UNLOCKING THE WORLD OF BIASED ITERATED PRISONER'S DILEMMA GAMES

Maximilian M. Nguyen

April 27, 2025

ABSTRACT

We generalize the family of 2 x 2 iterated prisoner's dilemma (IPD) games that are governed by a Markovian transition matrix by removing restrictions on the form of the matrix elements that are imposed historically by assuming unbiased transitions between game states. In doing so, we can more easily contextualize and categorize the variety of IPD games that allow for zero-determinant (ZD) strategies. This perspective generalizes the line of thought initially gleamed by Press and Dyson and opens up the possibility for other surprising strategies (like ZD strategies) to be discovered in the future in biased IPD games.

Keywords prisoner's dilemma, zero-determinant strategies, game theory, Markov transition matrix

1 Introduction

In a remarkable and innovative paper by Press and Dyson [1], the authors unexpectedly found and introduced the possibility of zero-determinant (ZD) strategies within the context of the repeated game known as the iterated prisoner's dilemma (IPD). It is natural to ask then, given the history of the famous IPD problem in both the scientific and the public eye [2–6], why did it take so long for these special strategies to be discovered?

In this Perspective, we can begin to deduce an answer to this question by revisiting the framework and assumptions underlying the IPD game. In doing so, we will notice an important assumption that is made regarding game transitions that are inherently unbiased versus biased. By analyzing this assumption, we will illuminate a broader family of IPD games, of which the Press-Dyson game is a single variety. In doing so, we hope to illuminate why ZD strategies can naturally arise based on the assumptions of the structure of the game, and hopefully pave the way for other interesting strategy sets to be discovered in the family of Markovian IPD games.

2 Results

We start by assuming a framework for a 2 x 2 IPD game with Markovian transitions. In this setup, two players A and B play a game where they simultaneously choose to either cooperate (c) or defect (d) with one another. If both players cooperate, both are given a reward payoff (R). If only one player defects, the defector is given a temptation payoff (T), while the cooperator is given the sucker payoff (S). If both players defect, they are both given the punishment payoff (P). For simplicity of notation but that will not effect the message of this paper, we will assume here that the value for each type of payoff is the same for both players. This 2 x 2 game takes on the interesting behavior of the prisoner's dilemma when assuming that $T > R > P > S$ and that $R > \frac{T+S}{2}$ [2]. The game is then repeated in discrete time steps indefinitely, resulting in an iterated prisoner's dilemma game. Typically, the IPD is then posed as a question of identifying strategies that would maximize a player's total payoff across all iterations of the game.

A critical component of the setup of IPD is deciding how the game progresses from one time step to the next. In the Markovian setup that we consider here, the next state of the game is only dependent on the current state of the game (and thus independent of the prior history of the game). Since the game can be in one of 4 possible states at each time step (cc, cd, dc, dd), where each duplet is the joint choices of player A and player B respectively, the dynamics of the game are given by a 4 x 4 Markov transition matrix with matrix elements $e_{i\to j}$, representing the probability to

		Next Game State (j)			
		cc	cd	dc	dd
Current Game State (i)	cc	$e_{CC\to CC}$	$e_{CC\to CD}$	$e_{CC\to DC}$	$e_{CC\to DD}$
	cd	$e_{CD\to CC}$	$e_{CD\to CD}$	$e_{CD\to DC}$	$e_{CD\to DD}$
	dc	$e_{DC\to CC}$	$e_{DC\to CD}$	$e_{DC\to DC}$	$e_{DC\to DD}$
	dd	$e_{DD\to CC}$	$e_{DD\to CD}$	$e_{DD\to DC}$	$e_{DD\to DD}$

Figure 1: Under the Markovian assumption, transitions between the current and next step of the IPD game are governed probabilistically by the elements of a transition matrix.

