The Evolution of Cooperation in Self-Interested Agent Societies: A Critical Study

Lisa-Maria Hofmann
Karlsruhe Institute of
Technology
Karlsruhe, Germany
lisamariahofmann@gmail.com

Nilanjan Chakraborty
School of Computer Science
Carnegie Mellon University
Pittsburgh, USA
nilanjan@cs.cmu.edu

Katia Sycara
School of Computer Science
Carnegie Mellon University
Pittsburgh, USA
katia@cs.cmu.edu

ABSTRACT

We study the phenomenon of evolution of cooperation in a society of self-interested agents using repeated games in graphs. A repeated game in a graph is a multiple round game, where, in each round, an agent gains payoff by playing a game with its neighbors and updates its action (state) by using the actions and/or payoffs of its neighbors. The interaction model between the agents is a two-player, two-action (cooperate and defect) Prisoner's Dilemma (PD) game (a prototypical model for interaction between self-interested agents). The conventional wisdom is that the presence of network structure enhances cooperation and current models use multiagent simulation to show evolution of cooperation. However, these results are based on particular combination of interaction game, network model and state update rules (e.g., PD game on a grid with imitate your best neighbor rule leads to evolution of cooperation). The state-of-the-art lacks a comprehensive picture of the dependence of the emergence of cooperation on model parameters like network topology, interaction game, state update rules and initial fraction of cooperators. We perform a thorough study of the phenomenon of evolution of cooperation using (a) a set of popular categories of networks, namely, grid, random networks, scale-free networks, and small-world networks and (b) a set of cognitively motivated update rules. Our simulation results show that the evolution of cooperation in networked systems is quite nuanced and depends on the combination of network type, update rules and the initial fraction of cooperating agents. We also provide an analysis to support our simulation results.

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Multiagent systems

General Terms

Experimentation

Keywords

Emergent behavior, Evolution of Cooperation, Repeated Games on Graphs

1. INTRODUCTION

The emergence of cooperation in a system of interacting self-interested agents has been studied in social science [3], evolutionary biology [13] and physics [18]. Examples of evolution of cooperation can be seen in natural systems including cellular structures like RNA [20], microbial organisms [6], animals [17], and humans [2]. The interaction model is a Prisoners' Dilemma (PD) game, which is a well-known game-theoretical model to study social dilemma situations among rational, self-interested, utility maximizing agents. Each player has two actions (or strategies): cooperate and defect. Defect is a dominant action, i.e., the payoff for playing defect is higher irrespective of the opponent's action. Thus, in the one-shot version, both players should always choose to defect, which is the only Nash equilibrium of the game. However, the equilibrium is not Pareto-efficient, i.e., both players would be better off by choosing to cooperate. Hence, a *social dilemma* arises. This contradicts the observed phenomenon of cooperation in human experiments. Repeated interaction was shown to be one of the factors for evolution of cooperation in two-agent PD games [3]. However, in multiagent interaction, evolutionary game theory has shown that in a big (or infinite) population, if players have repeated random encounters, a population of cooperators cannot resist invasion by defectors, and thus cooperation cannot survive. Defection is the only evolutionary stable strategy. Subsequently, it was shown that if the interaction between the players has a network structure, cooperation emerges and can be sustained. This phenomenon was initially shown (via multiagent simulations) for grids [13] and later for scale-free networks [16] or graphs with adaptive topology [22]. In this paper, we perform simulation studies to critically examine the following question: *Under what conditions does cooperation emerge in a network of interacting agents?*

A repeated PD game proceeds in multiple rounds. In each round, an agent plays the game with all its neighbors and earns the aggregate payoff of all the games. The agent uses the payoff of its neighbors (including self) to decide the action for the next round. Nowak and May [13] used this model to show evolution of cooperation in a system of agents organized in a grid and used *imitate-best-neighbor* as a deterministic update rule. Subsequent work showed emergence of cooperation with the agents organized according to different network structure and using different update rules (see [16], and [18] for a review). However, these results are based on particular combination of interaction game, network model, and state update rules. The state-of-the art lacks a compre-

hensive picture of the dependence of the emergence of cooperation on the model parameters like the network topology, the update rules and the initial fraction of cooperators.

Our motivation for studying evolution of cooperation is two-fold. First, we want to understand the reasons behind evolution of cooperation in self-interested agents in natural systems. The complementary sociological question of emergence of conflict in a society of humans can also be studied in the same framework [11]. The second motivation comes from the design of autonomous artificial multiagent societies (e.g., an autonomous robot colony operating on extra-planetary surfaces). Social dilemma situations where an individual robot objective is in conflict with the social objective may arise and it is impractical for a designer to foresee every possible situation. An alternate way is to design protocols that ensure cooperation among the agents in social dilemma situations. In this paper, we will not concern ourselves with the applications aspect. We will perform simulation studies to characterize the parameters and provide basic understanding of situations under which cooperation emerges in a multiagent society.

