Contents lists available at SciVerse ScienceDirect



Journal of Theoretical Biology



journal homepage: www.elsevier.com/locate/yjtbi

Limited memory can be beneficial for the evolution of cooperation

Gergely Horváth^{a,*}, Jaromír Kovářík^b, Friederike Mengel^{c,d}

^a School of Public Administration, Southwestern University of Finance and Economics, 610074 Chengdu, Sichuan, China

^b Dpto. Fundamentos Análisis Económico I & BRiDGE, Universidad del País Vasco, Av Lehendakari Aguirre 83, 48015 Bilbao, Spain

^c School of Economics, University of Nottingham, University Park Campus, Nottingham NG7 2RD, United Kingdom

^d Department of Economics (AE1), School of Business and Economics, Maastricht University, PO Box 616, 6200MD Maastricht, The Netherlands

ARTICLE INFO

Article history: Received 7 September 2010 Received in revised form 20 January 2012 Accepted 23 January 2012 Available online 1 February 2012

Keywords: Evolution Reputation Bounded memory Cooperation

1. Introduction

In this paper we study the effect of limited memory on the evolution of cooperation. Memory can be crucial for the emergence of cooperation through reputation-based mechanisms, where people form expectations about others by learning from their own experience as well as from third-party information (see e.g. Silk, 2006). Ex ante it seems that the larger the human memory capacity the better the chances for cooperation, simply because being able to process more information should allow agents to learn more effectively about other agents' types. This is all the more so, if agents rely on indirect reputation-building as well, which can increase the effectiveness of reputation mechanisms, but also has a larger demand for memory capacity than first-hand experience alone.¹ Sufficient memory capacity thus seems to be a key requirement for allowing reputation-based mechanisms to successfully establish cooperative societies (Trivers, 1971; Nowak and Sigmund, 1998).

Since Miller's (1956) "magical number seven +/- two" it has been widely accepted, though, that human working memory is very limited (Cowan, 2001).² Hence, if larger memory provides an evolutionary advantage either individually or at the population level (by

ABSTRACT

In this study we analyze the effect of working memory capacity on the evolution of cooperation and show a case in which societies with strongly limited memory achieve higher levels of cooperation than societies with larger memory. Agents in our evolutionary model are arranged on a network and interact in a prisoner's dilemma with their neighbors. They learn from their own experience and that of their neighbors in the network about the past behavior of others and use this information when making their choices. Each agent can only process information from her last h interactions. We show that if memory (h) is too short, cooperation does not emerge in the long run. A slight increase of memory length to around 5–10 periods, though, can lead to largely cooperative societies. Longer memory, on the other hand, is detrimental to cooperation in our model.

© 2012 Elsevier Ltd. All rights reserved.

building more cooperative societies through more effective use of reputation-based mechanisms), why do not we observe an evolution toward larger working memories in humans?

In this study we ask whether there is some evolutionary advantage of limited memory capacity (other than the obvious saving of energy or complexity costs). In particular we analyze the effect of working memory on the evolution of cooperation and show a case in which societies with strongly limited working memory achieve higher levels of cooperation than societies with larger memory. Agents in our evolutionary model are arranged on a network and interact in a prisoner's dilemma. Individuals learn from their own experience and that of their neighbors in the network about the past behavior of others and use this information to make their choices. Each agent can only process information from the last *h* interactions. Our agents are heterogenous in terms of the strategies they employ, and evolution selects for strategies with higher evolutionary fitness. Evolutionary fitness is determined by the order of payoffs in the prisoner's dilemma. We show that if memory is too short, cooperation does not emerge in the long run. A slight increase of memory length to around 5-10 elements, though, can effectively build up largely cooperative societies. Longer memory, on the other hand, is detrimental to cooperation.

The paper is organized as follows. In Section 2 we present the model. In Section 3 we discuss our results concerning memory. In Section 4 we consider populations with heterogenous memory and discuss individual level selection of *both* memory and behavioral types. Section 5 discusses related literature. Section 6 is dedicated to a discussion of our results and of some of the modeling assumptions. Some additional tables, information about

^{*} Corresponding author. Tel.: +86 138 801 79956.

E-mail address: horvathgergely@gmail.com (G. Horváth).

¹ Direct reputation is built through own experience (Axelrod, 1984; Trivers, 1971), while indirect reputation is also learned through communication with others (Nowak and Sigmund, 1998; Sommerfeld et al., 2007).

² In a related framework, Milinski and Wedekind (1998) find working memory capacities of the same order of magnitude to be used by human subjects in an experiment.

^{0022-5193/\$ -} see front matter \circledcirc 2012 Elsevier Ltd. All rights reserved. doi:10.1016/j.jtbi.2012.01.034

simulations as well as extensions of the model can be found in Appendix A.

2. The model

The prisoner's dilemma: There is a population of individuals, $i \in \{1, 2, ..., N\}$, who are repeatedly matched to play a prisoner's dilemma game. The payoff of an agent when she chooses action a_i and her opponent chooses a_j is given by

$a_i \mid a_j$	С	D
С	а	0
D	1	d

with 1 > a > d > 0. We assume that agents are of one of these types:

- *D*: defectors always defect.
- *CC*: conditional cooperators cooperate whenever they believe that the probability that their current round match cooperates is large enough (at least $\frac{1}{2}$ given the parameters used in our main simulation).³
- A: altruists always cooperate.

Matching: Agents are organized on a fixed undirected network, which mediates who meets whom in the population. In each period agents play the prisoner's dilemma game (1) with (some of) their direct neighbors in the network. Each player plays the game at least once in each period. And for each player, there is an equal probability to meet each of her neighbors.

Reputation and beliefs: We assume that agents have bounded working memory and are only able (or willing) to remember and (simultaneously) process information from the last h interactions.⁴ They then form beliefs about the behavior of their match using their experience from these last h interactions. In addition, they can use information they get from their direct neighbors in the network.

Denote by $\gamma_{ij}(h)$ the fraction of times (between 0 and 1) that j cooperated with i in an interaction between i and j that took place among i's last h interactions. Hence $\gamma_{ij}(h) = 1$ if j cooperated always when he met player i in and if they met at least once i's last h interactions. $\gamma_{ij}(h) = 0$ if i and j interacted at least once among i's last h interactions, but j never cooperated. And if there was no interaction between i and j among i's last h interactions (which are all those that i can cognitively process), then we set $\gamma_{ij}(h) < 0$ as a matter of convention. Hence, γ refers to own experience. Denote by $\overline{\beta}_{ij}(h)$ the average of the respective statistics among i's neighbors: the average of number of times j cooperated with a neighbor of i divided by the number of times they interacted in the neighbor's last h interactions. Hence, β refers to information from neighbors. All neighbors with some

Table 1Benchmark parameter values.

Parameter	Description	Value
Ν	Population size	100
а	Payoff from mutual cooperation	0.80
d	Payoff from mutual defection	0.50
λ	Weight put on own experience	0.50
k	Selection parameter	20
ρ	Connection radius	4
h	Memory lengths	$\{1, 2, \dots, 20\}$
θ	Rewiring probabilities	0,0.01,0.05,1

information are weighted equally and neighbors without information are not considered when taking the average. Again we set $\overline{\beta}_{ij}(h) < 0$ if none of *i*'s neighbors has interacted with *j* in their last *h* interactions. The reputation that player *j* has for player *i* (what *i* thinks about *j*) at time *t* is then given by

$$r_{ij}^{t} = \begin{cases} \lambda \gamma_{ij}(h) + (1-\lambda)\overline{\beta}_{ij}(h) & \text{if } \gamma_{ij}(h) \ge 0 \land \overline{\beta}_{ij}(h) \ge 0\\ \gamma_{ij}(h) & \text{if } \gamma_{ij}(h) \ge 0 \land \overline{\beta}_{ij}(h) < 0\\ \overline{\beta}_{ij}(h) & \text{if } \gamma_{ij}(h) < 0 \land \overline{\beta}_{ij}(h) \ge 0\\ \overline{\sigma}^{t-1} & \text{if } \gamma_{ij}(h) < 0 \land \overline{\beta}_{ij}(h) < 0 \end{cases}$$
(2)

In words, the reputation that *j* enjoys with *i* is a weighted average of her direct reputation with $i(\gamma_{ij}(h))$ and her indirect reputation communicated from *i*'s neighbors ($\overline{\beta}_{ij}(h)$). A high value of λ means that she relies mostly on her own experience and low λ that she forms judgements based mainly on the information from others. If *i* has not met *j* in the last *h* periods, but at least one of her neighbors has, she relies on neighbors' experience alone and viceversa. If nothing is known about *j* (i.e. if neither *i* nor any of her neighbors have information about *j*) there is "no reputation". In this case agents use the average rate of cooperation in period t-1 ($\overline{\sigma}^{t-1}$), which is assumed to be always known to all players.

