

Final Year Project

Toby Devlin

March 18, 2018

Contents

1	Introduction	3
1.1	Iterated Prisoners Dilemma	3
1.2	Machine Learning & Computer Intelligence	3
1.2.1	Genetic Algorithms	3
1.3	Brief Task Overview	4
2	Literature Review	5
2.1	Background	5
2.2	Strategy Structures	5
2.2.1	Equivalent strategies	6
2.3	Strategies Of Interest	6
2.3.1	TitForTat	6
2.3.2	Alternator	6
2.3.3	Grudger	6
2.3.4	Random	6
2.3.5	EvolvedFSM16	6
2.3.6	CollectiveStrategy	6
2.3.7	ZDExtortion	6
2.3.8	Cycler	7
2.4	Genetic Algorithms	7
3	Task Background	8
3.1	Notation	8
3.2	Solution Form	9
3.3	Development Paradigms	9

3.3.1	Building The Algorithm	9
4	Developing The Codebase	11
4.1	Libraries And External Modules	11
4.2	Setting Up A Research Environment	12
4.3	Codebase Contributions	13
4.3.1	Version Control	13
4.3.2	Testing	14
5	Implementation Of Sequence Discovery	16
5.1	Changing Initial Population Size	17
5.2	Generation Length Analysis	19
5.3	Changing Mutation Rate	23
5.3.1	Changing Mutation Potency	24
5.3.2	Changing Mutation Frequency	25
5.4	Mitigating local maximum solutions	26
5.4.1	Ineffective Approach of Altering Crossover And Mutation	29
5.5	Altering Initial Population	32
5.5.1	Population Size	33
5.5.2	Generation Length	34
5.5.3	Mutation Potency	34
5.5.4	Mutation Frequency	34
5.6	Stochastic Opponents	34
5.7	Conclusion of approach	36
6	Results and Discussion	40
7	Practical Applications for Solution Sequences	41
8	Summary and Future Research	42
A	Code Appendix	43
B	List of Axelrod Opponents	44

Chapter 1

Introduction

In this report we will be looking at how strategies of the Iterated Prisoners Dilemma (IPD) can be ‘countered’ or ‘worked’ in order to give a better overall score for a player on the other side of the game. We will look into explicit sequences which should be played for single games of length 200 against a single opponent. These sequences will then be compared and contrasted to understand how some opponents could be grouped together even though they have no structure in common.

The concept of ‘working’ a strategy is more fitting in this research as we will be looking into boosting the score of our player, which in most cases (due to the nature of the IPD) will also boost the score of our opponent. ‘working’ also has a better real world interpretation than ‘countering’ as the real goal of increasing the score for our player is to generate as many Cooperation moves to Defect against without consequences as possible. The IPD game itself is described in Section 1.1 and the real world research is discussed in Chapter 2.

1.1 Iterated Prisoners Dilemma

See <https://vknight.org/gt/chapters/09/>

1.2 Machine Learning & Computer Intelligence

Machine Learning is a field of computer science that has existed since the first computers have been around. Most famously the questions posed by Alan Turing in 1950 [Tur50] asks ‘can machines think’, a question that has been refined and analysed to this day. The field of computer intelligence is rich in its complexities and has recently been making breakthroughs [Kni17] on topics which would traditionally be considered ‘thinking’. Along with this, recently there has also been record levels of funding [Chu17] put in to companies which operate in this field, producing results in areas that would usually seem ‘solved’.

1.2.1 Genetic Algorithms

This report will cover one of the forms of machine learning called genetic algorithms (GAs). Genetic Algorithms fall under a branch of machine learning called evolutionary feature selection algorithms. More generally, genetic algorithms are put into a class of unsupervised reinforcement learning algorithms. These are techniques of using genetic algorithms for generating solutions to problems that typically revolve around heuristically improving members of a population who represent these solutions. The concept of a genetic algorithm, and more generally an evolutionary algorithm, comes from nature; like nature we create a survival of the fittest selection [DB09] competition to evaluate a population then kill off the weakest members. After this cull we create offspring from the most successful population or introduce new members

from a predefined source. This process is then repeated until we stop it, or forever in the case of nature.

We say a genetic algorithm is structured in the following way. Given a population, P , each with unique genes (genes and member properties are interchangeable), and a number of generations, $G \in \mathbb{N}$, the algorithm will create G cycles of scoring and potentially removing each of the members of the population, $p_i \in P$. It does this by using a mapping from a member of the population to an ordered set, for example $f(p_i) \mapsto \mathbb{R}$, $p_i \in P$. This function, f , is defined beforehand in a way which describes the goal of our investigation. Defining a cut-off or bottleneck $b < |P|$, such that on conclusion of completing any cycle, the top b ranking members by score can be kept and the rest discarded. By doing this we are saving the more successful candidates allowing us to rebuild the population using a series of crossovers and mutations (and possibly introducing new members into the population) with the genes which were successful.

- Crossovers take in 2 members of the population and return a new member based on some parameters of the 2 ‘parents’. For example, our crossover takes the first half of a sequence from one and the second half from the other, merging them to form the third.
- Mutations allow a (possibly targeted¹) change in a single member of the population. A mutation has 2 parameters, a potency $M_p \in \mathbb{R} > 0$ and a frequency $M_f \in [0, 1]$. M_p describes how strong the mutation is, the higher it is the larger change to the member occurs. M_f explains the percentage of how many members of the population are mutated.

Figure 1.2.1 shows a flow diagram of a generic genetic algorithm cycle.

Put a figure of generic cycle here.

This is a general example of a generic algorithm. The specific algorithm we will be using is described in Subsection 3.3.1.

1.3 Brief Task Overview

This document will be looking at the creation of sequences to beat given opponents in The Iterated Prisoners Dilemma. Analysis will be focused on looking into just the single opponent use case, but the idea of designing a sequence for a tournament for a given number of opponents is a potential follow on to this work. Our task is as follows:

Problem:

When playing a given Iterated Prisoners Dilemma strategy, O , as an opponent, what is the best ordered solution sequence of moves, S , to play in order for us to obtain the highest possible average score per move across the game.

For example an opponent known as Tit For Tat, which cooperates on its first move and then copies your last move on subsequent moves, will have a solution sequence of moves that are all cooperation apart from the last. This is an obvious example and a simple strategy to calculate for, the ultimate goal of this investigation is to look into every strategy as defined in the Axelrod Library. These are listed in appendix Section B and definitions can be found in the axelrod codebase[pd17].

¹For example using intuition and targeting specific genes, or allowing another algorithm to improve the targeting of this meta function.

Chapter 2

Literature Review

2.1 Background

The Prisoners Dilemma a large area of repeated games in Game Theory and has applications in the real world. This is due to the game being a good example of strategies that give a cooperative benefit to repeated games. It has been used to describe actions of people and governments in situations stretching from warfare [TC88, AH92], finance [CS97] and politics [Sni85] to sex and relationships [Low15]. Because of its applicability to real world scenarios there is a strong desire to understand how strategies are beaten to either exploit flaws or counter opponents in real world scenarios.

The Prisoners dilemma became a large research area in the combination of mathematics and computer science after Robert Axelrod published his work named "Effective Choice In The Prisoners Dilemma" [Axe80]. In it he makes an introduction to how tournaments are run and the properties of successful results. His method of experimenting became the standard for handling the IDP problem. After the tournament is complete he describes what successful strategies had in common; It turns out the majority of strategies have properties of niceness and forgiveness. This allowed them to thrive in the tournament and have overall scores that rose above strategies without niceness or forgiveness.

Since Axelrods' original tournament there have been many research papers on what makes a successful strategy for a specific objective. For example [PD12] looks at how to remove an opponents moves to reach a high score and [MD09] looks at a range of different objectives at once. These specific objectives are really the core part of applying Game Theory and The Prisoners Dilemma in the really word. Objectives are the wrapper for which we can work with real world scenarios, for example in a the cold war it was not the goal to get as many missiles as possible to pass an oppositions defence but to not allow any missiles through your own defence; i.e.minimising your opponents score. In a football tournament where winning is the number or goals your team scores it doesn't really matter how many goals you get in so long as you score the most overall. There are also some very useful applications such as rent splitting or work assignments [GP15], all of which were programmed and deployed to a web server for general use online¹

2.2 Strategy Structures

Strategies can be defined in multiple ways [HKJ⁺17]. Each method of representing a strategies has its benefits and drawbacks, though there is no method (or strategy in that case) that is best in all cases. There are methods of creating strategies that remove the way opponents play in order to create overall scores that desired. Subsection 2.3.7 discusses this in more detail. Ways of structuring a strategy are shown below:

- LookerUpper, find examples
- Gambler

¹<http://www.spliddit.org>

- Neural Networks
- Finite State Machines
- Hidden Markov Models
- Explicit Move List
- Mixtures

2.2.1 Equivalent strategies

When performing analysis of an opponent using a GA it can sometimes be useful to brute force the starting positions of our feature selection to create rare² starting points. Looking at certain opponents there are occasions some strategies that look indistinguishable from others; for example Alternator and Random (0.5) can very often create the same output. One of the ways we are able to identify strategies is the process of fingerprinting [AKK04, AK08].

Another way is to look up their structure definitions in the same model (for example...)

2.3 Strategies Of Interest

2.3.1 TitForTat

Tit For Tat was the

2.3.2 Alternator

See Subsection 2.3.4.

2.3.3 Grudger

See Subsection 2.3.4.

2.3.4 Random

For all the strategies () given above we are able to logically deduce that the best counter is a totality of Ds, no matter what their definition structure is. This leads us to consider what the common theme is between the strategies as to be beaten in the same way.

2.3.5 EvolvedFSM16

2.3.6 CollectiveStrategy

2.3.7 ZDExtortion

This is the paper [PD12]

²For example of a random sequence that is a totality of Cs is incredibly rare, with around 6.2^{-61} chance of occurring

2.3.8 Cykler

2.4 Genetic Algorithms

This work will focus on Genetic Algorithms (GAs) which are a specific form of Machine Learning (ML). Advanced techniques of machine learning can be combined and used together³ in many situations, lots of these techniques combine genetic algorithms or some sort of fitness testing within a larger scope. The field of mathematics research is one which has plenty of examples of GAs in action; for example the same techniques as we will using is used in [CB97] to solve the Generalised Assignment Problem. Another example, [BLM95], used neural networks when approaching the state regulation problem.

In Game Theory GAs are used in a couple of areas, most prominently in the study of the iterated prisoners dilemma.

³Techniques for teaching and versioning static algorithms such as building a ‘clever’ game AIs, where the core concept of the AI is fine tuned using GA in an development environment but isn’t implemented into the game [BSVdH09].

Chapter 3

Task Background

This chapter will cover sections on how each part of the project will be carried out. Useful notation is given in Section 3.1 however in some areas of discussion sequences are more easily denoted by explicitly describing this sequence. Section 3.3 on code techniques will cover techniques used in the analysis and a walkthrough of setting up an environment to conduct similar analysis is given in Section 4.2.

3.1 Notation

Let $S \in \{C, D\}^L$ where C, D represents a cooperation, defection respectively. $L = 200$ is used throughout this report. We can split up sequences into blocks of consecutive move elements of the same type. We will use B_i to denote the block after i changes of move type from the explicitly stated starting move type.

- Every move in a block is of the same type; the type is implicit based on whether i is even or not and what the starting move type was.
- We can use the notation $|B_i|$ to denote the length of the i th block in the sequence. $|B_i| \in \mathbb{Z}$

This means we can write a sequence as a series of blocks:

$$S = B_1 B_2, \dots, B_n$$

A Sequence can also be defined shorthand by specifying the starting move and the length of subsequent blocks:

$$S = C : [|B_1|, |B_2|, \dots, |B_n|] \Rightarrow S = \overbrace{C \dots C}^{|B_1|} \overbrace{D \dots D}^{|B_2|} \dots \overbrace{(C|D) \dots (C|D)}^{|B_n|}$$

We can also construct sequences from repetitions of a sequence of blocks when it makes sense:

$$C : [(|B_1|, |B_2|, \dots, |B_m|)^k] \Rightarrow \overbrace{C \dots C}^{|B_1|} \overbrace{D \dots D}^{|B_2|} \dots \overbrace{D \dots D}^{|B_m|} \text{ }^{k\text{-times}}$$

The two notations can be combined to add starting and ending blocks to a repeating sequence (shown in the examples).

It is also possible to define sets of sequences by adding variables to parameters of the sequence. Appropriate selection of parameters mean the length of sequences shouldn't grow.

$$\{C : [i, l-i]\} \quad i \in [a, b] \Rightarrow \{ \underbrace{C \dots C}_a \overbrace{D \dots D}^{l-a}, \underbrace{C \dots C}_{a+1} \overbrace{D \dots D}^{l-(a+1)}, \dots, \underbrace{C \dots C}_b \overbrace{D \dots D}^{l-b} \}$$

For long sequences where there is no recognisable pattern it is typically easier to describe the solution. Otherwise we use the notation to describe a sequence. Examples:

$$C : [1, 4, 3, 2] = CDDDDCCDD \quad (3.1)$$

$$D : [(1, 1)^5] = DCDCDCDC \quad (3.2)$$

$$C : [1, (2, 1)^2, 2, 1] = CDDCDDCDDC \quad (3.3)$$

$$\{D : [i, 5 - i]\} \quad i \in [2, 4] = \{DDCCC, DDDCC, DDDCC\} \quad (3.4)$$

$$(3.5)$$

3.2 Solution Form

In this research our goal will be to use the algorithm, described in Subsection 1.2.1, to produce an arbitrary sequence for an opponent, S_o . This sequence will represent the moves we should play against the opponent to get our largest potential score per turn for a single game of 200 turns.