For the multiagent society, we assume simple agents that are myopic, of bounded rationality, organized according to a graph with fixed topology and repeatedly play a PD game with each other. We study evolution of cooperation using (a) a set of popular categories of networks, namely, grids, random, scale-free, and small-world networks and (b) a set of cognitively motivated state (or action) update rules. The rules we use are both deterministic and stochastic in nature. Cooperation is said to evolve in a society if the initial fraction of cooperators is lower than the final fraction of cooperators. We show by simulation that the phenomenon of evolution of cooperation is quite nuanced and depends on the graph topology, the initial fraction of cooperators, and the state update rule. In particular, we show that using the imitate-best-neighbor rule (as used in [13]), cooperation evolves in grids or scale-free networks for $d > 0.3$ but not in random or small-world networks (where d is the initial fraction of cooperators). We also show that the stochastic update rule used in [16] works only for scale-free networks and not for other types. This is significant because it shows that using the same update rule may not work across all network topologies. The update rules that show uniform performance irrespective of the network topology are (a) imitate the best action in your neighborhood (BS) and (b) win stay, lose shift (WSLS). BS ensures emergence of cooperation for $d \geq 0.6$, whereas WSLS ensures evolution of cooperation for $d \leq 0.5$. Moreover, for a given network type WSLS leads to the same final fraction of cooperators irrespective of the initial fraction. Although WSLS was shown to be a winning strategy update rule in two-player games, to the best of our knowledge, this rule has not been used in multiplayer repeated games. We believe that our characterization of the conditions under which cooperation evolves gives a more complete picture about emergence of cooperation for repeatedly interacting networked agents. This is the primary contribution of our work.

This paper is organized as follows: In Section 2, we discuss the relevant literature and in Section 3, we define our mathematical model including the network structures and state update rules used in the paper. In Section 4, we describe our simulation setup and in Section 5 we present our findings. In Section 6 we present our conclusions and outline future research directions.

2. RELATED WORK

The literature on using repeated games for studying evolution of cooperation among self-interested agents, can be classified according to the the number of players, interaction game model, and the interaction structure of the players. Game play can be between two players or between multiple players. In the multiagent setting, the agents may form an unstructured population where players randomly interact with each other or there may be structured interaction between them. For structured interaction, the interaction network may be of fixed or variable topology [22]. Both PD and the snowdrift game [7] has been used as the interaction model between agents, although (arguably) the PD game is more popular. For two-player games Axelrod first showed in a computer tournament that state update rules that rely on reciprocal altruism, such as *tit-for-tat*, where a player starts with cooperation and then imitates its opponent, can lead to the evolution of cooperation [2]. Similar results have been obtained for *win-stay, lose-shift* [12]. In this work, we concentrate on repeated PD games in population of agents whose interaction network has a fixed topology. Therefore we will restrict our review to repeated PD games in graphs.

Nowak and May [13] first demonstrated that cooperation evolves for memoryless agents playing repeated PD game with their 8 neighbors in a two-dimensional grid. The update rule used was deterministic imitate-best-neighbor. They show that cooperation evolves over a wide range of payoff parameters and the final fraction of cooperators is independent of the initial fraction. They also note that cooperators and defectors exist in clusters (or patterns) and the patterns are unstable against small random perturbations [10]. Subsequent research has tried to replicate the evolution of cooperation in different networks and using different update rules [16, 19, 5, 1]. A comprehensive review on evolutionary games in graphs including repeated games in graphs is given in [18].

Santos et al. [16] investigate the influence of Barábasi-Albert scale-free networks on cooperative behavior in comparison to complete, single-scale and random scale-free networks and show a clear rise in the final fraction of cooperators with the heterogeneity of the degrees. The update rule is a stochastic imitation rule (rule SA in Section 3). Tang et al. [5] demonstrates that there exist optimal values of the average degree for each kind of network leading to the best cooperation level. They test random, Barábasi-Albert scale-free, and Newman-Watts small-world graphs under a stochastic update rule that depends on the normalized payoff difference to a randomly chosen neighbor. They show via simulation that there is an optimal degree for cooperation in each network which is quite constant over a certain range of T (payoff for defecting when the opponent cooperates). Cooperation is highest for small average degrees ranging from 3 to 8. However, this is only done for an initial fraction of 0.5, a stochastic update rule, and 10 different realizations of the particular graph. The results on evolution of cooperation have been usually obtained on different networks using a particular state update rule. The concern that changing the state update rule may affect the evolution of cooperation has not been addressed in the literature. Therefore, we study the evolution of cooperation across a variety of networks with different update rules.

There has also been work on repeated PD games in graphs with variable topology [22, 8]. In [22], the initial graph is assumed to be a random network and the agents are allowed to (stochastically) break links with their neighbors playing defect and form a new link with their neighbor's neighbor. The authors show that this boosts cooperation in the society. In this paper, we do not consider variable graph topology. A study similar to ours can be done for networks with variable topology and we keep this as a future work.

3. PROBLEM MODEL

3.1 Network Models

The agent interactions can be encoded as an undirected graph $G = (V, E)$ where $V = \{v_1, v_2, \ldots, v_n\}$ are a set of n nodes (or agents) and $E \subseteq V \times V$ is a set of edges. The graph topology is fixed throughout the game. Two agents v_i and v_j are neighbors if $(v_i, v_j) \in E$. $\mathcal{N}(i) = \{v_j | (v_i, v_j) \in E\} \subset V$ is the set of v_i's neighbors and $|\mathcal{N}(i)|$ is the *degree* of node v_i. $\mathcal{N}^+(i) = \mathcal{N}(i) \cup \{v_i\}$.

We use four different graph types for the simulations:

Scale-free network: In a scale-free graph, the distribution of node degree follow a power law, $N_d \propto d^{-\gamma}$, where N_d is the number of nodes of degree d and $\gamma > 0$ is a constant (typically $\gamma \in [2, 3]$). We use the Barabási-Albert model with average degree 4 [4].

