Survey of strategies used in Iterated Prisoner's Dilemma Problem

Priyanka Gupta
M. Tech. Student University School of Information Technology,
GGS Indraprastha University Kashmere Gate, Delhi-6
Email: priya7rg5@gmail.com,

Abstract- *Game theory is a decision making process with two or more players. In the game theory each player has an intention to partly or completely conflict with each other. The attitude of each player has an important role to take the optimal decision .The Iterated Prisoner's Dilemma is widely considered as a standard model in the evolution of game theory. The payoff structure associated with Iterated Prisoners Dilemma having a various different kind of strategies as evevaluated with real life. This paper gives the idea of various different kinds of strategies related with the Iterated Prisoner's dilemma Problem and survey of techniques used to solve problem in various scenarios.*

Key words: Game Theory, cooperative game, Iterated Prisoner's Dilemma strategies, non cooperative game, payoff

1, INTRODUCTION

In the game theory strategic interaction among the players produce the outcome with respect to the player's preference [3] In the complex environment individuals are not able to evaluate the situation and calculate the optimal strategy Instead they want to adept their strategy based on the situation what has been affective or what has not. [4] Most of us realize there is a robust relationship between the game theory and the real world. [5] In the real word it is difficult to immediately analyze the strategy, so here we are trying to gives an idea about the strategies associated with the prisoner's dilemma that can also be applied in the real world situation.

The Iterated Prisoner's Dilemma is widely regarded as a standard model in the evolution of game theory. It constitutes a problem – Imagine there are two criminals, arrested under the suspicion that they have done the crime together but there is no sufficient proof against them. The prisoner's are isolated from each other, a supervisor visit each of them and offer a deal- the one who give the evidence against another one will be freed. If none of them accept the deal and cooperate then they both get few year of jail. If one of them gives the proof against another one then the defector will be freed and another one gets the jail. If both prisoner's betray then both get the full punishment because the supervisor got the sufficient proof against them .Since both the Prisoner's are isolated from each other therefore no one can judge the what the other one do. [6]

As in Game theory each player (Prisoner's) wants to maximize his /her own payoff without any concern with the other player's (Prisoner's) payoff. The dilemma says, only two actions are possible here either they can cooperate (C) or defect (D). If one of the Prisoner's defects then the other Prisoner's has an option either to cooperate acquire sucker's payoff S or defect acquire the punishment P for mutual defection. The

opponent Prisoner obtains Temptation T for the defect and if both cooperate then they receive reward R for the cooperation. The satisfactory condition for Payoff is [10]
T>R>P>S

With $R \geq \frac{(S+T)}{2}$

The above relationship can also be represented in the following form
DC>CC> DD> CD

Here DC shows One Prisoner defect and the other Prisoner cooperate similarly for CD one Prisoner cooperates and the other one defect and CC means both the Prisoner's do the cooperate, DD both Prisoner do the defect. The above relationship are shown below by the payoff Matrix

Prisoner (2) / Prisoner (1)	Cooperate (C)	Defect (D)
Cooperate (C)	R, R	S, T
Defect (D)	T, S	P, P

In the **iterated prisoner's dilemma** both the Prisoner's interact with each other repeatedly. Thus each Prisoner has an opportunity to punish the other Prisoner for previous non-cooperative behavior. If the number of steps is known by both Prisoners in advance, economic theory says that the two players should defect again and again; no matter how many times the game is played. Only when the players play an indefinite or random number of times can cooperation be an economic equilibrium. In this case, the incentive to defect can be overcome by the threat of punishment. [7] whet they both interact with each other repeatedly then there is no way to judge how the cooperation can be achieve form the non-cooperative state .In this paper we are trying to define all the strategies to solve this IPD Problem. In many setting, both the prisoner's may meet more then once. If an individual can recognize a previous interact ant and remember some aspects of previous outcome then the strategic situation become iterated prisoner dilemma. [11]

Iterated Prisoner's Dilemma strategies:

1.0 Dominant Strategy (Always Defect Strategies)

This strategy gives the best outcome to the player regardless of the decision of the other player .The Prisoner's always defect regardless the behavior of the other player. In this case Nash Equilibrium will be difficult to achieve. Defection is a best strategy in a single round IPD.

1.1 Always cooperate Strategies (*ALL C)*

In this case the Prisoner always try to cooperate they never goes for the defect part. The main characteristic of this strategy is **non-envious**, i.e. not striving more than the opponent.

1.2 Randomly choose strategies *Always set 50% probability.*

In this case each prisoner does randomly select the strategies. Sometimes he cooperates or sometimes he defect. If so many time he cooperate then he can attempt to defect for a once or more than one or from the longer time he are doing defect then he can attempt to cooperate or he can do C or D in the periodic manner. In this

strategies prisoner behavior can be entirely random, periodic random or partially random.

