Memory-n strategies of direct reciprocity

Christian Hilbe[a,1,2], Luis A. Martinez-Vaquero[b,1], Krishnendu Chatterjee[a], and Martin A. Nowak[c,d,e]

[a]Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria; [b]Institute of Cognitive Sciences and Technologies, National Research Council of Italy, 00185 Rome, Italy; [c]Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138; [d]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; and [e]Department of Mathematics, Harvard University, Cambridge, MA 02138

Edited by Kenneth W. Wachter, University of California, Berkeley, CA, and approved March 21, 2017 (received for review December 23, 2016)

Humans routinely use conditionally cooperative strategies when interacting in repeated social dilemmas. They are more likely to cooperate if others cooperated before, and are ready to retaliate if others defected. To capture the emergence of reciprocity, most previous models consider subjects who can only choose from a restricted set of representative strategies, or who react to the outcome of the very last round only. As players memorize more rounds, the dimension of the strategy space increases exponentially. This increasing computational complexity renders simulations for individuals with higher cognitive abilities infeasible, especially if multiplayer interactions are taken into account. Here, we take an axiomatic approach instead. We propose several properties that a robust cooperative strategy for a repeated multiplayer dilemma should have. These properties naturally lead to a unique class of cooperative strategies, which contains the classical Win–Stay Lose–Shift rule as a special case. A comprehensive numerical analysis for the prisoner's dilemma and for the public goods game suggests that strategies of this class readily evolve across various memory-*n* spaces. Our results reveal that successful strategies depend not only on how cooperative others were in the past but also on the respective context of cooperation.

evolutionary game theory | reciprocity | repeated games

In repeated social dilemmas, humans often show conditionally cooperative behaviors (1–3). When there is a temptation to defect at the expense of other group members, subjects consider whether they or others defected before, and react accordingly. However, modeling conditional cooperation is not straightforward, as it is difficult to capture how humans actually make their decisions. Economic models often consider rational subjects who remember all past interactions, and who follow a predefined equilibrium plan (4). Evolutionary models, on the other hand, often take the opposite approach. With a few notable exceptions (5–8), evolutionary models focus on naive subjects who can only choose from a restricted set of strategies (9–13), or who do not remember anything beyond the outcome of the very last round (14–21).

Both approaches represent idealizations, which serve the purpose of making the models computationally tractable. Already for the simplest example, the repeated prisoner's dilemma, calculations are greatly simplified if one assumes that the players' strategies depend on the last round only. These so-called memory-1 strategies represent a four-dimensional space, which can be explored systematically (e.g., refs. 16 and 22). Previous studies identified a number of successful memory-1 strategies, including Tit-for-Tat (TFT) (9), Win–Stay Lose–Shift ($WSLS$) (15, 23) or the class of generous zero-determinant (ZD) strategies (24–28). However, once we allow subjects to remember more than one round, the number of possible strategies increases dramatically. There are 65,536 pure memory-2 strategies and $1.84 \cdot 10^{19}$ pure memory-3 strategies, such that evolutionary simulations soon become incapable of exploring the full strategy space. Obviously, this problem is even further aggravated when we allow for social dilemmas with more than two players.

To identify successful memory-n strategies for multiplayer dilemmas, we thus take an alternative approach. Based on past research on the prisoner's dilemma (9, 15, 23, 29), we suggest three simple principles that stable cooperative strategies should obey. These three principles require that a strategy is (*i*) mutually cooperative, (*ii*) able to correct errors, and (*iii*) sufficiently retaliating against defectors. We show that all strategies that satisfy these requirements belong to a unique class of all-or-none strategies, which includes the well-known $WSLS$ strategy as a special case. All-or-none strategies cooperate against themselves, and they are stable provided the players' memory is sufficiently long. Conversely, we show numerically that, if a pure memory-2 strategy for the prisoner's dilemma with errors is cooperative and stable, then it necessarily exhibits all-or-none behavior. Evolutionary simulations further support these findings for various strategy spaces for the prisoner's dilemma and for the public goods game.

Results

Repeated Dilemmas with Memory-*n* Strategies. We consider a repeated game in a group of m players. In each round, players can either cooperate (C) or defect (D). If there are j cooperators among the other group members, the payoff of a cooperator is A_j and the payoff of a defector is B_j. Herein, we focus on social dilemmas, such that payoffs are assumed to satisfy the following three properties (30, 31): (*i*) Players prefer their coplayers to cooperate, $A_{j+1} \geq A_j$ and $B_{j+1} \geq B_j$ for all j. (*ii*) A defector's payoff always exceeds the payoff of a cooperator, $B_j > A_{j-1}$ for all j. (*iii*) For the whole group, mutual cooperation is beneficial, $A_{m-1} > B_0$. Most of the well-known examples in the literature satisfy these criteria. For example, in a prisoner's dilemma, $m=2$ and payoffs are $A_1 = R$ (the reward for mutual cooperation), $A_0 = S$ (the sucker's payoff), $B_1 = T$ (the temptation to defect), and $B_0 = P$ (the punishment for mutual defection), such that $T > R > P > S$. Herein, we will often use a specific parametrization of the prisoner's dilemma, where cooperation means paying

Significance

Direct reciprocity is one of the fundamental mechanisms for cooperation. It is based on the idea that individuals are more likely to cooperate if they can expect their beneficiaries to remember and to return their cooperative acts in future. Previous computational models, however, often had to restrict the number of past rounds subjects can memorize. Herein we suggest an alternative approach. We propose general properties that robust cooperative strategies ought to have. Then we characterize all memory-*n* strategies that meet these properties, and we show that such strategies naturally emerge across different evolutionary scenarios. Our results are applicable to general social dilemmas of arbitrary size. For some dilemmas, longer memory is all it takes for cooperation to evolve.

