

Final Year Project

Toby Devlin

November 28, 2017

Contents

1	Introduction	2
1.1	Background	2
1.1.1	Iterated Prisoners Dilemma	2
1.1.2	Machine Learning Concepts	3
1.2	Brief Overview	3
2	Literature Review	5
2.1	Background	5
2.2	Strategies Of Interest	5
2.2.1	Tit for Tat	5
2.2.2	Cycler	5
2.2.3	Other	5
3	Approach To The Problem	6
3.1	In Depth Definition Of Task	6
3.2	Finding Solutions	6
3.3	Solution Form	6
3.4	Initial Research	7
3.4.1	Changing Initial Population Size	8
3.4.2	Generation Length Analysis	13
3.4.3	Changeing Our Mutation Rate	17
3.4.4	Using crossover and mutation to remove local maximum solutions	19
3.4.5	Altering Initial Population	21
3.5	Conclusion of approach	21
4	Implementation Of Sequence Discovery	22
5	Results and Discussion	23
6	Practical Applications for Solution Sequences	24
7	Summary and Future Research	25

Chapter 1

Introduction

General discussion on what game theory is and what the PD is. What this topic is and how it fits into the big picture.

The Prisoners Dilemma is a classic game theory topic... Its important to game theory because... Creating general sequences to strategies is important because... Models in the real world that follow certain strategies... How we can leverage these to peruse goals.

1.1 Background

fill this section with background on game theory in when I do the course?

1.1.1 Iterated Prisoners Dilemma

The Prisoners Dilemma is a well known game theory problem based on the example of a pair of prisoners and their subsequent interrogation. The game is as follows:

Something about the PD

The single game itself is very basic and is modelled in the following way:

give – a – model – here

The Iterated Prisoners Dilemma is the iterated version of the Prisoners Dilemma¹. The iteration of the game is what makes the game an interesting concept, as now **learn the technical stuff and put it here!!!!** we are able to create strategies² that look to gain an upper hand based on **Something here**

¹reference this stuff dude, come on..

²When referring to ourselves, we will describe our moves as a strategy. When referring to

1.1.2 Machine Learning Concepts

This section will briefly provide a background to machine learning algorithms implemented in the axelrod-doj³. This is by no means a comprehensive look into these subjects but will provide sufficient background on technical discussion later on.

Genetic Algorithms

Genetic Algorithms are a description of techniques for generating solutions complex problems such as searching and, in our case, optimization⁴. The basis of a genetic algorithm is focused on a cycle of evolution. Like nature, we create a survival of the fittest concept⁵ to evaluate a population, kill off the weakest members and create offspring from the most successful population.

Starting with a set of randomly generated sequences, we will have each one play the opponent and return with a score. These sequences will be ranked and the lowest pairings% will be discarded, resulting in a fitter, but smaller, population than before. This smaller population will then create offspring using a —X TBD method X— pairing algorithm before mutating with —X TBD method X—. This new set of offspring will be included in the next scoring round and the process repeats for k number of rounds

Put a figure of the cycle here.

Initially we create a heuristic function, say our fitness function, which is a measure of how successful a candidate in our population is. Then we run our whole population through this function, ranking each one by how successful their score is. At this point we can create a cut off⁶ to decide which of the population not to put through to the next round.

Bayesian Optimization

1.2 Brief Overview

In this document I will be looking at the creation of sequences to beat given players in The Iterated Prisoners Dilemma⁷. My research looked into just the single opponent use case, but the idea of designing a sequence for a given number of opponents is looked at in the further study of the report. This task is the

an opponent we can use the term opponent and strategy interchangeably.

³referencing opportunity here

⁴Mitchell, Melanie (1996). An Introduction to Genetic Algorithms. Cambridge, MA: MIT Press. ISBN 9780585030944. learn to reference soon

⁵need to reference Darwin?

⁶Can often be referred to as the bottleneck

⁷Reference this for some background

Problem:

Given a certain opponent, O , (with a provided strategy, S) what is the best possible sequence of moves, in a game of n turns, made by my strategy to maximise my players score?

Chapter 2

Literature Review

2.1 Background

- book1
- book2

2.2 Strategies Of Interest

2.2.1 Tit for Tat

this is a sentence which has been changed[A⁺87]

2.2.2 Cyclor

this is a type of opponent

2.2.3 Other

add at least two here

Chapter 3

Approach To The Problem

3.1 In Depth Definition Of Task

What is the task? *When playing an Iterated Prisoners Dilemma strategy what is the best ordered series of moves to play such for us to obtain the highest possible average score per move across the game.*

What is the motivation for the task?

