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 desirable norms will emerge in various social interaction settings. Human beings appear to easily recognize the importance of a coordinated turn-taking behaviour as a mean to realize a fair outcome. We show turn-taking norm or alternating reciprocity emerge among networked agents who play dispersion games. We also investigate the co-evolutionary dynamics of networked agents in different network topologies and discuss the effects of the network topologies on evolution of such desirable norm. We show the symmetric local network under which each agent play with the same number of the closest neighbours fosters to emerge such desirable turn-taking norm.

1. Introduction

The undesirable outcomes that none would have chosen may occur when collective actions of interacting agents lead to a result which is not *optimal*. This problem is often referred as a *coordination failure*. The reason why uncoordinated activities of agents pursuing their own interests often produce outcomes that all would seek to avoid is that each agent's action affect 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*. Social interaction with externalities raises two basic questions: The first is how the aggregate outcome actually comes to exist, and the second one concerns what the ideal outcome should look like. From the perspective of a planner, social interactions with externalities will result in an outcome that is not socially optimal.

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. Many sphere of social interactions are governed by *norms* such as *reciprocity* and *equity*. Norms and conventions are the glue that holds a society of self-interested agents together. When outcomes appear to be coordinated, a kind of norms can potentially serve useful constructs to coordinate human behaviour. 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 that characterize social interactions of interest. It is a strategy that assigns a rule to each individual that is an optimal in the sense no one has an incentive to deviate from it.

However, there is little theory of norm development. The introduction of genetic algorithms enabled researchers to investigate the natural selection of norms using repeated games. The challenge for evolutionary approach is to identify mechanisms of emergent properties for a rich class of social interactions and to allow agents for more intelligent choices in competitive situations.

Evolutionary game theory offers one approach to the study of emergent properties in systems of interacting agents. Darwinian dynamics which are 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 a population 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 so-called arms race with no end. Evolving agents might continually adapt to each other in more and more specialized ways, never stabilizing at a desirable outcome. Of particular interest is the question how social interactions can be managed so that agents are free to choose their actions while avoiding outcomes that none would have desired.

Previous research on norm development has mainly focused on social interactions described as a 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. Wu and Axelrod (1997) have shown by computer simulations that reciprocal strategies such as Tit for Tat (TFT) tend to evolve, resulting in widespread joint cooperation. The evolutionary model also explains how TFT could have evolved as norm, given that natural selection operates at the individual levels.

Ellison (1993) studied the evolution of play in a population of agents continually interacting with coordination games. Agents are repeatedly and randomly matched against opponents, and they revise their actions at discrete random moments. Individual rationality alone would suggest that some kind of coordination would occur. But this model limits 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 equilibrium selection is universal. Hanaki *et al.*, (2005), for instance, have shown that in at least some models where matching is endogenous, payoff-dominant equilibrium 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 agent 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. If this type of social interaction is repeated, then the agents benefit in terms of accumulated payoffs by *coordinated alternation* by taking turns in choosing one of the two actions. There is also an evidence that this type of *turn-taking behaviour* occurs quite commonly in nature. Give and take or alternation is norm that is intuitive and simple, but even so it is beyond the scope of most traditional learning or evolutionary models to explain how such norm emerges.

Browning and Colman (2004) investigated through agent-based simulation how coordinated turn-taking or alternating cooperation can evolve without any communication between agents. Using a genetic algorithm incorporating mutation and crossing-over, they showed that coordinated turn-taking evolves in battle of sexes games. For each outcome of the game, each agent receives one of four payoffs, and she remembers three past outcomes. Since there are 4^3 different outcomes with 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 payoffs 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. By alternating coordination the agents benefit from it, however, how agents evolves alternating coordination without communication is not fully explained by their works.

Hanaki (2006) used adaptive models to understand the dynamics that lead to efficient and fair outcomes in the repeated battle of sexes game. He developed 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 also 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. In this paper we study how norm of coordinated turn-taking in a range of games with asymmetric situations may emerge and attempt to find a mechanism to explain how agents evolve such norm.

2. Social Games and Classification of Social Interactions

The problem of collective action arises in various contexts of social interactions. For example, each person’s enjoyment of driving their car is inversely related to others’ enjoyment. If too many of them drive, everybody becomes stuck in congested traffic, and 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.