Author contributions: C.H., L.A.M.-V., K.C., and M.A.N. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]C.H. and L.A.M.-V. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: christian.hilbe@ist.ac.at.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1621239114/-/DCSupplemental.

EVOLUTION

SOCIAL SCIENCES

a cost $c > 0$ for the coplayer to derive a benefit $b > c$. As a result, the payoffs become $R = b - c$, $S = -c$, $T = b$, and $P = 0$. In *SI Appendix*, we show that our results are robust with respect to other payoff specifications (*SI Appendix*, Fig. S7). Further examples of social dilemmas include the snowdrift game (32), the stag hunt game (33), and the public goods game (13, 34).

For a given group member i, we refer to the player's past n actions as the player's n-history, and we write $h_n^i = (a_{-1}^i, \ldots, a_{-n}^i)$. The elements $a_{-t}^i \in \{C, D\}$ represent the player's action t rounds ago. For example, for a player who always cooperated except for the last round, the respective n-history is given by $(D, C, \ldots, C)$. A game's n-history is the tuple $\mathbf{h}_n = (h_n^1, \ldots, h_n^m)$ that contains the individual histories of all players. For m-player interactions, the set of all possible n-histories H_n contains $|H_n| = 2^{mn}$ different n-histories.

In repeated games, a player's decision to cooperate in one round may depend on the entire history of the game so far. Herein, we assume that players base their decision on the previous n rounds. A memory-n strategy is a vector $\mathbf{p} = (p_{\mathbf{h}})_{\mathbf{h} \in H_n}$. The entries $p_{\mathbf{h}} \in [0, 1]$ give the player's probability of cooperating in the next round if the current n-history is $\mathbf{h} \in H_n$. For $n = 1$ and $m = 2$, this definition of memory-n strategies recovers the typical format of memory-1 strategies for the prisoner's dilemma (22), $\mathbf{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$. We say that a memory-$n$ strategy $\mathbf{p} = (p_{\mathbf{h}})$ is pure if all its entries are either 0 or 1. In contrast, the strategy is stochastic if there is at least one entry $p_{\mathbf{h}}$ for which $0 < p_{\mathbf{h}} < 1$. We assume that the players' actions in any round can be subject to implementation errors [i.e., players have a "trembling hand" (35)]. Specifically, when a player intends to cooperate (defect) in a given round, there is a probability $\varepsilon > 0$ that an error leads the player to defect (cooperate) instead. In the main text, we entirely focus on infinitely repeated games without discounting of the future. In such games, the assumption of occasional errors allows us to ignore the players' actions in the first n rounds (when no complete n-history is yet available). The payoffs of the players can be calculated independent of the initial history of play (for details, see *SI Appendix*).

In the limit of rare errors, a player with some given strategy may never experience certain n-histories. For example, in a prisoner's dilemma where player 1 uses *TFT*, the 2-history $\mathbf{h} = (h^1; h^2) = (CC; CD)$ cannot arise: It would require that the *TFT*-player cooperated in the last round although the coplayer defected the round before. Motivated by this observation, we say that an n-history $\mathbf{h}$ is consistent with respect to the focal player 1's strategy $\mathbf{p}$ if there are strategies $\mathbf{q}^j$ for the remaining group members such that, in a game without errors, the history $\mathbf{h}$ is revisited with positive probability if it was reached once (for formal definitions, see *SI Appendix*). We refer to the set of all consistent n-histories with respect to $\mathbf{p}$ as $\Phi_n(\mathbf{p}) \subseteq H_n$. The consistent n-histories of a strategy $\mathbf{p}$ are those histories that a $\mathbf{p}$-player will generically experience. We say two memory-n strategies $\mathbf{p} = (p_{\mathbf{h}})$ and $\mathbf{q} = (q_{\mathbf{h}})$ are equivalent, and write $\mathbf{p} \sim \mathbf{q}$, if both strategies have the same consistent n-histories $\Phi_n(\mathbf{p}) = \Phi_n(\mathbf{q})$, and if they prescribe the same action for all those consistent n-histories, $p_{\mathbf{h}} = q_{\mathbf{h}}$ for all $\mathbf{h} \in \Phi_n(\mathbf{p})$. As the error rate approaches zero, equivalent strategies $\mathbf{p} \sim \mathbf{q}$ become more and more indistinguishable. No matter which strategies the remaining group members use, the expected response of a $\mathbf{p}$-player will (almost) always coincide with the expected response of a $\mathbf{q}$-player.