Small-world network: A small-world graph shows a high clustering coefficient (as defined in [21]) and a short average path length. We use the Watts-Strogatz model with average degree 4 [21]. First, a ring is built and each node is connected to the 2 neighboring sites on each side. Then, links are randomly released and reconnected to other nodes. We set the rewiring probability to 0.2, which leads to an average degree of roughly 4.

Random network: A network where a link between nodes is set with a predefined probability p. The probability that a vertex v_i has k_i neighbors follows a binomial distribution $B(n - 1, p)$. For large n and $p \leq 0.05$ the degree distribution can be approximated by a Poisson distribution $Prob(k_i = k) = \exp(\lambda) \cdot \frac{\lambda^k}{k!}$ with $\lambda = n \cdot p$. We set $p = 0.05$ to ensure connectedness. The clustering coefficient is usually low.

Grid: A grid is a two-dimensional lattice where each inner player has 4 neighbors, each boundary player 3 and each corner player 2. The clustering coefficient is 0.

3.2 Repeated PD Games in Networks

A PD game is a two-player game where each agent has two actions, $\mathcal{A} \hat{=} \{\text{cooperate}(1), \text{defect}(0)\}$. The payoffs for two players are symmetric with the payoff matrix entries

	1	0
1	R	S
0	T	P

For a PD, $T > P > R > S$ holds and for repeated PD games $T + S < 2R$. We assume $R = 1, P = 0.1, S = 0$ with the incentive to defect T being the only parameter.

In a repeated PD game in a network, there are n-players that form the nodes of the graph and the game proceeds in

rounds. Each round has two phases: (a) In the game playing phase the players play the PD game with all their neighbors with a fixed strategy and compute their total payoff. (b) In the strategy update phase, each player updates its strategies according to the same *action update rule*. Such a rule might be a function of the neighbors' states, payoffs and/or the agent's own state and payoff. In our model the action update rule is synchronous.

Let $s_i(t)$ denote the state of player i at round t. The total payoff, $p_i(t)$, is the sum of the payoffs of the separate games in player i's neighborhood $\mathcal{N}(i)$:

$$p_i(t) = \sum_{j \in \mathcal{N}_i} [Rs_i(t)s_j(t) + T(1 - s_i(t))s_j(t) +$$
$$S(1 - s_j(t))s_i(t) + P(1 - s_i(t))(1 - s_j(t))]$$

3.3 State Update Rules and Convergence

In each round, the agents update their states according to a common state update rule. The rules that we use can be classified along two axes: innovative or non-innovative and deterministic or stochastic. Rules that use states already existing in the neighborhood are non-innovative (e.g., imitate-best-neighbor or imitate-best-strategy) whereas rules that can switch to a strategy not in their neighborhood are called innovative rules (e.g., win-stay, lose-shift). We use the following rules:

Imitate-best-neighbor (IB): Each agent imitates the action of the wealthiest agent (including itself) in the next round. If two or more players have the same payoff, the agent chooses randomly between them. The state update for agent i can be formalized as

$$s_i(t) = s_j(t - 1) \text{ where } j = \arg \max_{k \in \mathcal{N}^+(i)} (p_k(t - 1))$$

Imitate-best-strategy (BS): An agent copies the strategy that accumulates the highest payoff in its neighborhood. Each agent sums up the payoff of all cooperating as well as the payoff of all defecting neighbors including itself.

Let agent i play strategy s_1 in round $t - 1$ and have q_i neighbors. We denote its neighbors playing strategy s_1 and i itself by G_1, where $|G_1| = n_1$. The neighbors playing s_2 are denoted by G_2, where $|G_2| = n_2$. It holds that $G_1 \cup G_2 = \mathcal{N}_i^+$ and $n_1 + n_2 = q_i + 1$. Let w be the probability of switching. The update rule in any round t is as follows:

$$w = \begin{cases} 1 & \text{if } \sum_{i \in G_1} p_{t-1}(i) < \sum_{k \in G_2} p_{t-1}(k) \\ 0 & \text{otherwise} \end{cases}$$

Win-stay, lose-shift (WSLS): In a multiplayer setting, a strategy is maintained only if the current payoff p is at least as high as in the former round. We need to introduce a short-term (one-round) memory in order to calculate the payoff difference. In our case, there are only two possible strategies s_1 and s_2. In any round t the update rule is

$$w = \begin{cases} 1 & \text{if } p_{t-1} < p_{t-2} \\ 0 & \text{if } p_{t-1} \geq p_{t-2} \end{cases}$$

