

# Friendly-rivalry solution to the iterated $n$ -person public-goods game

Yohsuke Murase<sup>1</sup> and Seung Ki Baek<sup>2\*</sup>

**1** RIKEN Center for Computational Science, Kobe, Hyogo 650-0047, Japan

**2** Department of Physics, Pukyong National University, Busan 48513, Korea

\* seungki@pknu.ac.kr

## Abstract

Repeated interaction promotes cooperation among rational individuals under the shadow of future, but it is hard to maintain cooperation when a large number of error-prone individuals are involved. One way to construct a cooperative Nash equilibrium is to find a ‘friendly rivalry’ strategy, which aims at full cooperation but never allows the co-players to be better off. Recently it has been shown that for the iterated Prisoner’s Dilemma in the presence of error, a friendly rival can be designed with the following five rules: Cooperate if everyone did, accept punishment for your own mistake, punish defection, recover cooperation if you find a chance, and defect in all the other circumstances. In this work, we construct such a friendly-rivalry strategy for the iterated  $n$ -person public-goods game by generalizing those five rules. The resulting strategy makes a decision with referring to the previous  $m = 2n - 1$  rounds. A friendly-rivalry strategy inherently has evolutionary robustness in the sense that no mutant strategy has higher fixation probability in this population than that of neutral drift, and our evolutionary simulation indeed shows excellent performance of the proposed strategy in a broad range of environmental conditions.

## Author summary

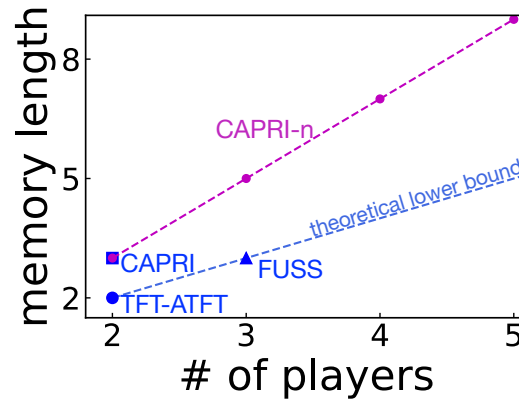
How to maintain cooperation among a number of self-interested individuals is a difficult problem, especially if they can sometimes commit error. In this work, we propose a strategy for the iterated  $n$ -person public-goods game based on the following five rules: Cooperate if everyone did, accept punishment for your own mistake, punish others’ defection, recover cooperation if you find a chance, and defect in all the other circumstances. These rules are not far from actual human behavior, and the resulting strategy guarantees three advantages: First, if everyone uses it, full cooperation is recovered even if error occurs with small probability. Second, the player of this strategy always never obtains a lower long-term payoff than any of the co-players. Third, if the co-players are unconditional cooperators, it obtains a strictly higher long-term payoff than theirs. Therefore, if everyone uses this strategy, no one has a reason to change it. Furthermore, our simulation shows that this strategy will become highly abundant over long time scales due to its robustness against the invasion of other strategies. In this sense, the repeated social dilemma is solved for an arbitrary number of players.

# Introduction

The success of *Homo sapiens* can be attributed to its ability to organize collective action toward a common goal among a group of genetically unrelated individuals [1], and this ability is becoming more and more important as the world is getting close to each other. Researchers have identified several mechanisms to promote cooperation in terms of evolutionary game theory [2,3]. For example, the folk theorem holds that repeated interaction can establish cooperation through reciprocal strategies, and this mechanism is called direct reciprocity [4]. Yet, how to resolve a conflict between individual and collective interests is a hard problem, especially when a large number of players are involved and they are prone to error [5–7], because an individual player has very limited control over co-players.

In this respect, the discovery of the zero-determinant (ZD) strategies in the iterated prisoner’s dilemma (PD) has been deemed counter-intuitive because a ZD-strategic player can unilaterally fix the co-player’s long-term payoff or enforce a linear relationship between their long-term payoffs [8]. For instance, one can design an *extortionate* ZD strategy, with which the player’s long-term payoff will increase by  $\chi \geq 1$  whenever the co-player’s does by one unit payoff. Another counter-intuitive aspect of the ZD strategy is that it is a memory-one strategy referring only to the previous round, so that such a simple strategy can perfectly constrain the co-player’s long-term payoff regardless of the co-player’s strategic complexity. Of course, the excellent performance in a one-to-one match does not necessarily mean evolutionary success: It is difficult for an extortionate strategy to proliferate in a population because, as its fraction increases, two extortionate players are more likely to meet and keep defecting against each other [9–12]. For this reason, especially in a large population, selection tends to favor a *generous* ZD strategy whose long-term payoff does not exceed the co-player’s [11]. A generous ZD strategy does not aim at winning a match, but it is efficient by forming mutual cooperation when they meet each other.

The important point in this line of thought is that a player’s strategy can unilaterally impose constraints on the co-player’s long-term payoff, so that we can now characterize strategies according to the constraints that they impose. One such meaningful characterization scheme is to ask if a strategy works as a ‘partner’ or as a ‘rival’ [13,14]: By ‘partner’, we mean that the strategy seeks for mutual cooperation, but that it will make the co-player’s payoff less than its own if the co-player defects from it. It has also been called ‘good’ [15,16], and the generous ZD strategies can be understood as an intersection between the ZD and partner strategies [11]. On the other hand, a rival strategy always makes its long-term payoff higher than or equal to the co-player’s, so it has been called ‘unbeatable’ [17], ‘competitive’ [13], or ‘defensible’ [18,19]. A trivial example of a rival strategy is unconditional defection (AllD), and an extortionate ZD strategy also falls into this class. Most of well-known strategies in the iterated PD game are classified either as a partner or as a rival [14]. However, which class is more favored by selection depends on environmental conditions such as the population size and the benefit-to-cost ratio of cooperation: If the population is small and cooperation is costly, it is better off to play a rival strategy than to play a partner strategy, and vice versa [11,14,20]. If a single strategy acts as a partner and a rival simultaneously, it has important implications in evolutionary dynamics because it possesses *evolutionary robustness* regardless of the environmental conditions, in the sense that no mutant strategy can invade a population of this strategy with greater fixation probability than that of neutral drift [11,20–22]. To indicate the partner-rival duality, such a strategy will be called a ‘friendly rival’ [22]. Tit-for-tat (TFT), a special ZD strategy having  $\chi = 1$ , is a friendly rival in an error-free environment [14], but a friendly rival generally requires a far more complicated structure in the presence of error. So far, the existence of friendly-rivalry strategies has been reported by a brute-force enumeration method in



