Evolutionary Algorithms [Wikipedia]

An evolutionary algorithm (EA) is a subset of evolutionary computation, a generic population-based metaheuristic optimization algorithm.

An EA uses mechanisms inspired by biological evolution, such as reproduction, mutation, recombination, and selection. Candidate solutions to the optimization problem play the role of individuals in a population, and the fitness function determines the quality of the solutions (see also loss function).

Evolution of the population then takes place after the repeated application of the above operators.

Genetic Algorithm

Components

Table 1.1: Simple Genetic Algorithm

Representation	Bit strings
Recombination	1-point crossover
Mutation	Bit flip
Parent Selection	Fitness proportional
Survival Selection	Generational

Chromosome: Representation

Individuals / chromosomes can be represented as a string of values.

Binary Representation Issues

Table 1.2: Chromosome A

Variable/Gene = 512									
1	0	0	0	0	0	0	0	0	0

Table 1.3: Chromosome A

Variable/Gene = 0									
0	0	0	0	0	0	0	0	0	0

Small, one bit, change can lead to drastic fitness changes. Solution: use Gray Code.

Gray Code

The reflected binary code (RBC), also known as reflected binary (RB) or Gray Code (named after Frank Gray – Bell Labs), is an ordering of the binary numeral system such that two successive values differ in only one bit (binary digit).

Table 1.4: Gray Code

Decimal	Binary	Gray	Decimal	Binary	Gray
0	0000	0000	7	0111	0100
1	0001	0001	8	1000	1100
2	0010	0011	9	1001	1101
3	0011	0010	10	1010	1111
4	0100	0110	11	1011	1110
5	0101	0111	12	1100	1010
6	0110	0101	13	1101	1011

Hamming Distance

The Hamming distance between two equal-length strings of symbols is the number of positions at which the corresponding symbols are different.

Gray code: subsequent numbers → Hamming distance of 1

Integer Representation

- Some problems naturally map (8-queens) to integer representations
  - solution is an assignment variable = integer value
- Unrestricted: any value is permissible
- Restricted: only values from a certain set / domain
  - for example $\{0, 1, 2, 3\}$ for {North, East, South, West}
- Consider:
  - is there any relationship between values (e.g. 3 is more like 4 than 789 → ordinal relationship) or not ({North, East, South, West})
  - Choose your recombination / mutation strategy accordingly

Permutation Encoding

In permutation encoding, every chromosome is a string of numbers that represent a position in a sequence.

Permutation encoding is useful for ordering problems. For some types of crossover and mutation corrections must be made to leave the chromosome consistent (i.e. have valid sequence in it) for certain problems.

Genetic Algorithm: Other Selection Mechanisms

Elitism

Elitism (Elitist selection) is a strategy in genetic algorithms (in practice: evolutionary algorithms in general) where one or more most fit individuals (the elites) in every generation, are inserted into the next generation without undergoing any change.
This strategy usually speeds up the convergence of the algorithm.

Steady State Selection

In every generation, few chromosomes are selected (good - with high fitness) for creating a new offspring.

Then some (bad - with low fitness) chromosomes are removed and the new offspring is placed in their place.

The rest of population survives to new generation.

Fitness Proportional Selection: Issues

Fitness Proportional Selection approach has issues:

- outstanding individuals take over the entire population quickly; this is called premature convergence
- potential solutions
- when fitness values are all very close together, there is almost no selection pressure [= selection is almost uniformly random] $\rightarrow$ performance increases slowly
- Potential solutions:
  - windowing (slide and subtract some value based on history)
  - Goldberg's sigma scaling: $f'(x) = \max\left(f(x) - (f_{avg} - 2 \times \sigma_f), 0.0\right)$
  - ranking

Rank Based Selection

In rank selection, the selection probability does not depend directly on the fitness, but on the fitness rank of an individual within the population. The exact fitness values themselves do not have to be available, but only a sorting of individuals according to quality.

- Linear ranking:

$$P(R_i) = \frac{1}{n}\left(sp - (2sp-2)\frac{i-1}{n-1}\right)\ 1 \le i \le n,\ 1 \le sp \le 2 \text{ with } P(R_i) \ge 0,\ \sum_{i=1}^{n} P(R_i) = 1$$

  sp – selection pressure (the degree to which the better individuals are favored: the higher the selectrion pressure, the more the better individuals are favored) which can take values between 1.0 (no selection pressure) and 2.0 (high selectrion pressure)

- Exponential ranking

Replace worst

- In this scheme, the worst k members of the population are selected for replacement
- This can lead to rapid improvements in the average population fitness, but can also cause premature convergence
  - it will focus on most fit individual
- Typically used with large populations

Age-Based Replacement

- Rather than look at fitness of the individual, pick the oldest (in iterations) first to replace
- A FIFO queue will be needed

Other Recombination Methods

"Cut and Crossfill"

1. Pick a random crossover point
2. Cut both parents in two segments after this position
3. Copy the first segment of Parent 1 into Child 1 and the first segment of Parent 2 into Child 2
4. Scan Parent 2 from left to right and fill the second segment of Child 1 with values from Parent 2, skipping those that are already contained in it
5. Do the same for Parent 1 and Child 2

Multiparent Recombination

- The idea: recombine genes of more than 2 parents
- Some strategies:
  - Allele frequency-based: ρ-sexual voting
  - segmentation and recombination of parents-based: the diagonal crossover
  - Based on numerical operations on real-values alleles

Other Recombination Options

- Integer representation:
  - same approaches as for binary
- Floating-point representation:
  - simple recombination
  - simple arithmetic recombination
  - whole arithmetic recombination
    * Child 1: $\alpha \times x + (1 - \alpha) \times y$ and Child 2: $\alpha \times y + (1 - \alpha) \times x$

Integer representation

Same approaches as for binary

Floating-point representation:

- Simple recombination
- Simple arithmetic recombination
- Whole arithmetic recombination

$$\text{Child 1: } \alpha x$$

Other Mutation Mechanisms

Scramble mutation

A subset of genes is chosen and their values randomly shuffles/scrambled.

Inversion mutation

A subset of genes, a substring is selected and inverted.

Interchage (order changing) mutation

Randomly select two positions of the chromosome and interchange the values.

Selected Problems

8-Queens Problem

Artificial Neuron (Perceptron)

A (single-layer) perceptron is a model of a biological neuron. It is made of the following components:

- inputs x_i - numerical values representing information
- weights w_i

Genetic Programming

Traditional Programming vs . . .

Traditional Programming:

$$\text{Input Data} + \text{Program (Rules)} = \text{Output}$$

. . . :

$$\text{Input Data} + \text{Output} = \text{Program (Rules)}$$