Emergence of Desired Collectives with Evolving Synchronized Coupling Rules

Akira Namatame,
Dept. of Computer Science, National Defense Academy,
Yokosuka, 239-8686, JAPAN
Email: nama@nda.ac.jp

Abstract

Norms are essential for human that are to display behavior comparable to social behavior or collaborate with other humans, because the use of norms is the key of human social intelligence There is also an increasing interest in the role of norms in the artificial intelligence community. An interesting problem is under what circumstances will a collection of agents evolve desirable norms so that they could realize efficient outcomes. This question will depend crucially on how they interact each other. In examining collective, we shall draw heavily on the individual adaptive behavior. Although individual behavior is nested within important to understand, it is not sufficient to describe how a collection of self-interested agents realizes a specific desired collective.

We consider a collection of agents and ask the following question: how they group its way towards a collectively desired outcome in an imperfect world when self-interested agents learn each? The envisioned research object is quite novel, since it requires harmonic and synchronized behavior among self-interested agents.

Agents are modeled to interact with each other based one of the following underlying games such as dilemma games, coordination games, hawk-dove games, or dispersion games (minority games). It is well known that the best-response dynamics and the replicator dynamics based on natural selection lead a collection of agents to inefficient equilibrium.

In this paper, we focus on co-evolutionary dynamics based on coupled learning among agents. Each agent is modeled to learn the coupling rule rather strategy. We show a society of agents can realize and sustain an efficient and equitable situation by co-evolving synchronized coupling rules. The space of all coupling rules is so complex, however, all agents can be categorizes by their learnt coupling rules into a few classes. By investigation these aggregated coupling rules, we show that how harmonic behaviors or and dynamic synchronized behavior will be emerged among self-interested agents.

The performance of the collective system consisting of interacting agents depends on how agents are properly coupled. The approach of coupled learning is very much at the forefront of the general topics of design of desired collectives in terms of efficiency, equity, and sustainability. The mission of coupled learning is to harness the emerging undesirable collective behavior to design collective systems that serve to secure a sustainable relationship among agents in an attainable manner.

Collective means any complex system of autonomous agents, together with a performance criterion by which we rank the behavior of the overall system. There are also growing interests for studying of collective phenomena including the dynamics of markets, the emergence of social norms and conventions, and daily phenomena such as traffic jams. There is no presumption that a collective action of interacting agents leads to a collectively satisfactory result without any central authority. How well agents do for it in adapting to their environment is not the same thing as how satisfactory an environment they collectively create. In this paper, I address the following questions: How should interacting agents learn collectively to reach desired collectives? This question depends on how they interact and adapt their behavior. Agents normally react to others' decisions, and the resulting volatile collective action is often far from being efficient. There are two closely related issues concerning collectives, the forward and inverse problems. The forward problem is to investigate what a collective of interacting agents determines its complex emergent behavior. The inverse problem is to investigate how should agents behave and learn to generate a desirable collective behavior. In this paper, I will show that it is important for self-interested agents to establish proper relationship with others by learning coupling rules.

Keywords: forward problem, inverse problem, collective learning, coupling rule, sharing superior rule

1. Introduction

Today we have many challenges for designing large-scale and complex systems consisting of multiple physically and geographically distributed processors. Collective means any pair of a complex system of autonomous system components, together with a performance criterion by which we rank the behavior of the overall system. The design of desired collectives from bottom up also becomes an important issue in many areas [9][18]. The performance of the collective system which consists of many interacting agents should be described on two different levels: the microscopic level, where the decisions of the individual agents occur and the macroscopic level where collective behavior can be observed. To understand the role of a link between these two levels also remains one of the challenges of complex system theory [15].

Natural evolution has ceased a multitude of systems in which the actions of interacting agents give rise to coordinated global information processing. An interesting aspect of interacting agents is emergent property. Emergent property is surprising because it can be hard to anticipate the full consequences of even simple forms of interaction. Emergence also refers to the appearance of global information-processing capabilities that are not explicitly represented in the system's elementary components or in their interconnection. Insect colonies, cellular assemblies, the retina, and the immune system have been often cited as examples of having the emergent properties. In order to investigate emergent property, we have to specify how agents, system components, interact, respond, adapt, or learn from each other. We attempt to probe deeper understanding this issue by considering decision-makings of heterogeneous agents.

We address the following question: how do the worlds of heterogeneous agents generate global macroscopic orders as the whole? There is no presumption that a population of self-interested agents in an imperfect world leads to a collectively satisfactory result. How well agents do for it in adapting to their environment is not the same thing as how satisfactory an environment they collectively create. An agent behaves based on not only his preference but also others' actions. It is also important to consider with whom an agent interacts and how each agent decides her action depending on others' actions.

Wolpert and Tumer propose that the fundamental issue is to focus on improving our formal understanding of two closely related issues concerning collective [18]:

(1) The forward problem of how the fine-grained structure of the system underlying a collective determines its complex emergent behavior and therefore its performance.

(2) The inverse problem of how to design the structure of the system underlying a collective to induce optimal performance.

We specify how the agents interact, and then observe properties that occur at the macro level. The connection between micro-motivation and macro-outcomes will be developed through agent-based simulation, in which a population of agents is instantiated to interact according to fixed or evolving rules of behavior. An interesting problem is then under what circumstances will a collection of agents realizes some particular stable situations, and whether they satisfy the conditions of efficiency? An agent wishes to maximize her own utility and a system designer wish to implement a decentralized algorithm for maximizing the whole utility of a group. While all agents understand the outcome is inefficient, acting independently is powerless to manage this collective about what to do and also how to decide. In previous works, the standard assumption was that agents use the same adaptive rule [1][3][4]. In this paper, we depart from this assumption by considering heterogeneity in term of their payoffs. We focus on locally interacting agents and show how agents interact with the others and their aggregate. We formalize our idea by modeling a population of heterogeneous agents in which agents are repeatedly matched within a long-time period to play games. There are many parameters to be considered such as payoff matrix, population structure, configuration, the number of agents and so on. Among these parameters, we examine the heterogeneity of payoffs and the configuration of locating heterogeneous agents.

