Vol 457 | 1 January 2009 | doi:10.1038/nature07601

LETTERS

Indirect reciprocity provides only a narrow margin of efficiency for costly punishment

Hisashi Ohtsuki[1,2], Yoh Iwasa[3] & Martin A. Nowak[4]

Indirect reciprocity[1–5] is a key mechanism for the evolution of human cooperation. Our behaviour towards other people depends not only on what they have done to us but also on what they have done to others. Indirect reciprocity works through reputation[5–17]. The standard model of indirect reciprocity offers a binary choice: people can either cooperate or defect. Cooperation implies a cost for the donor and a benefit for the recipient. Defection has no cost and yields no benefit. Currently there is considerable interest in studying the effect of costly (or altruistic) punishment on human behaviour[18–25]. Punishment implies a cost for the punished person. Costly punishment means that the punisher also pays a cost. It has been suggested that costly punishment between individuals can promote cooperation. Here we study the role of costly punishment in an explicit model of indirect reciprocity. We analyse all social norms, which depend on the action of the donor and the reputation of the recipient. We allow errors in assigning reputation and study gossip as a mechanism for establishing coherence. We characterize all strategies that allow the evolutionary stability of cooperation. Some of those strategies use costly punishment; others do not. We find that punishment strategies typically reduce the average payoff of the population. Consequently, there is only a small parameter region where costly punishment leads to an efficient equilibrium. In most cases the population does better by not using costly punishment. The efficient strategy for indirect reciprocity is to withhold help for defectors rather than punishing them.

Human societies are organized around cooperative interactions. But why would natural selection equip selfish individuals with altruistic tendencies? This question has fascinated evolutionary biologists for decades. One answer is given in terms of direct reciprocity[26–29]. There are repeated encounters between the same two individuals: I help you, and you help me. More recently, indirect reciprocity has emerged as a more general model: I help you, and somebody helps me. Indirect reciprocity is based on reputation[5]. People monitor the social interactions within their group. Helping others establishes the reputation of being a helpful individual. Natural selection can favour strategies that help those who have helped others[5–17]. The consequences for widespread cooperation are enormous. Direct reciprocity is like an economy based on the exchange of goods, whereas indirect reciprocity resembles the invention of money. The money that feeds the engines of indirect reciprocity is reputation. For direct reciprocity, my strategy depends on what you have done to me; for indirect reciprocity, my strategy also depends on what you have done to others. Direct and indirect reciprocity are mechanisms for the evolution of cooperation[30].

Punishment refers to an action that implies a cost for the punished person. Costly punishment means that the punisher also pays a cost for exercising punishment. In certain experimental situations costly punishment has been called 'altruistic punishment', because the punishers cannot expect any material gain from their action[20,21]. In reality, however, most punishment actions among humans are associated with the expectation of a delayed material gain; they are therefore not altruistic.

The suggested idea for the evolution of cooperation is that people might be more willing to cooperate under the threat of punishment. However, we note that costly punishment is not a separate mechanism for the evolution of cooperation but a form of direct or indirect reciprocity. If I punish you because you have defected with me, then I use direct reciprocity. If I punish you because you have defected with others, then indirect reciprocity is at work. In the setting of direct reciprocity, punishment is a form of retaliation[25]. For indirect reciprocity, punishment works through reputation and also includes third-party actions, which means that observers of an interaction are willing to punish defectors at a cost to themselves[21]. Therefore, any discussion of the evolution of costly punishment brings us immediately into the framework of direct or indirect reciprocity.

In general, the reputation score could be a continuous variable[5], but here we consider a simple model with binary reputation. People have either a good reputation (G) or a bad reputation (B). At times, two random players are chosen from the population, one in the role of donor, the other in the role of recipient. The donor can either cooperate (C), defect (D) or punish (P). Cooperation means the donor pays a cost c, and the recipient gets a benefit b. Punishment implies that the donor pays a cost α and the recipient incurs a cost β. For defection there is no cost and no benefit.

