Collective Evolution of Turn-taking Norm in Repeated Dispersion Games

Akira Namatame, Noriko Fujii (Tanoura)
Dept. of Computer Science,
National Defense Academy, Yokosuka, JAPAN
nama@nda.ac.jp

ABSTRACT. Using a game-theoretic model combined with the evolutionary model, we investigate the conditions under which the norm will emerge and sustain in various social interaction settings. We use adaptive models to understand the dynamics of networked agents that lead to efficient and fair outcomes in repeated dispersion games. We find that the efficient and fair outcome emerges relatively quickly in our model. The essential idea is to show how norms can emerge spontaneously at the collective level from the pair-wise interactions of evolving agents. Human subjects appear to easily recognize the possibility of a coordinated turn-taking behavior as a means to generate an efficient and fair outcome where desirable outcomes are achieved by asymmetric strategic choices with coordinated alternating reciprocity.
1. Introduction

The undesirable outcomes that none would have chosen may occur when social interactions of agents lead to a result which is not **optimal**. This problem is often referred to as a **coordination failure**. The reason why uncoordinated activities of agents pursuing their own ends often produce outcomes that all would seek to avoid is that each agent’s action affects the others and these effects are often not included in whatever optimizing process made by other agents. These unaccounted effects on others are called **externalities**. An externality is usually considered as an unaccounted side effect of activities by some agents on unrelated other agents.

An externality also occurs when individuals care about others’ choices and each individual’s choice affects others’ choices. For instance, when deciding which movies to visit, which new technologies to adopt or which job candidates to be decided, we often have little information with which to evaluate the alternatives. Therefore we rely on the recommendation of friends or simply pick the one to which most people are enjoying. Even when we have access to plentiful information, we often lack the ability to make sense of it and we rely on the advice of trusted friends or colleagues.

Social interaction with externalities raises two basic questions, a positive and a normative one. The first is how the outcome actually comes to exist, and the second one concerns what the ideal outcome should look like. From the perspective of a social planner, social interactions with externalities will result in an outcome that is not socially optimal.

Many sphere of social interactions are governed by **norms** such as **reciprocity** and **equity**. Social norms and conventions are the glue that holds a society of self-interested agents together. When social outcomes appear to be coordinated, the cause is most often not a remarkable concerted collective action, but individual adherence to well-understood and shared guides to normative behaviour. Although social norms can potentially serve useful constructs to understand human behaviour, there is little theory of norm development.

Durkheim (1982) writes that society is not the mere sum of individuals, but the system formed by their association represents a specific reality which has its own characteristics. Undoubtedly no collective entity can be produced if there are no individual consciousnesses: this is a necessary but not a sufficient condition. In addition, these consciousnesses must be associated and combined, but combined in a certain way. It is from this combination that social life arises and consequently it is this combination which explains it. By aggregating together, by interpenetrating, by fusing together, individuals give birth to a being, psychical if you will, but one which constitutes a psychical individuality of a new kind. Making norm obedience an argument of a utility function simply pushes back the explanatory problem. How did norms get there? Why do some social prescriptions become normative while others do not? When social interaction is repeated, reciprocity is rational if
individuals are sufficiently patient. But there are many distinct self-consistent social formations.

Norms are self-enforcing patterns of behaviour: it is in everyone’s interest to conform given the expectation that others are going to conform. If the situation is repeated, norm is a kind of equilibrium of the underlying games. It is a strategy choice rule that assigns a rule to each individual that is an optimal in the sense no one has an incentive to deviate from it. The introduction of genetic algorithms enabled researchers to investigate the natural selection of social norms using repeated games. The challenge for evolutionary approach is to identify emergent properties for a rich class of interaction mechanisms and to allow for more complex individual behaviours. The promise of this work is an account of the emergence of social norms that that satisfies the constraint of methodological individualism but which is not trivially reductive.

Evolutionary game theory offers one approach to the study of emergent properties in systems of interacting agents. In the context of evolutionary game theory, social norms are often understood as solutions to coordination problems which arise in a population of individuals large enough that most members never directly interact with one another, thus precluding the possibility of effective collective action. Darwinian dynamics based on mutation and selection form the core of models for evolution in nature. Evolution through natural selection is understood to imply improvement and progress. If multiple populations of agents are adapting each other, the result is a co-evolutionary process. The problem to contend with in co-evolution based on the Darwinian paradigm, however, is the possibility of an escalating arms race with no end. Competing agents might continually adapt to each other in more and more specialized ways, never stabilizing at a desirable outcome. Of particular interests is the question how social interactions can be restructured so that agents are free to choose their own actions while avoiding outcomes that none would have chosen.

Previous research has mainly focused on the Prisoner’s Dilemma (PD) game. The conflict between individual rationality and collective rationality problem has been largely solved by the theories of reciprocal altruism and indirect reciprocity, and computer simulations have shown that reciprocal strategies such as Tit for Tat (TFT) tend to evolve, resulting in widespread joint cooperation. The evolutionary modeling is to explain how TFT could have evolved as social norms, given that natural selection operates at the level of the individuals.

Kandori et al., (1993) studied the evolution of play in a population of players continually interacting with coordination games. Players are repeatedly and randomly matched against opponents which is known as random matching model. Players revise their strategic choices only at discrete random moments, each player independent of the others. The model is built from Poisson processes in continuous time, but individuals interact at discrete random moments. Individual rationality alone would suggest that some kind of coordination would occur. But the dynamics
of social interaction limit what kind of coordination can occur. For the pair-wise coordination problems studied, the only emergent state is risk-dominant coordination. However, there is no reason to believe that risk-dominant selection is universal. Hanaki et al., (2005), for instance, have shown that in at least some models where matching is endogenous, payoff-dominant selection can obtain.