**Fig 1.** Memory length  $m$  required for each of currently known friendly-rivalry strategies in the  $n$ -person PG game [18, 19, 22]. The dashed blue line depicts a theoretical lower bound  $m = n$  for friendly rivalry [19], and the strategy proposed in this work, called CAPRI- $n$ , has  $m = 2n - 1$ .

the iterated PD game [18, 22, 23] and the three-person public-goods (PG) game [19]. However, it is not straightforward to extend these findings to the general  $n$ -person PG game. For example, a naive extension of a solution in the iterated PD game fails to solve the three-person PG game because the third player cannot tell if one of the co-players is correcting the other's error with good intent or just carrying out a malicious attack [19]. To resolve the ambiguity, a strategic decision must be based on more information of the past interactions: In fact, if a player refers to the previous  $m$  rounds to choose an action in the  $n$ -person PG game, we can show that  $m$  must be greater than or equal to  $n$  as a necessary condition to be a friendly rival [19]. Unfortunately, the existing brute-force approach then becomes simply unfeasible because the number of possible strategies expands super-exponentially as  $2^{2^{mn}}$ . For example, in the three-person game ( $n = 3$ ), it means that we have to enumerate  $2^{512} \sim 10^{154}$  possibilities to find an answer. Although the symmetry among co-players reduces this number down to  $2^{288} \sim 10^{86}$ , it is still comparable to the estimated number of protons in the universe.

In this work, by using an alternative method to generalize behavioral patterns of a friendly rival for the iterated PD game [22], we construct a friendly-rivalry strategy for the  $n$ -person PG game. This approach makes use of the fact that it greatly lessens the computational burden if we only check whether a given strategy qualifies as a friendly rival. The required memory length of our strategy is  $m = 2n - 1$ , which satisfies the necessary condition  $m \geq n$  as shown in Fig. 1. We will also numerically confirm that it shows excellent performance in evolutionary dynamics due to its evolutionary robustness.

## Materials and methods

### Public-goods game

Let us consider the  $n$ -person public-goods (PG) game, in which a player may choose either cooperation ( $c$ ), by contributing a token to a public pool, or defection ( $d$ ), by refusing it. Let the number of cooperators be denoted as  $n_c$ . The  $n_c$  tokens in the public pool are multiplied by a factor of  $\rho$ , where  $1 < \rho < n$ , and then equally redistributed to the  $n$  players. We assume that the tokens are infinitely divisible. A

player's payoff is thus given as

$$\begin{cases} \frac{\rho n_c}{n} & \text{when the player chooses } c, \\ 1 + \frac{\rho n_c}{n} & \text{when the player chooses } d. \end{cases} \quad (1)$$

Clearly, it is always better off to choose  $d$  regardless of  $n_c$ , so full defection is the only Nash equilibrium of this one-shot game. In this study, this game will be repeated indefinitely with no discounting factor to facilitate direct reciprocity. Every player can choose an action between  $c$  and  $d$  by referring to the previous  $m$  rounds. At the same time, a player can make implementation error, e.g., by choosing  $d$  while intending  $c$  and vice versa, with small probability  $e \ll 1$ .

## Axiomatic approach

Let us consider a strategy profile  $\mathcal{P} = \{\Sigma_1, \Sigma_2, \dots, \Sigma_n\}$  of  $n$  players. Player  $X$ 's long-term payoff is defined as

$$\Pi_X \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \pi_X^{(t)}, \quad (2)$$

where  $\pi_X^{(t)}$  is player  $X$ 's instantaneous payoff in round  $t$ . If  $e > 0$ , the Markovian dynamics of strategic interaction for a given strategy profile  $\mathcal{P}$  converges to a unique stationary distribution, from which  $\Pi_X$  can readily be calculated [24, 25]. In terms of the players' long-term payoffs, we wish to propose the following three criteria that a successful strategy  $\Omega$  should satisfy [18, 19, 22, 26].

1. Efficiency: Mutual cooperation must be achieved with probability one as  $e \rightarrow 0$  if all the players have adopted  $\Omega$ . In other words, this criterion requires  $\lim_{e \rightarrow 0^+} \Pi_X = \rho$  when  $\mathcal{P} = \mathcal{P}_\Omega \equiv \{\Omega, \Omega, \dots, \Omega\}$ .
2. Defensibility: It must be guaranteed that none of the co-players can obtain higher long-term payoffs against  $\Omega$  regardless of their strategies and initial states when  $e = 0$ . It implies that  $\lim_{e \rightarrow 0^+} (\Pi_X - \Pi_Y) \geq 0$ , where player  $X$  is using strategy  $\Sigma_X = \Omega$  and  $Y$  denotes any possible co-player of  $X$ .
3. Distinguishability: If  $\Sigma_X = \Omega$  and all the co-players are unconditional cooperators (AllC), player  $X$  can exploit them to earn a strictly higher long-term payoff than theirs. That is,  $\Pi_X > \Pi_Y$  when  $Y$  is an AllC player.

When a strategy satisfies defensibility and efficiency, the strategy is a friendly rival. A symmetric strategy profile which consists of a friendly-rivalry strategy forms a cooperative Nash equilibrium [18, 19, 22]. The third criterion is a requirement to suppress invasion of AllC due to neutral drift in the evolutionary context [27–29]. We call a strategy 'successful' if it meets all the above three criteria. Depending on the definition of successfulness, one could choose a different set of axioms for an alternative characterization [30].