Desirable Properties of Memory-*n* Strategies. Within the set of memory-1 strategies for the prisoner's dilemma, evolutionary processes often lead to a particular cooperative strategy, *WSLS* (15, 16). A *WSLS* player only cooperates if both players chose the same action before, $(p_{CC}, p_{CD}, p_{DC}, p_{DD}) = (1, 0, 0, 1)$. *WSLS* has several qualities (22, 23). It is fully cooperative against a player with the same strategy; it is robust with respect to occasional errors; it is immune to invasion by unconditional altruists; and, if the benefit-to-cost ratio satisfies $b/c > 2$, *WSLS* is also stable against defectors. Herein we suggest that the same qualities may also prove useful for memory-n strategies in arbitrary m-player games. In the following, we thus formalize and generalize a few of the properties of *WSLS* (see Fig. 1*A* for a visual

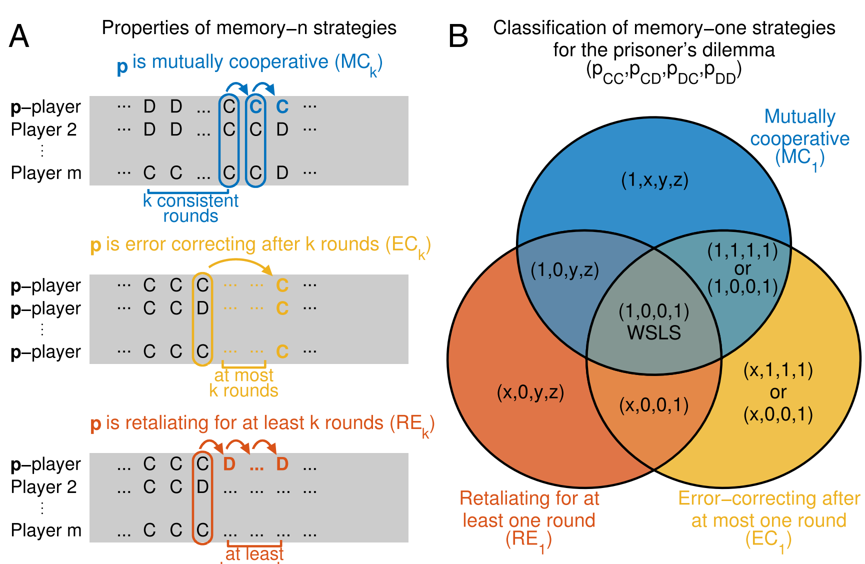

Fig. 1. Illustration of some important properties of memory-*n* strategies. (*A*) We propose several useful properties for memory-*n* strategies: A strategy $\mathbf{p}$ is called mutually cooperative if it cooperates after mutual cooperation for all generic histories. It is called error-correcting after at most k rounds if it takes a homogeneous group of $\mathbf{p}$-players at most k rounds before they revert to mutual cooperation. The strategy is retaliating for at least k rounds if, after any round in which the focal player cooperated whereas a coplayer defected, the focal player defects for the following k rounds. (*B*) Among the memory-1 strategies for the prisoner's dilemma, *WSLS* is the only strategy that satisfies the three properties (MC_1), (EC_1), and (RE_1). Conversely, all three properties are needed to uniquely select *WSLS* among all memory-1 strategies. The parameters x, y, z in the specification of memory-1 strategies can take any value between 0 and 1.

description, and see *SI Appendix* for formal definitions). We will start with the property of mutual cooperativeness.

(MC_k). A strategy is mutually cooperative if there are histories for which the strategy prescribes to cooperate, and if it continues to cooperate after rounds with mutual cooperation (provided the last k actions of the focal player were actually consistent).

The property (MC_k) follows from the general scope of our paper: We are exactly looking for strategies that, in principle, allow for fully cooperative interactions. We do not require such players always to cooperate after mutual cooperation, but they need to do so for all generic histories that are reached with positive probability as the error rate goes to zero. In addition to *WSLS*, there are many other well-known memory-1 strategies that satisfy (MC_k), including *AllC*, *TFT*, or *Grim* (22). However, the latter two strategies have a well-known weakness. They do not cope well with errors: If one player defects by chance, cooperation quickly breaks down. Thus, a natural additional property is the following.

(EC_k). A strategy $\mathbf{p}$ is error-correcting after at most k rounds if, after any history, it generally takes a group of $\mathbf{p}$ players at most $k+1$ rounds to reestablish mutual cooperation.

The property (EC_k) is especially relevant when considering simulations for stochastic strategies. In such simulations, mutations rarely introduce strategies that perfectly satisfy (MC_k). Instead, players may only apply approximations to mutually cooperative strategies. If a cooperative strategy is to be successful, it thus needs to find effective ways to cope with this noise.

When all players use the same (EC_k) strategy, it seems beneficial that k is as small as possible: The smaller k, the sooner players are able to reestablish cooperation after an error. However, a small k can also make a strategy vulnerable: *AllC* reverts to cooperation immediately, and hence satisfies (EC_0), but is easily exploited by defectors. Thus we ask for the following property.

(RE_k). A strategy $\mathbf{p}$ is retaliating for at least k rounds if, after rounds in which the focal player cooperated while at least one coplayer defected, the strategy responds by defecting the following k rounds.

Condition (RE_k) is important for two reasons. First, by construction, such strategies show some robustness against exploitation by defectors. Second, and maybe less obviously, such strategies are also more resistant against invasion by more forgiving strategies (which, in turn, would be susceptible to defectors). For example, in the prisoner's dilemma, a *WSLS* population [which is (RE_1)] cannot be invaded by *AllC* [which is (RE_0)]: If a *WSLS* player defects by error, *AllC* forgives immediately, and the *WSLS* player starts to take advantage of the cooperative opponent (22). An *AllC* player can thus only improve by switching to an (RE_1) strategy as well: The player needs to become less forgiving.

