NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

TOURNAMENT-WINNING STRATEGY FOR ITERATED OPTIONAL PRISONER'S DILEMMA

by

Ahmed A. Shamma

September 2020

Thesis Advisor:	Joshua A. Kroll
Second Reader:	Guillermo Owen

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE September 2020	3. REPORT TYPE AND DATES COVERED Master's thesis

4. TITLE AND SUBTITLE TOURNAMENT-WINNING STRATEGY FOR ITERATED OPTIONAL PRISONER'S DILEMMA	5. FUNDING NUMBERS
6. AUTHOR(S) Ahmed A. Shamma	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	12b. DISTRIBUTION CODE A - Approved for public release.

13. ABSTRACT (maximum 200 words)

Iterated optional prisoner's dilemma (IOPD) is an adversarial game that can be used to model several real-world scenarios, from mutual grooming between primates to alliances between business firms. This study utilizes simulation techniques to determine winning strategies for IOPD tournaments in a variety of initial conditions. Machine learning techniques are used to iteratively improve upon the winning strategy, culminating in a single undefeated strategy. The outcome of this study is a single strategy that we claim is likely to win an IOPD tournament for most reasonable initial conditions.

14. SUBJECT TERMS game theory, iterated optional prisoner's dilemma, IOPD	15. NUMBER OF PAGES 59
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

TOURNAMENT-WINNING STRATEGY FOR ITERATED OPTIONAL PRISONER'S DILEMMA

Ahmed A. Shamma
Lieutenant, United States Navy
BS, U.S. Naval Academy, 2015

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

from the

NAVAL POSTGRADUATE SCHOOL
September 2020

Approved by: Joshua A. Kroll
Advisor

Guillermo Owen
Second Reader

Gurminder Singh
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Iterated optional prisoner's dilemma (IOPD) is an adversarial game that can be used to model several real-world scenarios, from mutual grooming between primates to alliances between business firms. This study utilizes simulation techniques to determine winning strategies for IOPD tournaments in a variety of initial conditions. Machine learning techniques are used to iteratively improve upon the winning strategy, culminating in a single undefeated strategy. The outcome of this study is a single strategy that we claim is likely to win an IOPD tournament for most reasonable initial conditions.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION 1

II. BACKGROUND 3
- A. PRISONER'S DILEMMA 3
- B. ITERATED PRISONER'S DILEMMA 6
  - 1. The Unquestioned World Champion, Tit-for-Tat 6
  - 2. Deterministic Memory-One Strategies 7
  - 3. Vocabulary of Strategies 9
  - 4. Clarity 11
  - 5. Importance of Initial Conditions 11
- C. OPTIONAL PRISONER'S DILEMMA 11
- D. PURPOSE AND METHODOLOGY 15

III. MEMORY-ONE EXPERIMENT 17
- A. TOURNAMENT RESULTS 17
- B. LIMITATION OF THIS TOURNAMENT 19

IV. TOP HUNDRED EXPERIMENTS 21
- A. 101 TOP HUNDRED EXPERIMENTS 21
- B. GENETIC LEARNING TOURNAMENTS 22
  - 1. Method for Genetic Learning 23
  - 2. Results of Genetic Learning 23

V. AXELROD REPEATED 25
- A. STRATEGIES 26
  - 1. Deterministic Memory-One Strategies 26
  - 2. Complex Strategies 28
- B. TOURNAMENT RESULTS 29
- C. FINAL CONFIRMATION 31

VI. CONCLUSION 33
- A. TOPICS FOR FURTHER RESEARCH 33
- B. IMPLICATIONS 34

APPENDIX: MEMORY-ONE EXPERIMENT RESULTS 35

SUPPLEMENTALS 39

A. SUPPLEMENTAL 1: MEMORYONE.....39
B. SUPPLEMENTAL 2: GENETIC TOURNAMENT, MEANPLAYERS, AND NICEPLAYERS 39
C. SUPPLEMENTAL 3: AXELROD REPEATED.....39

LIST OF REFERENCES.....41

INITIAL DISTRIBUTION LIST.....43

LIST OF TABLES

Table 1. Prisoner's Dilemma Payoffs 3

Table 2. Cold War Arms Race 4

Table 3. Coke and Pepsi Price War....5

Table 4. Prisoner's Dilemma, Abstract Payoffs....5

Table 5. Memory-one strategy table 7

Table 6. Tit-for-Tat 8

Table 7. Memory-One IPD Strategies....8

Table 8. Equivalent All Defect Strategies....10

Table 9. Optional Prisoner's Dilemma 12

Table 10. Memory-One IOPD Strategies....13

Table 11. OPD Tit-for-Tat 14

Table 12. Nicer Tit-for-Tat 14

Table 13. Mole 17

Table 14. 2nd, 3rd and 4th Place Memory-One Tournament Strategies 18

Table 15. Viper 22

Table 16. Turtle....22

Table 17. Rattlesnake....24

Table 18. Snail 24

Table 19. Tit-for-Tat 26

Table 20. Nicer Tit-for-Tat 26

Table 21. Grim Trigger 27

Table 22. Abra....27

Table 23. Mole 27

Table 24. Turtle 28

Table 25. Snail 28

Table 26. Tournament Results for the Axelrod Repeated Tournament 29

LIST OF ACRONYMS AND ABBREVIATIONS

ALL C	Always Cooperate
ALL D	Always Defect
IOPD	Iterated Optional Prisoner's Dilemma
IPD	Iterated Prisoner's Dilemma
OPD	Optional Prisoner's Dilemma
TF2T	Tit-for-Two-Tats
TFT	Tit-for-Tat
WSLS	Win-Stay-Lose-Shift

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to first thank my thesis advisor, Professor Joshua Kroll, for the numerous thesis video chats, which rarely stayed on topic but were always entertaining and informative. I learned more than I ever imagined I would about cryptography and jellyfish by writing a thesis about machine learning and game theory. More sincerely, I would also like to thank him for his patience, understanding and forbearance throughout the difficulties that the coronavirus inflicted upon this project. I could not have asked for a better advisor.

I would also like to thank Nicky Case for creating "evolution of trust," the model that spurred my interest in prisoner's dilemma and helped me to appreciate how game theory helps us to understand human nature and interactions. Kindness, trust, forgiveness, justice—these are not social constructs; they are core parts of our identity as a species. They are central to the moral cooperation that helped some bald apes go from scraping the dirt of the savanna looking for bugs to landing biochemistry labs on the surface of Mars in just a few short millennia.

As well, I would like to thank my partner, my companion, and the love of my life, Angela Yu, who cheerfully and compassionately took on the burdens of day-to-day life as the stress and pressure of grad school and my thesis bled into my personal life. I could not have done this without you, and I am honored and excited to share my life and my future with you. Thank you, my angel.

Finally, I would like to thank my father, Professor Shihab Shamma, for instilling me with a lifelong passion for math and science. When I was a child, he would take every other Friday off from work to take me to the zoo or to the Smithsonian museums and share his love of the natural world with me. I would be a different and lesser person without those experiences. I love you, Baba.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Simple adversarial games provide useful models for complex strategic and tactical conflicts in the real world. Game-theoretic formalization provides the means to bring the tools of mathematics to bear on analogies of real-world phenomena.

Iterated prisoner's dilemma (IPD) is one such game, which describes a variety of real-life scenarios that involve trust, benefits from cooperation, and vulnerability to defection. This game has been used by evolutionary biologists, social scientists, and many other experts to analyze strategy in a diverse range of scenarios, from social feeding among vampire bats (Dawkins, 1976), to nuclear arms races (Majeski, 1984), to price competition between opposing firms (Wayland, 1991).

One strategy, Tit-for-Tat, has long been known to be the "unquestioned world champion" in IPD (Nowak, 2006, p. 78). This strategy was first examined in depth by the University of Michigan Ann Arbor political scientist Robert Axelrod, who ran two tournaments in 1980 in which he invited several experts in a variety of fields to submit strategies to play against one another in a round-robin style. Tit-for-Tat won both tournaments, and has been shown since to be a robust and extremely successful strategy for IPD (Axelrod, 2006).

