Evolutionary Reputation Games On Social Networks

Chen Avin[*]
Communication Systems Engineering Department,
Ben-Gurion University of The Negev,
Beer-Sheva 84105, Israel

David Dayan-Rosenman[†]
Department of Political Science,
University of California, Los Angeles,
Los Angeles, CA, 90095-1472, USA
and
Weill Cornell Medical College
1300 York Ave, New York, NY, 10021, USA

We adapt an evolutionary model based on indirect reciprocity to the context of a structured population. We investigate the influence of clustering on the dynamics of cooperation in social networks exhibiting short average path length and various levels of locality. We show empirically how, as expected, a higher degree of locality, measured by clustering, can promote cooperation in a game involving reputation. More surprisingly, we show that a higher degree of locality results in slower convergence times for the population. These results show the existence of a trade-off between the need for higher cognitive abilities (understood as a longer memory of past interactions and/or the ability to keep tabs on a larger number of people) and the convergence time needed to reach a cooperative equilibrium. A population of individuals will need higher cognitive abilities to achieve a faster convergence time; on the other hand, a population with lower cognitive abilities may be able to reach the cooperative equilibrium but will get there slower. The trade-off between the rate of convergence and the need for higher cognitive abilities can be controlled by tuning the amount of locality in the graph (the clustering). These results shed some light on two facts: (1) Successful groups that do not rely on institutional enforcement of social norms tend to present a high degree of clustering; (2) Groups that experience rapid changes in membership tend to present low clustering.

1. Introduction

Humans distinguish themselves from most social animals by their propensity to cooperate with each other in highly varied situations and

[*]Electronic mail address: avin@cse.bgu.ac.il.
[†]Electronic mail address: dsr2004@med.cornell.edu (corresponding author).

often outside of kin groups [1]. Waging war, ensuring the operation of financial markets as well as controlling the international spread of an epidemic like *influenza* are tasks requiring that humans organize and cooperate with each other in order to carry out some form of collective action. This can be achieved by setting up public institutions designed to facilitate and lend structure to complex cooperative efforts for the common good. However, the benefits of institutions tend to be shared by most while the costs associated with their operation is primarily borne by those who participate in them. Why would each rational individual engage in this collectively beneficial but individually costly cooperative behavior? While it is easy to see how modern states can help bring about such cooperative behavior through the use of reward and punishment it just shifts the problem one level up to figuring out how modern states came to be.

Identifying and explaining the mechanisms that make cooperation possible have been one of the most vexing problems in the social sciences and have been studied extensively across fields (see, e.g., [2, 3]). Approaches based on non-cooperative game theory and the rational choice paradigm have been applied to the problem with some success. Yet they are limited in that they do not shed light on the process leading to cooperation.

An alternative paradigm is that of evolutionary game theory: People are ascribed a certain "type" that defines their behavior, which they proceed to change by imitating their more successful neighbors. This approach results in an adaptation dynamic that mimics that of natural selection. In this context, cooperation can evolve from direct reciprocity as individuals are often willing to forfeit immediate benefits for the promise of greater benefits through sustained future cooperation—a phenomenon Axelrod coined the "shadow of the future" [4]. However, when people do not entertain the prospect of future interactions, as is the case in one-shot interactions, and they only consider their short-term gains, the shadow of the future disappears and models of cooperation based on direct reciprocity unravel. Models of indirect reciprocity have been suggested to remedy that shortcoming by introducing the idea of "reputation" [5–11]. In these models, an individual's propensity to cooperate with others will determine his/her reputation, and subsequently, an individual's reputation will be considered by others in future interactions.

Most of these models formalize social interactions in one of two ways: the population is either modeled as being unstructured, where everyone can interact with everyone else (the population is well mixed and can be represented as a complete graph); or it is modeled as a collection of rather rigid groups (*cliques*) that are competing in a multilevel evolutionary fashion [12]. However, these models do not closely consider the structure of the population in which these interactions occur and, as significant parts of what constitutes the structure of "real-life" groups are abstracted away, these models run the risk of overlooking significant effects attributable to the social structure of such groups (see [13] for an example of how "social viscosity" can affect evolutionary dynamics). In particular, the idea of how reputation gets affected by social structure has been largely left out of these models.

The question we investigate stems from the observation that a few groups stand out by minimizing their reliance on institutional enforcement by the state. Groups like the Amish, or the New York City based orthodox Jews involved in the diamond trade, tend to rely instead on reputation and social pressure to enforce social norms of cooperation. If an individual member of one of these groups deviates from the accepted social norm, they risk being shunned or otherwise ostracized. One property that these groups share is the fact their members are likely to know everyone else in their insular communities, and are rather secluded from people outside of their group. In other words, they tend to display a much greater cohesiveness than most other groups.

Questions of interest, model, and results overview

In this empirical study we explore questions regarding the role of population structure on evolutionary dynamics. In particular, we investigate the effect of the clustering coefficient both on the amount of cooperation and on the speed of the evolutionary process in a game involving reputation.