1.3 Nucleus -cooperative Strategies, in this strategies Prisoner's always behave cooperatively. The other Prisoner's behavior with never effect his Strategies

1.4 Aggressive Strategies In this strategies Prisoner's always behave non-cooperatively. in the very first move he behave cooperatively and in the other next move he always behave non cooperatively .(start with C and then DDDDDDDDDDD)

1.5 Shield – Pragmatic Strategies, in this strategies prisoner behave protectively. He wants to save his life (years of life). In this strategy he always keep record of the other Prisoner previous move and then take the decision. If in the previous move the other prisoner cooperate then he will cooperate or for non cooperate behavior sometimes he cooperate or sometimes not cooperate.

Deterministic Strategies

In this strategy Prisoner keep track record of other prisoner previous move and then take the decision accordingly. This strategy work if the average chance to meet to the other prisoner more than the Θ times in a row is w^{Θ} (where $0<w<1$). Here prisoner behave cooperate (C) or defect (D) according to the history of interaction so far.[8]

1.6 Tit –for-Tat

"Be nice, but punish any defections"

Prisoner (2) / Prisoner (1)	Cooperate (C)	Defect (D)
Cooperate (C)	R= 2 R=2	S=0 T=3
Defect (D)	T=3 S=0	P=1 P=1

Fig: The payoff matrix for the prisoner dilemma

The two prisoners have two options either cooperates or defect .If one of them defect against the cooperation of other prisoner then he got the maximum year of punishment of 3 years (instead of getting 2 years punishment) and he got the 1 year punishment (instead of getting no punishment) against the defect of other prisoner. If both the prisoners cooperate then they both get the 2 years of punishment.

This strategy is entirely based on the history of the previous interaction. In the first interaction prisoner's always start to behave with cooperate and then always copy the strategy of the other prisoner in the previous move. Initially start with C and then D if the opponent prisoner defect or C if the opponent prisoner cooperates. Starts by cooperating and, after that always does what the other player did on the previous round [9]

Prisoner 1	C C D C D C C C C D D C D D C . . .
Prisoner 2 using TFT	C C C D C D C C C C D D C D D . . .

Fig: Shows how this strategy works. Tit-for Tat strategy has three characteristics. **Nice** - it will not defect before its opponents do this characteristic make it top scoring strategy, **Retaliatory** - punishes with defection if opponent does the defection in the previous move, this characteristic (not be a blind optimistic) make it successful strategy and **Forgiving** -once again fall back to cooperating if the opponent does not continue to defect. In the long run, TFT outcompeted its rivals and went to fixation

$$W \geqslant \max\left\{\left\{\frac{T-R}{T-P}\right\}; \left\{\frac{T-R}{T-S}\right\}\right\}$$

Trigger strategy
This strategy is employed in the non-cooperative environment .the prisoner employed this strategy initially play with cooperate but punish the opponent if a certain level of defection occur.

1.7 Tit-for-Two Tats (TF2T)
Allowing two consecutive Defects before retaliating
This strategy is also entirely based on the history of the previous interaction. In the first interaction prisoner's always start to behave with cooperate and then always copy the strategy of the other prisoner in the previous move but the difference is Initially start with C and then D if the opponent prisoner defect twice and once again fall back to cooperating if the opponent does not continue to defect.

Prisoner1	C C D D C C D C C C D D C C D D C . . .
Prisoner2 using TF2T	C C C C D D C C C C C C D C C C D . . .

1.8 Grim Trigger

Start with the cooperate and punish the opponent for non-cooperate and never retort to cooperate (punish the opponent indefinitely) once the opponent do the defect

Prisoner 1	C C C C C C D D C D C D C D C. .
Prisoner2 Grim trigger	C C C C C C C D D D D D D D D. . .

1.9 Suspicious Tit –for-Tat
STFT, playing D in the first interaction
This strategy is the identical to the Tit-for-Tat strategy accept first move start with the defection instead of cooperation and then copy the opponent previous move.

Prisoner 1	C C D C D C C C C C D D C D D C . . .
Prisoner2 STFT	D C C D C D C C C C C D D C D D . . .

2.0 Observer Tit-for-Tat
Gathering the information about future interaction and reduce the defector pay off
Observer Tit-for-Tat is analogous to the Tit-for-Tat accept that it can be start with defect, If this partner was seen defecting on another. OTFT can do better than the TFT if W (0<w<1) is relative small and the proportion of defector is sufficiently high.[13]

2.1 Tit For Tat / Random
Repeat the opponent last choice and sometime skewed by random setting
This strategy is also entirely based on the history of the previous interaction but sometime twisted by random selection.

Prisoner 1	C C D C D C C C C C D D C C D C . . .
Prisoner 2 TFT/ random	C C C D C D C C *D D C D D* C D . . .