Stochastic imitate-best-neighbor (stIB): This rule represents a stochastic version of the IB rule. Each agent i picks the best neighbor j in $\mathcal{N}^+(i)$ and imitates its strategy with a probability w depending on the payoff difference $\Delta p_{i,j}$: $w = 1/(1 + exp(-\beta \Delta p_{i,j}))$. In test runs, $\beta = 0.75$ gives a
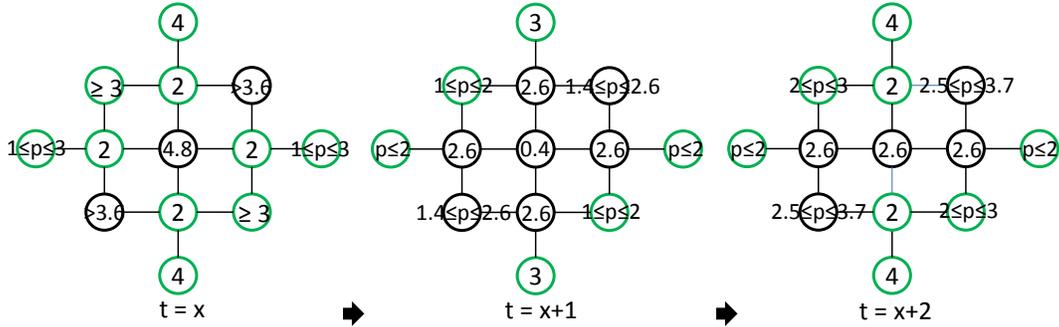
Figure 1: A typical oscillatory state with the IB rule. Green nodes represent cooperators, black defectors, the number in a node corresponds to its payoff p. We only display relevant part of the grid. The state of period x will be repeated in period $x + 3$.

reasonable trade-off between the payoff difference and the probability to switch.

Stochastic imitate-best-strategy (stBS): The strategy that yields a higher payoff in a neighborhood is imitated with probability of its total payoff divided by the total payoff in the neighborhood. Otherwise, the player keeps its current strategy.

Stochastic win-stay, lose-shift (stWSLS): If a player's payoff deteriorates in round t, the player will switch strategies with a probability w depending on the difference of its current and last payoff $\Delta p_{t-1,t}$: $w = 1/(1 + exp(-\beta \Delta p_{t-1,t}))$ with $\beta = 0.75$.

Stochastic imitate-random-neighbor (SA): The rule is described and used in [16]: for each i, one neighbor j among all k_i neighbors is picked at random. Only if $p_j > p_i$, i imitates j's strategy with probability $(p_j - p_i)/k_> D_>$, where $k_> = \max(k_i, k_j)$ and $D_> = \min(T, 1) - \max(S, 0)$.

The stochastic rules that we use are counterparts of our deterministic rules (except for the SA rule, taken from [16]). The deterministic IB rule is taken from [13] and the WSLS is taken from [12]. These rules are simple heuristics that have been shown to be used by humans for decision making under certain circumstances. Note that we have not used the best response strategy because for our model it always leads to evolution of defection among all agents. The imitating rules also have an evolutionary biology interpretation. Instead of a player updating its state, we can say that in each round, neighbors are competing against each other for occupying the empty node in their middle. The player with highest payoff, i.e., *fittest* player wins and its strategy gets replicated (with some probability in stochastic updates).

Steady States: Since our repeated game model is a dynamical system and we will use simulations to study the evolution of cooperation it is important to understand the convergence properties of the system to design appropriate stopping criterion for simulations. Note that the all-cooperate and all-defect solutions are trivial steady states for all the state update rules. For deterministic rules, a steady state is reached if the concatenated strategy vector (state vector) of all agents repeats itself $\mathbf{s}_t = \mathbf{s}_{t-1}$. For our system, we can show that we may not reach a steady state. Figure 1 shows a simple example demonstrating that oscillations can occur in a repeated PD game in grids. Figure 1, shows the part of a grid network where players will keep changing strategies. For some boundary nodes we give ranges for their payoff. As long as these payoff requirements are fulfilled, we do not have to consider any further players. Simple calculation shows that the system will oscillate.

For stochastic rules, the notion of convergence is different as the state vector represents the realization of the current *probability to cooperate* of each player. Thus, in the strict sense, this probability has to stay within a certain range for each player over time to ensure convergence. From our results we see that this does not happen as the current probability in a particular round t does not depend on the one of the round $t - 1$, but on the realization of this probability. A more simple criterion could be that running averages of f_c for each player do not change much.

4. SETUP OF THE SIMULATIONS

We test three deterministic and four stochastic update rules on four different networks: scale-free, small-world, grid and random networks. Three stochastic rules are counterparts of the deterministic versions, the fourth comes from the literature. We call the combination of a particular update rule, graph and initial fraction of cooperators a *setting*. For each setting, we perform 100 runs, each with a different realization of the graph if there is a stochastic component in the setup (except for grids, where there is no stochastic component). We compute the average final fraction of cooperators, f_c, over all the runs and also compute the standard deviation over the final fraction of cooperators, σ. For all the results that we present σ is quite low except where we explicitly mention. For a given initial fraction of cooperators, each player is randomly assigned the action cooperate or defect such that the ratio of total number of cooperators to defectors is equal to the given fraction.

Stopping criterion for simulations: For deterministic rules, we simulate for a maximum of $t = 500$ rounds. If the simulation converges, we take the last state as final result. If not, we average over the last five rounds. For stochastic update rules, we simulate over $t = 5000$ rounds and average over the last 100 rounds. The number of rounds to average over was heuristically determined after finding that the deviation over the last few runs usually is very low.

We use values of ($T \in \{1.1, 1.2\}$) for most of our simulations, except for grids. Higher values of T will give rise to more defection. We test several initial fractions of cooperators d ($d \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$). In some cases we test additional values in order to determine more exact thresh-

olds or to point out differences between certain settings.