To do so we offer a new model for games of indirect reciprocity played on arbitrary graphs. In particular, we extend previous models for unstructured populations and integrate models of social networks. Our game presents localized dynamics where one agent can only observe and play with others in its neighborhood. This leads us to redefine the concept of reputation as a pair-wise property between two individuals as opposed to a population-wide property (as is the case in an unstructured population). Note that with this condition and when each agent knows all agents (i.e., a complete graph) we get the original model for an unstructured population. Finally, we offer a new kind of graph that enables exploring the significance of clustering coefficients by using agent-based simulations.

We find that higher clustering promotes cooperation in this game. More surprisingly, we find that a higher degree of locality results in slower convergence times for the population. These results show the existence of a trade-off between (1) the need for higher cognitive abilities (understood as a longer memory of past interactions and/or the ability to keep tabs on a larger number of people) and (2) the convergence time needed to reach a cooperative equilibrium: A population of individuals will need higher cognitive abilities to achieve a faster convergence time;

on the other hand, a population with lower cognitive abilities may be able to reach the cooperative equilibrium but will get there slower. We find that the trade-off between the rate of convergence and the need for higher cognitive abilities can be controlled by tuning the amount of clustering in the graph.

Outline

The rest of this paper is structured as follows: section 2 discusses evolutionary models of indirect reciprocity in unstructured populations, it introduces the model of Boyd and Panchanathan from [10]. Section 3 introduces our model and explains how it generalizes the previous model to account for population structure. In particular, this section tackles the question of how imitation dynamics and reputation dissemination can be modeled on a more general graph. It also introduces the properties that are specific to social networks. Section 4 is devoted to the specifics of our simulation setup. Finally, we present the results of our simulations in section 5, and conclude by discussing their significance in section 6.

2. Models of indirect reciprocity in unstructured populations

In [10], Boyd and Panchanathan revisited an indirect reciprocity game presented in [8] by Nowak, and originaly inspired by Sugden in [7]. They consider an infinite, unstructured population of agents interacting in the context of a public-aid game. In the first round, each agent acts as a potential donor to a randomly selected partner. If a donation is offered, the donor's *standing* is good in the next round and his fitness is decremented by c while the recipient's fitness is increased by b (with $b > c > 0$). If no donation is offered then we distinguish between two cases: (1) If the standing of the potential recipient was good then the donor looses his good standing; (2) If the standing of the potential recipient was bad then the standing of the donor is unaffected. In other words, an individual falls into bad standing when they fail to cooperate with a good standing partner (this reflects the idea that there should be no penalty associated with not helping a selfish person). The recipient's standing stays unaffected. Subsequent rounds of social interactions occur with probability $0 \leq w < 1$ and adaptation through payoff-biased imitation occurs after the agents stop interacting. We call the time between two adaptations a *step*. All agents begin in good standing at every step. The standing of each agent is known to all other agents in the population. An intended donation can fail with probability α; in this case the fitness of both donor and recipient are not affected, but the standing of the donor gets revised according to the rule defined above. A donation does not happen by error if it was not intended. Three strategies are considered: indiscriminate altruists (ALLC) always donate to a random partner, indiscriminate defectors (ALLD) never do, and reputation discriminators (RDISC) only donate to partners in good standing and do not donate to partners in bad standing [10, p. 116].

Boyd and Panchanathan show that RDISC and ALLD are the only evolutionary stable equilibria of this game.[1] Indirect reciprocity based on standing can be evolutionary stable and the RDISC strategy has a basin of attraction of a substantial size as long as w (and hence the expected number of rounds before payoff-biased imitation occurs) is high enough. Depending on the original population profile, the population will converge to being exclusively composed of either RDISC individuals or ALLD individuals.

The assumption of complete information is then relaxed (i.e., any individual can observe the actions of all others and thus infer their standing) by assuming instead that an individual knows the standing of their current partner with probability q. This is a reasonable thing to do when considering a well-mixed population but not in the case of a structured one. In that case, we expect both the social distance (the length of the shortest path between two nodes) and the possible acquaintances that individuals have in common to play a role in whether one individual knows the reputation of another. This concept is formalized in section 3.1.

3. Extending games of indirect reciprocity to arbitrary graphs

To test the influence of social structure on a game involves indirect reciprocity, so we adapted Boyd and Panchanathan's game to be played in the context of an arbitrary graph rather than that of an unstructured population. Our game is identical to Boyd and Panchanathan's when it is played on a clique, but we tackle differently the processes of reputation dissemination and of imitation when played on a general graph. We consider a population of agents represented as nodes of a graph $G = (V, E)$. An edge $(u, v) \in E : u, v \in V$ signifies that u and v are neighbors and can potentially engage in social interactions with each other. The agents interact following the social structure described by the graph G.

Each agent is ascribed an initial type $t_k \in \{\text{ALLC}, \text{ALLD}, \text{RDISC}\}$. Each agent can only observe and gain knowledge of social interactions they are a part of, or social interactions conducted between two of its neighbors. Each agent has a standing with respect to each of its neighbors as well as with respect to themselves. To enable us to explain the model more concisely, we define the *standing-neighborhood* of two agents as the set of all agents that can observe an interaction between these two agents.

[1] Other strategies may also be stable in the game with the three given strategies.