The fact that selfish behaviour may not achieve full efficiency at the aggregate level 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. Of particular interests is the issue how social interactions should be managed so that agents are free to choose their own actions while avoiding outcomes that none would choose.

We consider a population of N agents, each faces a binary choice problem between two behavioural types: C(Cooperate) or D(Defect). 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, whichever way she makes her choice, depends on the number of agents who choose on way or the other. In this case 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 binary decision itself can be considered a function solely of the *relative* number of other agents who are observed to choose one action (or the proportion) over the other. The payoffs to each agent choosing from C or D are given:

$$U(C) = (a-b)p + b, \qquad (2.1)$$
$$U(D) = (c-d)p + d$$

where p $(0 \le p \le 1)$ is the proportion of the agents to choose C. Fortunately, in this situation, social interactions among multiple agents are analyzed by decomposing into the underlying 2x2 games with the payoff matrix in Table 1.

Table 1. The payoff matrix of a social game

All others' choices Own choice	C *(p)*	D (*1-p*)
C	*a*	*b*
D	*c*	*d*

We can classify social games with the payoff functions in (2.1) into the following types depending on the payoffs, *a, b, c* and *d*.

<Category I>

(1) Prisoners' dilemma (PD) game: ($c > a, d > b, 2a > b + c$),

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 choosing D and then each receives *d*. On the other hand, if all agents choose C, each agent receives *a*. However if one agent chooses D, and all other agents choose C then she receives *c(c>a)*. Therefore no agent will be motivated to deviate unilaterally from choosing D.

(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 points at p=0 and p=1. If only a few choose C, they will switch to chooses D if they are rational, and if most agents choose C, the few agents who choose D will switch to choose C. If everyone chooses C or if everyone chooses D, then no one is motivated to switch.

In this case with multiple equilibria, the problem is to get a concerted choice. Since both the payoff curves in (2.1) 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 (p=1). If many agents choose D, no agent is motivated to choose C unless enough other agents do to switch beyond the intersection of the two payoff functions. Therefore the ratio at the intersection provides a *critical mass* (*threshold*) for the selection of collective efficiency. The direction in which collective behaviour will move depends on the initial proportion to choose C or D. Therefore it is necessary 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 the entire region of *p*. The Hawk-Dove game has the unique symmetric Nash equilibrium in the mixed population, the proportion of agents to choose C is $\theta = v/c$ and that of agents to choose D is $1-\theta = 1-v/c$. The payoff at the Nash equilibrium is the same whether the individual agent chooses C or D, and the expected payoff per agent is *(v/2){1-(v/c)}*. However if all agents choose D, each agent receives *v/2*, which is better than Nash equilibrium. Therefore collective maximum can occur at *p=1* where all agents choose D.

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 action.

<Category II>

(4) Dispersion 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. We define these games as the *dispersion games*.

The two payoff functions in (2.1) can be equated at the intersection and such p is obtained as

$$p = (b-d)/(b+c-a-d) \equiv \theta\,. \tag{2.2}$$

The dispersion game has the unique equilibrium in which both actions are used at the ratios θ 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 θN agents choosing C and the rest $(1-\theta)N$ agents choosing D.

However, such a mixed Nash equilibrium is not to be efficient. 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 sharper 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.

Collective efficiency (Pareto efficiency) is achieved at the 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)*. If 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

$$\begin{aligned} E(p) &= pU(e_1, p) + (1-p)U(e_2, p) \\ &= (a+d-b-c)p^2 + (b+c-2d)p + d \ . \end{aligned} \tag{2.3}$$

Collective efficiency is achieved at ***p***=*(p*, 1-p*)* where *p** is given as

$$p^* = (b+c-a-d)/2(b+c-a-d) \tag{2.4}$$

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.