A major shortcoming of the previous influential research is its focus on games in which cooperation in Prisoner’s Dilemma (PD) game or coordination in coordination games involves the agents acting similarly. There are games in which favourable payoffs are possible only if one player acts one way while the other acts the opposite way. For instance, to cooperate successfully, the agents have to alternate or take turns, out of phase with each other. A typical example is the battle of sexes game or labour-division game. If this type of interaction is repeated, then the agents benefit, in terms of natural selection, by coordinated alternation by taking turns in choosing one of the two strategies and there is evidence to show that this type of turn-taking occurs quite commonly in nature. Give and take or alternation is a strategy that is intuitive and simple, but even so it is beyond the scope of most traditional learning models.

Browning et al., (2004) investigate through agent-based simulation how this type of coordinated, alternating cooperation can evolve without any communication between players. Using a genetic algorithm, they showed that coordinated turn-taking can evolve in games with asymmetric Nash equilibria if the players benefit from it. However, precisely how coordination evolves without communication is not fully explained, although we have testable hypotheses about it. This project is designed to study the nature, properties and phenomena of coordinated alternating cooperation in a range of games with asymmetric equilibria and to attempt to find a mechanism to explain how it evolves.

Hanaki (2006) used adaptive models to understand the dynamics that lead to efficient and fair outcomes in the repeated battle of sexes game. He develops a model that not only uses reinforcement learning but also the evolutionary learning that operates through evolutionary selection. He found that the efficient and fair outcome emerges relatively quickly through turn taking. However, his model requires a long run pre-experimental phase before it is ready to take turn. Turn taking in the battle of the sexes game is just one of many game theoretic phenomena, and it raises an important general point for further studies.

2. N-person Games and Classification of Strategic Structures

The fact that selfish behaviour may not achieve full efficiency has been well known in the literature (Young 1998). It is important to investigate the loss of collective welfare due to selfish and uncoordinated behaviour. Recent research efforts have focused on quantifying this loss for specific environments, and the
resulting degree of efficiency loss is known as the price of anarchy. Of particular interests is the issue how social interactions should be restructured so that agents are free to choose their own actions while avoiding outcomes that none would choose.

The problem of collective action arises also in the context of interaction within whole groups of individuals. For example, each person’s enjoyment of driving their car is inversely related to others’ enjoyment: if too many of us drive, everybody becomes stuck in congested traffic. The result is a kind of social congestion. A collective action solution to social congestion would involve individuals voluntarily restricting their own consumption of the limited resources, but in the absence of enforcement, each individual again has an incentive to free ride on the prudence of others.

There are many social interactions in which the underlying game has multiple equilibria. A very simple example is games involving contributions to the public community. Let consider the following example: On Sunday morning we have to choose either to participate in volunteer work to clean up a public park or to stay at home for extra sleep. The underlying game of this example is described as a coordination game that has three Nash equilibrium situations. In the first round, we all choose the strategy of contributing the community. In the second round, we all choose to stay at home and nobody cleans up the public park. In the third round, we each toss a coin to decide whether to contribute. The third alternative may seem dubious, but if everybody else is randomizing her choice, tossing a coin to decide what to do is as good as any other method of selection. In this example, we face an equilibrium selection problem, and the condition of efficiency takes us some way towards solving this equilibrium selection problem.

We consider a population of \( N \) agents, each faces a binary choice problem between two behavioural types: Cooperate (C) or Defect (D). For any agent the payoff to a choice of C or D depends on how many other agents also choose C or D. Here we consider social interactions in which agents are identically situated in the sense that every agent’s outcome, which ever way she makes her choice, depends on the number of agents who choose on way or the other. The payoff to each agent is given as an explicit function of the actions of all agents, and therefore she has an incentive to pay attention to the collective decision. The payoffs to each agent choosing from Cooperate (C) or Defect (D) are given:

\[
P(C) = \frac{1}{2} + \frac{p}{2}, \quad P(D) = \frac{1}{2} - \frac{p}{2},
\]

where \( p \) is the proportion of the agents who choose D (defect).

The binary decision itself can be considered a function solely of the relative number of other agents who are observed to choose one alternative over the others. The outcome depends on the strategy choices of all agents. Fortunately, in certain strategic situations, interactions among multiple agents are analyzed by decomposing into the underlying 2x2 games with the payoff matrix in Table 1. We
can classify social games with the payoff functions in (2.1) into the following types depending on the payoff values, \( a, b, c \) and \( d \).

<table>
<thead>
<tr>
<th>Own choice</th>
<th>Cooperate(C) ((p))</th>
<th>Defect (D) ((1-p))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate (C)</td>
<td>(a)</td>
<td>(b)</td>
</tr>
<tr>
<td>Defect (D)</td>
<td>(c)</td>
<td>(d)</td>
</tr>
</tbody>
</table>

**<Category I>**

1. **Prisoners’ dilemma (PD) game:** \((a > c, b > d, 2a > b + c)\).

   The \(N\)-person *Prisoners’ Dilemma (PD) game* is a multi-person decision-making involving the clash of individual and collective interests. The all-\(D\) choice \((p=0)\) indicates a Nash equilibrium, and the all-\(C\) choice \((p=1)\) indicates collective efficiency. If all agents seek their individual rationality, they result in defect and then each receives \(b\). On the other hand, if all agents cooperate, each agent receives \(a\) \((a > b)\). However if one agent defects, and all other agents cooperate then she receives \(c\) \((c > a)\). Therefore no agent will be motivated to deviate unilaterally from defect.

2. **Coordination game:** \((a > c, d > b)\)

   A *coordination game* is unlike a prisoners’ dilemma game, in so far as the payoff function \(U(D)\) in (2.1) does not dominate the function \(U(C)\) across the entire region of \(p\). In this case, we have a different case with two stable Nash equilibria at the end points in the left and right, and one unstable a Nash equilibrium at the intersection point. If only a few cooperate, they will switch to defect if they are rational, and if most agents cooperate, the few agents who defect will switch to cooperate. If everyone cooperates or if everyone defects then no one is motivated to switch. The direction in which collective behaviour will move depends on the initial proportion to choose cooperate or defect.