## Strategy design

Let us construct a deterministic strategy with memory length  $m = 2n - 1$  and show that the proposed strategy indeed meets all of the above three criteria. In the following, we will take an example of three players ( $n = 3$ ) who are called Alice ( $A$ ), Bob ( $B$ ), and Charlie ( $C$ ), respectively, and choose Alice as a focal player playing this strategy.

Before proceeding, it is convenient to introduce some notations for the sake of brevity. The three players' history profile over the previous  $m = 5$  rounds can be represented as  $h_t = (A_{t-5}A_{t-4}A_{t-3}A_{t-2}A_{t-1}; B_{t-5}B_{t-4}B_{t-3}B_{t-2}B_{t-1}; C_{t-5}C_{t-4}C_{t-3}C_{t-2}C_{t-1})$ , where  $A_\tau$ ,  $B_\tau$ , and  $C_\tau$  denote their respective actions at round  $\tau$ . The last round of full cooperation will be denoted by  $t^*$ . In addition, we introduce a binary variable  $\lambda_X^{(t)}$  which equals one if  $X_t = d$  and zero otherwise for player  $X \in \{A, B, C\}$ . According to the payoff matrix [Eq. (1)], Alice's payoff in round  $t$  can be rewritten as

$$\pi_A^{(t)} = \rho \left[ 1 - \frac{1}{n} \sum_X \lambda_X^{(t)} \right] - [1 - \lambda_A^{(t)}], \quad (3)$$

which has linear dependence on  $\lambda_X^{(t)}$  for every  $X$ . This linearity implies that Alice's total payoff  $\sum_t \pi_A^{(t)}$  in the iterated game is fully determined by counting every player's total number of defections, i.e.,  $\sum_t \lambda_X^{(t)}$  for every  $X$ : For example, if all the players have defected the same number of times, their payoffs must be the same irrespective of the exact history. Let  $\Delta_A^{\tau_1, \tau_2}$  thus denote Alice's number of defections in  $[\tau_1, \tau_2]$ . Likewise, we can define  $\Delta_B^{\tau_1, \tau_2}$  for Bob and  $\Delta_C^{\tau_1, \tau_2}$  for Charlie. We also define  $N_d$  as the maximum difference among the players in numbers of defections over the previous  $m$  rounds:

$$N_d \equiv \max_{i,j \in \{A,B,C\}} \left| \Delta_i^{t-m, t-1} - \Delta_j^{t-m, t-1} \right|. \quad (4)$$

With these notations, we can now design a successful strategy satisfying all the three criteria simultaneously. To this end, we divide the set of history profiles into three mutually exclusive cases: The first case is that full cooperation occurred in the last round ( $t^* = t - 1$ ). The second case is that it is not in the last round but still in their memory ( $t - m \leq t^* < t - 1$ ). The third case is that no player remembers the last round of full cooperation ( $t^* < t - m$ ). Let us consider these cases one by one, together with adequate rules for each.

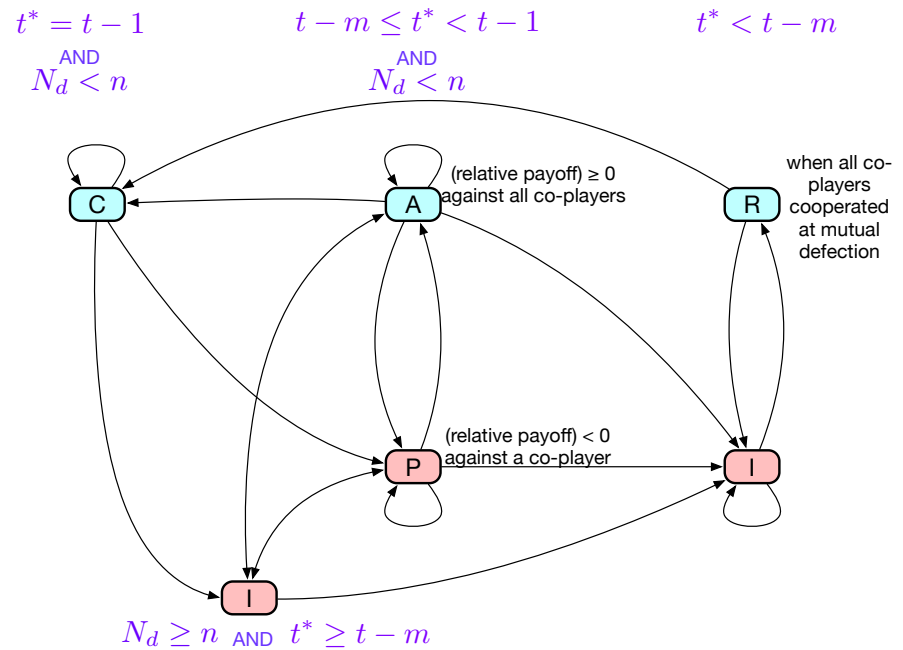
1.  $t^* = t - 1$

- Cooperate: If this is the case, Alice has to choose  $c$  under the condition that  $N_d < n$ . For example, the inequality is true for  $(ccccc; cccdc; ccccc)$ , for which  $N_d = 1$ . On the other hand, it is not true for  $(cdddc; cdddc; ccccc)$  because its  $N_d = 3$  is equal to  $n$ .

2.  $t - m \leq t^* < t - 1$

- Accept: Alice has to accept punishment from the co-players by choosing  $c$ , under the condition that  $\Delta_A^{t^*, t-1} \geq \Delta_B^{t^*, t-1}$  and  $\Delta_A^{t^*, t-1} \geq \Delta_C^{t^*, t-1}$  in addition to  $N_d < n$ . For example,  $c$  will be prescribed to Alice at  $(cccdc; ccccd, ccccc)$ , where we have  $t^* = t - 3$ ,  $\Delta_A^{t^*, t-1} = \Delta_B^{t^*, t-1} = 1$ ,  $\Delta_C^{t^*, t-1} = 0$ , and  $N_d = 1$ , which satisfies the above inequalities. On the other hand, the condition is not met by  $(cdddd; cdddd; ccccc)$  which gives  $N_d = 3$ .
- Punish: Alice has to punish the co-players by choosing  $d$ , under the condition that  $\Delta_A^{t^*, t-1} < \Delta_B^{t^*, t-1}$  or  $\Delta_A^{t^*, t-1} < \Delta_C^{t^*, t-1}$  in addition to  $N_d < n$ . For example,  $d$  is prescribed at  $(cccd; ccdd, ccccc)$  because  $N_d = 2$  and Alice has defected fewer times than Bob since the last round of full cooperation at  $t^* = t - 3$ .

