The Success of TIT FOR TAT in Computer Tournaments

Robert Axelrod, 1984 “THE EVOLUTION OF COOPERATION”

Presenter: M. Q. Azhar (Sumon)

ALIFE

Prof. SKLAR

FALL 2005

Topics to be discussed

- Some background
  - Author
  - Prisoner’s dilemma
  - Motivation of work
- Iterated Prisoner’s Dilemma
- The Computer Tournament
- The Results of the tournaments
- Comments on the results
- Conclusion.

About the Author and his work: Robert Axelrod and "The emergence of Cooperation"

- Robert Axelrod is a political scientist and a mathematician.
- Distinguished University Professor of Political Science and Public Policy at the University of Michigan.
- He is best known for his interdisciplinary work on the evolution of cooperation which has been cited in more than five hundred books and four thousand articles.
- His current research interests include complexity theory (especially agent based modeling) and International security.
- To find out more about the author go t0
  - **Robert Axelrod's homepage**
  - **http://www-personal.umich.edu/~axe/**

What is Prisoner’s Dilemma?

- **Where does it all start?**
  - The prisoner's dilemma is a game invented at Princeton's Institute of Advanced Science in the 1950's.
- **What is Prisoner’s Dilemma?**
  - The **prisoner's dilemma** is a type of non-zero-sum game (game in the sense of Game Theory). Non-zero sum game means the total score distributed among the players depends on the action chosen. In this game, as in many others, it is assumed that each individual player ("prisoner") is trying to maximize his own advantage, without concern for the well-being of the other players.
- **The heart of the Dilemma:**
  - In equilibrium, each prisoner chooses to defect even though the joint payoff would be higher by cooperating. Unfortunately for the prisoners, each has an individual incentive to cheat even after promising to cooperate. This is the heart of the dilemma.

Prisoner's Dilemma

- B and C are collectively charged with a crime and held in separate cells, with no way of meeting or communicating. The Detective told them that:
  - if B confesses and C does not, the B (confessor) will be freed, and C will be jailed for three years.
  - if both of them confess, then each of them will be jailed for two years
- Both of them know that if neither confesses, then they will each be jailed for one year

Pay off Matrix for Prisoner's Dilemma

- **1st Case (Top left):** If both of them defect, then both get punishment for mutual defection 2 years
- **2nd Case (Top right):** If B cooperates and C defects, B gets sucker's payoff of 3 years, while C gets free or 0 years.
- **3rd Case(Bottom left):** If C cooperates and B defects, Clyde get sucker's payoff of 3 years, while B gets 0 years
- **4th Case (Bottom right)**: Reward for mutual cooperation. Both of them cooperate, They get 1 years.

		B			
		defect		coop	
C	DEFECT		2		0
		2		3	
	COOP		3		1
		0		1	

Analysis of PD

- Thus there are **two choices**—
  - **to cooperate**, (in this scenario remain silent),
  - **or to defect**, which here means to confess.
- There are **four possible outcomes**, depending on one prisoner's move: the other prisoner's may:
  - 0, 1, 2 or 3 years in prison.
- **Results of Cooperation:**
  - The prisoners serve one or three years.
- **Results of Defection:**
  - Prisoners may serve 0 or 2 years.
- **The Dilemma**:
  - Because B does not know whether C is trust worthy. Besides, there is no opportunity to communicate when deciding the move.
  - So defection is the best response to all possible strategies. Most rational prisoners will choose to defect in order to maximize the upside (0 years) and minimize the downside (only 2 years instead of 3).
  - Although the outcome consistently is better for two cooperating prisoners (1 years) than for two defecting players (2 years).

What does Prisoner’s Dilemma signify?

- One may comes to the following conclusion from the PD analysis:
  - According to the game theory, prisoners should act rationally and cooperate to maximize their global profit. However, prisoner’s dilemma says otherwise! Does it mean that game theory’s notion of rational action is wrong?
  - Perhaps, the dilemma has flaws!
- Under which conditions the prisoners will cooperate?
  - The prisoners are not selfish.
  - The other prisoner is prisoner’s family (twin brother) or lover (bonnie and clyde)!
  - The shadow of the future that will force someone to cooperate.
    - (i.e if the prisoner doesn’t cooperate, something bad will happen when he gets released…mafia/organized crime?)

Introducing
Iterated Prisoner's Dilemma

- **Questions Remained:**
  - When should a person cooperate, and when should a person be selfish, in an ongoing interaction with another person?
  - Should a friend keep providing favors to another friend who never gives favor back in return?
- **Introducing Iterated Prisoner's Dilemma:**
  - The simple way to represent this type of situation is to play prisoner's dilemma game more than once (iteratively)!
  - If you know that you will be meeting your opponent again, then the incentive to defect appears to disappear.
  - The game allows the players to achieve mutual gains from co operation, but it also allows for the possibility that one player will exploit the other, or the possibility that neither will co operate.