In forming our graphs we ensure that all of them are connected. We simulate on graphs with 750 and 1000 nodes. The IB rule in random networks is the only setting where we find the scaling of a network to change results for f_c, because the average degree changes with the number of nodes and the number of neighbors matters in imitating rules. Therefore, simulations of this setting have be to made with caution and be tested for different levels of n and p. In all other settings, differences in f_c between $n = 750$ and $n = 1000$ are smaller than 5%. Research about the influence of the average degree states that an increasing average degree usually leads to less cooperation [14].
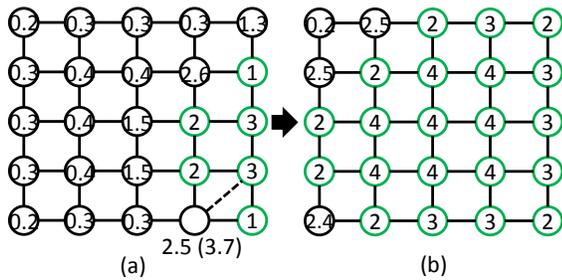
5. SIMULATION RESULTS

Figure 2: In a grid where the dotted link does not exist, the defector only yields a payoff of 2.5 **and cooperation will spread until it reaches state (b) with** $f_c = 84\%$ **after** 5 **rounds. Adding only one link to (a) and therefore increasing the clustering coefficient increases the defector's payoff to** 3.7 **and leads to** 100% **defection after** 3 **rounds (**$T = 1.2$**).**

We want to have a general insight about conditions under which there is *evolution of cooperation*, i.e., the final fraction of cooperators f_c is higher than the initial fraction d. Table 1 gives the general findings. Most of the literature focuses on whether cooperation emerges at all (by looking at the final fraction of cooperators), whereas we want to focus on the relation of the final fraction of cooperators to the initial fraction, which has not received much attention so far. Comparing the obtained results of $T = 1.1$ and $T = 1.2$, there is hardly any difference except for random networks. We will come back to this issue in 5.4. Thus, we only display results for $T = 1.2$ in Figure 3. In grids, there is no difference for the IB rule, because the interaction structure is very simple. Possible constellations of payoffs and strategies are very limited and stay the same for these two levels of T. If we set $T = 1.3$ in grids, one new constellation that helps defectors at the first glance actually leads to the collapse of bigger clusters of defectors and therefore to higher final fractions of cooperators. From simulations we see that this development abruptly ends starting from $T = 1.4$, where defectors are better off in most constellations and f_c drops drastically.

5.1 Scale-Free Networks

In scale-free networks, all the state-update rules that we study show emergence of cooperation for different ranges of the initial fraction of cooperators. From Figure 3, we see

that the IB rule leads to evolution of cooperation for both deterministic and stochastic versions. For the deterministic IB rule, we further observe, that even though most of the runs converge, there is a large standard deviation $\sigma = 0.15$ to 0.46. This is because the distribution of f_c is bimodal: either f_c drops to zero or reaches very high values. Averaging over 100 runs gives the obtained high levels. It has already been pointed out that the important factor for cooperation or defection in scale-free networks is the behavior of high-degree nodes [16]. If a high degree node defects it can exploit all linked cooperators and gain a high payoff. Imitating the best, all its neighbors will switch to defection. From this state onwards it is not likely that defectors find a wealthier cooperator as cooperators surrounded by defectors do not obtain payoffs. Additionally, a high-degree node still accumulates relatively high payoffs even for defect-defect links because of sheer number. In simulations where we apply a normalization of payoffs by the number of neighbors a drastic drop of f_c is observed and there is no evolution of cooperation. Another indicator revealing the importance of the heterogeneity in degrees is that we find the highest degree node cooperating in 100% of the runs with high final levels and defecting in all runs with a very low final percentage of cooperators. If there are cooperators in the cases with low f_c, they usually occur in a cluster around a wealthy cooperator.

For the SA rule used by Santos et al [15], we find a higher final fraction of cooperators than with any IB rule in scale-free networks. Like the deterministic IB rule, in this case, the standard deviation σ is very high, but unlike the IB rule the distribution is not bimodal. Although the SA rule performs well for scale-free networks, it does not lead to any cooperation in grids and random networks and only small to levels in small-world graphs.

5.2 Grids

In grids we find high cooperation rates with the IB rule. A closer look at the dynamics shows that clustering is the crucial factor for the success of cooperators in grids as already pointed out in [13]. A cluster of cooperators is a set of cooperating nodes that are connected to each other. Boundary players of clusters of cooperators are the nodes that are also linked to defectors. In settings with low values of T, a 2×2 cluster of cooperators already leads to propagation of cooperation through the whole network. There are several possibilities for defectors to survive, e.g., in corners, in a line of 4 players or in several spatial structures. However, too big clusters of defectors become unstable at low levels of T. For clusters of cooperators, the well-defined interaction with defecting neighbors originating from the typical grid structure, is helpful as defectors cannot exploit cooperators in the middle of clusters. Cooperators on the boundary will not turn into defectors as long as the neighboring defector has less than four cooperating neighbors. The cooperator in the middle of the cluster is the wealthiest player and backs up the boundary cooperators.

However, note that the stochastic IB rule does not lead to cooperation in grids. The success of cooperators depends on the formation of clusters. However, cooperators in the middle of clusters may randomly turn into defectors. A single defector surrounded by only cooperators can turn all its neighbors into defectors. The above intuition is also true for the SA rule and hence it does not lead to cooperation.

Table 1: Summary of the evolution of cooperation with different update rules and networks. *Yes* denotes that there is evolution of cooperation with the range of the initial fraction of cooperators d for cooperation to emerge given in parentheses. Thresholds for d are indicative and not exact.