Definition 1 (standing-neighborhood) The standing-neighborhood of two nodes $u, v \in G = (V, E)$ denoted $\sigma_G(u, v)$ or simply $\sigma(u, v)$ is

$$\sigma(u, v) = \{\Gamma(u) \cup u\} \cap \{\Gamma(v) \cup v\}$$

where $\Gamma(x)$ is the neighborhood of node x.

Every agent begins in good standing with respect to all of its neighbors. As before, in the first round of social interactions, each agent u acts as a potential donor to one randomly selected neighbor v. This interaction between u and v can only affect the standing of u, the potential donor, with respect to the agents in its neighborhood. In particular, we have the following two cases.

1. If a donation is offered, the donor u becomes in good standing with respect to all of its neighbors (all nodes in $\Gamma(u)$). The donor's fitness is decremented by c while the recipient's fitness is increased by b (with $b > c > 0$).
2. If no donation is offered then the donor u becomes in bad standing with respect to the agents of the standing-neighborhood of u and v, $\sigma(u, v)$, that held the recipient v in good standing and the donor's standing does not change with respect to the agents of $\sigma(u, v)$ that held the recipient v in bad standing. The fitness of both donor and recipient are not affected.

Subsequent rounds of social interactions occur with probability $0 \leq w < 1$. As before, an intended donation can fail with probability α and a donation does not happen by error if it was not intended. The same three strategies ALLC, ALLD, and RDISC are considered.

At the end of each step (when agents stop interacting, in expectation after $1/(1 - w)$ rounds), agents undertake a phase of adaptation through payoff-biased imitation (the process is detailed in section 3.2) and the whole process repeats until equilibrium is reached.

3.1 Dissemination of reputation

The major way in which our model differs from that of Boyd and Panchanathan concerns the mechanism of appraisal and dissemination of reputation (modeled as standing). In Boyd and Panchanathan's model, the standing of every individual in the population is known to every other individual. This does not seem to be a realistic assumption in the context of a structured population. They later relax this and assume instead that each individual knows the standing of a certain fixed proportion of the population. This relaxed assumption is equally not adapted to the context of a structured population as we expect agents to have more information about their close neighbors (agents whose social network overlaps with theirs) than about people that are socially further away.

We take the position that, in a structured population, standing is not a property associated with an individual as much as one associated with an interacting pair of individuals. In other words, every individual has a standing with respect to each of their neighbors. Alice's standing in Bob's eyes can be different from her standing in Charlie's eyes depending on the information Bob or Charlie have about Alice. We assume that individuals can form their image of a person's standing through direct experience or through observation of an interaction between two of their neighbors. Specifically, an individual A begins in good standing with individual B. A can lose their good standing with B in one of two ways: (1) refusing to honor a call for help issued by B, or (2) refusing to honor a call for help issued by a neighbor C of B and C is in good standing with B. To account for the possibility of error, an individual can regain good standing with a neighbor when he is observed helping a common neighbor. In particular, refusing to help someone in bad standing does not affect the standing of the individual that refused to help. The dynamics of the reputation system we use are a generalization of those from Boyd and Panchanathan to an arbitrary social structure, and are identical to their model in the case of a clique.

3.2 Replicator's dynamic

We use a structured population as opposed to a well-mixed population since individuals do not have the same probability of interacting with every other individual in the population—they only interact with their neighbors. Thus, instead of considering the payoffs and the frequency of each type in the whole population to determine the adaptation dynamics, we only consider payoffs in the neighborhood of each node. Our replicator dynamic is a localized version of the proportional fitness rule. It is consistent with dynamics of imitation since you cannot imitate what you are not aware of. The probability that a node n will adopt type $t_k \in \{\text{ALLC}, \text{ALLD}, \text{RDISC}\}$ is obtained by taking the quotient of the total fitness accumulated by nodes of type t_k in n's neighborhood (including n) by the total fitness accumulated by all nodes in n's neighborhood (including n). The replicator's dynamic can then be described by this equation:

$$p_i(t_k) = \frac{\displaystyle\sum_{\substack{j \in \{\Gamma(i) \cup \{i\}\} \\ \wedge\, \text{type}(j) = t_k}} f(j)}{\displaystyle\sum_{j \in \{\Gamma(i) \cup \{i\}\}} f(j)} \tag{1}$$

where $p_i(t_k)$ denotes the probability that node i will adopt type t_k in the next step and $f(j)$ denotes node j's fitness.

3.3 Playing the game on social networks

Although this game can be played on any type of graph, we are interested in running it on a graph representing a social network. Our intuition of social structure stems from the following two properties.

1. Two people arbitrarily drawn from a population can be linked to each other by a relatively short succession of acquaintances (see [14] for a description of the Milgram experiment). This can be modeled by graphs with a short average path length.
2. Two people that are acquainted to a third one are more likely to be acquainted with each other than two people taken uniformly at random from the population. This idea of locality can be captured by the *clustering coefficient*: a number between 0 and 1 that reflects the fraction of a vertex's neighbors which are neighbors themselves (see [15] for a formal definition).

Watts and Strogatz [16] offered the small-world network model to capture both the locality and the short average path length present in real social networks. The graph model that we introduce in section 4 shares many features of the small-world network but differs from it in substantial ways.

