$n-$bits reactive strategies

Nikoleta E. Glynatsi, Ethan Akin, Martin Nowak, Christian Hilbe

Abstract

We focus on repeated games and the strategies players can employ in these games. It is famously assumed in repeated games that strategies utilizing past history can be employed. Here, we concentrate on a specific set of strategies known as reactive strategies. Notably, our work stands out by exploring strategies with longer memory spans, in contrast to the prevailing norm in existing literature.

We demonstrate, how this set of strategies have an imidate effect of the co-oplayer and show that a history that is not shared does not befenit the longer strategy.

In the following we study repeated games and the strategies player's can derive in these games. Famously, in repeated game it is assumed that strategies that use the past history can be adopted. Here we focus on such a set, called the reactive strategies. In comparison to previous studies this work explores higher memory strategies, greater than one compared to the majority of the works in the literature. We demonstrate, how this set of strategies have an imidate effect of the co-oplayer and show that a history that is not shared does not befenit the longer strategy.

We then characterise partner strategies, which are strategies assuring mutual cooperation without being exploited. A recipe for evolutionary stability. We show that the class of Tit For Tat and Generous Tit For strategies even the delayed versions of these strategies are partner strategies.

For memory lengths of two and three we characterize all partner strategies amongst the reactive set. The conditions simple and yet newly found. For a specific class of reactive counting strategies, counting of defections/defections instead of remembering the actual occurance of the action, we can characterize partner strategies in all memory lengths.

We further test the evolutionary properties of partner strategies in higher memory. The results show that.

1 Introduction

2 Model

Repeated Game. We consider an infinitely repeated game with two players. In each round, player 1 and player 2, can choose to cooperate (C) or to defect (D). If both players cooperate, they receive a payoff R (the reward for mutual cooperation), and if both players defect, they receive a payoff P (the punishment for mutual defection). If one player cooperates, the cooperative player receives the sucker's payoff S, and the defecting player receives the temptation payoff T. We assume that the payoff are such that $T > R > P > S$ and $2R > T + S$. This game is known as the Prisoner's Dilemma. Here, we employ a specific parametrization of the Prisoner's Dilemma, where cooperation implies incurring a cost c for the co-player to derive a benefit $b > c$. Consequently, the payoffs are defined as follows: $R = b - c, S = -c, T = b, P = 0$. In the Appendix, we present results applicable to the general Prisoner's Dilemma.

We assume in the following, that the players' decisions only depend on the outcome of the previous n rounds. To this end, an *n-history for player $i \in \{1, 2\}$* is a string $h^i = (a^i_{-n}, \ldots, a^i_{-1}) \in \{C, D\}^n$ where an entry a^i_{-k}

corresponds to player i's action k rounds ago. Let H^i denote the space of all n-histories for player i where set H^i contains $|H^i| = 2^n$ elements. A *reactive−n strategy* for player 1 is a vector $\mathbf{p} = (p_h)_{h \in H^2} \in [0,1]^n$. Each entry p_h corresponds to the player's cooperation probability in the next round, based on the co-player's actions in the previous n rounds. Therefore, reactive-n strategies exclusively rely on the co-player's n-history, independent of the focal player's own actions. For $n = 1$, this definition of reactive-n strategies recovers the typical format of reactive-1 strategies [Baek et al., 2016, Wahl and Nowak, 1999, McAvoy and Nowak, 2019], $\mathbf{p} = (p_C, p_D,)$.

Another class of strategies we will be discussing in this work are, *self-reactive-n* strategies which only consider the focal player's own n-history, and ignore the co-player's. Formally, a self-reactive-n strategy for player 1 is a vector $\tilde{\mathbf{p}} = (\tilde{p}_h)_{h \in H^1} \in [0,1]^n$. Each entry $\tilde{p}_h$ corresponds to the player's cooperation probability in the next, depending on the player's own actions in the previous n rounds. We say that a reactive or self-reactive strategy is pure if all the entries of the strategy are either 0 or 1. We refer to the set of all pure self-reactive strategies as $\tilde{P}_n$.

