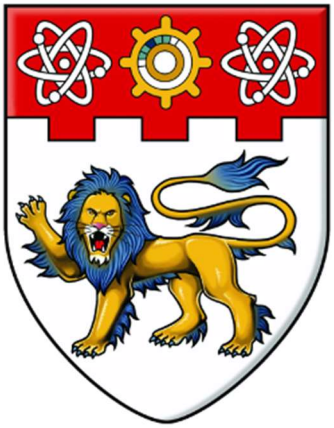

NANYANG TECHNOLOGICAL UNIVERSITY

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

CZ4046 – Intelligent Agents

ASSIGNMENT 2: 3 Prisoners' Dilemma

Name : Wu Rongxi
Matriculation Number: U2020719J

Table of Contents

Introduction

This report will be on the game play of a repeated version of Prisoners Dilemma, challenges and characteristics of the game will be discussed as well as a brief analysis of each type of given player. Then, we will discuss about how we created a player in attempt the win in the tournament of the Repeated Prisoners Dilemma.

In the Repeated Prisoners Dilemma, when a given game (often thought of in normal form) is played multiple times (possibly infinitely many times) by the same set of players. We compute the (average) reward of a player in a repeated game, to be (rj is the player's payoff in round j) lim k→∞ X k j=1 rj/k

The Score System

Actions	Payoffs
C C C	6 6 6
C C D	3 3 8
C D C	3 8 3
C D D	0 5 5
D C C	8 3 3
D C D	5 0 5
D D C	5 5 0
D D D	2 2 2

Fig 1: payoff matrix [1]

Social Welfare

	$\sum_{i \in Ag} u_i(\omega)$	Player 3			
		Defect		Cooperate	
		Player 2		Player 2	
		Defect	Cooperate	Defect	Cooperate
Player 1	Defect	6	10	10	14
	Cooperate	10	14	14	18

Fig 2: Social welfare matrix

Three prisoner's Dilemma

To defect…

Defecting may seem like the best strategy if we assume the worst for all our opponents that they are self-interest agents, however if all players all to defect, the social welfare will be low(Fig 2), and we do not want that.

...or we can Cooperate!

To maximize social welfare (highest being 18), it is best for all players to cooperate however, as the number of the players in the game of prisoners' dilemma increases, it becomes more difficult for

cooperation to emerge. This is mainly because a player is unable to clearly distinguish between who the defectors and who the cooperators are.

In the 2-player iterative prisoner's dilemma game, a player can easily reciprocate against a defector in a one-to-one situation, and therefore discourage defection. However, in the N-player iterative prisoner's dilemma, retaliation against a defector means punishment to everyone else in the game, including the cooperators.[1]

Analysis of Given Players

NicePlayer

The NicePlayer will always choose to cooperate with other players, regardless of what the others will choose.

- The NicePlayer will only win if it is playing against two other NicePlayer/TolerantPlayer/T4TPlayer, obtaining the payoff of 6 for every round.
- When NicePlayer plays against any player who chooses to defect during the tournament, the payoff obtained by it will be greatly compromised.

NastyPlayer

The NastyPlayer will always choose to defect, regardless of what the other players will choose.

- The NastyPlayer when played with other players will be able to achieve Nash equilibrium, neither player can obtain higher score than the NastyPlayer by changing their action when they are played against the NastyPlayers.
- The NastyPlayer will always get a payoff of either 2 or 5 or 8
- The NastyPlayer will never get a payoff of 0

RandomPlayer

The RandomPlayer defects and cooperates randomly for every game in the round.

- A Wild card, affecting the results at the end of the tournament
- Luck based

FreakyPlayer

The FreakyPlayer will randomly choose to defect or cooperate at the start of the match and use the same strategy throughout that match.

- A wild card, affecting the results at the end of the tournament
- Luck based

T4TPlayer(Tic for Tac)

The T4TPlayer will cooperate in the first game and then proceed to randomly pick one of the opponent's previous actions for the current game.

- The T4TPlayer will cooperate if both its opponent always to cooperate and defect if both opponents choose to defect
- The T4TPlayer is a forgiving strategy, unless there are more than one T4TPlayer in that match, once the third player defect, the T4Tplayer may both choose to defect in the next round, or both continues to cooperate because of the random element of picking a random opponent to follow.
- T4TPlayer will also act as wild card just like RandomPlayer
- **The T4TPlayer relies on the assumption that the opponent is trying to maximize their score** as well, however if it is put into a match with a mindless player like the RandomPlayer, the T4TPlayer will act like a RandomPlayer

TolerantPlayer

The TolerantPlayer is one of the most rational players amongst the 6 players, it will cooperate in the first round and then see if majority of its opponents' actions are to defect or cooperate and select the corresponding action based on that.