What are we expecting to come from the task?

Terminology?

3.2 Finding Solutions

The sequence archetype will use the *Cycler()* player for our strategy each time, only editing the input to improve our score. To this model we can apply an optimised input of length 200 to the player, this sequence, as per the design of the strategy will then be repeated until the games end (if $n = \text{length of game}$, we are just calculating the sequence for the whole game). The input sequence itself will be created using a genetic optimisation, see Section 1.1.2 for in depth explanations.

This sequence of Play-Rank-Create-LOOP will be the basis of creating the optimal strategy for each other opponent.

3.3 Solution Form

Our algorithm will, after its generations are concluded, an arbitrary sequence of the form:

$$X, X, X, \dots, X \text{ where } X \in C, D$$

This sequence will be represent what moves we should play against the opponent to get our largest potential score per turn. In this notation a C represents

a cooperation move and a D represents a defection move.

We are looking for sequences that will allow us to maximise our score overall, rather than just beating any given opponent. A nice analogy of this concept is a team playing a football tournament, but instead of a knockout competition our team is placed in the standings based off the total goals they have scored across the tournament. More applications of these results are discussed in Chapter 6.

3.4 Initial Research

Before conducting the bulk calculations for the set of opponents listed in the appendix we will test the algorithms' parameters to see what the best settings are for finding solution sequences. We will describe a series of sequences as 'converged' if the best score over the set number of generations has reached a stable point; described as when none of the moves in the sequences has changed over a number of generations. This stable point, however, may not be the optimal solution. we may have found a local maxima for the solution sequence rather than the global maximum.

The opponents we select are in some way interesting. They are all 'simple' and can be explained in a very brief sentence or two, but each one has a fundamentally different structure to how they work¹. We will look into these as we can confirm that the genetic algorithm will select the optimal sequence solution for the selected opponent.

Player	Optimal Sequence
axl.TitForTat()	<i>CCC...CD</i>
axl.Alternator()	<i>DDD...DD</i>
axl.Grudger()	<i>CCC...CD</i>
axl.Random()	<i>DDD...DD</i>
axl.EvolvedFSM16()	??
axl.CollectiveStrategy()	??

What we want to look at is how the best score rises over generations as we change certain features of the algorithm. Once the best score hits a maximum such that it wont change no mater how many more generations are run we call this the optimal solution sequence². If the algorithm finds a sequence with a score which doesn't change over a significant period of generations we can say that the sequence has converged to a solution sequence, however this solution is not optimal.

Once a solution has been found the generation number where this plateau occurs is called the solution sequence distance, or solution distance; minimising this is

¹See section 2.2 for detailed discussion.

²Proof of a unique solution sequence for an opponent is out of scope

that aim of this initial investigation. During the investigation we may find solutions that are not optimal, meaning that the algorithm will have found a sequence that will do well against an opponent but won't find the best sequence that will return can possibly get. These sub optimal solutions are due to the occurrence of local maximums in the set of scores of neighbouring sequences. Will look in depth into how to overcome the possibility of our algorithm finding a local, rather than, global maximum in Section 3.4.4. Some of the questions we will hope to be answering include:

- If we make each sequence more likely to mutate generation to generation what will happen?
- If we have a larger initial population sample to start with, will we reach our maximum best score earlier?
- What about increasing the generations, what is the optimal number of generations to run the algorithm for such that we always find a solution sequence.

We have selected our opponents for this section is such that we are able to know their optimal solution sequence as to identify local maximum solution sequences. The next sections are a discussion on how these parameters effect the outcome of our goals.

3.4.1 Changing Initial Population Size

The initial population size is the number of starting sequences we use in our algorithms first generation. Once this generation concludes the population will go through the series of phases outlined in figure 1.1.2; altering the population to keep the best performers against our opponent to continue on to subsequent generations. During any given generation the population defines the maximum potential range of scores that we can achieve against our opponent. For example, having 2 members with distinct sequences in our population would provide us with 2 distinct³ scores after a game. Because of this we can reasonably assume the larger our population the larger the number of distinct scores leading to a larger chance of finding the solution sequence with the optimal score; hence we should converge to the solution sequence in less generations.

The implementation of analysing a range of populations is simple, all we need to understand how the solution distance is affected as we increase our populations through a set of sizes, say population sizes of: [25, 50, 100, 150, 200, 250, 500].

