Cooperation and Coordination in the Turn-Taking Dilemma

Daniel B. Neill
Carnegie Mellon University
Department of Computer Science
5000 Forbes Avenue
Pittsburgh, PA 15213
neill@cs.cmu.edu

Abstract

In many real-world situations, "cooperation" in the simple sense of the Prisoner's Dilemma is not sufficient for success: instead, cooperators must precisely coordinate more complex behaviors in a noisy environment. We investigate one such model, the *Turn-Taking Dilemma*, a variant of the repeated Prisoner's Dilemma in which the highest total payoff is achieved not by simultaneous mutual cooperation, but by taking turns defecting (alternating temptation and sucker payoffs). The Turn-Taking Dilemma more accurately models interactions where players must take short-term losses for long-term gains: situations marked by the intricate give-and-take of bargaining and compromise. Using "evolutionary dominance" as a general measure of performance, we investigated which strategies are most successful in Turn-Taking Dilemma interactions. Our experiments demonstrate that turn-taking can be achieved in a noisy environment, even when agents have strict resource constraints (limited memory strategies). Top strategies such as $EXALT_2$ can effectively coordinate turn-taking under noise, while exploiting cooperators and resisting exploitation by defectors; these strategies are likely to achieve success in the variety of real-world interactions modeled by the Turn-Taking Dilemma.

1 Two models of cooperation

The Prisoner's Dilemma, in its various forms, is widely used to model the evolution of cooperation in interactions between individuals with partially conflicting goals. In particular, two variants of the problem have captured the attention of the research community. In the *Iterated Prisoner's Dilemma* (IPD), players simultaneously choose between *cooperation* for mutual benefit and *defection* for individual benefit, and this decision situation is repeated over a number of rounds (Axelrod and Hamilton, 1981). In the *Alternating Prisoner's Dilemma* (APD), players alternate turns, and on each player's turn he must choose whether to cooperate or defect (Nowak and Sigmund, 1994; Frean, 1994). For both of these models, the potential outcomes are given in the following payoff table:

		P2 C	P2 D
P1	C	R / R	S / T
	D	T / S	P / P

Table 1: Prisoner's Dilemma Payoffs to P1 / P2

If both players cooperate, each receives a high payoff R as a *reward* for mutual cooperation. If both defect, each receives a low payoff P as a *punishment* for mutual defection. If one player defects while the other cooperates, the defector receives a very high payoff T as a *temptation* to defect, and the cooperator receives a very low *sucker* payoff S. Despite the fact that mutual cooperation is preferred to mutual defection, each player scores higher if he defects regardless of the opponent's choice, and hence mutual defection is the only "rational" outcome of the single-shot Prisoner's Dilemma game.

In both repeated Prisoner's Dilemma models, however, reciprocal altruism can develop between two self-interested players, enabling the establishment of mutual cooperation to the benefit of both. These models have been applied to fields as diverse as biology, economics, and politics, and have sparked theories of the evolution of cooperation based on reciprocity (Maynard Smith, 1982; Axelrod, 1984; Nowak and Sigmund, 1994). However, it has been shown that the IPD and APD are distinct problems, and decision-making strategies which are successful in one variant may perform poorly in the other (Frean, 1994). This creates a significant dilemma for the researcher: which model should be used to describe a given real-world situation?

The clearest example of an Iterated Prisoner's Dilemma is the predator inspection behavior of the stickleback fish (Milinski, 1987; Nowak and Sigmund, 1994). Both fish simultaneously swim up to the predator in a series of short, halting moves; in this example, cooperation translates to swimming forward, and defection translates to lagging behind. The optimal result is that of continued mutual cooperation: both fish swim as close as possible to the predator while equally sharing the potential risk.

The clearest example of an Alternating Prisoner's Dilemma is the feeding behavior of the South American vampire bat (Wilkinson, 1984). Bats who have found a good meal will help hungry bats by donating some of their extra food. Only bats who have a surplus of food can make a decision whether to cooperate or defect on any given night; in this case, cooperation translates to donating to a hungry bat, and defection translates to hoarding the surplus. The optimal result is that of alternating reciprocal altruism: the sated bat donates to its hungry companions, and the other bats reciprocate during future feedings.

Other real-world situations seem to have elements of both IPD and APD, leading to debates over which model is more appropriate. The guarding behavior of the dwarf mongoose (Rasa, 1989) is one such situation. The subadult males of the group take turns watching for predators, increasing their own risk in order to protect the group. As discussed in Neill (2001), the guarding mongoose must choose between vigilant guard behavior (cooperation) or focusing on its own safety (defection), and thus the decision can be modeled as an Alternating Prisoner's Dilemma. But this example leads to some interesting questions: how does the group decide which mongoose takes a given watch? Do all the mongooses share the job equally, or do some take more watches than others? It is reasonable to model the mongooses' choices whether or not to guard as simultaneous choices between cooperation and defection, which suggests the Iterated Prisoner's Dilemma as a better model. However, the group does not need multiple males to take the same guard position simultaneously, so simultaneous mutual cooperation is not the optimal outcome. A better outcome would be one in which the subadult males *take turns* on guard duty; while one male is guarding the others are safe within the group. And this is exactly how they behave. (In actuality, this is an N-person Turn-Taking Dilemma: as is common in the Prisoner's Dilemma literature, we focus on the simplest case, $N = 2$.)

A careful examination of these three examples reveals that they differ in two main criteria: the simultaneity of the players' moves, and the desired form of cooperation. In the APD, only one player is making a choice at a given time, while in the IPD players must choose their actions simultaneously. Similarly, simultaneous mutual cooperation is the goal in the IPD, while in the APD "it makes no sense to cooperate simultaneously; the partners have to take turns" (Nowak and Sigmund, 1994). The stickleback demonstrates both simultaneous choices and the goal of simultaneous cooperation, and thus is a clear example of the IPD. The vampire bat demonstrates both alternating choices and the goal of alternating reciprocal altruism (turn-taking), and thus is a clear example of the APD. The mongoose demonstrates the combination of simultaneous choices and turn-taking, and thus seems to fit both IPD and APD models. We shall show, however, that it is substantially different from either model. To do so, we propose a third model which more accurately represents this decision situation.

2 The Turn-Taking Dilemma

An Iterated Prisoner's Dilemma game is mathematically defined by two relationships between its temptation, reward, punishment, and sucker payoffs. First is the obvious ordering of payoffs: $T > R > P > S$. Additionally, we require $2R > T + S$, so cooperation achieves more points than alternating T and S payoffs.

If the second inequality is reversed, we have a substantially different situation: mutual cooperation is no longer the Pareto optimal result in the repeated game. Both players do better if they take turns; in this interaction one player defects and the other cooperates, then they switch roles on the next round. We define a *Turn-Taking Dilemma* (TTD) as one in which these two conditions hold: $T > R > P > S$, and $T + S > 2R$.