There are many situations where interacting agents can benefit from coordinating their action. Coordination usually implies the increased effort by some agents may lead the remaining agents to follow suit, which gives to rise multiplier effects. These strategic situations are modeled as coordination games in which agents receive payoff if he selects the same strategy as the majority does. We call this type of interaction as symmetric interaction. In this case agents behave based on the logic of majority, since agents receive payoff if they select the same strategy as the majority does. Symmetric interaction usually implies that increased effort by some agents leads the remaining agents to follow suit, which gives multiplier effects [2]. On the other hand, in the route selection problem, agents receive payoff if they select the opposite strategy as the majority does. Let consider a competitive routing problem of networks, in which the paths from sources to destination have to be established by independent agents. For example, in the context of traffic networks, agents have to determine their routes independently. This type of interaction is distinguished by defining as asymmetric interaction. In this case agents behave based on the logic of minority, since agents receive payoff if they select the opposite strategy as the majority does. The EL Farol bar problem and its variants minority games may provide a clean and simple example of this type [5][12]. In this paper, we consider the generalized Rock-Scissor-Paper game as the underlying game of asymmetric interactions among agents.

Qualitative aspect of evolutionary dynamics have been investigated by many researchers, (i) by varying the payoff matrix of the game, (ii) by introducing spatial dimensions, and (iii) by introducing the mechanism of co-evolution. An important aspect of evolution is the learning strategy adapted by individuals. The term evolutionary dynamics often refers to systems that exhibit a time evolution in which the character of the dynamics may change due to internal mechanisms. In this paper, we also focus on co-evolutionary dynamics that may change in time according to certain local rules that are co-evolved over generation. Co-evolutionary models can be characterized both by the level at which the mechanisms are working and the dimensionality of the system. Therefore we need to describe co-evolutionary dynamics by specifying collective learning at microscopic levels. Agents myopically evolve their actions based on their own rules obtained as the function of their idiosyncratic utility and the actions of their neighbors The assignment of heterogeneous agents also becomes to be important. We investigate the relation between the collective behavior and the assignment of heterogeneous agents.

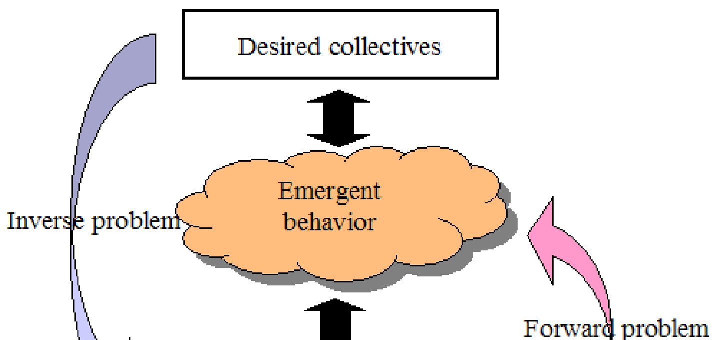


Figure 1. The forward and inverse problems of collectives

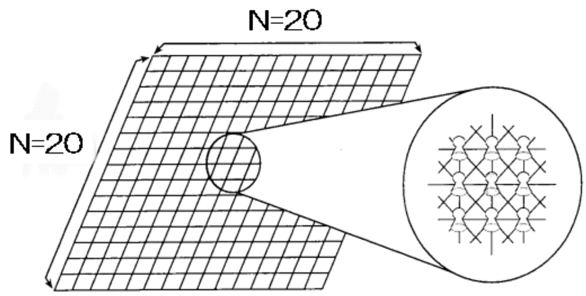


Figure 4. Social Games on the lattice

We consider the lattice structure as shown in Figure 4, in which each agent interacts with her 8 neighbors. We arrange a population of heterogeneous agents in the area of $20x20$ (400 agents in total) with no gap, and four corners and an edge of an area connect it with an opposite side. Each agent chooses an optimal strategy based on an information about what his neighbors behaved in the past. The consequence of their actions also gives an effect on agents with whom not directly linked. 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-field model (random matching) 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 model.

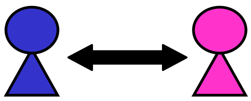
2. Coupling Rules

An important aspect of collectives is the learning strategy adapted by individuals. We make a distinction between evolutionary systems and adaptive systems. The adaptation may be at the individual level through learning, or it may be at the population level through differential survival and reproduction of the more successful individuals. Either way, the consequences of adaptive processes are often very hard to deduce when there are many interacting agents following rules that have nonlinear effects.

Several learning rules have been found to lead an efficient outcome when agents learn from each other. Many adaptive models have been proposed such as best-response dynamics or payoff improving learning, and soon. However, it is notoriously difficult to formulate adaptive dynamics that guarantee convergence to Nash equilibrium. We call a dynamical system uncoupled if the adaptive dynamic for each agent does not depend on the payoff functions of the other agents. Hart showed that there are no uncoupled dynamics that are guaranteed to converge to Nash equilibrium [7]. Therefore coupling among agents, that is, the adjustment of an agent's strategy depend on the payoff functions of the other agents is a basic condition for adaptive dynamics.

We take a different approach by focusing co-evolution. Co-evolutionary dynamics differ, in this sense, from the common use of the genetic algorithm, in which a fixed goal is used in the fitness function and where there is no interaction between individuals. In the genetic algorithm, the focus is on the final result what is the best or a good solution. In models of co-evolutionary systems, one is usually interested in the transient phenomenon of evolution, which in the case of open-ended evolution never reaches an attractor. We can make a distinction between evolutionary dynamics and adaptive dynamics. The equations of motion in an evolutionary dynamics reflect the basic mechanisms of biological evolution, i.e., inheritance, mutation, and selection. In an adaptive dynamics, other mechanisms are allowed as well, e.g., modifications of strategies based on individual forecasts on the future state of the system. But, increasing the possibilities for individualistic rational behavior does not necessarily improve the outcome for the species to which the individual belongs in the long run.