3.4 Metrics of interest

In order to measure the influence of clustering on the dynamics of evolution, we keep tabs on two particular metrics. The probability of reaching a cooperative outcome measures how likely it is that the population ends up with no agents of type ALLD. The time to equilibrium is a measure of the number of evolution steps needed until a "point of no return" is reached in the population profile.

1. *Probability of reaching the cooperative equilibrium.* We repeat the simulation a number of times for each specific graph. For each one of these runs, we record whether the population reaches the cooperative or the uncooperative equilibrium. We consider the proportion of these runs where the cooperative outcome was reached to be the probability of reaching that particular equilibrium.
2. *Time to equilibrium.* For each run, we record the number of steps until the population is either exclusively composed of type ALLD (the uncooperative equilibrium) or until all agents of type ALLD have disappeared and the population is a mix of ALLC and RDISC. Once the population is only composed of these two cooperative types, it is not strictly speaking in the cooperative equilibrium (where all agents are of type RDISC) but deterministically converges to that state since RDISC has a higher fitness than ALLC in the presence of errors. Both RDISC and ALLC commit errors with probability α but RDISC retaliates against ALLC when they fail to cooperate while ALLC does not retaliate. We will record the median time to equilibrium taken over a set of runs.

4. Simulation setup

As we are interested in the influence of clustering on indirect reciprocity, we considered small-world networks since they provide a family of graphs with varying amounts of clustering.[2] There are, however, simulation issues related to the original small-world networks: they have nodes with a distribution of degrees (nodes exhibit variance in the number of neighbors they have) that could add significant noise to our results. Therefore, in the next section we introduce a family of graphs called "tribal graphs" that are designed to vary the amount of clustering displayed by the graph while keeping the degree of the nodes constant.

4.1 Tribal graphs

Tribal graphs are regular graphs (i.e., all nodes have the same constant degree) whose amount of clustering can be tuned using a specific parameter. A tribal graph is built in a manner inspired by that of small-world networks. We start with a deterministic collection of cliques of equal size and proceed to rewire a given number of edges by circling through the cliques and applying the following algorithm for each: select an edge uniformly at random from that clique, select an edge uniformly at random from another clique and "cross the edges over" (see Figure 1 for an illustration of the process).

Formally, tribal graph $TG(t, s, r) = (V, E)$ is a graph with $|V| = n = t \cdot s$ nodes and $|E| = t\binom{s}{2}$ edges. The graph is obtained in two phases as follows: First, arrange the nodes into t cliques, each containing s nodes. Let $0 \leq i \leq t - 1$, let V_i be the set of nodes in the ith cliques such that $V = \cup V_i$, and E_i is the $\binom{s}{2}$ edges associated with V_i. Note that after the first phase $E = \cup E_i$, the degree of each node is $s - 1$, and the clustering coefficient is 1. In the second phase we execute r rewirings as described in Algorithm 1.

[2]We leave the treatment of scale-free networks to future work, as it is not yet clear how to construct scale-free networks with varying clustering.

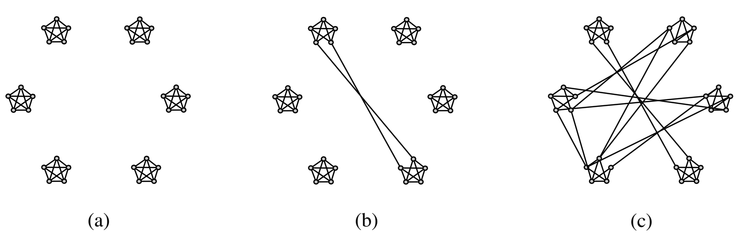

Figure 1. The construction of a tribal graph. We start with a collection of cliques (a) and cross pairs of edges (b) until we have rewired r of them (c).

Algorithm 1 Edge rewiring in $TG(t, s, r)$

```
 1: for k=1 to r do
 2:    pick an edge (i, j) uniformly at random from E_{k mod t}
 3:    pick x uniformly at random from {0, 1, ;t − 1} \ {k mod t}
 4:    pick an edge (u, v) uniformly at random E_x.
 5:    if (i, u) ∈ E or (j, v) ∈ E then
 6:       goto line 3
 7:    end if
 8:    remove edges (i, j) and (u, v) from E
 9:    add edges (i, u) and (j, v) to E.
10: end for
```

Tribal graphs are similar to small-world networks in their main features. First, they start with high clustering and as we increase the number of rewired edges the clustering coefficient decreases. Second, a low number of rewired edges is sufficient to guarantee that the graph has a small average path length, since the rewired edges function as "shortcuts" between nodes that would otherwise be more distant. Third, for a large number of rewired edges both graphs are similar to random graphs.[3]

The main difference is that in tribal graphs the degree of all the nodes does not change during the rewiring process. Since we cross two edges, each of the four nodes involved loses one neighbor and gains one neighbor in the process, keeping the degree invariant. By contrast, in small-world networks each rewiring involves a random node losing a neighbor and another random node gaining a neighbor, and this creates variance in the nodes' degrees. Therefore, tribal graphs allow us to obtain a "controlled environment" to test whether clustering (and not degree distribution) matters in the context of indirect reciprocity. As shown in Figure 2, the more edges that get rewired, the lower the resulting average clustering coefficient of the graph gets.