3.  $t^* < t - m$



**Fig 2.** Schematic diagram of the transition between states of CAPRI- $n$ . The five rules of the strategy can be identified with the player's internal states [26], each of which is represented as a node in this diagram. An exception is state I, which corresponds to two nodes to clarify the following point: when  $t^* \geq t - m$ , I may have outgoing connections to A and P. When  $t^* < t - m$ , on the other hand, the only possible next state is limited to R. A blue (red) node means that the player has to choose  $c$  ( $d$ ) at the internal state. We have omitted error-caused transitions for the sake of simplicity.

- Recover: Alice has to recover cooperation by choosing  $c$ , under the condition that all the players except one cooperated in the last round. For  $n = 3$ , it means  $(ddddd, ddddc, ddddc)$  and its permutations.

4. In all the other cases, defect.

A strategy of this sort for the  $n$ -person PG game will be called CAPRI- $n$  after the first letters of the five constitutive rules. Note that these five rules may be implemented in a number of different ways [22], and we take this way because it provides the most straightforward way to prove the three criteria. Each of the rules can also be regarded as the player's internal *state* consisting of multiple history profiles [26]. For example, Alice can find herself at state R, the abbreviation for 'Recover', when her history profile is  $(ddddd, ddddc, ddddc)$ , at which she must choose  $c$ . The connection structure of the above five states is graphically represented in Fig. 2, which is helpful for understanding how defensibility and efficiency are realized as shown below.

Let us begin by checking defensibility. Our CAPRI- $n$  player Alice cooperates only at states C, A, and R, so the question is whether she can be forced to visit one of these states repeatedly with giving a strictly higher payoff to one of her co-players. If Alice's state is C, it means that everyone cooperated at  $t - 1$ . If some of her co-players defect from this full cooperation at  $t$ , she will retaliate at  $t + 1$  with state P, so she experiences unilateral defection at most once. Full cooperation is already broken, so it must be only through state A or R if she comes back to C. The former case means that Alice has already been compensated for the payoff loss. In the latter case, the only possible

history profile is  $(ddddd, ddddc, ddddc)$  unless she made a mistake, which means the compensation has been done in the last round. Finally, state A can be accessed from states P and I, at both of which one cannot exploit Alice who chooses  $d$ . To sum up, it is impossible to have the unilateral cooperation of a CAPRI- $n$  player repeatedly.

The next criterion is efficiency. Provided that CAPRI- $n$  is employed by all the players, only full cooperation or full defection can be a stationary state, and we can verify this statement by checking each possible case:

- If  $t^* = t - 1$ , everyone have to cooperate again as prescribed at state C, so full cooperation will continue.
- If  $t - m \leq t^* < t - 1$  and  $N_d < n$ , some players must be at state A while the others are at state P. The latter players at P will keep defecting until satisfying  $\Delta_A^{t^*, t-1} = \Delta_B^{t^*, t-1} = \Delta_C^{t^*, t-1}$ . If they make it with keeping  $t^* \geq t - m$ , all of them should choose  $c$  as prescribed at state A, and the resulting mutual cooperation will continue. If they don't, the situation to everyone reduces to state I, at which they will defect over and over.
- The remaining state is R, but it is always transient.

In order to judge efficiency, we need to consider error-caused transition between these two stationary states, i.e., full cooperation and full defection. The transition from the latter to the former is possible only through state R, which occurs with probability of  $O(e^{n-1})$ . On the other hand, full cooperation can be made robust against every possible type of  $(n - 1)$ -bit error if  $m = 2n - 1$ : Imagine that a player, say, Bob, mistakenly defects from full cooperation at  $t = 1$ . He will have state A at  $t = 2$ , while the others have state P, so their payoffs should be equalized at  $t = 3$  as a result of punishment. Note that this simple recovery from a single-bit error takes only two rounds. However, if this is interrupted at  $t = 3$  by another mistake occurring to any of the players, it will need additional two rounds to reach full cooperation at  $t = 5$ . The following example shows how Bob's mistakes at  $t = 1$  and 3 are corrected:

$$\begin{array}{lllll} A : & ccccc & ccccd & cccdc & ccdcd & cdcdc \\ B : & ccccd & cccdc & ccdcd & cdcdc & dcdcc. \\ C : & ccccc & ccccd & cccdc & ccdcd & cdcdc \end{array} \quad (5)$$

Among all types of  $(n - 1)$ -bit error, the longest memory length is needed to correct this kind of error that occurs every other round, so it requires  $m = 2(n - 1) + 1$  in total, where the last bit has been added to memorize the last round of full cooperation. Therefore, with memory length  $m = 2n - 1$ , the transition probability from mutual cooperation to defection can be suppressed down to  $O(e^n)$ . Therefore, the players form full cooperation in the limit of  $e \rightarrow 0$ , fulfilling the efficiency criterion.

The last criterion is distinguishability. If the others are AllC players, our CAPRI- $n$  player will continue unilateral defection when she defected  $n$  consecutive times by error, as prescribed by I. One can escape from such a state with probability of  $O(e^n)$  due to the condition of  $N_d < n$  for the rule C, so this stationary state coexists with full cooperation in the limit of  $e \rightarrow 0$ .

## Evolutionary simulation

We consider a standard stochastic model proposed in [29], where a well-mixed population of size  $N$  evolves over time by an imitation process. A key assumption of this model is that the mutation rate is low so that at most one mutant strategy can exist in the resident population. In other words, the time that it takes to go extinct or occupy the whole population by selection is assumed to be much shorter than the time



scale of mutation. Let us assume that a mutant strategy  $x$  is introduced to a population of strategy  $y$ . The population dynamics is modeled by the frequency-dependent Moran process, in which the fixation probability of the mutant is given in a closed form:

$$\phi_{xy} = \left( \sum_{i=0}^{N-1} \prod_{j=1}^i \Gamma_j \right)^{-1} \quad (6)$$

with  $\Gamma_j \equiv P_{j,j-1}/P_{j,j+1}$ , where  $P_{j,j\pm 1}$  denotes the probability that the number of mutants increases or decreases from  $j$  by one.