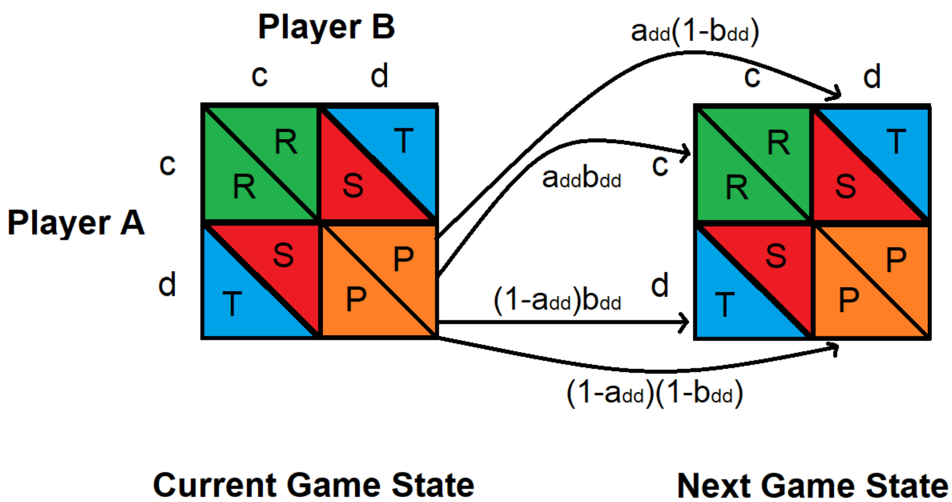

Figure 2: For an IPD game under an unbiased Markov transition matrix, the matrix elements are given by the product of player A's decision (a_i or $1 - a_i$) and player B's decision (b_i or $1 - b_i$), depending on the current and next game states. Shown here are the four possible transitions if the current state of the game is dd. Similarly, arrows can be drawn for transitions out of the other three game states.

transition from the current game state i to the next game state j (Figure 1). The two requirements of a Markov transition matrix are that its elements all lie on the unit interval (i.e. $0 \leq e_{i\to j} \leq 1$) and that the matrix is row stochastic (i.e. $\sum_j e_{i\to j} = 1$, or in words that each row of elements sum to 1), which ensures that all transitions can be described by a proper probability.

In the classical literature on IPD [1–4], an implicit assumption is that the next state of the game is solely determined by the joint choices of both players from the current state of the game. Let the current game state be denoted by i, where $i \in \{cc, cd, dc, dd\}$. Let a_i be the probability for player A to cooperate in the next time step given that the current state of the game is i, and similarly let b_i be the probability for player B to cooperate in the next time step given that the current state of the game is i. Then under the assumption that the transition element corresponds to the joint probability of each player making their decision independently, we see that each element of the Markov transition matrix is given by a product of either a_i or $1 - a_i$ with either b_i or $1 - b_i$, depending on the current and next states of the game (Figure 2).

In this form of IPD, each element of the transition matrix is a function of the joint probabilities of both players' choices. Phrased another way, no future state of the game is independent of either players' strategy (player A's strategy is given by the set $\{a_{cc}, 1 - a_{cc}, a_{cd}, 1 - a_{cd}, a_{dc}, 1 - a_{dc}, a_{dd}, 1 - a_{dd}\}$, while player B's strategy is given by the set $\{b_{cc}, 1 - b_{cc}, b_{cd}, 1 - b_{cd}, b_{dc}, 1 - b_{dc}, b_{dd}, 1 - b_{dd}\}$). Let us call the IPD game governed by this form of a transition matrix the $unbiased$ IPD game. We have chosen the terminology $unbiased$ to reflect that this form of the IPD game is the most fair, in the sense that transitions to every state jointly depend on both players' strategies and that transitions are not influenced by any factor outside of this joint probability. The situation is analogous to throwing $loaded$ versus $fair$ dice, and one could use those terms instead if preferred. Press and Dyson [1] showed that ZD strategies exist when the difference between the transition matrix and the identity matrix can be made unilaterally singular by a single player. Thus the central object to evaluate when determining if ZD strategies exist for a type of IPD game is the matrix obtained

		Next Game State			
		cc	**cd**	**dc**	**dd**
Current Game State	**cc**	$a_{cc}b_{cc}-1$	$a_{cc}(1-b_{cc})$	$(1-a_{cc})b_{cc}$	$(1-a_{cc})(1-b_{cc})$
	cd	$a_{cd}b_{cd}$	$a_{cd}(1-b_{cd})-1$	$(1-a_{cd})b_{cd}$	$(1-a_{cd})(1-b_{cd})$
	dc	$a_{dc}b_{dc}$	$a_{dc}(1-b_{dc})$	$(1-a_{dc})b_{dc}-1$	$(1-a_{dc})(1-b_{dc})$
	dd	$a_{dd}b_{dd}$	$a_{dd}(1-b_{dd})$	$(1-a_{dd})b_{dd}$	$(1-a_{dd})(1-b_{dd})-1$

Unbiased IPD Game

Figure 3: The difference between the Markov transition matrix for an unbiased IPD game and the identity matrix.