One could imagine that in a more sophisticated model agents form reputations by also considering the behavior or type of *j*'s matches (i.e. they may not judge defecting against a cooperator and defecting against a defector in the same way). See more on this point below.

Selection: Evolution selects among altruists, conditional cooperators and defectors. We are interested in which of the three types survive evolutionary selection, when cooperation will emerge, and how this depends on memory constraints. Evolutionary fitness is defined as the payoffs received in *one* of the games played in the last period. In Appendix A we describe how the algorithm selects such an interaction.⁵ The results are robust to considering the average payoffs from several periods or interactions in one period for fitness.

The selection process is modeled as follows. In each period *K* agents are called randomly for selection. For each of these agents k another agent $m \neq k$ is randomly chosen from the population and their payoffs from the last interaction are compared. If m has higher payoff, agent k adopts m's type. Note that this implies that only the order of payoffs matter, not the particular numerical values of the parameters a and d given in Table 1. This process could be interpreted as cultural evolution, which could take place e.g. via imitation learning. In this case the randomly chosen agent (see e.g. Mengel, 2008). However, since types are not observable,

³ This is the optimal behavior of e.g. an agent whose preferences are represented by matrix (1) but who suffers a psychological cost $w \in [1-a,d]$ each time she defects. Cooperation yields higher expected "payoffs" for such an agent whenever r*a > r(1-w)+(1-r)*(d-w), where r is the probability that the current round match cooperates. If w > d, then such an agent will behave like an altruist, i.e. will cooperate irrespective of the value of r. And if w < 1-a, then she will behave like a defector, i.e. will defect irrespective of the value of r. Under the assumption that a+d > 1, those are the only three possible types in such a model. See e.g. Mengel (2008).

⁴ We make no claim as to whether agents use only h periods to form their beliefs because (i) they are not able to remember or process more period or (ii) because they are not willing to do so because processing additional information is cognitively too costly given its additional information content. Hence, the limitation of h may also come from a sophisticated trade off between costs and benefits.

⁵ We chose to define fitness as the payoffs from one interaction rather than e.g. as the sum of payoffs from all interactions to rule out that degree of a player has an effect on fitness *per se.*



Fig. 1. Watts and Strogatz (1998) small-world networks (N=16 and ρ = 2) for three rewiring probabilities: θ = 0 (regular lattice, left), small but positive θ (small-world network, center), and θ = 1 (random network, right). See Section A.2, point 1.(a).ii for details about rewiring.

type changes can only be learned over time. Hence, we will keep the focus on the evolutionary interpretation of the model.

Networks: In the simulations, we use small-world networks (Watts and Strogatz, 1998). These networks are generated from an one-dimensional lattice, where agents have links to their neighbors up to a distance ρ , the connection radius. Starting from this network, each link is rewired with probability θ . As θ tends to 0, we get a regular lattice; as θ becomes positive, distances shorten dramatically, while the clustering coefficient is almost unaffected; while as θ tends to 1 a random network is obtained (see Fig. 1 showing network structures corresponding to these three cases). Since short distances and high clustering are typical for real-life social networks (Vega Redondo, 2007; Goyal, 2007), we focus on these "small-world transition" values (Watts and Strogatz, 1998). More details on small-world networks can be found e.g. in the textbooks by Vega Redondo (2007) or Goyal (2007).

3. Results

The dynamic process described in the previous section either converges to an absorbing state or to a class of states where only conditional cooperators are present (Proposition 1 in Appendix A). There are two types of absorbing states: one characterized by full cooperation and one by full defection. Altruists and defectors never survive together in the long run (Proposition 2). If only conditional cooperators survive, the system might or might not converge to an absorbing state.

In this section, we focus on the numerical analysis of the model. The benchmark parameter values used in the simulations are summarized in Table 1. (Note that – since we use only one interaction for fitness – the exact values for a and d do not matter. Only the order of payoffs in the PD does.) We run 100 simulations for each parameter setting. We start each simulation with the same number of individuals of each type. They are randomly allocated on the network. At the beginning, individuals do not have reputation for each other. Conditional cooperators' first action is to cooperate. In Appendix A, we provide a detailed description of the simulations.

The theoretical results are reflected in the simulations. For $\theta < 0.05$, we always observe a convergence toward any of the two absorbing states: either everybody cooperates or everybody defects in the population. Which state is reached in the long run depends on θ and h. Only if both θ and h are large enough we observe a small fraction of simulations that stabilizes around a positive fraction of cooperating individuals and positive fraction of defecting ones (see Fig. 4 in Appendix A).⁶ The general pattern in Fig. 4 is that the zero cooperation regimes emerge frequently



Fig. 2. The *y*-axis plots the average level of steady state cooperation across all runs; the *x*-axis shows the size of the memory constraint. The network structure used is a small-world network (Watts and Strogatz, 1998). In the simulations, we use four values for the rewiring probability: (i) $\theta = 0$ (blue), (ii) $\theta = 0.01$ (green), (iii) $\theta = 0.05$ (red), and (iv) $\theta = 1$ (light blue). The 95% confidence intervals are reported in Table 4 in Appendix A. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for low (1–2) and very high (above 10) levels of memory, while full cooperation regimes are more likely to emerge for intermediate memory lengths (between 2 and 10). States of intermediate cooperation levels appear only if the frequency of cooperative populations is very low.⁷ Since these scenarios are rare, the following results can in most cases be interpreted as the fraction of simulation runs that lead to full cooperation.

Fig. 2 shows the average level of long-run cooperation over the runs of each parameter constellation. We observe a non-trivial relation between steady state cooperation levels and memory length. Cooperation is rare for low memories. However, as the memory constraint increases from 2 to 5, there is a dramatic increase of cooperation. Hence, only a slight increase of memory span of agents is sufficient to achieve relatively cooperative societies. The stark increase of cooperation is followed by a relatively stable level for memory constraints between 5 and 10. For larger memory levels, cooperation starts to decrease. It seems

⁶ This occurs for h > 15 if $\theta = 0.05$ and for h > 9 if $\theta = 1$ (Fig. 4).

 $^{^7}$ In these cases, if θ = 0.05 the long-run cooperation levels lie between 20 and 90 depending on the run and remain stable, while for θ = 1 there is never more than 30% of cooperating individuals in the long run. There is no systematic formation of cooperative clusters that may resist the invasion of outside defectors, since the analyzed network topologies are small world networks. One of the characteristics of such networks is that there is large overlap of neighborhoods whenever $\theta < 1$.

to be optimal for a population if its members have (strong) bounds on their memory constraints. In fact, in (almost all) our simulations the memory level at which the highest levels of cooperation are achieved lies between 5 and 10 (see Appendix A).

The intuition behind the simulation results is as follows. If memory is too low, the vast majority of conditional cooperators cannot effectively learn the type of their opponents and hence are prone to exploitation by defectors. This is true as long as the amount of defection in the population does not exceed the critical level above which conditional cooperators will start to defect. In this case everyone in the society will defect except for some altruists who will be exploited by defectors and conditional cooperators and who will not survive evolutionary selection.

For intermediate levels of memory, conditional cooperators can prevent the extinction of altruists. Since, due to larger memory, conditional cooperators can largely differentiate between defectors and cooperative individuals, they will cooperate against altruists and other conditional cooperators and defect against defectors. As a consequence, cooperation can survive.

But why is even larger memory detrimental? As memory gets larger, cooperation is sustained by conditional cooperators. Long memory enables them to distinguish between altruists and defectors, but now there is a drawback. Initially, defectors have a selective advantage over altruists on average, because before reputations are established both types are treated equally by conditional cooperators. This means that there is a point at which the system consists of quite many defectors (see examples of this phenomenon in Section A.2 in Appendix A). This temporary selective advantage of defectors has long-lasting consequences now. The reason is that if there are many defectors which are identified as such, conditional cooperators will quite often defect against such defectors. But since memory is very long this will not easily be forgotten. Many conditional cooperators will earn a bad reputation among each other and will start defecting among themselves. Hence, large memory traps the system in full defection. Shorter memory is more forgiving and avoids this trap.