The interaction between the donor and the recipient is observed by the other members of the population (Fig. 1). The reputation of the donor is updated according to a social norm. First-order assessment depends only on the action of the donor; for example, cooperation leads to a good reputation, whereas defection leads to a bad reputation. Second-order assessment[12,13] depends both on the action of the donor and the reputation of the recipient: for example, it could be deemed 'good' to cooperate with a good recipient but 'bad' to cooperate with a bad recipient. Here we study social norms that use second-order assessment. The donor has three possible moves (C, D or P) and the recipient has one of two reputations (G or B). There are therefore six combinations and $2^6 = 64$ social norms with second-order assessment. All detailed calculations are shown in the Supplementary Information.

Any interaction leads to either a good or a bad reputation for the donor. We assume that this process of reputation updating is subject to errors. There may be wrong observations or the spread of false rumours. With probability μ an incorrect reputation is assigned and adopted by all. In the simplest model, everyone has the same opinion of everyone else. There are no private lists of reputation. Triggering a wrong reputation affects everyone equally. The parameter $q = 1 - 2\mu$ quantifies the ability of the population to distinguish between good

[1]Department of Value and Decision Science, Tokyo Institute of Technology, Tokyo 152-8552, Japan. [2]PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan. [3]Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka 812-8581, Japan. [4]Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, Department of Mathematics, Harvard University, Cambridge, Massachusetts 02138, USA.

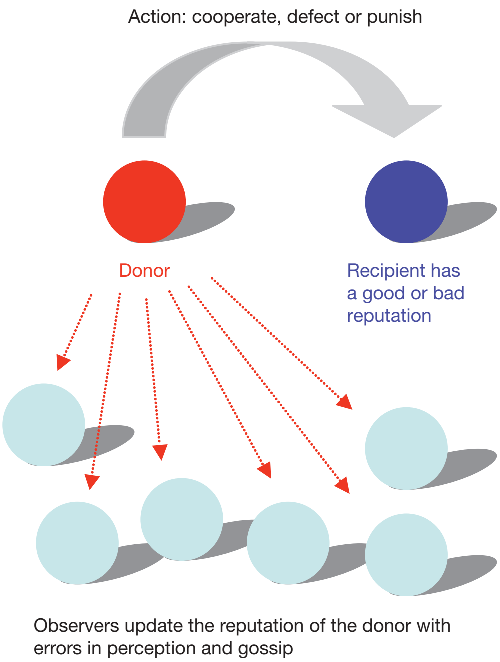

Figure 1 | Indirect reciprocity with costly punishment. The donor chooses one of three actions: cooperate, defect or punish. The recipient has a binary reputation, which is either 'good' or 'bad'. The donor's choice depends on the recipient's reputation. The donor's action is observed by other members of the population, who update the donor's reputation according to a social norm, which is shared by all. The donor is assigned an incorrect reputation with probability μ. The 'social resolution', $q = 1 - 2\mu$, is a key parameter of indirect reciprocity: it defines the ability to distinguish between good and bad.