Evolutionary process. To examine the evolutionary properties of reactive strategies, we perform an evolutionary study based on the framework of Imhof and Nowak [Imhof and Nowak, 2010]. The framework considers a population of size N where initially all members are of the same strategy. In our case the initial population consists of unconditional defectors. In each elementary time step, one individual switches to a new mutant strategy. The mutant strategy is generated by randomly drawing cooperation probabilities from the unit interval $[0, 1]$. If the mutant strategy yields a payoff of $\mathbf{s}_{M,k}$, where k is the number of mutants in the population, and if residents get a payoff of $\mathbf{s}_{R,k}$, then the fixation probability ϕ_M of the mutant strategy can be calculated explicitly,

$$\phi_M = \left(1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} e^{(-\beta(\mathbf{s}_{M,j} - \mathbf{s}_{R,i}))}\right)^{-1} \tag{1}$$

The parameter $\beta \geq 0$ is called the strength of selection, and it measures the importance of the relative payoff advantages for the evolutionary success of a strategy. For small values of β, $\beta \approx 0$, payoffs become irrelevant, and a strategy's fixation probability approaches $\phi_M \approx 1/N$. The larger the value of β, the more strongly the evolutionary process favours the fixation of strategies that yield high payoffs.

Depending on the fixation probability ϕ_M the mutant either fixes (becomes the new resident) or goes extinct. Regardless, in the elementary time step another mutant strategy is introduced to the population. We iterate this elementary population updating process for a large number of mutant strategies and we record the resident strategies at each time step.

3 Results

3.1 Self-Reactive Sufficiency

To predict which reactive-n strategies are partner strategies, we must first characterize which nice reactive-n strategies are Nash equilibria. Determining whether a given strategy, $\mathbf{p}$, is a Nash equilibrium is not straightforward. In principle, this would involve comparing the payoff of $\mathbf{p}$ to the payoff of all possible mutant strategies. However, in this work, we demonstrate otherwise. Specifically, we show that if a player adopts a reactive strategy, it is only necessary to consider mutant strategies that are self-reactive-n. In other words, we show the following result (see Appendix for proof),