The search for evolutionary foundations of game-theoretic solution concepts leads from the notion of an evolutionarily stable strategy to alternative notions of evolutionary stability to dynamic models of evolutionary processes. The commonly used evolutionary dynamics which model the evolutionary process as a system of a deterministic difference or differential equations may tell us little about equilibrium concepts other than that which Nash equilibrium is realized. We can attempt to probe deeper into these issues by modeling the choices made by the agents with their own internal models. We especially focus on co-evolutionary dynamics described by equations of motion that may change in time according to certain rules of agents, which can be interpreted as crossover operations. It is known that natural selection does not lead to social efficiency in these games. We show that all agents mutually learn to cooperate which result in social efficiency.



bit	previous strategy		next strategy
	Own	Opp	
4	0	0	#
5	0	1	#
6	1	0	#
7	1	1	#

Figure 6 The definition of a coupling rule

The 2x2 games with the payoff matrix in Table 1 or Table 2 is examined by how the four outcomes are rank ordered by the agents. Each agent is represented by a strategy specifying how the agent behaves as it interacts with other agents. Agent strategies are restricted to those employing just the previous one move or two with the other agent(s) to determine current choices. Each strategy is represented as a binary string so that the genetic operators can be applied. In order to accomplish this we treat each strategy as deterministic bit strings. We use a memory of one (or two), which means that the outcomes of the previous one or two moves are used to make the current choice. We represent S_1 by 0 and S_2 by 1. As Table 2 shows, there are four possible outcomes for each move between two agents, $(S_1,S_1)=(0,0)$, $(S_1,S_2)=(0,1)$, $(S_2,S_1)=(1,0)$, and $(S_2,S_2)=(1,1)$. By coupling rule we specify a deterministic strategy choice by recording what the strategy will do in each of the 4 different combinations of the strategy choices with the other. Since there are four combinations of memory 1, there exist $2^4=16$ possible coupling rules, which specify either 0 or 1 represented by # in Table 3. In a population of heterogeneous agents, there are two types of agents: (1) Type 1: Agent with the threshold $\theta_i \leq 0.5$. Such an agent gains more payoff by choosing S_1. Therefore the best strategy of agents of type 1 is always to choose S_1 and the efficient coupling rule is given [0,0,0,0]. On the other hand, (2) Agents of the type 2 with the threshold $\theta_i > 0.5$ gain higher payoff by choosing S_2. Therefore the best strategy of an agent of the type 2 is always to choose S_2 and the efficient coupling rule is given [1,1,1,1].

Since at the start, an extra two bits are needed to specify a hypothetical history.

Each rule with memory 1 can be defined by a 6 bit string as shown in Figure 7. Since no memory exists at the start, an extra two bits are needed to specify a hypothetical history. Each rule with memory 1can be defined by a 6-bit string as shown in Figure 7. In order to accomplish a coupling rule in Table 2, we treat each strategy as deterministic bit strings. Since no memory exists at the start, an extra 4 bits are needed to specify a hypothetical history as shown in Figure 7. At each generation, agents repeatedly play the game for T iterations. Agent i, $i \in [1 \ldots N]$ uses a binary string i to choose his strategy at iteration t ,$t \in [1 \ldots T]$. Each position of a binary string in Figure 7 as follows: The first position, p_1 encodes the action that agent takes at iteration $t = 1$. A position p_j, $j \in [2,3]$ encodes the memories that agent i takes at iteration $t - 1$ and his opponent. A position p_j , $j \in [4 \ldots 7]$, encodes the action that agent i takes at iteration $t > 1$, corresponding to the position p_j, $j \in [2,3]$. An agent i compares the position p_j, $j \in [2,3]$, and decision table as shown in Table.4, and then, an agent i decide the next action. Here is an example of binary string given the agent's action taken in the previous iteration.

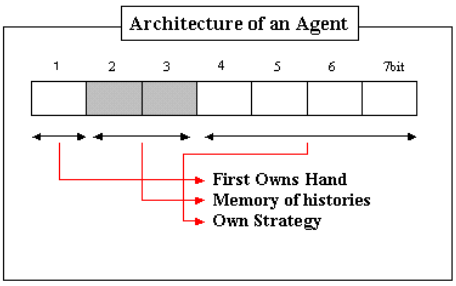


Figure 7 Representation of a coupling rule in Figure 6

3. Simulation Results

We arrange agents for an area of 20×20 (N = 400 agents) as the lattice model as shown in Figure 1(a) with no gap. Four corners and end of an area connect it with an opposite side. All agents of the 20x20 lattice play 2x2 games against their eight agents. The payoff of these eight repeated games are summed and the average payoff per game provides the individual’s score. A genetic algorithm is used to evolve interaction rules. After every agent has repeated games during one generation, the interaction rule of each agent is updated according to the crossover operator: a half of the interaction rule with 5 bits in Figure 8 (including the first bit which represents the initial strategy of the next generation) is replaced with the rule of the neighbor of the highest scoring.

We also consider the error at the choice of the strategy. Agents choose their strategy that is specified by the meta-rule. However, we assume there exists small probability of choosing the wrong strategy. We showed the simulation results without any error and with the error rate 5% in Figure 9. Consequently, we can conclude that evolution learning leads to a more efficient situation in the strategic environments. Significant differences were observed when agents have small chances of making mistakes. As shown Figure 9(a), the highest payoff and the lowest payoff become to be close, which imply that each agent acquires the almost the same. We can fully describe a deterministic strategy by recording what the strategy will do in each of the 16 different situations that can arise in the iterated game.