For  $n = 2$ , the fixation probability is calculated in the following way: Suppose that we have  $j$  individuals of the mutant strategy and  $N - j$  individuals of the resident strategy. If we randomly choose a mutant and a resident individual, their average payoffs are obtained as

$$\begin{cases} s_x = \frac{1}{N-1} [(j-1)s_{xx} + (N-j)s_{xy}] \\ s_y = \frac{1}{N-1} [(N-j-1)s_{yy} + js_{yx}], \end{cases} \quad (7)$$

respectively, where  $s_{\alpha\beta}$  is  $\alpha$ 's long-term payoff against  $\beta$ . According to the imitation process,  $x$  can change to  $y$  with probability  $f_{x \rightarrow y}$  defined as follows:

$$f_{x \rightarrow y} = \frac{1}{1 + \exp[\sigma(s_x - s_y)]}, \quad (8)$$

where  $\sigma$  means the strength of selection. Then, we have

$$\Gamma_j = \exp[\sigma(s_y - s_x)], \quad (9)$$

and the fixation probability is calculated as

$$\phi_{xy}^{-1} = \sum_{i=0}^{N-1} \prod_{j=1}^i e^{\sigma[(N-j-1)s_{yy} + js_{yx} - (j-1)s_{xx} - (N-j)s_{xy}]/(N-1)} \quad (10)$$

$$= \sum_{i=0}^{N-1} e^{\sigma i[(-i+2N-3)s_{yy} + (i+1)s_{yx} - (-i+2N-1)s_{xy} - (i-1)s_{xx}]/[2(N-1)]}. \quad (11)$$

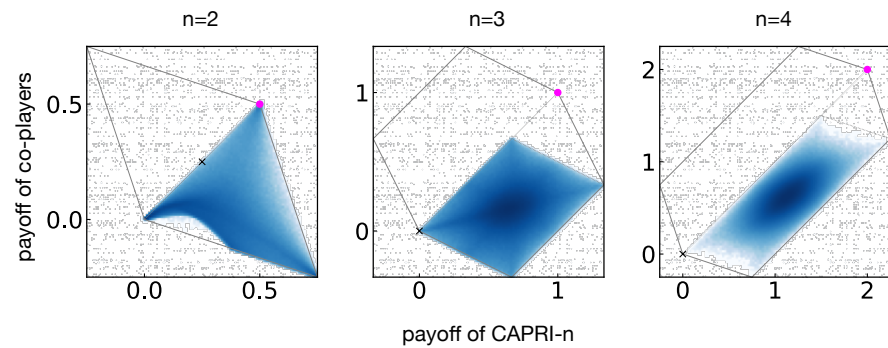
For  $n = 3$ , the fixation probability is calculated in a similar way. We randomly pick up three players from a well-mixed population, and the respective average payoffs of playing  $x$  and  $y$  can be written by using the binomial coefficients as follows:

$$\begin{cases} s_x = \frac{1}{(N-1)(N-2)} \left[ \binom{j-1}{2} s_{xxx} + \binom{j-1}{1} \binom{N-j}{1} s_{xxy} + \binom{N-j}{2} s_{xyy} \right] \\ s_y = \frac{1}{(N-1)(N-2)} \left[ \binom{j}{2} s_{yxx} + \binom{N-j-1}{1} \binom{j}{1} s_{yyx} + \binom{N-j-1}{2} s_{yyy} \right], \end{cases} \quad (12)$$

where  $s_{\alpha\beta\gamma}$  is player  $\alpha$ 's long-term payoff against co-players  $\beta$  and  $\gamma$ . Plugging these expressions into Eqs. (6) and (9), one can calculate the fixation probability  $\phi_{xy}$  for the three-person case as well.

We can interpret  $\phi_{xy}$  as transition probability from  $y$  to  $x$  from the viewpoint of the population. From the stationary distribution of this Markovian dynamics, we can thus calculate abundance of each available strategy in a numerically exact manner [31, 32]. For the sake of simplicity, we use the donation game as a simplified form of the PD game as well as its generalization to  $n$  players in the numerical calculation. That is, with the benefit of cooperation  $b > 1$ , each player can donate  $b/(n-1)$  to each co-player at the unit cost, which corresponds to  $\rho = nb/[b + (n-1)]$  up to scaling.





**Fig 3.** Distribution of long-term payoffs when a CAPRI- $n$  player meets co-players whose  $p_{\mu\nu}$ 's are randomly sampled from the unit interval. The multiplication factors for  $n = 2, 3$ , and  $4$  are  $1.5, 2$ , and  $3$ , respectively, and the solid lines indicate the region of feasible payoffs. In each case, the filled circle means the long-term payoffs when CAPRI- $n$  is adopted by all the players, whereas the cross shows those of TFT players as a reference point.