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

There remains one 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 action *C*. Let consider the case where the payoffs satisfy the following two conditions: $c > a > d > b$, and $b + c > a + d$. In this case collective efficiency is achieved at the mixed population of agents choosing C and D, since the average payoff per agent is also maximized at the point in (2.4). In this case the whole population gains a higher payoff if they allow some agents choosing D (defect) rather than all agents choosing C (cooperate). When collective efficiency occurs only when all choose C, all agents receive the same payoff. However, in this mixed population case, some agents (defect) gain more than the other agents (cooperate). The problem is then how to manage the situation where collective efficiency is achieved at an inequitable outcome. 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.

3. Social Interaction 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. We call a dynamical system *uncoupled* if an agent's learning model does not depend on the payoff functions of the other agents. Hart and Mas-Colell (2003) proved that there are no uncoupled dynamics that are guaranteed to converge to Nash equilibrium. Therefore, a proper coupling between agents 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. 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 outcome of interacting 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 (Young 2005). The research 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 modelings by focusing on collective evolution. Our approach also 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 finding better coupling rules among agents which leads to desirable joint actions.

The first question we must address is what agents know and what it is that they are learning about. In repeated games, agents repeatedly play an underlying game, each time observing their own payoff and the other's action. In the classic work on learning in game theory, the agents select their action in the next iteration of the game based on the result of the previous round using some adaptive mechanism. An important aspect of iterated games is the introduction of a rule by which an agent chooses her action (Lindgren 1997).

Table 2. An evolvable coupling rule (# represents 0 or 1)

Action site in Figure 1	Previous actions		Next action
	Own	Opponent	
4	0	0	#
5	0	1	#
6	1	0	#
7	1	1	#

Table 3. All possible behavioural rule s with history 1

Type 1: 0000	(ALL-C)	Type 9: 0001	
Type 2: 1000		Type10: 1001	
Type 3: 0100		Type11: 0101	(TFT)
Type 4: 1100		Type12: 1101	
Type 5: 0010		Type13: 0011	
Type 6: 1010		Type14: 1011	
Type 7: 0110	(PAVLOV)	Type15: 0111	(FRIEDMAN)
Type 8: 1110		Type16: 1111	(ALL-D)

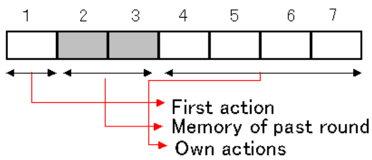


Figure 1 A genetic representation of a coupling rule based on the previous outcome

We will shift attention to coupled dynamics where agents play repeated games with some coupling rules. We make a distinction from adaptive dynamics and evolutionary dynamics. In an adaptive dynamics, other mechanisms are allowed as well, e.g., modifications of actions based on the distribution of the actions in the population. But, such adaptive dynamics do not necessarily improve the outcome to which the agents belong in the long run. Evolutionary dynamics refer to the systems based on the basic evolutionary mechanisms that allow inheritance, mutation, and selection. However, evolutionary dynamics may also converge to an inefficient outcome. On the other hand, coupled dynamics differ 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 outcome (history 1: h=1). A coupling rule must specify what action the agent should choose depending the previous outcomes. 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 action between two agents, which are represented as (0, 0), (0, 1), (1, 0), and (1, 1). Each coupling rule specifies the action the agent chooses based on each outcome of the previous game. Each string of 4 binary digits needs to specify a choice for 4 possible outcomes. Since no memory exists at the start, an extra bit is needed to specify a hypothetical action at the beginning. The coupling rule in a genetic code form in Figure 1 is evolvable by changing the bits from 4 to 7. A quick calculation shows that the number of possible coupling rules is 2^4 as shown in Table 3. With the increase of the histories of the past games, there are a huge number of rules. The hope is that agents would find a better coupling rule out of a large number of coupling rules after a reasonable number of repeated games.

4. Repeated Games on Social Networks

It is important to consider with whom an agent interacts and how she evolves her action depending on others' actions. In order to describe the interaction topology, we have two fundamental models, global interaction (or mean-field model) and local interaction. 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 dynamics. The lattice dimensions are shown in Figure 2(a). Agents allocated to each cell of a *50x50* lattice play an underlying game against their nearest neighbours. 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 more effective interaction rules evolve.

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

Recent studies on the structure of social networks have shown that they share salient features that situated them far from being completely regular or random. 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 1993), as shown in Figure 2(b). In a regular lattice model (local interaction), 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 (2004) 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. He found that collective behaviours corresponding to the small-world network enhances D choosers where cooperation (C) is the norm in the fixed regular network.

