INFO 6205, Spring 2019

Algorithm & Data Structure

Final Project

**Space Mining**

Section 5 Group 518

Tu, Siyu NUID 001424985

Chen, Yiyin NUID 001408372

Zheng, Yue NUID 001715577

Table of Contents

[1. Introduction 3](#_Toc6490504)

[2. Method 3](#_Toc6490505)

[2.1 Genotype 3](#_Toc6490506)

[2.2 Phenotype 3](#_Toc6490508)

[2.3 Fitness 3](#_Toc6490510)

[2.4 Selection 3](#_Toc6490512)

[2.5 Crossover 4](#_Toc6490515)

[2.6 Mutation 4](#_Toc6490517)

[3. Result and Discussion 4](#_Toc6490520)

[4. Conclusion 6](#_Toc6490525)

# Introduction

In the future, the galaxy is filled with high-value mines. Once the humans found out, they sent their spaceships out, hoping to find the biggest mine in the galaxy based on the signal strength they received. We approach the problem with a genetic algorithm and provide a solution to determining where the most valuable/biggest mine is. We modelled the galaxy as a cube with a dimension of 100-unit width, 100-unit length and 100-unit height. In it, the mines are randomly deployed, and the spaceships are assumed to warp into the galaxy at random locations. After generations and generations by selection based on signal fitness, finally the target mine is located by the humans.

# Method

### Genotype

### In the project, the representation of a spaceship’s genotype is a randomly generated thirty-digit binary number. It is in the form of

### Phenotype

Phenotype is rendered by first dividing the genotype into three ten-digit binary numbers. And then each of these binary numbers are converted to a decimal number whose value is within the range of 0 to 100. They represent the coordinates of the spaceship. Taking the example above, the phenotype of “001100101010101010101010101101” will now become

### Fitness

The fitness is defined to be a function of both the radiation of the mine, a constant representing the reserve of it, and the distance between the ship and the mine, as follows

where D is the distance and R is the radiation. A larger radiation and shorter distance will render a small value, which we define a better fitness. So when there are, for instance, ten mines in the galaxy, a ship should receive 10 distinct signals, from which the ship will choose the smallest value to go on with the mine-hunting.

### Selection

The selection process can be described as taking out two ships, choosing one with the better fitness and repeating the step until every ship has been screened.

The Tournament Selection is used for selection process because 1, it has lower complexity which is O(n); 2, it guarantees the optimal selection in each iteration and preserves the diversity of the gene pool at the same time, since in every iteration half of the ships that are less likely to find the target are eliminated and same amount of the ships are being bred by the selected ones in the meanwhile. So the total number of the ships maintains unchanged.

### Crossover

Crossover is where two arbitrary ships’ genes randomly mix up and bear a new gene (a new ship). Each digit in the genotype from both parent ships has a 50 percent chance to be inherited.

### Mutation

After so many selection and crossover, the ships’ genes become similar, and “marriage in proximity of blood” starts to take place. As a result, the genes stop evolving so the ships cluster to a local biggest mine and stop finding the possibly bigger ones. We call this local optimal solution. To prevent such a solution, the mutation is introduced.

However the chance of mutation is rare. When it happens to a ship, five percent of the gene digits will be reversed – if the digit is 0, it will become 1 after mutation.

### Parallelism

The galaxy is divided into eight sub-space, the spaceships into eight sub-fleet, and the mines into eight groups. Then after the sub-fleets find the target in the sub-spaces, we sort the eight sub-targets and return the biggest one.

We designed the algorithm to be such that there is a lower chance of getting a local optimal solution, thus increasing the chance of finding the real target. Also, because the eight sub-space is independent of one another, an eight-threaded parallelism is performed.

# Result and Discussion

We passed a few test cases where we change the number of both the mines and spaceships. We used different set of numbers to run the tests. We found that with more mines distributed across the galaxy, it becomes harder to find the target because it tends to find the local optimal solution. If we define finding a mine whose rank is at the top 10% of all the mines in the galaxy. Then in case 1, there are 8/10 qualified. Similarly, there are 7/10 in case 2 and 5/10 in case 3. The results are shown below.

1, With 10 spaceships and 40 mines:

2, With 20 spaceships and 80 mines:

2, With 20 spaceships and 80 mines:

# Conclusion

The project provides a solution to a genetic algorithm problem, space mining, where it tries to find the biggest mine in the galaxy. We found that the algorithm’s effectiveness starts to decrease when more mines are introduced even though the mine-ship ratio is maintained.

The work could be improved on increasing the realism in future endeavour by taking into account the fact that signal strength is possibly weakened by obstacle, which could be a planet, between a ship and mine. It would be the third factor in addition to D and R, where the fitness would be less relative to the value or R.