We use evolutionary learning where agents learn from the most successful neighbor. Each agent adapts the most successful strategy as guides for their own decision (individual learning). Hence their success depends in large part on how well they learn from their neighbors. If the neighbor is doing well, her strategy can be imitated by neighbors. A part of the list is replaced with that of the most successful neighbor.

In the simulation result shown in Figure 8(a) without any mistake, the average payoff per an agent was gradually increased to 0.75. We also show the highest payoff and lowest payoff, and there are lucky agents got the maximum payoff of 1 since they were and the unlucky agents got nothing since they were always belong to the majority side. We also consider the implementation error of the strategy. Agents choose their strategy specified by the coupling rule. We showed the simulation results with the error rate 5% in Figure 8(b). Consequently, we can conclude that co-evolution learning leads to a more efficient situation when agents have small chances of making mistakes. As shown Figure 8(b), the highest payoff and the lowest payoff become to be close by implying that each agent acquires the almost the same payoff in the long-run.

From this simulation result we can conclude that some mistakes in co-evolutionary learning of coupling rules that lead a collective behavior to be both efficient and equitable. For self-organizing systems, fitter usually means better or with more potential for growth. However, the dynamics implied by a fitness landscape does not in general lead to the overall fittest state, and the system has no choice but to follow the path of steepest descent. This path will in general end in a *local* minimum of the potential, not in the *global* minimum. Apart from changing the fitness function, the only way to get the system out of a local minimum is to add a degree of indeterminism by adding some noise to the co-evolutionary dynamics, that is, to give the system the possibility to make transitions to states other than the locally most fittest one. This can be seen as the injection of noise or random perturbations into the system, which makes it deviate from its preferred trajectory. Physically, this is usually the effect of outside perturbations. Such perturbations can push the system upwards, towards a higher potential. This may be sufficient to let the system escape from a local minimum.

6. Generalized Rock-Scissor-Paper Games

There are many situations where agents receive payoff if they select the opposite strategy as the majority does. The underlying game of his type of asymmetric interaction is modeled as anti-coordination game. In this section we generalize the Rock-Scissor-Paper (R-S-P) game as the underlying game of asymmetric social interactions. It is not surprising that games as absorbing as bridge and chess have their world federations and international unions. But not everyone knows that even a game as lowly as rock-scissor-paper has its own society. This game, which must surely be very old, can be explained to any toddler. Two players signal, on a given cue, either rock (fist), scissors (two fingers), ore paper (flat hand). If I display a flat hand and you show me your fist, I win, as 'paper wraps rock'. Similarly, scissors cuts paper, and rock smashes scissors. If both players make the same signal, the game ends in a draw.

We generalize the R-S-P game with the payoff matrix given in Table 3. If $\lambda=2$, this game is equivalent with the original R-S-P game. The game with the payoff matrix in Table 2 has the unique mixed Nash equilibrium. The mixed Nash equilibrium strategy is to select rock, scissor, and paper with the same probability 1/3. The expected payoff for each agent under this mixed Nash equilibrium strategy is $(\lambda+1)/3$.

With the generalized R-S-P game, we are especially interested in the effect of changing the parameter value λ. The most preferable situation is that each agent wins and lose alternatively by adopting the different strategy. We examine how they play if the parameter value λ is increased.

Table. 3. The payoff matrix of the generalized Rock-Scissor-Paper game $(\lambda \geq 2)$

opponen own	S_1 (Rock)	S_2 (Scissor)	S_3 (Paper)
S_1 (Rock)	1 1	0 λ	λ 0
S_2 (Scissor)	λ 0	1 1	0 λ
S_3 (Paper)	0 λ	λ 0	1 1

Each strategy in the repeated game is represented as a string so that the genetic operators can be applied. In order to accomplish this we treat each strategy as deterministic bit strings. We use a memory, which means that the outcomes of the previous one move are used to make the current choice. There exit nine possible outcomes for each move ((S_1,S_1), (S_1,S_2), (S_1,S_3), (S_2,S_1), (S_2,S_2), (S_2,S_3), (S_3,S_1), (S_3,S_2), (S_3,S_3)). We can fully describe a deterministic strategy by recording what the strategy will do in each of the nine different situations that can arise in the iterated game. Since no memory exists at the start, extra one bit is needed to specify a hypothetical history. If we assume that 0 = S_1, 1 = S_2, and 2 = S_3 then each strategy can be defined by a 3-bits string. As shown in Figure 12, each agent is composed of the binary strings which represents the rule of strategy choices. The strategy choice is determined by the combination of the strategies of her opponent of the previous time period. Agents co-evolve by swapping the part of the coupling rule for that of the most successful neighbor. Therefore, the efficient rule has a chance to be spread throughout a population from neighbor to neighbor [13][14].

An agent is represented as A_i, $i \in [1,\ldots,N]$ and his decision rule is represented by the temary string. At each generation $gen, gen \in [1,\ldots,lastgen]$, agents repeatedly play the game for T iterations. An agent A_i, $i \in [1,\ldots,N]$ uses a temary string to make a decision about his action at each iteration t, $t \in [1,\ldots,T]$. A temary string consists of 12 positions (genes). Each position p_j, $j \in [1,\ldots,12]$ is represented as follows: The first position p_1, encodes the action that the agent takes at iteration $t = 1$. A position p_j, $j \in [2,3]$, encodes the history of mutual hands (rock, scissor, or paper) that A_i took at iteration $t - 1$ with his neighbor (opponent). A position p_j, $j \in [4,\ldots,12]$, encodes the action that A_i takes at iteration $t > 2$, corresponding to the position p_j, $j \in [2,3]$. If A_i decides his rule by iteration $t-1$ and $t-2$, A_i's gene is represented. The position p_k, $k \in [1,2]$, encodes the action that the agent takes at $t = 1$ and $t = 2$. And p_k, $k \in [3,\ldots,6]$ encodes the history that an agent adopted at iteration $t-1$ and $t-2$, and p_k, $k \in [7,\ldots,87]$ does the action that A_i takes at iteration $t > 3$.