4.2 Parameters setup

Unless stated otherwise, our simulations are conducted on a population of 16 groups, each of them of size 16, we thus have 256 agents initially organized into cliques. The population starts with an equal number of agents of each type distributed randomly on the graph. The values $w = 0.95$, $\alpha = 0.05$, $c = 0.003$, and $b = 0.01$ are the same as in [10]. We vary the value of clustering by rewiring an increasing number of edges and creating tribal graphs with lower average clustering.[4] We

[3] Random regular graphs in the case of tribal graphs.

[4] With low probability, the graph that is generated can be composed of more than one connected component, potentially leading to the establishment of a different equilibrium in each of the components. We ignore these cases as they are just an artefact of the algorithm constructing the tribal graph.

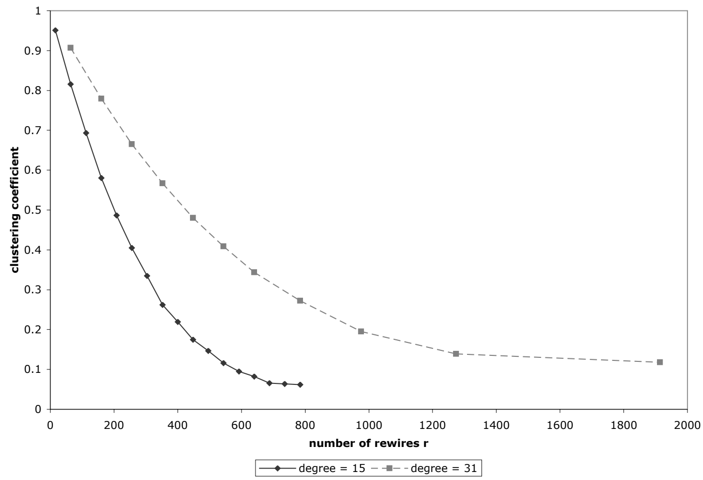

Figure 2. Average clustering coefficient as a function of the number r of edges that get rewired. Shown for $TG(16, 16, r)$ (degree 15 of each node) and $TG(8, 32, r)$ (degree 31 of each node). The relation is monotonically decreasing.

conduct 30 runs of each simulation on the same graph with each set of parameters to obtain each of the data points.

5. Results

Our main result is that the ability of a population to achieve and sustain the cooperative outcome exhibits a phase transition with respect to the clustering coefficient of the graph representing the population structure. We also find that the expected number of interactions before an adaptation step occurs (determined by the value of w) can be reduced when higher clustering is present and still lead to the productive outcome. Finally, we show that the possibility of a cooperative outcome increases with the degree of the nodes and that a smaller degree can be compensated for by a larger amount of clustering.

5.1 Likelihood of cooperation *versus* clustering

Figure 3 illustrates the fact that higher clustering makes achieving the cooperative outcome easier. The likelihood of a cooperative outcome exhibits a phase transition with respect to clustering. When clustering is very low, we obtain a random regular graph that displays a minimal amount of locality. Information about reputation does not spread well in such a graph and reputation discriminators cannot efficiently make use of previously observed interactions to guide their behavior. On the other extreme, when the population is highly clustered, the dissemination of

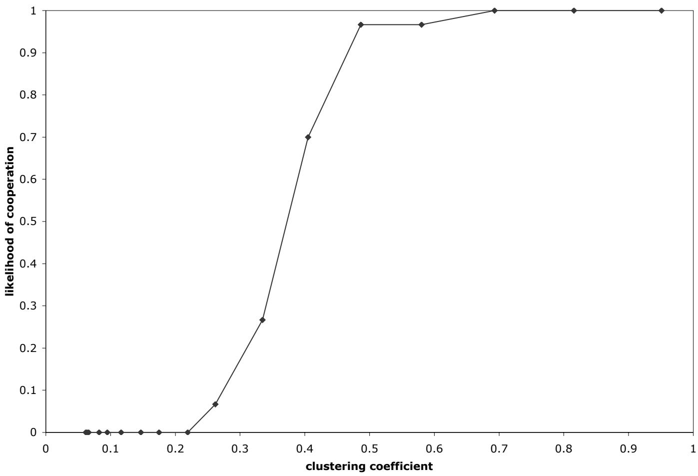

Figure 3. Probability that a population starting with an equal number of all three types of agent reaches a cooperative equilibrium.

reputation is effective since agents have many neighbors whose common interactions can be observed.

Figure 4 illustrates, for various starting points, the differences in the evolution of the population profiles. The dynamics for a particular initial population profile can be significantly affected by the amount of clustering. As expected, we can also see that a higher initial proportion of agents of type ALLC (indiscriminate altruists) helps agents of type ALLD (indiscriminate defectors) spread; while a higher initial proportion of agents of type RDISC (reputation discriminators) helps to achieve cooperation.

