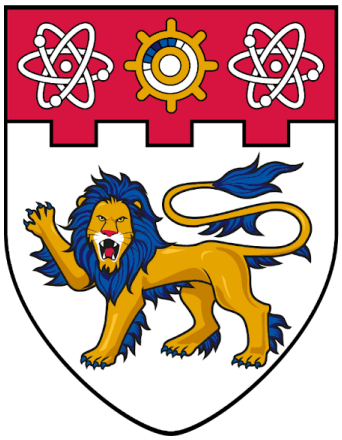

NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

CZ4046: Intelligent Agents

Name	Nur Dilah Binte Zaini
Matriculation Number	U2022478K
Coursework	Assignment 2

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
NANYANG TECHNOLOGICAL UNIVERSITY

Table of contents

Introduction

The main goal of the assignment is to implement a strategy for an agent in a 3-player repeated prisoner's dilemma. The 3 players will play against one another repeatedly in a match. The match consists of 100 rounds. The average of the payoffs from each round will be the match's score. The previous plays of the opponents will affect the action of Player 1.

The table representing the payoffs for Player 1 in a 3-player game is shown as follows:

C = Cooperate; D = Defect

	Payoffs
DCC	8
CCC	6
DDC	5
CDC	3
DDD	2
CDD	0

Figure 1: Payoffs for Player 1

According to Figure 1, the payoff will be the highest for player 1 if he defects while the rest of the opponents cooperate. However, the drawback of playing defect from the first round is opponents will not trust player 1 and decide to defect afterward. This leads to mutual defection. However, the payoff of mutual cooperation is higher than the payoff of mutual defection. The player should maximize cooperation while keeping an idea of the next actions of the opponents.

Strategies

This section explains the multiple strategies a player or the opponents can take as an action at each stage of the match.

Tit-For-Tat (T4T) Strategy

The strategy follows the simple principle of ‘an eye for an eye’ and ‘a tooth for a tooth’. The player cooperates at the beginning of the match. If one of the opponents decides to defect in the previous game, the player will choose to defect for the current game. The player choose cooperates as the default action and mimics the opponents' actions for the next rounds of the match.

Trigger Strategy

In this strategy, the player starts the match by cooperating as its first action. If his opponents choose to defect for consecutive rounds, the player will set the defect as his move for all the other rounds. The Trigger strategy forms a Nash equilibrium both with itself and with the T4T strategy.

Tolerant Strategy

The player keeps track of the opponents' history and calculates the total number of cooperations and defections they took. If the total number of cooperations is higher than the total number of defections, the player will cooperate.

Other Strategies

Other strategies are strategies that the player can choose that are not affected by the previous plays of the opponents:

Nasty Strategy

The player defects throughout the match regardless of the history of the opponent's plays.

Nice Strategy

The player cooperates throughout the match regardless of the history of the opponent's plays. This strategy may seem too naive as it has been proven that the best strategy for the player is to defect in Figure 1. The first initial thought if the player adopts a Nice strategy were to go against the opponents that adopt a Nasty strategy, the player using the Nice strategy will lose.

```
NastyPlayer scored 2.0 points, NastyPlayer scored 2.0 points, and NastyPlayer scored 2.0 points.
NastyPlayer scored 5.0 points, NastyPlayer scored 5.0 points, and NicePlayer scored 0.0 points.
NastyPlayer scored 8.0 points, NicePlayer scored 3.0 points, and NicePlayer scored 3.0 points.
NicePlayer scored 6.0 points, NicePlayer scored 6.0 points, and NicePlayer scored 6.0 points.

Tournament Results
NastyPlayer: 24.0 points.
NicePlayer: 24.0 points.
```

Figure 2: Nice Player Versus Nasty Player

Figure 2 shows that a Nasty player technically won every match against a Nice player but both players have the same total score. The effects of using either the Nice strategy or the Nasty strategy are equally the same. Therefore, this proves that the Nasty strategy does not outperform the Nice strategy.

Freaky Strategy

The player decides whether to cooperate or defect for the entire tournament at the beginning of the match and will stick with the decision made for the rest of the rounds in the match.

Random Strategy

The player chooses a random action regardless of the history of the opponent's plays. There is a 50% chance of choosing defect or cooperate. Hence, the player will cooperate half of the time and defect the other half.

Customized Strategy

The opponents may choose either of the strategies mentioned previously. Hence, I decided to implement a customized strategy that is able to give the best action the player can take in a match. However, the strategy the player decides to use will be competing against other players implemented by other students. It will not be able to cover all possible strategies that can be designed in an infinite number of ways. Therefore, the motive of the customized strategy is to counter the strategies mentioned in the Strategies section of the report.

The customized strategy follows a set of steps:

Step 1) Cooperate at the start of the match

For the first round, I decided that the player should be cooperative. The payoff of mutual cooperation is higher than the payoff of mutual defection. Even though there may be a chance that opponents may choose to defect in the first round, it is a better solution for the player to play ‘nice’ as compared to mutual defection.