This investigation focuses on sequences that will allow us to maximise our score overall, rather than just beating any given opponent. An 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 real world applications of these results are discussed in Chapter 7.

Each solution sequence is uniquely generated for each opponent. If there are two similar opponents, say Grudger and Collective Strategy, these will be independently analysed and individual solutions will be generated. The output sequences will be compared on how similar they are to each other, what type of opponents and their corresponding 'best scores' are and whether there are patterns to how to group strategies, Chapter 6 contains further details.

When working with stochastic players, we will be seeding them in order to determine the best sequence. Each stochastic player will be considered under a variety of seeds. The motive to allow this to happen is explained described in Section 5.6. Non-stochastic opponents will be seeded in the code, however there is no change in their behaviour because of this.

3.3 Development Paradigms

This section will discuss how we will leverage code to complete aspects of the investigation. We will be using Python because of its flexibility and it is the language the required libraries are written in.

3.3.1 Building The Algorithm

In order to generate a solution sequence we have to train against each Axelrod opponent against another Axelrod opponent. We will use the Cycler strategy¹ in the algorithm by editing the input parameter to improve our score against an opponent.

Our implementation of a genetic algorithm is more custom and has the following steps:

1. Start with a predefined population, supplemented with randomly generated member until to size.
2. Each member plays the given opponent with their sequence and returns with the average score per turn.
3. Members of the population are ranked by this average score per turn and the highest scoring 25% will be kept for the next round. The remaining 75% are killed off.

¹See Section 2.3.8 for description

4. The remaining population will then be copied and these copies mutated to create unique sequences before being merged back in to the main population.
5. The remaining 50% difference is then made up of mutated results of crossovers from members of the current population or random new members, depending on a random selection algorithm.²
6. A generation has now concluded. Repeat from step 2 until the desired number of generations are finished and a final best sequence is returned.

Figure 3.1 shows a flow diagram of our cycle. This is the algorithm we will use in Chapter 5 when analysing the algorithms parameters

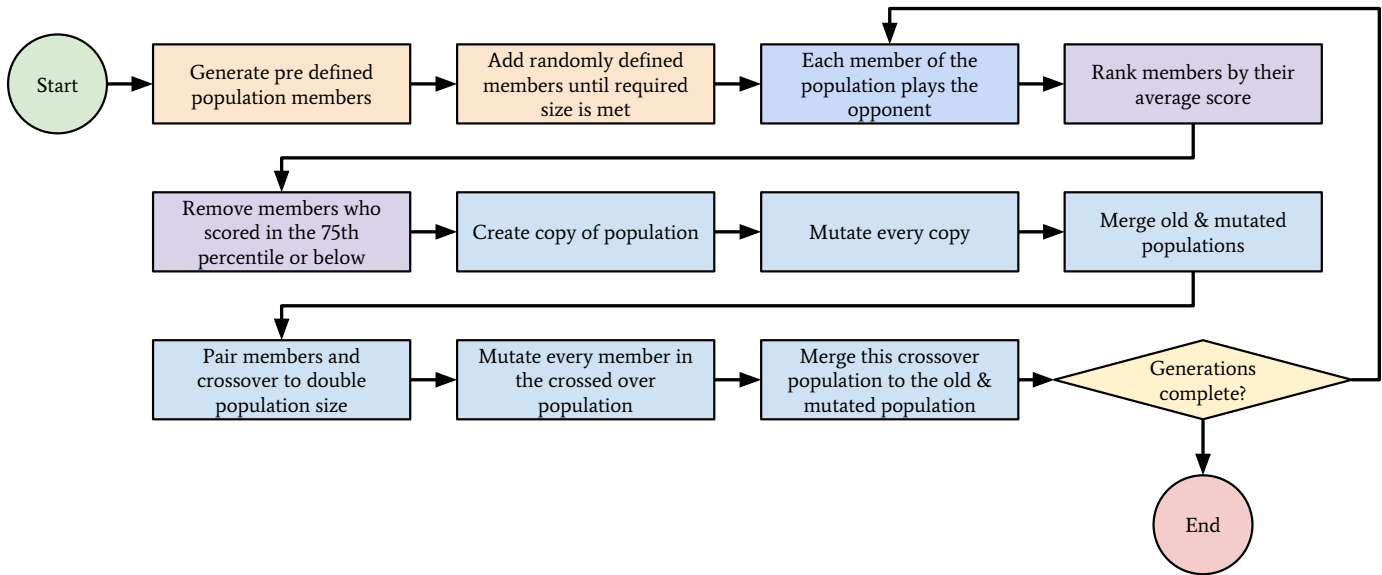


Figure 3.1: High Level Genetic Algorithm Cycle

This model means we can create a population of Cycler players and input a sequence of length 200 as a parameter to set off our genetic algorithm. The subsequent inputs for the populations Cycler players will be created using the genetic mutation and crossover techniques, see Section 1.2.1 for details.

The flow of Play-Rank-Create looping (see Figure 3.1) will be the basis of creating the optimal strategy for each other opponent. Each step is defined in the `Axelrod Dojo Population` and `CyclerParams` classes. Rather than store all the functionality in one place we are able split up aspects of the flow to allow for flexibility in what type of population can be used.

²This algorithm had a bug which would change the size of the population in the first generation. This was fixed after Section 5.7 was written. See: <https://github.com/Axelrod-Python/axelrod-dojo/issues/43>

Chapter 4

Developing The Codebase

For the analysis a mix of Jupyter Notebooks and integrated development environments¹ were used to write and execute code. The main analysis was run using a factory class, written by me, called ‘full_analysis’. This class was used to wrap a query to the Axelrod-Dojo functionality and subsequently to the Axelrod Library in such a way that was easy to control batch executions.

The analysis itself was done using native multi-threading on a Linux OS to improve individual opponent analysis run times and the overall scalability of the project. This will be discussed further in Subsection 4.2.

4.1 Libraries And External Modules

Main Research Libraries

- **Axelrod** — Used for the core of the prisoners dilemma and iterated prisoners dilemma functionality code.[pd17]
- **Axelrod-Dojo** — Applied machine learning techniques that revolve around generating solutions to questions relating to the Axelrod library.

Functional libraries Table 4.1 shows the external functional libraries used, while Table 4.2 shows the internal python built in modules that were leveraged during development. These are libraries which are not involved in the core functionality of the IPD.

Library	Reason
matplotlib pyplot	For plotting graphs and images with data
pandas	For data manipulation.
numpy	For reducing complexity of numerical calculations.

Table 4.1: Functional Python libraries for analysis

¹Pycharm Professional Edition and Microsoft VS Code.

Library	Reason
os	For operating system functionality.
time	For time calculations.
itertools	For easier iterations over data structures.

Table 4.2: Internal built in python modules used

4.2 Setting Up A Research Environment

This section will go into detail on the set up and reasoning behind using specific development environments. This tutorial will also assume you're working with a local development station; analysis on remote cloud instances of Jupyter Notebooks is possible but the set up is different.

Installing Basic Libraries Your first step should be to download and install the Anaconda distribution for your OS here: <https://www.anaconda.com/download/>. This will allow you to use the integrated c++ libraries python has to offer without needing to mess about too much. Anaconda also has them majority of Functional Libraries above and Jupyter Notebooks pre installed to make setting up much easier. From here, follow the instructions the install wizard has to add any environment variables to allow CMD/Bash access to binaries.

Installing the Axelrod and Axelrod-Dojo libraries uses the pip tool that already comes with Anaconda and should be ready to execute after the last step. Running 'pip install axelrod' then 'pip install axelrod-dojo' will install these.

Once this is installed the `full_analysis.py` file has to be downloaded from github², it can be found in the code directory. This can just be copied and pasted if needed all were interested in is the class to generate a sequence for an opponent.

Running a Test Figure 4.1 is some sample code that will run an analysis with the following settings

- Override the default mutation frequency of 0.1 to 0.3.
- Set the prefix for all the files to be 'example-'.
- Analysing 3 opponents. (Random will have multiple instances for different seeds.)

```
from full_analysis import NewAnalysisRun
import axelrod as axl

run = NewAnalysisRun(mutation_frequency=0.33)
run.save_file_prefix = "example-"

run.add_opponent(axl.TitForTat())
run.add_opponent(axl.Random())
run.add_opponent(axl.Grudger())

run.start()
```

Figure 4.1: Code to create a sequence result to optimise best score for 3 opponents

After it has run the output should be stored in the './output' directory. If the code fails to run there may be issues with this directory being created. There should be multiple output files, each with one opponents evolution stages through the generations.

²<https://github.com/GitToby/FinalYearProject>

Cloud Notebook Setup If you want to use a Cloud service, such as Azure Notebooks or AWS Sagemaker, the set up is similar to the above just executed differently. Installing Anaconda is not needed, the environment has the required installs. Using pip to install the Axelrod libraries and download the full analysis script can be done in an integrated web terminal or directly in a notebook. Figure 4.2 shows an example in azure, copy these in your top few cells of your jupyter notebook and it will work as required.

```
# ----- CELL 1 -----  
# The ! means 'run this as a bash script'  
! pip install axelrod  
! pip install git+https://github.com/Axelrod-Python/axelrod-dojo.git  
! wget https://raw.githubusercontent.com/GitToby/FinalYearProject/master/code/full_analysis.py  
  
# ----- CELL 2 -----  
import axelrod as axl  
import full_analysis as fa  
  
run = fa.NewAnalysisRun(population_size=40)  
  
run.add_opponent(axl.TitForTat())  
run.add_opponent(axl.Random())  
run.add_opponent(axl.Grudger())  
  
run.start()
```

Figure 4.2: Cells for creating the jupyter instance of a research environment

4.3 Codebase Contributions

Throughout the project I had split time between writing my own code and expanding the Axelrod Dojo codebase. In modern software development there is a commitment in the developer community to track and reuse as much code as possible, typically using a version control system such as Git or Subversion. The majority of open source community development is conducted on github[Tor] using Git, the Axelrod libraries, along with most other libraries I used, are hosted here.

When writing code in a professional environment there is a predetermined scope that all parties agree upon before the work commences. This, however, is not the case for research development which is more organic; final products that are created when conducting research for papers are highly personalised and typically cannot be used in other areas. This leads to more flexible products that operate more as platforms or tools for further research, this is why the Axelrod, Axelrod Dojo and other libraries mentioned in table 4.1 exist. My final code will only be used for researching the IPD in my projects specific direction, this meant extending the platforms to handle what I needed them to do, subsection 4.3.1 looks into how this was completed in my project.

4.3.1 Version Control

During the project I created a ‘fork’ of the axelrod dojo in order to add content to the open source community. This resulted in using Git to create a new ‘branch’ on which to write my code before asking the owners of the repository to pull my work back into the core product using a ‘pull request’(PR). The PR opened for my code³ resulted in changes to 457 lines of code and 14 files, adding classes and fixing bug that had been previously flagged. Figure 4.3 shows the initial scope of the PR and Figure 4.4 shows a section of the commits made before merging the branch.

After a request is opened there is a period of reviews by the owners, this is to ensure code quality and scope coverage. As

³<https://github.com/Axelrod-Python/axelrod-dojo/pull/45>

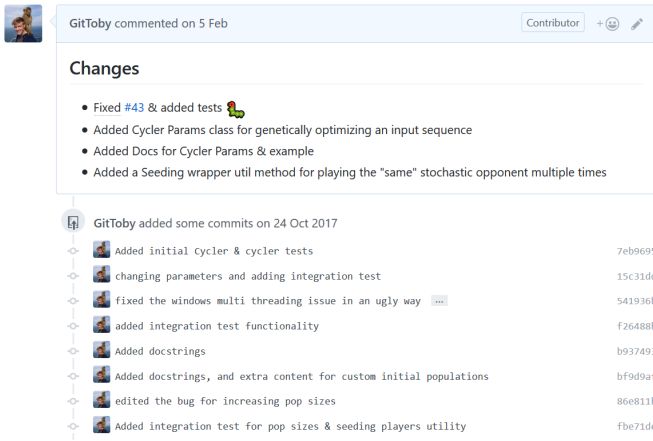


Figure 4.3: Description and commits for PR on Github

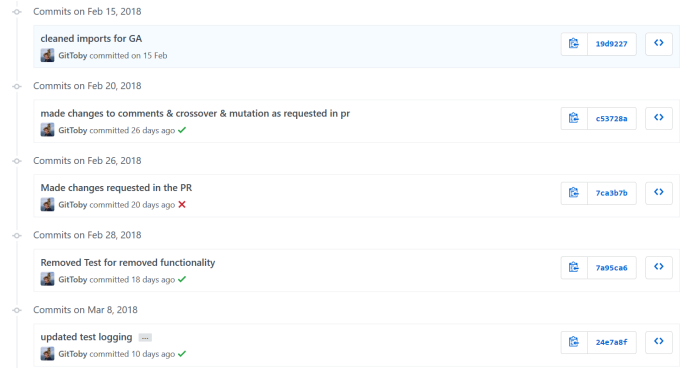


Figure 4.4: Tail of code commit log as shown on Github

my work progressed I continued to add to the PR which lead to more and more requests being added for features that fell outside the scope of the PR. Eventually we decided to create another fork of the PR that contained specialist code for my project that wouldn't benefit the codebase as a whole. Figures 4.5 & 4.6 show examples of requests and discussions around features and code quality. Because of this review process the overall quality of the codebase can be kept high and the owners can decide on how their platforms are developed.