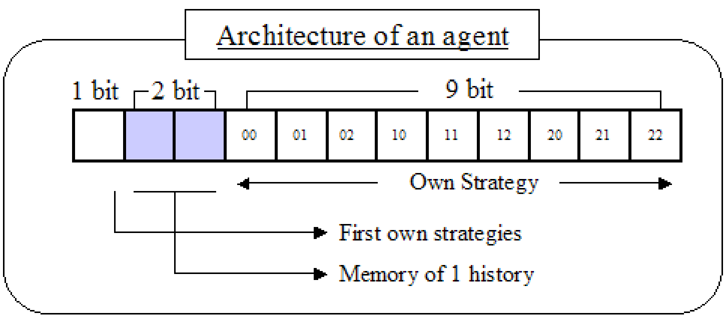


Figure 12 The agent's memory architecture (history of 1)

Each agent interacts with the agents on all eight adjacent squares and cross-overs the rule of any better performing one. In each generation, each agent attains a success score measured by its average performance with its eight neighbors. Then if an agent has one or more neighbors who are more successful, the agent cross over with the rule of the most successful neighbor. Each agent is matched several times with 8 neighbors, and the list of the strategies (rule of interaction) for the same partner is coded as the list. With partial mimicry, cross-over, a part of the list is replaced with that of the most successful agent. The neighbors also serve another function as well. If the neighbor is doing well, the behavior of the neighbor can be shared, and successful strategies can spread throughout a population from neighbor to neighbor [14].

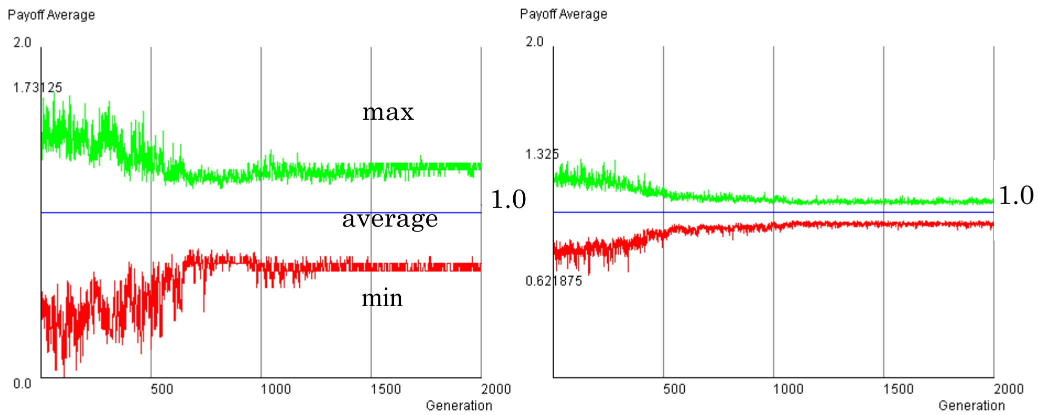
We also consider situations where there exists a small probability of choosing the wrong strategy, which is defined as an implementation error. Agents sometimes make error by choosing different strategy that is specified by the rule. We will compare the simulation results without any errors and with a small percentage of error. We will show significant differences will be observed at macroscopic level when agents have small chances of making mistakes at the microscopic level.

7. Simulation Results

In this section, we investigate the property of co-evolutionary learning among agents. We especially investigate what the learning influence from agents' payoff. Each agent adopts the most successful strategy as guides for their own decision (individual learning). Hence their success depends in large part on how well they learn from their neighbors. If the neighbor is doing well, her rule will be more likely imitated by others. We define the mutual exchange process of better rules as collective learning. In an evolutionary approach, there is no need to assume a rational calculation to identify the best strategy. Instead, the analysis of what is chosen at any specific time is based upon an implementation of the idea that effective strategies are more likely to be retained than ineffective strategies. Moreover, the evolutionary approach allows the introduction of new strategies as occasional random mutations of old strategies. The evolutionary principle itself can be thought of as the consequence of any one of three different mechanisms. It could be that the more effective individuals are more likely to survive and reproduce. A second interpretation is that agents learn by trial and error, keeping effective strategies and altering ones that turn out poorly. A third interpretation is that agents observe each other, and those with poor performance tend to imitate the strategies of those they see doing better.

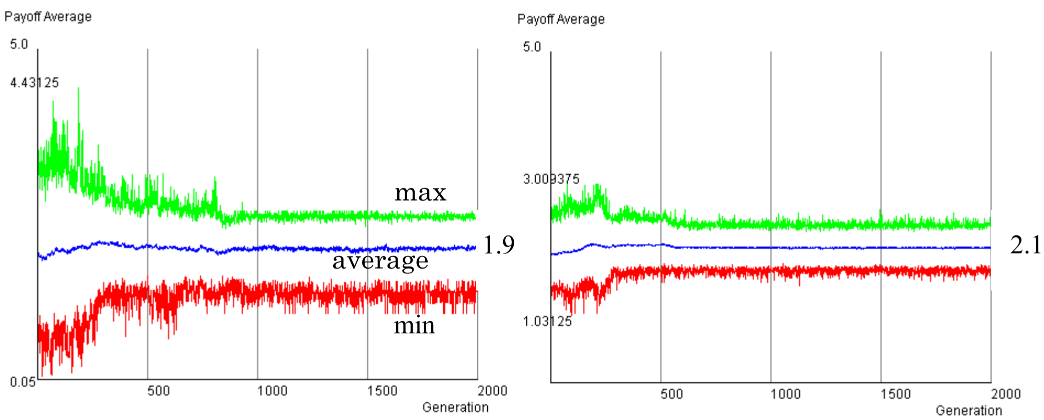
In this simulation, we consider the case in which each underlying interaction is modeled as generalized R-S-Paper game with the payoff matrix in Table 3. Number of agents are 400 (N=400). At each time period, each agent plays the game in Table 3 with his eight neighbors and they decide next strategy by referring 1-history or 2-history. At next time period, agents cross-over their rule with the one of the most successful neighbor who obtain the highest payoff.