Why do not we assume that agents form reputations (of, say, agent j) conditional on the past behavior or type of j's matches in our model? Remember first that types are unobservable.⁸ Hence, it is impossible to know the type of j's past matches. In addition, it is impossible to infer whether a defecting agent defects because he is a defector or a conditional cooperator who has no information about her match and "bad beliefs" about her environment. Hence in order to condition their behavior on the "type" of the opponent agents would need a theory about how other agents form their beliefs about their matches, etc. In our model this would require agents to sample not only from their own and their neighbors experience about j's behavior, but also try to find out what j knows about themselves via joint neighbors. This is not only cognitively extremely complex, but also impossible if agents have only limited information about the network.

The results are robust if we allow for heterogeneity in agent's memory by modifying the model as follows: assume that each agent remembers all her interactions of the last h periods where each period consists of N interactions in the population, but people are drawn to play with replacement. In this case, some agents can be chosen more often than others. This means that *on average* agents remember h interactions but some agents may remember more and others less (see Section A.2 for details). Allowing for heterogeneity in this sense does not affect our results (see Section A.3).

Our findings are also robust to the introduction of mutations in the following sense. Assume that there is a small probability that an agent chosen for selection, rather than adopting the type of the individual she compares with, adopts a random type. This does not change our results qualitatively. The analogue of Fig. 2 with mutations is reported in Appendix A (see Section A.5). The results are furthermore robust to changes of other parameters of the model and modifications of model attributes (see Appendix A). Among others we show in Section A.3 that the results are robust even if we allow for agents being matched with others located further away in the social network.

Finally one could ask whether the results of the paper could also be obtained from a simpler version of the model. Additional simulation results illustrate that if we remove any mechanism (direct reputation, indirect reputation, or network-based meetings) from the model the long-run levels of cooperation decline dramatically or disappear completely.⁹ If only direct reputation is removed, the patterns of cooperation-maximizing memory lengths persist. However, the percentage of runs converging to cooperation lies well below the magnitudes observed in Fig. 2 (see Section A.4).

4. Evolution of memory

In order to check whether limited levels of memory are optimal from the point of view of individual selection *within* a population, we inject into the model conditional cooperators with different working memory capacities. We choose three memory levels: (i) h=2 (corresponding to full defection states in Fig. 2), (ii) h=7 (cooperation-maximizing memory levels), and (iii) h=15 (decreasing levels of cooperation). Thus in total we have five types in the population. In the simulation initially each type represents 20% of the population. The selection process still works at the level of types, but each memory level is treated as a different type here. We report here the results for the model with mutations: with a probability 0.02 the randomly chosen agent's fitness is not compared to the fitness of someone else, but she adopts a type uniform randomly.

Table 2 shows the average final type distributions and cooperation levels for different network structures over 100 runs of each parameter constellation. We observe very high cooperation levels in the long run, as long as the networks exhibit the smallworld property (i.e. $\theta < 1$ in the table; Watts and Strogatz, 1998). Defecting states have a hard time emerging in these environments (between 10% and 14% of population), as opposed to conditional cooperators who tend to survive, irrespective of their memory level (more than 75%).

Conditional cooperators with different memory levels have about the same share in the population in the long run. Hence, in terms of individual selection there are no strong evolutionary pressures favoring any of the memory sizes.¹⁰ The cooperationmaximizing memory capacity is not evolutionary costly within groups, while from a population viewpoint bounded memory leads to the largest levels of cooperation. These results are robust to disregarding mutations and changes in the initial conditions (see Appendix A for the case where there initially are 1/3 of altruists, 1/3 of defectors, 1/9 of each type of conditional cooperators).

⁸ If they were observable, then it is well known that cooperation can survive. See e.g. Bester and Guth (1998).

⁹ In terms of our model, removing network-based meeting corresponds to a complete network. In such a case, in any period people are matched to play with *any* other member of the population with equal probability. ¹⁰ Note that in an absorbing state, which almost always emerge, there is no

¹⁰ Note that in an absorbing state, which almost always emerge, there is no evolutionary pressure on types (and therefore memory levels).

Table 2

Evolution of memory with mutations. The second column provides the average rates of cooperation for each rewiring probability. Each cell in the third through the last column shows the average fraction of individuals of a certain type (columns) and rewiring probability (rows). *A*, *D*, *CC*₂, *CC*₇ and *CC*₁₅ state for altruists, defectors, conditional cooperators with memory of length 2, 7 and 15, respectively. Reported averages over 100 runs of each parameter constellation.

θ	Cooperation	Α	D	CC ₂	CC ₇	CC ₁₅
0	79.68	0.1150	0.1003	0.2599	0.2622	0.2623
0.01	77.48	0.1124	0.1096	0.2564	0.2598	0.2617
0.05	68.63	0.1009	0.1397	0.2361	0.2633	0.2597
1	11.03	0.0400	0.1955	0.1983	0.2535	0.3125

5. Related literature

We organize our discussion of related literature as follows. We start by discussing reputation-based models of cooperation, then move to models of memory and cooperation and finally discuss related models of cooperation in networks.

Reputation: In indirect and direct reciprocity models, people condition their actions on the reputation. Assessing reputation through "image scoring" (Nowak and Sigmund, 1998), where individuals monitor and assign either a good or bad score to others according to their behavior toward third parties, has shown to stabilize cooperation. The literature provided an exhaustive analysis of different rules of reputation assessment (Ohtsuki and Isawa, 2007; Panchanathan, 2011; Corten and Cook, 2008; Raub and Weesie, 1990 and the references contained therein). Naturally, more complicated rules may require larger memory capacities. However, this literature has not explored the length of memory systematically and mostly focuses on the framework of indirect reciprocity. Roberts (2008) uses both direct and indirect reputation mechanisms. However, agents choose either direct or indirect reputation in his model, while individuals simultaneously use both types of information to assess partners' reputation in our model. We show in Appendix A (Section A.4) that relying only on direct reputation does not lead to the emergence of cooperation in our model, while relying on only indirect reputation leads to substantially lower rates of cooperation (while preserving the non-monotonic effect of memory). Unlike Roberts (2008) we also systematically vary the number of past interactions our agents remember to address our main research question-the effect of memory. Moreover, due to the role of the network in our study agents may be unaware of the reputation of a partner and the same individual may enjoy different reputation scores for different people.¹¹ These elements of social interactions relate to memory constraints in obvious and important ways. Both longer time intervals between potential cooperative encounters and having to recall the past play of multiple opponents are challenges for reputation-based mechanisms (Milinski and Wedekind, 1998; Stevens and Hauser, 2004). Other related work includes Nakamuru and Kawata (2004) who show that defectors can be identified in a model with noisy information and Sommerfeld et al. (2008) who study the evolution of trustworthiness in a model with gossip.

Memory and cooperation: The explicit role of working memory for the evolution of cooperation has already attracted some attention in the literature. Several studies have suggested that longer memory is beneficial (Hauert and Schuster, 1997; Kirchkamp, 2000). Qin et al. (2008) have investigated the effect of memory if agents are arranged on a square lattice. They have found that the density of cooperators was enhanced by an increasing memory effect for most parameters. Cox et al. (1999) explore whether memory about past interactions can compensate the fact that people do not have information about others. Aktipis (2006) studies the role of recognition memory for the emergence of cooperation. In her model agents can either remember agents that previously cooperated or agents that previously defected. She shows that the strategy of remembering cooperators (rather than defectors) may require less memory size to be able to invade the population. Janssen (2006) presents a reputation-based model under which agents have the possibility to provide feedback on positive or negative experiences. Unlike in our model whether feedback is communicated or not is endogenous and agents in his model play a prisoner's dilemma with an additional strategy called "withdraw". He conducts multiple simulations and shows that it is not likely that reputation scores alone will lead to high levels of cooperation. This is consistent with our result reported in Section A.4 in Appendix A, where we show that each of our model ingredients is crucial for the emergence of cooperation.