We note that, for the prisoner's dilemma, *WSLS* satisfies all three properties: (MC_1), (EC_1), and (RE_1). Conversely, *WSLS* is the only memory-1 strategy with these properties (Fig. 1*B*). In *SI Appendix*, we show that similar strategies exist for any memory length n and for any group size m. Specifically, for $k \le n$, let us consider the memory-n strategy $\mathbf{p} = (p_{\mathbf{h}})$ with

$$p_{(h_n^1,\ldots,h_n^m)} = \begin{cases} 1 & \text{if } h_k^i = h_k^j \text{ for all players } i,j \\ 0 & \text{if } h_k^i \neq h_k^j \text{ for some players } i,j \end{cases}. \qquad [1]$$

A player with this strategy only cooperates if all players used exactly the same actions in the past, that is, if, in each of the last k rounds, either everyone cooperated or no one did. We refer to this behavior as an all-or-none strategy, or as AON_k. All-or-none behaviors were observed in multiple scenarios: (*i*) The *WSLS* strategy for the prisoner's dilemma corresponds to AON_1. (*ii*) Strategies resembling AON_2 were observed by Hauert and Schuster (6) and by Lindgren (5) when simulating the prisoner's dilemma with memory-2 players. (*iii*) Pinheiro et al. (34) were first to use the name "all-or-none strategy," to describe the strategy AON_1 that emerged in their simulations on public good games among memory-1 players. Our results generalize these previous findings and make them applicable to general multiplayer dilemmas and to players with arbitrary memory.

AON_k strategies satisfy all three properties defined above; conversely, we prove, in *SI Appendix*, that any strategy satisfying the above properties needs to be equivalent to AON_k. Moreover, we show that AON_k is a subgame perfect equilibrium if

$$k \ge \frac{B_{m-1} - A_{m-1}}{A_{m-1} - B_0}. \qquad [2]$$

When condition **2** holds, no mutant strategy can gain a higher payoff, even if mutants were allowed to use stochastic strategies or if they had access to higher memory. For the prisoner's dilemma, this condition simplifies to $b/c \ge (n+1)/n$, which becomes increasingly simple to satisfy as players remember more rounds.

Although subgame perfection guarantees that no single mutant can have a higher payoff, it does not imply that AON_k strategies are evolutionary stable in the sense of Maynard Smith (36). Evolutionary stability requires that any rare mutant strategy is at a disadvantage. In contrast, AON_k populations can be neutrally invaded by strategies that always choose the same action as AON_k, but for a finite number of rounds (in infinitely repeated games without discounting, these finitely many deviations have no negative payoff consequences for the mutant). Previous research found that evolutionary stability is generally difficult to achieve; in repeated games without errors, no strategy is evolutionary stable (37, 38). Moreover, the neutral invasion by one mutant strategy can catalyze the subsequent invasion of other mutant strategies (7, 39). There are two ways that this generic instability can be addressed. On the one hand, it has been shown that evolutionary stability is possible in games with errors when mutants are restricted to strategies of finite complexity (40), or when future payoffs are discounted (41). We thus show, in *SI Appendix*, that AON_k strategies become evolutionary stable in repeated games with discounting (provided that k is sufficiently large and that future payoffs are sufficiently important). On the other hand, one may consider alternative notions of stability. Stewart and Plotkin (27, 42) call a strategy evolutionary robust if a single mutant in a population of size N has a fixation probability of at most $1/N$ (the fixation probability of a neutral mutant). In *SI Appendix*, we provide simulation results suggesting that AON_k can be evolutionary robust even if *AllD* is not (*SI Appendix*, Figs. S2 and S5).

Finally, it is also worth noting that cooperative strategies do not necessarily need to have the form of AON_k to be stable. For example, the condition (RE_k) is rather restrictive. It does not require only that unilateral defection is punished for k rounds but also requires that this punishment begin immediately after the defection occurs. Instead, one may imagine strategies that punish the coplayer with some delay. Players may simply ignore the last round and only react to whatever happened in the second-to-last round. However, in the following section, we provide numerical results suggesting that AON_k strategies and delayed or equivalent versions thereof in fact represent all pure memory-2 strategies that can maintain cooperation in the repeated prisoner's dilemma.

Stability of Pure Memory-2 Strategies with Errors. As an application of our previous results, let us consider memory-2 strategies for the repeated prisoner's dilemma. Depending on the player's own two moves in the previous two rounds, and depending on the two moves of the coplayer, there are $2^4 = 16$ possible 2-histories $\mathbf{h}$. A player with a memory-2 strategy needs to determine whether to cooperate or defect for each possible 2-history, and thus there are $2^{16} = 65,536$ pure memory-2 strategies. However, there are only two all-or-none strategies: AON_2

EVOLUTION

SOCIAL SCIENCES

punishes unilateral defection for two rounds, and reverts to cooperation thereafter. AON_1 punishes defectors only once (and hence coincides with the memory-1 strategy *WSLS*). In addition, one may also consider a delayed version of AON_1, which waits for one round until it punishes unilateral defection. AON_2 is predicted to be stable for $b/c \geq 1.5$, whereas the other two strategies require $b/c \geq 2$.

We performed an exhaustive numerical analysis to identify all strict Nash equilibria if players are restricted to pure memory-2 strategies. To this end, we considered a small error rate of $\varepsilon = 0.01$, and we computed, for each of these strategies, whether any other pure strategy can yield at least the same payoff (see *SI Appendix* for a detailed description of the method). This analysis shows that, for typical benefit-to-cost ratios (with $1 < b/c < 10$), there are 11 Nash equilibria that yield the mutual cooperation payoff against themselves; four of those are equivalent to AON_2, four others are equivalent to AON_1, and the remaining three are delayed versions of AON_1 (Table 1). In particular, all 11 strategies are mutually cooperative, revert to mutual cooperation after at most two rounds, and punish defectors for at least one round (possibly with one round delay).