Motivation:
A new Approach for IPD

- Prisoner's dilemma is everywhere from international relation to personal relations.
- It would be useful to know how best to act when in this type of setting.
- But there is no one best strategy. What is best depends on what other player is likely to be doing.
- Further, what the other is likely to be doing will depend on what the other player expects you to do.
- There have been lot of literature written.
- A lot of models of important political, social and economic processes have the Prisoner's Dilemma as their foundation.
- Unfortunately, none of the past literatures (during 80s) reveals very much how to play the game well.
- Therefore, author recognizes the need for a new approach

Motivation (Cont.)
Computer Tournaments

- The new approach would have to draw on people who have
  - rich understanding of the strategic possibilities inherent in a non-zero-sum game setting, a situation in which the interests of the participants partially coincide and partially conflict.
- The two important facts about non-zero-sum settings would have to be taken into account.
  - First, what is effective depends on not only upon the characteristics of a particular strategy, but also upon the nature of the other strategies with which it must interact.
  - Second point follows directly from the first. An effective strategy must be able at any point to take into account the history of interaction as it has developed so far.
- A computer tournament for the study of effective choice in the iterated prisoner's dilemma meets these needs.

Computer Tournament for IPD (Round 1)

- **Selection Procedure:**
  - Experts in **game theory** had been invited to submit programs like a computer Prisoner's Dilemma Tournament.
  - Most of the entrants were recruited from those who has published articles on game theory in general or the Prisoner's Dilemma.
- **Introducing the Players:**
  - The fourteen entries submitted came from five disciplines: **psychology, economics, political science, mathematics and sociology**.
- **Rules of Engagement:**
  - Each program had the **access to history of interaction** so far.
  - Each program could make choices (I.e Cooperate or Defect) based on this history.
  - There were f**ourteen entries** and a **random rule . RANDOM** program randomly cooperates and defects with equal probability.
  - The game was **structured as round robin**, meaning each entry was paired with each other entry.
  - Each entry was also paired with its own **twin** and with **RANDOM**.

The Tournament Setup (cont)

- **Rules of Engagement (cont):**
  - Each game consists of exactly 200 moves.
  - Entire round robin tournament was played five times to get a more stable estimates of the score for each player.
  - Each computer game was played against each other for five games.
  - There were **120,000 moves, making for 240,000 separate choices.**

Payoff Matrix for the tournament

		PLAYER “A”			
		defect		coop	
PLAYER “B”	DEFECT		1		0
		1		5	
	COOP		5		3
		0		3	

Meet the Winner (Round 1)

TIT-FOR-TAT:

- Submitted by: Professor Anatol Rapoport of the University of Toronto.
- This was simplest (only five lines of Fortan code) and it turned out to be winner.
- This strategy is based on the following rules:
  - Cooperate on the first round.
  - On round t>1, do what your opponent did on round t-1.
  - This decision rule is probably the most widely known and most discussed rule for playing PD.

Analysis of the results

(Round 1)

- Analysis of the results show that the following did not contribute to the rule's relative success:
  - Discipline of the author
  - Shortness of the program
  - Length of the program
- Then What does?

Analysis of the results (cont.)

(Round 1)

- Interpretation of the Scores

- **Good performance!**
  - In a game of 200 moves, a useful benchmark for very good performance is 600 points. This score was attained by a player when both sides cooperate with each other.
- **Bad performance!**
  - A useful benchmark for bad performance is 200 point. This score was attained by a player when both sides never cooperate with each other.
- **Range**
  - Most scores range between 200 and 600. The scores from 0-1000 points were possible.
- **Winner's Average**
  - The winner TIT FOR TAT, averaged 504 points per game.

Analysis of the results (cont.)

(Round 1)

- **BE NICE! BE NICE!..**
  - Surprisingly, The property of **“BEING NICE”** is the single property which **distinguishes** the relatively **high-scoring entries** from the **relatively low-scoring entries.**
  - Being nice means that a nice rule will not first to defect!
  - For this tournament, nice rules are those who were not the first to defect before the last few moves (i.e 199^{th} moves).
  - Each of the eight top ranking rules (entries) is nice.
  - None of the other entries is!
  - The nice entries received tournament averages between 472 and 504, while the best entries that were not nice received only 401 points.
  - Nice rules did well because there were enough of them to raise substantially each other’s average score.
  - There are also the Kingmakers who play an important role determining the ranking among the top contenders.

Analysis of the results (cont.)