   In this case with multiple equilibria, the problem is to get a concerted choice. Since both the payoff curves have the same direction, there is no ambiguity about which equilibrium is superior one. The problem is then how to achieve the most efficient situation. If many agents defect, no agent is motivated to cooperate unless enough other agents do to switch beyond the intersection of the two payoff functions. Therefore the ratio at the intersection provides a crucial *mass parameter (threshold)* for the selection of collective efficiency. It is enough merely to get agents to make the right choice at the beginning.

3. **Hawk-Dove game:** \((a=(v-c)/2, b=v, c=0, d=v/2)\).

   A *Hawk-Dove game* is also unlike the NPD in so far as the function of \(D\) does not dominate the function of \(C\) across its entire region of \(p\). The Hawk-Dove game
has the unique symmetric Nash equilibrium in the mixed population, the proportion of agents to cooperate is \( \theta \) and that of agents to defect is \( 1-\theta \). The payoff at the Nash equilibrium is the same whether the individual agent cooperates or defects, and the expected payoff per agent is \( (v/2)/(1-(v/c)) \). However if all agents cooperate, each agent receives \( v/2 \), which is better than Nash equilibrium. Therefore collective maximum can occur at \( p=1 \) where all agents cooperate.

In summary, all games in the category 1 have the same property: both maximum efficiency and equity is achieved when all agents choose the same strategy C.

**Category II**

(4) Dispersion (or chicken) game: \( (c > a, b > d) \),

A more studied class of games is the coordination games in which agents gain high payoffs when they choose the same action. A complementary class that has received relatively little attention is the games in which agents gain payoffs only when they are disperse by choosing distinct actions. These games are sometimes called the dispersion games.

The payoff at the Nash equilibrium is the same whether agents choose C or D. The two payoff functions can be equated at the intersection and such is obtained as

\[
\theta \frac{v}{2} \left( 1 - \frac{v}{c} \right) = \frac{v}{2}.
\]

The dispersion game has the unique equilibrium in which both strategies are used at the ratios \( \theta \) and \( 1-\theta \). Under this mixed population, the payoffs of all agents are equal. If the game is repeated, there will be a tendency towards a stable equilibrium with \( \theta N \) agents choosing C and the rest \( (1-\theta)N \) agents choosing D.

In the context of the two-person symmetric game with the payoff matrix in Table 1, a pair of mixed strategy \( (p, p) \) constitutes a Nash equilibrium. A Nash equilibrium of the two-person game is defined in terms of how the payoff is affected when she switch to the other strategy when the other agent sticks to the current mixed strategy \( p=(p, 1-p) \).

However, such a mixed equilibrium situation is not to be socially efficient. However, this Nash equilibrium is not at collective maximum. Any agent who chooses C or D gains if some choosing D will shift and choose C. The collective maximum occurs to the left of the intersection. Since the slope of the payoff function C is shaper than that of the payoff function of D, if fewer agents than the ratio at the intersection choose C, agents who choose C will gain more than the loss of the agents who choose D. If the collective maximum does not occur at the intersection, there is a payoff difference between a choice of C and a choice of D. For instance, if the collective maximum occurs to the left of the intersection, agents who choose D gain less than those who choose C. This is a big difference from a Nash equilibrium at the intersection in which all agents receive the same payoff.
Collective efficiency (Pareto efficiency) is achieved at the strategy distribution where the average payoff per agent is maximized. We denote the average payoff per agent by $E(p)$ when the strategy distribution is $p=(p, 1-p)$. Since the proportion of agents to choose $C$ is $p$ and that of agents to choose $D$ is $1-p$, the average payoff per agent is

$$E(p) = pE(C) + (1-p)E(D).$$

(2.3)

Collective efficiency is achieved under the strategy population $p=(p^*, 1-p^*)$ where $p^*$ is given as

$$p^* = \frac{a-d}{b+c-a-d}.$$

(2.4)

There is often more than one Pareto outcome, not every Pareto outcome will be regarded as desirable. In general there are many Pareto efficient allocations, some of which are very bad from the point of view of equity, and there is no connection between Pareto efficiency and equity. In particular, a Pareto efficient outcome may be very inequitable. For example, consider a dictatorship run solely for the benefit of one person. This will, in general, be Pareto optimal because it will be impossible to raise the welfare of anyone except the dictator without reducing the welfare of the dictator. Nevertheless, most people (except the dictator) would not see this as a desirable outcome. There is a conflict between Nash equilibrium and efficiency in social interaction with negative externalities. There is also an efficiency-equity tradeoff.

(5) A variant of Prisoners’ Dilemma game: $(a > c, b > d, b + c > 2a)$.

There remains one definitional question as follows: how they do behave when an agent is better off if the more there are among the others who choose their inferior cooperative strategy $C$. Let consider the case the payoff parameters satisfy the following conditions: $c > a > d > b$, and $b + c > a + d$.

In this case collective efficiency is achieved at the mixed population of cooperators and defectors, since the average payoff per agent is also maximized at the point in (2.4). In this case the whole population gains higher payoff if they allow some defectors rather than all agents cooperate. When collective efficiency occurs only when all choose $C$, all agents receive the same payoff. However, in this mixed population case, some agents (defectors) gain more than the other agents (cooperators). The problem is then how to maximize collective efficiency at such an inequitable situation. It may become hard to devise a scheme to split agents into two groups, where agents in one group receive less than the other group.