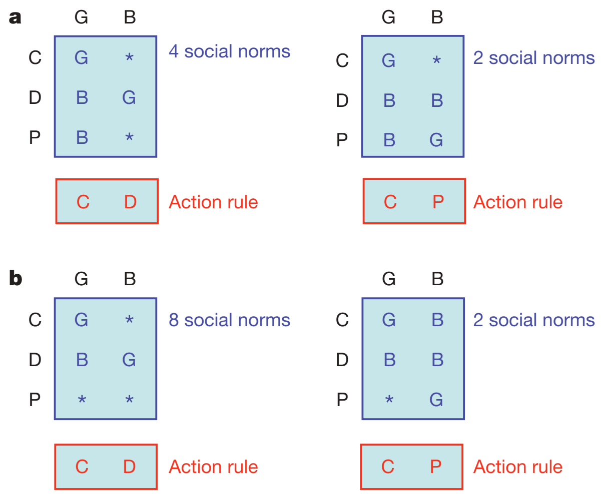

Figure 2 | Social norms of cooperation. We have determined all social norms that allow the evolutionary stability of action rules prescribing cooperation with good recipients. There are two such action rules, CD and CP. The former 'punishes' bad recipients by defection (D); the latter uses costly punishment (P). In **a** the cost of cooperation exceeds the cost of punishment; in **b** the cost of cooperation is less than the cost of punishment. There is an intuitive summary of all successful social norms: following the action rule maintains a good reputation; deviating from it can lead to a good or bad reputation; a deviation that is less costly than the prescribed action must lead to a bad reputation. An asterisk denotes G or B. The detailed parameter requirements for evolutionary stability are given in the text and in the Supplementary Information.

and bad. We call q the 'social resolution'. If $\mu = 1/2$ then reputation is assigned at random, and there is no ability to distinguish between good and bad ($q = 0$).

Games of indirect reciprocity contain social norms and action rules. The action rule specifies for the donor whether to cooperate, defect or punish a recipient who is either good or bad. For example, the action rule CD prescribes cooperation with a good recipient and defection with a bad recipient; this rule does not use costly punishment. In contrast, the action rule CP prescribes cooperation with good recipients and punishment of bad recipients. The action rules CC, DD and PP encode, respectively, unconditional cooperation, defection and punishment. In total, there are nine possible action rules.

For each of the 64 social norms we study the competition of all nine action rules. We assume that everyone in the population has the same social norm, and we evaluate whether this norm allows the evolutionary stability of action rules that specify cooperation with good recipients. There are only two candidates for such action rules, CD and CP, because CC is not stable against invasion by defectors (DD). Figure 2 shows all social norms that allow the evolution of cooperation. The action rule DD is evolutionarily stable for any social norm.

Social norms that stabilize the CD action rule have the following properties: (1) cooperation with a good recipient leads to a good reputation; (2) defection against a good recipient leads to a bad reputation; and (3) defection against a bad recipient leads to a good reputation. The three remaining positions in the norm can be either G or B. If the cost of cooperation is greater than the cost of punishment ($c > \alpha$), then punishing a good recipient must lead to a bad reputation; otherwise a donor can keep a good reputation by using the cheaper punishment option instead of the more expensive cooperation move.

Social norms that stabilize the CP action rule have the following properties: first, cooperation with a good recipient leads to a good reputation; second, defection always leads to a bad reputation; and third, punishing a bad recipient leads to a good reputation.

CD action rules are evolutionarily stable, if the social resolution exceeds the cost-to-benefit ratio ($q > c/b$). In contrast, CP action rules are evolutionarily stable if $q > \max\{c,\alpha\}/(b + \beta)$. Note that costly punishment can stabilize cooperation even if $q < c/b$. Thus, costly punishment can in principle extend the stability range of cooperation. DD action rules are always evolutionarily stable.

We have performed computer simulations in heterogeneous populations of finite size to test the validity of our analytical calculations. We find that the CD and CP action rules are stable against invasion by other action rules under the appropriate social norms and given the right parameter region. In the simplest simulations everyone has the same information about the reputation of others. In the extended simulations we drop this assumption. Now there are individual errors in assigning reputation. Consequently everyone has a private list of the reputation of others. These errors can destroy indirect reciprocity unless there is a mechanism for re-establishing coherence. Gossip is such a mechanism. We assume that individuals talk to each other and sample each other's opinions (as in a 'voter model'). If there are enough communication events, then we observe the evolutionary stability of our strategies as predicted. We have also studied errors resulting in execution of the wrong action ('trembling hand') or recalling an incorrect reputation ('fuzzy mind'). Our results are robust as long as these errors are not too frequent. All simulations are described in the Supplementary Information.