(Round 1)

- Meet the Kingmaker
- **The most important Kingmaker (DOWNING) :**
  - **DOWNING** is one of the Kingmaker rules which did not do well for themselves but they largely determine the rankings among the top contenders.
  - **DOWNING** is based on an “outcome maximization” principle originally developed as a possible interpretation of what human subjects do in the Prisoner’s Dilemma lab experiments (Downing 1975).
- **DOWNING’s strategy:**
  - It will try to understand its opponent and then make the choice to maximize its score in the long run.
  - If the other player is not responsive to what DOWNING was doing, DOWNING will try to get away with whatever it can by defecting.
  - If the other player seems responsive, it will cooperate.

Analysis of the results (cont.)

(Round 1)

<u>Where did DOWNING fail?</u>

- Although it's a sophisticated rule, DOWNING's implementation has some flaw.
- It initially assumes that the other player is unresponsive. DOWNING is doomed to defect on the first two moves.
- These first two defections led many other rules to punish DOWNING. So, things usually got off to a bad start.
- This is why DOWNING served so well as a Kingmaker. Paying with first ranking "TIT-FOR-TAT" and the second ranking "TIDEMAN AND CHIERUZZI), It learned quickly that defection does not pay but the co operation does.
- All of the other nice rules went downhill with DOWNING.

Analysis of the results (cont.)

(Round 1)

- **BE FORGIVING!!**
  - Forgiveness of a rule can be informally described as its propensity to cooperate in the moves after the other player has defected.
  - The least forgiving rules among the nice rules scored the lowest.
  - FRIEDMAN, a totally unforgiving rule that employs permanent retaliation. It is never first to defect, but once the other defects ever once, FRIEDMAN defects from then on.
  - In contrast, the winner, TIT FOR TAT, is unforgiving for one move, but thereafter totally forgiving of that defection. After one punishment, it lets bygones by bygones!

Analysis of the results (cont.)

(Round 1)

- Where did “not so nice rules” fail?
- Most of the “not so nice rules” were not forgiving.
- For instance, **JOSS:**
  - a sneaky rule that tries to get away with occasional defection.
  - It always defects immediately after the other players defect.
  - But instead of cooperating after the other player cooperates, 10% of the time it defects after the other player cooperates.
  - Although, it’s a variation of TIT FOR TAT, its overall performance was terrible.
  - The problem was a combination of an occasional defection after the other’s cooperation by JOSS, combined with lack of forgiveness by both sides.
- Moral of this result:
  - If both sides retaliate in the way that JOSS and TIT FOR TAT did, its not worth being as greedy as JOSS was.

Major lesson of this tournament

- Importance of minimizing echo effects in an environment of mutual power.
- A single defection can set off long string of recriminations and counter recriminations and both sides suffer.
- More sophisticated analysis of choice must go at least three levels deep:
  - **Level 1 analysis: *Direct effect of a choice***. Defection earns more than co operation.
  - **Level 2 analysis**: ***Considers the indirect effect***. Take into account that the other side may or may not punish a defection.
  - **Level 3 analysis:** Although defection can be justified by level 1 and level 2 analysis, a lot of the players did not take into account the ***tertiary effect when one's isolated defections can turn into unending mutual recriminations***.

Rules that could have been winner in First Tournament

- **TIT FOR TWO TAT**
  - More forgiving.
  - Defects only if the other player defected on the previous two moves.
- **LOOK AHEAD**
  - Inspired by techniques used in AI programs to play chess.
  - This was the rule that won preliminary tournament
- Modified DOWNING
  - If DOWNING had started with the initial assumptions that the other player would be responsive rather than unresponsive.
  - Then it would too had won by a large margin.
  - A Kingmaker could have been King.

End Game Effect

- **<u>Backwards Induction:</u>**

  Both player knows that they will play the game exactly *n* times. On round *n - 1*, player have an incentive to defect, to gain that extra bit of payoff…But this makes round *n – 2* the last "real", and so the player have an incentive to defect there, too and so on.This is the **backwards induction** problem.
- Playing the prisoner's dilemma with a fixed, finite, pre-determined, commonly known number of rounds, defection is the best strategy.

The Computer Tournament

(Round 2)

- **Motivation:**
  - The effectiveness of a particular strategy depends not only on its own characteristics, but also the nature of the other strategies with which it must interacts.
  - For this reason, the results of a single tournament are not definitive.
  - The second round provided a chance to both test the validity of the themes developed in the analysis of the first round
  - To develop new concepts to explain successes and failure**.**
- **Selection Procedure:**
  - The contestants were largely recruited through announcements in journals for users of small computers.
  - **Game theorists** participated in the first round of the tournament were also invited to try again.