We simulated several cases by changing the parameter λ. Figure 13, 14, and 15 shows transition of payoff average to generation in cases when we set $\lambda = 2$, 5, or 10, and shows the distribution of payoff average which agents acquired in the final generation. And we simulated to the each case to be no noise and each ten percent. It played a game 20 times a generation and the 2000 generations was repeated to see the changing society enough.



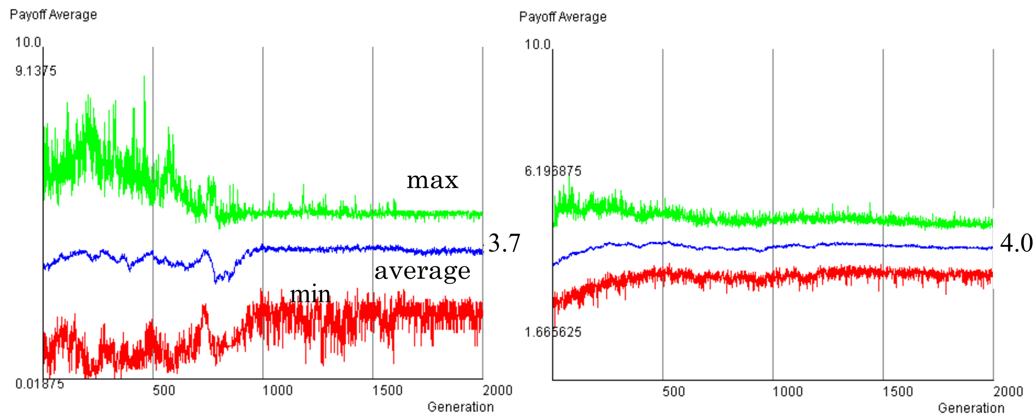
(a) The average payoff of each generation: error rate: 0%: (b) error rate: 10%

Figure 13 Simulation result with $\lambda = 2$ in Table 3



(a) error rate: 0% (b) error rate: 10%

Figure 14. Simulation result with λ=5



(a) error rate: 0% (b) error rate: 10%

Figure 15. Simulation result with λ=10

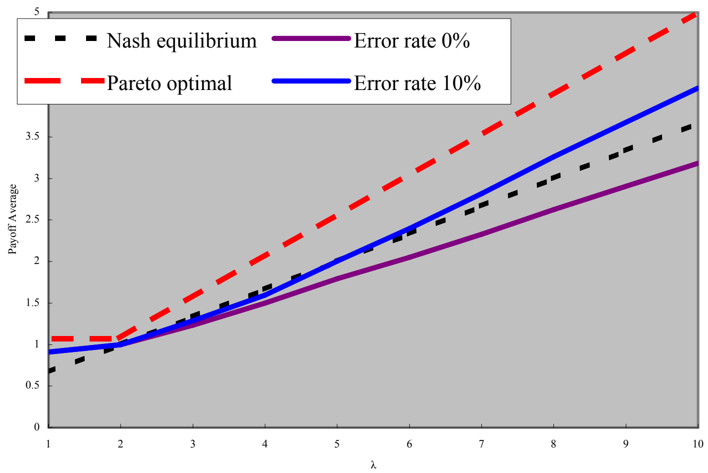


Figure 16. The average payoff by changing the value of λ in Table 2

As shown in Figure 16, all agents receive the same average payoff if λ=2. If the payoff by winning the game increases (λ>2), the cooperative relationship collapses, and resulting in lower average payoff than at Nash equilibrium. It means the agent society becomes a Nash equilibrium. Swinging is caused in the inferior state of a Nash equilibrium when the noise is added ten percent and the society as a whole faces settling in the state of Pareto optimum.

8. What Types of Coupling Rules Did Agents Co-evolve?

In this section, we investigate what types rules are learned by agents. When agents remember the mutual moves in previous one steps, one bit is needed for the initial strategy, 2 bits are for the memory of the past history and 9 (3x3) bits are for the rule of the strategy choices. Thus, 12 bits are required for each agent. Table 4 shows the different types of rules learned by 400 agents (20 x 20 in a lattice). At the beginning each agent has a different rule. The number of different types of rules learned by 400 agents after the co-evolution process is shown in Table 4. This is the simulation result for $\lambda=10$, and without any implementation error and error with the rate of 10%, Without any implementation error, 400 different rules were aggregated into 238 rules after the 2000-th generation.

With the implementation error at 10%, however, 400 different rules were aggregated into 8 rules. The number in the bracket in Table 5 represents the number of agents who converged to same strings. These 8 type rules are shown in Table 5, and they also have some commonality. Each rule has the same bits in common place that correspond to

Table 4 The number of different learning rules among 400 agents

Generation	Number of different rules	
	Error rate: 0%	Error rate: 10%
500	400	400
1000	400	250
1500	368	30
2000	238	8

Table 5 Some commonality among learned coupling rules ($\lambda=10$, error rate 10%)

	Initial_Strategy	00	01	02	10	11	12	20	21	22	
Rule Type 1:	1	2	2	2	0	2	1	1	0	2	(126)
Rule Type 2:	1	2	2	2	0	2	1	1	0	0	(76)
Rule Type 3:	1	1	2	2	0	2	1	1	0	2	(58)
Rule Type 4:	1	1	2	2	0	2	1	1	0	0	(54)
Rule Type 5:	1	2	2	2	0	0	1	1	0	2	(40)
Rule Type 6:	1	1	2	2	0	0	1	1	0	2	(19)
Rule Type 7:	1	2	2	2	0	0	1	1	0	0	(17)
Rule Type 8:	1	1	2	2	0	0	1	1	0	0	(10)

the combination of the previous strategies, (0,1),(0,2),(1,0),(1,2),(2,0),(2,1). We investigate characteristics of those aggregated 8 types of rules using the state diagrams. Although there are 8 different types of rules, the state diagrams are classified only two types: the state diagram for any two agents with the same type of rule, and for any two agents with the different type of rule.

