

GLAD: A simple adaptive strategy that yields cooperation in dilemma games*

Steffen Huck[†]

University Collge London

Hans-Theo Normann

Jörg Oechssler

Royal Holloway, University of London

University of Bonn

April 29, 2004

Abstract

We introduce a simple adaptive rule where agents choose a cooperative effort on a grid. Agents can adjust this effort step by step and **G**ains and **L**osses **A**djust **D**irections. We show that this process converges to the cooperative outcome in a two-person Prisoners' Dilemma game, and we provide simulations showing that the results also holds with a larger number of agents.

PACS codes: 87.23.Ge, 89.65.Gh.

Keywords: cooperation, prisoners' dilemma, adaptive dynamics.

*Financial support from DFG grants HU 787/1-1 and OE 198/1-1 is gratefully acknowledged.

[†]Department of Economics & ELSE, University College London, Gower Street, London WC1E 6BT, UK, email *s.huck@ucl.ac.uk*, phone +44 20 7679 5895, fax +44 20 7916 2774.

1 Introduction

Cooperative behavior is widespread among both, humans and animals, and has even been documented among viruses [20]. Models explaining the emergence of cooperation range from models of direct reciprocity (variations of “tit-for-tat”) [1, 2, 6, 10, 12, 21, 22] via indirect reciprocity models [13, 14] and models of spatial interaction [8, 9, 18] to aspiration-based models [7, 15, 16, 17, 11]. Most of these models view behavior as adaptive and rules driving adaptation typically require substantial cognitive capabilities. Reciprocity, for example, requires that individuals have an understanding of how others affect their payoff, i.e., they require that individuals have a mental model of the world [19].

Here we introduce an adaptive rule (GLAD) that is much simpler and could help to explain why cooperation is observed in non-human species. In particular, GLAD does not require any knowledge about how own payoffs depend on others. In fact, it does not even require knowledge about the existence of others. Yet, it does yield cooperation. GLAD prescribes that individuals who choose a cooperative effort, increase (decrease) their effort step by step as long as their payoff increases. If it decreases, the direction of adjustments is reversed. This process converges to cooperation in dilemma games.

Players engage in a symmetric game in which the strategy space can be ordered along one dimension. This could be the “degree” or “intensity” of cooperation but, in general, it could also be something which affects the degree of cooperation only indirectly, for example, the time spent hunting

or caring for members of a group. As in the case of time, the action variable might be continuous but we assume that chosen actions are taken from a grid (that can be arbitrarily fine though). Actions are adjusted by moving either one step up or down the grid (in discrete time). The adaptive rule we introduce prescribes the direction of adjustments. It says: Go another step in the same direction if the last step increased your payoff. If not, switch directions.

As Gains and Losses Affect the Direction of adjustments we shall call the rule GLAD. GLAD is related to (aspiration-based) win-stay, lose-shift strategies that implement Pavlovian ideas [11].¹ If recent payoffs are above a certain threshold, a win-stay, lose-shift rule prescribes to repeat the same action again. If recent payoffs are below the threshold, the rule prescribes to shift to another action (which, if there is more than one alternative, may be chosen randomly). Our rule works similarly. But instead of implementing the win-stay, lose-shift principle on the domain of actions, it implements the principle on the domain of adaptations. If the last adaptation was successful, pick the same change again. If not, shift back.

Our model is related to models by Fort [4, 5] who also studies agents who increase or decrease their level (here a probability) of cooperation depending on how good the last payoff was in comparison to some other variable. However, in Fort's models the comparison variable is not the previously obtained payoff but a payoff expectation based on the assumption that all

¹Pavlovian strategies that can induce cooperation are also discussed by Brauchli *et al.* [3] who find that they do particularly well in spatially structured populations.

other agents use the same probability of cooperation. Insofar, Fort's models require far more sophisticated agents than ours. Moreover, our agents achieve higher levels of cooperation than Fort's.