This numerical approach is not restricted to cooperative strategies; we also used this method to record all other stable memory-2 strategies for the repeated prisoner's dilemma. We find that, besides the class of cooperative strategies, one can distinguish three additional classes of stable behaviors. First, there are Nash equilibria that lead to mutual defection, containing 15 elements, including *AllD* and *Grim* (*SI Appendix*, Table S2). Second, we also identified a class of stable self-alternating strategies, containing eight strategies in total (*SI Appendix*, Table S3). When applied by both players, these self-alternating strategies lead to a deterministic switch between rounds of mutual cooperation and rounds of mutual defection. Finally, the last class of Nash equilibria consists of strategies that have two absorbing states, for example, mutual cooperation and mutual defection (*SI Appendix*, Table S4). When two such players interact, they defect for a large number of rounds; however, after a specific sequence of erroneous moves, players begin to cooperate until cooperation again breaks down due to errors. Thus, the set of memory-2 strategies allows for multiple equilibria that differ in their prospects for cooperation. Which of these equilibria is most relevant may thus depend on how likely they are to emerge in natural evolutionary processes.

Evolutionary Dynamics Among Memory-*n* Players. Based on the previous equilibrium analysis, we may predict the following: (*i*) For intermediate b/c ratios, cooperation should more readily evolve among memory-2 strategies than among memory-1 strategies. (*ii*) If cooperation evolves, it is due to strategies with an all-or-none character (i.e., strategies that are particularly likely to cooperate if players chose the same actions during the previous rounds). In the following, we test these predictions by simulating a simple imitation process based on the dynamics of Imhof and Nowak (17) for stochastic memory-1 and memory-2 strategies (the setup of these simulations is outlined in *Materials and Methods*).

Our simulation results support both predictions. When players are allowed to use memory-2 strategies, the evolving cooperation rates sharply increase once $b/c > 1.5$ [i.e., when AON_2 becomes stable (Fig. 2*A*); as the evolving cooperation rate is a continuous function of the b/c ratio, we do not expect full cooperation when b/c is only slightly above 1.5]. In contrast, when players are restricted to memory-1 strategies, cooperation increases more gradually, and substantial cooperation rates are only achieved when $b/c > 2$. To gain some insights on which strategies are particularly successful, we recorded the most abundant strategy for several independent simulation runs (that is, for each simulation run, we recorded which memory-1 or memory-2 strategy was adopted by the population for the longest time). The most abundant strategies resemble the predicted AON strategies reasonably well (Fig. 2 *B* and *C*). When both players cooperated in all remembered rounds, players are almost certain to cooperate in the next round. Moreover, players are most likely to cooperate if the players' actions in the last rounds coincided, consistent with all-or-none behavior.

Among memory-2 players, AON_2-like strategies seem to be preferred over strategies that resemble AON_1. In particular, although memory-2 players could make use of the classical *WSLS* strategy, the most abundant strategy does not resemble *WSLS* (in Fig. 2*C*, this would require that the first four bars are close to 1, because a *WSLS* player always cooperates if both players cooperated in the previous round). These findings suggest that memory-2 players consider the full 2-history of play when deciding whether to cooperate in the next round.

Fig 2 *B* and *C* also suggests that the evolving memory-1 strategies yield somewhat better approximations to AON behavior than the evolving memory-2 strategies. This may be partly due to the differential selection pressure on each of the $p_{\mathbf{h}}$ values. All 1-histories $\mathbf{h}$ are consistent with respect to AON_1, and hence an AON_1 player will frequently be challenged to give an optimal response to each possible 1-history. This selection pressure is somewhat damped for AON_2 strategies, because each individual 2-history is less likely to occur over the course of a game (in fact, some of the 2-histories will typically only occur after a rather particular sequence of errors).

To further corroborate our theoretical predictions, we show, in *SI Appendix*, that all-or-none strategies are also predominant when we simulate the dynamics among memory-1 strategies for

Table 1. Stable and cooperative memory-2 strategies for the prisoner's dilemma

Strategy description	Previous two rounds: player 1, player 2																Cooperation rate against itself	Minimum *b/c* ratio
	CC, CC	CC, CD	CD, CC	CD, CD	CC, DC	CC, DD	CD, DC	CD, DD	DC, CC	DC, CD	DD, CC	DD, CD	DC, DC	DC, DD	DD, DC	DD, DD		
Strategies equivalent to AON_2	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0.952	1.526
	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0.951	1.526
	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0.951	1.526
	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	1	0.952	1.526
Strategies equivalent to AON_1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0.971	2.041
	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	0.971	2.041
	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0.970	2.041
	1	0	0	1	0	0	0	0	0	0	0	0	1	1	1	1	0.971	2.041
Delayed versions of AON_1	1	0	0	1	0	0	0	1	0	0	0	1	1	0	0	1	0.952	2.083
	1	0	0	1	1	0	0	0	1	0	0	0	1	0	0	1	0.970	2.021
	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	0.971	2.041

All strategies prescribe to cooperate if (*i*) both players cooperated in both rounds, (*ii*) both players defected in both rounds, and (*iii*) both players cooperated in the last round, but defected in the second-to-last round. Moreover, all strategies defect if the actions of the two players were different from each other in both previous rounds. The last column gives the threshold that the b/c ratio needs to exceed for the respective strategy to be a strict Nash equilibrium among the finite set of pure memory-2 strategies. The table suggests the existence of a tradeoff: Strategies equivalent to AON_2 require a lower b/c ratio to be stable, but they also have a slightly lower cooperation rate against themselves.