We distinguish two types of externalities: strategic compatibility and strategic complementarity. If social interactions are characterized to have strategic compatibility, agents’ payoffs are increasing in the number of agents taking the same action. A typical example is the situation where the increased effort by some agents leads the remaining agents to follow suit, which gives multiplier effects. In
this case, each agent receives a high payoff if she selects the same action as the majority does. Instead, if social interactions are characterized to have strategic complementarity, things are better off if agents distribute themselves among the possible actions. But even if everyone prefers to be mixed, it often turns out that most agents become to take the same action. The problem of coordination failure arises in both contexts of social interactions with externalities. Social interactions under which the underlying games belong to Category 1 are classified as strategic compatibility and those in Category 2 are strategic compatibility (Namatame 2006).

3. A Strategy Choice with a Coupling Rule

The literature on learning in the game theory is mainly concerned with the understanding of learning procedures that if adopted by interacting agents will converge in the end to the Nash equilibrium of the underlying game. The main concern is to show that adaptive dynamics lead to a rational behaviour, as prescribed by a Nash equilibrium strategy. We call a dynamical system uncoupled if an agent’s learning model does not depend on the payoff functions of the other agents. Hart (2003) proved that there are no uncoupled dynamics that are guaranteed to converge to Nash equilibrium. Therefore, a coupling between agents, that is, the adjustment of an agent’s strategy depends on the payoff functions of the other agent, is a basic condition for convergence to Nash equilibrium.

The learning algorithms themselves are not required to satisfy any rationality requirement. Instead, they converge to a rational behaviour if it is adopted by all agents (Young 2005). In addition, Nash equilibrium cannot make precise predictions about the outcome of repeated games. Nor can it tell us much about the dynamics by which a collective of agents moves from an inefficient equilibrium to a better outcome. These limitations, along with concerns about the cognitive demands of forward-looking rationality, have motivated many researchers to explore alternatives backward-looking learning models. Most of these efforts have been invested in evolutionary dynamics. The research agenda of evolutionary dynamics is to explore non-equilibrium explanations of equilibrium in repeated games to view equilibrium as the long-run outcome of a dynamic learning process.

In this paper, we will take a different approach from the previous evolutionary dynamics by focusing on collective evolutionary dynamics. This approach differs from the common use of the genetic algorithm, in which the goal is to optimize a fixed fitness function. In the genetic algorithm, the focus is also on the best final result or on a good solution. In collective evolution, we are interested in better coupling among agents, which leads to desirable joint actions.

The first question we must address is what individuals know and what it is that they are learning about. In repeated games, agents repeatedly play an underlying game, each time observing their payoff and other agents’ strategies. In the classic work on learning in game theory, the agents select their strategy in the next iteration
of the game based on the result of the previous play using some updating rule. In the repeated model, agents engage in a series of games with different rules at each stage. In fact, the nature of each game depends on the results of the previous game, and this means the strategy choice depends on agents’ joint action in the previous rounds of games.

An important aspect of iterated games is the introduction of a coupling rule by which an agent can decide her strategy (Lindgren 1997). We will shift attention to coupled dynamics among agents with coupling rules. We make a distinction between adaptive or evolutionary dynamics and coupled dynamics. In an adaptive dynamics, other mechanisms are allowed as well, e.g., modifications of strategies based on the strategy distribution of the population. But, such adaptive dynamics do not necessarily improve the outcome to which the individual belongs in the long run. Evolutionary dynamics refer to the systems based on the basic mechanisms of biological evolution, that allows, inheritance, mutation, and selection. However, evolutionary dynamics based on natural selection also converge to an inefficient outcome. Coupled dynamics differ, in this sense, from evolutionary dynamics, in which a fixed goal is used in the fitness function and where there is no coupling among agents.

Let assume that each agent remembers the past outcomes (history). A coupling rule must specify, for each history, what strategy the agent should choose. We represent C by 0 and D by 1. In Table 2, we show all possible coupling rules with h=1. There are four possible outcomes for each move between two agents: (0, 0), (0, 1), (1, 0), and (1, 1). A quick calculation shows that the number of possible coupling rules with the outcome of only the previous round is $2^4$. With the increase of the memories of the past rounds, there are a huge number of coupling rules. The hope is that agents would find a better coupling rule out of the overwhelming number of possible rules after a reasonable number of repeated games.

Each coupling rule specifies the strategy choice based on the outcome of the previous round. Agent strategies are restricted to those employing only the previous move with the other agent to determine the next choice. Each agent has an evolvable rule as shown in Figure 1. Since no memory exists at the start, an extra one bit are needed to specify a hypothetical history.

<table>
<thead>
<tr>
<th>Strategy site in Figure 1</th>
<th>Previous strategies</th>
<th>Next strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Own</td>
<td>Opp</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2: A learnable coupling rule (# represents 0 or 1)
Table 3. All possible coupling rules with history 1

<table>
<thead>
<tr>
<th>Type</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL-C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Type 6</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Type 7</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Type 8</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Type 9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Type 10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TFT</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PAVLOV</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Friedman</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ALL-D</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 1 A coupling rule based on the previous round ($h=1$)

4. Repeated Games on Social Networks

It is important to consider with whom an agent interacts and how each agent decides her action depending on others’ actions. In order to describe interaction at the individual level, we may have two fundamental models, global interaction (or mean-field model) and local interaction. The introduction of spatial dimensions, so that individuals only interact with those in their neighborhood, may affect the dynamics of the system in various ways. The possibility of space-temporal structures may allow for global stability where the mean-free model would be unstable. The presence of these various forms of space-temporal phenomena may, therefore, also alter the evolutionary path compared with the mean-field case and we may see other strategies evolve.