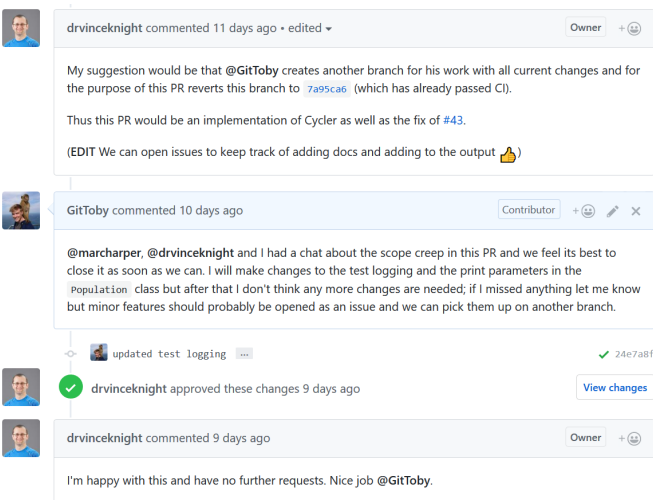


Figure 4.5: Feature discussions in PR on Github

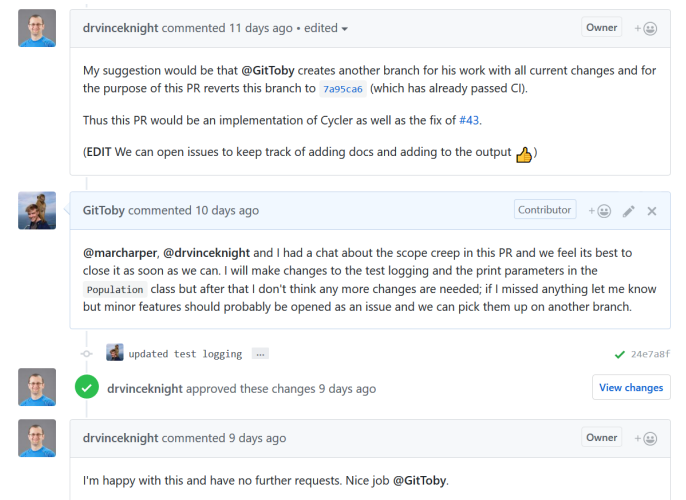


Figure 4.6: Scope Discussion in PR on Github

4.3.2 Testing

Code testing and version control go hand in hand. When new releases of code are made it is important to ensure that the new changes dont break previous functionality. This is kept in check by the presence of continues integrations (CI) and test environments. As of this project the CI for the Axelrod Dojo keeps the libraries tests running on a linux environment and the output fed to the Github page. During my development I had to implement tests for any functionality I wrote to ensure the library still worked as intended. Following the mixed practice of test driven design (TDD) and behaviour driven design (BDD) the code written to extend the Axelrod Dojo had tests to cover examples of what happens in a production environment. Snippet 4.7 shows an example of a test I wrote.

```
def test_creation_seqLen(self):
    axl.seed(0)
    test_length = 10
    self.instance = CyclorParams(sequence_length=test_length)
    self.assertEqual(self.instance.sequence, [D, C, C, D, C, C, C, C, C, C])
    self.assertEqual(self.instance.sequence_length, test_length)
    self.assertEqual(len(self.instance.sequence), test_length)
```

Figure 4.7: An Example of a test in the Axelrod Dojo

Chapter 5

Implementation Of Sequence Discovery

In this chapter, consideration will be given to the process of finding the optimal sequence of moves against another player. The various approaches used and a detailed analysis of the optimisation procedures and parameters will be described.

Before conducting the bulk calculations for the set of opponents listed in the appendix we will test values for 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 mostly 'simple' and can be explained in a sentence or two, but each one has a fundamentally different structure to how they work. We will look into these as we can confirm that the genetic algorithm will select the optimal sequence solution for the selected type of opponent.

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 per turn hits a maximum such that it wont change no mater how many more generations are run; as described in Section 3.2 this the optimal solution sequence and it is unique. NOTE: unique solution sequences only exist for non stochastic opponents.

Once a solution has been found the generation number where this plateau occurs is called the solution sequence distance, or solution distance; one of the goals of this initial investigation is to see how parameters affect the distance. 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 wont find the best sequence that will return can possibly get. These sub optimal solutions are due to the occurrence of local maxima in the set of scores of neighbouring sequences. Genetic algorithms are designed in way to avoid local maxima's; the property of mutating (i.e jumps of their features) allow members of the population to potentially remove themselves from these local maximas. We will look in depth into how to overcome the

Player	Optimal Sequence	Representation
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()	TODO need to find	
axl.CollectiveStrategy()	TODO need to find	
axl.Champion()	TODO need to find	

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
...

Table 5.1: Output data table

possibility of our algorithm finding a local, rather than, global maximum in Section 5.4. Some of the questions we will hope to be answering include:

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

5.1 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 3.1; 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 distinct new sequences after mutation. 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 requires us to understand how the solution distance is affected as we run our algorithm through a set of population sizes, say $|P| \in [25, 50, 100, 150, 200, 250, 500]$.

EFFICIENCY 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 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 Snippet 5.1 is an implementation how we go about analysing and storing the tests on generation sizes listed. It leverages the use of the function ‘runGeneticAlgo’ show in appendix Snippet ???. This code will output data in the form of Table 5.1.

By grouping this data by the population we observe how initial populations affect different opponents. Its clear that from figure 5.2 that the initial population size has a significant effect on finding better sequences. We can see if there is a larger

```

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_one, 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 5.1: code to check multiple populations

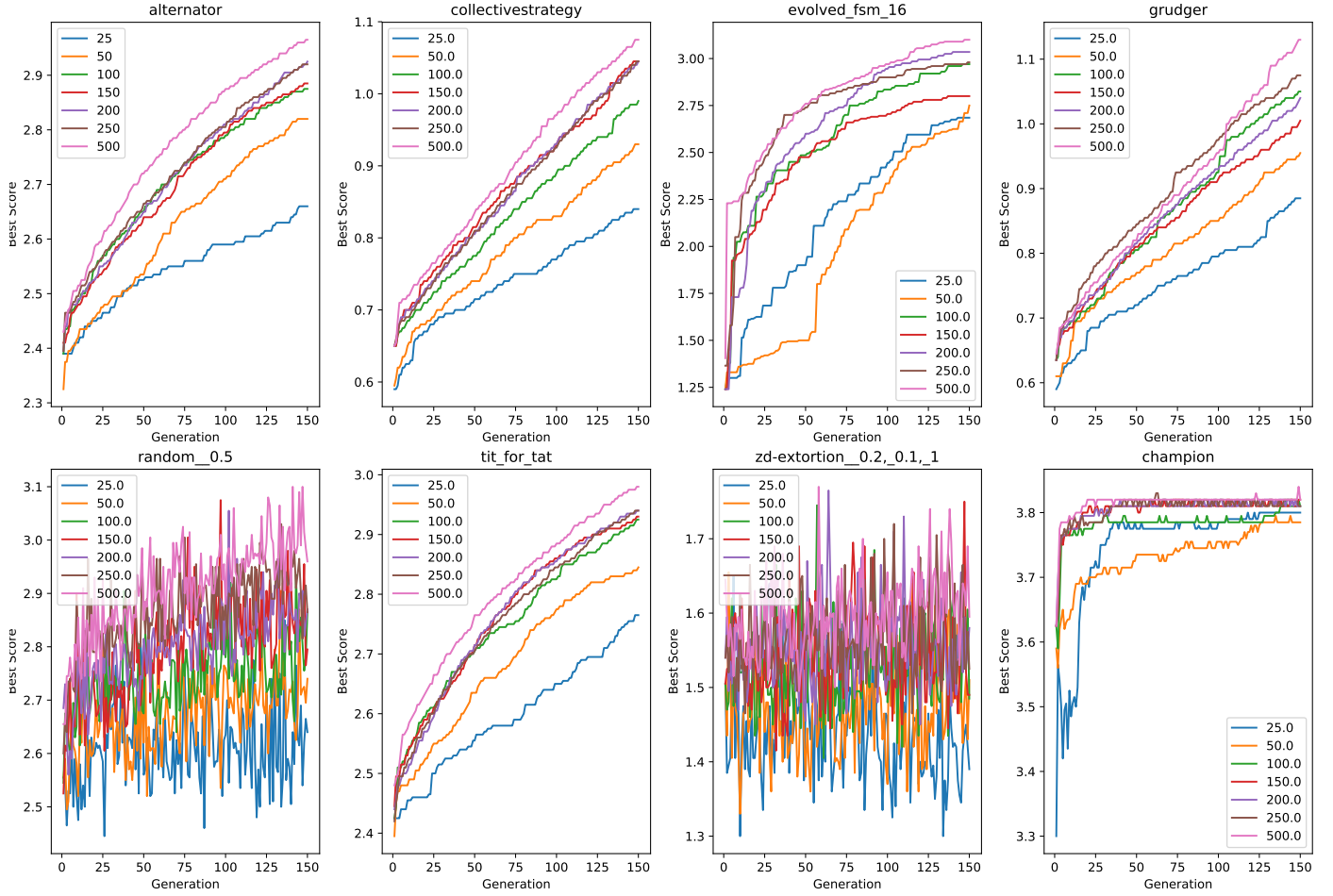


Figure 5.2: Best score per turn vs generation for different initial population sizes

initial population there is typically a higher best score shown once concluding all of the generations. This can also be shown in figure 5.3 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-linear 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.

5.2 Generation Length Analysis

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

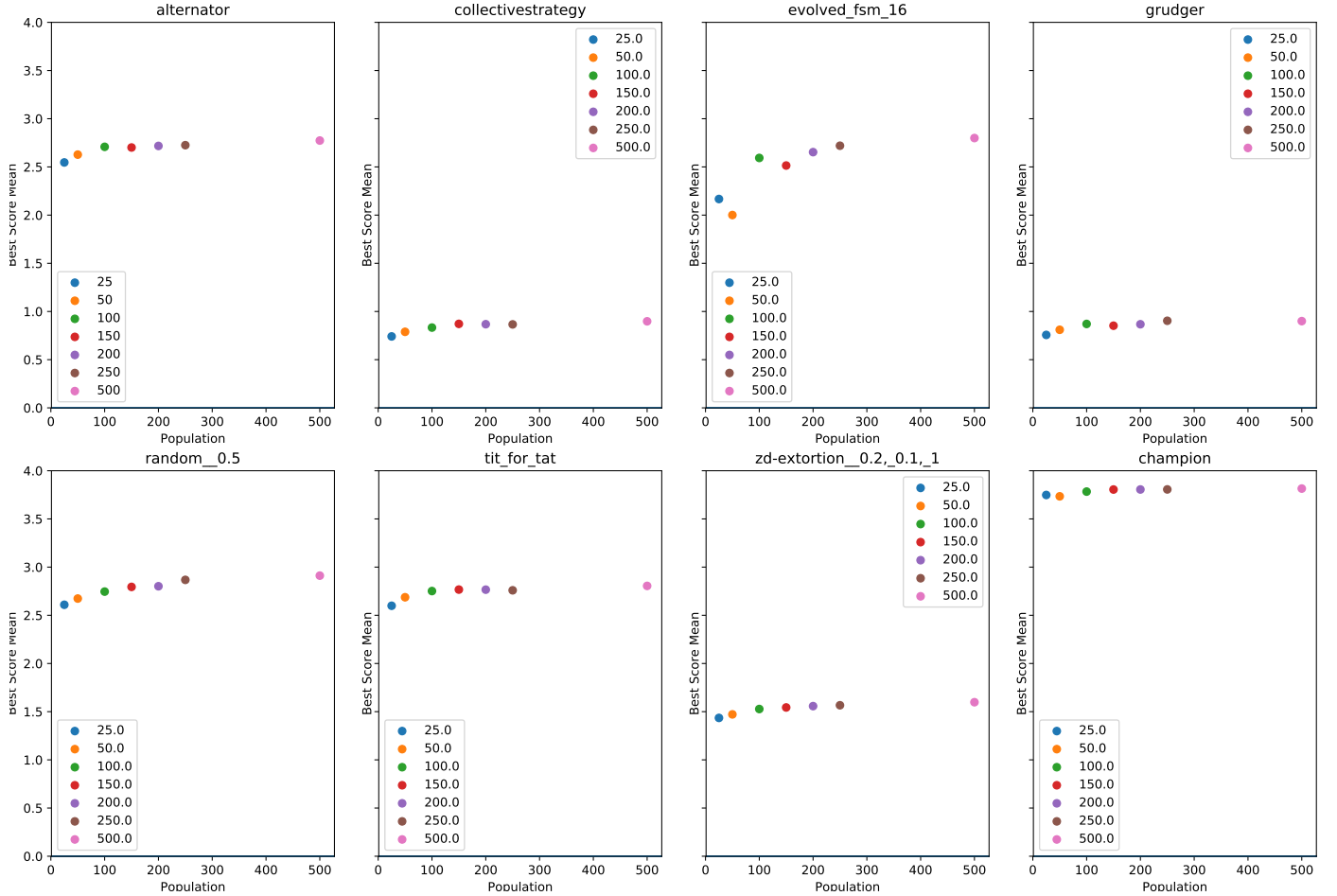


Figure 5.3: Scatter of max best score vs different initial populations

- The total combinations of features, (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 set of sizes for our parameter to run our analysis for each size, here we will take generation lengths $G \in [50, 150, 250, 350, 450, 500]$.¹ The code in Snippet 5.4 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 figure 5.5, mean best score difference against the number of generations, 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; when we are close to a maximum it is more difficulty to randomly select which element in the sequence needs changing to improve our score. On this result we can conclude as we increase generations there is less and less benefit per generation. There is, however, still a benefit to extending the generations but we may have better performance by altering another parameter of the algorithm. There may be a

¹We will be using a population of 150 as this was the best average for score vs computation time for analysis.