This strategy has certain key characteristics that contribute to its success. It is "nice," meaning that it starts by cooperating and is never the first to defect. It is "retaliatory," meaning that it retaliates against defection by its opponent. However, it is also "forgiving," meaning that after it retaliates, the Tit-for-Tat strategy can go back to cooperating. Formal definitions for these characteristics are given in the Background section.

Iterated optional prisoner's dilemma (IOPD) is a related game to IPD. It is also useful for modelling conflicts, but has received comparatively less attention and analysis. IOPD models situations where, in addition to defecting or cooperating, there is an option simply not to participate (hence the inclusion of "optional" in its name). The addition of the third option to abstain makes IOPD a more complicated game and leads to richer dynamics than in IPD.

No unquestioned champion like Tit-for-Tat has been definitively shown to exist for IOPD. In this thesis, we attempt to find such a champion strategy. In this thesis project, we performed three experiments in an effort to find just such a strategy, should one exist. Our initial expectation was that such a strategy would exist, and that it would share characteristics with Tit-for-Tat, i.e., it would be nice and forgiving.

To find such a strategy, we describe the results of three sets of tournaments, each similar in structure to the original Axelrod tournaments for IPD. Our first tournament narrows down the strategic space by playing 59049 simple strategies against one another to determine the best performers. In a second set of related tournaments, we play 200 successful strategies from the previous tournament against one another in various different initial conditions, and use genetic learning to allow successive generations to improve upon the previous round's winners, so that they ultimately produce a single winning strategy. Finally, one last tournament adapts some of the strategies from Axelrod's original IPD tournament to IOPD to see if any of those strategies could beat the winning strategies from the previous tournaments. In our conclusion, we argue that the winner of the last two tournaments, "Snail," is likely to be a canonical champion strategy for IOPD.

Although no single strategy won all of the tournaments, high performing strategies in all four tournaments shared three common characteristics. For each of the tournaments, the winning strategy was nice, as we predicted it would be. However, contrary to our predictions, it was not forgiving. Additionally, the winning strategies were not strictly retaliatory like Tit-for-Tat, but instead would exercise their option to abstain rather than face repeated mutual defection. We called strategies that have this property "defensive."

II. BACKGROUND

A. PRISONER'S DILEMMA

Prisoner's dilemma is among the most exhaustively studied games in game theory. Though simple in formulation, the game provides far-reaching analogies with implications for many other fields, including economics, politics, and ethics.

Prisoner's dilemma was originally proposed in 1950 by Merrill Flood and Melvin Drescher, and was given its name in 1951 by Albert W. Tucker (Mero, 2012). In Tucker's (1983) formulation, two suspects are charged with joint commission of a crime. The authorities offer each a reward worth 1 unit to tattle on his partner. However, their partner will receive a fine worth 2 units if the other confesses. Therefore, if they both confess, then they will each pay a net fine of 1 unit. If both stay silent, then they will both go free with no fine. These outcomes are summarized on Table 1.

Table 1. Prisoner's Dilemma Payoffs

	Prisoner B stays silent	Prisoner B confesses
Prisoner A stays silent	Both A and B go free	B gets reward of 1, A is fined 2
Prisoner A confesses	A gets reward of 1, B is fined 2	Both pay a net fine of 1

Examining the payoff table, we see that regardless of what Prisoner B does, Prisoner A is better off confessing. Since the game is symmetrical, the same applies to Prisoner B. So, since both players should confess, optimal play will result in both prisoners receiving a fine of one unit. Notice, though, that this outcome in the lower right corner (both pay a net fine of 1) is strictly worse than the outcome in the upper left corner (both go free with no fine), where each prisoner serves a sentence of just one year. Paradoxically, the outcome of both players playing their optimal move in Prisoner's Dilemma optimally results in an outcome strictly worse than another possible outcome for both players.

Many real-world scenarios mirror the prisoner's dilemma, not just the example of two criminals lying to police. Generally speaking, prisoner's dilemma is useful for modelling any cooperative scenario where betraying the other player provides a reward and cooperating with the other player involves taking on a risk.

One example is the case of a nuclear arms race between two nations (Majeski, 1984). If one country elects to make nuclear weapons and the other does not, then the nuclear nation will enjoy a massive military and diplomatic advantage over the other state. But if both nations develop weapons, then they will have both expended resources without gaining any comparative advantage. Table 2 illustrates this example as applied to the United States and the Soviet Union in the Cold War.

Table 2. Cold War Arms Race

	USSR Builds Food	USSR Builds Weapons
USA Builds Food	Global peace and plenty (Good outcome for both)	USSR global hegemony (Best outcome for USSR, Worst outcome for USA)
USA Builds Weapons	USA global hegemony (Best outcome for USA, Worst outcome for USSR)	Cold war stalemate (Bad outcome for both)

Another example is price wars between competing firms selling a similar product. Either firm might have the option to lower their prices, thus generating more sales than their competitor, but generating less profit per item sold. However, if both firms lower their prices, then neither will improve their relative market share, and thus both reduce their total revenue. Rebecca Wayland of the Harvard Business Review applied this analysis to price wars between Coca-Cola and Pepsi-Cola, summarized in Table 3.

Table 3. Coke and Pepsi Price War

	Coke Maintains Prices	Coke Lowers Prices
Pepsi Maintains Prices	Maintain revenue (Good outcome for both)	Coke dominates the market (Best outcome for Coke, Worst outcome for Pepsi)
Pepsi Lowers Prices	Pepsi dominates the market (Best outcome for Pepsi, Worst outcome for Coke)	Reduced revenue for both (Bad outcome for both)

Real life often contains more options than just Cooperation and Defection, but reducing conflicts to simple games is a useful analytic tool. It can help to give concise answers to questions such as "why did the USA and the USSR spend billions building weapons that they never used, instead of investing in their own citizens?" This tool can help with decision-making as well as analysis. A simple model lends itself to mathematical analysis and complete solutions, providing a formal best solution for even complex predicaments.

The Prisoner's Dilemma can be viewed abstractly for game theory analysis. It is an adversarial game where players may choose to either Cooperate or Defect. If they both Cooperate, they receive a good payoff ("Reward"). If they both Defect, they receive a smaller payoff ("Punishment"). If one Cooperates, and the other Defects, then the player who defected receives the best payoff ("Temptation") while the player who Cooperated receives the worst payoff ("Sucker").[1]

Table 4. Prisoner's Dilemma, Abstract Payoffs

	Cooperate	**Defect**
Cooperate	R, R	S, T
Defect	T, S	P, P

[1] All the terms here are relative. In some cases, all the outcomes can be differing degrees of good or bad, as in the original case of the two prisoners who will serve some amount of jail time.

$$S < P < R < T, S + T < 2R$$

Table 4 summarizes the description in the previous paragraph, where the first character represents the payoff for the left column player, and the second represents the payoff for the top row player. The exact amounts assigned to each value can vary for different versions of the game, but the strict inequalities below the table must hold for the game to be Prisoner's Dilemma.[2] For purposes of this study, we will use the conventional values: S = 0 ("Sucker"), P = 1 ("Punishment"), R = 3 ("Reward"), T = 5 ("Temptation") (Press & Dyson, 2012).

In game theory, a "strictly dominant" strategy is one that is always strictly better than the other options, no matter what the opponent chooses. When a strictly dominant strategy is proven to exist in a game, that game is said to be "solved." Since such a strategy exists in Prisoner Dilemma (always Defect), Prisoner's Dilemma is a solved game.