Various studies have also 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.

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 and Eguiluz (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 actions of the agents and the interaction links defining the network.

Goyal and Vega-Rendondo (2005) also 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 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 an action in the game. Under these conditions they showed that risk dominant equilibrium looses 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.

In this paper, we compare and identify the effects of the manner of interaction by considering the lattice network, small-world network, and random network as shown

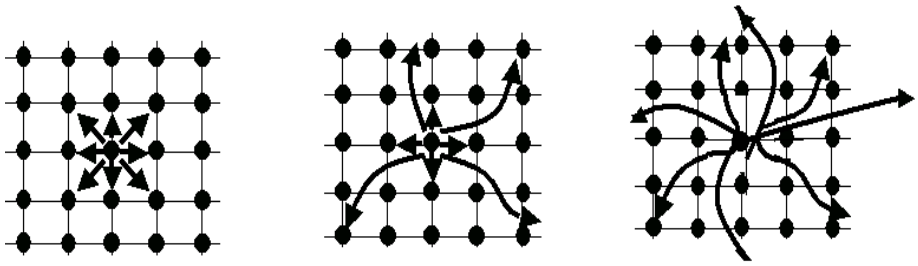


Figure 2 Repeated games on social networks. (a) Games on a lattice network: each agent interacts with her nearest neighbors. (b) Games on a small-world network: each agent interacts with her four nearest neighbors and four randomly chosen agents. (c) Games on a random network: each agent interacts with eight randomly chosen agents from the population.

in Figure 2. In a lattice network in Figure 2(a), the locally networked agents play the game with her nearest 8 neighbours. In a small-world network in Figure 2(b), the locally networked agents will be compared to a half-mixed population in which a half of the population is again modelled by a lattice, but in this case each agent interacts with four agents that are nearest neighbours and four randomly chosen agents from the population. Games on a random network in Figure 2(c), each agent interacts with eight randomly chosen agents from the population.

5. Simulation Results on Dispersion Games

(1) Symmetric Dispersion Game

At first, we consider the case in which all agents repeatedly play the symmetric dispersion game in Table 4. This game has two equilibria with the pairs of the pure strategies (C, D), (D, C). At either equilibrium, both agents receive the same payoff of 1. There is one more equilibrium of the mixed strategy, and both agents receive the payoff of 0.5.

The average payoff per agent at each generation is shown Figure 3. Figure 3 shows the same experiments using different network frameworks. Each agent becomes to receive 0.8, and we can observe that an efficient collective outcome is realized at the macro level. There is little difference in the graphs using local and small-world networks and the graph obtained using the non-spatial random network.

The advantage of agent-based modelling is that we can investigate coupling rules learned by all agents. In Table 5, we show all coupling rules evolved by 2,500 agents. In the beginning, they are endowed with randomly chosen coupling rules.

However, these different rules were updated through evolution, and these were finally aggregated into a few types shown in Table 5. The numbers in the right hand column represent the number of agents who have the same rule. Those aggregated rules also have the commonality in the strategy site #5 and #7 specified in Table 2. The action choices between two agents with the rules in Table 5 can be shown as the state transition process in Figure 4. If agents choose C(0) and their opponent chooses D(1) at the previous round, then they choose C(0). If agents choose D(1) and their opponent chooses C(0) at the previous round, then they choose D(1). These rules represent the following behavioural rule: if they gain then they repeat the same winning action. This is the same as the principle of reinforcement learning.