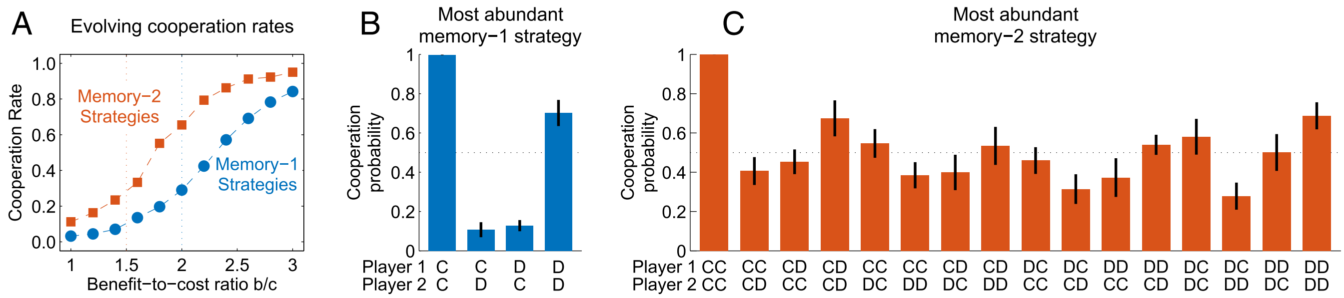

Fig. 2. Comparing the evolving cooperation rates and the most abundant strategies for memory-2 and memory-1 strategies. (*A*) To assess the impact of memory on the evolution of cooperation, we ran simulations based on ref. 17 for different benefit-to-cost ratios. As expected, higher b/c values lead to more cooperation in both strategy spaces. Memory-2 players, however, require a lower benefit-to-cost ratio to achieve substantial cooperation. (*B* and *C*) Using a benefit $b=3$, we ran 10 independent simulations and recorded the most abundant strategy for each run. Colored bars show average values, and error bars represent the SE across the 10 simulation runs. In both cases, the most abundant strategy achieves a high cooperation rate against itself. Players almost certainly cooperate if both players were cooperative in all remembered rounds. Moreover, players are more likely to cooperate if the players' decisions in the past coincided. Parameters are as follows: population size $N=100$, cost $c=1$, strength of selection $s=1$, and error rate $\varepsilon=0.001$; each simulation was run for 10^7 mutant strategies.

the public goods game (see *SI Appendix*, Fig. S4, using a strictly larger strategy set than in refs. 13 and 34). Similarly, we show that behaviors reminiscent of all-or-none strategies evolve in the prisoner's dilemma when players only remember how often (but not when) players cooperated during the previous $n \leq 3$ rounds (*SI Appendix*, Fig. S3). Among these simplified memory-n strategies, the most abundant strategies tend to cooperate if both players were equally cooperative in the previous n rounds, and they defect otherwise. However, higher memory no longer leads to substantially higher cooperation rates (*SI Appendix*, Fig. S3*A*). These results suggest memory is not only important to assess how cooperative other group members have been but is also an important mechanism to reach coordination among like-minded players. Such coordination attempts are most successful if players remember both the degree of cooperation and its timing.

Discussion

Previous research often used tournaments and evolutionary contests to distill properties that successful reciprocal strategies ought to have. Herein, we take the converse approach. We formulated three simple principles that we can expect well-performing strategies to obey. Based on these principles, we derived a successful class of cooperative strategies for general multiplayer dilemmas. Each of our principles seems to be psychologically intuitive. Our first principle of mutual cooperativeness is motivated by the observation that most subjects are most likely to cooperate in fully cooperative groups (3, 43). The principle of retaliation is based on findings that people are willing to fight back when being exploited, sometimes even if this comes at a cost to themselves (44, 45). Our last principle acknowledges that mutual cooperation can only be sustained in noisy environments if we are able to forgive others for their occasional failures to cooperate (46). The importance of these principles was noted before (9, 22, 23, 29, 47). However, the application of these principles has been usually limited to specific two-player games, assuming that players are subject to rather severe constraints on their cognitive capabilities. Here we show that the above principles can be used to construct strategies that can be applied in any multiplayer dilemma, and where subjects may remember an arbitrary number of past events.

Similar to the well-known *WSLS* rule in the prisoner's dilemma, all-or-none strategies AON_k follow a Pavlovian pattern. If players obtained different payoffs in the last round, AON_k players will repeat a successful action for the next k rounds (if they received the higher payoff of a defector within the mixed group), or they will abandon their inferior action for the next k rounds (if they received the lower payoff of a cooperator). However, if all group members obtained the same payoff in each of the previous k rounds, AON_k players cooperate. AON_k strategies are self-synchronous: Independent of the previous history, all members of a group of all-or-none players will choose the same action in every round. Such self-synchronicity can greatly facilitate mutual coordination toward the cooperative equilibrium once the other players' similarity is perceived (48).

Higher memory is not necessary if a strategy only needs to resist invasion by defectors. As an example, let us consider the generous ZD strategy (24–26, 28, 42) that always reciprocates cooperation, but forgives the coplayer's defection in the prisoner's dilemma with probability $1/(k+1)$. A player with such a strategy retaliates against *AllD* for an expected number of k rounds, and hence the generous ZD strategy is stable against *AllD* if $b/c>(k+1)/k$, just like the AON_k strategy. However, the generous ZD strategy is susceptible to indirect invasions: Unlike AON_k, it can be easily subverted by unconditional cooperators, who, in turn, promote the emergence of defectors (7, 39). In line with this observation, the most successful strategies in our simulations indeed made use of their full memory capabilities; players did not rely on cooperative memory-1 strategies when they had access to higher memory strategies.