B. ITERATED PRISONER'S DILEMMA

Iterated Prisoner's Dilemma (IPD) is a variant on prisoner's dilemma, wherein the two players play prisoner's dilemma many times successively, instead of just once. In the context of IPD, each of these successive rounds of prisoner's dilemma is referred to as a "move," while the entire match of IPD is called a "game." Unlike prisoner's dilemma, IPD is not solved. Players who play repeated games against each other can benefit in the long run from mutual Cooperation. Therefore, the optimal strategy is more complicated than simply always Defecting.

1. The Unquestioned World Champion, Tit-for-Tat

The landmark study into IPD was performed by the University of Michigan Ann Arbor political scientist Robert Axelrod. In this study, Axelrod invited interested parties from all over the world to submit strategies to compete against each other. Fourteen

[2] There are games called Stag Hunt (Skyrms, 2001) and Chicken (Snyder, 1971), which are similar in formulation to Prisoner's Dilemma but with different inequalities governing the payout. They are useful for modelling other situations of trust and cooperation but are beyond the scope of this thesis.

respondents submitted computer programs implementing a variety of strategies, which were played against one another in sequence.

The winner, submitted by mathematical psychologist Anatol Rapoport, was also the simplest strategy. Called "Tit-for-Tat" (TFT), it started by Cooperating and then mirrored the opponent's move in the previous round. That is, if the opponent Defected, then the strategy would Defect in return in the next round. If the opponent Cooperated, then TFT would continue to Cooperate. Axelrod (1980a) published the results of this tournament along with analysis in a 1980 paper "Effective Choice in the Prisoner's Dilemma." In the same year, he ran a tournament with the same rules but 62 entrants. Tit-for-tat won again (Axelrod, 1980b). In the years since, more tournaments like Axelrod's have been run by Axelrod and others, and although there have been cases where Tit-for-Tat has not won (Axelrod, 2006, p. 205), it is typically the winner in a large variety of initial conditions, to the extent that it has been called by one expert the "unquestioned world champion in the repeated PD" (Nowak, 2006, p. 78).

2. Deterministic Memory-One Strategies

TFT is one of a class of strategies called deterministic memory-one strategies, the simplest class of strategies that is successful in IPD tournaments. A "memory-one" strategy makes its next move based only upon the most recent move (Press & Dyson, 2012). A "deterministic" strategy is one that does not utilize any form of randomness in its decision making, whereas a "mixed" strategy is one that does (Heuer, 2019).

In IPD, a deterministic memory-one strategy can be defined by its first move, as well as its response to the outcomes of the prior move. Notating such a strategy requires five binary variables, which we can efficiently display in a table:

Table 5. Memory-one strategy table

Strategy Name	*CC*	*CD*
firstplay	*DC*	*DD*

The top left cell is the name of the strategy, while the other five cells are the five binary variables representing the strategy. The bottom left cell of the table ("*first_play*") is the first play that the strategy makes. The other four cells are the plays that the strategy will make based upon the moves of the previous round. The move in the cell marked "*CC*" is played when both players Cooperated in the previous round, "*DC*" when the strategy described Defected and its opponent Cooperated, "*CD*" when the strategy Cooperated and the opponent Defected, and "*DD*" when both players Defected.

As an example, Tit-For-Tat would be as follows:

Table 6. Tit-for-Tat

Tit-for-Tat	C	D
C	C	D

A surprising amount of strategic complexity is possible using only deterministic memory-one strategies. Some other memory-one IPD strategies include

- **All Cooperate (ALL C):** always Cooperates regardless of its opponent's moves
- **All Defect (ALL D):** always Defects
- **Grim Trigger:** starts by Cooperating, but then Defects and permanently continues Defecting if the opponent Defects even once
- **Win-Stay-Lose-Shift (WSLS):** repeats its previous move if the opponent Cooperated but changes its move if the opponent Defected.

Table 7. Memory-One IPD Strategies

All Defect	D	D
D	D	D

Grim Trigger	C	D
C	D	D

All Cooperate	C	C
C	C	C

WSLS	C	D
C	D	C

3. Vocabulary of Strategies

There is a common vocabulary for discussing possible strategies, to quickly identify and describe properties that some of them share. This vocabulary is not strictly limited to deterministic or memory-one strategies, and can be extended to all possible strategies of IPD.

a. Nice

Strategies that are never the first to defect are called "nice" (Axelrod, 2006, p. 33). It is easy to tell whether a deterministic memory-one strategy is nice, by inspection of its table. If a strategy table has a "C" value in its first play cell and it the CC cell, then the strategy is nice. Otherwise, it is not nice. All Cooperate, Grim Trigger, TFT and WSLS are all nice strategies. All Defect is not. Nice strategies have consistently been shown to perform well, particularly when paired among many other nice strategies (Axelrod 2006).

b. Mean

In this thesis, we use the term "mean" to refer to strategies that Defect on the first move. The other four variables describe the response to the play made in the previous round. It is conceivably possible for a strategy to be neither "nice" nor "mean" in IPD, if it Cooperates on the first move, but then Defects on following moves.

c. Forgiving and Unforgiving

"Unforgiving" strategies are those which, once an opponent has Defected, will never again Cooperate. "Forgiving" strategies are those which, when an opponent has Defected, can Cooperate again in future rounds of the same game. Grim Trigger is an unforgiving strategy, while TFT, WSLS and All Cooperate are forgiving. Note that there is no relationship between whether a strategy is nice and whether it is forgiving. That is, a strategy could be nice yet unforgiving, like Grim Trigger. An example of a strategy that is forgiving, but not nice is JOSS, a mixed memory-one strategy submitted by mathematician Johann Joss (Axelrod, 2006) in the original Axelrod tournament. It plays as TFT and is therefore forgiving but has a 10% chance of Defection any time it would otherwise Cooperate, which makes it not nice.

d. Retaliatory

Axelrod defined a strategy as "retaliatory" if it "immediately defects after an 'uncalled for' defection from the other" (Axelrod, 2006, p. 44). He did not precisely define what "uncalled for" meant, but the point is that a strategy will Defect in response to the other player Defecting. By inspection, we can tell if a strategy is "retaliatory" by looking at its table. If the far right column contains only "D," then the strategy is retaliatory.

e. Equivalent Strategies

In Section II.B.2, we showed that deterministic memory-one IPD strategies can be represented with five binary variables. Five binary variables would imply that there are 32 possible deterministic memory-one strategies, but this calculation is misleading since four strategies are functionally equivalent to All Cooperate, and four strategies are functionally equivalent to All Defect. We can see an example of equivalent strategies in Table 8.

Table 8. Equivalent All Defect Strategies

All Defect #1	D	D
D	D	D

All Defect #2	C	C
D	D	D

Clearly, the strategy on the left will always Defect, regardless of the moves made by the opponent. However, the strategy on the right will also always Defect. The strategy will never have Cooperated in the previous move, so the contents of the top two cells *CC* and *CD* are irrelevant. These cells never govern the play of this strategy regardless of what the opponent does.

For two strategy representations to be functionally equivalent, a row of their table must never be activated. If a row is never activated, then its contents do not matter, and thus all possible permutations of that row will be functionally equivalent. Because there are only two rows in a strategy table in IPD, there are two cases which give rise to functionally equivalent strategies. In one, the Cooperate row is never activated, and another where the Defect row is never activated. Respectively, these are the equivalence classes of All Defect

and All Cooperate. It follows that the other 24 strategies are unique, for a total of 26 equivalence classes of strategies.

4. Clarity

Most of the strategies submitted to the Axelrod tournaments were not memory-one strategies and attempted to use cleverer and more sophisticated strategies in order to win. JOSS, for example, was a strategy similar to Tit-for-Tat, but would attempt to sneak an occasional cheat in. If the opponent Cooperated, there was a 10% chance that JOSS would Defect in the subsequent move, to try to cheat the opponent. Several strategies also attempted to characterize the opponent's strategy by alternately Defecting and Cooperating to see what the opponent would do. However, despite being among the simplest strategies submitted in both tournaments, TFT was also the most successful. Axelrod considered clarity to be an advantage of TFT, that it did not attempt to be too clever (Axelrod, 2006).