Rule/Graph	Scale-free	Small-world	Grid	Random
IB	yes ($d > 0.3$)	no	yes ($0.3 \leq d < 0.9$)	no
BS	yes ($d \geq 0.6$)			
WSLS	yes ($d \leq 0.5$)			yes ($d \leq 0.7$)
stIB	yes ($d > 0.5$)	yes ($d > 0.5$)	no	no
stBS	yes ($d > 0.5$)	yes ($d > 0.2$)	yes ($d > 0.2$)	yes ($d > 0.7$)
stWSLS	yes ($d < 0.45$)	yes ($d < 0.5$)	yes ($d < 0.45$)	yes ($d < 0.65$)
SA	yes	no	no	no

However, for the BS rule, a single defector surrounded by cooperators cannot destroy the cluster and so the stochastic BS rule leads to evolution of cooperation for $d \geq 0.2$. In fact, the stochastic version outperforms the deterministic BS rule. The propagation of defection by single defectors can also account for the fact that we do not find evolution of cooperation in grids with IB for a very high initial fraction of cooperators ($d = 0.9$). Defectors most likely do not appear in clusters and can therefore exploit all cooperating neighbors at once, which gives them a high payoff and leads to defection in their neighborhoods.

5.3   Small-World Networks

In Table 1 we see that there hardly is emergence of cooperation in small-world networks with the IB rule. This is especially interesting because the main graph features as the average degree and its standard deviation are almost the same as in grids, where the IB rule leads to cooperation. To discover reasons for the differences we have to look at the clustering coefficient c. c is very high in small-worlds ($c \approx 27\%$) in comparison to grids, where $c = 0$. We have seen before that well-defined boundaries between groups of cooperators and defectors in grids help to propagate cooperation. The Watts-Strogatz model constructs a small-world graph starting from a ring. The rewiring process creates shortcuts between different neighborhoods and can turn inner players into boundary players if the shortcut links them to defectors. Thus, some small clusters of cooperators that would have grown in grids cannot grow in small-worlds. Figure 2 gives an example how a slightly higher clustering coefficient leads to $f_c = 0$ instead of $f_c = 84\%$ for $c = 0$.

The final fraction of cooperators usually slightly drops from the initial fraction in small-worlds. Even if f_c increases slightly over the starting point for medium levels of d, we do not consider this as evolution of cooperation because the standard deviation ranges from 0.04 to 0.16. The stochastic BS rule yields 100% cooperation with small initial fractions and turns out to be the most successful rule in small-worlds. Starting from a medium level of d, the deterministic version leads to evolution of cooperation, too. The deterministic and the stochastic WSLS yield a medium level of cooperation.

5.4   Random Networks

Random networks are the only network where simulations with $T = 1.1$ and $T = 1.2$ yield different results. However, we do not consider results for $T = 1.1$ to be reliable because of high standard deviations and low convergence rates. Therefore, we discuss the results for $T = 1.2$ which show 100% convergence and a lower σ.

The WSLS rule yields the highest levels of cooperation in random graphs compared to other networks. The IB rule hardly leads to cooperation in random networks. The results for the BS rule are drastic, as f_c turns out to be either 0 or 1. We see from the simulation results that the jump occurs between $d = 0.55$ and $d = 0.6$. We will give the explanation in the discussion of the BS rule. The stochastic version yields the best result for random graphs.

5.5   The Imitate-Best-Strategy Rule

In all settings, we find that the BS rule does not lead to evolution of cooperation for any initial fraction, $d \leq 0.5$, whereas it takes place for all $d \geq 0.6$. This phenomenon is extraordinarily strong in random networks, where the final fraction of cooperators, f_c jumps from 0 to 1. Further simulations indicate that there is a threshold for the evolution of cooperation that occurs between $d = 0.55$ and $d = 0.65$.

We now estimate analytically the threshold value of the parameter d for cooperation to emerge in scale-free graphs, small-worlds, and grids, under some simplifying assumptions. In our model, each player i has 4 neighbors on average. Let x be the fraction of cooperating neighbors of a node i and let each neighbor in $\mathcal{N}(i)$ have 4 neighbors with y the fraction of cooperators in their neighborhoods (i's two-step neighborhood) a constant (but need not be the same as x). We assume that x is representative of the whole network, i.e., the fraction of cooperators in the network at that round is x. Since the average degree is 4, we assume that x can take the values $0, 0.25, 0.5, 0.75, 1$. Recall that in this rule, a player decides which strategy to take according to the wealthiest strategy in its neighborhood. In general, i will cooperate if

$$x(yR + (1-y)S) > (1-x)(yT + (1-y)P) \qquad (1)$$

which is $0.56(0.55)$ for $x = y$ and $T = 1.2(1.1)$. From Equation 1, we also find that i will always prefer to play defect for $x \leq 0.5$ (irrespective of the value of y). However, for $x = 0.75$, i will play cooperate in every case where $y \geq 0.25$ (which is very likely, as x is representative of the whole network). Thus, from this simple analysis, we predict a threshold value of $d = 0.55$ for cooperation to emerge. Note that the argument above accounts for a simple average-case because it assumes the degree of each node to be 4 and an even distribution of cooperators in all two-step neighborhoods. As the actual x and y for a given neighborhood can differ from the average we find our simulation threshold to be slightly different from the predicted value of $d = 0.55$.