2 The model

To analyze how GLAD induces cooperation we introduce a simple continuous n -person prisoners'-dilemma (PD) game (though our findings can be extended to a larger class of games). In our model each individual i chooses an effort $x_i \geq 0$ which directly benefits all other players $j \neq i$. However, player i has to incur costs of βx_i^α with $\alpha \geq 1$, $0 < \beta \leq 1$ and $\beta < 1$ if $\alpha = 1$. Thus, a player's payoff function can be written as

$$\Pi_i(x_1, \dots, x_n) = \sum_{j \neq i} x_j - \beta x_i^\alpha,$$

and the non-cooperative (Nash) equilibrium is clearly given by zero cooperation, i.e., by $x_i^N = 0$ for all i . The cooperative solution is found by maximizing

$$\sum_{i=1}^n \Pi_i(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j \neq i}^n x_j - \beta \sum_{i=1}^n x_i^\alpha$$

with respect to all x_i 's. An interior solution exists for all $\alpha > 1$ and is given by $x_i^c = \left[\frac{n-1}{\alpha\beta} \right]^{1/(\alpha-1)}$ for all i .

As a special case our continuous prisoners' dilemma (PD) contains the usual PD by setting $\alpha = 1$, $n = 2$ and letting players choose strategies from $\{0, 1\}$, where 1 stands for cooperate and 0 for defect. This yields the

following two-player game.

	C	D
C	$1 - \beta, 1 - \beta$	$-\beta, 1$
D	$1, -\beta$	$0, 0$

For $\alpha = 1$, $\beta < 1$, $n = 2$, and $x_i \in [0, 1]$, our game reduces to the Wahl and Nowak's [21] continuous PD where the cooperative solution is on the boundary, $x_i^c = 1$.

Individuals choose their actions as multiples of δ with $\delta > 0$ being the arbitrarily small grid size. For convenience, we assume that the cooperative action is a multiple of δ . Our rule is formalized as follows. Let σ_i^t be an indicator for success of individual i in period t ,

$$\sigma_i^t := \begin{cases} 1 & \text{if } \Pi_i^t - \Pi_i^{t-1} \geq 0 \\ -1 & \text{if } \Pi_i^t - \Pi_i^{t-1} < 0 \end{cases} .$$

The process is specified by $x_i^{t+1} = \max \{0, x_i^t + \delta s_i^t\}$ where the direction of movement s_i^t is $s_i^t := \sigma_i^t \text{sign}(x_i^t - x_i^{t-1})$. Given some initial conditions specifying for each player i a tuple (x_i^0, s_i^0) , we can now determine the complete path of the adaptive dynamics.

Consider first what happens when both individuals move up by δ . Individual i will improve his payoff (and will, therefore, repeat this direction) if and only if $\Delta \Pi_i = \delta - \beta(x_i + \delta)^\alpha + \beta x_i^\alpha \geq 0$ or

$$\delta + \beta x_i^\alpha \geq \beta(x_i + \delta)^\alpha. \tag{1}$$

For $\alpha = 1$ and $\beta < 1$, (1) is satisfied independently of x_i . For $\alpha > 1$, consider a Taylor expansion of the right hand side of (1) at $\delta = 0$. We see that (1)

holds for small δ if and only if $\alpha\beta(x_i + \delta)^{\alpha-1} \leq 1$ or $x_i < (\alpha\beta)^{-1/(\alpha-1)} - \delta$. That is, when both individuals move up, the payoff for i increases if and only if x_j is below $x_j^c - \delta$. Likewise, it can be shown that when both individuals move down, i 's payoff increases if and only if $x_j > x_j^c + \delta$. To complete the picture we only have to note that if one player moves up while the other moves down, the former always decreases his payoff while the latter increases it, which results in both moving down one period later.

<Figure 1>

Figure 1 is drawn according to this logic (for the case of $\alpha > 1$). The strategy space in Figure 1 is partitioned into four large regions and two “corridors” meeting at the cooperative outcome. In Region A both players are below the critical value of $x_j^c - \delta$. Denoting by $(+-)$ a move up by the first player and a move down by the second, etc., we get the following transitions

$$\begin{array}{c} (+-) \rightarrow (--) \leftarrow (-+) \\ \downarrow \\ (++)^\circ \end{array}$$

Thus, in Region A , players continuously increase their degree of cooperation. Just the opposite happens in Region C where both players decrease their cooperation levels. For Region B it is easy to derive the following transitions.

$$\begin{array}{c} (+-) \rightarrow (--) \rightleftarrows (-+) \\ \uparrow \\ (++) \end{array}$$

Thus, the process zigzags as indicated in Figure 1. In the mirror image Region D zigzagging occurs between $(--)$ and $(+-)$. It remains to consider the movements inside the two corridors. Straightforward analysis shows that the process moves inside the corridors in the direction of x^c and eventually falls into a cycle at the intersection of the two corridors.