(1) Homogeneous agents with the same type of the coupling rule

The state diagram between two agents who has the same type of rule *i*, i=1,2,…8, becomes the same as shown in Figure 17. The state diagram in Figure 17 is characterized by one absorbing state at (2,2), and one limiting cycle. If both agents plays either 0(rock) or 1(scissor), they eventually converge to the sate (2,2), and they play with 2((paper) forever. In this case both agents draw receiving the payoff of 1. The other limiting cycle is characterized the cycles such that one agent win three times by choosing 0(rock), 2(paper) and 1(scissor), then she lose three times by choosing 1(scissor), 0(rock), 2(paper). On average each agent gains the payoff of $\lambda/2$.

They eventually converged to the sate (2,2), and they play with 2((paper), one cycle which eventually absorbing into (22), and one l in the figure has an absorbing state via state of 00 or 11. On the other hand, right cycle has no absorbing states and it shows efficient cycle that agents win 3 times and lose 3 times each other.

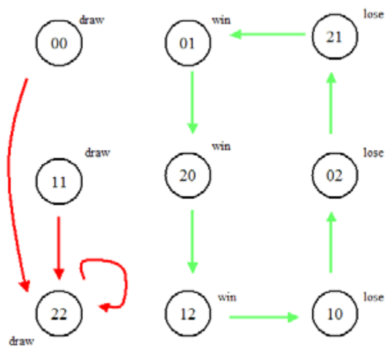


Figure 17 The state diagram of play between agents with the same rule type

(2) Heterogeneous agents having the different types of the coupling rules

The state diagram of two agents with the different rules is shown in Figure 18. Figure. 18 is basically the same as the state diagram in Figure 17, except that there are chances to move into the efficient cycle from (0,0) and (2,2). Therefore both agents have chances to receive higher payoffs by avoiding draw. Agents can move to efficient cycle without implementation errors.

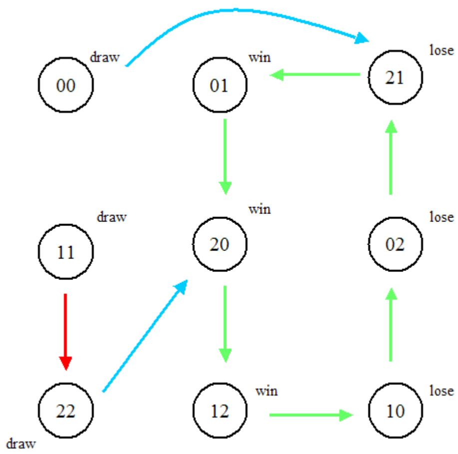


Figure 18 The state diagram of agents with the different rule

If different type agents play, they can move to efficient cycle without implementation errors. However, they could not stay in the efficient cycle by implementation errors with probability *p*, and go to inefficient cycle. In this situation, implementation errors would interrupt acquiring efficiently the payoff. Therefore, if homogeneous agents rule play each other, they can get efficiently payoff because they start at draw cycle, else heterogeneous agents can stay the most efficient cycle. How agents with the different rule structure is the most important for achieving efficient equilibrium.

The interaction of heterogeneous agents produces some kind of coherent and systematic behavior. We investigated the macroscopic patterns arising from strategic interactions of heterogeneous agents who behave based on local rules. The hypothesis we employed in this paper reflects limited ability of agent to receive, decide, and act according to the information they get in the course of interactions. Emergent efficient collective behavior is based on the concept of give-and-take at individual levels that produce both efficient and equitable behavior. These results implicate that we have a tools for examination how conventions evolve in a society that begins in an amorphous state where there is no established custom and individuals only rely on hearsay to determine what to do. From the discussion we confirm locality and the low probability of error are the key of emerging a sort of give-and-take strategies.

9. Equilibrium Analysis vs. Evolutionary Design Analysis

By surveying recent works on multi-agent learning, Shoham characterized many models on learning into three basic agenda [16]. The basic agenda is equilibrium analysis, and which is descriptive It asks how humans learn in the context of other learners.

Researchers propose various dynamics that are a perceived as plausible in one sense or another, and proceed to investigate whether those converge to equilibria. This is a key concern for game theory, since a successful theory would support the notion of Nash equilibrium, which play a central role in non-cooperative game theory. The second agenda is so called 'the AI agenda'. This agenda asks what the best learning strategy is for a given agent for a fixed class of the other agents in the game. The third agenda is more prescriptive. They ask how agents should learn. This is a problem of distributed control; a central designer controls multiple agents, but cannot or will not design an optimal policy for them. Instead it assigns them each an adaptive procedure that converges to an optimal policy. In this case there is no role for equilibrium analysis; the agents have no freedom to deviate from the prescribed algorithm.

Much of the literature on economic design deals with the search for mechanisms where the agents will have dominant strategies that lead to desired behavior, e.g. maximizing social efficiency. However, it can be shown that such mechanisms rarely exist. One alternative is to consider mechanisms where there exists an ex-post equilibrium: a strategy profile of the agents in which it is irrational to deviate from each agent's strategy, assuming the other agents stick to their strategies, and *regardless* of the state of the system. M Tennenholtz introduced the concept of efficient learning equilibrium (ELE), a normative approach to learning in non-cooperative settings [17]. In ELE, the learning algorithms themselves are required to be in equilibrium. In addition, the learning algorithms must arrive at a desired value after polynomial time, and a deviation from the prescribed ELE become irrational after polynomial time. We prove the existence of an ELE (where the desired value is the expected payoff in a Nash equilibrium) and of a Pareto-ELE (where the objective is the maximization of social surplus) in repeated games.