Some other studies have already suggested that more effective memory might actually be detrimental to cooperation. Both Qin et al. (2008) and Alonso-Sanz (2009) observe a non-monotonic effect of memory on cooperation. However, memory is used in a very different manner in their studies. It is not used to assess the reputation of others as in our study. Instead, agents imitate others that have the highest average payoff over a number of past periods. The non-monotonic effect in their model comes from the trade off of having more information with longer memory which is also less accurate. This is very different from our model, where agents use memory to learn about each other's type. The trade off in our model comes from the fact that longer memory can lead to a stigmatization of conditional cooperators as defectors. The reason is that with longer memory one wrong assessment of a conditional cooperator as defector can have large consequences, since conditional cooperators will start defecting against each other and hence assess each other as defectors. Shorter memory is more forgiving and avoids these traps. In addition Qin et al. (2008) observe this effect only for a very particular parameter constellation and this effect disappears in Alonso-Sanz (2009) if people solely remember the last two rounds of play, while this effect is very robust in the present study. More closely related, Janssen (2006) detects a gualitatively similar effect of memory length on cooperation. In his model, if memory is too short cooperation does not survive, whereas too large memory spans can lead to a modest decline of cooperation (see Fig. 1 in Janssen, 2006).¹² Since his model aims to test the reputation system in e-Bay online actions, reputation of individuals is common knowledge and people are matched randomly to play the game. Most importantly, in his framework agents have the option of not to play the game with an opponent. This option changes the strategic structure of prisoner's dilemma game and is known to enhance cooperation on its own (e.g. Izquierdo et al., 2010). All these studies suggest that the detrimental effect of more efficient memory abilities for the evolution of cooperation may be a more general phenomenon.

Cooperation in networks: There is a substantial body of literature on the emergence of cooperation in *fixed* social networks. Direct reputation-building within simple network architectures, such as circles and lattices, has been studied by Boyd and Richerson (1989), Eshel et al. (1998), and Nakamaru et al. (1997). More recently Ohtsuki et al. (2006) and Ohtsuki and Nowak (2007) have analytical results for regular graphs and simulation results for random and scale-free networks. Santos et al. (2006) stress the role of scale-free

¹¹ See also Ohtsuki and Isawa (2007).

¹² Quantitatively, the level of memory with largest level of cooperation in his model is an order of magnitude larger than ours.

networks. These studies conclude that network structures facilitate cooperation. Abramson and Kuperman (2001) study the evolution of cooperation on small-world networks, such as the networks studied in our model. However, there is no reputation-building in their model and they do not study the effect of memory.

Concerning indirect reputation, Mohtashemi and Mui (2003) explicitly focus on the effect of social information that travels through network on cooperation. Direct links mutually share information in their model, spreading reputation of individual agents. In contrast to the present model, they rule out repeated interactions. They show that non-direct assessment of reputation itself can stabilize cooperative behavior. However, they model a growing network, in which everybody ends up knowing everybody, which is crucial for the survival of cooperation in their setting (see Mohtashemi and Mui, 2003, p. 527).

Raub and Weesie (1990) demonstrate the existence of network effects in reputation-based systems by comparing the extreme cases of "atomized" (direct reputation) with "perfectly embedded" (apart from direct reputation assessment, all actors are immediately informed about all interactions of their partners with third parties) interactions. They show that efficiency, i.e. mutual cooperation in a prisoner's dilemma, is more easily obtained in embedded systems. This is consistent with our analysis in Section A.4 where we show that cooperation breaks down in the absence of either a network structure or an indirect reputation mechanism.

The co-evolution of cooperation and networks structure is studied in Hanaki et al. (2007), Biely et al. (2007), Corten and Cook (2008)Nakamaru (2006), Fosco and Mengel (2011), Zimmermann et al. (2004), or Ebel and Bornholdt (2002) among others. Fosco and Mengel (2011) study cooperation in an endogenous network theoretically and show that in absorbing states there is either "separation" of defectors, i.e. two disconnected components emerge one with defectors and one with cooperators or there is "marginalization" of defectors, i.e. one connected component emerges in which cooperators are in more central positions than defectors. More closely related is the simulation study by Corten and Cook (2008). They also model agents interacting in an endogenous network (just as Fosco and Mengel, 2011) and like in this paper their agents form expectations about the behavior of others via direct and indirect reputation. They find that reputation does not always foster cooperation and show that network cohesiveness is more likely a consequence of cooperation rather than a cause. The latter result is consistent with evidence found in Fosco and Mengel (2011). Unlike in this paper, their networks are endogenous and they do not study the effect of memory. However, their result that reputation does not always foster cooperation is consistent with our robustness analysis conducted in Section A.4 in Appendix A.

6. Discussion

Our results provide some new viewpoints on the co-evolution of cooperation and memory or more loosely speaking on why evolution did not make us "infinitely" smart. Undoubtedly, there are other obvious reasons (such as energetic/reasoning costs) for why human memory is limited, but we show one example where long memory need not even be optimal in the absence of such costs. We have seen that societies with limited memory length may achieve higher rates of cooperation than others. Hence, while we do not model group selection explicitly (see e.g. Boyd and Richerson, 1990), limited memory in our model could emerge from the conflict of two populations endowed with different memory capacity. This would only be true however if societies that achieve higher rates of cooperation outperformed others with lower rates of cooperation. Since human memory capacity is in reality limited, our results are suggestive for future research studying how and whether human cognitive capacities, such as memory, may have coevolved with cooperation. Tomasello (2008) for example argues that human capacities such as communication have evolved jointly with cooperation.

Note that there are multiple kinds of memory and other forms of memory (such as short-term/working vs. long-term memory, sensory memory, episodic memory or recognition memory) may play a role for the evolution of cooperation.¹³ In the present model we define memory in a way that is guite standard in Economics and Game Theory (see e.g. Mailath and Samuelson, 2006: Sarin, 2000). Our definition of memory corresponds most closely to short term or working memory. Working memory is the ability to actively hold information in the mind needed to do complex tasks such as reasoning, comprehension and learning. Its capacity has been shown to be limited in a number of studies. Miller (1956) for example conducted studies, where he showed that the memory for chunks of information such as strings of letters or digits is limited. Cowan (2001) has proposed that working memory has a capacity of about four chunks in young adults (and fewer in children and old adults). The capacity of long term memory is typically considered to be immeasurably large. Sensory memory on the other hand refers to approximately the first 200-500 ms after an item is perceived. Clearly neither longterm memory nor sensory memory are the type of memory that we have in mind but rather short term or working memory. There are even finer categorizations of memory types.

In the present framework memory is semantic rather than episodic in that it concerns facts independent of context instead of relating them to a particular time and place, but this is simply because "time and place" are abstracted from in our model. In applications of the model one would probably want to think about episodic memory instead, because usually one would be thinking of personal memories rather than abstract notions. We do not model episodic memory more explicitly, because our model is abstract and we feel this would be orthogonal to our research question.

It should also be noted that there is little in the model that hinges on this definition of memory. The essential insight is that with larger memory there can be drawbacks to reputation based systems because an initial fitness advantage of defectors will entrap conditional cooperators in defection not only against defectors, but also among themselves. This essential insight seems to go through even if we had modeled, e.g. episodic memory instead of working memory capacity. In future research, however, it could be very interesting to model how these different types of memory interplay in fostering cooperative relations.

We also rely on a rather simple form of reputation. Cognitively more complicated strategies of reputation assessment can have larger requirements on memory capacity. Hence, it would be interesting to study the co-evolution of memory constraints and strategies which possibly require different memory levels.

We have also found that there is a certain memory threshold, under which large-scale cooperation does not emerge. Primatologists agree that primates have lower working memory capacity than humans (Kawai and Matsuzawa, 2000; Premack and Premack, 2003). How much cooperation there is in other animals is an area of controversy (see e.g. the discussion in Silk, 2006 or examples of cooperation even among the simplest animals in Crespi, 2001; West et al., 2006). Interestingly, though, some evolutionary psychologists believe that the era in which our ancestors seem to have surpassed the working memory capacity

¹³ We thank Athena Aktipis for pointing this out.

of today's primates coincides with the rise of more complex forms of social organization around 80,000 years ago (see e.g. Read, 2008), which would naturally require different levels of organization and cooperation. Since longer memory provides no evolutionary advantage in terms of individual selection in our model but can be detrimental for a society, there may be good reasons why human memory is strongly limited. Of course, human memory capacity also has non-social functions and potential trade-offs have to be taken into account. In addition, other forms of memory (such as e.g. episodic memory) are relevant for human cooperation. Hence one has to be careful when interpreting our results in the light of different findings from the cognitive sciences. There is large scope for future research in this area.