Table 4. Payoff matrix of the symmetric dispersion game

Strategy of the other agent / Own strategy	*C*	*D*
C	0 0	1 1
D	1 1	0 0

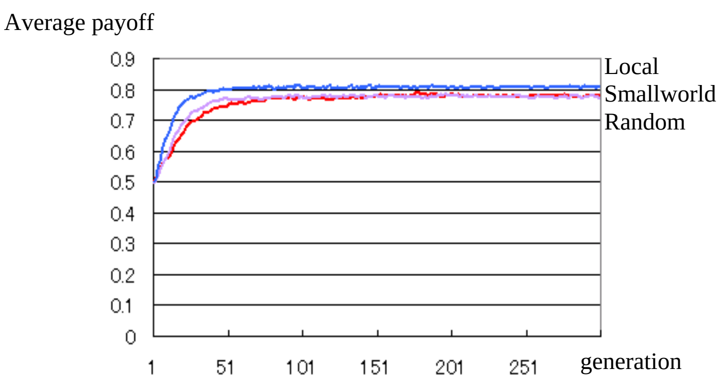


Figure 3. The average payoff per agent over generation

(2) Asymmetric Dispersion Game

Next, we investigate the case in which the underlying game is the asymmetric dispersion game in Table 5. This asymmetric game has two equilibria with the pairs of the pure strategies

Table 5. Coupling rules learned by 2,500 agents in the symmetric dispersion game in Table 4: a lattice, network, a small-world network, and a random network.

Rule type	Initial action	Action site				Number of agents
	1	4	5	6	7	
5	0 or 1	0	0	1	0	699
6	0 or 1	1	0	1	0	805
13	0 or 1	0	0	1	1	383
14	0 or 1	1	0	1	1	613

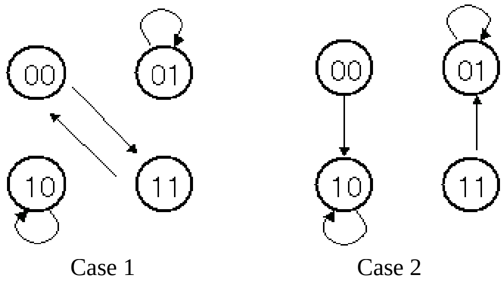


Figure 4. State transition of the two evolved coupling rules in Table 5: Case1: Two agents with the same coupling rule, Case2: Two agents with different coupling rules

of the pure strategies (C, D), (D, C). At either equilibrium, the two agents receive the different payoffs: one agent receives 1 and the other agent receive 3. There is one more equilibrium of the mixed strategy, and in this case both agents receive the same payoff, 0.75.

The average payoff per agent at each generation is shown Figure 5. Figure 5 shows the same experiments using different network frameworks. Figure 5 shows the average payoff per agent was gradually increased to 1.6 in the local network. But there are a couple of important differences in this graph and the graphs obtained using the small-work network and the random network. Under these two networks, the average payoff per agent is only about 0.6.

Table 6 Payoff matrix of the asymmetric dispersion game

Strategy of the other agent / Own strategy	*C*	*D*
C	0 0	3 1
D	1 3	0 0

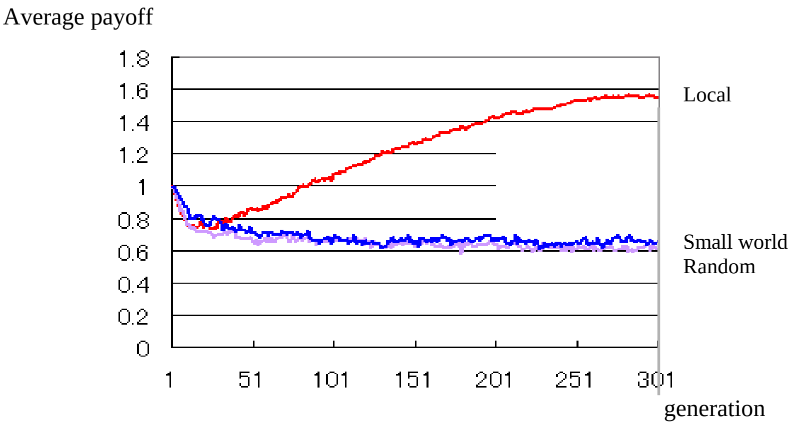


Figure 5 The average payoff per agent over generations

(3) A variant of dilemma game

We set the payoff parameters in Table 1 as *a=1*, *b=-1*, *c=4*, *d=0* as shown in Table 7. The average payoff will be 1 in a completely cooperating population. However a mixed population with agents choosing C and D is better than a completely cooperating population, since the average payoff will be greater than 1 in a mixed population. Figure 6 shows the same experiments using different network frameworks. After 100 generations the average payoff is close to 1.2, and an efficient collective outcome is realized at the macro level. There is little difference in the graphs using local and small-world networks and the graph obtained using the non-spatial random network.