Step 2) Compute the percentage of opponent history’s cooperation

In a 3-player repeated prisoner’s dilemma, opponent histories are known to the players. The player would know if the opponent defect or cooperate in the previous round of the match. Taking this knowledge to the player’s advantage, I am able to compute the percentage of the opponent’s history of cooperation. The player will be able to gauge when the opponents will be willing to cooperate and achieve mutual cooperation. The threshold for cooperation will be 95. If the percentage of both opponents’ cooperation is more than 95%, the player will cooperate.

This step is designed to counter T4T Strategy and Trigger Strategy. If the opponent adopts a Trigger Strategy, the player is able to adapt and choose to defect once the opponent decides to defect as the cooperation percentage will not be high. For T4T Strategy, it succeeds if it is put against other opponents that do not cooperate. The player does not tend to always cooperate and will defect when it needs to.

```
for(int i = 0; i < oppHistory1.length;i++){
    if (oppHistory1[i] == 0) {
        oppCoop1 = oppCoop1 + 1;
    }

    if (oppHistory2[i] == 0) {
        oppCoop2 = oppCoop2 + 1;
    }
}

float oppCoop1Percentage = (oppCoop1 / oppHistory1.length) * 100;
float oppCoop2Percentage = (oppCoop2 / oppHistory2.length) * 100;

if(oppCoop1Percentage >= 95){
    if(oppCoop2Percentage >= 95) {
        return 0;
    }

}
```

Figure 3: Code Snippet of Step 2

Step 3) Defect for the last few rounds of the match

Even though the player tries to cooperate for the majority of the time, the dominant strategy for the player is still defective.

	Description
DCC > CCC	The player's best action is defect when both opponents cooperate
DDD > CDD	The player's best action is defect when both opponents defect
DCD > CCD	The player's best action is defect when 1 opponent cooperates while the other opponent defects

Figure 3: Providing examples to show that the player's dominant strategy is defect

In the tournament, the number of rounds is randomly selected from 90 as the minimum to 110 as the maximum. To maximize the player's payoffs, the player is implemented to defect after the 90th round.

Step 4) By default, the player will defect

If the conditions of the repeated prisoner's dilemma did not meet any of the steps mentioned above, the player will defect. The reason behind it follows the same reason Step 3 has provided. The dominant strategy for the player is defect hence this will provide maximum utility when the player is unsure of the opponent's actions.

Player Implementation

The player is created based on the steps mentioned shown in the code snippet.

```
if (n == 0)  {
    return 0;
} else if (n > 90) {
    return 1;
} else {
    for(int i = 0; i < oppHistory1.length;i++){
        if (oppHistory1[i] == 0) {
            oc1 = oc1 + 1;
        }

        if (oppHistory2[i] == 0) {
            oc2 = oc2 + 1;
        }
    }
    // Calculate the percentage of cooperation by getting the total number of cooperations the
    // opponent did over the moves the opponent has already done in the history * 100
    oc1Percentage = (oc1 / oppHistory1.length) * 100;
    oc2Percentage = (oc2 / oppHistory2.length) * 100;
    // If the opponents has been cooperating, the player will cooperate
    if(oc1Percentage >= 95){
        if(oc2Percentage >= 95) {
            return 0;
        }
    }
}
// By Default it will defect
return 1;
```

Figure 4: Code Snippet of player's implementation following the 4 steps

Evaluation

The ranking of the players in the tournament varies after every run of the program. Hence, I decided to calculate the percentage of wins for each player for different numbers of runs for the tournament to have a proper representation of how each player did.

The percentage of wins is calculated: (total number of wins for a player / total number of runs for the tournament) * 100

Figure 5: Player performance rankings based on the percentage of wins

The player Zaini_NurDilah_Player dominates the rankings in all cases. Based on the results above, it shows that the player does not always win in every tournament. T4T player and Tolerant player have their fair share of wins in the tournaments. There are a couple of instances, in which the positions of T4T player, Tolerant player, and Zaini_NurDilah_Player were swapped among themselves. However, most of the time, Zaini_NurDilah_Player will take first place in the rankings. Other than the 2 competitors, the Nice player had small wins (eg. when the tournament rounds are equaled to 50, Nice Player achieved 2.0% of wins) while Nasty Player did not achieve any wins. This shows that it is important to cooperate and not to always defect. Due to the calculation of the percentage of opponent history's cooperation, Zaini_NurDilah_Player will always beat the Random player and Freaky player. Other than that, our player will also always beat the Nice player and Nasty player as it knows when it should cooperate or defect. The Zaini_NurDilah_Player seems like the best choice as it can beat all the mentioned strategies majority of the time,

Conclusion

In conclusion, the player is designed to counter strategies like Tolerant player and T4T player. Taking into account other strategies for the implementation of the player, the player is able to outperform the strategies most of the time.

The player will play nice in the first round as default. Among the agents participating in the tournament, the player will not be the first to defect. Afterward, the player's actions depend on the percentage of opponent histories cooperation. By doing so, the player will be able to deviate its action to defect if the percentage of the opponents' histories cooperation is not high. If the opponents play nice, the player will choose to cooperate all the time till it reaches the 90th round which the player will decide to defect to maximize its own payoffs. Even though the player can outperform all these strategies, there are still infinite strategies that the design for the player did not take into account.