```

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 5.4: code to check multiple generation lengths.

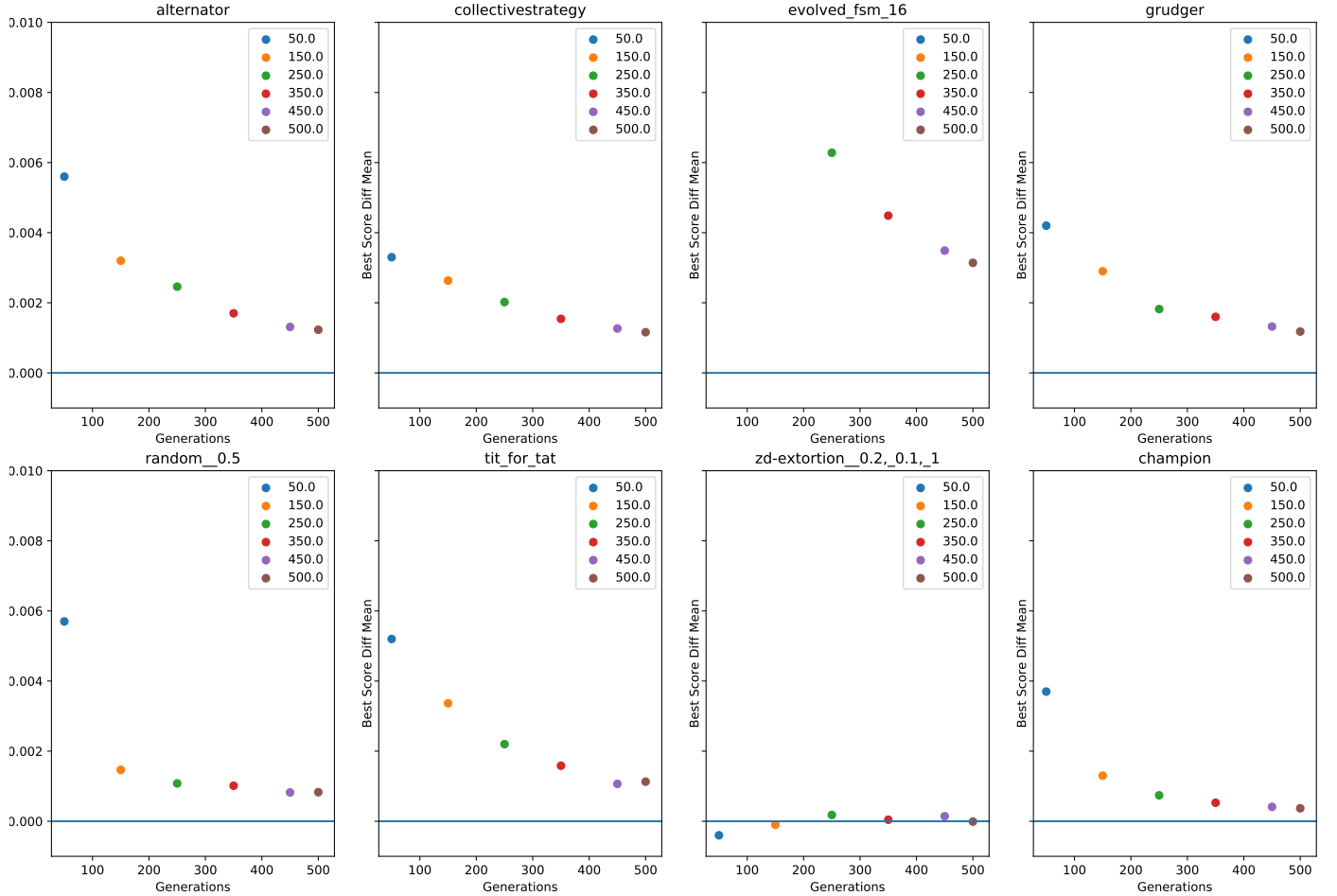


Figure 5.5: Mean Best Score diff vs total generation lengths

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. The probability of finding a solution as we narrow in on a maximum decreases due to the number of elements that, when changed, will provide a better score. Increasing the mutation frequency at this point means that there will be more members of the population that could potentially mutate the elements needed to improve the sequence.

Figure ?? 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.

After 250 generations we *seem* to have reached a solution state for our opponents Tit for tat and alternator but not for grudger. Against Grudger we see an example, we have only reached an average score per turn of ____, which is obviously far from its optimal sequence. From the combination of the plots, having more generations means that there is, on average, less of an improvement per generation. It is clear that a higher number of generations is required to find a better solution sequence for an opponent. From now on, 250 is the number of generations we will use to find our solution sequence during the analysis.

For most of the opponents 250 generations seems reasonable to reach a solution sequence as shown in the Alternator and Tit For Tat. However, there are clear signs of local maximums occurring in the Grudger example. Figure ?? has reached a better sequence in 450 generations than 500²; meaning that increasing the generation length doesn’t necessarily mean

²These are independent trials and have different sequences.

```

def mutate(self):
    # if the mutation occurs
    if random.rand() <= self.mutation_probability:
        mutated_sequence = self.get_sequence()
        for _ in range(self.mutation_potency):
            index_to_change = random.randint(0, len(mutated_sequence))
            # Mutation - change a single gene
            if mutated_sequence[index_to_change] == C:
                mutated_sequence[index_to_change] = D
            else:
                mutated_sequence[index_to_change] = C
        self.sequence = mutated_sequence

```

Figure 5.6: The mutation code as given in the axelrod-doj

a local maximum. The complexities with local maximums during the generations lie with mutation rates and crossovers. We will cover this in Section 5.4

In this investigation we will want to find the optimal solution and so, from these results, we will want to extend the generation length as far as possible. An infinite number of generations would be preferable, but we don't have an eternity so a selection of a relatively large generation size will be adequate when coming to the final series of tests.

5.3 Changing Mutation Rate

By changing the way in which we mutate our elements within a sequence, we might be able to more effectively narrow in on an optimal solution sequence. The default settings are a frequency of 0.1; meaning for every 10 members of our population that continue into the next generation one of these has some elements in its sequence changed, and a potency of 1; meaning that every sequence that was altered only has 1 element altered. Here we will look into these 2 different concepts and see how they might improve our distance to an optimal sequence and whether we can escape local maximums.

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

These are two separate questions, so first we will look at increasing the potency of our mutation. Once we have found some information on how this effects our solution, we can look into the frequency of our mutations with the new potency as a permanent setting. As shown further on, there is not much of an improvement on our algorithm to changing either of these. The mutation algorithm is shown in Snippet 5.6

EFFICIENCY NOTE: This approach allows for an $O(1)$ factor of scaling. This makes changes in mutation a great candidate for an approach to reduce our solution sequence distance compared with other approaches, for example increasing the population size.


```

def mutationPotencyChecker(opponent):
    file_name = "data/" + str(opponent) \
        .replace(" ", "_") \
        .replace(":", "_") \
        .lower() + "_mutation_potency.csv"
    if not os.path.isfile(file_name):
        df_main = pd.DataFrame(data=None, columns=col_names)
        for potency in mutatuon_potency_list:
            start_time = time.clock()
            pot_run = runGeneticAlgo(opponent,
                                    population_size=150,
                                    number_of_game_turns=200,
                                    cycle_length=200,
                                    generations=250,
                                    mutation_probability=0.1,
                                    mutation_potency=potency,
                                    reset_file=True)

            end_time = time.clock()
            tmp_df = pd.read_csv(pot_run[0], names=col_names)
            tmp_df["mutation_potency"] = potency
            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 5.7: Mutation potency code

5.3.1 Changing Mutation Potency

Changing the potency of the algorithm will mainly generate the noise in our sequence generation to generation, increasing the distance³ between the mutated sequence from the original.

This potentially could create an algorithm that is too ‘jumpy’ for narrowing in on a solution. We can imagine a sequence as a vector in 200 dimension space then a mutation for element S_i is the same as changing the vector in its i^{th} dimension. Shortening this example to a vector in 3 dimensions (or a sequence of length 3) then a mutation is much more easily visualised. It is clear that a mutation potency should be kept low as to keep consecutively mutated sequences more similar; we will only be looking at mutating our sequences at up to 10 percent of their elements. We will look into having mutation potencies $M_p \in [1, 2, 3, 5, 10, 15, 20]$

Using the data generated from Snippet 5.7 of code we are able to look at how our best score and our best score diff is affected as we increase the number of positions.

³Distance concept has been taken from coding theory; $d(s_1, s_2)$ = the number of differing positions between 2 sequences s_1 and s_2 . For example: $d(111, 110) = d(CCC, CCD) = 1$

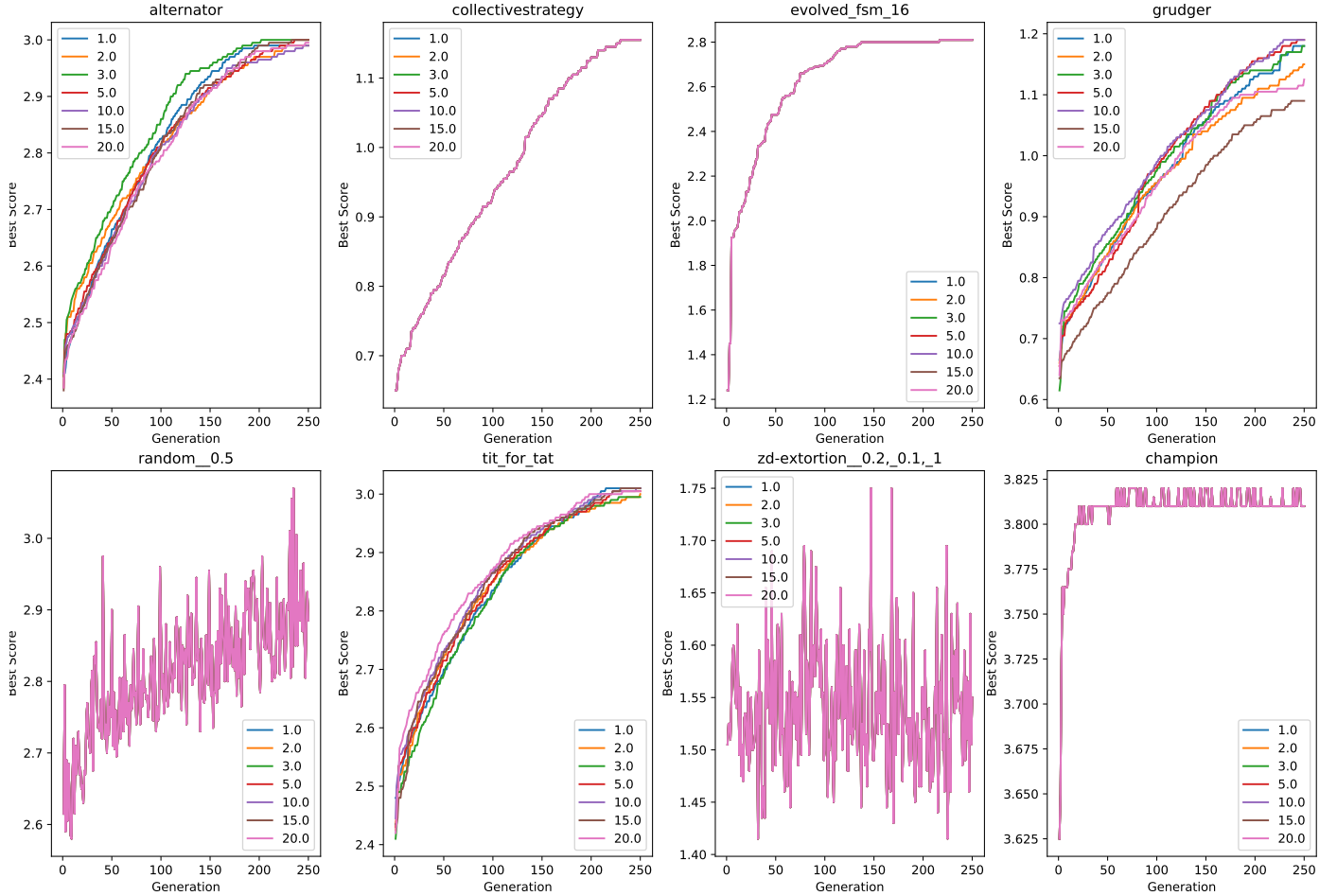


Figure 5.8: Best score vs generation for different mutation potencies

Figure 5.8 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 changing only 1. This may be down to chance (and if the test is rerun this may disappear), however looking at more opponents than just Grudger we find there is no clear benefit to increasing the mutation potency with respect to 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 looking at how our average best score difference changes as we increase the mutation potency there is no sign that there is a significant improvement to our sequence. The increase in mean best score difference is not substantial and could be down to chance.

5.3.2 Changing Mutation Frequency

In contrast to changing the mutation potency, increasing the frequency should allow us to generate more unique sequences generation to generation. We will look at what happens when we run the genetic algorithm on a set of mutation frequencies $M_f \in [0.1, 0.2, 0.3, 0.4, 0.5]$. The code in Snippet 5.10 shows the algorithm that will complete this analysis.

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 Figure 5.11 we can see that the mutation frequency of 0.1 and 0.2 produced higher scoring solutions than the other

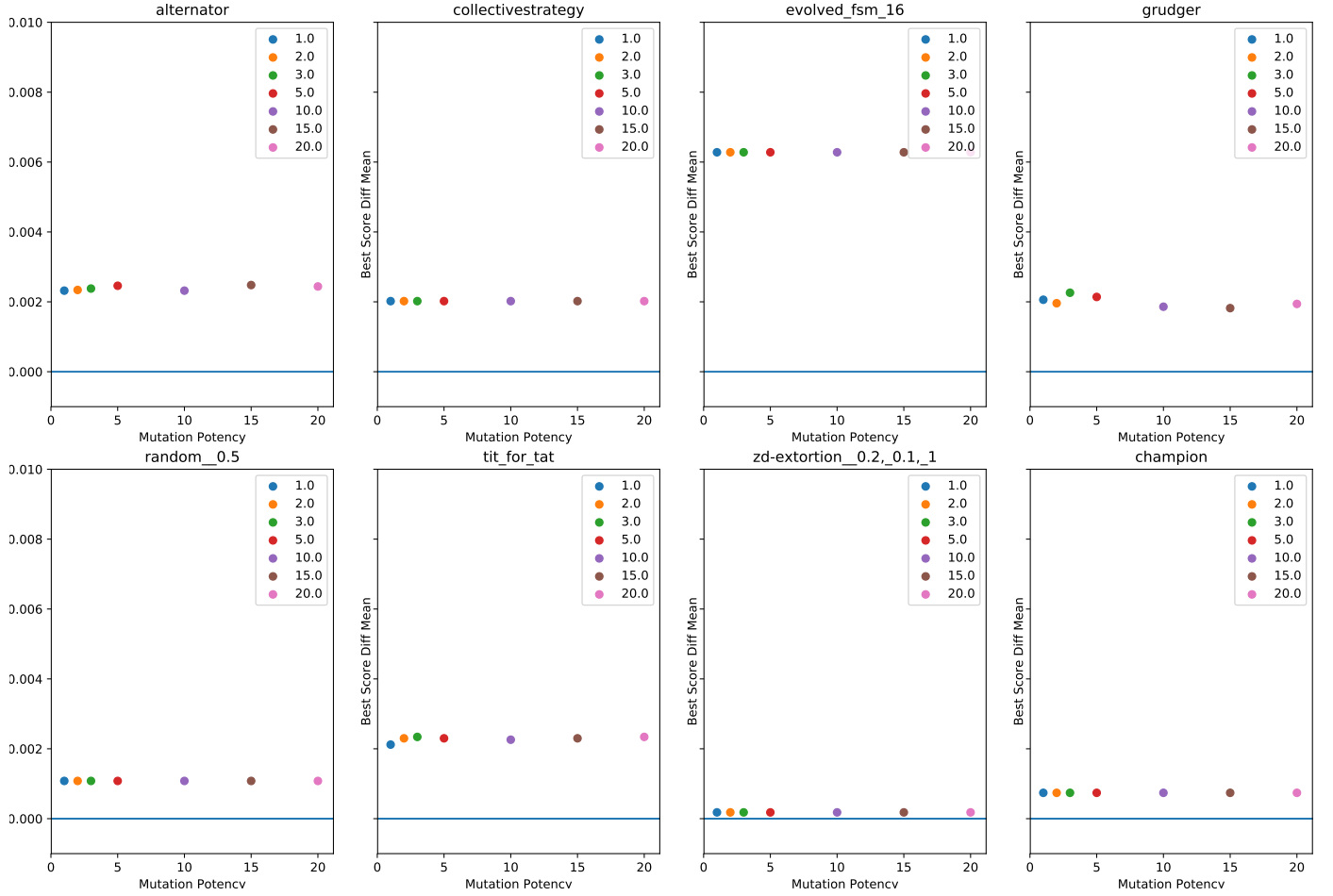


Figure 5.9: Average best score diff vs mutation potencies

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.

5.4 Mitigating local maximum solutions

The occurrence of local maximums is something that has only occurred in sequences against for the Grudger opponent so far. Figure 5.11 shows that there are clearly 2 distinct plateaus that are reached in terms of score, and the overall scores are much lower than the optimal, around 1.4.

The difference between the Grudger and the other opponents have been looking at is that the Grudger has a singularity point where its behaviour changes. The change in behaviour is not uncommon, Tit For Tat works in a similar way and in both cases our algorithm has managed to identify this behaviour and adapt to overcome its negative effects.

Grudger is an opponent which it is possible to attain a local maximum score and be ‘trapped’ in this solution sequence. From the output sequences below it appears that the sequence is being changed into becoming 2 sections of opposing moves. If we look at a random start sequence and then the end sequence after 250 generations we can see that the genetic