Intuitively, a similar logic should apply for larger number of players $n > 2$. However, obviously we cannot use the same techniques for showing the convergence to the cooperative outcome for arbitrary n . Instead we use simulations to study the behavior of GLAD for larger n .

<Figure 2>

Figure 2 shows the results of simulations for the parameters $\delta = 10^{-3}$, $\alpha = 2$, $\beta = 1/2$ and $n \in \{2, 3, 4, 5, 6\}$. Each simulation starts by assigning a random initial conditions for each player i , (x_i^0, s_i^0) , where $x_i^0 \in [0, 2n - 2]$ and $s_i^0 \in \{-1, 1\}$. Then the GLAD adjustment process was triggered. The simulations stopped when the process had converged² to the cooperative solution which is $x_i^c = n - 1$ for the above parameters.

The simulations show that the GLAD process does converge to the cooperative solution for $n > 2$. Depending on the initial conditions, x_i often monotonically move towards x_i^c but non-monotonicities do occur. In the simulations with $n \leq 4$ shown in the figure, once players have reached x_i^c they stayed there independently of the actions of the other players. With $n \in \{5, 6\}$ more complicated patterns occurred.

²Convergence was given in round t if $|\bar{x}_i^k - x_i^c| < n\delta, \forall k \in \{t - 100, \dots, t\}$, where $\bar{x}_i^k = \frac{1}{100} \sum_{m=k-100}^k x_i^m$.

3 Conclusion

We have introduced a simple adaptive rule where agents choose a cooperative effort on a grid. Agents can adjust this effort step by step and Gains and Losses Adjust Directions. We have shown that this process converges to the fully cooperative outcome in an n -person Prisoners' Dilemma game.

There are a number of interesting open questions. First of all, how robust is this result? Does GLAD also achieve cooperation in other games? To that question we have at least a partial answer. Whenever there is a fully cooperative outcome and agents use GLAD there exists a cooperative cycle. To see this, suppose all players are one grid point below x_i^c and, next, move upwards. Then, clearly, all players' payoffs will increase. Hence, they will increase their degree of cooperation again. Now, however, all players' payoffs will decrease as they have moved beyond the optimal level of cooperation and now waste resources. Given our adaptive rule, all players will now reverse their direction. Hence, they will be back in the cooperative outcome in the next period which means that payoffs will have increased again. Accordingly, each player will further reduce their level x_i . This, now will reduce all players payoffs and they will go up again. And so forth. It is clear that such a cycle always exists—regardless of the actual payoff function and this alone seems an intriguing property of our rule.

A further question concerns how well GLAD would do against other strategies. For thinking about this question it seems worthwhile to make a couple of observations about how GLAD performs in simple (non-stochastic) optimization problems. There it is straightforward to see that an individual

relying on this rule will converge to the neighborhood of a local maximum where it will cycle around it. If the profit function is concave, the individual will converge to the global maximum. While these observation may seem utterly trivial, they might be important from the evolutionary perspective. A simple rule that helps individuals to solve maximization problems and, additionally, induces cooperation in dilemmata seems predestined to evolutionary success.

References

- [1] Axelrod, R. (1984) *The Evolution of Cooperation*, New York: Basic Books.
- [2] Binmore, K., and L Samuelson (1992) Evolutionary stability in repeated games played by finite automata, *Journal of Economic Theory* 57, 278-305.
- [3] Brauchli, K., T. Killingback, and M. Doebeli (1999) Evolution of cooperation in spatially structured populations, *Journal of Theoretical Biology* 200, 405-417.
- [4] Fort, H. (2003) Exploring the cooperative regimes in an agent-based model: Indirect reciprocity vs. selfish incentives, *Physica A* 326, 286-298.
- [5] Fort, H. (2003) Cooperation and self-regulation in a model of agents playing different games, *Physical Review E* 68, art. no. 026118.

