AI Assignment 2. Report

First generation:

The algorithm at each step works with a generation (population) that contains 100 crosswords.
The first population contains 100 crosswords, the words in which have random parameters (coordinates and orientation). At the same time, no word can go beyond the boundaries of the crossword puzzle. This is done to simplify the code so as not to write an unnecessary check in the fitness function.
The population is stored as an array of pairs, where the first element of the pair is the value of the fitness function for the crossword puzzle, and the second is the crossword itself. This is done for the convenience of sorting everything by population by the value of the fitness function. The crossword puzzle is stored as an array of words. Each word has 3 parameters (the word itself, the coordinates of the origin, orientation).

Crossover:

In the crossover operation, I take all possible pairs of crosswords from the previous generation and get a new pair from this pair, which is formed as follows: I swap the parameters of a single word with the parameters of the same word from the second crossword with a certain chance(50%). This way I get a new pair of crosswords.
After I have received all the new pairs of crosswords, I sort them by the value of the fitness function and then select 80% of the population size, that is, the 80 best crosswords that will go into the new generation. I also perform a mutation operation on these 80 crosswords.
I take only 80 crosswords, not 100 in order to keep 20% of the best crosswords from the previous population, so I will not violate the principle of elitism.

Mutation:

The mutation operation is performed on a single crossword puzzle. I run through all the words in the crossword puzzle and, with a certain chance(25%), completely change all the parameters of the word (except, of course, the word itself).
After I have performed this operation, I recalculate the value of the fitness function.

Fitness function:

The fitness function checks whether a single crossword puzzle is suitable. The value of the fitness function is not a positive number. If the fitness function of the crossword

is 0, then the crossword is absolutely suitable, that is, it fits all the conditions that are specified in the task.
The fitness function checks the following situations in the crossword puzzle:
1) Superimposition of letters. A situation where there are two or more letters that are not equal to each other on one cell of the crossword puzzle. This situation is the simplest and most common, so I decided that it weighs the least points(1).
2) Incoherence of a crossword puzzle: a situation where a crossword puzzle has two or more separate parts. I find the number of individual parts using a depth-first search algorithm. If, as a result, I have k separate connectivity components, then I reduce the value of the fitness function by (k-1) * 30, that is, for each individual part of the crossword puzzle, I subtract 30 starting from the second.
3) Parallelism or overlapping of parallel words: a situation when there is not enough distance between parallel words or when parallel words are on the same line (column) and at the same time pile up. The weight of this error is 20.
4) No word intersection: a situation where a single word does not intersect with any other word in a crossword puzzle. This error, in fact, is checked in the second paragraph, but a separate check will not hurt. The weight of this error is 7.
5) Touching perpendicular words: a situation where two words with different orientations touch, but do not intersect. This error weighs 15.

Approximate running time of the program on different tests:

The above program on 100 tests with a different number of words (and with different words in each of them) has the following average running time:
5 words: 0.72 seconds
6 words: 5.95 seconds
7 words: 106.83 seconds
8 words: 258.99 seconds