For some parameter regions, multiple action rules are evolutionarily stable. We therefore ask the following question: for all possible parameter regions, which of the three action rules CD, CP and DD are stable, and which one is the most efficient in the sense of leading to the highest average payoff at equilibrium? We obtain the following answer. (1) If $q > c/b$, then CD is most efficient. (2) If $c/b > q > c/(b + \beta)$ then CP is stable and more efficient than DD, if the following two conditions hold:

$$q > \frac{\alpha}{b+\beta} \quad \text{and} \quad q > \frac{\alpha+\beta-b+c}{\alpha+\beta+b-c} \qquad (1)$$

Otherwise DD is more efficient than CP. If $b < c$, then DD is always more efficient than CP. (3) If $c/(b+\beta) > q$, then only DD is evolutionarily stable.

Thus, if the accuracy of assigning the correct reputation, q, is too low, then only DD is efficient. If q is sufficiently large, then CD is efficient. For intermediate values of q there can be a small window where CP is efficient. However, the existence of this parameter region depends on whether the key parameters b, c, α and β fulfil the constraints given by equation (1). Let us consider a numerical example. If $b=2$, $c=1$, $\alpha=1/2$ and $\beta=2$, then CD is efficient for $q>1/2$, whereas CP is efficient for $1/2>q>3/7$ and DD is efficient for $3/7>q$. If we increase the effect of punishment to $\beta=5/2$ (or larger), there is no region left where CP is efficient. Intuitively, if CD is evolutionarily stable, it is always the most efficient equilibrium. If it is not stable, the remaining parameter region where CP is stable and more efficient than DD is very small or non-existent. Figure 3 illustrates the narrow margin of efficiency of costly punishment.

These considerations of efficiency do not imply that all populations will evolve towards punishment-free action rules. It is possible that a population is stuck at an inefficient equilibrium for a long time. A model with contingent movement allows us to study the competition of different social norms. We examine a simple model in which two groups have two different social norms. One norm stabilizes CD, whereas the other norm stabilizes CP. People interact only within their own group, but sometimes they compare their payoff with individuals from the other group. People might move to the other group and adopt its social norm if they find that its members have a higher payoff. We observe rapid selection of the efficient equilibrium (see Supplementary Information).

In an experimental study, the observers of a Prisoner's Dilemma game between two other people sometimes punish defectors at a cost to themselves[21]. This behaviour is a form of indirect reciprocity. In another experiment[23], a public goods game is followed by one round of punishment and then by one round of cooperation or defection. This setup is not directly comparable with our model, but the observation is that adding the third round reduces the amount of punishment that is being used in the second round. This particular finding is in agreement with our result: other possibilities of indirect reciprocity reduce the amount of costly punishment. In the context of our theory it would be important to extend both experiments to permit reputation-building over multiple rounds of interaction and a choice between cooperation, defection and costly punishment in every round. We predict that such an experiment will show that costly punishment is an inefficient behaviour for most parameter regions.

We have studied the effect of costly punishment in an explicit model of indirect reciprocity. We have analysed all social norms that use binary reputation and second-order assessment. We find that both CD and CP action rules can stabilize cooperation. These rules reward good recipients with cooperation and 'punish' bad ones with either defection (CD) or costly punishment (CP). If both CD and CP action rules are evolutionarily stable, the use of costly punishment leads to a lower equilibrium payoff and is therefore inefficient. It is even possible that costly punishment yields a lower payoff than all-out defection (DD). Costly punishment maximizes the group average payoff for only a very limited parameter region. This narrow margin of efficiency requires fine-tuning of the key parameters. If the social resolution exceeds the cost to benefit ratio ($q > c/b$), CD rules are always more efficient than CP rules. The evolution of improved mechanisms of indirect reciprocity therefore leads to societies in which costly punishment between individuals is not an efficient behaviour for promoting cooperation.