5.2 Higher clustering can compensate for lower *w*

As w gets larger, the expected number of rounds of interactions on which the individuals base their imitation decisions is also getting larger. In some way, w is thus a measure of the number of observations, or of the learning time before a decision is made as to whom to imitate. It can also be interpreted as a measure of the memory of the agents. A higher number of rounds before adaptation takes place favors the advent of the cooperative outcome, as it allows for reputation information about agents to more accurately reflect their actual type. It thus allows RDISC to discriminate more efficiently between cooperative and uncooperative types. Figure 5 shows that when w is higher the probability of achieving a cooperative outcome is heightened since the phase transition occurs at lower clustering. Relatedly, it also shows how an equivalent level of cooperation can be achieved with a lower w by increasing clustering.

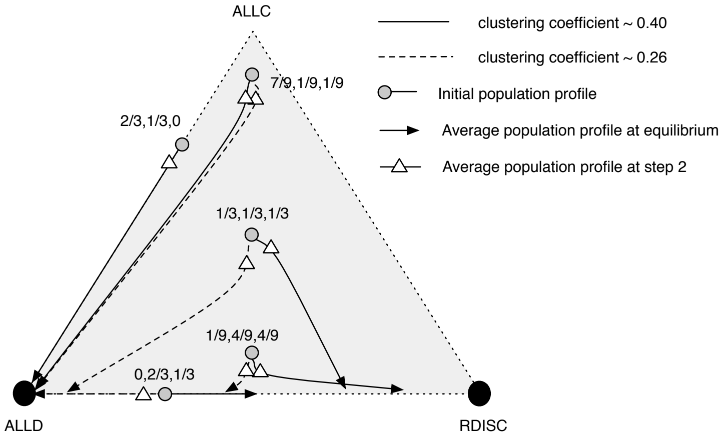

Figure 4. The evolution of five different initial population profiles. Each data point is averaged over 30 runs. The dynamics are shown for two different values of clustering: 0.40 (corresponding to 256 rewires) and 0.26 (corresponding to 352 rewires).

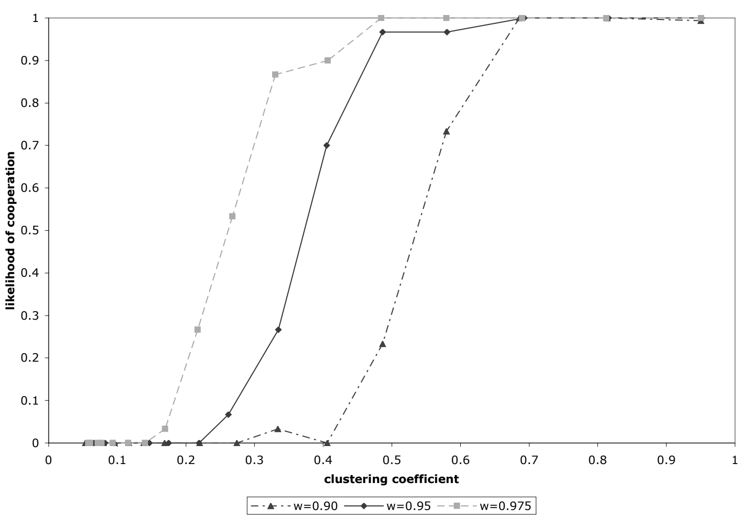

Figure 5. Probability that a population starting with an equal number of all three types of agent reaches a cooperative equilibrium. Graphs with different values of w are shown.

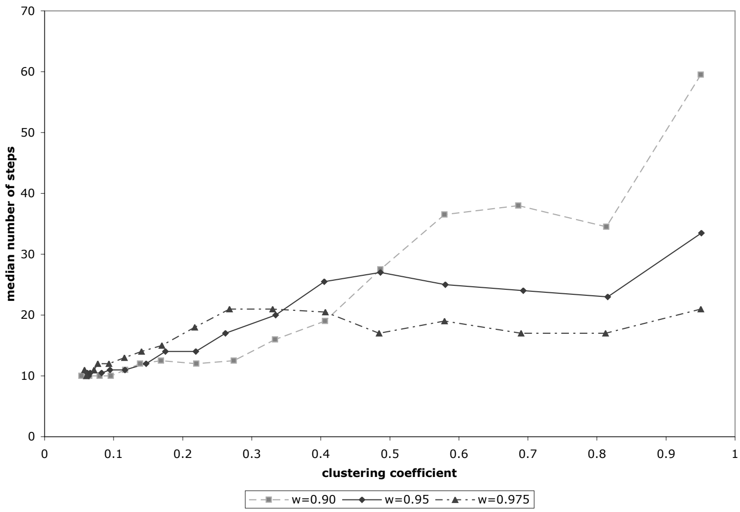

Figure 6. Median time to equilibrium for graphs with different values of w. Higher clustering generally results in longer time to equilibrium. The effect is more marked for runs with lower w, corresponding to a shorter period of observation.

Figure 6 shows that higher clustering, however, generally necessitates more time to reach equilibrium, pointing to a slower adaptability of the overall population to a changing environment (in particular pointing to a slower response to an invasion of defectors). Combining these two observations suggests that, in order to achieve a cooperative outcome, a trade-off between necessary memory and reactivity to change might be struck by varying the amount of clustering in the population. In other words, we can increase the likelihood of getting a cooperative outcome by increasing the clustering, but this results in detrimental effects on the pace of adaptation.