The Turn-Taking Dilemma can be thought of as representing two types of situations. In the first type, there is a benefit available, but at most one of the two players can have it. Defection corresponds to “taking a turn,” and cooperation corresponds to “not taking a turn,” if the result of taking the turn is a benefit. In this case, both players do better if they resist the benefit than if they give in to its temptation, but the optimal solution is for one to collect the benefit and one to wait his turn. One example of this is found in economics, when two competing companies must decide when to release a “new and improved” product or service. During each time period, each company may choose to release a new version (defection) or wait (cooperation). We assume that there is some cost X associated with releasing and marketing the new version. If one company releases a new version and the other does not, we assume that the new product captures the entire market: the releasing company makes a huge profit Y while the other company makes no profit. If neither company’s product is “new,” or if both companies’ products are “new,” they each obtain a partial market share with a much smaller associated profit Z. Thus the payoff table is the following:

		P2 C	P2 D
P1	C	Z / Z	0 / $Y-X$
	D	$Y-X$ / 0	$Z-X$ / $Z-X$

Table 2: Profits for companies P1 / P2 in TTD situation

Assuming that the release cost $X = 2$, the profit for total market share $Y = 9$, and the profit for partial market share $Z = 3$, we obtain a payoff table with $T = 7$, $R = 3$, $P = 1$, and $S = 0$. These are the payoffs we consider to be standard for the TTD problem, and we use this payoff table throughout our experiments. Note that $2R = 6$ and $T + S = 7$, so alternating product releases (alternating defections) is more advantageous to both companies than either never releasing products (mutual cooperation) or always releasing products (mutual defection). Like the Prisoner’s Dilemma, though, there is the opportunity for exploitation: if a company is too nice, its competitor will usurp all of the “turns” (continually dominating the market) rather than waiting for its turn to release a product.

In the second type of Turn-Taking Dilemma, there is an impending harm, and at least one of the two players must face it. Cooperation corresponds to “taking a turn,” and defection corresponds to “not taking a turn,” if the result of taking the turn is a harm. If both players refuse to face the danger, both will be harmed; if both face the danger they will suffer less, but the optimal solution is for one to face the harm and one to remain safe. The guarding behavior of the dwarf mongoose is described well by this model: taking a turn carries a significant risk of being eaten by predators (S payoff), though this risk is reduced somewhat if both stand guard (R payoff). If the other mongoose stands guard, the mongoose in the tribe is very safe (T payoff), but if both shirk their duty the entire tribe is at great risk (P payoff).

An interesting combination of these two types can occur in the phenomenon of “gift-giving.” In this interaction, each player has two choices: to give a gift (cooperation) or not to give a gift (defection), and we assume that gifts are more valuable to the recipient than their cost to the donor. For example, if the gift costs the donor 4 points and gains the recipient 8 points, we have the payoff table ($T = 8$, $R = 4$, $P = 0$, $S = -4$). Since $2R > T + S$, this is clearly a Prisoner’s Dilemma; repeated gift-giving may be modeled as either an IPD (if the gift-giving is simultaneous, ex. Christmas), or an APD (if the gift-giving alternates, ex. birthdays). But what if the simultaneous giving of gifts by the two players diminishes the value of both gifts? A striking example of this is found in O. Henry’s short story, “The Gift of the Magi.” This is a fable of “two foolish children in a flat who most unwisely sacrificed for each other the greatest treasures of their house” (Henry, 1905): Della sells her hair to buy a fob for Jim’s watch, while Jim sells his watch to buy combs for Della’s hair. While in this case the value of the gifts is hugely reduced, even less dramatic examples will transform gift-giving from a Prisoner’s Dilemma to a Turn-Taking Dilemma: duplicate gifts that must be returned, or incompatible gifts (such as tickets to two events occurring at the same time). Assuming in the above example that duplicate gifts cause a 3 point hassle, we have the payoff table ($T = 8$, $R = 1$, $P = 0$, $S = -4$). Since $T + S > 2R$, this is a Turn-Taking Dilemma.

The Turn-Taking Dilemma requires us to redefine our intuitive notions of reciprocal altruism and exploitation formed from study of the Prisoner’s Dilemma. In the IPD and APD, reciprocal altruism is continued cooperation (whether simultaneous or alternating), and any defection represents exploitation of the other player for one’s own benefit. In the TTD, each player defects half the time in a reciprocal relationship; exploitation occurs when a player attempts to take more than his fair share of turns (if the turn is beneficial) or less than his fair share (if the turn is harmful).

Thus the TTD differs from the IPD and APD in two important respects. First, while in the Prisoner's Dilemma two cooperating players receive high payoffs every turn, in the Turn-Taking Dilemma the achievement of mutual help requires each player to frequently take very low payoffs. Hence the Turn-Taking Dilemma more accurately models the variety of situations in which one must take short term losses in order to maximize their long term gain. This game of give and take, the willingness to receive suboptimal payoffs in exchange for future reciprocation, is the basis for all compromise.

Second, since moves are made simultaneously, coordination of turns becomes an important issue. Even if two players strongly desire to take turns, they must still coordinate their turn-taking in order to avoid situations where both, or neither, take the turn. In the presence of noise (a non-zero probability that a player will defect when he meant to cooperate or vice versa), this coordination may be extremely difficult to achieve. In some real-world situations, coordination can be achieved by communication between the players: they may decide in advance who is responsible for which turns, or change their minds about their action when they discover that both, or neither, are taking a turn. Communication may destroy the simultaneous nature of the game: in these cases the situation is better modeled by an Alternating Prisoner's Dilemma. In other cases, choices must be made too quickly for a joint decision to be made. This may be true in the mongoose guarding example, and is certainly true in related situations: for example, if the tribe is ambushed by a predator too large to fight, and one mongoose must distract it while the tribe flees to safety. Alternatively, aspects of the environment may prevent communication: in the example of the product releases by competing companies, sharing information may be prohibited by antitrust laws. The Turn-Taking Dilemma considers issues of both cooperation and coordination, so it is a better model of these decision situations.

3 Optimality criteria for the Turn-Taking Dilemma

A strategy for the Turn-Taking Dilemma is a method of deciding whether to cooperate or defect on any given round. As in the Prisoner's Dilemma, the two simplest strategies are *ALLC* (the strategy which always cooperates), and *ALLD* (the strategy which always defects), but an infinite variety of other strategies are possible. The remainder of this paper attempts to answer the question of what strategies are most successful in Turn-Taking Dilemma interactions. We first discuss several techniques for evaluating the performance of strategies, and then use these techniques to explore the space of "memory one" and "memory two" strategies. In doing so, we hope to find what properties make a strategy successful. The question of strategy for the Turn-Taking Dilemma requires us to consider questions of both cooperation and coordination: what proportion of the turns do we take, and how do we coordinate this turn-taking in the presence of noise? To be successful in a Turn-Taking Dilemma, a strategy should be able to coordinate turn-taking with a wide variety of strategies. Additionally, it should resist exploitation, preventing other strategies from receiving more than their fair share of turns, as well as taking advantage of strategies who are exploitable in this way. Following Neill (2001), we define three optimality criteria for the Turn-Taking Dilemma, letting $w(X|Y)$ represent the expected value of the average payoff to strategy X against strategy Y in an infinitely long game. We make the assumption of "infinitesimal noise": mistakes can occur, but the probability of error is close to zero.

