

Genetic Algorithms Lab

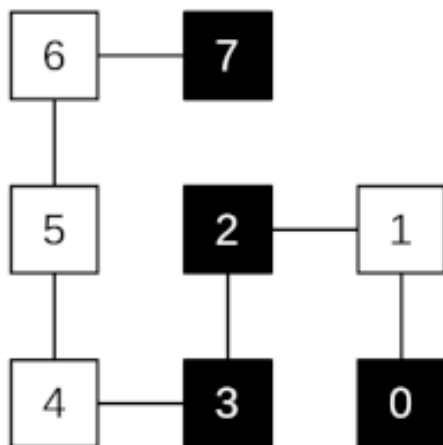
You are to implement a genetic algorithm (GA) which finds good foldings of an underlying amino acid sequence in the 2d HP model. You implement your GA with the programming language Java. Successful participation is required to be eligible to participate in the Genetic Algorithms exam. The following milestones outline, what is expected from you during the individual lab sessions.

Milestone 1

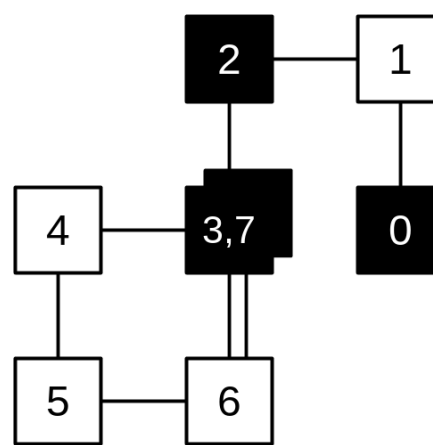
You defined a data structure to represent foldings, which in our context can be considered as solution candidates. You implemented a method („fitness function“) which takes an instance of a folding as its input and computes its fitness. The fitness is expressed as a positive floating point number and takes hydrophobic/hydrophobic (H/H) bonds, as well as overlappings into account. You debugged your implementation and computed fitness values for some sample foldings which you created.

Acceptance for Milestone 1 (due at session 1 or session 2)

Encode the folding 1 and compute its fitness. Display the fitness, as well as the number of H/H bonds and the number of overlappings. For each H/H bond and for each overlapping display the index of the involved amino acids. Now do the same for folding 2, where amino acid 3 overlaps with amino acid 7. Which values does the fitness function compute for folding 2?



Folding 1



Folding 2

Milestone 2

You implement a GA, starting with implementing fitness proportionate selection and integrating your fitness function from the first milestone. Further, you are able to visualize the phenotype of a given solution candidate. This visualization shows which amino acid is hydrophilic and which one is hydrophobic. It also shows the index of the amino acid in the underlying primary sequence. Hint: If you use Java, then use the class `BufferedImage` and save the diagram as a file in PNG format. You must also output the fitness value, the number of overlappings and the number of H/H bonds. In addition to this visualization, your GA writes a log file. This log file is structured as a simple comma or TAB separated data file, and for each generation, one line is appended. Each line contains six rows which contain the (1) generation number, (2) the average fitness of the population in this generation, (3) the fitness of the best candidate in this generation, (4) the fitness of the overall best known candidate and (5) its number of H/H bonds and (6) the number of its overlappings.

Acceptance for Milestone 2 (due at session 2 or session 3)

Start your GA with a population of 100 candidates and let it run 100 generations. Use fitness proportionate selection, but no mutation and no crossover. Visualize the results of the log file by importing it in a spreadsheet and explain the dynamics.

Milestone 3

You continue with the implementation of the mutation and crossover operators. Now you should be able to run your GA. It uses fitness proportionate selection, one point crossover, and point mutation. You can visualize the best folding and import the log file data into a spreadsheet to analyze and visualize it. Test your GA with the provided benchmark sequences and explain the results in the log files by importing and visualizing them in a spreadsheet.

Milestone 4

You implemented the tournament selection as an alternative to the fitness proportionate selection. Also, you implemented a generation dependent control of the mutation rate. For this feature you need to find appropriate criteria, under which the mutation rate is increased and under which the mutation rate is decreased. Add the current mutation rate in each generation as the seventh data point to the log file. You found a parameterization of your algorithm which provides reasonable results with the provided benchmark sequences.

Milestone 5, Final Acceptance

Success! You got the Genetic Algorithms Lab certificate. To do so, you presented your GA and showed, that your GA provides good results with the benchmark sequences. In addition, you are able to provide answers to questions related to the implementation of your GA.