- The TolerantPlayer is a revengeful player, it will look at what other players have done previously, and act based on it
- The TolerantPlayer will act based on what the majority does.
- However, when playing with another Tolerant or NicePlayer, the Tolerant will always cooperate since the majority of actions of both opponents are to cooperate. Giving the NastyPlayers an edge to significantly higher payoff.

Interesting Players

After doing some research, I found some interesting type of players used in prisoners' dilemma apart from the given players. One example will be the Grim Trigger Player, a extremely unforgiving player.

Grim Trigger Player

The Grim Trigger Player cooperates in the first round and in the subsequent rounds as long as his opponent does not defect from the agreement. Once the player finds that the opponent has betrayed in the previous game, he will then defect forever.

Testing of Strategies for optimal results

After a quick run of the initial program given, I recognised that the player that topped the scoreboard, being not surprisingly the most rational player: TolerantPlayer (see Fig 3)

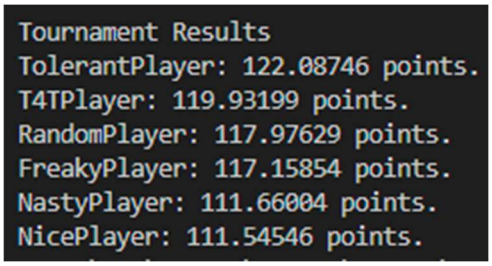

Fig 3

Being someone who really wants to win this tournament, I did the first thing anyone who wants to win will do, analyze the results of all the players, take the best player and try to improve its strategy to beat itself, I improvise on it such that it will win against both Nasty and TolerantPlayer. And the first version of my player, RX was created (is named after the first two initials of my given name).

Here is a brief explanation of my initial player:

1. Instead of simply calculating the total number of defects and cooperates of our opponents, RX will calculate the percentage of defects and cooperates up till the current round by each player including itself.
2. RX will always cooperate in the first round.
3. RX will attempt to identify if there is NicePlayer or NastyPlayer in the match by finding out if the percentage of cooperate by either or both of the players is 0% or 100%.
4. RX would cooperate if it suspected that there is one NicePlayer and the other is NastyPlayer
5. RX would defect if it suspected that both are NicePlayers.
6. If the percentage of opponents cooperating is not 0 or 100%, RX will suspect they are likely to be a TolerantPlayer. In order to maximize pay off, RX will attempt to trick the TolerantPlayer by always ensuring the percentage of cooperate for RX and the other opponent is above 50%.
7. RX will defect the moment the combined percentage of itself and one of the other opponents is above 50%
8. RX will defect if both opponents cooperate more than 85% of the time, it could be both opponents are TolerantPlayers
9. RX will cooperate if both opponents cooperate more than 65% of the time but less than 85% of the time, in order to ensure the TolerantPlayer will always choose to cooperate.
10. If RX suspect to be playing with two NastyPlayers, such that the rate of defects is more than 40% of the time, it will choose to defect as well to ensure a non-zero payoff.

As anticipated, the results were pretty satisfying(see Fig 4).