5. Importance of Initial Conditions

The initial conditions of a tournament can also influence the winner. Most of the submissions to the Axelrod tournament were nice, but in a tournament where all the entries are All Defect, no other strategy can do better. In Axelrod's words, "ALL D is collectively stable" (Axelrod, 2006, p. 63). For entries other than All Defect to invade and succeed under such circumstances, they must do so in "clusters" of sufficient size. Clusters of Tit-for-Tat that are added to such a tournament and mutually cooperate can gain enough by doing so to offset the few points that ALL D will gain in the first round against them.

C. OPTIONAL PRISONER'S DILEMMA

A variation of the prisoner's dilemma game is optional prisoner's dilemma (OPD). OPD is identical to prisoner's dilemma but introduces a third option of Abstention. If either player Abstains, then both players receive a "Loner" payoff, worse than Reward but better than Punishment. For purposes of this study, we use L=2 (Cardinot, O'Riordan, & Griffith, 2016). Optional prisoner's dilemma, unlike prisoner's dilemma, is not a solved game,

because no move strictly dominates all other moves.[3] If the player knows that their opponent will Defect, then they should Abstain. If the player knows that their opponent will Cooperate, then it is better to Defect. If their opponent Abstains, then all three possible moves result in the same outcome. Table 9 summarizes this description.

Table 9. Optional Prisoner's Dilemma

	Cooperate	**Abstain**	**Defect**
Cooperate	R, R	L, L	S, T
Abstain	L, L	L, L	L, L
Defect	T, S	L, L	P, P

$S<P<L<R<T,\ S + T < 2R$

OPD can be extended to many real-world scenarios in the same way that prisoner's dilemma can be. Oftentimes, real-world agents can opt out of the interaction altogether, rather than risk exposure to Defection. Batali and Kitcher (1995) used OPD to model social grooming among primates, where Cooperate means grooming a partner thoroughly, Defect was grooming a partner lazily, and Abstain was refusing mutual grooming and self-grooming instead. Arend and Seale (2005) used OPD to model alliances between business firms, where the Cooperate move represents making a costly investment into the partner firm, Defect represents not investing while profiting off the other firm's capital and labor, and Abstain reflects terminating the business relationship and operating solo. One final example might be whether to work alone or with a partner on a graduate thesis. Suppose there are two students who are considering working together on a joint thesis, which entails as greater amount of work. Either student has the option to either commit to work with a partner and work hard (Cooperate) or to commit to work with a partner and slack off (Defect). But there is a third option, which is to simply elect to do a thesis alone (Abstain).

[3] Note, however, that Defect does weakly dominate Cooperate. One move weakly dominates another when its payoff is greater *or equal* in all cases.

Working alone is more work than working with a diligent partner, but less work than working with a lazy one.

OPD is useful for modelling scenarios like those that prisoner's dilemma models. They key difference is that OPD models scenarios where entering the game is optional, or where Cooperation and Defection are both active choices and there exists a third passive choice to Abstain. We can see that OPD does not apply to the classic example of the two prisoners in separate rooms. They clearly do not have the choice to opt out of the interaction altogether, as they have both already been arrested. Similarly, in the case of the arms race, Cooperation models a passive choice to simply not build nuclear weapons. So, OPD would not apply to the arms race either.

Like Prisoner's Dilemma, OPD can also be played in successive rounds against the same opponent. This game is called iterated optional prisoner's dilemma (IOPD). Deterministic memory-one strategies exist in IOPD and can be notated similarly, but using 10 trinary variables instead of 5 binary variables, as in Table 10:

Table 10. Memory-One IOPD Strategies

Strategy Name	*CC*	*CA*	*CD*
firstplay	*AC*	*AA*	*AD*
	DC	*DA*	*DD*

The cell "*AC*" is the move that would be made after the player Abstains and the opponent Cooperates, "*AA*" is the move after mutual abstention, and so on. The strategic logic of Tit-for-Tat can be extrapolated to iterated OPD. A strategy that repeats the opponent's last play in the subsequent round would be notated as follows:

Table 11. OPD Tit-for-Tat

Tit-for-Tat	C	A	D
C	C	A	D
	C	A	D

Another equivalent summary of Tit-for-Tat's logic in prisoner's dilemma is "Cooperate unless your opponent Defected, in which case Defect once." Though that summary gives rise to the same outcomes as "play the opponent's last move" in prisoner's dilemma, in OPD it creates a different strategy, Nicer Tit-for-Tat, notated in Table 12:

Table 12. Nicer Tit-for-Tat

Nicer Tit-for-Tat	C	C	D
C	C	C	D
	C	C	D

By the same logic as the proof in Section II.B.3.e, it is provable that some of these representations are equivalent. We can count the number of strategies via a counting proof.

- There are three equivalence classes of strategies that always do the same thing (All Cooperate, All Abstain and All Defect), each of which contains 3^6 different representations.

- Additionally, there are 336 equivalence classes of strategies that only do two things (e.g., only Abstain and Cooperate, but never Defect). Each of these equivalence classes contains 3^3 strategic representations.

- Working backwards, we can subtract all of these equivalent representations from the total number of representations to find the number of unique

representations. $3^{10} - 3*3^6 - 336 * 3^3 = 47790$, which means 47,790 strategies have a unique representation

- Adding up all of the previous equivalence classes, we get 3 + 336 + 47790 = 48129, which means that there are a total of 48,129 strategies.

D. PURPOSE AND METHODOLOGY

No canonical best strategy has been found in iterated optional prisoner's dilemma. The purpose of this thesis is to determine a strategy for iterated optional prisoner's dilemma that performs well in a number of tournament structures, in much the same way that Axelrod was able to establish Tit-for-Tat as a favored strategy for IPD.

The strategy space of IOPD is much larger than the strategy space of IPD. Each deterministic memory-one strategy representation in IPD is defined by 5 parameters, each of which can take one of two possible states, for a total of 32 possible strategy representations. By contrast, we represent IOPD deterministic memory-one strategies using 10 parameters, each of which can take 10 possible states, for a total of 3^{10}, or 59049 possible representations.

To reduce this strategy space down to a more manageable size, we first played a tournament of 59049 players playing each of the possible memory-one strategy representations. The results of this tournament are described in Chapter III. A "tournament" of *n* players consists of a round-robin pairing of each player playing a game of IOPD against every other player. All games of IOPD in this study consist of 200 moves. After each move, the payoff for that move is added to each player's score, as described in Sections II.A and II.C. In a tournament of *n* players, each player played *n*-1 games, one against each other player in the tournament. A player's score and strategy are carried on from one game to the next. However, their memory is reset for each new game, so a player's first move is dictated by their strategy alone. At the end of the tournament, the players are ranked by their score, and the player with the highest score in a tournament is the winning player.

After that, we ran a second set of tournaments, with results described in Chapter IV, in which we took the top 100 nice strategies and the top 100 mean strategies from the first tournament, and pit them against one another in 101 different tournaments with different initial conditions, to see if any repeating winners emerged. Then, we replaced the worst strategies in each tournament with strategies closely related to the winning strategy (identical in representation but with a single cell changed). We repeated this process iteratively for ten generations to converge upon a winning strategy.

Finally, in the last tournament, the results of which are described in Chapter V, we wanted to determine if there were any strategies other than memory-one strategies that could beat the winners of the previous tournaments. To that end, we adapted some of the strategies from Axelrod's original tournament to IOPD. None of these strategies were able to beat any of the three winners of the previous tournaments, which scored in the top three.

These top three strategies all had in common that they were nice, but also, contrary to our expectations, they were all unforgiving. They also had a third attribute in common, which is that they had a mechanism to prevent themselves from becoming stuck in a state of repeated Defection. We called this third attribute "defensive."