## Results

### Friendly rivalry

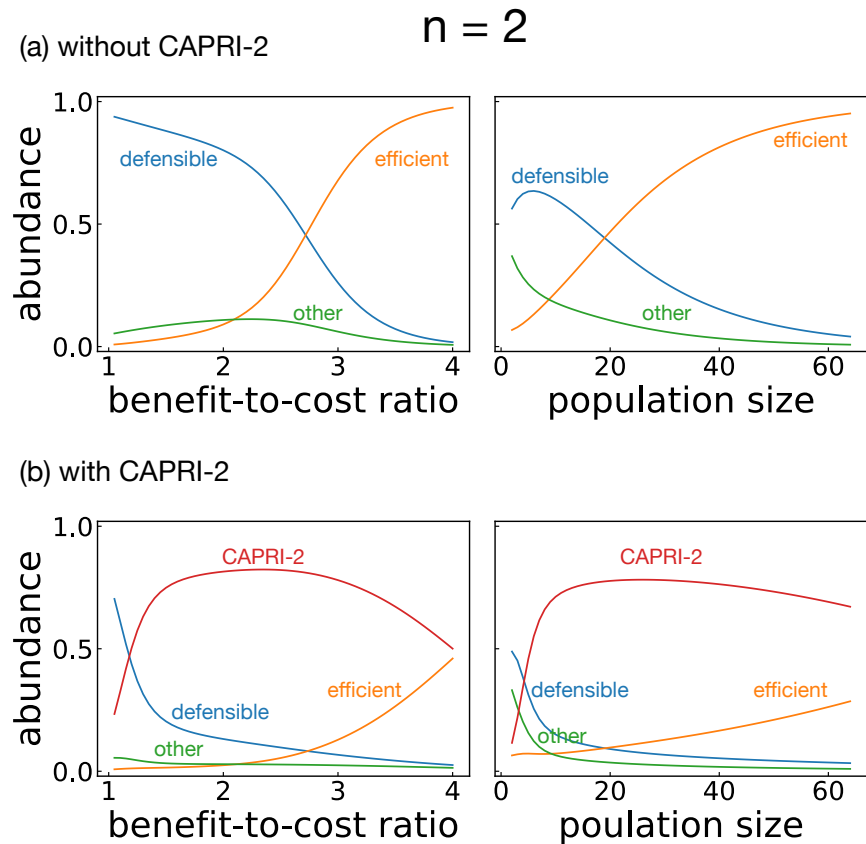
To check the validity of our construction, we examined the three criteria by using graph-theoretic calculations [19, 22, 33]. For  $n = 2$ , we directly confirmed that CAPRI- $n$  is indeed a successful strategy satisfying all the three criteria. For  $n = 3$ , we conducted mapping to an automaton to obtain a simplified yet equivalent graph representation [26], and the resulting automaton indeed passed all the criteria. For  $n = 4$ , the required amount of calculation to directly check the criteria was beyond our computational resources, so we employed a Monte Carlo method to simulate the game. The Monte Carlo method was also used to double-check the performance of CAPRI-2 and CAPRI-3.

The Monte Carlo calculation was performed as follows: Let us denote a memory-one strategy as  $(p_{cc}, p_{cd}, p_{dc}, p_{dd})$  where  $p_{\mu\nu}$  means the player's probability to cooperate when the player and the co-player did  $\mu$  and  $\nu$ , respectively, in the previous round. The initial  $\mu$  and  $\nu$  can be omitted in the strategy description because they are irrelevant to the long-term payoff as long as  $e > 0$ . Figure 3 shows the distribution of payoffs when Alice used CAPRI- $n$  whereas each of her co-players' strategies was composed of four  $p_{\mu\nu}$ 's randomly sampled from the unit interval. The co-players' payoffs never exceeded Alice's, as required by defensibility.

We also calculated the probability of full cooperation for  $n = 2, 3$  and  $4$  when CAPRI- $n$  was adopted by all the players in order to check efficiency. By using linear-algebraic [18, 19] or Monte Carlo calculation, For  $e = 10^{-4}$ , we obtained  $0.999, 0.997, 0.978$  for  $n = 2, 3$ , and  $4$ , respectively, which supports the conclusion that they all satisfy the efficiency criterion.

### Evolutionary robustness

Before checking the evolutionary performance of our proposed strategy, we conducted simulations without CAPRI- $n$  for comparison. Figures 4(a) and 5(a) show the results when the strategies were sampled from deterministic memory-one for  $n = 2$  and  $3$ . When  $b$  was low and/or  $N$  was small, defensible strategies such as AllD tended to be favored by selection, and the resulting cooperation level was low. On the other hand,

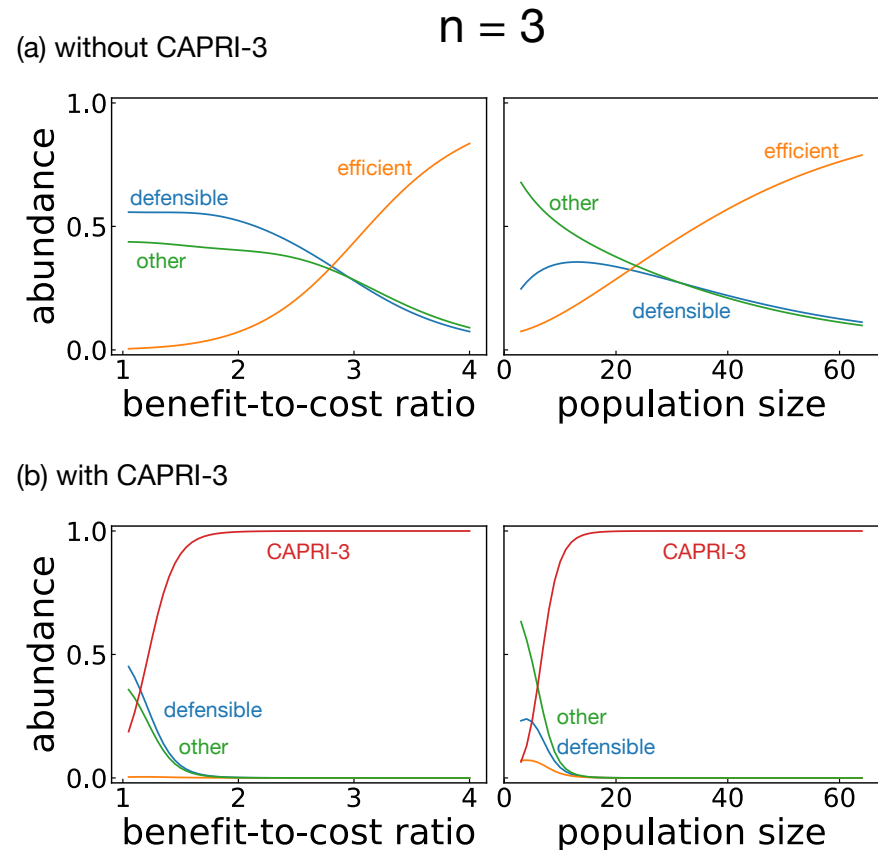


**Fig 4.** Abundance of strategies for  $n = 2$  as the benefit-to-cost ratio  $b$  and the population size  $N$  vary. The default values were  $b = 3$  and  $N = 30$  unless otherwise specified. The strength of selection and the error probability were set to be  $\sigma = 1$  and  $e = 10^{-4}$ , respectively. (a) Simulation result with 16 memory-one deterministic strategies, classified into three categories, i.e., efficient, defensible, and the other strategies. (b) Effect of CAPRI-2 when it was added to the available set of strategies.

when  $b$  or  $N$  was large, efficient strategies were favored, and they achieved a high level of cooperation. The reason is that cooperative strategies maintained high payoffs by interacting with many other cooperators even if they were exploited by a small number of aggressive mutants.