Acknowledgements

We are grateful to Hubert János Kiss, Marco van der Leij, Nick Vriend, an anonymous reviewer, Athena Aktipis and many seminar participants. Jaromír Kovářík acknowledges the financial support from the Basque Government (IT-223-07) and the Spanish Ministry of Science and Innovation (ECO2009-09120) and Friederike Mengel thanks the Dutch Science Foundation (NWO) for financial support.

Appendix A

A.1. Theoretical results

100

90

80

70

60

50

40

30

20

10

0

0

50

100

Time

150

In this subsection we would like to (partially) characterize absorbing states of our model analytically. To do this we need to introduce some notation. First note that at each point in time teach player is entirely characterized by her type, the vector of her

Altruists

Defectors Cooperation

Cond Cooperators

Memory = 5

reputations for others $r_i(t) := (r_{i_1}^t, r_{i_2}^t, \dots, r_{i_N}^t)$ as well as those of her first-order neighbors in the network. Consequently the population state at time *t* is given by the $n \times (n+2)$ matrix

$$\Xi(t) = (\tau(t) \| \pi(t) \| R(t))$$

where $\tau(t) = (\tau_1(t), \ldots, \tau_N(t))^T$ is the $n \times 1$ vector indicating the types of all players, $\pi(t)$ is the $n \times 1$ vector indicating each agent's selection relevant payoff (i.e. her payoff from the last interaction as player *i*) and $R(t) = (r_1(t), \ldots, r_N(t))'$ is the $n \times n$ matrix indicating all the player's reputations for each other, where $(r_1(t), \ldots, r_N(t))'$ indicates the transpose of $(r_1(t), \ldots, r_N(t))$. Note that since $\Xi(t)$ is completely determined by $\Xi(t-1)$ and the realizations of the random variables at *t*, the associated transition matrix describes a finite Markov chain on the state space $S := T \times \Pi \times \mathcal{R}$ (where \mathcal{R} is the set of all possible reputation matrices and Π is the payoff space). Denote the probability to reach state *s'* from state *s* by q(s,s'). We have the following definition.

Definition. State *s* is absorbing $\iff q(s,s) = 1$.

Our first result shows that the stochastic process does indeed converge to one of these absorbing states.

Proposition 1. Starting from any s, the stochastic process described above converges almost surely to either an absorbing state or to a recurrent class of states which contains only conditional cooperators.

Proof. We will show that there exists a number $K \in \mathbb{N}$ and a probability \hat{q} s.t. from any $s \in S$ the probability is at least \hat{q} to converge within K periods to an absorbing state or a recurrent class which contains only conditional cooperators. K and \hat{q} are time independent and state independent. Hence, the probability of not reaching an absorbing state after at least nK periods is at most $(1-p)^n$ which tends to zero as $n \to \infty$. Consider an arbitrary

Memory = 9



200

80

70

60

50

40

30

20

10

0

0

20

40

Time

60

80

state *s*(*t*). Denote by $i_{max}(t)$ the player with the highest payoff $\pi_i(t)$ from her last (selection relevant) interaction.

Assume first that $\tau(i_{\max}(t)) = D$. Afterwards, we will consider $\tau(i_{\max}(t)) = A$ or $\tau(i_{\max}(t)) = CC$. Assume $i_{\max}(t)$ was last matched with a defector. Then, all agents in the population must be either of type *CC* (choosing defection) or of type *D*. (Else $i_{\max}(t)$ cannot have the highest payoff. The reason is that any player matched with an agent that cooperates will have a higher payoff than $i_{\max}(t)$ irrespective of the action that player chooses.) Since all agents defect and all receive the same payoff, no agent will change their type. The reputation

matrix will converge and we reach an absorbing state with probability 1. Assume now that $i_{max}(t)$ was last matched with a cooperator denoted by $j(i_{max}(t))$. There is positive probability that all agents drawn for selection are cooperators, that furthermore they are the cooperators at largest geodesic distance from $i_{max}(t)$ and that all players k are matched with i_{max} for selection. Then, obviously at $t+1: \tau(k) = D \forall k$. There is positive probability that all agents (including i_{max}) will be matched with the same agents again during the next T periods, where T is chosen s.t. after T periods all agents choose defection. There is also positive probability that all players k are



Fig. 4. Frequency of cooperation scenarios over 100 runs. Blue line: zero cooperation, green line: full cooperation, red line: intermediate cooperation.



Fig. 5. Effect of locality of encounters ($\rho = 4$).

matched with i_{max} for selection in each of those *T* periods. Since in each selection period the cooperators at largest geodesic distance from $i_{max}(t)$ change their type, $i_{max}(t)$ will keep being the player with the highest average payoff during the transition. Then, all agents are choosing the same action, implying that all agents will obtain the same per interaction payoff. In finite time the reputation matrix will converge and we will have reached an absorbing state. Denote the probability with which this happens by $q_s > 0$.

Now consider the states where $\tau(i_{\max}(t)) = A$. Then $i_{\max}(t)$ must have been matched with either an altruist or a conditional cooperator that was cooperating. If not $i_{max}(t)$ could not have the highest possible payoff. There is positive probability that all agents drawn for selection are defectors, that furthermore they are the defectors at largest geodesic distance from $i_{max}(t)$ and that all players k are matched with i_{max} for selection. Then, obviously at t+1: $\tau(k) = A \forall k$. There is positive probability that all agents (including i_{max}) will be matched with the same agents again during the next T periods, where T is chosen s.t. after T periods all agents choose cooperation. There is also positive probability that all players k are matched with i_{max} for selection in each of those T periods. Since in each selection period the defectors at largest geodesic distance from $i_{max}(t)$ change their type, $i_{max}(t)$ will keep being the player with the highest average payoff during the transition. Then, all agents are choosing the same action, implying that all agents will obtain the same per interaction payoff. In finite time the reputation matrix will converge and we will have reached an absorbing state. Denote the probability with which this happens by $q_s > 0$ also for those states. Finally assume that $\tau(i_{max}(t)) = CC$. A very similar argument holds in this case except that it is possible that all agents are of type CC but do not end up choosing the same action in which case the reputation matrix may not converge. However, note that during such a transition the agent with the highest payoff will always be of type CC. If not we are in one of the two cases described above. This implies that starting from any state s the process either converges to an absorbing state or to a state where all agents are of type CC with probability $q_s > 0$. Let $\hat{q} = \min_{s \in S} q_s$. This completes the proof. \Box

Proposition 2. Generically, every absorbing state in which there is some altruist must be a state of full cooperation and every absorbing state in which there is some defector must be a state of full defection.



Fig. 6. Cooperation rates by heterogenous memories ($\rho = 4$). $\theta = 0$ (blue line), $\theta = 0.01$ (green), $\theta = 0.05$ (red), $\theta = 1$ (light blue). Confidence intervals are reported in Table 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



Fig. 7. Cooperation rates when mechanisms are removed.

Proof. Take any state s(t) in which some agents cooperate and some defect. Generically a cooperator and a defector will receive different payoffs. There is also positive probability that any cooperator and any defector are randomly chosen for selection. The agent with the lower payoff will change her type, unless both agents are conditional cooperators. Hence $q(s,s) \neq 1$. \Box

A.2. Details of the simulations

In this section we provide further details regarding the simulations of the model. We describe the sequence of events happening in a run of the computer program:

- 1. Initialize the model.
 - (a) Generate a small-world network of *N* agents:
 - i. Generate a ring of N agents, each connected to her ρ nearest neighbors by a link.
 - ii. Rewire each existing link with probability θ :
 - A. Choose an agent and the link that connects it to its nearest neighbor in a clockwise sense.
 - B. With probability ρ , reconnect this link to another agent chosen uniformly at random.
 - C. Consider each agent moving clockwise around the ring until one lap is completed and with probability ρ rewire her first link in the same way.
 - D. Next, make a similar circle but now rewire the link which connects the agent to it's second nearest neighbor clockwise.
 - E. Continue this process proceeding outward to more distant neighbors until each link in the original lattice has been considered once for rewiring.
 - (b) Assign types: each agent's initial type is uniform randomly drawn from the type space {A,CC,D}.
 - (c) Set initial belief to 2/3 (conditional cooperators play 'cooperate' in the first period).
- 2. In each period t = 1, 2, 3...:
 - (a) Interactions: repeat N times.
 - i. Draw a random agent *i* from the set of all *N* agents with [*without*] replacement,¹⁴