We first define a "self-alternating" strategy in terms of its self payoff $w(X|X)$. To be self-alternating in the TTD, a strategy should receive the temptation payoff half the time, and the sucker payoff half the time, in an infinitely long game against its clone:

Definition 1 *A strategy X is* self-alternating *if* $w(X|X) = \frac{T+S}{2}$.

More generally, we define a strategy's "self-performance" $\sigma_S = \frac{w(X|X)-P}{\frac{T+S}{2}-P}$. Thus $\sigma_S = 1$ for strategies which achieve self-alternation, $\sigma_S = 0$ for strategies which defect continually against their clone, and $0 < \sigma_S < 1$ otherwise.

Next we define a "C-exploiting" strategy in terms of its payoff $w(X|ALLC)$. To be C-exploiting in the TTD, a strategy should defect continually against *ALLC*:

Definition 2 *A strategy X is* C-exploiting *if* $w(X|ALLC) = T$.

More generally, we define a strategy's "C-performance" $\sigma_C = \frac{w(X|ALLC)-R}{T-R}$. Thus $\sigma_C = 1$ for strategies which defect continually against *ALLC*, $\sigma_C = 0$ for strategies which cooperate continually against *ALLC*, and $0 < \sigma_C < 1$ otherwise.

Next we define a "D-unexploitable" strategy in terms of its payoff $w(X|ALLD)$. To be D-unexploitable in the TTD, a strategy should defect continually against *ALLD*:

Definition 3 *A strategy X is* D-unexploitable *if* $w(X|ALLD) = P$.

More generally, we define a strategy's "D-performance" $\sigma_D = \frac{w(X|ALLD)-S}{P-S}$. Thus $\sigma_D = 1$ for strategies which defect continually against *ALLD*, $\sigma_D = 0$ for strategies which cooperate continually against *ALLD*, and $0 < \sigma_D < 1$ otherwise.

Strategies may be said to "totally" or "partially" meet each of these three criteria. For example, consider a strategy with $w(X|X) = 3.5$, $w(X|ALLC) = 5$, and $w(X|ALLD) = 0.75$, assuming the standard (7,3,1,0) payoff table. We calculate $\sigma_S = 1$, $\sigma_C = .5$, and $\sigma_D = .75$. This strategy is said to be totally self-alternating, 50% C-exploiting, and 75% D-unexploitable.

It is very likely that strategies with high values of σ_S, σ_C, and σ_D are able to coordinate turn-taking with a large number of other strategies, while exploiting overly cooperative strategies and resisting exploitation by others. Thus these strategies are likely to be extremely successful in a wide variety of Turn-Taking Dilemma interactions.

4 Performance measures for the Turn-Taking Dilemma

We will use the criteria of "self-alternation," "D-unexploitability," and "C-exploitation" as guides and illustrations in our exploration of the most successful TTD strategies. But how do we determine which strategies are the "most successful"? To do so, we require a general measure of the performance of a strategy. Following Neill (2001), and Kraines and Kraines (2001), we focus on two such measures: the round-robin tournament and the dominance score. In a round-robin tournament, the performance of a strategy X is measured by its average score $w(X|Y)$ against all strategies Y in the tournament. This method has been used commonly as a measure of Prisoner's Dilemma performance since the computerized tournament conducted by Robert Axelrod (Axelrod and Hamilton, 1981). Newer measures have focused on the "evolutionary" performance of a strategy: how does the strategy perform in a changing environment where high-performing strategies survive and reproduce, while low-performing strategies die off? Neill (2001) provides an overview of various evolutionary Prisoner's Dilemma models, and argues that three properties are essential for an evolutionary model. First, the model must consider not only a strategy's resistance to invasion ("evolutionary stability") but also its ability to invade other strategies ("evolutionary potency"). Second, it must be able to evaluate a strategy's performance against a large space of strategies, not only its evolutionary kin. Third, it must define a time-invariant measure of performance, depending only on the set of strategies being considered.

Neill (2001) discusses the various evolutionary models in detail, and concludes that "none of the commonly used models of the evolutionary APD lend themselves easily to a general measure of evolutionary performance." However, he proposes a new measure, the "dominance criterion," based on the invasion criteria of Maynard Smith (1982). From Maynard Smith, strategy X invades strategy Y if $w(X|Y) > w(Y|Y)$, or $w(X|Y) = w(Y|Y)$ and $w(X|X) > w(Y|X)$. If strategy X invades strategy Y, we write $X > Y$. Otherwise, we write $X \not> Y$. The independence of $X > Y$ and $Y > X$ leads to four possibilities:

I. $X > Y$ and $Y \not> X$. In this case, any initial proportion of strategy X can take over, and completely wipe out, strategy Y. We say that X *dominates* Y, and write $X \gg Y$.

II. $X \not> Y$ and $Y > X$. In this case, any initial proportion of strategy Y can take over, and completely wipe out, strategy X. We say that Y *dominates* X, and write $Y \gg X$.

III. $X > Y$ and $Y > X$. In this case, no matter what the initial proportions of strategies X and Y, the two strategies reach a stable equilibrium where the proportion of strategy X is given by:

$$p = \frac{w(X|Y) - w(Y|Y)}{w(X|Y) - w(Y|Y) + w(Y|X) - w(X|X)}$$

In this case, we write $X \overset{p}{\leftrightarrow} Y$.

IV. $X \not> Y$ and $Y \not> X$. This is a bistable equilibrium, in which either strategy X or strategy Y will take over the population depending on initial proportions. X will take over if its initial proportion is higher than:

$$m = \frac{w(Y|Y) - w(X|Y)}{w(X|X) - w(Y|X) + w(Y|Y) - w(X|Y)}$$

In this case, we write $X \overset{m}{\Leftrightarrow} Y$.

These definitions allow the definition of the dominance score $\mathrm{dom}(X|Y)$, which is a measure of the relative evolutionary performance of strategies X and Y (Neill, 2001).

$$\mathrm{dom}(X|Y) = \begin{cases} 1 & \text{if } X \gg Y \\ 0 & \text{if } Y \gg X \\ p & \text{if } X \overset{p}{\leftrightarrow} Y \\ 1-m & \text{if } X \overset{m}{\Leftrightarrow} Y \end{cases}$$

Thus the dominance score of X against Y is 1 if X dominates Y, 0 if Y dominates X, and between 0 and 1 otherwise. To use the dominance score, we can conduct a round-robin tournament in a format similar to Axelrod's (1984), but using dominance scores $\mathrm{dom}(X|Y)$ rather than payoffs $w(X|Y)$ for each pair of strategies. Alternatively, we can compute each strategy's dominance score against all strategies in a given space. Successful strategies will have very high dominance scores against most other strategies, and thus we can use a strategy's average dominance score as a general measure of its evolutionary performance. Our main focus in this paper will be to discover strategies that are optimal with respect to the dominance measure, and to explain what characteristics make these strategies highly successful in Turn-Taking Dilemma interactions.