When CAPRI- $n$  was introduced, it occupied a large amount of the population as shown in Figs. 4(b) and 5(b). Whereas each memory-one strategy flourished depending on the environmental parameters  $b$  and  $N$ , CAPRI- $n$  was found abundant in the entire parameter region. In particular, it is striking that CAPRI-3 overwhelms all the other strategies in the three-person PG game for any moderate sizes of  $b$  and  $N$  [Fig. 5(b)].

It is nevertheless worth pointing out that CAPRI-2 gave more and more room to efficient strategies in the iterated PD game as  $b$  or  $N$  increases [Fig. 4(b)], and this is due to neutral drift: Although CAPRI-2 earns a strictly higher long-term payoff than ALLC = (1, 1, 1, 1) and Win-Stay-Lose-Shift (WSLS) = (1, 0, 0, 1), it does not with respect to (1, 1, 1, 0), which can, in turn, be invaded by WSLS. For this reason, WSLS can become abundant in the presence of (1, 1, 1, 0) when the environmental conditions are favorable.



**Fig 5.** Abundance of strategies for  $n = 3$  as the benefit-to-cost ratio  $b$  and the population size  $N$  vary. The default values were  $b = 3$  and  $N = 30$  unless otherwise specified. The strength of selection and the error probability were set to be  $\sigma = 1$  and  $e = 10^{-4}$ , respectively. (a) Simulation result with 64 memory-one deterministic strategies, classified into three categories, i.e., efficient, defensible, and the other strategies. (b) Effect of CAPRI-3 when it was added to the available set of strategies.

## Discussion

In summary, we have constructed a friendly-rivalry strategy for the iterated  $n$ -person PG game. It maintains a cooperative Nash equilibrium in the presence of implementation error with probability  $e \ll 1$ , and it has evolutionary robustness regardless of the environmental conditions such as the population size and the strength of selection. In this sense, the  $n$ -person social dilemma is solved. The strategy requires memory of the previous  $m = 2n - 1$  rounds and consists of the following five rules: Cooperate if everyone did, accept punishment for your own mistake, punish others' defection, recover cooperation if you find a chance, and defect in all the other circumstances.

Although we have considered only implementation error, perception error can also be corrected if it occurs with sufficiently low probability: The disagreement between the players' history profiles due to the perception error will soon be removed at full defection, and the players will escape from mutual defection with probability of  $O(e^n)$ . Unless another perception error perturbs this process, the players will eventually arrive at full cooperation, overcoming the perception error.

Another important solution concept to the  $n$ -person dilemma can be derived from a different set of criteria: By requiring mutual cooperation, error correction, and retaliation with a time scale of  $k$  rounds, one can characterize the all-or-none (AON- $k$ ) strategy, which is defined as prescribing  $c$  only when everyone cooperated or no one did in each of the previous  $k$  rounds [30, 34, 35]. For example, WSLs = (1, 0, 0, 1) is equivalent to AON-1. For each  $k$ , one can find a threshold of the multiplication factor above which AON- $k$  constitutes a subgame-perfect equilibrium [30]. AON- $k$  performs well in evolutionary simulation because it prescribes  $d$  as the default action, just as CAPRI- $n$  does in state I, unless the players have synchronized their behavior over the previous  $k$  rounds. As a result, it earns a strictly higher payoff against a broad range of strategies.

In general, CAPRI- $n$  with  $m = 2n - 1$  can repeatedly exploit the other co-players playing AON- $k$  if  $k < m - 1$ , which means that an AON- $k$  population can readily be invaded by CAPRI- $n$  unless  $k$  is large enough. Considering the condition for AON- $k$  to be subgame perfect, one could speculate that AON with small  $k$  can be abundant in an environment with a high multiplication factor. However, our finding implies that such a simple solution may not be sustained when CAPRI- $n$  is available. This is especially crucial when population size is not large enough because AON- $k$  lacks defensibility. Still, AON- $k$  remains as a strong competitor to CAPRI- $n$  in evolutionary simulation: For example, although WSLs earns a strictly less payoff against CAPRI-2, it circumvented the difficulty of fixation with the aid of a third strategy (1, 1, 1, 0).

From a practical point of view, it is worth noting that the five rules of CAPRI- $n$  mostly refer to two factors: One is the players' last action at  $t - 1$ , and the other is the differences in the players' respective numbers of defections over the previous  $m$  rounds. In other words, exact details of the history profile are irrelevant, and this point greatly reduces the cognitive burden to play this strategy. In fact, according to a recent experiment, people assign reputation to their co-players based mainly on their last action and their average numbers of defection [36]. This could explain the reason that such a delicate relationship called friendly rivalry can develop spontaneously and unwittingly among a group of people. How to keep such a relation healthy and productive has so far been acquired as tacit knowledge surrounded by anecdotes and experiences, and CAPRI- $n$  expresses its essential how-tos in a form of explicit knowledge which can be designed, analyzed, and transmitted systematically.

## Acknowledgments

Y.M. acknowledges support from MEXT as “Exploratory Challenges on Post-K computer (Studies of multi-level spatiotemporal simulation of socioeconomic phenomena)” and from Japan Society for the Promotion of Science (JSPS) (JSPS KAKENHI; Grant no. 18H03621). S.K.B. acknowledges support by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1I1A2071670). This research used computational resources of the K computer provided by the RIKEN Center for Computational Science through the HPCI System Research project (Project ID:hp160264). OACIS was used for the simulations in this study [37].