For random networks also, any agent i will cooperate if Equation 1 holds, i.e., an agent i will cooperate for $x \geq 0.56$

when $x = y$ and $T = 1.2$. Since the average degree is not 4 the calculation of the values of x for which cooperation emerges is more complicated for $x \neq y$. Note, that the number of cooperators in i's neighborhood follows a binomial distribution $\#\text{coop} \sim \text{bin}(n-1, d)$. Using Equation 1, for a given x, we can calculate the value of y required for i to be cooperating. Thereafter using the binomial distribution, we can compute the probability that such a fraction y will exist in i's two-neighborhood. For example, if $x = 0.47$, y should be ≥ 0.58. However the probability that $y \geq 0.58$ is equal to 0.046. Thus it is unlikely for agent i to play cooperate. However, for any $x \geq 0.56$ cooperation is very likely. For example, for $x = 0.61$, y should be more than 0.47 and the probability for $y \geq 0.47$ is 0.93. Thus, in this case also a threshold value of $d = 0.56$ would ensure cooperation, which is in good agreement with our simulations.

5.6 The Win-Stay, Lose-Shift Rule

Although the WSLS rule was shown to perform very well in two-agent settings it has not been investigated in multi-player settings. An interesting aspect of WSLS is that for every network, it leads to the same f_c irrespective of the initial d. However, the actual value of f_c reached depends on the type of network. Another surprising aspect is that WSLS always leads to evolution of cooperation, if $d \leq 0.5$ and is the only strategy to do so across all the types of networks studied. We note that the update for an agent depends on the own payoff over time and therefore indirectly on strategy distributions of the neighbors. Furthermore, the rule is innovative, such that defectors surrounded by defectors are still able to change to cooperation, which is never possible in imitating rules. We found examples of how parts of a network can easily turn from all-defection to all-cooperation and vice versa in several time steps only.

We note that for the WSLS strategy most of the simulations do not converge. For the deterministic WSLS, we do not find convergence except in random networks for $T = 1.2$, where all runs converge (usually within 20 rounds). However, even though the runs do not converge, the standard deviation usually is lower than 0.016. Thus, we can have reasonable confidence about the correctness of our findings. We note that this update rule is especially helpful in systems where one does not have an influence on the initial fraction of cooperators but wants to ensure a medium level of cooperation. The stochastic WSLS yields slightly higher levels of cooperation in grids, scale-free, and small-world networks and here also the final fraction of cooperators is constant (independent of the initial fraction). Here, the standard deviation is lower than 0.01.

6. CONCLUSIONS

In this paper, we performed a comprehensive simulation study of the phenomenon of evolution of cooperation in self-interested multiagent societies. Our research shows that general statements on evolution of cooperation in networked multiagent systems cannot be made. The emergence of cooperation depends on the type of network, the state update rule, and the initial fraction of cooperators. We find a high dependency of final results on the initial fraction especially in imitating, non-innovative rules. We observe that the evolution phenomenon do not depend on the size of the network as long as the network is large enough to show its typical properties and crucial network parameters do not change

with the number of nodes. Our main findings are as follows:

- In scale-free networks, almost all the state update rules lead to evolution of cooperation. However, the deterministic imitation rule and stochastic imitation rule of [16] perform better.

- For small-world networks stochastic BS performs best.

- For grids the deterministic IB performs the best and most stochastic rules (except stochastic BS) do not perform well.

- For random networks WSLS performs the best.

- WSLS gives the interesting result that for every type of network we studied, the final fraction of cooperators reaches a constant value. Further, this is the only rule that ensures evolution of cooperation for low initial fraction of cooperators. This result holds across all types of networks.

- The BS rule also has the interesting property of supporting evolution of cooperation above a threshold value of initial fraction of cooperators across all networks.

We also find that stochastic versions of deterministic rules usually perform slightly better. The final results still highly depend on the network: e.g., rules that work very well in scale-free graphs do not have to be successful in grids. Furthermore, results for different stochastic rules can vary greatly in the same setting. In most cases we find them to yield similar results as the deterministic versions.

Future Work: In this paper we have considered the PD game as an interaction model with a fixed topology of interaction. An important future direction of research is to relax the assumption of fixed topology. Although versions of this problem has been studied [22], there is no restriction placed on the topology of the graph, except that it remains connected. An interesting extension would be to study the evolution of cooperation in variable topology graph where the statistical properties of the graph is maintained (i.e., a scale-free graph remains scale-free). Another future research agenda is to give a broad understanding of rules and networks for emergence of cooperation in the Snowdrift game. A first glance at pilot simulations also shows different behavior for different settings [9].

From the theoretical perspective there are a few interesting directions that can be pursued. The results obtained from WSLS seem to indicate some universal underlying phenomenon for the rule. Theoretical understanding of why there is a uniform final fraction of cooperators for WSLS in a given type of network is an important research direction. Moreover, here we have prescribed rules and tried to analyze whether the rules lead to evolution of cooperation. Designing a rule that guarantees a certain level of cooperation irrespective of the network topology is an important problem that we wish to pursue.

Acknowledgements

7. REFERENCES

[1] G. Abramson and M. Kuperman. Social games in a social network. *Phys. Rev. E*, 63(3):030901, Feb 2001.

[2] R. Axelrod. *The evolution of cooperation.* Basic Books, New York, NY, 1984.