		Next Game State			
		cc	**cd**	**dc**	**dd**
Current Game State	**cc**	$a_{cc}b_{cc}-1$	a_{cc}	b_{cc}	$(1-a_{cc})(1-b_{cc})$
	cd	$a_{cd}b_{cd}$	$a_{cd}-1$	b_{cd}	$(1-a_{cd})(1-b_{cd})$
	dc	$a_{dc}b_{dc}$	a_{dc}	$b_{dc}-1$	$(1-a_{dc})(1-b_{dc})$
	dd	$a_{dd}b_{dd}$	a_{dd}	b_{dd}	$(1-a_{dd})(1-b_{dd})-1$

IPD Game
Where Both Players Can Access ZD Strategies

Figure 4: An IPD game where both players can access ZD strategies. The transition matrix reflects the game analyzed originally by Press and Dyson [1]. The transition matrix is obtained by starting with the unbiased transition matrix, then adding column 1 to both column 2 and 3, and then subtracting column 1 twice from column 4. The matrix shown is the difference between the resulting transition matrix and the identity matrix.

by taking the difference between the Markov transition matrix and the identity matrix (Figure 3). Since both players are agnostic of the other player's strategy in a prisoner's dilemma, inspection of this matrix shows that neither player can guarantee the matrix will become singular with their choice of strategy set as no column of elements can be forced to zero unilaterally. Thus, no guaranteed ZD strategies exist for this unbiased game.

However, let us now peel back this assumption of an $unbiased$ game. While the unbiased case is a natural setup for IPD, it is a very special case of IPD games governed by a Markov transition matrix. Given that the elements of this Markov transition matrix can be drawn from the unit interval, one can formulate infinitely many other forms of IPD that obey a Markov transition matrix in principle. So long as the the elements lie on the unit interval and the elements of each row sum to unity, any combination of row elements result in a mathematically valid IPD game. Thus, the matrix elements need not only be a function of both players' previous choices (as in the unbiased IPD game). In principle, the elements could be a function of only a single player's previous choice, or even a function of neither players' choices (i.e. given by a fixed constant).

In contrast to the unbiased IPD game, an IPD game on the other end of the bias spectrum would be a game that deterministically transitions from one state to another state (i.e. each row has one element with value 1 and the other elements are 0). In this case, players no longer have control over their choice of action. Obviously, this extreme example is a less interesting case to analyze from a social behavioral standpoint as the players do not have decision autonomy.

In between these two extremes of bias lie an infinite spectra of IPD game possibilities. To illustrate, let us take the alternative IPD game originally posed by Press and Dyson as a starting example. To obtain their transition matrix, they started with the unbiased transition matrix and subsequently added column 1 to both column 2 and 3. Furthermore, to ensure the transition matrix remains row stochastic matrix, one needs to conserve probability, so column 1 needs to be subtracted twice from column 4. The difference between the transition matrix for this IPD game and the identity matrix is given in Figure 4. While these simple column operations of adding and subtracting columns are shears on the column space that preserve the matrix determinant, the magnitude of the leading eigenvalue at 1, and the associated vector of all ones for the right eigenvector, in general, they do not preserve the associated left eigenvector (which gives the stationary distribution of the matrix). Thus, the stationary distribution in this new biased IPD game will generally be different than the stationary distribution associated with the unbiased game.

The existence of ZD strategies follows from Press and Dyson's results once a transition matrix is prescribed with at least one column under the control of a single player, as that player can unilaterally force the determinant of the

		Next Game State			
		cc	**cd**	**dc**	**dd**
Current Game State	**cc**	$b_{cc}-1$	$a_{cc}-b_{cc}$	$(1-a_{cc})b_{cc}$	$(1-a_{cc})(1-b_{cc})$
	cd	b_{cd}	$a_{cd}-b_{cd}-1$	$(1-a_{cd})b_{cd}$	$(1-a_{cd})(1-b_{cd})$
	dc	b_{dc}	$a_{dc}-b_{dc}$	$(1-a_{dc})b_{dc}-1$	$(1-a_{dc})(1-b_{dc})$
	dd	b_{dd}	$a_{dd}-b_{dd}$	$(1-a_{dd})b_{dd}$	$(1-a_{dd})(1-b_{dd})-1$