5 Memory one strategies

We begin our search by examining the space of "memory one" strategies. A memory one strategy makes its decision whether to cooperate or defect based on its payoff in the previous round. The strategy is defined by a 4-tuple (p_R p_S p_T p_P) representing its probabilities of cooperation after the R, S, T, and P payoffs respectively. Most of the strategies commonly discussed in the Prisoner's Dilemma literature are memory one strategies (in the Alternating Prisoner's Dilemma, these translate to "two ply" strategies, where one "ply" is a single player's move).

These strategies include *ALLD* (0 0 0 0), the strategy which always defects; *ALLC* (1 1 1 1), the strategy which always cooperates; *RAND* (.5 .5 .5 .5), which moves randomly; and *ALT* (0 0 1 1), which alternates between cooperation and defection. The Tit for Tat (*TFT*) strategy, which echoes its opponent's last move (1 0 1 0), has achieved great success in many variants of the Prisoner's Dilemma game (Axelrod, 1984); we consider both this strategy and the generous Tit for Tat (*GTFT*) strategy (1 g 1 g). The Pavlov (*PAV*) strategy (1 0 0 1) is able to achieve self-cooperation in the Iterated Prisoner's Dilemma with noise (Kraines and Kraines 1989, 1993), while the Firm But Fair (*FBF*) strategy (1 0 1 g) is able to achieve self-cooperation in the Alternating Prisoner's Dilemma with noise (Frean, 1994).

We first conducted a round-robin TTD tournament between eight strategies: *ALLD*, *ALLC*, *RAND*, *ALT*, *TFT*, *GTFT*(.5), *PAV*, and *FBF*(.5). The standard payoff table ($T = 7$, $R = 3$, $P = 1$, $S = 0$) was used, and a noise level $\varepsilon = .01$ was assumed. The top strategy in this tournament was *ALLD*, with an average score of 3.44; *ALT* was second with 3.15, and *TFT* was third with 2.92. These results suggest that none of the best known Prisoner's Dilemma strategies are also successful in the Turn-Taking Dilemma.

To confirm this result, we computed the average dominance score of each of the eight strategies against 10000 randomly selected "memory two edge strategies." A memory two strategy is defined by a 16-tuple (p_{RR} p_{RS} p_{RT} p_{RP}; p_{SR} p_{SS} p_{ST} p_{SP}; p_{TR} p_{TS} p_{TT} p_{TP}; p_{PR} p_{PS} p_{PT} p_{PP}). Each of these values is a probability of cooperation based on the past two rounds: for example, p_{RS} is the probability of cooperating if a sucker payoff was received last round, and a reward payoff was received the previous round. An edge strategy is one for which each value is either 0 (always defect), .5 (cooperate half the time), or 1 (always cooperate). Similarly, a corner strategy is one for which each value is either 0 or 1. There are $3^{16} = 43046721$ memory two edge strategies, of which $2^{16} = 65536$ are memory two corner strategies.

Strategy (p_R p_S p_T p_P)	Dominance score
ALLD (0 0 0 0)	.8687
FBF (1 0 1 .5)	.6929
PAV (1 0 0 1)	.6577
RAND (.5 .5 .5 .5)	.6133
ALT (0 0 1 1)	.5751
TFT (1 0 1 0)	.5740
GTFT (1 .5 1 .5)	.1535
ALLC (1 1 1 1)	.0570

Table 3: Dominance scores of some memory one strategies

When the average dominance score of each strategy was computed, *ALLD* outperformed the other strategies by a substantial margin, with a dominance score of .8687, and no other strategy had a dominance score above .70. This confirms that none of the high-performing Prisoner's Dilemma strategies are also successful in the TTD. To find a cooperative strategy which outperforms the unconditional defector *ALLD*, we consider the three optimality criteria for the Turn-Taking Dilemma. In particular, a successful strategy must be able to achieve turn-taking in games against its clone. For memory one strategies (p_R p_S p_T p_P), a turn-taking strategy should cooperate after receiving a temptation payoff, and defect after receiving a sucker payoff, allowing the turn-taking to continue after it has been achieved:

D C D C ...
C D C D ...

This implies $p_S = 0$ and $p_T = 1$. What should p_R and p_P be? Assuming the (7,3,1,0) payoff table, a strategy receives an average of 3.5 points per round ($\frac{T+S}{2}$) when turn-taking, 3 points per round when mutually cooperating, and 1 point per round when mutually defecting. We assume that the goal is to achieve alternation as quickly as possible after an error: each round of mutual cooperation costs the strategy 0.5 points, and each round of mutual defection costs the strategy 2.5 points, in a game against its clone. The easiest way to resume alternation after an *R* or *P* payoff is for both players to move randomly: $p_R = p_P = .5$. This means that each turn there is a 50% chance of resuming alternation, and alternation will be resumed in an average of two turns. Alternatively, we can compute the number of points lost from a CC or DD error. A CC error is when the player who cooperated last turn accidentally cooperates again (a lower case letter is used to represent an error):

D C D C c
C D C D C

A DD error is when the player who defected last turn accidentally defects again:

D C D d
C D C D

We can use Markov chains to compute the cost of CC or DD errors. From the CC or DD state there is a 50% chance of resuming alternation (0 points lost), a 25% chance of mutual cooperation (.5 points lost, return to CC/DD state), and a 25% chance of mutual defection (2.5 points lost, return to CC/DD state). This gives us the equation: $A = .25(A+.5)+.25(A+2.5) \longrightarrow A = 1.5$. This implies a CC error loses $A+.5 = 2$ points, and a DD error loses $A+2.5 = 4$ points.

Thus we have found a strategy that achieves self-alternation ($\sigma_S = 1$), and takes only a short time to recover from CC or DD errors. We call this (.5 0 1 .5) strategy $RALT_1$: it is a "memory one random alternator" that alternates if possible and moves randomly otherwise. We compute the dominance score of $RALT_1$ against the memory two edge strategies, and discover that $RALT_1$ has a dominance score of .8917, higher than *ALLD*. Can we do better? To answer this question, we consider the second and third optimality criteria: to be successful in the TTD, a strategy should be able to exploit *ALLC* and resist exploitation by *ALLD*. $RALT_1$ can exploit *ALLC* only a third of the time ($\sigma_C = \frac{1}{3}$), and is exploited by *ALLD* a third of the time ($\sigma_D = \frac{2}{3}$). A self-alternating strategy with higher σ_C and σ_D is likely to outperform $RALT_1$.