The Computer Tournament

(Round 2)

- **Introducing the Players:**
  - The **sixty two** entries submitted came from **six countries**.(US, CANADA, Great Britain, Norway, Switzerland and New Zealand)
  - Contestants ranged from a **ten-year-old computer hobbyist to professors of computer science, physics, economics, psychology, mathematics, sociology,political science and evolutionary biology.**
  - There were 3969 ways the 63 rules (including RANDOM) were paired in the round robin tournament.
  - There were over a **million moves** in the second round.

The Computer Tournament

(Round 2)

Rules of Engagement:

- **Same as round 1 except the following changes:**
  - **Minor end game effects have been eliminated.**
  - **The length of the game was determined probabilistically. No one knew when the last move would come.**
  - **So end-game effects were successfully eliminated.**

- **Other Factors:**
  - Second round were given all detailed analysis of the first round, including discussion of the supplemental rules that would have done very well in the first round (i.e TIT FOR TWO TAT etc)
  - The entrants also drew their own lessons from the experience of the first round.
  - They were aware not only of the outcome of the first round, but also of the concepts used to analyze success, and the strategic pitfalls that were discovered.
  - They each knew that other knew these things.
  - Thus, second round began at a much higher level of sophistication.
  - It was interesting to see how entries based on different lessons actually interact.

Rules for some strategies from the Second Tournament!

- TESTER:
  - Submitted by David Gladstein and ranked 46th tournament.
  - On 1st round, defect. If the opponent retaliated, then play TIT-FOR-TAT. Otherwise intersperse cooperation and defection,
  - It provides low scores for some easy going rules.
  - TIT FOR TAT vs TESTER:
    - TESTER never defects twice in a row.
    - For instance, TIT FOR TWO TATS always cooperate with TESTER and gets badly exploited for its generosity.
  - THREE REVISED DOWNING vs TESTER:
    - All three were exploited by TESTER even though they would have looked promising in the first round.
    - They all calculated that the best thing to do with a program that cooperated just over half the time after one's own cooperation was to keep on cooperating.

Rules for some strategies from the Second Tournament!

- TRANQUILIZER:
  - Submitted by Craig Feathers and came in 27th in the tournament.
  - It first seeks to establish a mutually rewarding relationship with the other player,
  - only then does it cautiously try to see if it will be allowed to get away with something.
  - It does not defect twice in succession if it maintains at least average pay off at least 2.5 points per move. It does not defect more than one quarter of the time.
  - Its not greedy.

Analysis of the results
(Round 2)

- Factors that did not effect the game:
  - Personal attributes of the contestants.
  - Use of different programming language (FORTAN vs BASIC).
  - Short programs vs long programs.
- Factors that effected the game:
  - Being first to defect was costly.
  - Substantial correlation between whether a rule was nice and how well it did
  - Of the all top fifteen rules, all but one were nice (and that one ranked the eighth.)

Analysis of the results

(Round 2)

- How promptly and how reliably they responded to a challenge by the other player.
- The winner is still “TIT FOR TAT”.
- It is nice, forgiving and retaliatory.
- **Question remained:**
  - Would the results of the second round have been much different if the distribution of entries have had been substantially different?
  - Does TIT FOR TAT do well in a wide variety of environment? Is it Robust?

More experiments!

- Construct a series of hypothetical tournaments, each with a very different distribution of the types of rules participating.
  - The results were that TIT FOR TAT won five of the six variants of the tournaments and came in second in the sixth. A strong test to show how robust TIT FOR TAT is.
- Simulation using evolutionary biology
  - TIT FOR TAT still came number one!

An Axelrod’s guide for Success in IPD

- Axelrod suggests the following rules for succeeding in his tournament:
  - **Don’t be too competitive:** Don’t play the game as if it were zero sum game. Remember the fact that both can be a winner!
  - **Be nice, Be Nice..:** Start by cooperating, and reward those who cooperate.
  - **Be Forgiving:** Always reward cooperation immediately
  - **Revenge with moderation:** Always punish defection immediately by defecting, but be moderate. (hints: remember level 3 analysis)

Conclusion.

- There is no absolutely best rule independent of the environment.
  - What can be said for the successes of TIT FOR TAT is that it is a very robust rule.
  - Part of its success might be that other rules anticipate its presence and are designed to do well with it.
- TIT FOR TAT’s robust success is its combination of being nice, retaliatory, forgiving and clear!
  - Its **niceness** prevents it from getting into unnecessary trouble.
  - Its **retaliation** discourages other side from persisting whenever defection is tried.
  - Its **forgiveness** helps restore mutual cooperation.
  - Its **clarity** makes it intelligible to other player, thereby eliciting long-term cooperation.