IPD Game
Where Only 1 Player Can Access ZD Strategies

Figure 5: An IPD game where only a single player can access ZD strategies. The transition matrix is made through column manipulation of the unbiased game. The matrix is obtained by starting with the unbiased transition matrix, then adding column 3 to both column 1, and then subtracting column 3 twice from column 2. The matrix shown is the difference between the resulting transition matrix and the identity matrix. In this game, a zero determinant strategy can only be implemented by player B.

		Next Game State			
		cc	**cd**	**dc**	**dd**
Current Game State	**cc**	$(a_{cc}b_{cc})^2-1$	$(a_{cc}(1-b_{cc}))^2$	$((1-a_{cc})b_{cc})^2$	$1-a_{cc}^2-b_{cc}^2+2a_{cc}^2b_{cc}+2a_{cc}b_{cc}^2-3a_{cc}^2b_{cc}^2$
	cd	$(a_{cd}b_{cd})^2$	$(a_{cd}(1-b_{cd}))^2-1$	$((1-a_{cd})b_{cd})^2$	$1-a_{cd}^2-b_{cd}^2+2a_{cd}^2b_{cd}+2a_{cd}b_{cd}^2-3a_{cd}^2b_{cd}^2$
	dc	$(a_{dc}b_{dc})^2$	$(a_{dc}(1-b_{dc}))^2$	$((1-a_{dc})b_{dc})^2-1$	$1-a_{dc}^2-b_{dc}^2+2a_{dc}^2b_{dc}+2a_{dc}b_{dc}^2-3a_{dc}^2b_{dc}^2$
	dd	$(a_{dd}b_{dd})^2$	$(a_{dd}(1-b_{dd}))^2$	$((1-a_{dd})b_{dd})^2$	$1-a_{dd}^2-b_{dd}^2+2a_{dd}^2b_{dd}+2a_{dd}b_{dd}^2-3a_{dd}^2b_{dd}^2-1$

IPD Game With Modified Transition Rates and No ZD Strategies

Figure 6: An IPD game where the probability of transitioning to one state is enhanced, while transitioning to the other three states is diminished compared to the unbiased game. This transition matrix is made through squaring each of respective element of the first three columns of the unbiased game. The last column is simply the the difference between unity and the sum of the first three column to conserve probability. The matrix shown is the difference between the resulting transition matrix and the identity matrix. In this particular IPD game shown, the probability of transitioning to the dd state is greater than in the unbiased game.

transition minus identity matrix to zero (and by extension influence dependent quantities such as expected payoffs) by picking appropriate strategy values [1]. However, the game posed in the original Press and Dyson paper is only one way to obtain such a condition. There are many different combinations of column operations that will result in a future state being under the control of a single player. Consider for instance, a Markov transition matrix obtained by starting with the unbiased transition matrix, then adding column 3 to column 1, and then subtracting column 3 from column 2. The resulting IPD game puts the cc state under the sole control of player B (Figure 5). In contrast to the Press-Dyson game, where both players were able to force a zero determinant, in this game, only player B has such power. Various other column manipulations can create other IPD games with different states under the control of different players.

One also need not restrict the transition matrix to be simple column manipulations of the unbiased IPD game. Consider for instance, a game where the elements of the first three columns are taken from the first three columns of the unbiased transition matrix and then squared. After adjusting the final column to conserve probability, the resulting IPD game would be governed by the transition matrix shown in Figure 6. Assuming the elements of the first three columns of the unbiased transition matrix are neither 0 nor 1, then squaring them produces values that are strictly less than their original values in the unbiased game. As a result, the dd column in this alternative IPD game will be strictly larger than the unbiased IPD game. Consequently, one would expect the optimal strategy to be a new type of strategy that factors this bias in. In this game, since each column is still a function of both players' strategies, ZD strategies cannot exist for this game.

Discussion

As we have seen, by relaxing the assumption on the form of the transition matrix, we now have access to a much larger (in fact, infinitely large) family of IPD games.