III. MEMORY-ONE EXPERIMENT

The first experiment performed was to simulate a tournament similar to the Axelrod tournament but for iterated Optional Prisoner's Dilemma instead of Prisoner's Dilemma, with a player for each of the 3^{10} = 59049 representations of deterministic memory-one strategies. As with all tournaments in this study, each player played one game against each other player, with games consisting of 200 simultaneous moves.

Overall, this would require 1,743,362,676 total games of IOPD for this tournament (59049 choose 2). However, these deterministic memory-one strategies have only ten states, so eventually every game settles into a repeating pattern of moves. It was possible to identify these repeating patterns of moves and thus extrapolate the outcome of each game without having to individually simulate every single move.

A. TOURNAMENT RESULTS

For analysis, players of this tournament were sorted by score, from best performing to worst performing. It is useful to consider the top 100 unique strategies from the Memory-One Tournament; players that employed functionally equivalent strategies would have the same score were collapsed into the same strategy. The full table of the top 100 unique strategies, and the scores of the player that employed them, are included in Appendix A.

Our expectation, that the winning strategy would be nice and forgiving like Tit-for-Tat, was wrong. Tit-for-Tat did not even do particularly well, finishing in 20,308th place out of 59,049. The winning strategy, which we assign the shorthand nickname "Mole," was:

Table 13. Mole

Mole	C	D	D
C	D	D	D
	D	D	A

Overall, Mole is more similar to Grim Trigger than it is to Tit-for-Tat. Like Grim Trigger and Tit-for-Tat, it is a nice strategy that Cooperates on the first play and continues to Cooperate as long as the opponent does. However, it is an unforgiving strategy, in that as soon as the other player Defects or even Abstains, it stops Cooperating for the remainder of the game, just like Grim Trigger, and begins to Defect.

Unlike Grim Trigger, rather than get stuck in an endless loop of mutual Defection, the strategy will Abstain if both players Defected in the previous move. This tactic prevents permanent mutual Defection, which is a poor outcome. This property, of Abstaining after mutual Defection, will turn out to be important to successful strategies in the experiments to follow, and we shall refer to this behavior as "defensive."

A defensive strategy has Abstain in the lower right-hand corner of its tabular representation. Interestingly, defensive strategies are not , because their far-right column does not contain all Ds. In other words, defensive strategies can be Defected against without Defecting in response. This was significant way in which Mole different not only from TFT but also from Grim Trigger.

Mole was a nice strategy. So were the 2nd, 3rd and 4th place strategies, which were all relatives of Mole.

Table 14. 2nd, 3rd and 4th Place Memory-One Tournament Strategies

2nd Place	C	D	D
C	A	D	D
	D	D	A

3rd Place	C	D	D
C	D	A	D
	D	D	A

4th Place	C	D	D
C	D	D	A
	D	D	A

Surprisingly, though, the 5th place strategy was "mean" (it Defected on the first move), and so were 90 of the top 100 strategies. This result may indicate that when Abstention is possible, there is less of an incentive to play nice.

The four top strategies were nice, and most of the top 100 strategies were mean, but Abstaining on the first play was a particularly unsuccessful strategy. This came as some surprise, since intuitively, one might think that being cautious for the first play, and thus seeing how the opponent behaved before committing to an action would be a smart strategy. At least in this tournament of IOPD, it was not.

The top 5 strategies were all defensive, having Abstain in the bottom right corner of their strategy table. Of the top 100 strategies, none had Defect in the bottom right corner. In other words, it was important to avoid getting caught in an endless cycle of mutual Defection.

B. LIMITATION OF THIS TOURNAMENT

As a method for determining a champion strategy for IOPD, the specific methodology used in this first experiment could fairly be criticized in several ways. Mainly, though, in the original Axelrod tournaments, the submissions were all made by experts in the field and were therefore, at least in principle, good and well-considered strategies. By contrast, most of the 59,049 possible memory-one iterated OPD strategies are rather poor. The main utility of this tournament was to narrow down the range of all the possible strategies into a workable subset. So, in order to see if Mole was in fact a broadly high-performing strategy, another tournament had to be run that evaluated its performance in a wider range of conditions.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. TOP HUNDRED EXPERIMENTS

The main limitation of the Memory-One experiment was that every single possible strategy was included, not just the effective ones. Some players in the Memory-One tournament were utilizing poor strategies that no serious competitor would ever submit to a tournament. To build more realistic tournaments, we performed a second set of experiments that focus on the best-performing strategies, that is, the most successful strategies from the Memory-One tournament.

A. 101 TOP HUNDRED EXPERIMENTS

As previously discussed in Section I.B.6, the initial conditions of a tournament can affect the winner. For example, in Axelrod's tournament, nice strategies needed to have a critical mass of other nice strategies in order to succeed, which Axelrod referred to as a "cluster" (Axelrod 66). In order to determine the size of this necessary cluster, 101 tournaments were run, each with a different ratio of nice and mean strategies.

First, the top 100 nice representations and the top 100 mean representations from the Memory-One Tournament were put into two separate lists.[4] Next, 101 tournaments were run. These 101 tournaments were labeled "tournament N_x," where x is an integer between 0 and 100, indicating the number of nice players in the tournament. For example, Tournament N_1 was populated by the top-scoring nice strategy (Mole) and the top 99 mean ones. Tournament N_2 had the top two nice players and the top 98 mean ones, and so on. Strategies that Abstain on the first move were particularly unsuccessful in the Memory-One tournament and were thus not explored further.

Two strategies emerged victorious from this set of 101 tournaments. One ("Viper") was a mean strategy that won tournaments N0 to N15. The other ("Turtle") was a nice strategy that won tournaments N16 to N101. We find that the minimum cluster size at which nice strategies

[4] Note that these are lists of strategy representations, not unique strategies, and could thus contain strategies that were functionally equivalent.

beat mean ones is 16, consistent with the expectation that a nice strategy would be most successful in a tournament in which there was a cluster of other nice strategies.

Table 15. Viper

Viper	D	A	A
D	D	D	A
	D	D	C

Table 16. Turtle

Turtle	C	D	D
C	D	A	A
	D	D	A

The nice strategy, Turtle, is more interesting for our purposes, since mean strategies tend to fail in environments where the majority of other strategies are nice. Turtle, like Mole, is a variant on Grim Trigger, where it tries to Cooperate, but never forgives an opponent for Defecting, and will never again Cooperate for the rest of the game if the opponent Defects once.

Turtle differs from Mole in that it has an Abstain in its AA spot and its AD spot. This means that Turtle will Abstain continuously against an always-Defecting opponent. Since Abstaining is the play that yields the highest payoff when the opponent Defects, this strategy seems more intuitive than Mole.

B. GENETIC LEARNING TOURNAMENTS

We made use of genetic learning techniques to attempt further improvements in strategy performance. Genetic learning is a form of machine learning that proceeds by repeating a process in successive iterations called generations. At each generation, the most

successful agent produces offspring that are similar but mutated in some small way, while the least successful agents are eliminated.

For purposes of this application, each strategy's "genotype" is its table representation. As we saw in the definition of equivalence (I.B.4.d), two representations can be of the same strategy equivalence class but have different genotypes. Distinguishing between genotypes and strategies is important as we start using genetic learning to converge upon a canonically best strategy, because although two equivalent strategies make all the same decisions in a game of IOPD, they will have different sets of offspring if they reproduce. We say two strategies are "related" or "relatives" if and only if their genotypes differ in exactly one cell.

1. Method for Genetic Learning

The same 101 tournaments in the previous experiment were repeated, but in each successive generation, the bottom-scoring strategies from the tournament in the previous generation were replaced with relatives of the winning strategy, mutated according to the rules below, and the tournament was re-run.

Offspring were generated by modifying exactly one of the ten parameters of the winning strategy. IOPD strategies have ten parameters, each of which can take one of three states. So, an IOPD strategy has 20 genotypically distinct offspring. If a given offspring was identical to one already in the tournament, then it was not added, and one fewer strategy would be dropped, thus keeping the next tournament at 100 players. Thus, every strategy in every tournament was genotypically unique.