¹⁴ We ran two distinct sets of simulations. One where we draw agents with replacement and one where we draw agents without replacement. Drawing

- ii. Draw an opponent *j*: a random agent among *i*'s neighbors.
- iii. Determine *j*'s reputation for *i*:
 - A. Compute the number of times that *j* cooperated with *i* in *i*'s last *h* interactions divided by all interactions they (*i* and *j*) had among *i*'s last *h* interactions ($\gamma_{ij}(h)$)—set $\gamma_{ij}(h) < 0$ if there was no interaction.
 - B. Compute the average of the respective statistics among *i*'s neighbors ($\overline{\beta}_{ij}(h)$): the number of times *j* cooperated with a neighbor of *i* divided by the number of times they interacted in the neighbor's last h interactions—set $\overline{\beta}_{ij}(h) < 0$ if there was no interaction with any of the neighbors.
 - C. Compute the population share of individuals who cooperated in the previous period ($\overline{\sigma}^{t-1}$).
 - D. Determine reputation according to the following equation:

$$\underset{ij}{t} = \begin{cases} \lambda \gamma_{ij}(h) + (1-\lambda)\overline{\beta}_{ij}(h) & \text{if } \gamma_{ij}(h) \ge 0 \land \overline{\beta}_{ij}(h) \ge 0 \\ \gamma_{ij}(h) & \text{if } \gamma_{ij}(h) \ge 0 \land \overline{\beta}_{ij}(h) < 0 \\ \overline{\beta}_{ij}(h) & \text{if } \gamma_{ij}(h) < 0 \land \overline{\beta}_{ij}(h) \ge 0 \\ \overline{\sigma}^{t-1} & \text{if } \gamma_{ij}(h) < 0 \land \overline{\beta}_{ij}(h) < 0 \end{cases}$$

(3)

where λ is the weight put on own experience.

- iv. Determine *i*'s action based on *j*'s reputation.
- v. Determine *i*'s reputation for *j*: in the same way as in (iii) exchanging the roles of *i* and *j*.
- vi. Determine *j*'s action based on *i*'s reputation.
- vii. Realize payoffs based on the two players' actions.
- viii. Save *i*'s payoff for comparison in the selection process.
- ix. Change *i*'s memory: new information: *j*'s action.
- (b) Count how many agents cooperated in the position of agent *i*.
- (c) Selection: k times.

r

- i. Draw a random agent *i*.
- ii. Draw a random agent *j*.
- iii. If *j*'s payoff from the last interaction is higher than *i*'s payoff from the last interaction, *i* adopts *j*'s type.
- (d) Update memory: for every agent: delete the oldest information (from period t-h).
- (e) Update the default belief $(\overline{\sigma}^{t-1})$ using the number of cooperating agents.

All the results reported in this paper are the averages of the steady state cooperation rates over 100 runs.

A.2.1. Sample runs

Fig. 3 shows how the type shares and the level of cooperation evolve during the course of two typical runs. On the left panel the system evolves to full cooperation for memory h=5. We can see that at the beginning of the run altruists are exploited and their share is declining. In contrast, defectors soar in the population. However, after about 100 periods the share of defectors starts to decrease and altruists gain a higher share. In 100 periods

(footnote continued)



Fig. 8. Average cooperation rates in the case of mutations for different memory values and rewiring probabilities.

Table 3

Evolution of memory with mutations (2%). Initial conditions: A, 1/3; D, 1/3; CC_2 , 1/9; CC_7 , 1/9; CC_{15} , 1/9.

θ	Cooperation	Α	D	CC ₂	CC ₇	<i>CC</i> ₁₅
0	79.4	0.1143	0.102	0.2582	0.2626	0.2629
0.01	77.38	0.113	0.11	0.257	0.26	0.26
0.05	67.88	0.1	0.143	0.238	0.26	0.258
1	11.03	0.0398	0.1959	0.1993	0.2491	0.316

conditional cooperators had the chance to learn the types of their opponents. Hence defectors earn low payoffs when matched with conditional cooperators while altruists are able to gain high payoffs.

The right panel of Fig. 3 shows a different case for memory h=9. Here the system ends up in full defection. Again, defectors have advantage at the beginning of the run while the altruists' share is decreasing. However, in this case we cannot observe the reversal of these trends after some periods of running. Due to the long memory, early defections are remembered for long, thus conditional cooperators gain bad reputation and the system is trapped in full defection.

A.2.2. Frequency of cooperation scenarios

In our simulations, the long-run cooperation levels can be classified into three scenarios: (1) every agent cooperates, (2) no agent cooperates, and (3) the cooperation stabilizes around some intermediate value. This third scenario happens only when both altruists and defectors die out (see Proposition 2).¹⁵ Fig. 4 shows the frequency of these three possible scenarios over 100 runs for each parameter setting. For lower values of the rewiring probability, the system always converges to one of the corner cases. Intermediate cooperation values appear only if both the rewiring probability and the memory size are large enough.

without replacement ensures that each agent remembers exactly the last h interactions. This corresponds to the benchmark model in the main text. Drawing with replacement allows for heterogeneity because some agents may be drawn repeatedly and hence remember more than h interactions. On average, though, agents will remember h interactions also in that case. See Section A.3.

¹⁵ Note that if both altruists and defectors die out, the system still might converge to one of the corner cases of full or zero cooperation.

Table 4

Av. rates of cooperation (without mutation) and 95% conf. intervals (see Fig. 2 in the main text).

	θ											
Mem	0			0.01			0.05			1		
	Av.	_	+	Av.	_	+	Av.	_	+	Av.	-	+
1	0	0	0	0	0	0	0	0	0	0	0	0
2	8	2.66	13.34	4	0.14	7.86	0	0	0	0	0	0
3	32	22.81	41.19	24	15.59	32.41	13	6.38	19.62	0	0	0
4	41	31.31	50.69	36	26.54	46.46	25	16.47	33.53	0	0	0
5	41	31.31	50.7	40	30.35	49.65	37	27.49	46.51	0	0	0
6	40	30.35	49.65	41	31.31	50.59	37	27.49	46.51	0	0	0
7	50	40.15	59.85	35	25.60	44.4	39	29.39	48.61	1	0	2.96
8	46	36.18	55.82	31	21.89	40.11	42	32.28	51.72	0	0	0
9	29	20.06	37.94	35	25.6	44.4	36	26.54	45.46	0	0	0
10	25	16.47	33.53	31	21.89	40.11	35	25.6	44.4	2.52	0	5.32
11	21	12.98	29.02	28	19.16	36.84	37	27.49	46.51	1.05	0.26	1.84
12	20	12.12	27.88	21	12.98	29.02	26	17.36	34.64	4.3	0.85	7.75
13	24	15.59	32.41	14	7.16	20.84	30	20.97	39.03	5.13	1.69	8.57
14	16	8.78	23.22	16	8.78	23.22	21	12.98	29.02	4.89	1.89	7.89
15	14	7.16	20.84	21	12.98	29.02	24	15.59	32.41	4.12	1.17	7.07
16	15	7.97	22.03	15	7.97	22.03	16.5	9.35	23.65	3.9	1.64	6.16
17	12	5.60	18.4	19	11.27	26.73	16.85	9.51	24.19	3.45	2.12	4.78
18	9	3.36	14.64	12	5.6	18.4	11.34	5.32	17.36	6.42	2.89	9.95
19	12	5.60	18.4	13	6.38	19.62	13.47	7.23	19.71	3.5	2.27	4.73
20	15	7.97	22.03	15.96	8.76	23.16	11.91	6.55	17.27	2.8	1.5	4.1

A.3. Extensions

A.3.1. Meeting strangers: can cooperation still survive?

In the benchmark model, people only interact with their firstorder neighbors. This assumption can be too extreme, since in real life people meet more socially distant individual or even complete strangers. In such a case, the combination of limited memory and casual meeting of unknown individuals may suggest that cooperation cannot arise. We explore an alternative setup here. We assume that people can meet *anybody* in their component, being the probability to interact with a particular player proportional to the shortest "geodesic distance" that separates the two players. Geodesic distance between any two nodes is the minimum number of links that connects the two nodes.