It is also easy to see now under which variety of IPD games that ZD strategies will exist. As originally hinted by the analysis of Press and Dyson [1], when there are matrices where there is at least one column that is a function of only one player's strategy for both players, then both players can access ZD strategies. When the matrix is such that only one player that has a column that is solely a function of their strategy, then only that player can access ZD strategies. And in the unbiased case and the cases where all matrix elements are a function of either both players' strategies or neither players' strategies, then guaranteed ZD strategies will not exist.

Finally, this brings us to the question of why ZD strategies were not discovered sooner. Given the naturalness, fairness, and simplicity of posing an IPD in its unbiased formulation, it makes sense that people might default to the unbiased form. However, since Press and Dyson, the literature on ZD strategies seems to have conflated ZD strategies with the unbiased IPD transition matrix. The origin of this conflation may stem from when Press and Dyson performed the column operations on the unbiased game in their original paper. While the column operations did preserve the determinant of that transition matrix, Press and Dyson did not mention that the matrix transformation alters the stationary distribution of the matrix which creates a fundamentally different IPD game. This explains then why people did not discover ZD strategies sooner, as unknowingly Press and Dyson were the first to analyze a biased IPD game. Their pioneering work revealed the possibility of ZD strategies just from running into the tip of the iceberg in terms of the possible biased IPD games that exist. And through standing on their shoulders, we can now see much further into the vast horizons of the landscape of IPD games.

In terms of realism, it is also not unreasonable to believe that scenarios in real life can introduce various biases that might distort the transitions from the unbiased case. The generalization of the transition matrix forces us to consider the IPD game as being composed of two equally important pieces: the prisoner's dilemma game being played at each step, and also the process for iterating between time steps. This latter piece seems to have been largely simplified by taking the unbiased assumption in older literature and its relative importance was ignored until the work of Press and Dyson. Here we emphasize that the transitions are not only a function of the action probabilities given by each players' strategies, but they are also a function of any form or rules imposed on the transition matrix itself.

In future work, it would be interesting to see what new types of strategies emerge for transition matrices that have not yet been studied. ZD strategies emerged as a result of a specific variety of games in this larger family of IPD games. For more obscure forms of the IPD game with more intricate behavior and transition dynamics, it is not unreasonable to assume the optimal strategies in those cases will also be interesting and unintuitive. While here we have explicitly looked at IPD games of the 2 x 2 variety, the lesson of generalizing the transition matrix extends to arbitrary m x n games. We hope the generalization here will inspire future work to investigate the vast family of IPD games that are out there.

3 Acknowledgments

I was very fortunate to meet Freeman Dyson at the IAS in the fall of 2019, a couple of months before he passed away. Although I did not know of his involvement in this topic at the time, I am sure it would have been fun to talk to him about it. I will remember him for his boldness in the pursuit of scientific inquiry and his playful spirit. I hope this perspective encourages others to continually re-evaluate what we think the truth to be.

References

1. Press, W. H. & Dyson, F. J. Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences* **109.** Publisher: Proceedings of the National Academy of Sciences, 10409–10413. https://www.pnas.org/doi/10.1073/pnas.1206569109 (2024) (June 26, 2012).
2. Axelrod, R. & Hamilton, W. D. The Evolution of Cooperation. *Science* **211.** Publisher: American Association for the Advancement of Science, 1390–1396. https://www.science.org/doi/abs/10.1126/science.7466396 (2024) (Mar. 27, 1981).
3. Axelrod, R. & Dion, D. The Further Evolution of Cooperation. *Science* **242.** Publisher: American Association for the Advancement of Science, 1385–1390. https://www.science.org/doi/abs/10.1126/science.242.4884.1385 (2024) (Dec. 9, 1988).
4. Nowak, M. & Sigmund, K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* **364.** Number: 6432 Publisher: Nature Publishing Group, 56–58. ISSN: 1476-4687. https://www.nature.com/articles/364056a0 (2024) (July 1993).
5. Nowak, M. A. Five Rules for the Evolution of Cooperation. *Science* **314.** Publisher: American Association for the Advancement of Science, 1560–1563. https://www.science.org/doi/full/10.1126/science.1133755 (2024) (Dec. 8, 2006).

6. Kendall, G., Yao, X. & Chong, S. Y. *The Iterated Prisoners' Dilemma: 20 Years on* Google-Books-ID: A_VoDQAAQBAJ. 273 pp. ISBN: 978-981-270-697-3 (World Scientific, 2007).