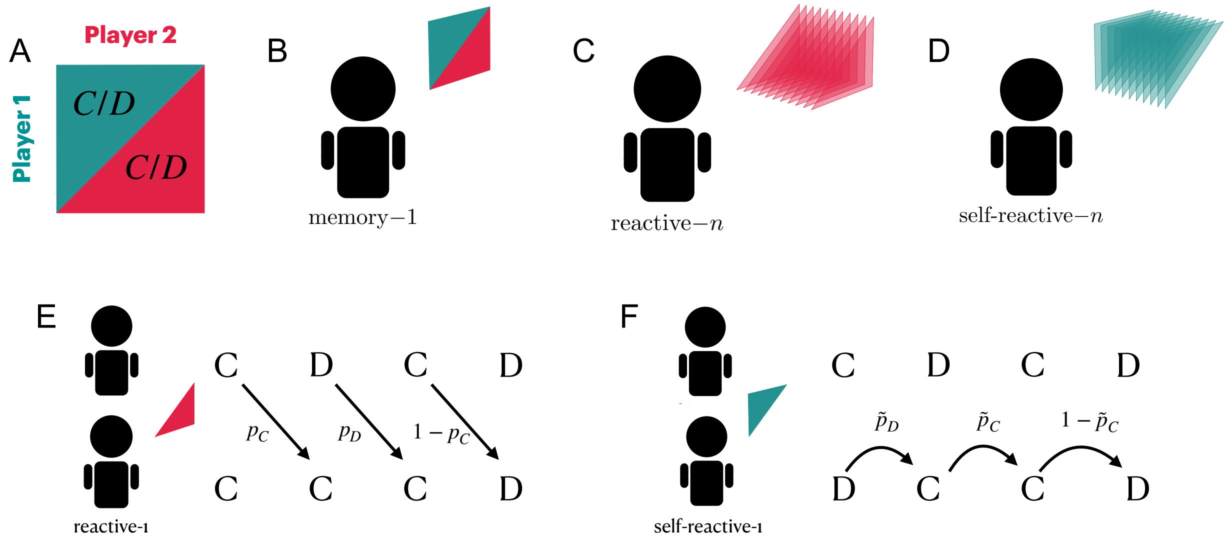

Figure 1: **Model.** **A.** In each turn of the repeated game, players 1 and 2 decide on an action, denoted as C (cooperate) or D (defect), respectively. We assume, that the information that a player can use in subsequent turns is limited to the actions taken by both players in the current turn. **B.** Memory-1 strategies, a well-studied set of strategies, utilize the actions of both players in the previous turn to make decisions. In the graphical representation of memory-1 strategies, we use a single square to illustrate this concept. **C.** This work primarily focuses on reactive-n strategies, which take into account only the actions of the co-players. **E.** For the case of $n = 1$, a reactive-1 strategy is represented as a vector $\mathbf{p} = (p_C, p_D)$, where p_C is the probability of cooperating given that the co-player cooperated, and p_D is the probability of cooperating given that the co-player defected. In the example shown, the bottom player employs a reactive-1 strategy. They cooperate with a probability p_C in the second round because the co-player cooperated in the first round. In the second round, the player cooperates with a probability p_D since the co-player previously defected. Finally, the player defects in the third round with a probability of $1-p_C$, considering that the co-player cooperated. **D.** Another set of strategies we consider is that of self-reactive-n strategies, which rely solely on a player's own previous n actions. **F.** For the case of $n = 1$, a self-reactive-1 strategy is represented as a vector $\tilde{\mathbf{p}} = (\tilde{p}_C, \tilde{p}_D)$, where $\tilde{p}_C$ is the probability of cooperating given that the player's last action was cooperation, and $\tilde{p}_D$ is the probability of cooperating given that the player's last action was defection. In the example shown, the bottom player employs a self-reactive-1 strategy. They cooperate with a probability $\tilde{p}_D$ in the second round given that they defected in the first. In the second round, the player cooperates with a probability $\tilde{p}_C$ since they cooperated in the previous round. Finally, the player defects in the third round with a probability of $1-\tilde{p}_C$, considering that they cooperated in the previous round.

Lemma 3.1. Let $\mathbf{p}$ be a reactive$-n$ strategy for player 1. Then, for any memory$-n$ strategy $\mathbf{m}$ used by player 2, player 1's score is exactly the same as if 2 had played a specific self-reactive memory-n strategy $\tilde{\mathbf{p}}$.

Our result aligns with the findings of [Press and Dyson, 2012]. They explored a scenario where one player uses a memory-1 strategy while the other employs a longer memory strategy. They demonstrated that the payoff of the player with the longer memory is exactly the same as if they had used a specific shorter-memory strategy, disregarding any history beyond what is shared with the short-memory player. Our results hint at a more general insight: if one player does not observe a part of the history, the co-player gains no advantage by considering the unshared history.

Lemma 3.1 allows us to focus on self-reactive strategies when evaluating whether a reactive strategy is a

Nash equilibrium. This significantly reduces the search space for potential mutants. Our next result further reduces this space:

Lemma 3.2. A reactive-n strategy $\mathbf{p}$, is a Nash strategy if and only if the payoff when playing against itself is greater than or equal to any payoff that a pure self-reactive-n strategy, $\tilde{\mathbf{p}} \in \tilde{P}$, can achieve against it.

See Appendix for proof.

In the case of $n = 3$ there are are $2^8 = 256$ possible self-reactive strategies. As opposed to memory-3 strategies where there are $2^8 \times 2^8 = 65536$. Thus, the number of strategies we need to check against is reduced by a factor of 256.

3.2 Partner Strategies Amongst Reactive-2 and Reactive-3 Strategies

Previous studies have characterized subsets of partner strategies for $n = 2$, and while we also focus on characterizing a subset, our contribution extends to the complete set of reactive-2 strategies. Furthermore, we extend the analysis to the case of $n = 3$. We begin by characterizing partner strategies amongst the set of reactive-2.

Theorem 3.3 ("Reactive-2 Partner Strategies")**.** A nice reactive-two strategy $\mathbf{p}$, is a partner strategy if and only if, the strategy entries satisfy the conditions:

$$p_{DD} < 1-\frac{c}{b} \quad and \quad \frac{p_{CD} + p_{DC}}{2} < 1 - \frac{1}{2} \cdot \frac{c}{b}. \tag{2}$$

These conditions can be summarized as follows: For the strategy to be Nash, the strategy ALLD must not be able to invade ($p_{DD} \leq 1 - \frac{c}{b}$ ensures this), and the average cooperation rate following a defection must be less than half of the cost-benefit ratio (c/b).