Note: Increasing the size of our population will have an impact on computation time; each generation must process the full population in a linear fashion causing a computation overhead of $O(n)$. For an increase to be useful in any time

³Would this be 'potentially distinct'? Maybe prove it?

restricted scenario our algorithm would need to show a higher order benefit in our distance to convergence, or in our average score per turn. However We are not working in a time restricted scenario, and so we should just select the best overall initial population size independent of computation overhead. In a perfect world where everything was time independent we would brute force every possible sequence combination

The code in 3.4.1 is an implementation how we go about analysing and storing the tests on generation sizes listed. It leverages the use of the function ‘run-GeneticAlgo’ show in appendix code ??

This output will provide us with a table with the following form⁴:

best score	gen.	mean score	population	sequence	std dev.	time taken
2.425	1	2.264	25.0	DD...	0.067	6.646
2.425	2	2.343	25.0	DD...	0.046	6.646
2.425	3	2.393	25.0	DD...	0.038	6.646
...
2.830	102	2.782	100.0	CC...	0.112	28.425
...
2.980	150	2.911	500.0	CC...	0.158	152.684
...

By grouping this data by the population we observe how initial populations affect different opponents. Its clear that from figure 3.2 that the initial population size has a significant effect on finding better sequences. We can see if there is a larger initial population there is typically a higher best score shown once concluding all of the generations. This can also be shown in figure 3.3 that... It doesn’t, however, ensure that we find the solution sequence; as is shown in the lack of long plateaus in the lines.

The improvement’s from this effect are non-linier from observation. The change in overall final best score for a population of 50 compared with a population of 100 is huge in comparison to the same relative increase from 200 to 250. This may suggest there are more effective approaches to improving our score after a certain size of initial population than to continuing to increase it further.

None of these results have found a solution sequence (or at least we cant tell from the graph). It is clear that larger initial populations do, on a relative scale, much better than small ones. There are no large plateaus for the graph, so as we continue our research the initial population size will be increased to 150 to keep computation times manageable.

⁴Data has been concatenated for this summary table.

```

def populationChecker(opponent):
    # make a nice file name
    file_name = "data/" + str(opponent).replace(" ", "_")
                                                .replace(":", "_")
                                                .lower()
                                                + "_pop.csv"
    # if the file exists don't run, it takes forever, make sure it exists
    if not os.path.isfile(file_name):
        df_main = pd.DataFrame(data=None, columns=col_names)
        for pop_size in populations:
            start_time = time.clock()
            pop_run = runGeneticAlgo(opponent,
                                    population_size=pop_size,
                                    number_of_game_turns=200,
                                    cycle_length=200,
                                    generations=150,
                                    mutation_probability=0.1,
                                    reset_file=True)

            end_time = time.clock()
            tmp_df = pd.read_csv(pop_run[0], names=col_names)
            tmp_df["population"] = pop_size
            tmp_df["time_taken"] = end_time - start_time
            df_main = df_main.append(tmp_df, ignore_index=True)
        df_main.to_csv(file_name)
        print("List Complete:", file_name)
        return df_main
    else:
        print("file already exists, no calcs to do.")
        file_df = pd.read_csv(file_name)
        # remove first column
        file_df = file_df[list(file_df)[1:]]
        return file_df

```

Figure 3.1: code to check multiple populations



Figure 3.2: Best score per turn vs generation for different initial population sizes



Figure 3.3: Scatter of max best score vs different initial populations

3.4.2 Generation Length Analysis

Another major component parameter of a genetic algorithm is the number of generations it will run before concluding on a final sequence. The number of generations has an influence on a number of different things within the algorithm:

- The total combinations of features⁵ (sequence elements) that the algorithm can test.
- The number of low performers we can remove in our population.

For our goal of finding the optimal solution sequence for each opponent it would be useful to extend the generations as far as possible; this would provide the most combinations of features possible. Here we will look into how close to a solution sequence we get when we increase the generations the algorithm runs for. Like in previous experiments with other variables we will use a range of sizes for our parameter to run our analysis over⁶; say, generation lengths of: [50, 150, 250, 350, 450, 500]. The code in 3.4.2 shows how we will approach the analysis.

Generation size differs from other parameters in the fact this is purely performance based. A genetic algorithm with 1 generation is just a series of tests; with the results split into 2 sets — winners and losers. As we extend the generations we would be more focused on what happens to certain averages of results across the whole run, rather than absolute improvement. If we look at the scatter of mean best score difference against the number of generations in figure 3.5 we can observe how, on average, the number of generations has a declining effect the overall change in our best score per generation.