5.3 Clustering *versus* node degree

As shown in Figure 7, graphs with a higher node degree are more likely to reach a cooperative outcome. These graphs correspond to populations where agents have more neighbors and an overall higher number of edges in the graph results in a better flow of information related to reputation. Correspondingly, we can get an equivalent likelihood of achieving the cooperative outcome with a lower degree node by increasing the clustering coefficient. Figure 8, however, shows that higher clustering results in longer time to equilibrium, particularly for lower degree graphs. These two observations once more point to a trade-off: clustering can compensate for lower degree but this results in detrimental effects on the pace of adaptation of the population.

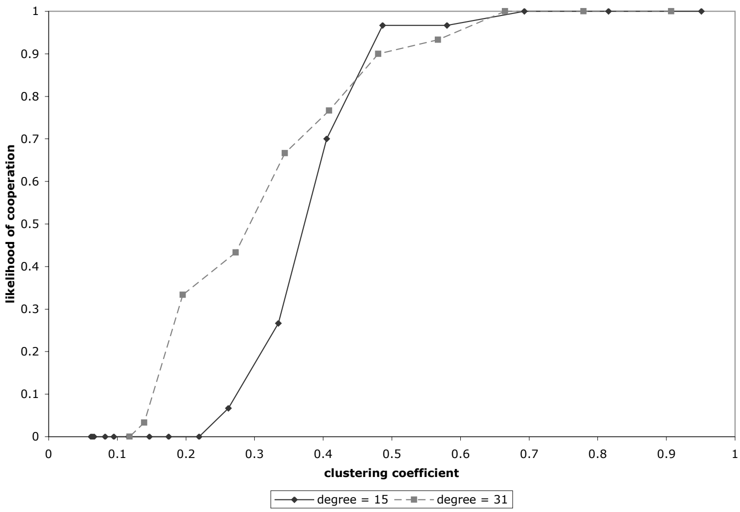

Figure 7. Probability that a population starting with an equal number of all three types of agent reaches a cooperative equilibrium. Graphs with two different degrees are shown. The higher degree graph transitions to the cooperative outcome at a lower clustering coefficient than the lower degree graph does. A higher degree thus makes cooperation easier.

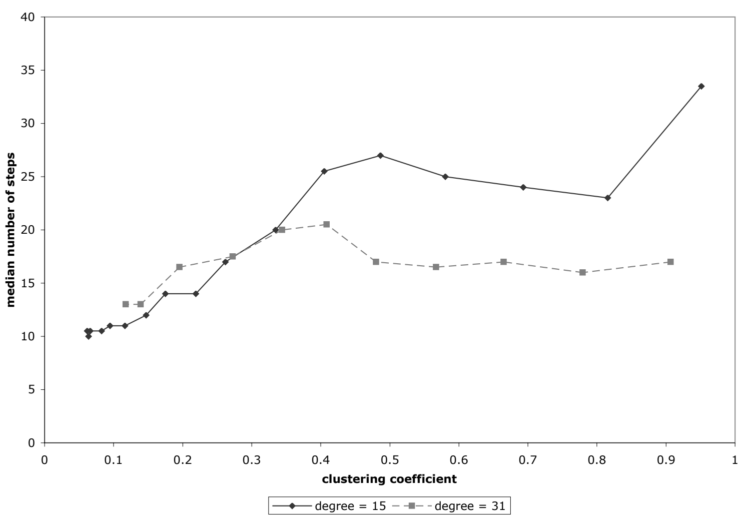

Figure 8. Time to equilibrium for graphs with different degrees. Higher clustering results in longer time to equilibrium. The effect is more marked for graphs with a lower node degree.

6. Discussion

In this paper, we showed that higher average clustering generally helps to promote cooperation in the context of a game of indirect reciprocity involving reputation. In particular, higher clustering allows preserving cooperation despite a reduced number of rounds before adaptation (corresponding to a smaller memory of the agents). In addition, higher clustering can also compensate for lower node degree (corresponding to a smaller number of social connections per agent). Our findings demonstrate that this increased likelihood of achieving cooperation comes, however, at the price of a slower progression toward equilibrium (a larger number of adaptation steps). This suggests that higher clustering results in a slower overall reactivity of the population to repeated exogenous shocks and, in particular, a lower resilience to repeated invasions by defectors. A higher level of clustering can seem at first to allow the population to get more information with a lesser cognitive burden. Agents are able to achieve the same likelihood of cooperative outcomes while keeping tabs on a lower number of other agents (corresponding to a lower node degree), and with shorter periods of observation (corresponding to a lower expected number of rounds per adaptation step) as long as the graph exhibits a high enough level of clustering. On a closer look though, this is not so much a bargain as it is a trade-off involving slower overall reactivity of the population.