We can characterizing partner strategies amongst the set of reactive-3. In the case of $n = 3$, a nice reactive-3 strategy is given by a vector

$$\mathbf{p} = (p_{CCC}, p_{CCD}, p_{CDC}, p_{CDD}, p_{DCC}, p_{DCD}, p_{DDC}, p_{DDD}).$$

Theorem 3.4 ("Reactive-Three Partner Strategies")**.** A nice reactive-three strategy $\mathbf{p}$, is a partner strategy if and only if, the strategy entries satisfy the conditions:

$$\begin{aligned}
p_{DDD} &< 1-\frac{c}{b} \\
\frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} &< 1-\frac{1}{3} \cdot \frac{c}{b} \\
\frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} &< 1-\frac{2}{3} \cdot \frac{c}{b} \\
\frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} &< 1-\frac{1}{2} \cdot \frac{c}{b} \\
\frac{p_{CDC} + p_{DCD}}{2} &< 1-\frac{1}{2} \cdot \frac{c}{b}
\end{aligned} \tag{3}$$

Increasing the memory we allow strategies by one results to five conditions instead of two. Inherently, these conditions still exhibit some symmetry with the previous case. The strategy should not be invaded by the strategy ALLD, resulting in the condition $p_{DDD} < 1 - \frac{c}{b}$. Additionally, the average cooperation following a single defection must be lower than 2/3 of the cost-benefit ratio, and the average cooperation following two defections must be smaller than 1/3 of the cost-benefit ratio. There are two additional conditions that do not appear to have clear interpretations. We hypothesize that as the memory space we allow increases, the number of conditions will also increase, and some of the conditions will deviate from the symmetry.

The proofs for both theorems can be found in the Appendix. We can prove the results of this in two independent ways. One leverages the findings of Lemma 3.2, where we explicitly derive the payoff expressions against all pure self-reactive strategies. The second method utilizes the techniques and results presented in [Akin, 2016]. In the Appendix, we demonstrate how one of the central results from Akin's work can be generalized.

3.3 Partner Strategies Amongst Reactive Counting Strategies

A special case of reactive strategies is reactive counting strategies. These are strategies that respond to the co-player's actions, but they do not distinguish between when cooperations/defections occurred; they solely consider the count of cooperations in the last n turns. A reactive-n counting strategy is represented by a vector $\mathbf{r} = (r_i)_{i \in \{n, n-1, \dots, 0\}}$, where the entry r_i indicates the probability of cooperating given that the co-player cooperated i times in the last n turns.

In the case of $n = 1$ a reactive-1 strategy and a counting strategy are equivilent. Since both strategies consider that a defection or cooperation occured in the previous turn. A reactive-2 counting strategies are denoted by the vector $\mathbf{r} = (r_2, r_1, r_0)$, and we can characterise partner strategies among the reactive-2 counting strategies by simply setting $r_2 = 1$, and $p_{CD} = p_{DC} = r_1$ and $p_{DD} = r_0$ in conditions (2). This gives us the following result.

Corollary 3.4.1. A nice reactive-2 counting strategy $\mathbf{r} = (1, r_1, r_0)$ is a partner strategy if and only if,

$$r_1 < 1 - \frac{1}{2} \cdot \frac{c}{b} \quad \textit{and} \quad r_0 < 1 - \frac{c}{b}. \tag{4}$$

Note that even thought the strategies themselves are not equivalent, the conditions for partner are. In both cases cooperating after full defection and the average. The different is that there less counting strategies. They fall in the same space but are a plane where areas counting strategies are. A graphical representation of this is given in Appendix.

Reactive-3 counting strategies are denoted by the vector $\mathbf{r} = (r_3, r_2, r_1, r_0)$. We can characterise partner strategies among reactive-3 counting strategies by setting $r_3 = 1$, and $p_{CCD} = p_{CDC} = p_{DCC} = r_2, p_{DCD} = p_{DDC} = p_{CDD} = r_1$ and $p_{DDD} = r_0$ in conditions (3). This gives us the following result.