The introduction of spatial interactions leads to the development of spatial games in which agents are located in the nodes of a fixed regular network of interaction, displaying rich spatio-temporal dynamics. The spatial dimensions are shown in Figure 2. Agents allocated to each cell of a 50x50 lattice play an underlying game against their nearest neighbours. In the lattice model each agent may choose the strategy of the most successful agent among her immediate neighbours. This may not be the most successful strategy in the position of this agent, since her neighbours may interact with different neighbours. The summed payoff of each game provides the agent’s fitness. After every individual has played the game with her neighbours, each rule of the agents is updated according to the general evolutionary rules based on the principle of natural selection. Each agent is replaced by an offspring of the highest scoring individual of the nearest neighbours. These offspring play the same
strategy as their ancestors, unless a mutation occurs, which happens at a small mutation rate. If a mutation occurs, the offspring's strategy is not its parent's strategy but a new strategy chosen randomly.

The main effect of the spatial structure in the repeated PD, for instance, is that cooperative strategies can build clusters in which the benefits of mutual cooperation can outweigh losses against defectors. Thus, clusters of cooperators can invade groups of defectors that prevail in non-spatial populations. The selection pressure of such an arrangement is clearly lower, since individuals are only assessed on a local level, not in a global fashion. This allows for individuals, which may have been eliminated if assessed against all players, to survive in a niche eventually be fit individuals or contribute genetic material to fit individuals as the environment changes. More importantly, from the viewpoint of a truly evolutionary system (one in which the individuals are evolving, as well as the population), the use of this spatial arrangement is thought to increase the genetic diversity by preserving apparently less fit strategies in niches.

Recent studies on the structure of social, technological, and biological networks have shown that they share salient features that situated them far from being completely regular or random. Social interactions are rarely well described by random or regular networks. Therefore, we also need to study the influence of the topological aspects of networks by exploring the different network topology. The topology of social networks is much better described by what has been called a small-world network (Watts 2003), as shown in Figure 2(b).

In a regular lattice model, agents interact with the nearest neighbours. In the version of a small-world network, a fraction of the neighbours is replaced by breaking interactions. An equal number of new agents are selected from outside of the current neighbours. These new agents for interaction are selected randomly from the rest of the population. Kirley (2000) studied an evolutionary version of the prisoner's dilemma game, played by agents placed in a small-world network. Agents are able to change their strategy, imitating that of the most successful neighbour. They found that collective behaviours corresponding to the small-world network enhances defection where cooperation is the norm in the fixed regular network.

Another important issue to consider is that networks are dynamic entities that evolve and adapt driven by the actions of agents that form a network. Zimmermann et al., (2004) studied the evolution of the social network. Initially, each agent plays a prisoner’s dilemma game with fixed neighbours. The network of interaction links evolves, adapting to the outcome of the game. They analyzed a simple setting of such an adaptive and evolving network, in which there is co-evolution of the state of the agents and the interaction links defining the network. The network of interaction evolves into a hierarchical network structure that governs the global dynamics of the system. However, the resulting network has the characteristics of a small-world network when a mechanism of local neighbour selection is introduced.
Various studies have examined the impact of different network structures on equilibrium selection in the context of iterated coordination games. If agents can choose the partners with which to interact, then they will form networks that lead to efficient Nash equilibrium play in the underlying coordination game. Ellison (1993) analyzed the role of local interactions for the spread of particular strategies in coordination games, showing how play converges to risk-dominant equilibrium if agents are located on a circle and interact with their two nearest neighbours. Similarly, Blume (1993) and Kosfeld (2002) proved the convergence to the risk-dominant equilibrium in a population of agents located on a two-dimensional lattice.

Goyal et al., (2005) studied the formation of networks among agents who are bilaterally involved in coordination games. In addition to specifying which pairs of agents in the population play the game, the network structure also determines how strategic information diffuses among the agents and how coordination among the agents is found. They showed that once agents are allowed to choose their partners, the situation is very different. They introduce a number of locations where agents can meet and play the coordination game with each other. Thus, at any time, agents choose both a location and a strategy in the game. Under these conditions they showed that risk dominance loses its selection force and that the population is most likely to coordinate on the Pareto efficient equilibrium. Since agents can freely choose their interaction partners, they are able to find partners that choose the Pareto efficient equilibrium strategy, and at the same time they can avoid agents that choose the risk-dominant inefficient strategy.

We will compare and identify the effect of the manner of interaction by considering the lattice network, small-world network, and random network as shown in Figure 2. In small-world network, for instance, the locally networked agents will be compared to a half-mixed population in which a half of the population is again modeled by a lattice, but in this case each agent interacts with four partners that are nearest neighbors and four partners that are randomly chosen from the population.

Figure 2  Repeated games on a grid. (a) Local model: Each agent interacts with her nearest neighbors.(b) Games on a small-world network. Each agent interacts with her four nearest neighbors and four other agents randomly chosen from the population.
5. Simulation Results on Dispersion Games

(1) Symmetric Dispersion Game

First of all, we consider the case in which all agents repeatedly play the symmetric dispersion game with the payoff matrix in Table 4. This game has two equilibria with the pairs of the pure strategies (C, D), (D, C) and one equilibrium of the mixed strategy.

The average payoff per agent at each generation is shown Figure 3. Figure 3 shows the same experiment, this time using different network frameworks. There is little difference in the graphs using local and small-world networks and the graph obtained using the non-spatial random environment.

The advantage of agent-based modeling is that we can investigate coupling rules learned by all agents that lead to a desirable collective outcome at the macro level.

Table 4. Payoff matrix of the symmetric dispersion game

<table>
<thead>
<tr>
<th>Own strategy</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3. The average payoff per agent over generation
In Table 5, we show the coupling rules learned by 2,500 agents. In the beginning, they are endowed with randomly chosen coupling rules. However, these different rules were updated through collective evolution. The 2,500 rules were finally aggregated into a few types, as shown in Table 5. The numbers in the right-hand column represent the number of agents who share the same rule. Those aggregated rules have the commonality. We show the strategy choices between two agents with the rules in Table 5 as the state transition process as shown in Figure 4. If agents choose C and their opponent chooses D at the previous time period, then they choose C. If agents choose D and their opponent chooses C at the previous time period, then they choose D. These rules represent the following action: if they gain then they repeat the same winning strategy, and this is the same principle of reinforcement learning.