To find a better strategy, we calculate the dominance score of each of the $3^4 = 81$ memory one edge strategies. Of these strategies, the top performer is (0 0 1 .5), with a dominance score of .9203. This strategy is similar to $RALT_1$, but always defects after CC. Thus the strategy can still resume alternation after an error, but must reach the DD state before alternation, even after a CC error. It is still self-alternating ($\sigma_S = 1$), but takes slightly longer to recover from an error. From the DD state there is a 50% chance of resuming alternation (0 points lost), a 25% chance of mutual defection (2.5 points lost, return to DD state), and a 25% chance of mutual cooperation followed by mutual defection (3 points lost, return to DD state). This gives the equation: $A = .25(A+3)+.25(A+2.5) \longrightarrow A = 2.75$. This implies that a CC error loses $A+3 = 5.75$ points, and a DD error loses $A+2.5 = 5.25$ points.

Thus the strategy loses 5.5 points per error, as opposed to 3 points per error for $RALT_1$ (an average of one extra round of mutual defection after an error). This disadvantage is minor compared to the resulting increase in exploitation of unconditional cooperators: $\sigma_C = \frac{1}{2}$, as opposed to $\sigma_C = \frac{1}{3}$ for $RALT_1$. Thus the optimum memory one edge strategy is not the best coordinator of turn-taking, but a slightly more exploiting variant that is also able to achieve self-alternation.

6 Memory two strategies

To improve on this result, we can consider the space of memory two strategies. As discussed above, a memory two strategy chooses whether to cooperate based on its payoffs (R, S, T, or P) in the last two rounds, and thus a memory two strategy is defined by a 16-tuple representing probabilities of cooperation after each combination of payoffs: (p_{RR} p_{RS} p_{RT} p_{RP}; p_{SR} p_{SS} p_{ST} p_{SP}; p_{TR} p_{TS} p_{TT} p_{TP}; p_{PR} p_{PS} p_{PT} p_{PP}).

We now consider what combinations of values will result in a successful TTD strategy. In order to achieve self-alternation, a strategy should cooperate after a temptation payoff, but only in the case of turn-taking; if it can exploit the opponent, the strategy may be better off defecting after a temptation payoff. If the strategy received a temptation payoff last round, and the opponent received a temptation payoff the previous round, turn-taking is going as planned, and it is the strategy's turn to cooperate. This implies that $p_{ST} = 1$. In any case, the strategy should defect if the opponent received a temptation payoff last turn: it does not matter whether it is simply "taking its turn" or responding to its opponent's attempt at exploitation. This implies $p_{RS} = p_{SS} = p_{TS} = p_{PS} = 0$. Similarly, the strategy must defect after *SP*: otherwise an opponent can defect continually against the strategy, exploiting it every other round. This implies that the top strategy should be of the form (p_{RR} 0 p_{RT} p_{RP}; p_{SR} 0 1 0; p_{TR} 0 p_{TT} p_{TP}; p_{PR} 0 p_{PT} p_{PP}).

To attain values for the other ten parameters, we consider how to deal with DD and CC errors: when, in the course of alternation, either the last defector or the last cooperator accidentally repeats his move. When a DD error occurs, the last defector has accidentally defected a second time, depriving the other player of his "turn" to receive the temptation payoff. It makes sense for the accidental defector to cooperate and allow the other player his turn:

D	C	D	d	C	D	…
C	D	C	D	D	C	…

This implies that a strategy should cooperate after a punishment payoff if the previous payoff was temptation (resuming the turn-taking after a DD error), and should cooperate after a temptation payoff if the previous payoff was punishment (continuing the turn-taking after a DD error). Hence we set $p_{TP} = p_{PT} = 1$.

Now we must deal with a CC error. In this case, the last cooperator has accidentally cooperated a second time, not taking the temptation payoff to which he is entitled. Who should take the next turn? This suggests three options. In the first option, the other player waits for him to take his turn:

D	C	c	D	C	…
C	D	C	C	D	…

For this to occur, the strategy must cooperate after TR, defect after SR, and cooperate after RT. This is a "patient" memory two alternator strategy, which we denote by $PALT_2$: (p_{RR} 0 1 p_{RP}; 0 0 1 0; 1 0 p_{TT} 1; p_{PR} 0 1 p_{PP}).

In the second option, the accidental cooperator is assumed to have forfeited his turn: he cooperates, and allows the other player to receive the temptation payoff:

D	C	c	C	D	...
C	D	C	D	C	...

For this to occur, the strategy must defect after TR, cooperate after SR, and cooperate after RT. This is an "impatient" memory two alternator strategy, which we denote by $IMPALT_2$: (p_{RR} 0 1 p_{RP}; 1 0 1 0; 0 0 p_{TT} 1; p_{PR} 0 1 p_{PP}).

In the third option, both players move randomly after a CC error, similar to the $RALT_1$ strategy. For this to occur, we assume $p_{TR} = p_{SR} = .5$, and $p_{RT} = 1$ (to continue alternation once it is resumed). This is a "random" memory two alternator strategy, which we denote by $RALT_2$: (p_{RR} 0 1 p_{RP}; .5 0 1 0; .5 0 p_{TT} 1; p_{PR} 0 1 p_{PP}).

To examine these three types of "alternator" strategies, we find the most successful example of each among the $3^5 = 243$ edge strategies fitting that pattern. For each, the strategy with the highest dominance score against the memory two edge strategies was selected. For all three strategy classes, the most successful strategy had $p_{TT} = 0$, allowing it to continually exploit an unconditional cooperator if this state was reached. The most successful strategy in each class also had $p_{RR} = p_{PR} = .5$: if it is not known which strategy took the last turn, both strategies move randomly after mutual cooperation in order to achieve the higher payoff of turn-taking.

The top $PALT_2$ strategy was (.5 0 1 0; 0 0 1 0; 1 0 0 1; .5 0 1 0), with a dominance score of .9386. The top $IMPALT_2$ strategy was (.5 0 1 0; 1 0 1 0; 0 0 0 1; .5 0 1 .5), with a dominance score of .9605. The top $RALT_2$ strategy was (.5 0 1 0; .5 0 1 0; .5 0 0 1; .5 0 1 .5), with a dominance score of .9511.

Comparing these results to the top memory one edge strategy (0 0 1 .5) with dominance .9203, we note two main facts. First, as in the Prisoner's Dilemma (Neill, 2001), higher memory strategies perform substantially better than the low-memory strategies commonly studied in the context of these problems. Second, it appears that the "impatient" method of responding to a CC error is most successful.

To understand why these strategies perform better than (0 0 1 .5), we analyze them in light of our optimality criteria. All three are self-alternating ($\sigma_S = 1$), and respond well to CC or DD errors. For $PALT_2$ and $IMPALT_2$, alternation is restored immediately after either kind of error: a CC error loses only .5 points, and a DD error loses only 2.5 points. For $RALT_2$, a DD error loses 2.5 points. To compute the number of points lost for a CC error, we let A represent the number of points lost in the R state and B represent the number of points lost in the PP state. This gives us: $A = .25(A+.5)+.25(B+5)$ and $B = .25(A+.5)+.25(B+2.5)$. Solving these equations, we get $A = \frac{39}{16}$; the number of points lost for a CC error is $A+.5 = 2.9375$.