```

def mutationFrequencyChecker(opponent):
    file_name = "data/" + str(opponent).replace(" ", "_") \
        .replace(":", "_") \
        .lower() + "_mutation_frequency.csv"
    if not os.path.isfile(file_name):
        df_main = pd.DataFrame(data=None, columns=col_names)
        for freq in mutation_frequency_list:
            start_time = time.clock()
            pot_run = runGeneticAlgo(opponent,
                                    population_size=150,
                                    number_of_game_turns=200,
                                    cycle_length=200,
                                    generations=250,
                                    mutation_probability=freq,
                                    mutation_potency=1,
                                    reset_file=True)

            end_time = time.clock()
            tmp_df = pd.read_csv(pot_run[0], names=col_names)
            tmp_df["mutation_frequency"] = freq
            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 5.10: Mutation potency code

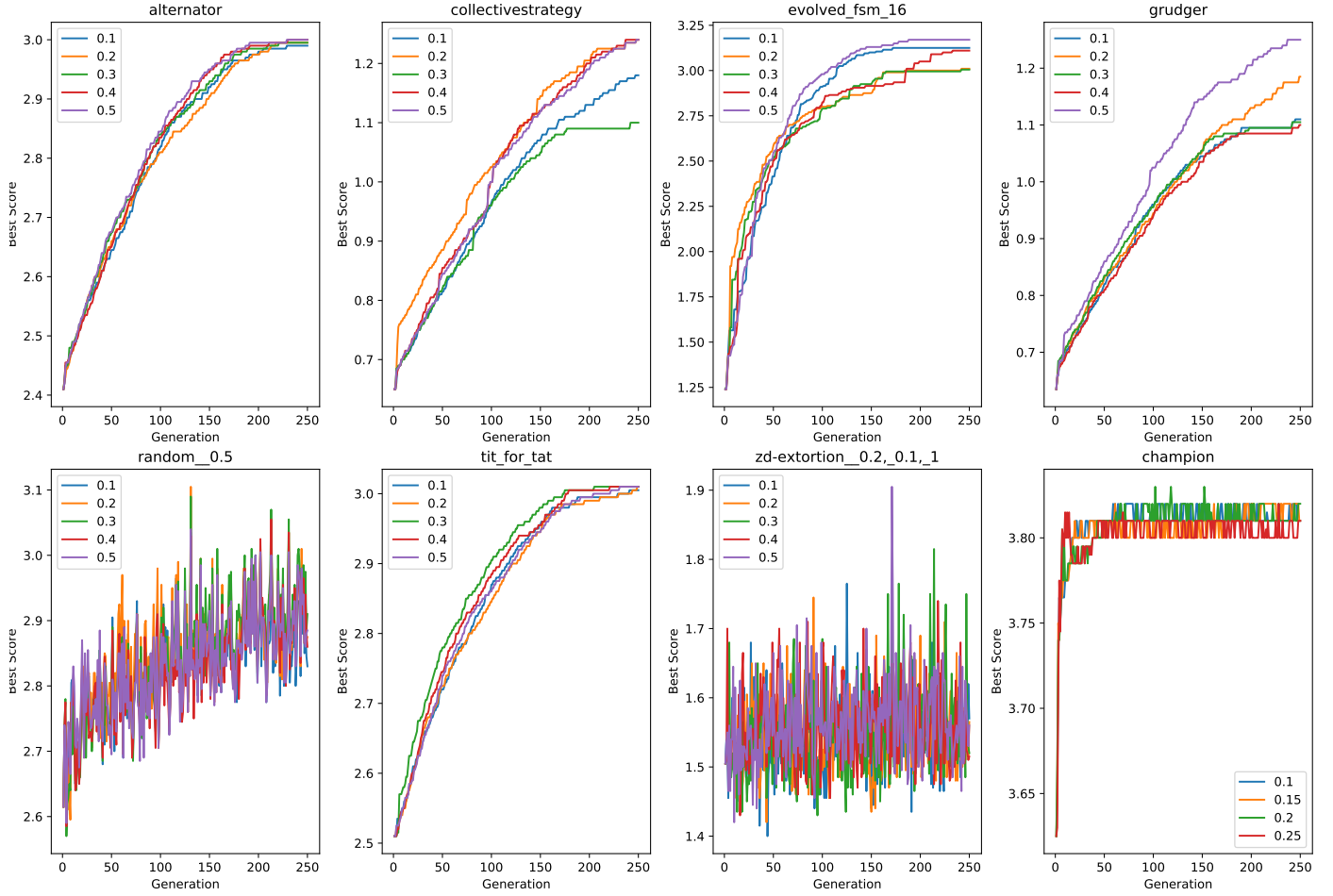


Figure 5.11: Best score vs generation for different mutation frequencies

algorithm is learning defect after some point and cooperate before. This is due to the fact a good solution will end in lots of defections after it has already defected, but cooperate beforehand to avoid this harsh behaviour.

Grudger best start: CCCCCDDDD... (No pattern is obvious.)

Grudger best end: $C : [22, 178]$

We may be able to identify a way to mitigate this singularity effect by looking at the difference between 2 similar solutions. Grudger and Tit For Tat differ in their strategies in two ways:

- Grudger never changes its mind. There is one change in behaviour for the entire game, unlike Tit For Tat. The algorithm then only has a single opportunity to observe this per population per generation meaning the behaviour is much less frequently encountered.
- Grudger will not become 'smart' again. This means that the move genetic algorithm picks up the effect of a single defection in its sequence it starts playing a 'dumb' opponent. A random start of Cs and Ds puts the likelihood of at least 1 defection occurring in the first 10 moves at above 99.99%. This swap from 'smart' to 'dumb' will, most likely, always occur in the first 10 moves; the only way of extending the 'smart' player is to add a cooperation to the end of the starting Cs.