These findings suggest new insight on two specific questions. First, they can contribute to explaining the rationale for the tribe or clan-like organization of early human groups: A high level of clustering can help bring about a cooperative outcome more easily in the context of lesser cognitive abilities (such as smaller memory and capacity for a lower number of social connections). As humans' cognitive abilities improved, they gravitated toward social organizations displaying a lower amount of clustering but exhibiting higher reactivity to changes. Second, these findings can explain why specific groups such as the New York City based Orthodox Jews involved in the diamond trade and the Amish display such high levels of clustering. These groups tend not to rely on institutional forms of enforcement for social norms and high clustering leads to some degree of distributed collective memory. The population as a whole benefits from the limited local information regarding standing that is being held by individuals about their neighbors. The lower reactivity to change that comes with higher clustering is not as much of an issue in these groups as it is in the general population. This is because these groups exist in a rather static environment with fewer opportunities for potentially uncooperative outsiders to join the group. Other groups in the general population do not need this information as they rely on institutional memory (through court records, for example).

Finally, it is interesting to note the resurgence of social interactions involving reputation systems. The development of the Internet; and in particular, the development of e-commerce with sites like *ebay*, gives new relevance to the appraisal of reputation. New developments in peer-to-peer technology also make use of such a system of accountability, as free-riding is a common problem on these networks. These new developments present new challenges, however, as the difficulty of establishing the reputation of participants is compounded by the fact that it is easier to shed an identity associated with a bad reputation in cyberspace than it is in the real world.

Acknowledgment

The authors would like to thank Dana Landis and the anonymous reviewers for their comments and corrections.

References

[1] Ernst Fehr, "Don't Lose Your Reputation," *Nature*, **432** (2004) 449–450.

[2] Mancur Olson, *Logic of Collective Action: Public Goods and the Theory of Groups* (Harvard University Press, 1971).

[3] Thomas C. Schelling, *The Strategy of Conflict* (Harvard University Press, 1980).

[4] Robert M. Axelrod, *The Evolution of Cooperation* (Basic Books, 1985).

[5] Paul Milgrom, Douglass North, and Barry Weingast, "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs," *Economics and Politics*, **3** (1990) 1–23.

[6] R. L. Trivers, "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology*, **46** (1971) 35–57.

[7] R. Sugden, *The Evolution of Reciprocal Altruism* (Basil Blackwell, Oxford, 1986).

[8] Martin A. Nowak and Karl Sigmund, "Evolution of Indirect Reciprocity by Image Scoring," *Nature*, **224** (1998) 573–577.

[9] Martin A. Nowak and Karl Sigmund, "The Dynamics of Indirect Reciprocity," *Journal of Theoretical Biology*, **194** (1998) 561–574.

[10] Karthik Panchanathan and Robert Boyd, "A Tale of Two Defectors: The Importance of Standing for Evolution of Indirect Reciprocity," *Journal of Theoretical Biology*, **224**(11) (2003) 115–126.

[11] Karthik Panchanathan and Robert Boyd, "Indirect Reciprocity Can Stabilize Cooperation without the Second-order Free Rider Problem," *Nature*, **432**(7016) (2004) 499–502.

[12] R. Boyd and P. J. Richerson, "Group Beneficial Norms Can Spread Rapidly in a Structured Population," *Journal of Theoretical Biology*, **215**(3) (2002) 287–296.

[13] H. Ohtsuki, C. Hauert, E. Lieberman, and MA Nowak, "A Simple Rule for the Evolution of Cooperation on Graphs and Social Networks," *Nature*, **441**(7092) (2006) 502–505.

[14] Stanley Milgram, "The Small World Problem," *Psychology Today*, **2** (1967) 60–67.

[15] Duncan J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness* (Princeton University Press, 1998).

[16] Duncan J. Watts and Steven H. Strogatz, "Collective Dynamics of 'Small-World' Networks," *Nature*, **393**(4) (1998) 440–442.

Appendix

Algorithm for the public-aid game

Algorithm 2 Playing a round of the public-aid game on a general graph. δ is the potential donor, ρ is the potential recipient, $\text{standing}_y(x)$ is the standing of x with respect to y, $\Gamma(x)$ is the neighborhood of node x, and α is the probability of error. rand generates a number between 0 and 1 uniformly at random

```
for δ = 1 to n do
  Choose ρ in Γ(δ) uniformly at random {choose a potential recipient}
  if type(δ)=ALLC then
    if rand() ≤ 1 − α then
      give b to ρ at a cost c to δ {make an unconditional donation}
      update standing_y(δ) to good for all agents y in σ(δ,ρ)
    else
      update standing_y(δ) to bad for all agents y in σ(δ,ρ) that consider ρ in
      good standing.
    end if
  else if type(δ)=ALLD then
    update standing_y(δ) to bad for all agents y in σ(δ,ρ) that consider ρ in
    good standing.
  else if type(δ)=RDISC then
    if rand() ≤ 1 − α then
      if standing_δ(ρ)=good then
        give b to ρ at a cost c to δ {make an unconditional donation}
        update standing_y(δ) to good for all agents y in σ(δ,ρ)
      else
        update standing_y(δ) to bad for all agents y in σ(δ,ρ) that consider ρ
        in good standing.
      end if
    else
      update standing_y(δ) to bad for all agents y in σ(δ,ρ) that consider ρ in
      good standing.
    end if
  end if
end for
```