The probability for *i* to interact with *j* given that *i* has been chosen as row player is the following:

$$P_{ij} = \begin{cases} \frac{e^{-\alpha d_{ij}}}{\sum_{k \neq i} e^{-\alpha d_{ik}}} & \text{if } d_{ij} > 0\\ 0 & \text{if } d_{ij} = 0 \text{ or } d_{ij} = \infty \end{cases}$$

Observe that the higher α the more likely it is to meet neighbors and the more unlikely to meet distant individuals. Hence, α parameterizes the effect of the network for matching. As $\alpha \rightarrow \infty$, everybody meets exclusively her neighbors (the original model in the main text). If $\alpha = 0$ matching is completely random and cooperation does not survive. If $d_{ij} = \infty$ (agents in disconnected components), $P_{ij} = 0$.

Fig. 5 shows the average of the steady state cooperation rate over the 100 runs.¹⁶ For a given values of *h* and θ , the cooperation increases monotonically in α . As α gets higher, agents are matched most frequently with agents from the neighborhood, which limits the number of possible opponents and facilitates the learning about their types. More importantly, the overall pattern of cooperation as a function of memory is robust to changes in the locality of matching.

A.3.2. Heterogeneity in memory

In the benchmark model, people are drawn to play the game *without* replacement. This ensures that each agent interacts only once per period. Hence *all* agents remember their last *h* interactions. In this section, we allow for heterogeneity in the number of interactions people remember. To this aim, in each round we draw the agents *with* replacement. As a result, some people might have more and some less than *h* interactions in the last *h* periods. This creates heterogeneity with respect to the amount of information people have to assess reputation of their future opponents. Fig. 6 shows that the results still hold under this specification of the model.

A.4. Removing mechanisms

We also analyze whether cooperation emerges if we remove one of the model mechanisms (direct reputation, indirect reputation or network-based meetings). Fig. 7 shows the simulation results. If agents rely only on own experience ($\lambda = 1$), cooperation rates are very low for all memory values. If individuals only use the information of their neighbors ($\lambda = 0$), cooperation emerges but it is dramatically lower than in the baseline case. If we remove the matching role of the social network and maintain both reputation mechanisms, agents are matched randomly and cooperation never emerges (non-reported in Fig. 7).

A.5. Mutations

We introduce the possibility of mutations into the selection process. With a probability 0.02 the randomly chosen agent's fitness is not compared to the fitness of someone else, but she adopts a type uniform randomly.¹⁷ This way, every behavioral type has the chance to be reintroduced to the population.

In this case, the model does not converge to absorbing states. Hence we run the model for 30 000 periods and computed the average cooperation rate over the 30 000 periods as the outcome of one simulation. We run 100 such simulations and take the average

 $^{^{16}}$ In the computer program we change step 2.(a).ii: we draw the opponent based on the distance using the described function.

¹⁷ This modifies the computer program at the point 2.(c).

cooperation rate over this sample. The results are shown in Fig. 8 and Table 6 for different values of memory (*h*) and network structure parameter (θ). Cooperation is slightly larger around the cooperation-maximizing memory levels than in the figure in the main text. This causes a starker decrease of cooperation levels as memory grows larger. Most importantly, we can conclude that the results are robust against the inclusion of mutations.

A.6. Evolution of memory: robustness to initial type distribution

We also check whether the results of Section 4 are robust to changes of the initial conditions. In the main text, we show the case when each of the five types represents 20% in the initial state of the population. Here we report results for a case in which altruist and defectors initially represent one-third of the population, while each type of conditional cooperators represents 1/9. The results are reported in Table 3. Observe that the long-run levels of cooperation and type distributions are virtually the same as in the main text.

A.7. Tables: group-level cooperation

Average level of cooperation is computed over 100 runs. The 95% confidence intervals are calculated using the formula

Table 5

Average rates of cooperation (without mutation) and 95% confidence intervals in the case of 'heterogeneity of memory' (see Fig. 6).

Mem	θ														
	0			0.01	0.01			0.05			$\begin{array}{c ccccccccccccccccccccccccccccccccccc$		1		
	Av.	_	+	Av.	_	+	Av.	-	+	Av.	-	+			
1	0	0	0	0	0	0	0	0	0	0	0	0			
2	1	0	3	1	0	3	0	0	0	0	0	0			
3	19	11.2	26.7	17	9.6	24.4	2	0	4.8	0	0	0			
4	38	28.4	47.6	28	19.2	36.8	14	7.2	20.8	0	0	0			
5	56	46.2	65.8	44	34.2	53.8	37	27.5	46.5	0	0	0			
6	50	40.1	59.8	38	28.4	47.6	41	31.3	50.7	0	0	0			
7	49	39.1	58.8	44	34.2	53.8	34	24.7	43.3	0	0	0			
8	36	26.5	45.5	49	39.2	58.8	37	27.5	46.5	1.0	0	3.0			
9	41	31.3	50.7	33	23.7	42.3	43	33.2	52.8	2.0	0	4.8			
10	35	25.6	44.4	30	21.0	39.0	34	24.7	43.3	2.2	0	5.0			
11	27	18.2	35.7	28	19.2	36.8	34	24.7	43.3	2.5	0	5.3			
12	23	14.7	31.3	23	14.7	31.3	29	20.1	37.9	5.8	1.6	10.2			
13	23	14.7	31.3	22	13.8	30.2	25	16.5	33.5	4.0	1.2	7.0			
14	21	12.9	29.0	22	13.8	30.2	26	17.4	34.6	3.2	1.1	5.5			
15	30	20.9	39.0	20	12.1	27.9	24	15.6	32.4	1.2	0.5	2.0			
16	24	15.5	32.4	21	13.0	29.0	24	15.6	32.4	2.3	1.2	3.4			
17	22	1.8	30.2	20	12.1	27.9	15.7	8.6	22.9	6.0	2.5	9.6			
18	15	7.9	22.0	13	6.4	19.6	14.7	7.8	21.6	4.3	2.8	5.9			
19	14	7.1	20.8	14	7.2	20.8	16.8	9.7	24.0	11.2	6.5	16.0			
20	12	5.5	18.4	20	12.1	27.9	14.1	7.6	20.8	11.2	6.2	16.3			

Table 6

Av. rates of cooperation with mutation	(and 95% conf. intervals, see Fig. 2
--	--------------------------------------

Mem	θ											
	0			0.01			0.05			1		
	Av.	_	+	Av.	-	+	Av.	_	+	Av.	-	+
1	0	0	0	0	0	0	0	0	0	0	0	0
2	9.6	6.1	13.2	7.9	4.4	11.5	3.1	1.5	4.7	0	0	0
3	40.6	33.3	47.8	39.9	32.2	47.5	31.3	24.8	37.8	0	0	0
4	54.8	46.9	62.8	63.3	55.6	71.1	51.8	44.4	59.3	0	0	0
5	67.1	60.1	74.0	54.1	46.0	62.3	53.0	45.2	60.8	0.6	0	1.1
6	60.2	52.4	68.0	56.0	48.1	63.9	56.9	49.2	64.6	2.1	1.1	3.1
7	61.2	53.5	69.0	54.5	46.5	62.4	60.8	53.2	68.4	3.2	2.0	4.4
8	49.9	42.1	57.8	49.1	40.9	57.2	59.2	51.5	66.9	4.6	3.3	6.0
9	49.1	41.9	57.1	56.0	48.3	63.8	51.7	43.8	59.7	6.4	4.4	8.3
10	49.2	41.4	57.0	40.3	32.1	48.4	50.8	42.8	58.7	8.1	5.5	10.6
11	34.2	26.8	41.6	49.0	41.4	56.6	41.0	33.7	48.4	9.9	7.4	12.4
12	34.4	27.3	41.4	32.2	25.3	39.2	36.4	29.5	43.4	9.5	7.3	11.8
13	23.6	17.6	29.5	24.5	18.6	30.4	30.2	24.1	36.3	11.8	9.2	14.4
14	16.9	11.9	21.9	16.1	11.1	21.1	27.54	21.5	33.5	10.9	8.8	13.0
15	11.8	8.2	15.4	13.4	9.5	17.4	19.5	15.2	23.9	12.9	10.1	15.6
16	10.3	7.1	13.5	13.7	10.1	17.2	14.5	10.6	18.4	9.1	7.1	11.1
17	8.7	5.7	11.7	7.0	4.7	9.3	15.3	11.3	19.3	7.5	5.7	9.2
18	6.1	4.2	7.9	6.2	4.3	8.1	7.1	5.1	9.1	7.3	5.6	9.1
19	4.6	3.1	6.1	5.1	3.6	6.5	5.9	4.3	7.6	7.5	6.0	9.1
20	5.5	3.8	7.2	3.9	2.6	5.2	7.4	5.3	9.6	6.3	4.9	7.8

 $[\hat{\mu}-z_{0.975}(\sigma/\sqrt{n}); \hat{\mu}+z_{0.975}(\sigma/\sqrt{n})]$ where $z_{0.975} = 1.96$ comes from the normal distribution, n=100 is the sample size and σ is the standard deviation computed from the sample.