2. Results of Genetic Learning

For each of the 101 initial conditions, we ran ten generations of new tournaments.[5] Another pair of victorious strategies emerged at the end of ten generations across all 101

[5] The number ten was chosen since, starting from any strategy, any other strategy can be obtained within 10 generations. Therefore, it was assumed that the best strategy would emerge within 10 generations. To verify this intuition, another genetic experiment with 20 generations was also run, and indeed the winner of the tournament was always the same for all 101 initial conditions.

initial conditions, each a genotypic relative of either Viper or Turtle. These strategies, Rattlesnake and Snail, respectively, are shown in Tables 17 and 18.

Table 17. Rattlesnake

Rattlesnake	D	C	A
D	D	D	A
	D	D	C

Table 18. Snail

Snail	C	D	A
C	D	A	A
	D	D	A

Again, the nice strategy, Snail, is more relevant to our search for the highest-performing strategy. Mole, Turtle and Snail are all very similar strategies. They have in common that they are all nice strategies, but have a “grim trigger” condition in that once their opponent Defects or Abstains, they will never again Cooperate for the remainder of the IOPD game. As well, the three strategies are defensive; they all avoid entering into a loop of constant mutual Defection by Abstaining after mutual Defection.

Turtle and Snail take this defensive tactic a step further against a repeatedly Defecting opponent, by Abstaining repeatedly. Therefore, if Turtle played against All Defect, it would Abstain for all except the first two moves. Snail is even more effective, Abstaining against a Defecting opponent after the first move. The fact that Turtle and Snail are able to converge on repeating the optimal move against an aggressive opponent is a big advantage that helps explain their success.

V. AXELROD REPEATED

Unlike the original Axelrod tournament, these studies considered all and only the memory-one strategies. It is certainly possible that a more sophisticated strategy could have won any or all of the previous tournaments. Thus, we played the winners of the previous tournaments against a sophisticated set of strategies based on expert strategy development in the literature, specifically the original Axelrod tournaments.

But the strategy space for even the deterministic memory-one strategies in IOPD is large compared to IPD, and for each possible degree of freedom added in adding strategic complexity, the strategy space increases exponentially. The desire to test more intricate strategies needed to be balanced against the practical ability to define and test these strategies. The main repository for complex strategies created by experts was the original Axelrod prisoner's dilemma tournament.

In the final experiment, the strategies from the original Axelrod tournament were adapted to optional prisoner's dilemma to the greatest extent possible. For some of these strategies, there was more than one possible way to adapt them to IOPD, so all possible variants were created and included as entries in the tournament. Other strategies in the original Axelrod tournament had logic that could not be adapted for IOPD, and thus had to be discarded.

Some of the strategies that had to be discarded were "tester" strategies, that would Defect in early moves to try to determine their opponent's strategy and exploit it. However, the mechanisms for both testing and exploitation were all programmed with the assumption of just two options. Other strategies that had to be discarded were ones such as Tideman & Chieruzzi, whose logic was based upon managing ratio of likelihood to either Cooperate or Defect. But where there are three options to either Cooperate or Defect, the logic of managing a ratio becomes more complicated. Rather than make barely functional copies, or grossly extrapolate the designer's intent, it seemed in the spirit of the tournament to remove these strategies. This had the net result of mostly leaving the simpler strategies in the tournament while removing complex ones, but that was an unavoidable side-effect.

A. STRATEGIES

Ultimately, this approach yielded 17 strategies. Nine were memory-one strategies and eight were more complex. They are listed below:

1. Deterministic Memory-One Strategies

- **All Cooperate:** Always Cooperates.
- **All Defect:** Always Defects.
- **Tit-for-Tat**: Starts by Cooperating, then replays the opponent's last move.

Table 19. Tit-for-Tat

Tit-for-Tat	C	A	D
C	C	A	D
	C	A	D

- **Nicer Tit-For-Tat**: Starts by Cooperating. Then replays the opponent's last move, unless it was "Abstain," in which case it Cooperates.

Table 20. Nicer Tit-for-Tat

Nicer Tit-for-Tat	C	C	D
C	C	C	D
	C	C	D

- **Grim Trigger**: Cooperates until the other player does not Cooperate, then Defects for the rest of the game.

Table 21. Grim Trigger

Grim Trigger	C	D	D
C	D	D	D
	D	D	D

- **Abra**: As Axelrod's Grim Trigger, but Abstains instead of Defecting.

Table 22. Abra

Abra	C	A	A
C	A	A	A
	A	A	A

- **Mole**: The winning strategy of the Memory-One Tournament.

Table 23. Mole

Mole	C	D	D
C	D	D	D
	D	D	A

- **Turtle**: The winning strategy from the Top Hundred tournament

Table 24. Turtle

Turtle	C	D	D
C	D	A	A
	D	D	A

Snail: The winning strategy of the Genetic Tournament.

Table 25. Snail

Snail	C	D	A
C	D	A	A
	D	D	A

2. Complex Strategies

These more complex strategies were either mixed strategies, or had memory greater than one, or both. They are not deterministic memory-one strategies and as such, the tabular representation we have been using thus far does not apply to them.

- **All Random**: Has a uniform random chance of choosing any one of the three moves. Axelrod included this strategy as a contestant in all of his tournaments.
- **True Joss**: As True Tit-for-Tat, but if it would Cooperate, then it has a 10% chance of Defecting instead.
- **Nicer Joss**: As Nicer Tit-for-Tat, but if it would Cooperate, then it has a 10% chance of Defecting instead.

- **Tit-for-Two-Tats**: Always Cooperates, unless opponent Defects twice in a row, in which case it Defects once (then goes back to Cooperating).
- **Grudging Tit-for-Two-Tats**: Always Cooperates, unless opponent Defects twice without Cooperating in between, in which case it Defects once (then goes back to Cooperating).
- **Davis**: Plays All Cooperate in the first 10 rounds, then switches to Grim Trigger.
- **Grofman**: Always Cooperates, unless the players did different things on the previous move, in which case it Defects with a 5/7 probability and Cooperates with a 2/7 probability.
- **Shubik**: Cooperates until the other player Defects, and then Defects for a number of consecutive moves equal to the sum total of previous Defections the opponent has made thus far in the game.

B. TOURNAMENT RESULTS

The results of the tournament are displayed in Table 26.

Table 26. Tournament Results for the Axelrod Repeated Tournament

Names	Scores
Snail	8997
Turtle	8979
Mole	8891
Grudging Tit-for-2-Tats	8866
Abra	8819
Tit-for-Two-Tats	8801
Grofman	8735

Names	Scores
All Cooperate	8642
Davis	8400
Grim Trigger	8390
Shubik	8383
True TFT	8282
Nicer TFT	8249
All Random	6330
True Joss	6219
Nicer Joss	5696
All Defect	5638

The results of this tournament lead us to several conclusions. First, the three strategies identified by the three methods employed in earlier experiments were the top 3 performers in this tournament. These nice, defensive, and unforgiving strategies clearly all perform extremely well in an IOPD setting, even against expert-crafted strategies.

Beyond the top three strategies, one other strategy in the tournament had the property of being nice, defensive and unforgiving: the Abra strategy. Abra behaves like Grim Trigger, but Abstains rather than Defecting once an opponent no longer cooperates. This strategy was considerably gentler to those that had wronged it than Snail, Mole and Turtle, since it would not Defect, only Abstain. Abra was thus giving up a few points in the tournament in places where it Abstained but could have Defected against an opponent that was Cooperating. Still, its presence in the top 5 reinforces the importance of these three qualities.