This mean increase of score per generation trend is to be expected; it gets harder to randomly select which element in the sequence needs changing to improve our score when we are close to a maximum. On this result we can conclude as we increase generations there is less and less benefit per generation; there is however still a benefit. There may be a benefit from increasing the mutation rates when we get close to one of these maximums; the more noisy our algorithm is for sequences could improve our chance of finding the correct solution.⁷

Figure 3.6 shows the proximity the optimal solution sequence once the analysis has concluded. A good score is a score of 3 or more; this can change from player to player, and is never explicitly obvious. We can see that after a number of generations that solutions sometimes get ‘stuck’ in a local maximum score.

⁵Section 1.1.2 explains what we mean by feature selection

⁶We will be using a population of 150 as this was the best average for score vs computation time for analysis.

⁷Due to using a prebuilt algorithm in this code there is no way to dynamically increase mutation rates, or other parameters, at points in our algorithm. This may be a potentially interesting research topic on how to build an ‘intelligent’ or preprogrammed genetic algorithm.

```

def generationSizeChecker(opponent):
    file_name = "data/" + str(opponent).replace(" ", "_")
                                                .replace(":", "_")
                                                .lower()
                                                + "_generation.csv"
    if not os.path.isfile(file_name):
        df_main = pd.DataFrame(data=None, columns=col_names)
        for gens in generation_list:
            start_time = time.clock()
            pop_run = runGeneticAlgo(opponent,
                                    population_size=150,
                                    number_of_game_turns=200,
                                    cycle_length=200,
                                    generations=gens,
                                    mutation_probability=0.1,
                                    reset_file=True)

            end_time = time.clock()
            tmp_df = pd.read_csv(pop_run[0], names=col_names)
            tmp_df["generations"] = gens
            tmp_df["time_taken"] = end_time-start_time
            tmp_df["opponent"] = str(opponent)
            tmp_df["best_score_diff"] = np.append([0], np.diff(tmp_df["best_score"]))
            df_main = df_main.append(tmp_df, ignore_index=True)
        df_main.to_csv(file_name)
        print("List Complete:", file_name)
        return df_main
    else:
        print("file ", file_name, " already exists, no calcs to do.")
        file_df = pd.read_csv(file_name)
        \# remove first column
        file_df = file_df[list(file_df)[1:]]
        return file_df \

```

Figure 3.4: code to check multiple generation lengths.



Figure 3.5: Mean Best Score diff vs Total Generation Lengths



Figure 3.6: Max best score vs total number of generations

population that continue into the next generation one of these has a single gene, or action, that is changed in its sequence.

Here we will look at 2 different concepts:

- Is changing one or more actions of a members' sequence the best way of mutating a candidate (More potent mutation)
- Is it beneficial for more/less than 1 in 10 members to be mutated generation to generation? (More frequent mutation)

These are two separate questions, so first we will look at increasing the potency of our mutation. Then, once we have found some information out on this, we can look into the frequency of our mutations with the new potency as a permanent setting. This approach allows for an $O(1)$ factor of scaling making this a great candidate for an approach to reduce our solution sequence distance compared with other approaches, for example increasing the population size.

Changing Mutation Potency

mutation_potencylist = [1, 2, 3, 5, 10, 15, 20]

def mutationPotencyChecker(opponent) : filename = "data/" + str(opponent).replace("'", " ").replace(":", " ").

This single graph shows no clear benefit from increasing the mutation potency. We can see that having changed 15 genes in our sequence each time we are still not improving our score as much as using 1. This may be down to chance (and if the file is regenerated this may disappear) looking at more opponents than just grudger we may find a clear improvement. Now we can look at how our best score is effected against other opponents.

From these graphs there is no clear benefit to increasing the mutation potency to affect the overall best score value against an opponent. If we instead look at what our average increase of score per mutation is we may observe a useful result.

From further analysis there is not much of an improvement by increasing your mutation potency with regards to score or for average increase of score. The increase in mean best score difference is not substantial and could be down to chance

Changing Mutation Frequency

mutation_frequencylist = [0.1, 0.15, 0.2, 0.25]

def mutationFrequencyChecker(opponent) : filename = "data/" + str(opponent).replace("'", " ").replace(":", " ").

The results on changing the mutation rate don't obviously effect that generations until convergence from this overview. There is an interesting result that can be seen on the grudger plot; the algorithm has found 2 different maximums.

From this there we can see that the mutation frequency of 0.1 and 0.2 produced higher scoring solutions than the other mutation frequencies. This shows that we have found 3 different solution sequences (in freqs .15, .2, .25) with the solution for .1 continuing to improve as the generations ended.

As an additional point, we can also observe that increasing the mutation frequency means that there is less variation in the best scoring sequences. (TODO: Does this have an impact on escaping local maximums?)