All three types of strategies have similar values for σ_D: $\sigma_D = 1$ if $p_{PP} = 0$, and $\sigma_D = \frac{3}{4}$ if $p_{PP} = .5$. There is an interesting tradeoff here: a strategy with $p_{PP} = 0$ is totally D-unexploitable, but a strategy with $p_{PP} = .5$ has slightly higher self-payoff (since it cannot fall into a rut of continued defection with its clone). For $PALT_2$, the D-unexploitability is of greater benefit, while for $RALT_2$ and $IMPALT_2$, the higher self-payoff is more significant.

When we examine σ_C, however, we notice substantial differences between the three types of strategy. While $PALT_2$ "patiently" waits for $ALLC$ to take its turn, $IMPALT_2$ (and sometimes $RALT_2$) will take advantage of $ALLC$ immediately. Hence $IMPALT_2$ is the most C-exploiting ($\sigma_C = \frac{3}{5}$), $PALT_2$ is the least C-exploiting ($\sigma_C = \frac{1}{3}$), and $RALT_2$ is moderately C-exploiting ($\sigma_C = \frac{3}{7}$). Thus $IMPALT_2$ is most successful because it is able to best exploit unconditional cooperators while still maintaining high self-cooperation and D-unexploitability.

7 Dominance of memory two corner strategies

In order to achieve a more comprehensive examination of the space of memory two strategies, we conducted a round-robin dominance tournament for the $2^{16} = 65536$ memory two corner strategies. As discussed above, a corner strategy is a strategy for which each probability of cooperation is either 0 or 1; the average dominance score of each strategy against all 2^{16} strategies was computed, and the ten strategies with the highest dominance scores were recorded.

Strategy	dom vs. mem 2 corner	dom vs. mem 2 edge
(0 0 0 0; 1 0 1 0; 0 0 0 1; 0 0 1 0)	.9463	.9612
(1 0 0 0; 1 0 1 0; 0 0 0 1; 0 0 1 0)	.9463	.9610
(0 0 0 1; 1 0 1 0; 0 0 0 1; 0 0 1 0)	.9406	.9593
(1 0 0 1; 1 0 1 0; 0 0 0 1; 0 0 1 0)	.9403	.9589
(0 0 0 0; 1 0 1 0; 0 0 0 1; 1 0 1 0)	.9403	.9570
(0 0 1 0; 1 0 1 0; 0 0 0 1; 0 0 1 0)	.9402	.9533
(1 0 1 0; 1 0 1 0; 0 0 0 1; 0 0 1 0)	.9400	.9536
(1 0 1 0; 1 0 1 0; 0 0 0 1; 1 0 1 0)	.9396	.9532
(0 0 1 0; 1 0 1 0; 0 0 0 1; 1 0 1 0)	.9393	.9528
(0 0 0 0; 0 0 1 0; 1 0 0 1; 0 0 1 0)	.9391	.9519

Table 4: Dominance scores of top memory two corner strategies

All of these ten strategies are variants of the top strategy, (0 0 0 0; 1 0 1 0; 0 0 0 1; 0 0 1 0). This strategy has $\sigma_S = 1$, $\sigma_C = 1$, and $\sigma_D = 1$, and thus fulfills all of the optimality criteria. How does this strategy work? Since it has $p_{TS} = 0$ and $p_{ST} = 1$, it can take turns with its clone. It handles DD errors in the obvious manner: the accidental defector cooperates, allowing the other player to take his turn.

D	C	D	d	C	D	…
C	D	C	D	D	C	…

Thus it loses only 2.5 points for a DD error; this is the best possible handling of the error. Its response to a CC error is quite interesting; it is "impatient," but also reaches a single round of mutual defection before it resumes alternation:

D	C	c	D	D	C	D	…
C	D	C	C	D	D	C	…

Thus it loses 3 points for a CC error. Though its defection after *RT* prevents it from immediately restoring alternation after a CC error, this also allows it to exploit *ALLC* continually, giving it $\sigma_C = 1$. This strategy is a memory two C-exploiting alternator, which we denote by $CALT_2$. We note that it is possible for $CALT_2$ to fall into a rut of mutual defection with its clone, since it has $p_{PP} = 0$:

D	C	D	d	d	D	D	…	D	D	c	D	C	…
C	D	C	D	D	D	D	…	D	D	D	C	D	…

However, we note that it takes two errors to fall into the DD rut (probability $O(\varepsilon^2)$), and only one error to escape from the DD rut (probability $O(\varepsilon)$). Hence it is in the DD rut with probability on the order of ε, and thus it is totally self-cooperating ($\sigma_S = 1$) for infinitesimal noise. Its self-payoff is $3.5 - 13\varepsilon$: it loses 2.5ε for its handling of CC errors, 3ε for its handling of DD errors, and 7.5ε for falling into DD ruts.

Thus the $CALT_2$ strategy not only meets our three optimality criteria, but has the highest dominance score of all memory two corner strategies. Against the memory two edge strategies, it also performs very well, with a dominance score of .9612. This is better than the top $PALT_2$, $IMPALT_2$, and $RALT_2$ strategies. While no memory two corner strategy does better, if we extend our search to the memory two edge strategies some simple improvements can improve this strategy's dominance score.

First, increasing p_{RP} to .5 improves the dominance score. The resulting strategy, (0 0 0 .5; 1 0 1 0; 0 0 0 1; 0 0 1 0), has a dominance score of .9647. If we reduce p_{PT} to .5, this further increases the dominance score. The resulting strategy, (0 0 0 .5; 1 0 1 0; 0 0 0 1; 0 0 .5 0), has a dominance score of .9754. Like $CALT_2$, it has $\sigma_S = \sigma_D = \sigma_C = 1$, but its chance of defection after *PT* results in an average of one extra round of punishment against its clone:

D	C	d	D	D	C	D	D	C	…
C	D	D	C	D	D	D	C	D	…

Thus it loses 5 points for a DD error, and 5.5 points for a CC error, resulting in a self-payoff of $3.5-18\varepsilon$. However, its chance of defection after PT also improves its performance against strategies which are 50% D-exploitable, such as Pavlov. These strategies will cooperate half the time whether they are given their turn or not: thus the best strategy is continual defection, which will result in payoff $\frac{T+P}{2}$, rather than alternation, with payoff $\frac{T+S}{2}$. This strategy does not defect continually against 50% D-exploitable strategies, but exploits them more often than $CALT_2$. Thus, as in the case of memory one strategies, we find that performance is increased by a strategy which is a slightly worse turn-taker, but compensates for this by improved exploitation of (unconditional and 50%) cooperators.

8 Dominance of memory two edge strategies

In our exploration so far, we have found strategies with dominance scores up to .9754 against the space of memory two edge strategies. Can we do better? To find out, we conducted a dominance tournament for the $3^{16} = 43046721$ memory two edge strategies. The average dominance score of each strategy against 100000 randomly selected memory two edge strategies was computed, and the ten strategies with the highest dominance scores were recorded. Multiple tests were run in order to insure a large enough sample size: this was confirmed by identical results for each of the tests.