Table 7 Payoff matrix of the variant Prisoner's dilemma game

Strategy of the other agent / Own strategy	*C*	*D*
C	1 1	4 -1
D	-1 4	0 0

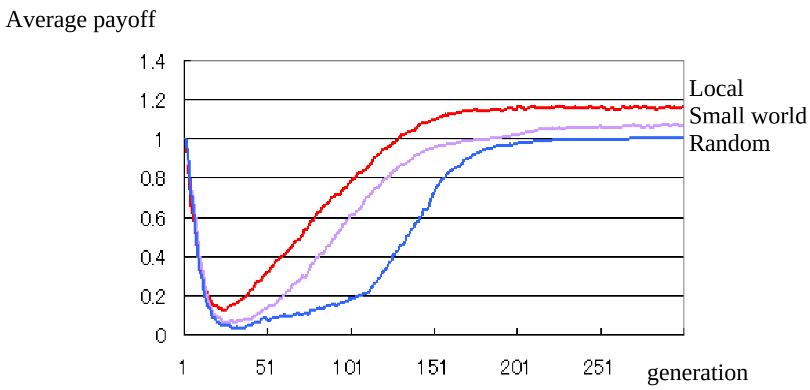


Figure 6 The average payoff per agent over generations

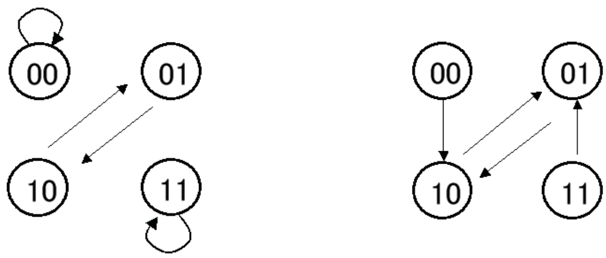
6. Effects of Network Topology

In this section we discuss the impact of different network structures on evolution of interaction rules in the context of various social interaction settings. In the local model, each agent interacts her nearest eights agents as shown in Figure 2(a). In the small-world network, each agent interacts with four agents that are nearest neighbours and four agents that are randomly chosen from the population as shown in Figure 2(b). In the random network, each agent interact with eight randomly chosen agents from the population as shown in Figure 2(c). We compare and identify the effects of the manner of interaction by investigating the properties evolved coupling rules under the different network topologies..

In Table 8, we show the coupling rules learned all agents. In the beginning, they are endowed with randomly chosen coupling rules they were finally aggregated into a few types in Table 8. The right-hand column represents the numbers of agents with the same rule. Those rules also have the commonality at the strategy sites 5 and 7.

Table 8 Coupling rules learned by 2,500 agents: the asymmetric dispersion game under the local network

Rule type	Initial action	Action site				Number of agents
	1	4	5	6	7	
3	1	0	1	0	0	335
4	1	1	1	0	0	430
11	0 or 1	0	1	0	1	919
12	1	1	1	0	1	816



Case 1 Case 2

Figure 7. State transition of evolved coupling rules in Table 8. Case1: Two agents with the same coupling rule, Case2: Two agents with different coupling rules

We show the strategy choices between two agents with the rules in Table 8 as the state transition process in Figure 7.

There is a significant difference depending on two agents whether they learn to have the same rule or different rules. If two agents with the same rule interact (Case 1 in Figure 7), there are two absorbing states at (0,0) or (1,1), and one limiting cycle of visiting (0,1) or (1,0) alternatively. In the latter situation, the turn-taking behaviour is emerged. 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) and the other agent also shifts to C(0). 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) and the other agent shifts to D(1).

In the case the two agents have different rules (Case 2 in Figure 7), they converge to only the limiting cycle of visiting (0,1) or (1,0) alternatively, and the turn-taking behaviour is also emerged. If one agent chooses C(0) and their opponent chooses D(1) at the previous round, then they choose C(0). If agents choose D(1) and their opponent chooses C(0) at the previous round, then they choose D. Therefore the locally networked agents succeeded in learning the coupling rules that have the following property: If the agent gains the payoff (success), then they change their strategies. This coupling rule is characterized as the behavioural principle based on so-called the *give-and-tak*e.