## References

1. Nowak MA, Highfield R. Supercooperators. New York: Free Press; 2011.
2. Nowak MA. Five rules for the evolution of cooperation. *Science*. 2006;314(5805):1560–1563.
3. Sigmund K. The Calculus of Selfishness. Princeton: Princeton Univ. Press; 2010.
4. Fudenberg D, Tirole J. Game Theory. Cambridge, MA: MIT Press; 1991.
5. Molander P. The optimal level of generosity in a selfish, uncertain environment. *J Conflict Resolut.* 1985;29(4):611–618.
6. Boyd R, Richerson PJ. The evolution of reciprocity in sizable groups. *J Theor Biol.* 1988;132(3):337–356.
7. Gintis H. Behavioral ethics meets natural justice. *Politics Philos Econ.* 2006;5(1):5–32.
8. Press WH, Dyson FJ. Iterated Prisoner’s Dilemma contains strategies that dominate any evolutionary opponent. *Proc Natl Acad Sci USA.* 2012;109(26):10409–10413.
9. Hilbe C, Nowak MA, Sigmund K. Evolution of extortion in Iterated Prisoner’s Dilemma games. *Proc Natl Acad Sci USA.* 2013;110(17):6913–6918.
10. Hilbe C, Nowak MA, Traulsen A. Adaptive dynamics of extortion and compliance. *PloS one.* 2013;8(11):e77886.
11. Stewart AJ, Plotkin JB. From extortion to generosity, evolution in the Iterated Prisoner’s Dilemma. *Proc Natl Acad Sci USA.* 2013;110(38):15348–15353.
12. Adami C, Hintze A. Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature communications.* 2013;4(1):1–8.
13. Hilbe C, Traulsen A, Sigmund K. Partners or rivals? Strategies for the iterated prisoner’s dilemma. *Games Econ Behav.* 2015;92:41–52.
14. Hilbe C, Chatterjee K, Nowak MA. Partners and rivals in direct reciprocity. *Nat Hum Behav.* 2018;2(7):469–477.
15. Akin E. What you gotta know to play good in the iterated Prisoner’s Dilemma. *Games.* 2015;6(3):175–190.

16. Akin E. The Iterated Prisoner's dilemma: good strategies and their dynamics. In: Assani I, editor. *Ergodic Theory, Advances in Dynamical Systems*. Berlin: de Gruyter; 2016. p. 77–107.
17. Duersch P, Oechssler J, Schipper BC. Unbeatable imitation. *Games Econ Behav*. 2012;76(1):88–96.
18. Yi SD, Baek SK, Choi JK. Combination with anti-tit-for-tat remedies problems of tit-for-tat. *J Theor Biol*. 2017;412:1–7.
19. Murase Y, Baek SK. Seven rules to avoid the tragedy of the commons. *J Theor Biol*. 2018;449:94–102.
20. Stewart AJ, Plotkin JB. Collapse of cooperation in evolving games. *Proc Natl Acad Sci USA*. 2014;111(49):17558–17563.
21. Stewart AJ, Plotkin JB. Small groups and long memories promote cooperation. *Sci Rep*. 2016;6:26889.
22. Murase Y, Baek SK. Five rules for friendly rivalry in direct reciprocity; 2020. arXiv preprint arXiv:2004.00261.
23. Baek SK, Kim BJ. Intelligent Tit-for-Tat in the iterated prisoner's dilemma game. *Phys Rev E*. 2008;78(1):011125.
24. Nowak M. Stochastic strategies in the prisoner's dilemma. *Theor Popul Biol*. 1990;38(1):93–112.
25. Nowak MA, Sigmund K, El-Sedy E. Automata, repeated games and noise. *J Math Biol*. 1995;33(7):703–722.
26. Murase Y, Baek SK. Automata representation of successful strategies for social dilemmas; 2019. arXiv:1910.02634.
27. Imhof LA, Fudenberg D, Nowak MA. Evolutionary cycles of cooperation and defection. *Proc Natl Acad Sci USA*. 2005;102(31):10797–10800.
28. Imhof LA, Fudenberg D, Nowak MA. Tit-for-tat or win-stay, lose-shift? *J Theor Biol*. 2007;247(3):574–580.
29. Imhof LA, Nowak MA. Stochastic evolutionary dynamics of direct reciprocity. *Proc R Soc B*. 2010;277(1680):463–468.
30. Hilbe C, Martinez-Vaquero LA, Chatterjee K, Nowak MA. Memory- $n$  strategies of direct reciprocity. *Proc Natl Acad Sci USA*. 2017;114(18):4715–4720.
31. Jeong HC, Oh SY, Allen B, Nowak MA. Optional games on cycles and complete graphs. *J Theor Biol*. 2014;356:98–112.
32. Baek SK, Jeong HC, Hilbe C, Nowak MA. Comparing reactive and memory-one strategies of direct reciprocity. *Sci Rep*. 2016;6:25676.
33. Hougardy S. The Floyd–Warshall algorithm on graphs with negative cycles. *Inf Process Lett*. 2010;110(8-9):279–281.
34. Hauert C, Schuster HG. Effects of increasing the number of players and memory size in the iterated Prisoner's Dilemma: a numerical approach. *Proc R Soc B*. 1997;264(1381):513–519.

35. Lindgren K. Evolutionary Dynamics in Game-Theoretic Models. In: Brian Arthur W, Durlauf SN, Lane D, editors. *The Economy as an Evolving Complex System II*. Upper Saddle River, NJ: Addison-Wesley; 1997. p. 337–368.
36. Cuesta JA, Gracia-Lázaro C, Ferrer A, Moreno Y, Sánchez A. Reputation drives cooperative behaviour and network formation in human groups. *Sci Rep*. 2015;5(1):1–6.
37. Murase Y, Uchitane T, Ito N. An open-source job management framework for parameter-space exploration: OACIS; 2018. arXiv:1805.00438.