Strategy	Dominance
(.5 0 1 0; 1 0 1 0; 0 0 0 0; .5 0 0 .5)	.9863
(.5 0 1 0; 1 .5 1 0; 0 0 0 0; .5 0 0 .5)	.9862
(.5 0 1 .5; 1 0 1 0; 0 0 0 0; .5 0 0 .5)	.9854
(.5 0 1 .5; 1 .5 1 0; 0 0 0 0; .5 0 0 .5)	.9853
(1 0 1 0; 1 0 1 0; 0 0 0 0; .5 0 0 .5)	.9853
(1 0 1 0; 1 .5 1 0; 0 0 0 0; .5 0 0 .5)	.9853
(0 0 1 0; 1 0 1 0; 0 0 0 0; .5 0 0 .5)	.9845
(0 0 1 0; 1 .5 1 0; 0 0 0 0; .5 0 0 .5)	.9844
(1 0 1 .5; 1 0 1 0; 0 0 0 0; .5 0 0 .5)	.9842
(1 0 1 .5; 1 .5 1 0; 0 0 0 0; .5 0 0 .5)	.9842

Table 5: Dominance scores of top memory two edge strategies

All of the top ten strategies are variants of the top strategy, (.5 0 1 0; 1 0 1 0; 0 0 0 0; .5 0 0 .5). This remarkable strategy has a dominance score of .9863, more than .02 higher than $CALT_2$. How does it work? It has $p_{TS} = 0$ and $p_{ST} = 1$, so it can take turns with its clone. It resolves a CC error by the "impatient" method; the accidental cooperator loses his turn, and alternation continues:

D	C	c	C	D	...
C	D	C	D	C	...

Thus a CC error results in a loss of only 0.5 points; this is the best possible handling of a CC error. On the other hand, a DD error results in a long string of defections. Since the strategy defects after TP and SP, the DD error drives it first into the PP state. From the PP state, both players move randomly, but they cannot resume alternation since the strategy defects after both PS and PT. Instead, it has a 25% chance of moving into the PR state from the PP state, and a 50% chance of resuming alternation from the PR state. To compute the number of points lost from as CC error, let A be the number of points lost in the PR/RR state and B be the number of points lost in the PP state. This gives us the equations: $A = .25(.5+A) + .25(5+B)$ and $B = .25(.5+A) + .25(2.5+B) + .5(5+B)$. Solving these equations, we get A=9.25, B=22.25, and B+5=27.25 points lost for a DD error; this is approximately 11 punishments per error. Hence the strategy takes a long time to restore cooperation after a DD error. It has a self-payoff of $3.5-27.75\varepsilon$, and thus it may perform poorly for large amounts of noise. Nevertheless, the strategy is self-cooperating under infinitesimal noise: it has $\sigma_S = 1$. Since it cooperates half the time after PP, it can be exploited by $ALLD$ an average of once every four turns, and has $\sigma_D = .75$. It also exploits $ALLC$ three-fourths of the time, giving it $\sigma_C = .75$.

The strategy has lower σ_D and σ_C than $CALT_2$, and though both have $\sigma_S = 1$, it takes much longer to recover from an error. Why, then, does it perform so much better? The answer lies in its ability to exploit 50% cooperators such as Pavlov: since it defects after both PT and TP, it will defect continually if its opponent alternates cooperation and defection:

D	D	D	D	D	D	...
D	C	D	C	D	C	...

Against these strategies, it achieves a payoff of $\frac{T+P}{2}$. This is significantly higher than the highest "cooperative" payoff against these strategies, $\frac{T+S}{2}$. With the standard (7,3,1,0) payoff table, the strategy averages 4 points per round against 50% cooperators, while turn-takers can only score 3.5 points per round. Also, since the highest possible self-payoff for a strategy is 3.5 points per round, this strategy will invade all 50% cooperators that it can exploit in this manner, resulting in high dominance scores. Thus we name the strategy $EXALT_2$: not only for its "exalted" place among the memory two edge strategies, but because it is a "memory two exploiting alternator," capable of exploiting unconditional and 50% cooperators.

The success of $EXALT_2$ demonstrates that the three optimality criteria as proposed for the Turn-Taking Dilemma are insufficient: in addition to achieving turn-taking with its clone under noise, resisting exploitation by *ALLD*, and exploiting *ALLC*, a successful strategy should also be able to exploit the much larger class of 50% cooperators. Any strategy which meets the three optimality criteria, as well as achieving high payoffs against 50% cooperators, is likely to achieve great success in the Turn-Taking Dilemma with noise.

How important are these four criteria relative to each other? We prioritize the optimality criteria by examining the characteristics of all top strategies. We suggest the following list of rules, arranged in descending order of importance:

1. A successful strategy must have $\sigma_S = 1$ to achieve turn-taking with its clone; a strategy which takes turns with its clone is also likely to achieve turn-taking against a wide variety of other strategies.

2. A successful strategy must have $w(ALLD|X) < \frac{T+S}{2}$ to prevent invasion by defectors. With the (7,3,1,0) payoff table, this implies $\sigma_D > \frac{7}{12}$.

3. A successful strategy must have $\sigma_C > 0$ to exploit unconditional cooperators.

4. A successful strategy should be able to exploit both unconditional and 50% cooperators a significant fraction of the time: higher values of σ_C, and higher scores against 50% cooperators, are better.

5. A successful strategy should resist exploitation by unconditional defectors a significant fraction of the time: higher values of σ_D are better.

6. A successful strategy should have a self-payoff that is not reduced too sharply as a function of ε; if the self-payoff can be expressed in terms of the noise level as $\frac{T+S}{2} - a\varepsilon$, lower values of a are better.

9 Conclusions

The Turn-Taking Dilemma is a variant of the repeated Prisoner's Dilemma in which the optimal payoff is achieved not by simultaneous mutual cooperation, but by taking turns. Though it shares some characteristics with both the Iterated Prisoner's Dilemma and the Alternating Prisoner's Dilemma, all three models are distinct, and model different sets of real-world situations. In particular, the Turn-Taking Dilemma is a better model of "compromise" situations, in which one must take short-term losses to achieve a high average payoff in the long run. Also, it describes interactions which require not only a willingness to cooperate, but also a careful coordination of decisions. Examples of the Turn-Taking Dilemma can be found in fields ranging from economics to biology: it occurs frequently in everyday business decisions, in animal behavior, and even in literature.