When agents face the symmetric dispersion game with the payoff matrix in Table 4, 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 game 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 learned how to realize efficient and equitable outcomes by visiting these two pure Nash equilibria alternatively.

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

Those rules in Table 9 are the same as the evolved rules shown in Table 5 when they play the symmetric game in Table 6. The action choices between two agents with the rules in Table 9 can be shown as the same state transition process in Figure 4. If agents choose C(0) and their opponent chooses D(1) at the previous round, then they choose C(0). If agents choose D(1) and their opponent chooses C(0) at the previous round, then they choose D(1). These rules represent the following behavioural rule: if they gain then they repeat the same winning action, and this is the same principle of a reinforcement learning model.

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

Rule type	Initial strategy	Strategy site				Number of agents
	1	4	5	6	7	
5	1	0	0	1	0	719
6	1	1	0	1	0	380
13	0	0	0	1	1	842
14	0	1	0	1	1	559

Therefore we can conclude that the efficient and fair outcome emerges relatively quickly in symmetric lattice networks where each agent plays the game with the same number of agents. In symmetric networks, agents appear to easily recognize the possibility of coordinated turn-taking or alternating reciprocity norm as a means to generate an efficient and fair outcome.

Although norms have been conceptualized in various ways by a variety of researchers, definition in review articles is as follows: 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 (Kandori, 1993, Young, 1993). 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 norms as core issues for the theory of norms. That is, to elaborate the norm construct fully, we need to understand how norms can emerge voluntarily through agents' interaction without external regulating forces. This perspective is shared by many 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 learning is a sufficient mechanism for norm development. Although social 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. 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 mechanism of explain norm development.

7.Selective Interactions under 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 may 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 a action choice depending on an agent's ability to select partners in the framework of game theory. A different kind of collective behaviour arises when agents change those with whom they interact before they make up their mind how to behave.

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. If agents can choose the partners with which to interact, then they will form networks that lead to an efficient outcome in the underlying games. Selective interaction in dilemma games was introduced in some studies (Tesfatsion 1996). 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.

The work by Skyrms and Pemantle (2000) triggered a lot of interest with respect to the question of the selection of interaction networks. 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 action choice, they showed that risk dominance looses 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 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 study a model in which unsatisfied agents with lower average payoff than a threshold move to new sites and interact with new agents. 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. 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, agents repeatedly play the dispersion games in Table 4, Table 6 and Table 7. In our model, agents repeatedly play these underlying dispersion games with the current neighbours and myopically adapt their actions 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 decide to remain in the same place. On the other hand, dissatisfied or unsuccessful agents move to new places 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 place 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.

Figure 8 shows the average payoff per agent under the symmetric dispersion game in Table 4, the asymmetric dispersion game in Table 6, and under the variant PD game in Table 7. Until 400 generations, there is no movement and all agents co-evolve their coupling rules as discussed in the section 5. After the agent movement is allowed, and unsatisfied agents with lower payoff than the threshold move to new places. The average payoff after movement was suddenly increased to Pareto-optimal of each underlying game.

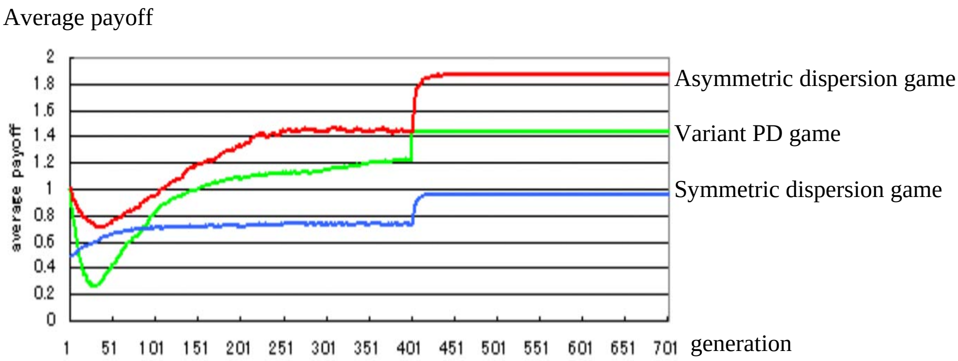


Figure 8. Average payoff under agents' movements: after the 400 generation, agents are free to move

