**Genetic algorithm in repeating prisoner dilemma**

Wenhao Li

Introduction:

In 1980, Dr. Axelrod R. analyzed the most effective strategies for repeating prisoner dilemma using computer programs. Unlike ordinary prisoner dilemma where betrayal is the better strategy, in repeating prisoner dilemma, corporation is preferred since your opponent will remember your last move. Among the 14 strategies that participants in the tournament, one strategy out-shine all the others, and it was named as ‘tit for tat strategy (Axelrod 1980). It is a simple strategy that decides its move solely on what its opponent has chosen on the last turn. For example, it will betray its opponent if its opponent has chosen to betray itself last turn. ‘Tit for TAT’ is such a simple strategy that it can be described with just a few lines in a computer program.

Dr. John Holland is another pioneer in the evolution of strategy in a natural or artificial system. In his book published in 1975, he described a process of how the strategy could evolve in an environment in which the winner and loser can be quantified with a process called genetic algorithms (John 1975). Genetic Algorithm is an alternative approach to artificial intelligence. Unlike the deep mind, the genetic algorithm achieves the same goal by simulating the process of natural selection. He demonstrated the algorithm through a simple program which I like to call it ‘pac-man’ program. And in his demonstration, he had shown that the best possible strategy generated by the genetic algorithm is a complex strategy that could not be simplified and the only way to describe such a strategy is to list all the possible moves one by one.

Chart, scatter chart

Description automatically generated

*Figure 1: A picture that demonstrates the ‘pac-man’ program designed by Dr. John Holland. The rule of the game that John Holland used to analyze genetic algorithms is like this: there is a space with 10 \* 10 columns, surrounded by walls; 50 ‘seed’ is spread randomly in this space and a ‘player’ will be born at a random column; the sight of this player can only reach to the 4 adjacent columns and it can only move to one of them at a time; there are 7 possible moves for the ‘player’: up, down, left, right, don’t move, move randomly, and eat ‘seed’. And the score for each player is calculated as follows: +10 points for each ‘seed’ eaten; -1 point if ‘player’ chose to eat and there is no ‘seed’ in the column; -5 for hitting the wall.*

So here we run into a contradiction: where the winner of Axelrod’s experiment is a simple and elegant strategy, but John’s experiment gives the opposite result. Therefore, I wish to propose an experiment that runs Axelrod’s experiment with genetic algorithms to find out if there is a strategy that could be better than ‘tit for tat’.

Hypothesis:

The genetic algorithm can produce a complex strategy that could behave better than ‘tit for tat’ in Axelrod Experiment.

Program Development:

* The equation for decision making

In order to develop a genetic algorithm for Axelrod’s experiment, we need an equation that could simulate the decision-making process. I have made two assumptions for the decision-making process:

1. Decision is solely based on three things: background knowledge + opponent’s previous moves + random event
2. These three things are independent of each other

The resulting equation is as follows:

Equation 1: Probability equation for the decision-making process, variables colored in red are coefficients for the three components of the equation: from left to right are background knowledge, the opponent’s previous move, and random event. Variable an represent the opponent's previous moves, which equals -1 when the opponent betrayed, equals 1 when the opponent corporate; and n represents the turn number, n = 1 for the last turn and n = 2 for the second last turn. The coefficient of the random event will be \* a random integer between -10 to 10. When the decision is positive, it will return corporation, and negative for returning betrayal. Because the competition will last for 200 turns, there will be 199 coefficients for opponent’s previous moves.

* Axelrod’s tournament:

The fitness value of each individual in one generation is decided by its score in Axelrod’s experiment. There are 11 preset strategies that acted as opponents for each individual strategy in genetic algorithms. The rule of competition is as follows: +5 points for betrayal when your opponent corporate; +3 points for corporate when your opponent corporate; +1 points for betrayal when your opponent betrayed; 0 points for corporate when your opponent betrayed. Each competition between two strategies will last 200 rounds, and the total score of 200 turns will be calculated. The fitness value of each individual will be the average score of 11 competitions between one strategy from the genetic algorithm and 11 preset strategies.

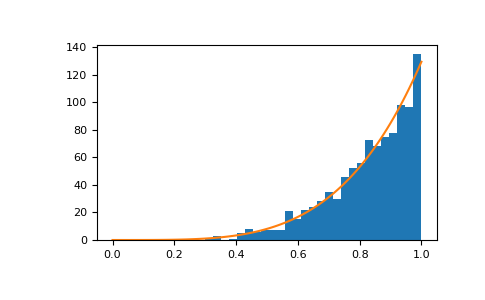
The resulting fitness score for each individual will be used in the process of natural selection and the generation of offspring.