The Turn-Taking Dilemma violates one of the two Prisoner's Dilemma inequalities, resulting in a situation in which alternating temptation and sucker payoffs achieve a higher score than continued reward payoffs. As a result of this, turn-taking "cooperation" occurs when each player receives half of the defecting (high-payoff) turns, and half of the cooperating (low-payoff) turns. Since each player is expected to defect half the time, exploitation occurs when one player takes more than his fair share of turns, harming the other player. Also, since mutual cooperation and mutual

Strategy	Dominance	σ_S	self-payoff (CC, DD, rut)	σ_C	σ_D	$w(X\|PAV)$
$EXALT_2$	.9863	1	$3.5-27.75\varepsilon$ (.5 27.25 0)	$\frac{3}{4}$	$\frac{3}{4}$	4.00
$CALT_{2a}$	.9754	1	$3.5-18\varepsilon$ (5.5 5 7.5)	1	1	3.67
$CALT_2$	.9612	1	$3.5-13\varepsilon$ (3 2.5 7.5)	1	1	3.10
$IMPALT_2$	.9605	1	$3.5-3\varepsilon$ (.5 2.5 0)	$\frac{3}{5}$	$\frac{3}{4}$	2.67
$RALT_2$	.9511	1	$3.5-5.44\varepsilon$ (2.94 2.5 0)	$\frac{3}{7}$	$\frac{3}{4}$	2.67
$PALT_2$	.9386	1	$3.5-6.33\varepsilon$ (.5 2.5 3.33)	$\frac{1}{3}$	1	2.67
$RALT_{1a}$	.9203	1	$3.5-11\varepsilon$ (5.75 5.25 0)	$\frac{1}{2}$	$\frac{2}{3}$	2.71
$RALT_1$	.8917	1	$3.5-6\varepsilon$ (2 4 0)	$\frac{1}{3}$	$\frac{2}{3}$	2.75
$ALLD$	.8687	0	$1+5\varepsilon$ (X X X)	1	1	4.00

Table 6: Table of top strategies

defection result in suboptimal payoffs, a failure to coordinate one player's cooperation with the other player's defection results in harm to both players.

Our aim in this paper was twofold: not only to present the Turn-Taking Dilemma as a new model of cooperative behavior, but also to explore the problem using the techniques discussed in Neill (2001), in order to find which strategies are most successful in Turn-Taking Dilemma interactions. Using the three "optimality criteria" (self-alternating, D-unexploitable, C-exploiting) as a guide, and "evolutionary dominance" as a general measure of evolutionary performance (combining evolutionary stability and the ability to invade other strategies), we explored the space of memory two edge strategies, discovering and examining a number of highly successful strategies. Table 6 summarizes these results, while Table 7 gives a description of each top strategy. We demonstrated that the top Prisoner's Dilemma strategies do not perform well in the Turn-Taking Dilemma; this is not surprising since they are only able to achieve continuous mutual cooperation rather than turn-taking. More importantly, we were able to show that turn-taking can be achieved even by strategies with very limited memory. All of the most successful Turn-Taking Dilemma strategies were able to achieve (and sustain) turn-taking with a wide variety of other strategies, while exploiting unconditional cooperators and resisting exploitation by defectors. Successful strategies differed in their methods of recovering from errors, and in their trade-offs between exploitation of cooperators and resistance to defectors; we examined these trade-offs in detail. In general, we found that the most successful TTD strategies are not necessarily the best coordinators of turn-taking, but slightly more exploiting variants that are also able to achieve turn-taking with their clones. The most successful memory two edge strategy, $EXALT_2$, was able to exploit not only unconditional cooperators but also "50% cooperators" (strategies which cooperate half the time against unconditional defectors), while maintaining high levels of self-alternation and D-unexploitability.

As in the Alternating Prisoner's Dilemma, a successful Turn-Taking Dilemma strategy must be "friendly" enough to cooperate with its clone, "pragmatic" enough to exploit cooperators, and "wary" enough to resist exploitation by defectors. Even within this simple framework, however, there are substantial differences. Since the strategy must achieve turn-taking with its clone (a more complex interaction than simple continued cooperation) it must be not only "friendly" but also "precise" in its ability to coordinate this alternation under noise. A "pragmatic" strategy in the Turn-Taking Dilemma will exploit not only unconditional cooperators, but also 50% cooperators, and hence a strategy must be "wary" of being exploited in this fashion as well. Strategies which can effectively coordinate turn-taking under noise, while exploiting cooperators and resisting exploitation by defectors, are likely to achieve great success in the variety of real-world interactions modeled by the Turn-Taking Dilemma.

The Turn-Taking Dilemma is one example of an interaction in which "cooperation" in the simple sense of the Prisoner's Dilemma is not sufficient for success. For these situations, we need a model of cooperation in complex behaviors, in which the cooperating players must precisely coordinate their actions in a changing and unpredictable environment. These include a variety of interactions where maximizing a player's long term payoff does not necessarily result from maximizing his payoff every turn: situations marked by the intricate give and take of bargaining and compromise. Thus our exploration of the Turn-Taking Dilemma is the first step toward a general model of cooperative behavior: a step toward understanding the more complex forms of cooperation which shape the behavior of individuals and societies.

Name	Strategy
$EXALT_2$	(.5 0 1 0; 1 0 1 0; 0 0 0 0; .5 0 0 .5)
$CALT_{2a}$	(0 0 0 .5; 1 0 1 0; 0 0 0 1; 0 0 .5 0)
$CALT_2$	(0 0 0 0; 1 0 1 0; 0 0 0 1; 0 0 1 0)
$IMPALT_2$	(.5 0 1 0; 1 0 1 0; 0 0 0 1; .5 0 1 .5)
$RALT_2$	(.5 0 1 0; .5 0 1 0; .5 0 0 1; .5 0 1 .5)
$PALT_2$	(.5 0 1 0; 0 0 1 0; 1 0 0 1; .5 0 1 0)
$RALT_{1a}$	(0 0 1 .5)
$RALT_1$	(.5 0 1 .5)
ALLD	(0 0 0 0)

Table 7: Descriptions of top strategies

10 Bibliography

• Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

• Axelrod, R. and Hamilton, W.D. (1981). "The evolution of cooperation," *Science* **211**, 1390-1396.

• Frean, M.R. (1994). "The prisoner's dilemma without synchrony," *Proc. Roy. Soc. Lond. B* **257**, 75-79.

• Henry, O. (1905). "The Gift of the Magi."

• Kraines, D. and Kraines, V. (1989). "Pavlov and the prisoner's dilemma," *Theory and Decision* **26**, 47-79.

• Kraines, D. and Kraines, V. (1993). "Learning to cooperate with Pavlov," *Theory and Decision* **35**, 107-150.

• Kraines, D. and Kraines, V. (2001). "Protocols for cooperation: cultural diversity of strategies for the alternating prisoner's dilemma," in preparation.

• Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

• Milinski, M. (1987). "Tit for tat in sticklebacks and the evolution of cooperation," *Nature* **325**, 433-435.

• Neill, D.B. (2001). "Optimality under noise: higher memory strategies for the alternating prisoner's dilemma," *J. Theor. Biol.* **211**, 159-180.

• Nowak, M.A. and Sigmund, K. (1994). "The alternating prisoner's dilemma," *J. Theor. Biol.* **168**, 219-226.

• Rasa, O.E.A. (1989). "The costs and effectiveness of vigilance behavior in the dwarf mongoose: implications for fitness and optimal group size," *Ethol. Ecol. Evol.* **1**, 265-282.

• Wilkinson, G.S. (1984). "Reciprocal food-sharing in the vampire bat," *Nature* **308**, 181-184.