2.2 Tit-for-Two Tats / Random

Opponent makes the same choice twice in a row it is reciprocated and also skewed randomly

This strategy is like Tit-For-Tat except that opponent select the same choice must make the same choice twice in a row and choice is skewed by random setting. Initially start with the C and then D if the opponent prisoner defect twice and once again fall back to cooperating if the opponent does not continue to defect sometimes randomly change the setting.[14]

Prisoner1	C C D D C C D C C C D D C C D D C . . .
Prisoner2 using TF2T	C C C C D D C C C D D D C C C C D. . .

2.3 Generous Tit-for-Tat

Reducing the probability of C and D specified by the outcome of the previous interaction from 1 to <1.In case of Infinitely IPD the first move will be irrelevant since it has been forgotten in long run. Let say p(q) (where 0>p ,q<1) is the probability for the cooperation after the cooperation(defection) move of opponent .for the value [8]

$$q = \min\left(1 - \frac{(T-R)}{(R-S)} \; ; \; \frac{(R-P)}{(T-P)}\right) = \frac{1}{3}$$

is said to provide the highest payoff.

2.4 Naive Prober *(Tit-for-Tat with Random-Defection)*

Repeat the opponent last choice but sometime skewed by random defection.

This strategy is also entirely based on the history of the previous interaction but sometime perverse by random selection of defection against the cooperation of other prisoner.

2.5 Remorseful Prober *Tit-For-Tat with random defections*

This strategy is also based on the record of previous move of opponent. This strategy is entirely base on Tit-for- Tat (repeat the opponent last move), but a little bit difference is sometime probe by defecting while the opponent is still cooperating and if the opponent defect in response to the probing then remorse by cooperating once.

Per C D	C C C D D D C C C C C D….
Tit-for-Tat	C C C C D D D C C C D C ….

2.6 Naive Peace Maker (Tit-for-Tat with Random cooperation)

Repeat the opponent last choice but sometime create peace by random cooperation

This strategy is entirely based on the history of the previous interaction but sometime develop a peaceful invironment by random selection of cooperation (C) against the defection (D) of other prisoner.

2.7 True Peace maker *hybrid of Tit-for-Tat and Tit-for-Two Tats with Random Co-operation*

Maintain the record of the opponent previous move, start with the cooperate unless the opponent defect twice in a row then defect ,some time make peace by random cooperation while the opponent still doing the defect.

2.8 Forgiving Strategies

This strategy improves the performance of Tit-for-Tat and other strategy by recognizing non-nice and naive strategies. The

performance of (CD)* against Tit-for-Tat is shown below

Per C D	C D C D C D C D C D C D C D . . .
Tit-for-Tat	C C C D C D C D C D C D C D . . .

This strategy is similar to Tit-for-Tat with vaguely modification i.e. use periodic strategies and forgive when interactions are spiraling into ongoing defection [9]as shown below

Prisoner 1	C C D C D C D D D D D D D D. . .
Prisoner 2 forgiving	C C C D C D C D D D D C C C C .. .

Stochastic Strategies

There is vast number of difficulty facing by cooperation in selfishly exploitative environment .A difficulty may occur due to noisy environment, Certain action may be misinterpreted in random break down in perception and transmit ion or during the transition various error massage may also be generated . [8]

2.9 Pavlov *repeats the last choice if the output is good*

In the repeated iterated prisoner's dilemma, if the adaptive strategy is not clear then the requirement of leading strategy is generated. Pavlov is proposed as a leading strategy in IPD. [16]In the noisy environment, when the other strategy like Tit-for- Tat does not perform well then this strategy good to select. This strategy is considered as a complex strategy. In this strategy prisoner not only notify the opponent last move but also notify own move and from the last move if he got the better result than repeats last choice. [19]

There are four dimensional vector (P_1, P_2, P_3, P_4) shows the probability (where $0<P<1$) of playing C after R, S, T, P and after a single D of opponent.
Pavlov = (1, 0, 0, 1)
Shows cooperate after R and P and defect after S and T – in other words, stay with the previous decision after scoring the higher payoffs R and T and switch after S and P. Pavlov have two important features over TFT, it can correct occasional mistakes and prevents invasion of strict cooperators by exploiting them in contrast to TFT. Pavlov loses in compare to ALLD because it interchanges between C and D.[17]

3.0 Pavlov/Random *repeats the last choice if the output is good and random*

Prisoner notifies own move and the opponent last move and if got the better output from the previous move repeat the previous move but sometime randomly make the choice. As compare to Pavlov four dimensional vector (P_1, P_2, P_3, P_4) shows the probability (where $0<P<1$) of playing C after R, S, T, P and after a single D of opponent
Pavlov = (1, 0, 0, 1)
Shows cooperate after R and P and defect after S and T – in other words, stay with the previous decision after scoring the higher payoffs R and T and switch after S and P but sometime randomly cooperate after S and T and randomly defect after R and P.