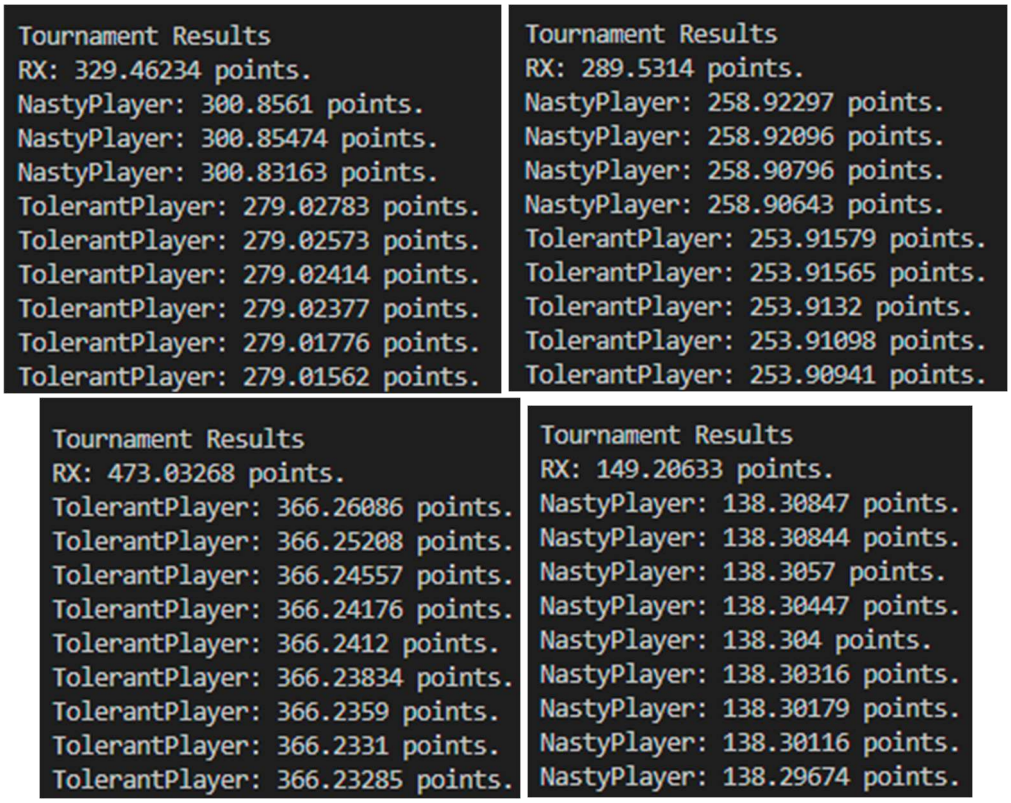

Fig 4

However, the results were disappointing when ran against majority T4TPlayers(see Fig 5)

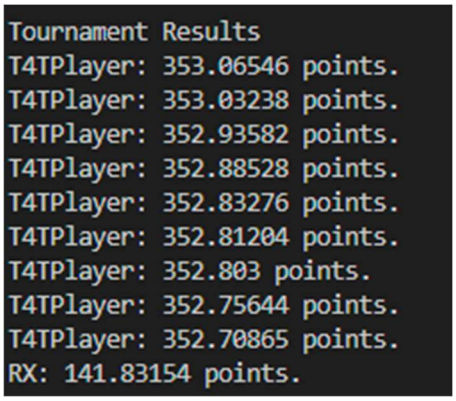

Fig 5

Identifying the Problem

After analyzing the results, I realize that my player RX, is only able to tackle a few types of players, the TolerantPlayer and NastyPlayer to be specific, it will lose when majority of the players are unforgiving ones such as Grim Trigger Player and T4Tplayers. I also recognised that my player is mainly self-interested and disregard the collective good of every player.

A Better Approach

After trying many trials and error of different strategies as well as research, it is better to aim for higher collective welfare than to only care about my own score. I decided to use a tit for tat kind of strategy and optimize on it by trying to achieve Program equilibrium.
The 4th version of my player is created: RX4

- My player will do the same as my opponents' previous action if both choose the same action in the previous round
- Else, my player will return my player's previous action.

This way, my player incorporates a tit for tat strategy as a mean to achieve program equilibrium, trying to maximize social welfare.

Tournament

Main Players played against **RX4** will be Nasty, T4T and Tolerant Players as Random players such as RandomPlayer and FreakyPlayer are completely irrational and would not show conclusive results in the short run if they are included in the game.

Game 1:

Number of match: 90 to 100
Number of players: 91
Players: 30 NastyPlayer, 30 TolerantPlayer, 30 T4TPlayer
Results:

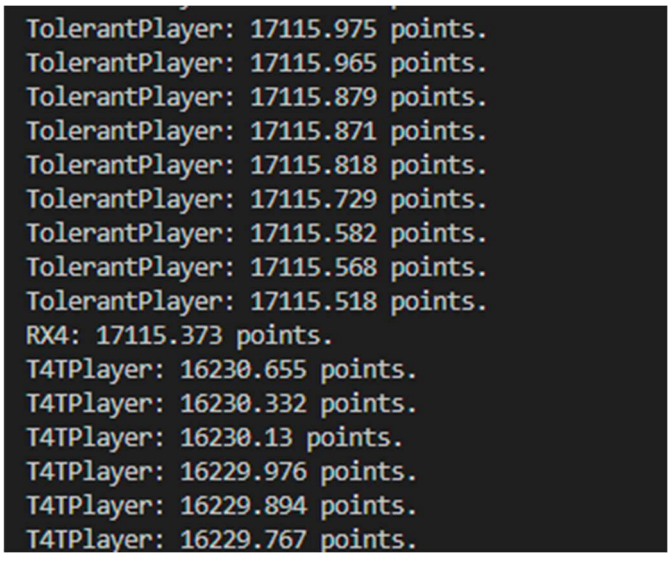

Fig 6

Game 2:

Number of match: 90 to 100
Number of players: 91
Players: 20 NastyPlayer, 20 TolerantPlayer, 50 T4TPlayer
Results:

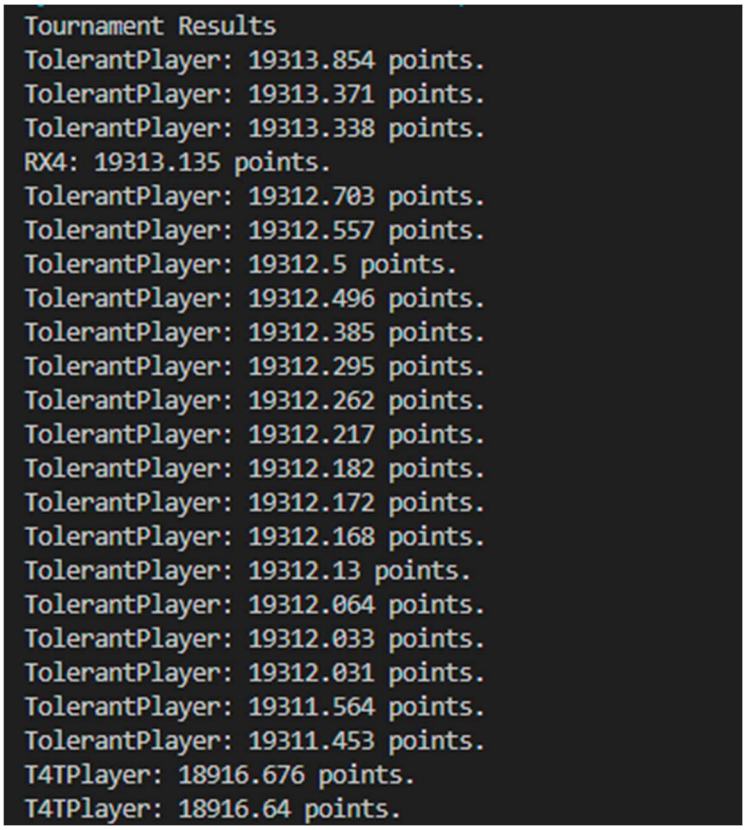

Fig 7

Game 3:

Number of match: 90 to 100
Number of players: 91
Players: 20 NastyPlayer, 50 TolerantPlayer, 20 T4TPlayer
Results:

```
TolerantPlayer: 19874.14 points.
TolerantPlayer: 19874.125 points.
TolerantPlayer: 19874.111 points.
TolerantPlayer: 19874.096 points.
RX4: 19874.092 points.
TolerantPlayer: 19874.014 points.
TolerantPlayer: 19873.95 points.
TolerantPlayer: 19873.928 points.
TolerantPlayer: 19873.871 points.
TolerantPlayer: 19873.86 points.
TolerantPlayer: 19873.836 points.
TolerantPlayer: 19873.836 points.
TolerantPlayer: 19873.834 points.
TolerantPlayer: 19873.83 points.
TolerantPlayer: 19873.82 points.
TolerantPlayer: 19873.812 points.
TolerantPlayer: 19873.812 points.
TolerantPlayer: 19873.81 points.
TolerantPlayer: 19873.797 points.
TolerantPlayer: 19873.785 points.
TolerantPlayer: 19873.781 points.
TolerantPlayer: 19873.756 points.
TolerantPlayer: 19873.748 points.
TolerantPlayer: 19873.729 points.
TolerantPlayer: 19873.723 points.
TolerantPlayer: 19873.693 points.
TolerantPlayer: 19873.684 points.
TolerantPlayer: 19873.646 points.
TolerantPlayer: 19873.629 points.
TolerantPlayer: 19873.598 points.
TolerantPlayer: 19873.568 points.
TolerantPlayer: 19873.512 points.
TolerantPlayer: 19873.39 points.
TolerantPlayer: 19873.217 points.
TolerantPlayer: 19872.773 points.
T4TPlayer: 18928.129 points.
```

Fig 8

My player RX4 managed to win against all the T4Tplayers and have roughly the same score as TolerantPlayers. (See Fig 6, Fig 7 & Fig 8)

Conclusion

In conclusion, it is possible to tackle against one type of player but it is almost impossible find a strategy that is able to win against all types of players completely, however, when developing our strategy it is best to keep in mind to aim for maximizing social welfare in a repeated game of N-player Prisoners' Dilemma, while at the same time not being exploited by self-interested players like NastyPlayer and the first player developed by myself.

Bibliography

[1] "Applicability of interactive genetic algorithms to multi ..." [Online]. Available: https://www.researchgate.net/publication/260530313_Applicability_of_Interactive_Genetic_Algorithms_to_Multi-agent_Systems_Experiments_on_Games_Used_in_Smart_Grid_Simulations. [Accessed: 16-Apr-2022].