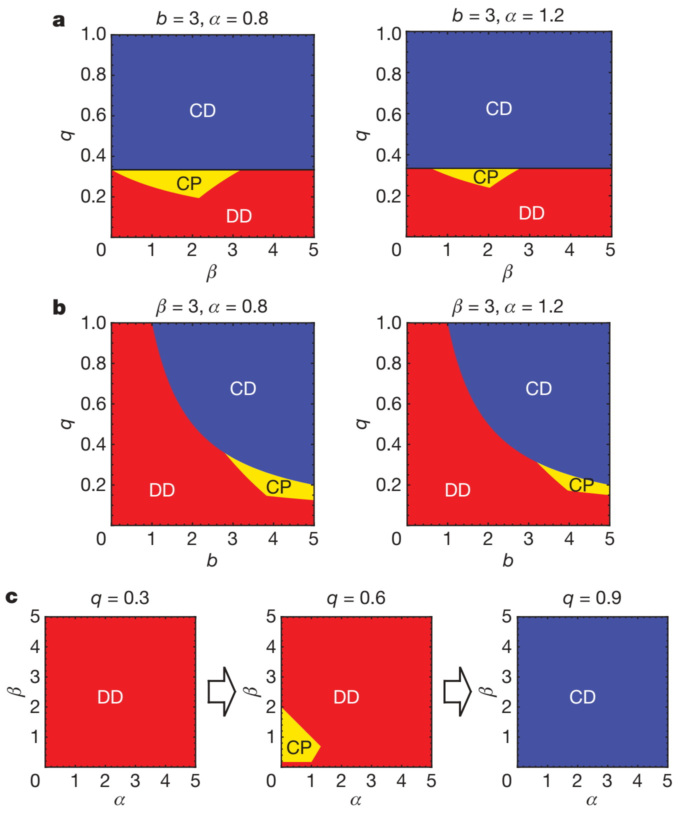

Figure 3 | The marginal efficiency of costly punishment. Projections of the five-dimensional (b, c, α, β and q) parameter space onto various two-dimensional planes. The parameters b and c denote the benefit and cost of cooperation. The parameters α and β denote the cost and effect of punishment. The social resolution of the system is given by q, the probability of distinguishing between good and bad in a world where errors in assignment of reputation are possible. The symbols CD, CP and DD represent the region where the corresponding action rule is the most efficient equilibrium, which means the evolutionarily stable strategy with the highest average payoff. CD means cooperation with good recipients and defection with bad ones. CP means cooperation with good recipients and punishment of bad ones. DD is unconditional defection. Costly punishment is an efficient equilibrium only for a very constrained parameter region (shown in yellow). **a**, Projections on the β–q plane for $b=3$ and $\alpha=0.8$ or 1.2. **b**, Projections on the b–q plane for $\beta=3$ and $\alpha=0.8$ or 1.2. **c**, Projections on the α–β plane for $b=1.5$ and $q=0.3$, 0.6 or 0.9. We always use $c=1$.

METHODS SUMMARY

An action rule s is formulated as a mapping from {G, B} (the recipient's reputation) to {C, D, P} (the prescribed action). A social norm n is a mapping from the product of {C, D, P} (the donor's action) and {G, B} (the recipient's reputation) to {G, B} (the donor's new reputation). We search for the combination of an action rule s and a social norm n that satisfies the following two properties: first, the monomorphic population in which all players adopt s and n achieves full cooperation in the absence of errors; and second, the action rule s is evolutionarily stable under the social norm n. We check these two criteria for each of all $9 \times 64 = 576$ possible combinations of action rule and social norm (s, n). From the first criterion, action rule s must use cooperation (C). Because of the symmetry in the binary labels G and B we can assume without loss of generality that the action rule prescribes cooperation to good recipients; that is, $s(G) = C$. To study the evolutionary stability of the action rule s, we use dynamic optimization. We assume that the social norm is n and that all players except the focal player adopt action rule s. Under this assumption we calculate the best-response action rule s^* of the focal player. If s^* exists uniquely and matches s, then s is evolutionarily stable under n. Coexistence of different action rules[10] is not within the scope of our analysis. See the Supplementary Information for further details.

Received 11 June; accepted 3 November 2008.

1. Sugden, R. *The Economics of Rights, Cooperation and Welfare* (Blackwell, 1986).
2. Alexander, R. D. *The Biology of Moral Systems* (Aldine de Gruyter, 1987).
3. Kandori, M. Social norms and community enforcement. *Rev. Econ. Stud.* **59**, 63–80 (1992).
4. Okuno-Fujiwara, M. & Postlewaite, A. Social norms and random matching games. *Games Econ. Behav.* **9**, 79–109 (1995).
5. Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577 (1998).