3.1 Grudger *cooperates, but only be a sucker once*
Notify the opponent last move if the opponent cooperate than cooperate until the opponent defect and once opponent defect then always defect and never riposte to the cooperate ,never forgive the opponent for behave badly.

Prisoner 1	C C D C D C C C C D D C …
Prisoner 2 gurdger	C C C D D D D D D D D . . .

3.2 Soft Grudger *cooperates until the opponent defect*

Notify the opponent last move if the opponent cooperates than cooperate until the opponent defect, once opponent inaugurate defect than punish the opponent but during the punishment sometimes riposte to the cooperate.

Prisoner 1	C C D C D C C C C D D C …
Prisoner 2 gurdger	C C C D D D D D D C C C . . .

3.3 Gradual *defect the total no times the opponent defect*

Start with the cooperate and continue until the opponent defects, once opponent inaugurate defect than punish the opponent and defect the total number of times the opponent has defect during the interaction follow by two cooperation .

3.4 Adaptive *start with the cooperate and take best choice that give the best average score*

Always notify own move, start with the incessant cooperate followed by the incessant defect , calculate the average score and then select between C and D which have give the best average score .Calculate the average score after every move .

3.5 Customized Strategy 1 *always cooperate unless sucker*

First set the default value of R,S,T,P such as R=1,P=1,T=1,S=0 and then start to play the game with cooperate and calculate the average score after every move. Always Play cooperates unless sucker (i.e. obtain average score = 0 point)

3.6 Customized Strategy 2 *always play alternatively defect/cooperate*

Always notify own move,, First set the default value of R,S,T,P such as R=0,P=1,T=1,S=0 and then start to play the game with cooperate and alternatively set the value of C and D..

Conclusion

A brief survey of various strategies used to solve game theory famous IPD has been presented here these has been widely implemented in the proposed solutions by many authors. Applicability of soft computing techniques in various instances or similar instances of the problem under uncertainty is still under investigation

Reference

1.Raj Mathew and M. R. Kaimal ,"A fuzzy approach to prisoner's dilemma game using fuzzy expected value model, *2003* lEEE

2.Kevin W. Li, Fakhreddine Karray,"Fuzzy Approaches to the Game of Chicken",2001 IEEE

3. John von Neumann, "Game Theory and the Puzzle of the Bomb William Poundstone" ,Anchor Books, Doubleday, 1993.

4.Robert Axelrod 1980a."effective choice in prisoner's dilemma ,"Journal of conflict resolution 1981

5.K. Arrow and S. Honkapohja,"What Is Game Theory Trying to Accomplish? ",Basil Blackwell, Oxford, 1985.

6.Axelrod, R.," The Evolution of Cooperation", ISBN ,1984

7.Robert Aumann, "Acceptable points in general cooperative n-person games", in R. D. Luce and A. W. Tucker (eds.), Contributions to the Theory 23 of Games IV, Annals of Mathematics Study ,ISBN 1994

8.Björn BrembsChaos, "Chaos, cheating and cooperation: potential solutions to the Prisoner's Dilemma", ISSN 1996
9.Colm O' Riordan ,"A forgiving strategy for the Iterated Prisoner's Dilemm",Journal of Artificial Societies and Social Simulation vol. 3, no. 4, 2000
10.Martin Novak and Karl Sigmmund ,"chaos, evolution of corporation",proc. Natl acad. Sci,usa vol 90, june 1993
11. Robert Axelrod a."evolution of coorporation ,"new yark basic book 1980
12.Boyd, R. and Lorberbaum, J. P.. No pure strategy is evolutionary stable in the repeated prisoner's dilemma game. 1987
13. Pollock, G. and Dugatkin, L. A. . ,"Reciprocity and the emergence of reputation",. - J. theor. Biol. 1992
14.David M. Chess," Simulating the evolution of behavior: the iterated prisoners' dilemma problem. Complex Systems", 1988
15.Nowak, M. A. and Sigmund, K. ,"Tit for tat in heterogeneous populations. – Nature," 1992
16.Nowak, M. A. and Sigmund, K. (a). ,"Chaos and the evolution of cooperation,". - USA 1993
17.Nowak, M. A. and Sigmund, K.(b). ,"A strategy of win-stay lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game,". - Nature , 1981
18.Nowak, M. A. and Sigmund, K. "Tit for tat in heterogeneous populations. – Nature" 1992.
19. Axelrod, R. and Hamilton, W. D., "The evolution of cooperation",. – Science,1981
20.Hamilton, W. D.. The genetical evolution of social behavior. - J. theor. Biol.1964
21.Dugatkin, L. A. and Wilson, D. ," Rover: a strategy for exploiting cooperators in a patchy environment", 1991