Table 5. Coupling rules learned by 2,500 agents in the symmetric dispersion game: lattice or small-world network

<table>
<thead>
<tr>
<th>Rule type</th>
<th>Initial strategy</th>
<th>Strategy site</th>
<th>Number of agents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>0 or 1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0 or 1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0 or 1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0 or 1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 4. State transition of evolved coupling rules in Table 5: Case1: Two agents with the same coupling rule, Case2: Two agents with different coupling rules
(2) Asymmetric Dispersion Game

Next, we investigate the case in which the underlying game is the asymmetric dispersion game with the payoff matrix in Table 5. Figure 5 shows the average payoff per agent and it was gradually increased to 1.6. Figure 5 shows the same experiment, this time using different network frameworks. There are a couple of important differences in this graph and the graph obtained using the non-spatial random environment. Under the random matching model, the average payoff per agent is about 0.6.

<table>
<thead>
<tr>
<th>Strategy of the other agent</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

In Table 7, we show the coupling rules learned by 2,500 agents. In the beginning, they are endowed with randomly chosen coupling rules. These different rules were updated through collective evolution and they were finally aggregated into a few types as shown in Table 7. The numbers in the right-hand column represent the number of agents who share the same rule. Those aggregated rules have the commonality. We show the strategy choices between two agents with the rules in Table 7 as the state transition process as shown in Figure 6.
Table 7 Coupling rules learned by 2,500 agents: the asymmetric dispersion game and variant PD game under local model

<table>
<thead>
<tr>
<th>Rule type</th>
<th>Initial strategy</th>
<th>Number of agents</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>0</td>
<td>816</td>
</tr>
<tr>
<td>10</td>
<td>0 or 1</td>
<td>919</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>430</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>335</td>
</tr>
</tbody>
</table>

There is a significant difference depending on two agents learn to have the same rule or different rules. If two agents with the same rule interact, they absorb in to (0,0) or (1,1) as shown in Figure 6 (case1). On the other hand, if they have different rules, the turn-taking behaviour is emerged as shown in Figure 6 (case2). If one agent chooses C and their opponent chooses D at the previous time period, then they choose C. If agents choose D and their opponent chooses C at the previous time period, then they choose D. These rules represent the following action: if they gain then they repeat the same winning strategy, and this is the same principle of reinforcement learning.

Most agents succeeded in learning the coupling rules that have the following property. If the agent gains the payoff (success), then they change their strategies. We define this behavioural rule based on the principle of give-and-take. When agents face symmetric dispersion games with the payoff matrix in Table 5, there is no difference in the payoff under outcome (C, D) or (D, C). Therefore, they learn the coupling rules to continue the same strategy if they gain. On the other hand, when they face asymmetric dispersion games with the payoff matrix in Table 6, the
payoffs to both agents at the two pure Nash equilibria (C, D) and (D, C) become asymmetric. Therefore, they learn to realize efficient and equitable outcomes by visiting the two pure Nash equilibria alternatively.

Browning et al., (2004) investigated how this type of coordinated, alternating cooperation can evolve without any communication between agents who play the dispersion game. Using a genetic algorithm incorporating mutation and crossing-over, they showed that coordinated turn-taking can evolves in battle of sexes and chicken games with asymmetric Nash equilibria. The procedure followed Wu et al., (1997). For each outcome of the game, each agent receives one of four payoffs, and she remembers three past outcomes. Since there are 4³ different three-move histories, each string of 64 binary digits suffices to specify a choice for every three-move history.

The offspring rules that played in each subsequent generation were formed from the most successful rule of the previous generation, using a genetic algorithm. The algorithm implemented the following five steps: (1) The payoff values were assigned according the underlying game. (2) An initial population was for each of the 20 randomly chosen rules. (3) In each generation, each of the 20 rules was paired with each of the others for the fixed number of repetitions with every other rule in the population (global interaction). (4) At the end of each generation, after each rule had played with each of the others, each rule’s mean payoff was computed, and it was as signed a mating probability proportional to its fitness score. (5) For each offspring strategy, two rules were randomly selected as parents, selection being proportional to mating probability scores.

They showed that about 85% of the plays in the population are characterized by coordinated turn taking. They study the nature, properties and phenomena of coordinated alternating cooperation in a range of dispersion games with asymmetric equilibria. By alternating coordination the agents benefit from it, however, how agents evolves alternating coordination without communication is not fully explained.

(3) A variant of dilemma game

We set the payoff parameters as \( a=1, b=-1, c=4, d=0 \) as shown in Table 8. The average payoff will be 1 in a completely cooperating population. However a mixed population with cooperators and defectors is better than a completely cooperating population, since the average payoff will be greater than 1 in a mixed population. In Figure 7, after 100 generations the average payoff is close to 1.2. Figure 7 shows the same experiment, this time using different network frameworks. There are a couple of important differences in this graph and the graph obtained using the non-spatial environment. The coupling rules learned by 2,500 agents are almost the same as the asymmetric dispersion game as shown in Table 7.
If one of the two agents chooses C(0) and her opponent chooses D(1) (in this case both agents gain the payoff), she changes her choice to D(1). On the other hand, if she chooses D(1) and her opponent chooses C(0) (in this case both agents also gain the payoff), then she also changes to C(0). In all other cases, that is, if both agents choose the same choices (in which case they do not gain the payoff), the likelihood of choosing the same strategy as in the previous round is equal to that of choosing the other strategy.