Overall, our study suggests that memory-n strategies are particularly valuable when the benefit of cooperation is small or intermediate (for $b/c \geq 2$, already, the simple memory-1 strategy *WSLS* allows for full cooperation). In such cases, we may well expect to observe selection for longer memory as suggested by previous simulations (8), provided that the expected gains are worth the higher cognitive costs (49). In principle, our results suggest that cooperation is feasible in any multiplayer dilemma, provided that the interaction is sufficiently relevant for subjects to memorize their coplayers' past actions.

Materials and Methods

In the following paragraphs, we describe the setup of our evolutionary process. Our evolutionary simulations are based on a simple pairwise imitation process, based on the dynamics described in ref. 17. We consider a population of size N. Initially, all members are unconditional defectors. In each elementary time step, one individual experiments with a new mutant strategy. This mutant strategy $\mathbf{q}=(q_\mathbf{h})$ is generated by randomly drawing 2^{2n} cooperation probabilities $q_\mathbf{h}$ from the unit interval [0,1]. If the mutant strategy yields a payoff of $\pi_M(j)$, where j is the number of mutants in the population, and if residents get a payoff of $\pi_R(j)$, then the fixation probability f_M of the mutant strategy can be calculated explicitly (50),

$$f_M=\left(1+\sum_{i=1}^{N-1}\prod_{j=1}^{i}\exp\left\{-s[\pi_M(j)-\pi_R(j)]\right\}\right)^{-1}. \quad [3]$$

The parameter $s \geq 0$ is called the strength of selection. It measures how important relative payoff advantages are for the evolutionary success of a strategy. When s is small, $s \approx 0$, payoffs become irrelevant, and the strategy's fixation probability approaches $f_M \approx 1/N$. The larger the value of s, the more strongly the evolutionary process favors the fixation of strategies that yield high payoffs. Once the mutant strategy either reaches fixation or goes to extinction, another mutant strategy is introduced to the resident population. We iterated this elementary population updating process for 10^7 mutant strategies per simulation run. This process provides a reasonable approximation on the dynamics among memory-n strategists when mutations are relatively rare (51, 52). In *SI Appendix*, however, we present further simulations suggesting that our qualitative results are independent of the assumption of rare mutations, and of the considered selection strength (*SI Appendix*, Fig. S6).

ACKNOWLEDGMENTS. This work was supported by the European Research Council Start Grant 279307: Graph Games (to K.C.), Austrian Science Fund (FWF) Grant P23499-N23 (to K.C.), FWF Nationale Forschungsnetzwerke Grant S11407-N23 Rigorous Systems Engineering/Systematic Methods in Systems Engineering (to K.C.), Office of Naval Research Grant N00014-16-1-2914 (to M.A.N.), and the John Templeton Foundation (M.A.N.). The Program for Evolutionary Dynamics is supported, in part, by a gift from B. Wu and Eric Larson. C.H. acknowledges generous support from the ISTFELLOW program, and L.A.M.-V. gratefully acknowledges support from the European Research Consortium for Informatics and Mathematics Alain Bensoussan Fellowship Program.