Snippet 5.13 shows two totality games against grudger, 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 random initial population. Because of this, the algorithm has

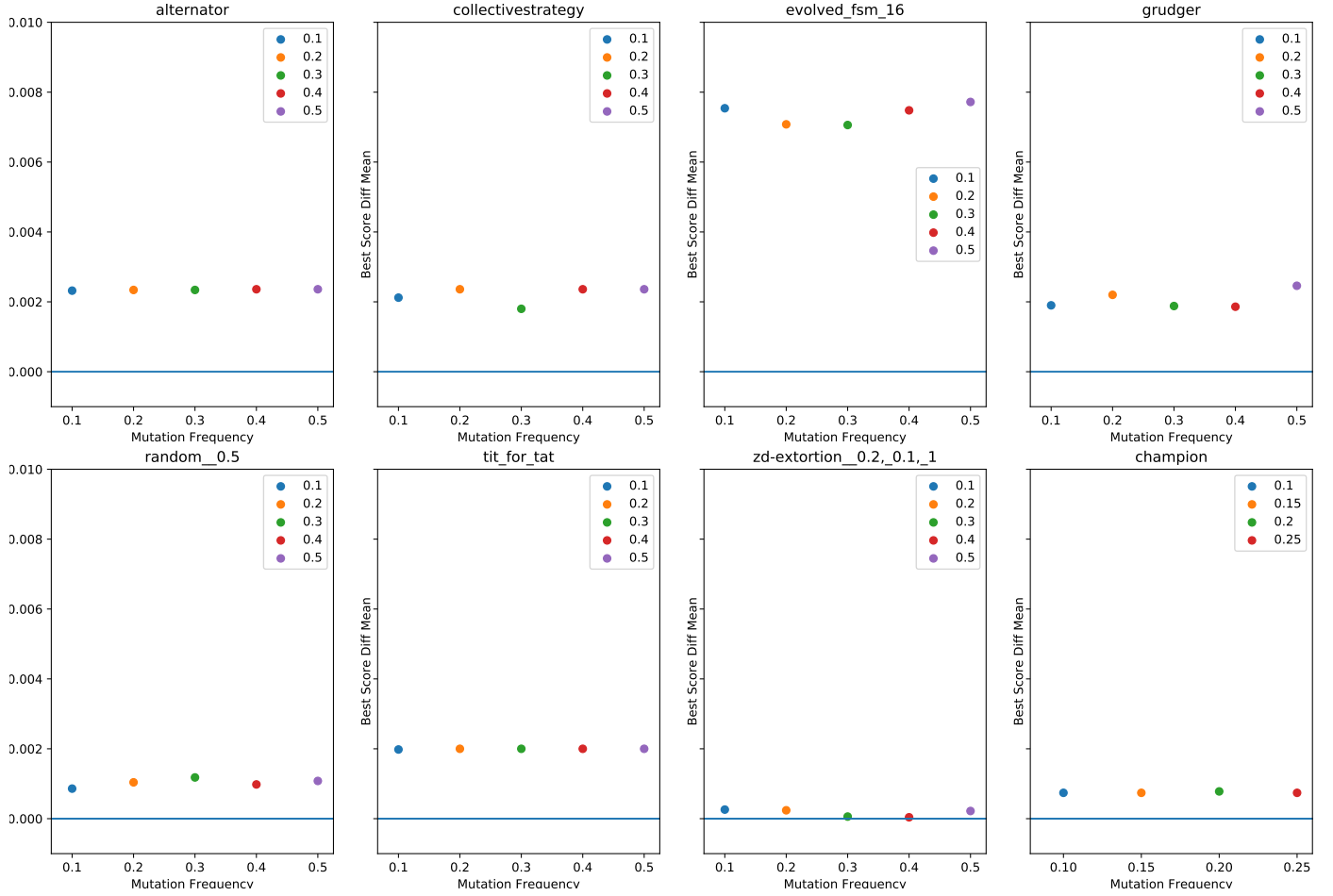


Figure 5.12: Average best score diff vs mutation frequencies

to shuffle towards the potential benefit of using these totalities rather than start with analysing them. Our case against grudger requires the algorithm to attempt this shuffle towards a totality after encountering the grudging effect. This would then require the algorithm to select (using random selection out of 200) the first defection move and change it (only 1 in 10 members are mutated) to a cooperation move. This likelihood is incredibly small, and starting at common or uniform sequences would be more beneficial, as posed in Section 5.5.

As Snippet 5.13 shows, we should be converging on a totality of Cs rather than what its doing; 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.

5.4.1 Ineffective Approach of Altering Crossover And Mutation

The process of converging to Ds when building a solution against grudger then sheds light on the process the algorithm takes to find a solution. If we are to find the optimal solution sequence, we must take a crossover and mutation path which doesn't cut off better paths as we work our way towards good solutions; 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

```

players = (axl.Grudger(), axl.Cycler("C"))
match = axl.Match(players, 200)
match.play()
print("final scores:", match.final_score())
print("per turn:", match.final_score_per_turn())

# >final scores: (600, 600)
# >per turn: (3.0, 3.0)

players = (axl.Grudger(), axl.Cycler("D"))
match = axl.Match(players, 200)
match.play()
print("final scores:", match.final_score())
print("per turn:", match.final_score_per_turn())

# >final scores: (199, 204)
# >per turn: (0.995, 1.02)

```

Figure 5.13: Grudger matches against totalities

```

def crossover_old(self, other_cycler):
    # single point crossover:
    crossover_point = int(self.get_sequence_length() // 2)
    # get half 1 from self
    seq_p1 = self.get_sequence()[0: crossover_point]
    # get half 2 from the other_cycler
    seq_p2 = other_cycler.get_sequence()[crossover_point: other_cycler.get_sequence_length()]
    crossed_sequence = seq_p1 + seq_p2
    return CyclerParams(sequence=crossed_sequence)

```

Figure 5.14: Old Crossover algorithm

large subsections of the sequence then allow the mutations to optimise these sub-sequences.

Currently we have the following design:

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:

$$\begin{array}{l}
 | \text{-----} | \text{and} | \text{+++++} | \\
 = | \text{-----} + \text{+++++} |
 \end{array}$$

to the mixture:

$$\begin{array}{l}
 | \text{-----} | \text{and} | \text{+++++} | \\
 = | \text{--} + \text{--} + \text{--} + \text{--} + \text{--} + \text{--} + \text{--} + \text{--} + \text{--} + \text{--} |
 \end{array}$$

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 shown in Figure 5.15. As shown, the algorithm splits the two sequences into 10 section and the new sequence is formed from Alternating sections. Figure 5.16 shows an example of the new crossover method.

```

def crossover(self, other_cycler):
    # 10 crossover points:
    step_size = int(len(self.get_sequence()) / 10)
    # empty starting seq
    new_seq = []
    seq1 = self.get_sequence()
    seq2 = other_cycler.get_sequence()
    i = 0
    j = i + step_size
    while j <= len(seq1) - step_size:
        new_seq = new_seq + seq1[i:j]
        new_seq = new_seq + seq2[i + step_size:j + step_size]
        i += 2 * step_size
        j += 2 * step_size
    return CyclorParams(sequence=new_seq)

```

Figure 5.15: New Crossover algorithm

```

seq1 = [1, 1, 1, 1, 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, 0, 0, 0, 0]
step_size = int(len(seq1) / 10)
i = 0
j = i + step_size
new_seq = []
while j <= len(seq1) - step_size:
    new_seq = new_seq + seq1[i:j]
    new_seq = new_seq + seq2[i + step_size:j + step_size]
    i += 2 * step_size
    j += 2 * step_size
print(seq1)
print(seq2)
print(new_seq)

# >[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
# >[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
# >[1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0]

```

Figure 5.16: Example of new crossover algorithm

We will look at how this new crossover algorithm works with the default mutation (freq=.1 and pot=1) to improve our local maximums with the selection off opponents.

There was no noticeable improvement from introducing this change of function. Each players average score per turn was in the expected score if the algorithms crossover method was unchanged. The problem of mitigating sub optimal solutions may be more efficiently solved using a predefined population. Section 5.5 discusses this in more detail.

5.5 Altering Initial Population

Altering the initial population allows us to select starting sequences that fit patterns we know will have good results. For example totalities, with heads/tails, are usually very effective against simple opponents; for example Tit For Tat, ‘dumb’ opponents or Grudger. Alternating and certain known solution optimal sequences are good starting sequences for an initial population, allowing a more intelligently distributed set of starting points for the random mutation process.⁴

This section discusses the results of working on an initial population that contain common solution sequences. We will start by creating a population of ‘neat’ starting members then allow the entropy of the genetic algorithm to alter these sequences. Deciding where to start our algorithm may mitigate potential sub-optimal solutions by reducing the distance between the starting sequences and optimal solutions. The list of ‘neat’ starting points are stated below:

Totalities

- $C : [200] — 1$ sequence

Single Change Sequences

- $\{C : [i, 200 - i]\} \quad i \in [1, 10], i \in \mathbb{Z} — 10$ sequences
- $\{C : [200 - i, i]\} \quad i \in [1, 10], i \in \mathbb{Z} — 10$ sequences

Matching Tail Sequences

- $\{C : [i, 200 - (i + j), j]\} \quad i, j \in [1, 5], i, j \in \mathbb{Z} — 25$ sequences

Alternating

- $C : [(i, i)^{100/i}] \quad i \in \{1, 2, 4, 5\}, i \in \mathbb{Z} — 4$ sequences

Handshakes

- $C : [i, j, k, 200 - (i + j + k)] \quad i, j, k \in \{0, 1, 2, 3\}, i \in \mathbb{Z} — 4$ sequences

For each of these sequence sets the inverse, with a defection move start, will also be added to the set of starting sequences. This in total gives 164 (after code 5.18) of sequences, we will then make up the difference to the population limit using a set of random sequences.

We can now look into how this has effected the previous experiments.

⁴This can be visualised as placing balls on a ‘lumpy’ 2d plane to try and find the minimum, starting with an educated guess means we wont get all the balls stuck in one vally which isnt the deepest.

5.5.1 Population Size

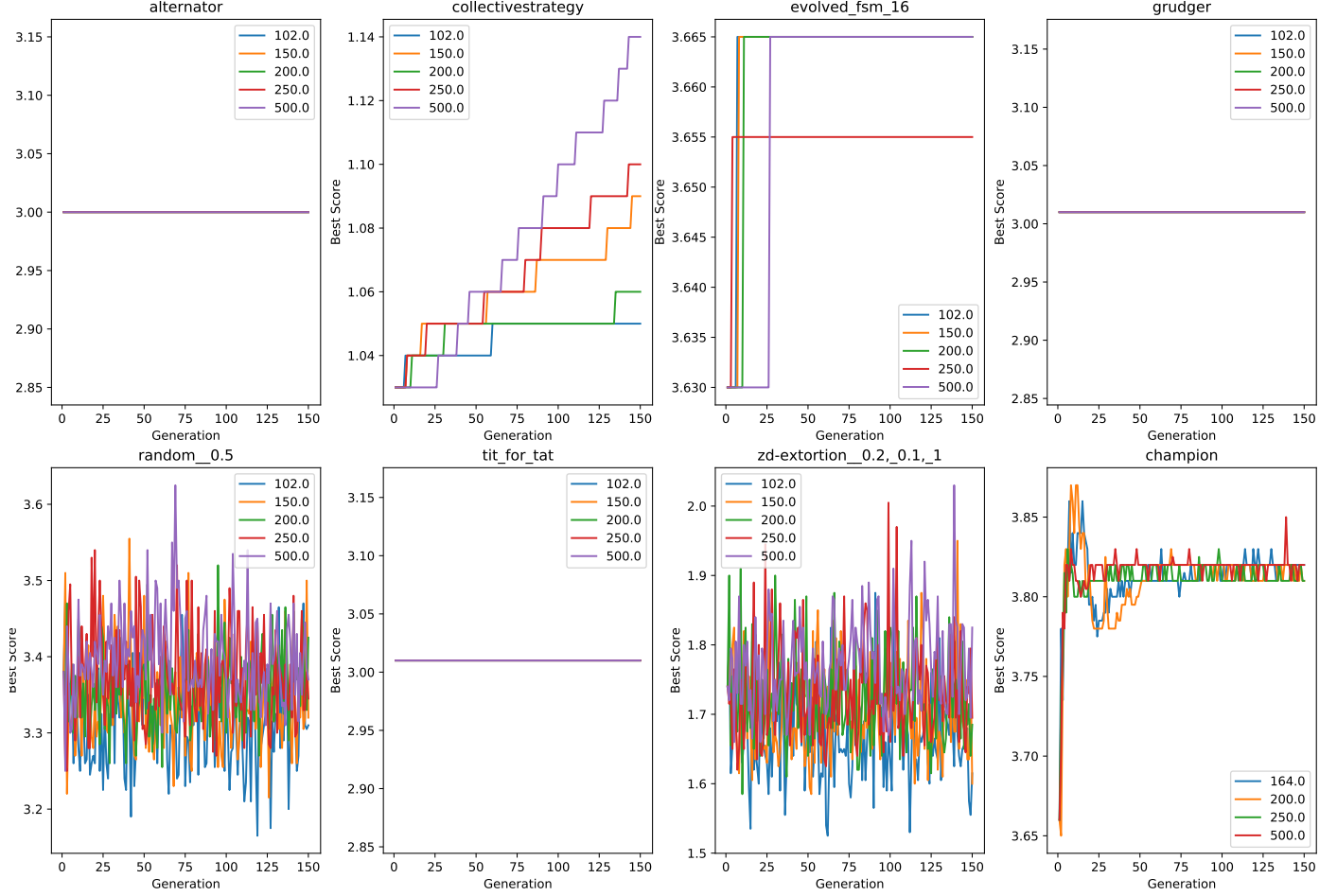


Figure 5.17: Best Score vs generations for preset initial populations on top of random sequences

Figure 5.17 shows the results of using a given initial population of sequences of size 102 along with supplementing this with random members. We will have $|P| \in [102, 200, 250, 500]$. The Collective Strategy run in figure looks as if there are much better results with an initial set population for the algorithm. Of the two that are not showing straight lines we can observe that the algorithm has found the optimal sequence for all but Random & Collective Strategy. This is probably due to 2 different reasons:

Random: The ‘dumb’ strategy should be beaten with a totality of D, the algorithm has in fact converged to almost this totality, but still has intermittent Cs. The reason for this is the scoring grade ‘score per turn’ will reflect on the number of intermittent Cs in the Random players sequence. More Cs in the Random sequence will allow the algorithm to score more. This leads to a solution sequence containing some random Cs not because they score better in some turns, but because the totality of Ds played against a Random sequence would ‘lose’ to a sequence containing some Cs in because its not a fair trial; the two sequences play against different Random opponent sequences. **Collective Strategy:** as described in chapter 2.3 this strategy is, basically, a handshake + Grudger. If we look into the sequence it selected we can see its found the handshake but then arrives on Grudger. After this encounter the same problem as we had before population selection occurs and the algorithms limits the damage by splitting into Cs then Ds. Solving the collective strategy (and handshakes in general) may be simple; we just put in all the possible n move handshakes followed by totalities and then set to work on the 2nd part of the sequence.