6. Wedekind, C. & Milinski, M. Cooperation through image scoring in humans. *Science* **288,** 850–852 (2000).
7. Dufwenberg, M., Gneezy, U., Güth, W. & van Damme, E. Direct vs indirect reciprocity: an experiment. *Homo Oecon.* **18,** 19–30 (2001).
8. Fishman, M. A. Indirect reciprocity among imperfect individuals. *J. Theor. Biol.* **225,** 285–292 (2003).
9. Ohtsuki, H. & Iwasa, Y. How should we define goodness?—reputation dynamics in indirect reciprocity. *J. Theor. Biol.* **231,** 107–120 (2004).
10. Brandt, H. & Sigmund, K. The logic of reprobation: assessment and action rules for indirect reciprocation. *J. Theor. Biol.* **213,** 475–486 (2004).
11. Bolton, G. E., Katok, E. & Ockenfels, A. Cooperation among strangers with limited information about reputation. *J. Public Econ.* **89,** 1457–1468 (2005).
12. Brandt, H. & Sigmund, K. Indirect reciprocity, image-scoring, and moral hazard. *Proc. Natl Acad. Sci. USA* **102,** 2666–2670 (2005).
13. Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity. *Nature* **437,** 1291–1298 (2005).
14. Suzuki, S. & Akiyama, E. Reputation and the evolution of cooperation in sizable groups. *Proc. R. Soc. Lond. B* **272,** 1373–1377 (2005).
15. Chalub, F. A. C. C., Santos, F. C. & Pacheco, J. M. The evolution of norms. *J. Theor. Biol.* **241,** 233–240 (2006).
16. Takahashi, N. & Mashima, R. The importance of subjectivity in perceptual errors on the emergence of indirect reciprocity. *J. Theor. Biol.* **243,** 418–436 (2006).
17. Pacheco, J. M., Santos, F. C. & Chalub, F. A. C. C. Stern-judging: a simple, successful norm which promotes cooperation under indirect reciprocity. *PLoS Comp. Biol.* **2,** 1634–1638 (2006).
18. Yamagishi, T. Seriousness of social dilemmas and the provision of a sanctioning system. *Soc. Psychol. Q.* **51,** 32–42 (1988).
19. Clutton-Brock, T. H. & Parker, G. A. Punishment in animal societies. *Nature* **373,** 209–216 (1995).
20. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415,** 137–140 (2002).
21. Fehr, E. & Fischbacher, U. Third-party punishment and social norms. *Evol. Hum. Behav.* **25,** 63–87 (2004).
22. Fowler, J. H. Altruistic punishment and the origin of cooperation. *Proc. Natl Acad. Sci. USA* **102,** 7047–7049 (2005).
23. Rockenbach, B. & Milinski, M. The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444,** 718–723 (2006).
24. Sigmund, K. Punish or perish? Retaliation and collaboration among humans. *Trends Ecol. Evol.* **22,** 593–600 (2007).
25. Dreber, A., Rand, D. G., Fudenberg, D. & Nowak, M. A. Winners don't punish. *Nature* **452,** 348–351 (2008).
26. Trivers, R. L. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46,** 35–57 (1971).
27. Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211,** 1390–1396 (1981).
28. Colman, A. M. *Game Theory and Its Applications in the Social and Biological Sciences* (Routledge, 1995).
29. Rutte, C. & Taborsky, M. The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): direct vs generalised reciprocity. *Behav. Ecol. Sociobiol.* **62,** 499–505 (2008).
30. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314,** 1560–1563 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Support from the John Templeton Foundation, the Japan Society for the Promotion of Science, the NSF/NIH joint program in mathematical biology (NIH grant R01GM078986) and J. Epstein is gratefully acknowledged.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.O. (ohtsuki.h.aa@m.titech.ac.jp).