Corollary 3.4.2. A nice reactive-3 counting strategy $\mathbf{r} = (1, r_2, r_1, r_0)$ is a partner strategy if and only if,

$$r_2 < 1 - \frac{1}{3} \cdot \frac{c}{b}, \quad r_1 < 1 - \frac{2}{3} \cdot \frac{c}{b} \quad \textit{and} \quad r_0 < 1 - \frac{c}{b}. \tag{5}$$

In the case of reactive-1 strategies, counting strategies are equivalent. However, even in the case of reactive-2 strategies the conditions do not changes. The ratio has to be smaller than. It's the the case of reactive-3 strategies that we observe the biggest difference. That is the are three conditions instead of five. The top conditions we can not account for are the conditions.

The properties of the reactive counting strategies are interesting. They allow us to character partner strategies in all memory lengths.

Corollary 3.4.3 ("Reactive-Counting Partner Strategies"). A nice reactive-n counting strategy $\mathbf{r}$, is a partner strategy if and only if:

$$r_{n-k} < 1 - \frac{k}{n} \cdot \frac{c}{b}, \text{ for } k \in \{1, 2, \ldots, n\}. \tag{6}$$

Regardless of the memory length, the conditions are the same. The ratio has to be smaller than the fraction of benefit and the tolerance or otherwise, the forgiveness which is measure by a higher probability of cooperation given a defection, decreases as the number of defections by the co-player increase.

3.4 Evolutionary Dynamics

Based on the previous equilibrium analysis, we know where the cooperative reactive-n nash strategies are. We can predict if these strategies will evolve are selected in an evolutionary process. Furthermore, what is not clear as to yet are the effects of the added memory, or even the effects of limiting to counting and no.

In the following, we test these predictions by simulating a simple imitation process based on the dynamics of Imhof and Nowak [Imhof and Nowak, 2010] for reactive-1 reactive-2 and reactive-3 strategies and for counting set. (the setup of these simulations is outlined in Methods).

In the previous sections we characterized partners strategies for reactive-2 and reactive-3, and we have also discussed the case of reactive counting strategies. Here we want to assess whether partner strategies are strategies that evolve, thus are beneficial to adopt in an evolutionary setting.

We then explore the type of strategies that evolve for each set of reactive strategies, Figure **??**. In all cases, the most abundant strategy achieves a high cooperation rate against itself. Notice that all most abundant strategies are the harsher when the co-player defects for the first time after a series of $n - 1$ cooperations. We can observe that in both the case of the two-bits and three-bits, the strategies are more forgiving towards two defections.

We initially test the evolving cooperation rates for different selection strengths, Figure **??**. To this end, we ran simulations for different b/c ratios. As expected, higher b/c values lead to more cooperation in all three spaces, and regardless of β's value. However, the more memory a strategy has it requires a lower benefit-to-cost ratio to achieve substantial cooperation. This verifies that the results of [Hilbe et al., 2017] also hold for reactive strategies.