Table 8 Payoff matrix of the variant Prisoner’s dilemma game

<table>
<thead>
<tr>
<th>Strategy of the other agent</th>
<th>Own strategy</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Average payoff

Figure 7 The average payoff per agent over generations

6. Effects of Network Topology

In this section we discuss the impact of different network structures on evolution of social norms in the context of iterated games. If agents can choose the partners with which to interact, then they will form networks that lead to efficient Nash equilibrium play in the underlying games.

In the local model, each agent interact his/her nearest eights agents (Figure 2(a)). In the small-world network (Figure 2(b)), the locally networked agents will be compared to a half-mixed population in which a half of the population is again
modeled by a lattice, but in this case each agent interacts with four partners that are nearest neighbors and four partners that are randomly chosen from the population. In the random matching model (Figure 2(c)), each agent interact with eight other agents chosen randomly. We compare and identify the effect of the manner of interaction by considering evolved coupling rules.

In Table 9, we show the coupling rules learned by 2,500 agents who repeatedly play the asymmetric dispersion games in Table 8 with randomly chosen eight neighbors (random matching). In the beginning, they are endowed with randomly chosen coupling rules. However, these different rules were updated through collective evolution. 2,500 rules were finally aggregated into four types, as shown in Table 8. The numbers in the right-hand column represent the number of agents who share the same rule. Those aggregated rules are different from those rules under local model and small-world network as shown in Table 7.

In Table 9, we show the coupling rules learned by 2,500 agents who repeatedly play the variant dilemma game in Table 8 with randomly chosen eight neighbors (random matching). In the beginning, they are endowed with randomly chosen coupling rules. However, these different rules were updated through collective

Table 9. Coupling rules learned by 2,500 agents: the asymmetric dispersion game under the small world network and random models.

<table>
<thead>
<tr>
<th>Rule type</th>
<th>Initial strategy</th>
<th>Strategy site</th>
<th>Number of agents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 10. Coupling rules learned by 2,500 agents: the variant PD game under random model

<table>
<thead>
<tr>
<th>Rule type</th>
<th>Initial strategy</th>
<th>Strategy site</th>
<th>Number of agents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
evolution. The 2,500 rules were finally aggregated into the two types, as shown in Table 10. The numbers in the right-hand column represent the number of agents who share the same rule. Those aggregated rules are different from those rules under local model and small-world network shown in Table 7.

Although social norms have been conceptualized in various ways by a variety of researchers, definition in a recent review article is as follows: Social norms are rules and standards that are understood by members of a group, and that guide and/or constrain social behaviour without the force of laws. These norms emerge out of interaction with others; they may or may not be stated explicitly, and any sanctions for deviating from them come from social networks, not the legal system. Put differently, norms are shared rules that emerge and are sustained through people's autonomous interaction without formal regulating authorities or forces such as laws.

This definition highlights emergence and sustainability of social norms as core issues for any theory of norms. That is, to elaborate the norm construct fully, we need to understand how social norms can emerge voluntarily through people's interaction without external regulating forces. This theoretical perspective is shared by other social scientists, yet we do not have a reasonable theory about norm development. Such a focus on arbitrary norms may have inadvertently led us to assume that social or cultural learning per se is a sufficient mechanism for norm development. Although social or cultural learning is vitally important for norm development, a more fundamental question may be why some beliefs are acquired socially and are maintained as a shared rule, while other beliefs are not.

A theory of norm development must address this question explicitly, explaining why particular behavioural or cognitive patterns proliferate in a society and are maintained as shared rules and standards. Given the fundamental fact that we are group-living species, the thesis that our minds are socially adaptive seems a reasonable meta-theoretical assumption. Social norms that link micro-level cognitions of individuals to a macro-level social condition, in a mutually constrained manner as we have demonstrated using the evolutionary game analysis capture an essential characteristic of such group life. In this sense, we believe that the notion of social norm can serve as one of the central and most useful constructs in social sciences in an integrated way.

7. Selective Interactions with Agents’ Movements

In many social interactions, agents consider not only which actions to choose, but also with whom they should interact. Similarly, in some social contexts, dissatisfied agents seek to break up some partnerships or alliances and to form new ones. This ability to rematch has strong implications for behaviour within social relationships. While this observation is a relatively obvious, we have no systematic method of modeling such choice behaviour depending on an agent’s ability to select partners in the framework of game theory. In this section, we introduce such a
methodology and examine a new class of social games in which agents also decide
with whom they will play the game.

A different kind of collective behaviour arises when agents change those with
whom they interact before they make up their mind how to behave. The work by
Schelling (1978) triggered a lot of interest with respect to the question of the
selection of neighbours. The aims of their models are to explain how social
integration or segregation may occur spontaneously, even if people do not intend for
them to occur. Individuals interact locally, having preference over their
neighbourhood. Taking the colour of an individual (for instance white or black) as
the criteria for discrimination, the problem faced by each individual is to choose a
location given an individual threshold of acceptance for the proportion of
individuals of different colour in their neighbourhood. They showed that a different
kind of social norm arises when people change those with whom they associate,
instead of changing how they behave given their associates. In other words, they
choose their neighbours instead of conforming to their neighbours.

This is called a sorting process. In their models, agents behave based on the
following simple rule. An agent agrees to stay in a neighbourhood with agents that
are mainly of the same colour. More specifically, the following behavioural rule is
used. An agent with one or two neighbours will try to move if there is not at least
one neighbour of the same colour. Under the assumption of a local behavioural rule
for each agent, a fully integrated structure is observed at equilibrium, where no
agent wants to move. However, they show that a slight perturbation is sufficient to
induce a chain reaction and the emergence of aggregate behaviour of segregation.
The agents move at random towards a new location in agreement with their own
preferences. The mobility of agents generates new discontented agents through a
chain reaction until a new equilibrium is reached, and finally spatial segregation
between two groups of agents with different colours often emerges. Thus, they show
that selective interactions are sufficient for the occurrence of complete segregation,
while it is not an attribute of the individual agents.

