness of best individual at P(t) is better than that of the best individual at P(t+1)

End Condition

- Max number of generations/iterations
- Fitness Stagnation

TP Objective:
Minimize the function:

$$f(x, y) = -\left|\frac{1}{2}x \sin\left(\sqrt{|x|}\right)\right| - \left|y \sin\left(30\sqrt{\left|\frac{x}{y}\right|}\right)\right|$$

where $x, y \in [10, 1000] \cap \mathbb{N}$.

Individuals and Fitness

Each individual, a solution of the optimization problem, should be represented as a binary sequence made up of two halves, representing the x and y coordinates of the individual.
Each coordinate will be made up of m=10 bits.
In order to compute the fitness, the use of a mapping function is needed:

$$map : x \mapsto \frac{x}{2^m}(b - a) + a$$

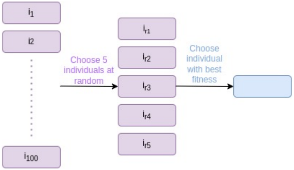
to ensure that $map(x) \in [a, b] \cap \mathbb{R}$.

Starting from the above formula, propose a mapping function to minimize $f(x, y)$, i.e., for $x, y \in [10, 1000] \cap \mathbb{N}$.

Selection, Crossover, and Mutation

Selection

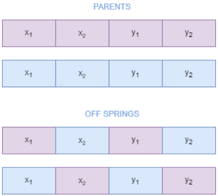
- 5 Tournament



- Repeat N times

Crossover

- One Point with Mid-Break



- $p_c = 0.6$
- Repeat $\frac{N}{2}$ times

Mutation

- $p_m = 0.01$
- $p_m = 0.1$

Set Population Size N to 100

GA Pseudo Code

```
generation = 0
Randomly initialize population
while not stop condition do
    generation ++
    Compute Individuals' Fitnesses
    Selection: 5 Tournament
    Crossover: One Point Mid-Break
    Mutation
    if no new best then
        insert best, remove worst
    end if
end while
output best
```

Work To Do

Function Visualization

Visualize the function to minimize in a 3D space to get an idea on how difficult is to find its global minimum, and localize its global minimum.

Tests

You should experiment the following:

- Selection 5 Tournament, One Point Mid-Break Crossover with p_c, Mutation with p_m=0.01
- Selection 5 Tournament, One Point Mid-Break Crossover with p_c, Mutation with p_m=0.1
- Selection 5 Tournament, no Crossover, Mutation with p_m=0.01
- Selection 5 Tournament, no Crossover, Mutation with p_m=0.1

Work To Do

Evaluations

Measure and present:

1. The cumulative empirical probability to reach the following solution qualities (optimum, relative distance to optimum $= 1.0\%$, relative distance to optimum $= 2.5\%$) over the number of evaluations.
2. Report the best, average and standard deviation of fitness among the populations for $10^3, 10^4$, and 10^5 fitness evaluations.
   The number of fitness evaluations can be estimated by the product of the population size with the number of generations.
   These statistics should be computed over several runs of the genetic algorithm (e.g., 10 runs).