8. Conclusion

The question of whether interacting agents self-organize desirable macroscopic behaviour 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. Although the individual decision problem is important to understand, it is not sufficient to describe how a collection of agents arrives at a specific desirable collective outcome. Therefore, we aim to discover the fundamental micro-mechanisms that are sufficient to generate the desirable macroscopic behaviours of interest. This type of self-organization is referred as emergence of desired order 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.

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 norms. Norms are also sets of socially agreed upon rules that we draw upon to structure agents' behaviour. A large literature testifies to the many ways in which norms shape behaviour and enable agents to coordinate with others.

Computer simulations on social interactions by evolving interaction rules have shown that, after thousands of repetitions of social games, norms such as such as turn-taking or alternating reciprocity emerge. 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 learns from 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 rule so-called turn-taking norm to break the asymmetry in payoffs from social interactions. In this case, the symmetric local network under which each agent play with the same number of the closest neighbours fosters to emerge such desirable turn-taking norm.

References

Browning, L., Colman, M., "Evolution of coordinated alternating reciprocity in repeated dyadic games". *Journal of Theoretical Biology* Vol.229, 2004, p. 549–557

Durkheim, E. *The Rules of Sociological Method*, The Free Press, 1982.

Ellison, G.,"Learning local interaction, and coordination", *Econometrica*, Vol.61, 1993, p.1047-1071.

Goyal, S., Vega-Redondo, F.,"Learning, network formation and coordination", *Games and Economic Behavior* Vol.50, 2005.

Hanaki, N., Erev,S.,I., Peterhansl, T., "Learning Strategies",. *Journal of Economic Behavior and Organization*, Vol. 56, 2005.p.523-542.

Hanaki, N.,"Individual and Social Learning", *Computational Economics*, 2006

Hart, S., Mas-Colell, A.,”Uncoupled dynamics do not lead to Nash equilibrium”, *American Economic Review,* Vol.93, 2003, p. 1830-1836.

Kandori, M., Mailath, G.,and Rob,B.,. “Learning, Mutation and Long Run Equilibrium in Games.” *Econometrica.* Vol.61, 1993, p.29–56.

Kandori, M. “Social Norms and Community Enforcement.” *The Review of Economic Studies,* Vol.59, No1. 1992, p.63–80.

Kirley, M.,”Compex networks and evolutionary games”. *Proc. of the 7th Asia-Pacific Conf. On Complex Networks,* 2004.

Kosfeld, M.,”Stochastic strategy adjustment in coordination games”, *Economic Theory*, Vol. 20, 2002. p. 321-339.

Lindgren, K.,”Evolutionary dynamics in game-theoretic models”, *The Economy As An Evolving Systems II*, Addison-Wesley, 1997.

Namatame, A, *Adaptation and Evolution in Collective Systems,* World Scientific, 2006.

Skyrms, B. and Pemantle,R., (2000), *A Dynamic Model of Social Network Formation*, PNAS. Vol.97, No.16, p. 9340-9346.

Tesfatsion, L., Ashlock, H., Smucker,.. “Preferential partner selection in an evolutionary study of prisoner's dilemma”, *BioSystems*, Vol.37,1996, pp. 9-25,

Young, H. P. “The Evolution of Conventions.” Econometrica Vol. 61, No.1,1993, p.57–84.

Young, H. P., *Individual Strategy and Social Structures*, Princeton University Press, 1998

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

Zimmermann.,M., and Eguíluz,V., “Cooperation, social networks, and the emergence of leadership in a prisoner’s dilemma with adaptive local interactions, Physical Review, E, Vol.72, 2005, p.056118-1-15

Watts, D. *Small Worlds*, Princeton University Press,1999.

Wu, J. and Axelrod, R. How to cope with noise in the iterated prisoner’s dilemma game, *Journal of Conflict Resolution*, 39, 1997, p. 183-189.