[3] R. Axelrod and W. Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.

[4] A.-L. Barabasi and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286(5439):509–512, 1999.

[5] C.-L. Tang, W.-X. Wang, and X. Wu. Effects of average degree on cooperation in networked evolutionary game. *Eur. Phys. J.*, 53(3):411–415, 2006.

[6] B. J. Crespi. The evolution of social behavior in microorganisms. *Trends Ecol. Evol.*, 16:178–183, 2001.

[7] M. Doebeli. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 428:643–646, 2004.

[8] H. Ebel and S. Bornholdt. Coevolutionary games on networks. *Phys. Rev. E*, 66, 2002.

[9] L.-M. Hofmann. The evolution of cooperation in self-interested agent systems. Master's thesis, Karlsruhe Institute of Technology, 2011.

[10] Y. F. Lim, K. Chen, and C. Jayaprakash. Scale-invariant behavior in a spatial game of prisoners' dilemma. *Phys. Rev. E*, 65(2), 2002.

[11] L. Luo, N. Chakraborty, and K. Sycara. Modeling ethno-religious conflicts as prisoner's dilemma game in graphs. *Computational Science and Engineering, IEEE International Conference on*, 4:442–449, 2009.

[12] M. Nowak and K. Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364, 1993.

[13] M. A. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature*, 359:826, 1992.

[14] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505, 2006.

[15] F. C. Santos and J. M. Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters*, 95(9), 2005.

[16] F. C. Santos, J. M. Pacheco, and T. Lenaerts. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl. Acad. Sci. USA*, 103:3490–3494, 2006.

[17] J. M. Smith and G. R. Price. The logic of animal conflict. *Nature*, 246(2):15–18, 1973.

[18] G. Szabo and G. Fath. Evolutionary games on graphs. *Physics Reports*, 446(4-6):97–216, 2007.

[19] M. Tomassini, E. Pestelacci, and L. Luthi. Social dilemmas and cooperation in complex networks. *International Journal of Modern Physics C*, 18:1173–1185, 2007.

[20] G. J. Velicer. Social strife in the microbial world. *Trends in Microbiology*, 11(7):330–337, 2003.

[21] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.

[22] M. G. Zimmermann and V. M. Eguíluz. Cooperation, social networks, and the emergence of leadership in a prisoners dilemma with adaptive local interactions. *Phys. Rev. E*, 72(5):056118, 2005.
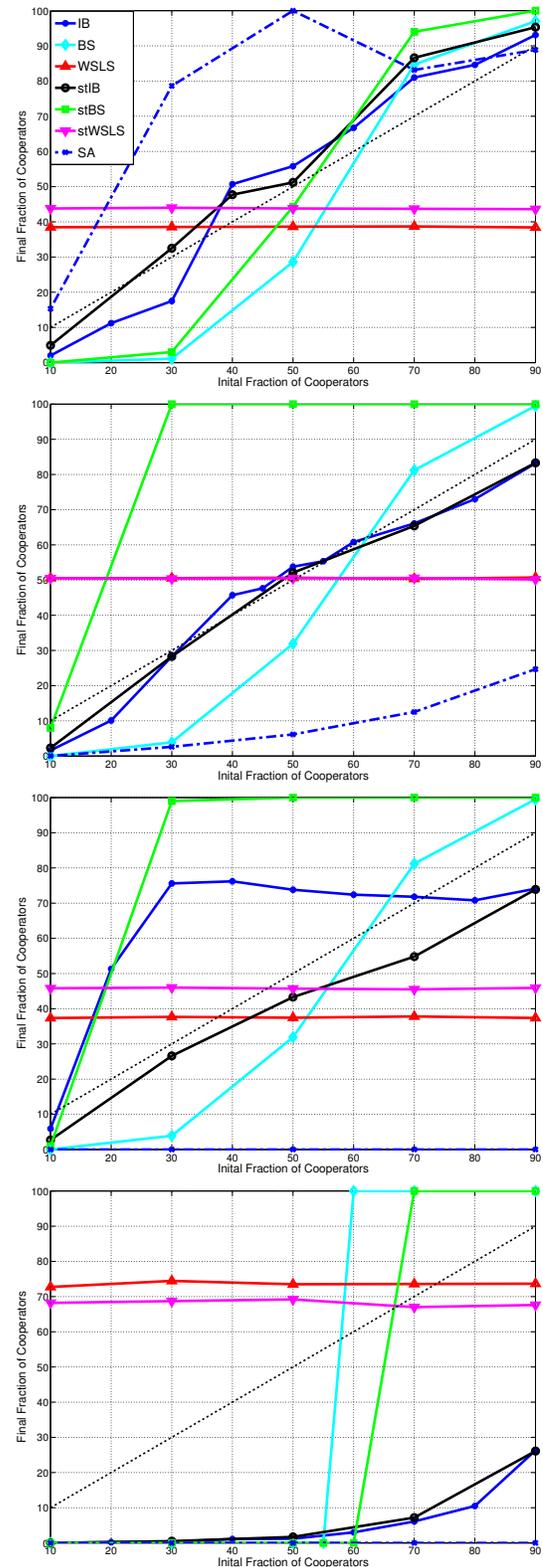
Figure 3: Final fraction of cooperators for $T = 1.2$, $n = 750$ and four different networks: scale-free, small-world, grid and random networks (in this order). Evolution of cooperation occurs where the final fraction is above the dotted black line.