- [6] Frean, M. (1996) The evolution of degrees of cooperation, *Journal of Theoretical Biology* 182, 549-559.
- [7] Macy, M.W., and A. Flache (2002) Learning dynamics in social dilemmas, *Proceedings of the National Academy of Sciences* 99, 7229-7236.
- [8] Masuda, N., and K. Aihara (2003) Spatial prisoner's dilemma optimally played in small-world networks, *Physics Letters A* 313, 55-61.
- [9] Nakamaru, H., H. Matsuda, and Y. Iwasa (1997) The evolution of cooperation in a lattice-structured population, *Journal of Theoretical Biology* 184, 65-81.
- [10] Nowak, M.A., and K. Sigmund (1992) Tit for tat in heterogenous populations, *Nature* 355, 250-253.
- [11] Nowak, M.A., and K. Sigmund (1992) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game, *Nature* 364, 56-58.
- [12] Nowak, M.A., and K. Sigmund (1994) The alternating prisoner's dilemma, *Journal of Theoretical Biology* 168, 219-275.
- [13] Nowak, M.A., and K. Sigmund (1998a) Evolution of indirect reciprocity by image scoring, *Nature* 393, 573-577.
- [14] Nowak, M.A., and K. Sigmund (1998b) The dynamics of indirect reciprocity, *Journal of Theoretical Biology* 194, 561-574.

- [15] Oechssler, J. (2002) Cooperation as a result of learning with aspiration levels, *Journal of Economic Behavior and Organization* 49, 405-409.
- [16] Posch, M. (1999) Win-stay, lose-shift strategies for repeated games: Memory length, aspiration levels and noise, *Journal of Theoretical Biology* 198, 183-195.
- [17] Posch, M., A. Pichler, and K. Sigmund (1999) The efficiency of adapting aspiration levels, *Proceedings of the Royal Society London B* 266, 1427-1435.
- [18] Sella, G., and M. Lachmann (2000) On the dynamic persistence of cooperation: How lower individual fitness induces higher survivability, *Journal of Theoretical Biology* 206, 465-485.
- [19] Stevens, J.R., and M.R. Hauser (2004) Why be nice? Psychological constraints on the evolution of cooperation, *TRENDS in Cognitive Sciences* 8(2):60-65.
- [20] Turner, P.E., and L. Chao (1999) Prisoner's dilemma in an RNA virus, *Nature* 398, 441-443.
- [21] Wahl, L., and M. Nowak (1999a) The continuous prisoner's dilemma: I. Linear reactive strategies, *Journal of Theoretical Biology* 200, 307-321.
- [22] Wahl, L., and M. Nowak (1999b) The continuous prisoner's dilemma: II. Linear reactive strategies with noise, *Journal of Theoretical Biology* 200, 323-338.

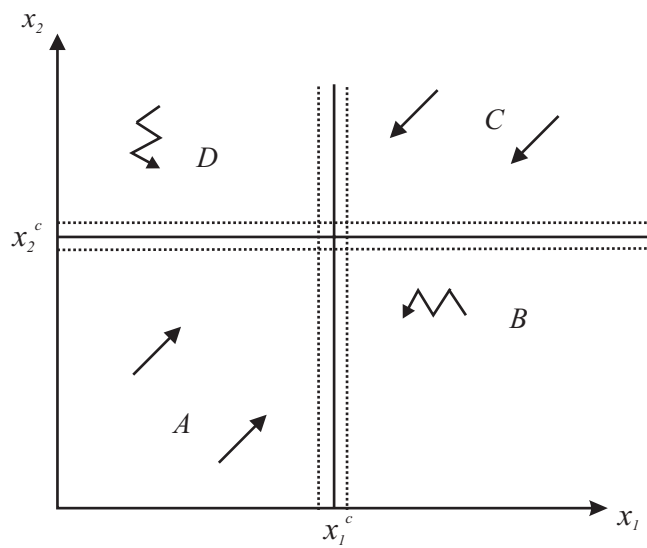


Figure 1: Adjustments of effort levels (for $\alpha > 1$).

