

Co-evolution of Sustainable Coupling Rules

Naoto Sato, Akira Namatame

Dept. of Computer Science,

National Defense Academy, Yokosuka, JAPAN

Phone: +81 468 41-3810 (ext. 2432) Fax: +81 468 44-5911

Email: {g42041, nama}@nda.ac.jp

Abstract

In this paper, we show how self-interested agents collectively learn “correlated rules” that sustain both efficient and equitable dynamic orders. The conventional zero-sum RSP game is generalized as non zero-sum games by modifying payoff matrix. We formulate a collection of agents in a lattice model repeatedly play the generalized rock-scissor-paper (RSP) games. Agents are modeled to play with neighbors by applying interaction rules. Those interaction rules are evolved by crossing over with the most successful neighbor. If the payoff for the winning the game increases, agents learn to win and lose in a coordinated way: they repeat these coordinated plays in order to realize the most efficient and equitable situation. We also investigate learning rules of agents. We show interacting agents co-evolve to learn sharing the same types of rule, which sustain the efficient and equitable society.

Keyword: *rock-scissor-paper game, co-evolution, local interaction, efficiency, equity*

1. Introduction

One of the central aims of ecology is to identify mechanisms that maintain biodiversity. Numerous theoretical models have shown that competing species can coexist if ecological processes such as dispersal, movement, and interaction occur over small spatial scales. In particular, this may be the case for non-transitive communities, that is, those without strict competitive hierarchies. The classic non-transitive system involves a community of three competing species satisfying a relationship similar to the children's game rock-paper-scissors, where rock crushes scissors, scissors cuts paper, and paper covers rock. Such relationships have been demonstrated in several natural systems. Some models predict that local interaction and dispersal are sufficient to ensure coexistence of all three species in such a community, whereas diversity is lost when ecological processes occur over larger scales.

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), or 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. And in case you think of it as a rather simple-minded pastime, you should take a look at the home page of the ‘World RPS Society’, which is a treat.

2. The Modified Rock -Scissor-Paper (RSP) Games

We generalize RSP games with the payoff matrix given in Table 1. If we set $\alpha = 2$, this game become equivalent with the original zero-sum game. The game with the payoff matrix in Table 1 has a Nash equilibrium. Nash equilibrium strategy is to select rock, scissor, and paper with the probability $1/3$. The expected payoff of each agent with this mixed Nash strategy obtained as $(\alpha + 1)/3$.

If the parameter I is not one, their payoff at equilibrium is asymmetric and the problem of the fairness may also arise. With this game, we are especially interested in the effect of changing the parameter value. The most preferable situation is that each agent adopts the different strategy, and we define they cooperate if they choose different strategies in this case.

opponent own	S1 (Rock)	S2 (Scissor)	S3 (Paper)
S1 (Rock)	1	0	0
S2 (Scissor)	0	1	0
S3 (Paper)	0	0	1

$(I \geq 2)$

Table 1. The payoff matrix of modified rock-scissor-paper game

3. Repeated Games with Interaction Rules

An important aspect of social evolution is the learning strategy adapted by individuals. In most game theoretic models, agents calculate their best strategy based on information about what other agents have done in the past. Then agents gradually learn the equilibrium strategy. A number of evolutionary models based on the iterated general non-cooperation games have been proposed. Many dynamical systems and evolutionary models have been constructed with the Prisoner's Dilemma game as a model for the interaction between individuals. Axelrod applied a genetic algorithm (GA) to the iterated Prisoner's Dilemma and used a bit-string representation of finite memory strategies.

We use the different approach suggested by [10]. Each strategy in the repeated game is represented as a ternary 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. As Fig.1 shows, there are nine possible outcomes for each move ((S1S1), (S1S2), (S1S3), (S2S1), (S2S2), (S2S3), (S3S1), (S3S2), (S3S3)). 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 = S1, 1 = S2, and 2 = S3 then each strategy can be defined by a 3-bits string.

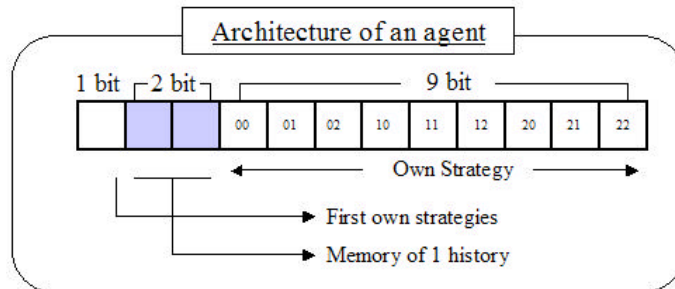


Fig. 1. The agent's memory architecture (1-history)

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 crosses with the strategy of the most successful neighbor. Partial mimicry which partially mimics the strategy of an opponent is known as cross-over in a genetic algorithm. Each agent is modeled to be matched several times with the 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 [10].