In fourth place is Grudging Tit-for-2-Tats, a complex strategy whose good performance indicates an area for further research. Unlike the other top five performers, it

is not unforgiving, nor is it strictly defensive. This strategy only Defects if its opponent Defects twice without Cooperating in between, in which case it Defects just once. This behavior does have the property of breaking infinite loops of mutual Defection, which is the point of defensiveness. Unlike regular Tit-for-Tat, when paired against Joss, this strategy will not be caught in a permanent loop of mutual Defection.

Davis and Grim Trigger are two very similar strategies that are nice and unforgiving, but not defensive. They came in 9th and 10th place out of 17, respectively, a middling performance that highlights the importance of defensiveness in OPD. Abstaining is the correct play against an opponent who is going to Defect, so it is important to exercise this option when it is the correct play.

Another area of further research would examine strategies that are nice and defensive, but not unforgiving. There were no such strategies in this final Axelrod Repeated tournament (although there were in the Memory-One tournament).

The four mixed strategies performed rather poorly, as they did in Axelrod's first tournament (Axelrod, 2005). True Joss and Nicer Joss both attempted to use randomness to sneak Defection in against their opponent, but ended up doing worse than the All Random strategy that behaved completely randomly. Grofman was the only mixed strategy with even middling performance, scoring 7th out of 17. Grofman scored 4th in the original Axelrod tournament, but was pushed lower in this tournament due to the success of Snail, Mole and Turtle.

C. FINAL CONFIRMATION

One final experiment was performed to verify Snail's dominance. Every single possible deterministic memory-one strategy was placed, in sequence, into the previous Axelrod tournament as an 18th player, and the tournament re-run with 18 players. Not a single strategy was able to beat Snail. This result, coupled with its performance in the first tournament and in the genetic top 100 tournament strongly suggests that Snail is the canonically superior strategy for IOPD, in the same way that TFT is for IPD.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSION

The naïve assumption, that Tit-for-Tat would be the optimal strategy for iterated optional prisoner's dilemma, was completely wrong. Tit-for-Tat did not do particularly well in any of the tournaments, finishing in 20308th place out of 59049 in the Memory-One Tournament, meaning it did not participate in our genetic learning experiments. And it finished 12th out of 17 in the adapted Axelrod tournament.

In every single IOPD tournament run in this thesis project, a trio of qualities separated the winners from the losers. The winning strategies were all nice, unforgiving and defensive. The capability to be defensive is what really distinguishes IOPD strategies from IPD strategies, since there is no option to Abstain in the traditional, non-optional Prisoner's Dilemma. Abstaining is the optimal move against a Defecting opponent, and a defensive strategy is powerful because it can make that optimal play.

If there were an Axelrod-style tournament for IOPD with open submissions, this author would submit Snail. Snail was the winner of both the Axelrod tournament and the genetic Top Hundred tournament, and it responds better than Turtle or Mole against a field of expert-developed strategies. Mole's victory in the first tournament probably indicates that Mole is better than Snail only when the opponents behave randomly and erratically, which would not be the case if the other contestants in the tournaments were experts in their fields.[6]

A. TOPICS FOR FURTHER RESEARCH

We assumed in this thesis that the optimal strategy of IOPD would be a memory-one deterministic strategy because the undisputed champion of IPD was, but this assumption might not be justified. Further work could examine strategies capable of more complex behavior, such as mixed strategies. The relatively good performance of the complex strategy Grudging Tit-for-Two-Tats suggests that perhaps, having the capacity to forgive a single Defection before entering the "grim trigger" state is worthwhile. Maybe a strategy that

[6] Interestingly, Snail placed in 3708th place in the Memory-One tournament, a decent showing out of 59,049 entrants but lower than one would expect out of the top-scoring strategy in two other tournaments.

combines this longer memory with the trifecta of nice, defensive and unforgiving qualities that were so successful would outperform Snail across a wide variety of tournament conditions.

The option to Abstain had a much bigger impact than was expected upon the nature of the high-performing strategies. The unforgiving nature of the most successful strategies poses interesting questions about how the option to simply disengage changes the dynamic of Optional Prisoner's Dilemma.

B. IMPLICATIONS

Real-world adversarial scenarios are often reducible to simple models in the form of strategic games. IOPD models a broad class of applications defined by the following properties: there is an advantage to cooperation, a temptation to defect, a substantial risk of the other agent defecting, and finally, the option to abstain.

The product of this thesis is a set of strong strategies that can perform well against a variety of actors in these cases, useful as both an analytic and decision-making tool for the myriad problems that have these properties. "Niceness" remains the default assumption in any IOPD situation as it is in IPD situations. One should trust their partner, and never be the first to betray their partner's trust. However, in an IOPD situation, once one's partner has betrayed them, it is best to either assume bad faith and avoid future interactions, or to assume that future overtures to cooperation are ultimately predatory, and to respond in kind.

APPENDIX: MEMORY-ONE EXPERIMENT RESULTS

These are the top 100 strategies from the Memory-One Experiment. 0 indicates Cooperation, 1 indicates Abstention and 2 indicates Defection. If a particular strategy has more than one representation, due to the fact that certain cells are never activated, then the irrelevant cells have been replaced with an **x**.

Rank	firstplay	CC	CA	CD	AC	AA	AD	DC	DA	DD	Score
1	0	0	2	2	2	2	2	2	2	1	32613726
2	0	0	2	2	1	2	2	2	2	1	32434320
3	0	0	2	2	2	1	2	2	2	1	32434320
4	0	0	2	2	2	2	1	2	2	1	32434320
5	2	**x**	**x**	**x**	2	2	2	2	2	1	32362926
6	2	2	2	1	1	1	1	2	2	0	32357264
7	2	2	2	1	1	1	2	2	2	0	32295560
8	2	2	2	1	1	2	1	2	2	0	32295560
9	2	2	2	1	2	1	1	2	2	0	32295560
10	0	2	2	2	2	2	2	2	2	1	32277633
11	2	0	2	2	0	2	2	2	2	1	32275122
12	2	0	2	2	2	2	0	2	2	1	32275122
13	2	0	2	2	2	0	2	2	2	1	32274825
14	2	2	2	2	0	2	2	2	2	1	32262486
15	2	2	2	2	2	2	0	2	2	1	32262486
16	2	2	2	2	2	0	2	2	2	1	32262390
17	2	1	2	2	0	2	2	2	2	1	32230113
18	2	1	2	2	2	2	0	2	2	1	32230113
19	2	1	2	2	2	0	2	2	2	1	32229917
20	2	2	2	1	1	2	2	2	2	0	32204876
21	2	2	2	1	2	1	2	2	2	0	32204876
22	2	2	2	1	2	2	1	2	2	0	32204876
23	2	2	1	1	2	2	2	2	2	0	32197586
24	2	2	1	1	1	2	2	2	2	0	32196452
25	2	2	1	1	2	1	2	2	2	0	32196452
26	2	2	1	1	2	2	1	2	2	0	32196452
27	2	0	2	2	0	2	0	2	2	1	32187318
28	2	0	2	2	0	0	2	2	2	1	32187021
29	2	0	2	2	2	0	0	2	2	1	32187021
30	2	0	2	1	1	1	2	2	2	0	32181404
31	2	0	2	1	1	2	1	2	2	0	32181404
32	2	0	2	1	2	1	1	2	2	0	32181404
33	2	2	1	2	0	2	2	2	2	1	32178246
34	2	2	1	2	2	2	0	2	2	1	32178246