References

- Abramson, G., Kuperman, M., 2001. Social games in a social network. Phys. Rev. E 63, 030901.
- Aktipis, C.A., 2006. Recognition memory and the evolution of cooperation: how simple strategies succeed in an agent-based world. Adaptive Behav. 14, 239–247.
- Alonso-Sanz, R., 2009. Memory versus spatial disorder in the support of cooperation. BioSystems 97, 90–102.
- Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.
- Bester, H., Guth, W., 1998. Is altruism evolutionary stable? J. Econ. Behav. Organ. 34, 193–209.
- Biely, C., Dragosits, K., Thurner, S., 2007. The prisoners' dilemma on dynamic networks under perfect rationality. Phys. D: Nonlinear Phenom. 228 (1), 40–48.
- Boyd, R., Richerson, P.J., 1989. Evolution of indirect reciprocity. Soc. Networks 11, 213–236.
- Boyd, R., Richerson, P.J., 1990. Group selection among alternative evolutionarily stable strategies. J. Theor. Biol. 145, 331–342.
- Cowan, N., 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav. Brain Sci. 24, 87–185.
- Corten, R., Cook, K., 2008. Cooperation and reputation in dynamic social networks. In: Poalucci, M. (Ed.), Proceedings of the First International Conference on Reputation: Theory and Technology—ICORE 09. Gargonza, Italy.
- Cox, S.J., Sluckin, T.J., Steele, J., 1999. Group size, memory, and interaction rate in the evolution of cooperation. Curr. Anthropol. 40, 369–377.
- Crespi, B.J., 2001. The evolution of social behavior in microorganisms. Trends Ecol. Evol. 16, 178–183.
- Ebel, H., Bornholdt, S., 2002. Coevolutionary games on networks. Phys. Rev. E 66, 056118.
- Eshel, I., Binmore, K., Shaked, A., 1998. Altruists, egoists, and hooligans in a local interaction model. Am. Econ. Rev. 88, 157–179.
- Fosco, C., Mengel, F., 2011. Cooperation through imitation and exclusion in networks. J. Econ. Dyn. Control 35, 641–658.
- Goyal, S., 2007. Connections: An Introduction to the Economics of Networks. Princeton University Press, NJ.
- Hanaki, N., Peterhansel, A., Dodds, P., Watts, D., 2007. Cooperation in evolving social networks. Manage. Sci. 53 (7), 1036–1050.
- Hauert, C., Schuster, H.G., 1997. Effects of increasing the number of players and memory steps in the iterated prisoner's dilemma, a numerical approach. Proc. R. Soc. B 264, 513–519.
- Izquierdo, S., Izquierdo, L.R., Vega-Redondo, F., 2010. The option to leave: conditional dissociation in the evolution of cooperation. J. Theor. Biol. 267, 76–84.
- Janssen, M., 2006. Evolution of cooperation when feedback to reputation scores is voluntary. J. Artif. Soc. Soc. Simulations 9 (1), 17.
- Kawai, N., Matsuzawa, T., 2000. Numerical memory span in a chimpanzee. Nature 403, 39–40.
- Kirchkamp, O., 2000. Spatial evolution of automata in the prisoner's dilemma. J. Econ. Behav. Organ. 43, 239–262.
- Mailath, G.J., Samuelson, L., 2006. Repeated Games and Reputations. Oxford University Press.
- Mengel, F., 2008. Matching structure and the cultural transmission of social norms. J. Econ. Behav. Organ. 67, 608–623.

- Milinski, M., Wedekind, C., 1998. Working memory constrains human cooperation in the prisoner's dilemma. Proc. Natl. Acad. Sci. U.S.A. 95, 13755–13758.
- Miller, G.A., 1956. The magical number seven plus or minus two some limits on our capacity for processing information. Psychol. Rev. 63, 81–97.
- Mohtashemi, M., Mui, L., 2003. Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altruism. J. Theor. Biol. 223, 523–531.
- Nakamuru, M., Kawata, M., 2004. Evolution of rumours that discriminate lying defectors. Evol. Ecol. Res. 6 (2), 261–283.
- Nakamaru, M., 2006. Lattice models in ecology and social sciences. Ecol. Res. 21 (3), 364-369.
- Nakamaru, M., Matsuda, H., Iwasa, Y., 1997. The evolution of cooperation in a lattice structured population. J. Theor. Biol. 184, 65–81.
- Nowak, M.A., Sigmund, K., 1998. The evolution of indirect reciprocity. Nature 437, 1291–1298.
- Nowak, M.A., Sigmund, K., 1998. The evolution of indirect reciprocity by image scoring. Nature 393, 573–577.
- Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M.A., 2006. A simple rule for the evolution of cooperation on graphs and social networks. Nature 441, 502–505.
- Ohtsuki, H., Isawa, Y., 2007. Global analysis of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. J. Theor. Biol. 244, 518–531.
- Ohtsuki, H., Nowak, M.A., 2007. Direct reciprocity on graphs. J. Theor. Biol. 247, 462–470.
- Panchanathan, K., 2011. Two wrongs don't make a right: the initial viability of different assessment rules in the evolution of indirect reciprocity. J. Theor. Biol. 277, 48–54.
- Premack, D., Premack, A., 2003. Original Intelligence: Unlocking the Mystery of Who We Are. McGraw-Hill, NY.
- Qin, S.M., Chen, Y., Zao, X.L., Shi, J., 2008. Effect of memory on the prisoner's dilemma game in a square lattice. Phys. Rev. E 78 (4), 041129.
- Raub, W., Weesie, J., 1990. Reputation and efficiency in social interactions: an example of network effects. Am. J. Sociol. 96 (3), 626–654.
- Read, D.W., 2008. Working memory: a cognitive limit to non-human primate recursive thinking prior to hominid evolution. Evol. Psychol. 6, 676–714.
- Roberts, G., 2008. Evolution of direct and indirect reciprocity. Proc. R. Soc. B 275, 173–179.
- Santos, F.C., Rodrigues, J.F., Pacheco, J.M., 2006. Graph topology plays a determinant role in the evolution of cooperation. Proc. R. Soc. B 273, 51–55.
- Sarin, R., 2000. Decision rules with bounded memory. J. Econ. Theory 90 (1), 151–160.
- Silk, J.B., 2006. Who are more helpful, humans or chimpanzees? Science 311, 1248–1249
- Sommerfeld, R., Krambeck, H.-J., Semmann, D., Milinski, M., 2007. Gossip as an alternative for direct observation in games of indirect reciprocity. Proc. Natl. Acad. Sci. U.S.A. 104 (44), 17435–17440.
- Sommerfeld, R.D., Krambeck, H.-J., Milinski, M., 2008. Multiple gossip statements and their effect on reputation and trustworthiness. Proc. R. Soc. B 275, 2529–2536.
- Stevens, J.R., Hauser, M.D., 2004. Why be nice? Psychological constraints on the evolution of cooperation. Trends Cognitive Sci. 8, 60–65.
- Tomasello, M., 2008. Origins of Human Communication. MIT Press.
- Trivers, R.L., 1971. The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35-57.
- Vega Redondo, F., 2007. Complex Social Networks. Cambridge University Press, UK. Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks.
- Nature 393, 409–410.West, S.A., Griffin, A.S., Gardner, A., Diggl, S.P., 2006. Social evolution theory for microorganisms. Nat. Rev.: Microbiol. 4, 597–607.
- Zimmermann, M.G., Eguíluz, V.M., San Miguel, M., 2004. Coevolution of dynamical states and interactions in dynamic networks. Phys. Rev. E 69, 065102-1.