The learning takes place by self-interested agents. Notice that the learning algorithms should satisfy the desired properties for the underlying game in a given class despite the fact that the actual game played is initially unknown. This system state may not be initially observable and might consist of various private inputs of the agents. In these situations, we may need to deviate from the spirit of equilibrium theory. As an alternative approach, we proposed an evolutionary design approach as shown in Figure 19. With this approach we try to find useful collective learning algorithms. We also provide the agents with learning algorithms to use. This kind of evolutionary device is not a designer who can enforce behaviors, and it does not possess any private knowledge or aim to optimize private payoffs. Therefore, the right way to view this design approach is as a kind of mediator or correlation device. It is up to each agent to decide whether to use it. However it is the proof that the algorithms are in social efficiency and equity that suggests that these suggested algorithms will be actually executed by both individually rational and social agents.

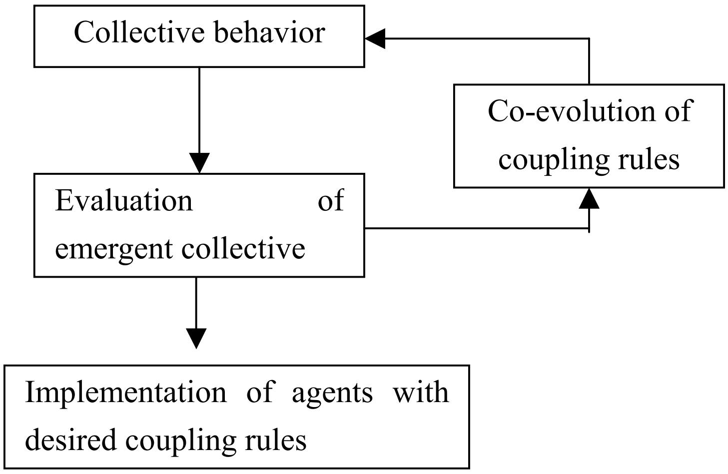


Figure 19 An evolutionary design with agent-based simulation

10.Conclusion

Many fundamental changes to a society result from collective behaviors of interacting agents. How do heterogeneous micro-world of individuals generate the global macroscopic orders and regularities of society? Much of hidden knowledge is underlying accumulated interactions. There are two closely related issues concerning collective, (1) the forward problem of how the fine-grained structure of the system underlying a collective determines its complex emergent behavior and therefore its performance, and (2) the inverse problem of how to design the structure of the system underlying a collective to induce optimal performance.

We considered two different types of interaction formulated as the coordination problem and the anti-coordination problem. We showed that the most crucial factor that considerably improves the performance of the system of interacting agents is the endogenous selection of the partners and reinforcement of preferences at individual levels. The interaction of heterogeneous agents produces some kind of coherent and systematic behavior. Emergent efficient collective behavior is based on the concept of give-and-take at individual levels that produce both efficient and equitable behavior. These results implicate that we have a tools for examination how conventions evolve in a society that begins in an amorphous state where there is no established custom and individuals only rely on hearsay to determine what to do. From the discussion we confirm locality and the low probability of error are the key of emerging a sort of give-and-take strategies.

References

[1] Axelrod R., *The Complexity of Cooperation*, Princeton Univ. Press, 1997.

[2] Banerjee. S., *Efficiency and Stability in Economic Networks*, mimeo: Boston University, 1999

[3] Batten,D.,"Co-Evolutionar Learning on Networks", in Beckmann, Johansson, B(eds), Knowledge and Networks in a Dynamic Economy, pp311-332,1998.

[4] Durlauf S.N. and Young H.P., *Social Dynamics*, Brookings Institution Press, 2001.

[5] Duffy, J., and Hopkins.E., "Learning, Information and Sorting in Market Entry Games, working paper, University of Pittsburgh, 2002

[6] Fudenberg, D., and Levine, D., The Theory of Learning in Games, The MIT Press, 1998.

[7] Hart S. and Mas-Colell,A.,"Uncoupled Dynamics Do Not Lead to Nash Equilibrium",memo, The Hebrew University of Jerusalem, <http://www.ma.huji.ac.il/~hart>

[8] Helbing D., *Quantitative Sociodynamics*, Kluwer Academic, 1995.

[9] Helbing D., "Dynamic Decision Behavior and Optimal Guidance through Information Services: Model and Experiments", (memo), 2003.

[10] Hofbauer, J., Sigmund, K., Evolutionary Games and Population Dynamics, Cambridge Univ. Press, 1998.

[11] Huberman B., Glance N., "Diversity and Collective Action", Interdisciplinary Approaches to Nonlinear Systems, Springer, 1993.

[12] Iwanaga S. and Namatame A., "Asymmetric Coordination of Heterogeneous Agents", *IEICE Trans. on Information and Systems*, vol. E84-D, pp.937-944, 2001.

[13] Iwanaga S. and Namatame. A., "The Complexity of Collective Decision", *Journal of Nonlinear Dynamics and Control*, vol.6, no.2, pp.137-158, 2002.

[14] Namatame,A, Sato, H, and Murakami,K," Co-evolutionary Learning in Competitive Environments", in *Recent Advances in Simulated Evolution and Learning*, World Scientific, 2004.

[15] Schweitzer, F. (ed), *Modeling Complexity in Economic and Social Systems*, World Scientific, 2002.
[16] Shoham,Y and Powers,R,” Multi-Agent Reinforcement Learning:a critical survey”(memo),2003
[17] Tennenholtz, ,”Effcient learning equilibrium”. In Proceedings of NIPS, 2002.
[18] Tumer, K and Wolpert, D.(Eds), *Collectives and the Design of Complex Sytems*, Springer, 2004.
[19] Weibull J. *Evolutionary Game Theory*, The MIT Press, 1996.