1. Fischbacher U, Gächter S (2010) Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *Am Econ Rev* 100:541–556.
2. Rand DG, Nowak MA (2012) Human cooperation. *Trends Cognit Sci* 117:413–425.
3. Grujic J, et al. (2014) A comparative analysis of spatial prisoner's dilemma experiments: Conditional cooperation and payoff irrelevance. *Sci Rep* 4:4615.
4. Fudenberg D, Tirole J (1998) *Game Theory* (MIT Press, Cambridge, MA), 6th Ed.
5. Lindgren K (1997) *Evolutionary Dynamics in Game-Theoretic Models in the Economy as an Evolving Complex System II*, eds Arthur WB, Durlauf SN, Lane DA (Addison-Wesley, Upper Saddle River, NJ).
6. Hauert C, Schuster HG (1997) Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: A numerical approach. *Proc R Soc B* 264: 513–519.
7. van Veelen M, García J, Rand DG, Nowak MA (2012) Direct reciprocity in structured populations. *Proc Natl Acad Sci USA* 109:9929–9934.
8. Stewart AJ, Plotkin JB (2016) Small groups and long memories promote cooperation. *Sci Rep* 6:26889.
9. Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science* 211:1390–1396.
10. Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428:646–650.
11. Kurokawa S, Ihara Y (2009) Emergence of cooperation in public goods games. *Proc R Soc B* 276:1379–1384.
12. Szolnoki A, Perc M, Szabó G (2009) Phase diagrams for three-strategy evolutionary prisoner's dilemma games on regular graphs. *Phys Rev E* 80:056104.
13. van Segbroeck S, Pacheco JM, Lenaerts T, Santos FC (2012) Emergence of fairness in repeated group interactions. *Phys Rev Lett* 108:158104.
14. Nowak MA, Sigmund K (1992) Tit for tat in heterogeneous populations. *Nature* 355:250–253.
15. Nowak MA, Sigmund K (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature* 364:56–58.
16. Kraines DP, Kraines VY (2000) Natural selection of memory-one strategies for the iterated prisoner's dilemma. *J Theor Biol* 203:335–355.
17. Imhof LA, Nowak MA (2010) Stochastic evolutionary dynamics of direct reciprocity. *Proc R Soc B* 277:463–468.
18. Martinez-Vaquero LA, Cuesta JA, Sanchez A (2012) Generosity pays in the presence of direct reciprocity: A comprehensive study of 2x2 repeated games. *PLoS One* 7:E35135.
19. Hilbe C, Nowak MA, Sigmund K (2013) The evolution of extortion in iterated prisoner's dilemma games. *Proc Natl Acad Sci USA* 110:6913–6918.
20. Dong Y, Li C, Tao Y, Zhang B (2015) Evolution of conformity in social dilemmas. *PLoS One* 10:e0137435.
21. Baek SK, Jeong HC, Hilbe C, Nowak MA (2016) Comparing reactive and memory-one strategies of direct reciprocity. *Sci Rep* 6:25676.
22. Sigmund K (2010) *The Calculus of Selfishness* (Princeton Univ Press, Princeton, NJ).
23. Kraines DP, Kraines VY (1993) Learning to cooperate with Pavlov an adaptive strategy for the iterated prisoner's dilemma with noise. *Theory Decis* 35:107–150.
24. Press WH, Dyson FD (2012) Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc Natl Acad Sci USA* 109:10409–10413.
25. Stewart AJ, Plotkin JB (2012) Extortion and cooperation in the prisoner's dilemma. *Proc Natl Acad Sci USA* 109:10134–10135.
26. Akin E (2013) Stable cooperative solutions for the iterated prisoner's dilemma. arXiv:1211.0969v2.
27. Stewart AJ, Plotkin JB (2013) From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proc Natl Acad Sci USA* 110:15348–15353.
28. Hilbe C, Traulsen A, Sigmund K (2015) Partners or rivals? Strategies for the iterated prisoner's dilemma. *Games Econ Behav* 92:41–52.
29. Neill DB (2001) Optimality under noise: Higher memory strategies for the alternating prisoner's dilemma. *J Theor Biol* 211:159–180.
30. Kerr B, Godfrey-Smith P, Feldman MW (2004) What is altruism? *Trends Ecol Evol* 19:135–140.
31. Hilbe C, Wu B, Traulsen A, Nowak MA (2014) Cooperation and control in multiplayer social dilemmas. *Proc Natl Acad Sci USA* 111:16425–16430.
32. Doebeli M, Hauert C (2005) Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecol Lett* 8:748–766.
33. Pacheco JM, Santos FC, Souza MO, Skyrms B (2009) Evolutionary dynamics of collective action in n-person stag hunt dilemmas. *Proc R Soc B* 276:315–321.
34. Pinheiro FL, Vasconcelos VV, Santos FC, Pacheco JM (2014) Evolution of all-or-none strategies in repeated public goods dilemmas. *PLoS Comput Biol* 10:e1003945.
35. Selten R (1975) Reexamination of the perfectness concept for equilibrium points in extensive games. *Int J Game Theor* 4:25–55.
36. Maynard Smith J (1974) Theory of games and the evolution of animal contests. *J Theor Biol* 47:209–221.
37. Boyd R, Lorberbaum J (1987) No pure strategy is evolutionary stable in the iterated prisoner's dilemma game. *Nature* 327:58–59.
38. Farrell J, Ware R (1989) Evolutionary stability in the repeated prisoner's dilemma. *Theor Popul Biol* 36:161–166.
39. Garcia J, van Veelen M (2016) In and out of equilibrium I: Evolution of strategies in repeated games with discounting. *J Econ Theor* 161:161–189.
40. Fudenberg D, Maskin E (1990) Evolution and cooperation in noisy repeated games. *Am Econ Rev* 80:274–279.
41. Boyd R (1989) Mistakes allow evolutionary stability in the repeated Prisoner's Dilemma game. *J Theor Biol* 136:47–56.
42. Stewart AJ, Plotkin JB (2014) Collapse of cooperation in evolving games. *Proc Natl Acad Sci USA* 111:17558–17563.
43. Traulsen A, Semmann D, Sommerfeld RD, Krambeck HJ, Milinski M (2010) Human strategy updating in evolutionary games. *Proc Natl Acad Sci USA* 107:2962–2966.
44. Sigmund K (2007) Punish or perish? Retaliation and collaboration among humans. *Trends Ecol Evol* 22:593–600.
45. Dreber A, Rand DG, Fudenberg D, Nowak MA (2008) Winners don't punish. *Nature* 452:348–351.
46. Fudenberg D, Dreber A, Rand DG (2012) Slow to anger and fast to forgive: Cooperation in an uncertain world. *Am Econ Rev* 102:720–749.
47. Molander P (1985) The optimal level of generosity in a selfish, uncertain environment. *J Conflict Resolut* 29:611–618.
48. Fischer I, et al. (2013) Fusing enacted and expected mimicry generates a winning strategy that promotes the evolution of cooperation. *Proc Natl Acad Sci USA* 110:10229–10233.
49. Milinski M, Wedekind C (1998) Working memory constrains human cooperation in the prisoner's dilemma. *Proc Natl Acad Sci USA* 95:13755–13758.
50. Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. *Phys Rev E* 74:011909.
51. Fudenberg D, Imhof LA (2006) Imitation processes with small mutations. *J Econ Theor* 131:251–262.
52. Wu B, Gokhale CS, Wang L, Traulsen A (2012) How small are small mutation rates? *J Math Biol* 64:803–827.