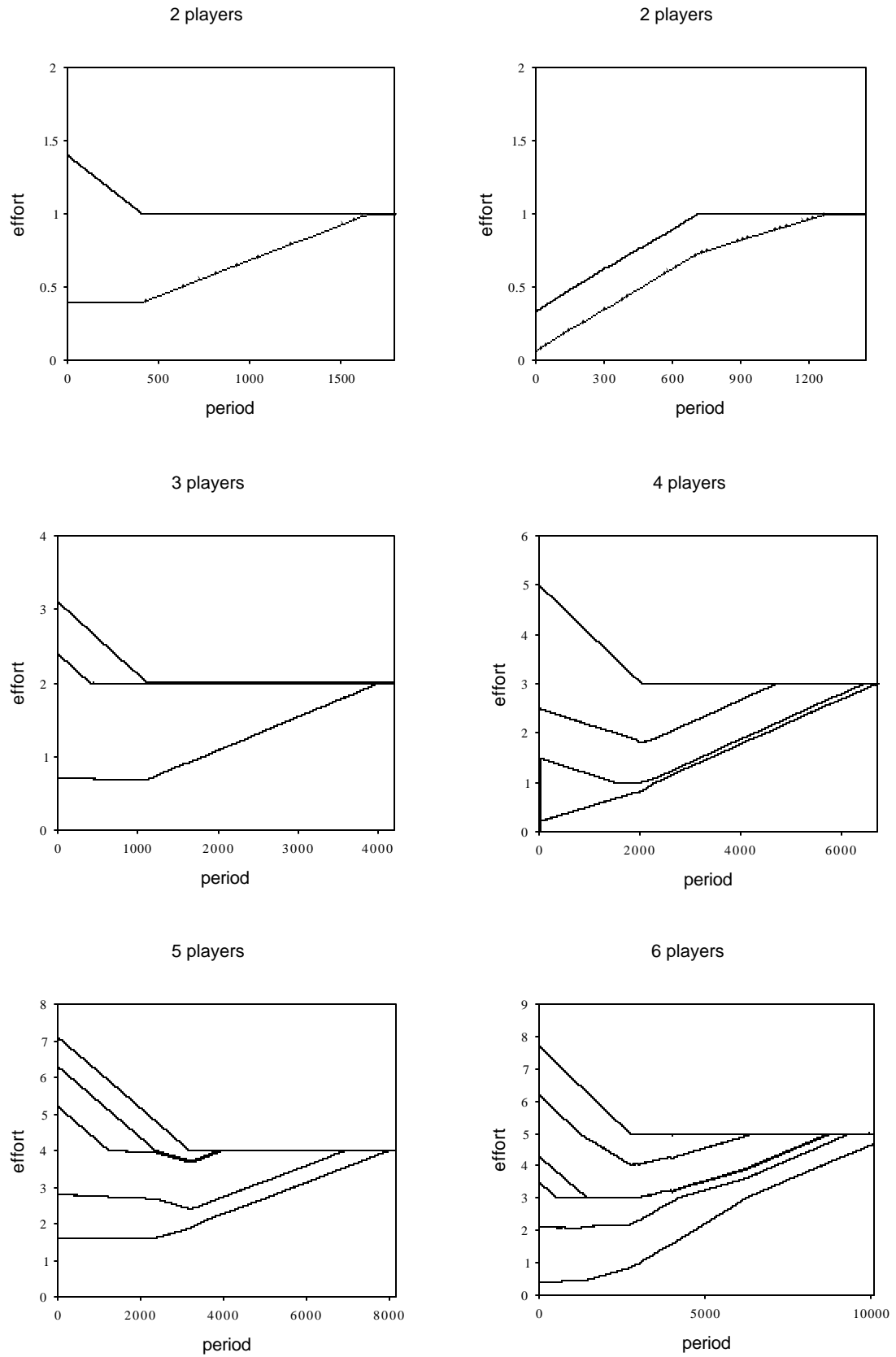


Figure 2: Simulations