3.4.4 Using crossover and mutation to remove local maximum solutions

The occurrence of local maximums is something that has only occurred for the Grudger opponent so far. The difference between the Grudger and the other opponents were looking at is that the Grudger has a singularity where its behaviour changes. The change in behaviour is not uncommon, Tit For Tat works in the same way, however this algorithm has managed to identify its behaviour and adapt to overcome its negative effects.

Grudger and Tit For Tat differ in their responsiveness in two ways:

- Grudger never changes its mind. There is one change in behaviour for the entire game. Unlike Tit For Tat, this means that the algorithm only has a single opportunity to observe this once every per population per generation meaning the behaviour is much less frequently observed.
- The Grudger also only does this change once no matter the games length. This means that the genetic algorithm picks up the effect of this choice as early in the match as the first defection in its random sequence. A random start of C and Ds puts the likelihood of at least 1 defection occurring in the first 10 moves at 99.99; this means our algorithm will, most likely, always encounter this grudging effect within the first 10 moves and will never score the full 600 points. (see below)

Below are two totality games, one of all Cs and one of all Ds. These are edge cases and would be incredibly rarely encountered as a starting point in the initial population. Because of this the algorithm has to shuffle towards the potential benefit of using these totalities rather than start with analysing them, and in our case the algorithm will probably first encounter the Grudging effect of out opponent before trying out [CCC..C] or [CCC..D] and so will probably never find the highest scoring solution.

```
players = (axl.Grudger(), axl.Cycler("C")) match = axl.Match(players, 200) match.play() print(match.final)
```

```
players = (axl.Grudger(), axl.Cycler("D")) match = axl.Match(players, 200) match.play() print(match.final)
```

Strangely, our solution sequence is set to find where we have the objective of "score" (see objective statement) which actually tries to improve the score per turn (*axelrod_dojos_utils.py* : 67); it should be converging on a totality of Cs rather than what its doing by finding the totality of Ds. This is probably because the algorithm initially limits its best score per turn once the first generation is complete and a cut-off has been established for each of the initial population. The crossover method between generations then doesn't provide enough of a mix up to allow the algorithm to escape the local minimum by switching a subsection with a sufficiently different potentially better subsection. Then when it comes to mutating, there is little any number of mutations can do to drastically change large sections of the sequence without having a huge effect on the score.

This then sheds light on the path the algorithm takes to find a solution. If we are to find the optimal solution, we must take a crossover and mutation path which doesn't cut off better paths as we work our way towards a solution; this is much easier said than put into practice due to the way the algorithm "cuts off paths". If we reverse this thinking and try to alter our crossover design and mutation rate such that instead of "cutting off" a path we are able to "build" new ones. We can re-design the crossover to switch up large subsections of the sequence then allow the mutations to optimise these sub-sequences.

currently we have the following design:

```
def crossover_old(self, other_cycler) : #boringsinglepointcrossover : crossover_point = int(self.get_sequence_len() / 2)
```

We want to allow the crossover to have more of an impact than just halving the sequence and optimizing each section. i.e. go from:

```
|-----|and|+++++|
+++++++=|-----++++++| to, say:
|-----|and|+++++|
+++++++=|---++--++--++--++--++|
```

This will allow the mutation rate to edit the subsections in a more interlaced manner, hopefully overcoming the pitfalls of sparse mutations to escape local maximums. our new crossover method is as follows:

```
def crossover(self, other_cycler) : #10crossoverpoints : step_size = int(len(self.get_sequence()) / 10) #emptysta
```

Below is an exapmple of the new crossover sequence:

```
seq1 = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1] seq2 = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0] step_size = 10
```

now we can look at how this new crossover algorithm works with the default mutation (freq=.1 and pot=1) to improve our local maximums with the grudger opponent:

3.4.5 Altering Initial Population

This section is to display the results of working on an initial population that contain common solution sequences. We will discuss entropy of a solution and that by creating "neat" starting points we can decide where to start on the plane, mitigating potential sub-optimal solutions.

From the conclusions in previous sections the problem with creating

3.5 Conclusion of approach

Chapter 4

Implementation Of Sequence Discovery

Chapter 5

Results and Discussion

Chapter 6

Practical Applications for Solution Sequences

maybe how this works with respect to pathfinding, who we should select in a set of opponents if we only have to play a single opponent in the set.

Chapter 7

Summary and Future Research

Bibliography

- [A⁺87] Robert Axelrod et al. The evolution of strategies in the iterated prisoners dilemma. *The dynamics of norms*, pages 1–16, 1987.