Rank	firstplay	CC	CA	CD	AC	AA	AD	DC	DA	DD	Score
35	2	2	1	2	2	0	2	2	2	1	32178150
36	2	0	1	2	0	2	2	2	2	1	32170605
37	2	0	1	2	2	2	0	2	2	1	32170605
38	2	0	1	2	2	0	2	2	2	1	32170308
39	2	**x**	**x**	**x**	1	2	2	2	2	1	32170146
40	2	**x**	**x**	**x**	2	2	1	2	2	1	32170146
41	2	**x**	**x**	**x**	2	1	2	2	2	1	32170047
42	2	2	2	2	0	2	0	2	2	1	32162046
43	2	2	2	2	0	0	2	2	2	1	32161950
44	2	2	2	2	2	0	0	2	2	1	32161950
45	2	1	1	2	0	2	2	2	2	1	32145873
46	2	1	1	2	2	2	0	2	2	1	32145873
47	2	1	1	2	2	0	2	2	2	1	32145677
48	2	0	2	1	1	1	1	2	2	0	32143514
49	2	2	0	2	0	2	2	2	2	1	32133156
50	2	2	0	2	2	2	0	2	2	1	32133156
51	2	2	0	2	2	0	2	2	2	1	32133060
52	0	0	1	2	2	2	2	2	2	1	32130561
53	0	0	2	1	2	2	2	2	2	1	32130561
54	2	0	2	1	1	2	2	2	2	0	32125100
55	2	0	2	1	2	1	2	2	2	0	32125100
56	2	0	2	1	2	2	1	2	2	0	32125100
57	2	2	2	1	2	2	2	2	2	0	32122256
58	2	0	2	2	0	0	0	2	2	1	32099217
59	2	0	1	1	2	2	2	2	2	0	32098649
60	2	0	0	2	0	2	2	2	2	1	32089065
61	2	0	0	2	2	2	0	2	2	1	32089065
62	2	0	0	2	2	0	2	2	2	1	32088768
63	2	2	1	1	1	1	2	2	2	0	32081234
64	2	2	1	1	1	2	1	2	2	0	32081234
65	2	2	1	1	2	1	1	2	2	0	32081234
66	2	1	0	2	0	2	2	2	2	1	32080155
67	2	1	0	2	2	2	0	2	2	1	32080155
68	2	1	0	2	2	0	2	2	2	1	32079959
69	0	2	2	2	1	2	2	2	2	1	32076429
70	0	2	2	2	2	2	1	2	2	1	32076429
71	0	2	2	2	2	1	2	2	2	1	32076329
72	2	2	2	1	0	2	2	2	2	1	32074512
73	2	2	2	1	2	2	0	2	2	1	32074512
74	2	2	2	1	2	0	2	2	2	1	32074416
75	2	2	0	1	1	1	1	2	2	0	32063522
76	2	2	2	2	0	0	0	2	2	1	32061510
77	2	0	2	2	0	2	1	2	2	1	32058825
78	2	0	2	2	1	2	0	2	2	1	32058825
79	2	0	2	2	0	1	2	2	2	1	32058726

Rank	firstplay	CC	CA	CD	AC	AA	AD	DC	DA	DD	Score
80	2	0	2	2	2	1	0	2	2	1	32058726
81	2	1	1	1	2	2	2	2	2	0	32058594
82	2	0	2	2	1	0	2	2	2	1	32058528
83	2	0	2	2	2	0	1	2	2	1	32058528
84	2	2	2	1	0	1	1	2	2	0	32055596
85	2	2	2	1	1	0	1	2	2	0	32055596
86	2	2	2	1	1	1	0	2	2	0	32055596
87	2	0	2	1	2	2	2	2	2	0	32053370
88	2	0	2	1	0	2	2	2	2	1	32047863
89	2	0	2	1	2	2	0	2	2	1	32047863
90	2	0	2	1	2	0	2	2	2	1	32047566
91	2	1	2	1	0	2	2	2	2	1	32042139
92	2	1	2	1	2	2	0	2	2	1	32042139
93	2	1	2	1	2	0	2	2	2	1	32041943
94	2	1	2	2	0	2	0	2	2	1	32036604
95	2	1	2	2	0	0	2	2	2	1	32036408
96	2	1	2	2	2	0	0	2	2	1	32036408
97	2	0	1	1	1	2	2	2	2	0	32035289
98	2	0	1	1	2	1	2	2	2	0	32035289
99	2	0	1	1	2	2	1	2	2	0	32035289
100	2	2	2	2	0	2	1	2	2	1	32032554

THIS PAGE INTENTIONALLY LEFT BLANK

SUPPLEMENTALS

Readers interested the following supplemental files the student should find links available with the main thesis's catalog entry in the Naval Postgraduate School Institutional Archive, Calhoun. For further information, please contact the NPS library.

A. SUPPLEMENTAL 1: MEMORYONE

MemoryOne.py is the experiment code described in Chapter IV. The Excel spreadsheet, MemoryOne Results_sorted.xls, is the sorted output of that program, which takes several hours to run.

B. SUPPLEMENTAL 2: GENETIC TOURNAMENT, MEANPLAYERS, AND NICEPLAYERS

GeneticTournament.py is the genetic learning experiment code described in Chapter III. The two Excel files, meanPlayers.xls and nicePlayers.xls are necessary inputs to run the code and must be in the same folder as GeneticTournament.py when it is run.

C. SUPPLEMENTAL 3: AXELROD REPEATED

AxelrodRepeated.py is the experiment code described in Chapter IV. Note that this program includes some amount of randomization and so its output might be slightly different than what is in the body of the thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Arend, R. J., & Seale, D. A. (2005). Modeling alliance activity: An iterated prisoners' dilemma with exit option. *Strategic Management Journal, 26*(11), 1057–1074. doi:10.1002/smj.491

Axelrod, R. (1980). Effective Choice in the Prisoner's Dilemma. *Journal of Conflict Resolution, 24*(1), 3–25. doi:10.1177/002200278002400101

Axelrod, R. (1980). More Effective Choice in the Prisoner's Dilemma. *Journal of Conflict Resolution, 24*(3), 379–403. doi:10.1177/002200278002400301

Axelrod, R. (2006). *The evolution of cooperation*. Basic Books.

Batali, J., & Kitcher, P. (1995). Evolution of altruism in optional and compulsory games. *Journal of Theoretical Biology, 175*(2), 161–171. doi:10.1006/jtbi.1995.0128

Cardinot, M., O'Riordan, C., & Griffith, J. (2016). The optional prisoner's dilemma in a spatial environment: coevolving game strategy and link weights. *Proceedings of the 8th International Joint Conference on Computational Intelligence, 1*, 86–93. doi:10.5220/0006053900860093

Case, Nicky. (2017) *The Evolution of Trust.* https://ncase.me/trust/

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Heuer, L., & Orland, A. (2019). Cooperation in the Prisoner's Dilemma: An experimental comparison between pure and mixed strategies. *Royal Society Open Science, 6*(7), 182142. doi:10.1098/rsos.182142

Majeski, S. J. (1984). Arms races as iterated prisoner's dilemma games. *Mathematical Social Sciences, 7*(3), 253–266. doi:10.1016/0165-4896(84)90022-2

Mero, L. (2012). *Moral Calculations: Game Theory, Logic, and Human Frailty*. Springer.

Nowak, M. A. (2006). *Evolutionary dynamics: Exploring the equations of life*. Belknap Press of Harvard University Press.

Porter, M. (1991, March 27). Coca-Cola vs. Pepsi-Cola and the soft drink industry. Retrieved September 01, 2020, from https://www.hbs.edu/faculty/Pages/item.aspx?num=12066

Press, W. H., & Dyson, F. J. (2012). Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences, 109*(26), 10409-10413. doi:10.1073/pnas.1206569109

Skyrms, B. (2001). The Stag Hunt. *Proceedings and Addresses of the American Philosophical Association, 75*(2), 31–41. doi:10.2307/3218711

Snyder, G. (1971). "Prisoner's Dilemma" and "Chicken" Models in International Politics. *International Studies Quarterly, 15*(1), 66–103. doi:10.2307/3013593

Tucker, A. (1983). The Mathematics of Tucker: A Sampler. *The Two-Year College Mathematics Journal, 14*(3), 228–232. doi:10.2307/3027092

Wayland, R. (1991). Coca-Cola vs. Pepsi-Cola and the Soft Drink Industry. *Harvard Business Review*.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California