4 Discussion

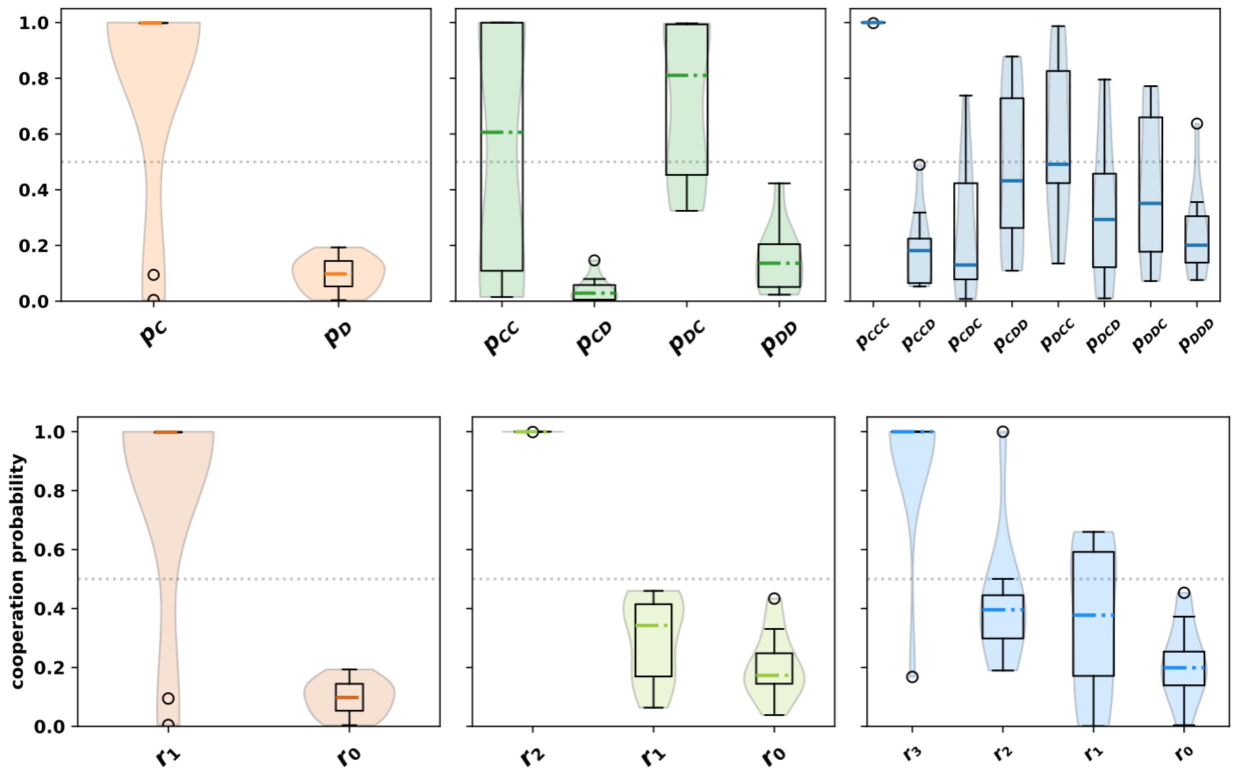

Figure 2: **Evolving strategies for** $n = 1, n = 2$ **and** $n = 3$**.** In the previous sections we characterized partners strategies for reactive-2 and reactive-3, and we have also discussed the case of reactive counting strategies. Here we want to assess whether partner strategies are strategies that evolve, thus are beneficial to adopt in an evolutionary setting. We ran simulations based on Imhof and Nowak. For a single run of the evolutionary process, we record the cooperating probabilities of the resident at each elementary time step. **A. Reactive Strategies.** We ran 10 independent simulations for each set of strategies and recorded the most abundant strategy for each run. The abundant strategy is the resident that was fixed for the most time steps. For the simulations we used $b = 1$ and $c = .5$. For n equal to 1 and 2, $T = 10^7$ and for $n = 3$ then $T = 2 \times 10^7$. **B. Reactive Counting Strategies**. In the top panel we show the most abundant strategies of the evolutionary process when the population can use any counting two-bit reactive strategy (r_0, r_1, r_2). The abundant strategies are the residents that were fixed for the most time steps. The most abundant strategies fall within the region of the partner strategies. In the bottom panel we look at the evolving cooperation rate closer. The average cooperation is calculated by considering the cooperation rate within the resident population. For a given β value we vary the cost $c \in [0, 1]$ whilst we have fixed $b = 1$. Three curves are shown, these are for different values of selection strength, $\beta \in 10^-1, 10^0, 10^1$.

References

S. K. Baek, H.-C. Jeong, C. Hilbe, and M. A. Nowak. Comparing reactive and memory-one strategies of direct reciprocity. *Scientific Reports*, 6(1):1–13, 2016.

C. Hilbe, L. A. Martinez-Vaquero, K. Chatterjee, and M. A. Nowak. Memory-n strategies of direct reciprocity. *Proceedings of the National Academy of Sciences*, 114(18):4715–4720, 2017.

L. A. Imhof and M. A. Nowak. Stochastic evolutionary dynamics of direct reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, 277(1680):463–468, 2010.