It is clear that having a larger population is good from the old analysis, but the initial population still has to be tweaked to improve its scores against certain handshake opponents. The Random opponent is a special case that requires more analysis to find the absolute optimal. The new initial population, now of size 164, will include all combinations of C & D

of length 5, followed by finishing on all Cs or Ds as shown in the Totalities & handshakes section of Snippet 5.18. This run will also supplement with random members for running test for $|P| \in [164, 200, 250, 500]$

After adding the extra members of the initial population we can see, from Figure 5.19, that now the handshake strategy (such as Collective Strategy) are solved much sooner. Also shown in Figure 5.19 is the Random and ZD extort players, these are examples of Stochastic players which are not finding optimal strategy's. These will be discussed further in Sections 5.5.4& ??.

5.5.2 Generation Length

Figure 5.20 shows the results from changing the number of generations the algorithm will run to $G \in [50, 150, 250, 350, 450, 500]$ whilst also running with a population size of $|P| = 200$.

The Figure 5.20 shows no real improvement from increasing the generations, we will use the results from analysis above sections for constructing the full analysis.

5.5.3 Mutation Potency

Figure 5.21 shows the results from changing mutation potency the algorithm will use to be $M_p \in [1, 2, 3, 5, 10, 15, 20]$ whilst also running with a population size of $|P| = 200$.

The Figure 5.21 shows no real improvement from increasing this variable of the algorithm, we will use the results from analysis in previous sections for constructing the full analysis.

5.5.4 Mutation Frequency

Figure 5.22 shows the results from changing mutation potency the algorithm will use to be $M_f \in [0.1, 0.15, 0.2, 0.25]$ whilst also running with a population size of $|P| = 200$.

The Figure 5.22 shows no real improvement from increasing this variable of the algorithm, we will use the results from analysis in previous sections for constructing the full analysis.

Discussion

The approach of adding a predefined set of members to the population before running the algorithm was incredibly successful. We have seen that for most opponents the solution sequence were looking for is given almost immediately. This result will mean that when calculating the solution sequence for the main population the predefined set, generated by Snippet 5.18, will be used to shorten analysis times.

Opponents such as Random, ZDExtort and Champion are showing that they're not finding the optimal solution. This is due to the fact these player belong to the class of stochastic opponents. Section 5.6 will look into these in more detail, putting forward possible approaches to simplify 'solving' these opponents.

5.6 Stochastic Opponents

Stochastic opponents form a problem with respect to continuity. The algorithm we're using continues to generate a new opponent sequence for every generation we run. This, essentially, leads to the algorithm playing a new version of the opponent every time, leaving very little opportunity to identify features which can be exploited.

```

def getCyclerParamsPrePop2(pop_size=200, mutation_prop=0.1, muatation_pot=1):
    pop = []
    if pop_size < 164:
        print("population must be 164+," )
        return

    # Totalities & Handshakes
    handshake_leng = 5
    for start in itertools.product("CD", repeat=handshake_leng):
        pop.append(axl_dojo.CyclerParams(list(start) + [C] * (200 - handshake_leng)))
        pop.append(axl_dojo.CyclerParams(list(start) + [D] * (200 - handshake_leng)))

    # 50-50
    pop.append(axl_dojo.CyclerParams([C] * 100 + [D] * 100))
    pop.append(axl_dojo.CyclerParams([D] * 100 + [C] * 100))

    # Single Change
    for i in range(1, 11):
        pop.append(axl_dojo.CyclerParams([C] * i + [D] * (200 - i)))
        pop.append(axl_dojo.CyclerParams([D] * i + [C] * (200 - i)))

    for i in range(1, 11):
        pop.append(axl_dojo.CyclerParams([C] * (200 - i) + [D] * i))
        pop.append(axl_dojo.CyclerParams([D] * (200 - i) + [C] * i))

    # Matching Tails
    for i in range(1, 6):
        for j in range(1, 6):
            pop.append(axl_dojo.CyclerParams([C] * i + [D] * (200 - (i + j)) + [C] * j))
            pop.append(axl_dojo.CyclerParams([D] * i + [C] * (200 - (i + j)) + [D] * j))

    # Alternating
    pop.append(axl_dojo.CyclerParams([C, D] * 100))
    pop.append(axl_dojo.CyclerParams([D, C] * 100))
    pop.append(axl_dojo.CyclerParams([C, C, D, D] * 50))
    pop.append(axl_dojo.CyclerParams([D, D, C, C] * 50))
    pop.append(axl_dojo.CyclerParams([C, C, C, C, D, D, D, D] * 25))
    pop.append(axl_dojo.CyclerParams([D, D, D, D, C, C, C, C] * 25))
    pop.append(axl_dojo.CyclerParams([C, C, C, C, C, D, D, D, D, D] * 20))
    pop.append(axl_dojo.CyclerParams([D, D, D, D, D, D, C, C, C, C] * 20))

    seq_len = 200
    while len(pop) < pop_size:
        random_moves = list(map(axl.Action, np.random.randint(0, 1 + 1, (seq_len, 1))))
        pop.append(axl_dojo.CyclerParams(random_moves))

    return pop

```

Figure 5.18: Initial Population Code

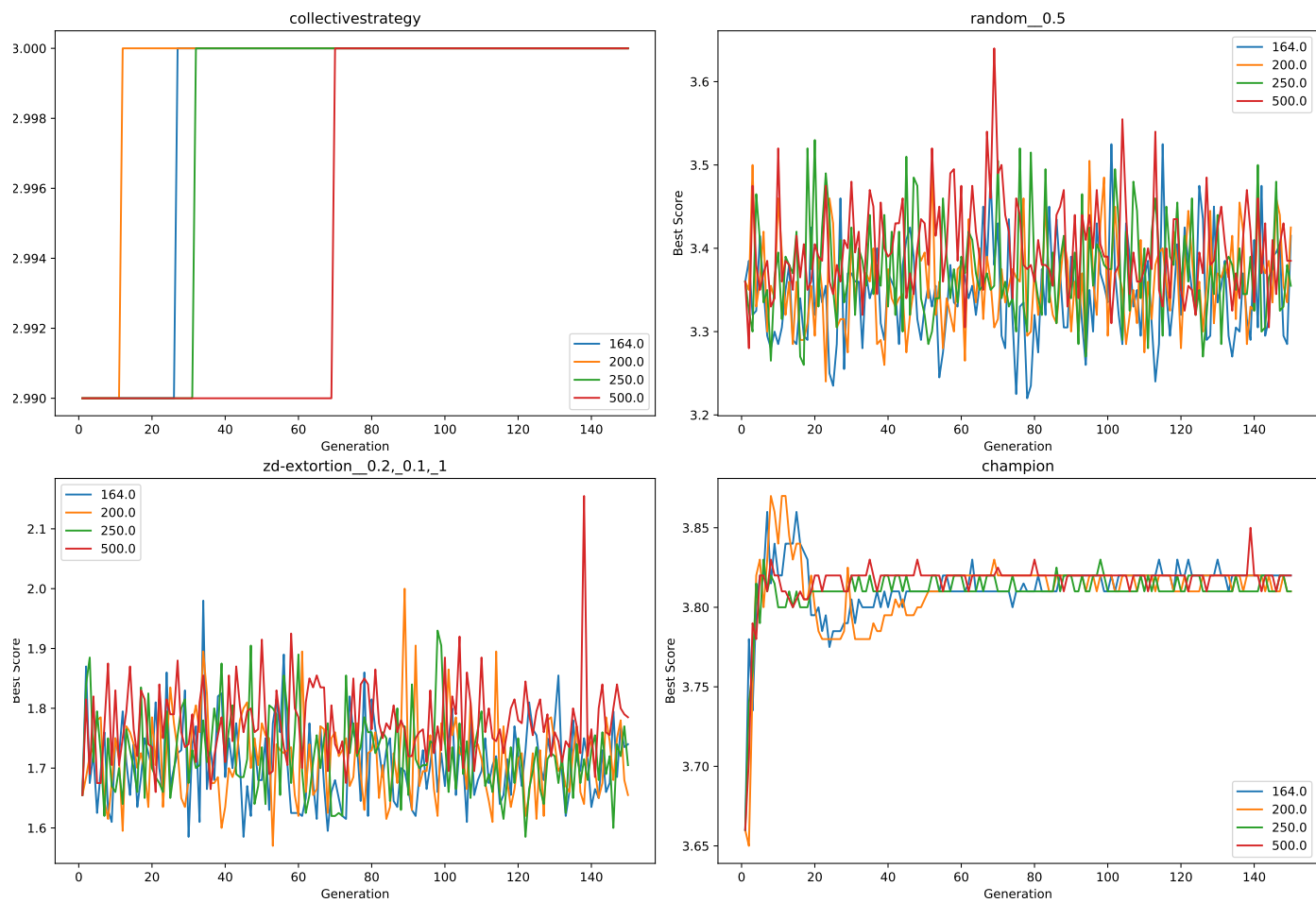


Figure 5.19: Non optimal sequence players after changing initial population

The concept of a stochastic opponent can be somewhat ‘overridden’ by seeding the pseudo random number generator that creates the parameters that define what moves our opponent will take. Currently the only way of doing this is by seeding the whole Axelrod library upon initialising the opponent instance, Snippet 5.23 shows the code for wrapping any given opponent with the global seed command.

When calculating the sequence solution for all of the opponents we will do so with seeded versions of the stochastic opponents. The parameters for stochastic opponents will be the default set by the creators and new instances will be made for any variations with real world applications. For example, default Random players will have a 0.5 chance of playing a C move; this parameter can optionally be changed, and will for certain opponents.

5.7 Conclusion of approach

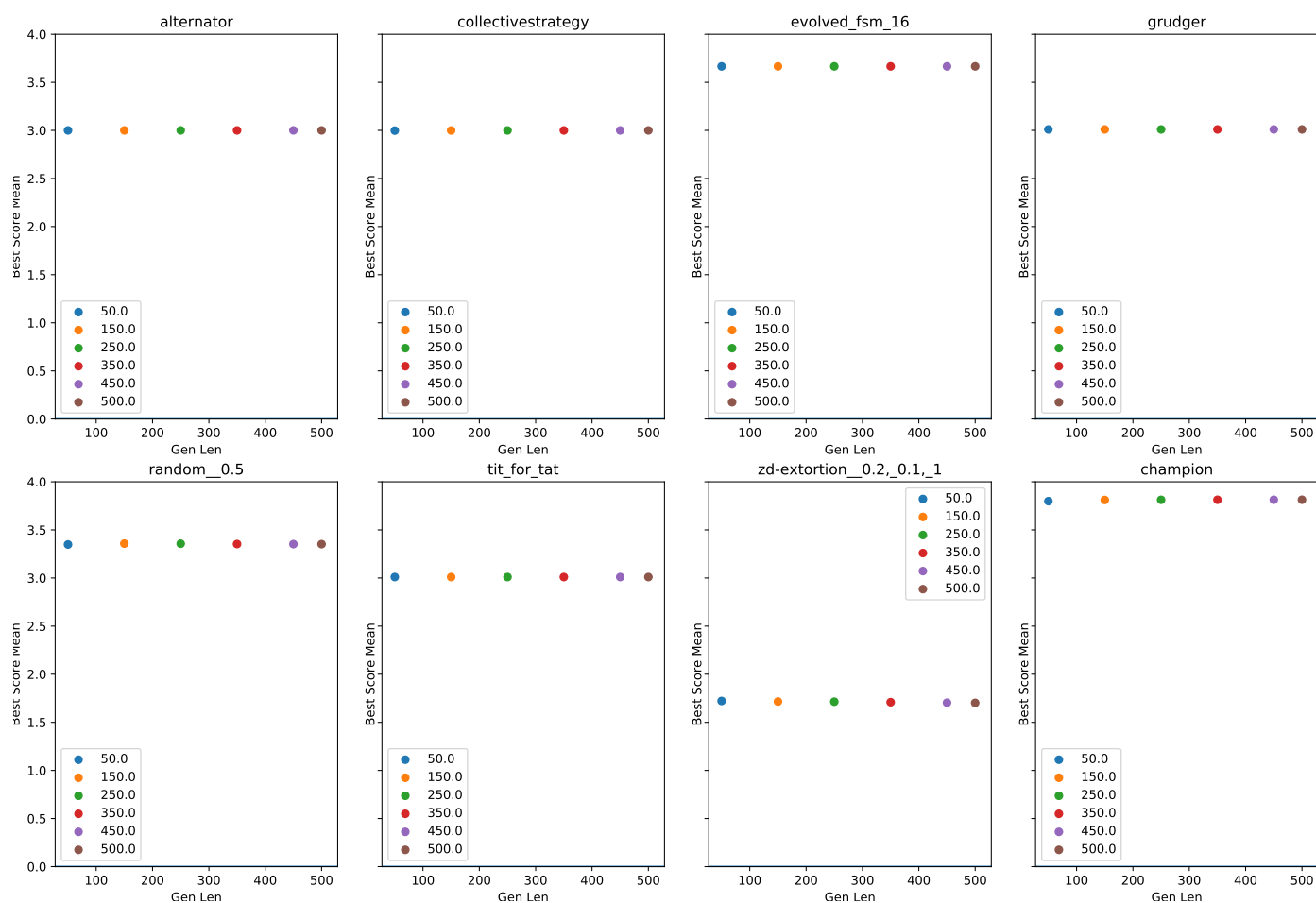


Figure 5.20:

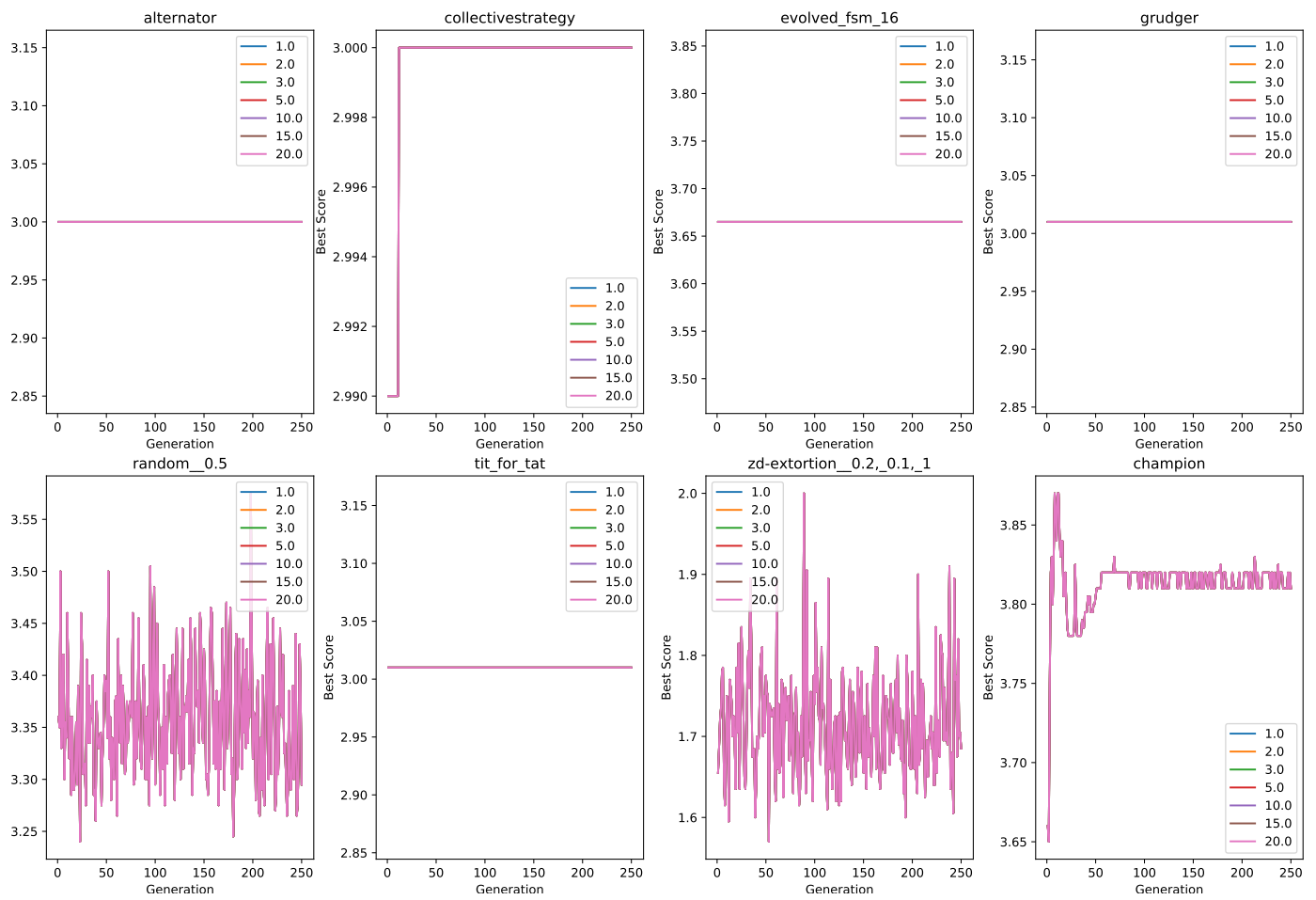


Figure 5.21:

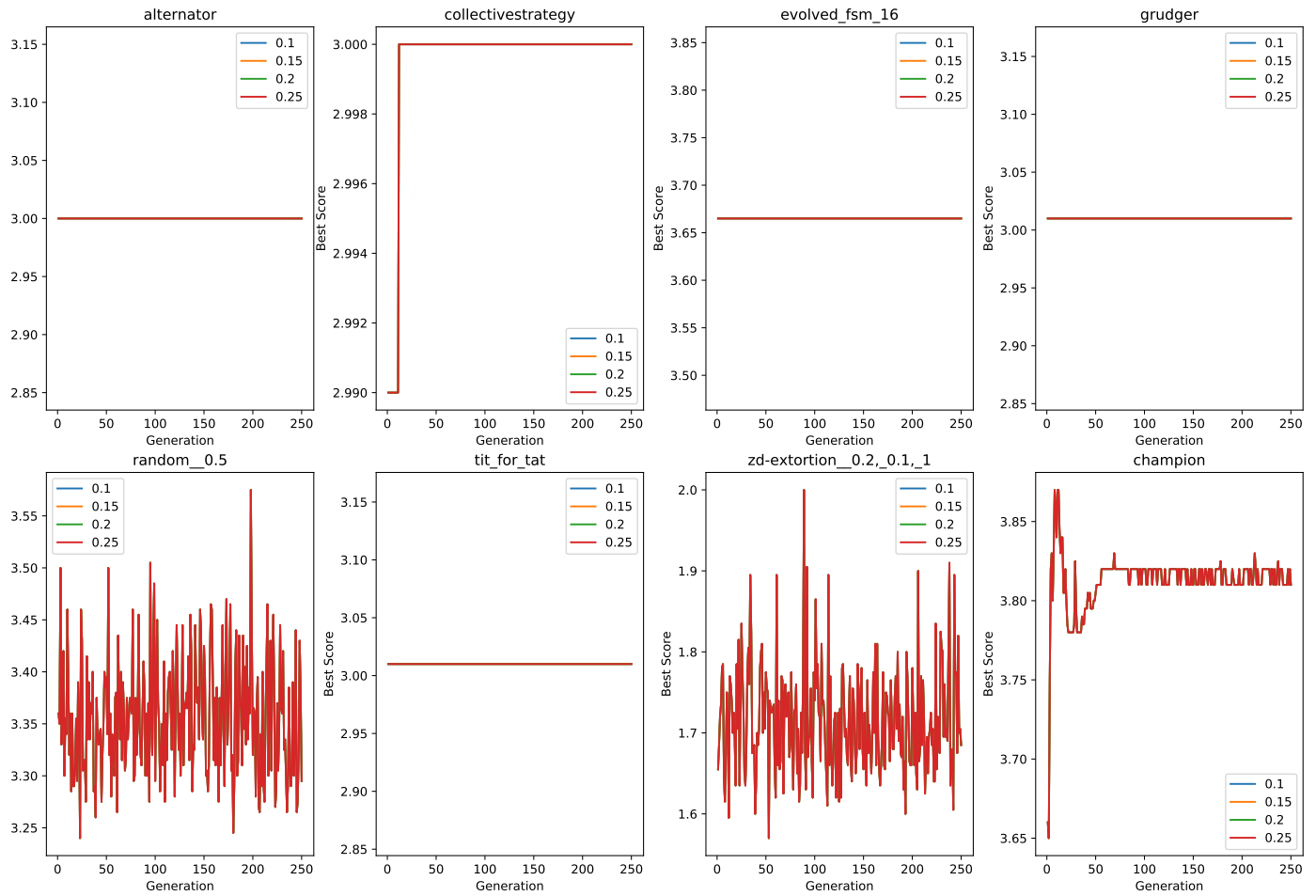


Figure 5.22:

```
def getSeededPlayer(player_class):
    class NewClass(player_class):
        def __init__(self, seed=0):
            axl.seed(seed)
            super().__init__()

    return NewClass
```

Figure 5.23: A function for wrapping a player with a global seed function call

Chapter 6

Results and Discussion

Chapter 7

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. If we can score an average of 3.8 against op1 or 3 against op2, we play op2 etc.

game shows to win money? We want to win as much money as possible but we don't mind how much our opponent wins, as long as we get lots.

What were actually doing is finding a sequence that will manipulate our opponent into providing us the most number of cooperation moves we can subsequently defect against. Failing that we are manipulating the opponent for cooperation moves we can cooperate against, then defection moves we can defect against, then defection moves we must cooperate for.

Chapter 8

Summary and Future Research

Appendix A

Code Appendix

Appendix B

List of Axelrod Opponents

List of Figures

3.1	High Level Genetic Algorithm Cycle	10
4.1	Code to create a sequence result to optimise best score for 3 opponents	12
4.2	Cells for creating the jupyter instance of a research environment	13
4.3	Description and commits for PR on Github	14
4.4	Tail of code commit log as shown on Github	14
4.5	Feature discussions in PR on Github	14
4.6	Scope Discussion in PR on Github	14
4.7	An Example of a test in the Axelrod Dojo	15
5.1	code to check multiple populations	18
5.2	Best score per turn vs generation for different initial population sizes	19
5.3	Scatter of max best score vs different initial populations	20
5.4	code to check multiple generation lengths.	21
5.5	Mean Best Score diff vs total generation lengths	22
5.6	The mutation code as given in the axelrod-dojo	23
5.7	Mutation potency code	24
5.8	Best score vs generation for different mutation potencies	25
5.9	Average best score diff vs mutation potencies	26
5.10	Mutation potency code	27
5.11	Best score vs generation for different mutation frequencies	28
5.12	Average best score diff vs mutation frequencies	29
5.13	Grudger matches against totalities	30
5.14	Old Crossover algorithm	30
5.15	New Crossover algorithm	31

5.16	Example of new crossover algorithm	31
5.17	Best Score vs generations for preset initial populations on top of random sequences	33
5.18	Initial Population Code	35
5.19	Non optimal sequence players after changing initial population	36
5.20	37
5.21	38
5.22	39
5.23	A function for wrapping a player with a global seed function call	39

List of Tables

4.1	Functional Python libraries for analysis	11
4.2	Internal built in python modules used	12
5.1	Output data table	17

Bibliography

- [AH92] Robert J Aumann and Sergiu Hart. *Handbook of game theory with economic applications*, volume 2. Elsevier, 1992.
- [AK08] Daniel Ashlock and Eon Youn Kim. Fingerprinting: Visualization and automatic analysis of prisoner’s dilemma strategies. *IEEE Transactions on Evolutionary Computation*, 12(5):647–659, 2008.
- [AKK04] Daniel Ashlock, Eun Youn Kim, and Warren Kurt. Finite rationality and interpersonal complexity in repeated games. 56(2):397–410, 2004.
- [Axe80] Robert Axelrod. Effective choice in the prisoner’s dilemma. *Journal of conflict resolution*, 24(1):3–25, 1980.
- [BLM95] Bir Bhanu, Sungkee Lee, and John Ming. Adaptive image segmentation using a genetic algorithm. *IEEE Transactions on systems, man, and cybernetics*, 25(12):1543–1567, 1995.
- [BSVdH09] Sander Bakkes, Pieter Spronck, and Jaap Van den Herik. Rapid and reliable adaptation of video game ai. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(2):93–104, 2009.
- [CB97] Paul C Chu and John E Beasley. A genetic algorithm for the generalised assignment problem. *Computers & Operations Research*, 24(1):17–23, 1997.
- [Chu17] Michael Chui. Artificial intelligence the next digital frontier? *McKinsey and Company Global Institute*, page 47, 2017.
- [CS97] Daniel M. Cable and Scott Shane. A prisoner’s dilemma approach to entrepreneur-venture capitalist relationships. *Academy of Management Review*, 22(1):142 – 176, 1997.
- [DB09] Charles Darwin and William F Bynum. *The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life*. AL Burt, 2009.
- [GP15] Jonathan Goldman and Ariel D Procaccia. Spliddit: Unleashing fair division algorithms. *ACM SIGecom Exchanges*, 13(2):41–46, 2015.
- [HKJ⁺17] Marc Harper, Vincent Knight, Martin Jones, Georgios Koutsovoulos, Nikoleta E Glynatsi, and Owen Campbell. Reinforcement learning produces dominant strategies for the iterated prisoners dilemma. *PloS one*, 12(12):e0188046, 2017.
- [Kni17] Will Knight. Alpha zeros alien chess shows the power, and the peculiarity, of ai. *MIT Technology Review*, 2017.
- [Low15] Bobbi S Low. *Why sex matters: A Darwinian look at human behavior*. Princeton University Press, 2015.
- [MD09] Shashi Mittal and Kalyanmoy Deb. Optimal strategies of the iterated prisoner’s dilemma problem for multiple conflicting objectives. *IEEE Transactions on Evolutionary Computation*, 13(3):554–565, 2009.
- [PD12] William H Press and Freeman J Dyson. Iterated prisoners dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26), 2012.
- [pd17] The Axelrod project developers. Axelrod: v4.0.0, 2017.
- [Sni85] Duncan Snidal. The game theory of international politics. *World Politics*, 38(1):25–57, 1985.

- [TC88] John Tooby and Leda Cosmides. The evolution of war and its cognitive foundations. *Institute for evolutionary studies technical report*, 88(1):1–15, 1988.
- [Tor] Linus Torvald. Git.
- [Tur50] Alan M Turing. Computing machinery and intelligence. *Mind*, 59(236):433–460, 1950.