Selective interaction in dilemma games was introduced in previous studies
(Tesfatsion 1996, Axelrod 1997). When agents interact with other agents, they begin
to develop a history of play. They keep track of how many times the other agent
defects. If the other agent defects more than a certain number of times in previous
interactions, then the agent will avoid interaction with that agent again. Another
crucial effect of selective interaction is that it allows agents to group together. An
agent can avoid interaction with other agents if she receives a payoff that is lower
than some threshold and moves to another site in order to have a chance to interact
with different agents. Because of the gain from cooperation, cooperators that are
surrounded by other cooperators can earn higher payoffs than defectors who are
primarily surrounded by other defectors. Thus, endowing agents with the capability
of selective interaction substantially increases the chances that cooperative agents
will survive and that cooperative behaviour will evolve.
Once agents are allowed to choose which partners to interact with, the situation is very different. They introduced a number of locations where agents can meet and play the coordination game with each other. Thus, at any time, agents choose both a location and a strategy. With the combination of the partner selection and the strategy choice, they showed that risk dominance loses its selection force and that the population of agents is most likely to coordinate to realize the Pareto-efficient equilibrium. The reason for this is intuitive. Since agents can freely choose their interaction partners, they are able to select neighbours who engage in the Pareto-efficient equilibrium strategy in order to gain higher payoff, and at the same time, they can avoid agents who choose the inefficient risk-dominant strategy to obtain a lower payoff.

We assume that agents are assumed to have the ability to move and interact selectively with other agents while making interaction mandatory for other agents. In this section, we study a model in which unsatisfied agents with lower average payoff than a threshold move to new sites and interact with new agents. An agent may need to select her neighbours to interact with while considering a tradeoff between joining a neighbourhood in which most agents share the same behavioural rule or another neighbourhood in which they have different rules. Agents also move because they prefer the neighbourhood they are moving into compared with the neighbours they are moving away from.

In our model, a collective of agents repeatedly plays the underlying 2x2 game, formulated as a dispersion game. The payoff to each agent represents the fitness in this case. We shall see that the combined model of the partner choice and preference evolution requires special analysis techniques. In our model, agents repeatedly play the underlying game with the current neighbours and myopically adapt their strategies with regard to neighbours in order to maximize their payoffs. After a number of repetitions of the game, they evaluate their performance in terms of the average payoff (fitness), and the successful agents increase the payoff parameter associated with the current strategy, and decide to remain in the same game. On the other hand, dissatisfied or unsuccessful agents move to new games in order to change the partners and interact with new neighbours.

More specifically, an agent decides to stop interaction with her current neighbours if she receives a payoff that is below some threshold, and moves to another game in order to interact with other neighbours. On the other hand, if her gain by choosing some specific strategy exceeds a certain threshold, then she continues to interact with the same neighbours and the preference (associated payoff parameter) of that strategy increases. Preference evolution therefore leads us to consider dynamics that run at two different speeds at once.

Figure 8 shows the average payoff per agent under the symmetric dispersion game in Table 6, the asymmetric dispersion game in Table 7, and under variant PD game in Table 8. Until 400 generations, all agents co-evolve their coupling rules as discussed in the previous sections, and after that, unsatisfied agents with lower
payoff than the threshold move to new sight. Then the average payoff after movement was suddenly increased to Pareto-optimal of each underlying game.

Average payoff

![Figure 8. Average payoff under agents' movements](image)

Although the individual decision problem is important to understand, it is not sufficient to describe how a collection of agents arrives at specific desirable collective outcomes. Therefore, we aim to discover the fundamental micro-mechanisms that are sufficient to generate the desirable macroscopic structures of interest. This type of self-organization is referred as the emergence of desired orders from the bottom up. The first priority for a desirable collective outcome is stability, which is crudely modeled using the idea of equilibrium of an underlying game. The next priority is efficiency, which is also defined as following Pareto optimality and is equivalent to the requirement that nobody can be made better off without someone else being made worse off. The third priority is equity.

The question of whether interacting agents self-organize desirable macroscopic behavior from bottom up depends on the type of social interaction as well as heterogeneity in agents. While agents may understand an outcome to be inefficient, by acting independently, they are powerless to manage the collective to overcome this inefficiency.

8. Conclusion

Social norms are self-enforcing patterns of social behaviour. It is in everyone’s interests to conform given the expectation that others are going to conform. Many spheres of social interactions are governed by social norms. Computer simulations have shown that, after thousands of repetitions of social interactions, social norms such as such as reciprocity, give and take, and so forth. Norms are sets of socially
agreed upon ‘rules’ that we draw upon to structure our behaviour. A large literature testifies to the many ways in which norms shape behaviour and enable us to cooperatively interact with others.

In an evolutionary explanation of norm development, there is no need to assume a rational calculation to identify the effective rule. Instead, the analysis of what is chosen at any specific time is based upon an implementation of the idea that effective rules are more likely to be retained than ineffective ones. Furthermore each agent mimics the most successful neighbour as guidance of improving her coupling rule. Their success depends in large part on how well they learn from their neighbours. If an agent gains more payoff than her neighbour, there is a chance her coupling rule will be imitated by others. The more successful agents are more likely to survive and reproduce effective coupling rules. However, agents also observe each other, and those agents with poor performance tend to imitate the rules of those they see doing better. This mechanism of collective evolution tends to evolve to both efficient and equitable outcomes. Furthermore, the asymmetry in payoffs from interaction induces agents to learn the behavioural rule, so-called turn-taking norm to break the asymmetry in social interactions.

References


Young, H. P. *Strategic Learning and Its Limits*, Oxford Univ. Press, 2005.