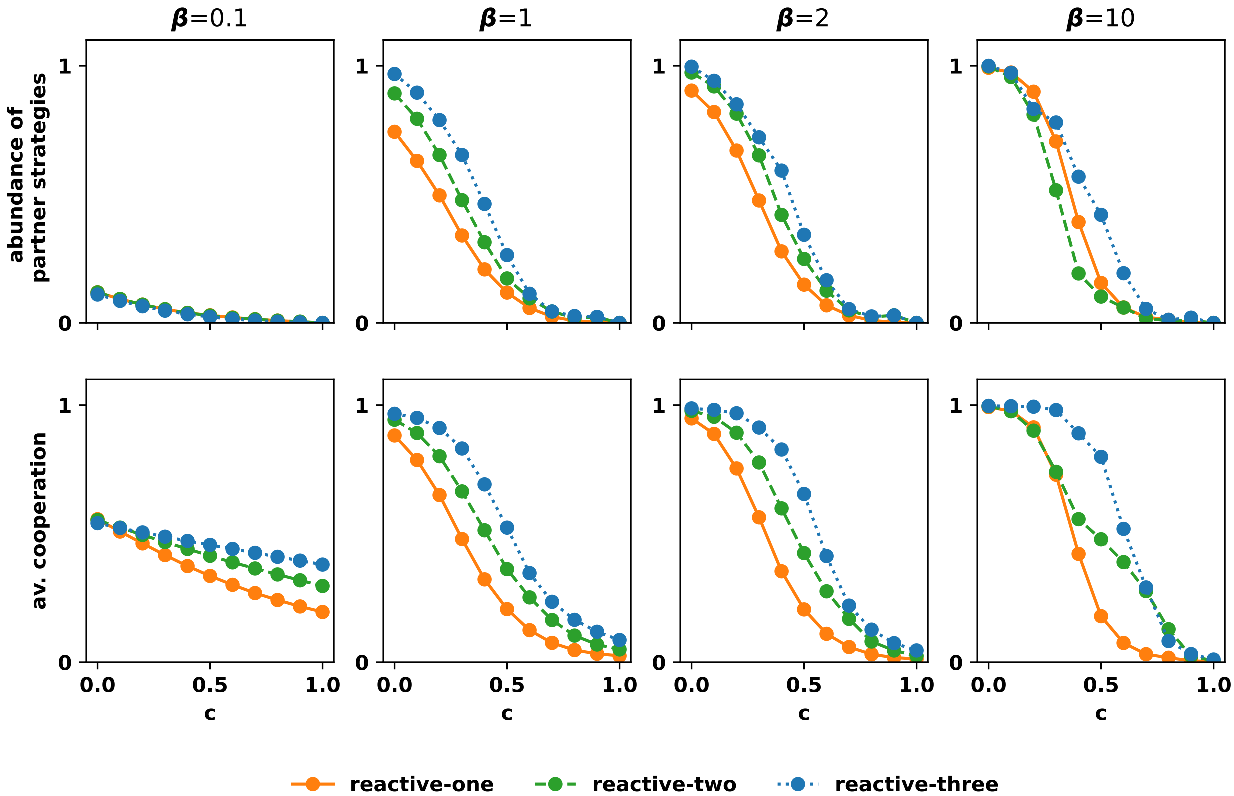

Figure 3: **Abundance of partner strategies.** We ran simulations based on Imhof and Nowak, varying the cost (c) and strength of selection (β). The cooperation benefit (b) is fixed at a value of 1. The results demonstrate that partner strategies evolve notably under strong selection ($\beta = 1$) and lower cost conditions. Furthermore, the abundance of partner strategies is consistently higher when individuals have access to greater memory. The bottom panel displays how these partner strategies lead to increased levels of cooperation.

A. McAvoy and M. A. Nowak. Reactive learning strategies for iterated games. *Proceedings of the Royal Society A*, 475(2223):20180819, 2019.

W. H. Press and F. J. Dyson. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26):10409–10413, 2012.

L. M. Wahl and M. A. Nowak. The continuous prisoner's dilemma: I. linear reactive strategies. *Journal of Theoretical Biology*, 200(3):307–321, 1999.