* Genetic algorithms

Diagram

Description automatically generated

*Figure 2: A basic flow chart for genetic algorithm. Each individual will be a list of 201 characters long, which represents each coefficient in Equation 1, and each character is randomly generated. Then the whole generation will go through a competitive process to assign a fitness value as described in subsection Axelrod’s tournament. Then the generation goes through the natural selection and offspring generation process. The individual that has reached a higher fitness score will have more offspring. The resulting generation with 200 offspring will go through the same process as the starting colony, and the process will be repeated for 50 rounds.*



*Figure 3: numpy-power distribution. This is the random number generation used for generate offspring. As shown in graph, most of the numbers lied between 0.5 to 1, which means the probability of reaching higher score is much higher then that of the lower score. The probability distribution fit well to the natural selection process where individual with higher fitness score will have more offspring.*

Result Analysis:

The highest score generated by genetic algorithms is 709 and the score for ‘tit for tat’ strategy is 492. It supports the hypothesis that the genetic algorithm can produce a complex strategy that could behave better than ‘tit for tat’ in Axelrod Experiment.

The maximum score for each generation is not improving. It is likely due to a limited number of variables in the genetic algorithm and the limited number of strategies involved. Both factors contribute to an environment that is not complex enough, which result in the first 200 randomly generated individual have already had some that have reached the highest possible score.

The average score of each generation does improve each round, as it has shown in Figure 4. The rate of improving have greatly slow down after ~25 generations (Figure 4), which is similar to Dr. John Holland experiment where the majority of progress was made in the first half of all generations (John 1975).

*Figure 4: the average fitness score for all individuals in each generation. The x-axis is the generation number, and the y-axis is the average score. The average score is improving, and the improvement slows down after ~25 rounds.*

The strategies that reach the highest score share some common features (Table 1):

The variable of background knowledge is relatively high

Random variables always weigh less than background knowledge

Both of these features show that this strategy is more likely to corporate than betray. However, there is no significant trend was found in the variable for the opponent’s previous moves, likely due to the same reason which resulted in the maximum score for each generation not improving.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Background knowledge | 100 | 93 | 84 | 95 | 56 | 94 | 88 | 90 |
| Random event | -9 | -5 | 4 | 6 | -2 | -3 | -6 | -7 |

*Table 1: list of some strategies generated by genetic algorithms that have reached the highest score, the first row is the coefficient for background knowledge and the second row is the coefficient for the random event. All of the variables for background knowledge are relatively high since it is set to be an integer between -100 to 100. Random variables always weigh less than background knowledge, even when the random value in equation 1 reach the highest value possible, it still cannot flip the decision to the other end.*

Limitations and further experiments:

The equation for the decision-making process is based on two assumptions:

1. Decision is solely based on three things: background knowledge + opponent’s previous moves + random event
2. These three things are independent of each other

However, both assumptions are flawed and can be improved in further experiments. First of all, the decision is not only based on the 3 factors listed above, such as words from families, friends, rumors, and so on. Secondly, these factors will interfere with each other in some ways, such as having a bad day (a random event) will degrade the coefficient for background knowledge or the opponent’s previous moves.

There are also limitations to Axelrod’s tournament that generate the fitness value for each individual. There are only 11 preset strategies defined, which has resulted in an environment that is not complex enough to make a distinction between some of the coefficients. And the competition is a fixed set of 200 rounds, further experiments can investigate the behavior of genetic algorithms when the number of turns is not fixed.

Genetic algorithms used in this experiment are also flawed. It does not count in the process of duplication and deletion: a very important genetic process that duplicates or deletes a whole gene. Duplication is important because it produces a new set of genes that is open for all kinds of manipulations. Further experiments could take these processes into account.

This experiment had always been on my mind since I heard about the two experiments run by Dr. Axelrod and Dr. John Holland. I have finally designed a program that could test my thought. This experiment will greatly improve my confidence in my studies in the cs program. However, there are still many aspects of the experiment that could be improved. I wish further courses will give me more insight into prefect this experiment.

References:

1. Axelrod, R. (1980). Effective Choice in the Prisoner’s Dilemma. Journal of Conflict Resolution, 24(1), 3–25.
2. John Holland. Adaptation in Natural and Artificial Systems (1975, MIT Press)