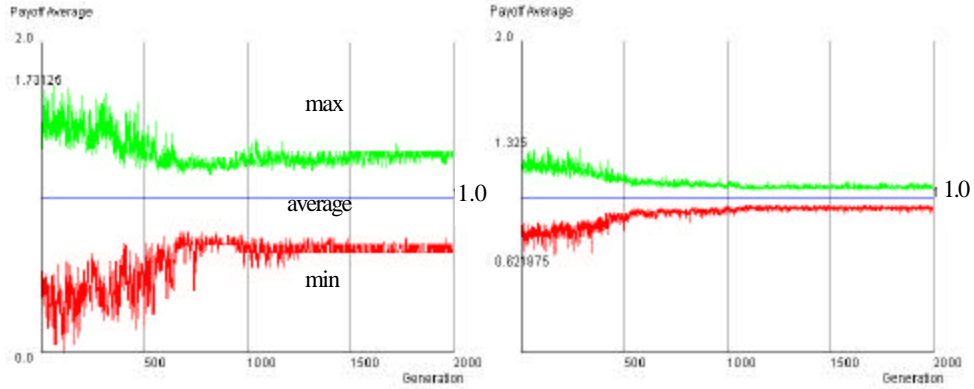
We also consider the implementation error specified by the rule and exists small probability of choosing the wrong strategy. We showed the simulation results without any errors and with ten percent of the error rate as noise. Significant differences will be observed when agents have small chances of making mistakes.

4. 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, his strategy can be imitated by others (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 [12]. 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 local interaction is modeled as games as shown in Table 1 Number of agents are 400 ($N=400$). At each time period, each agent plays the game in Table 1 with his eight neighbors and they decide next strategy by referring 1-history or 2-history. At next time period, each agent cross-over with the strategy of the most successful neighbor who obtain the highest payoff.

We simulated several cases by changing the parameter α . Fig. 2, 3, and 4 shows transition of payoff average to generation in cases when we set $\alpha = 2, 5, \text{ or } 10$, 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.



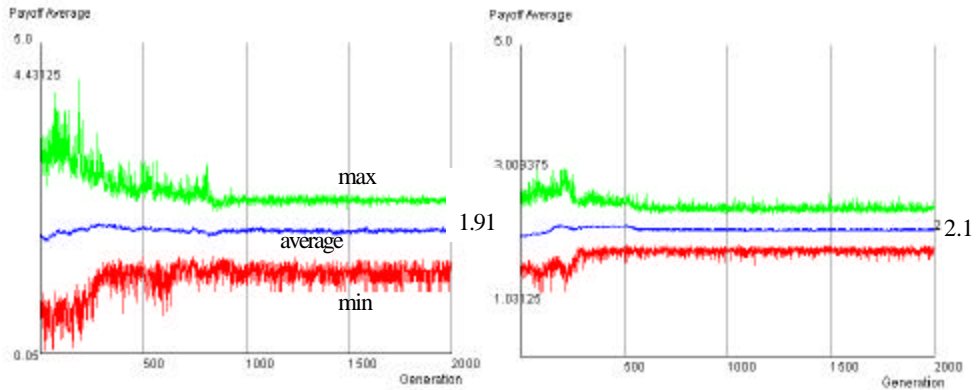
The average payoff of each generation when the winning payoff level

(a) error rate: 0%

The average payoff of each generation when the winning payoff level

(b) error rate: 10%

Fig. 2. Simulation result with $\alpha = 2$ in Table 1



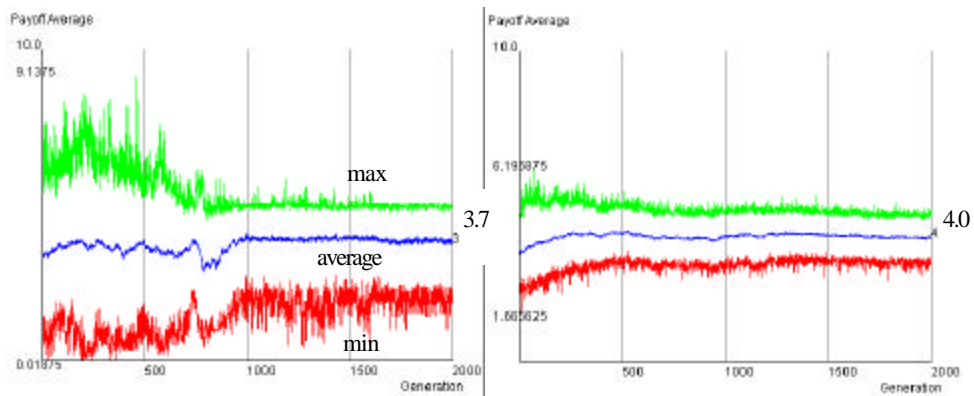
The average payoff of each generation when the winning payoff level

(a) error rate: 0%

The average payoff of each generation when the winning payoff level

(b) error rate: 10%

Fig. 3. Simulation result with $\alpha = 5$ in Table 1



The average payoff of each generation when the winning payoff level

(a) error rate: 0%

The average payoff of each generation when the winning payoff level

(b) error rate: 10%

Fig. 4. Simulation result with $\alpha = 10$ in Table 1

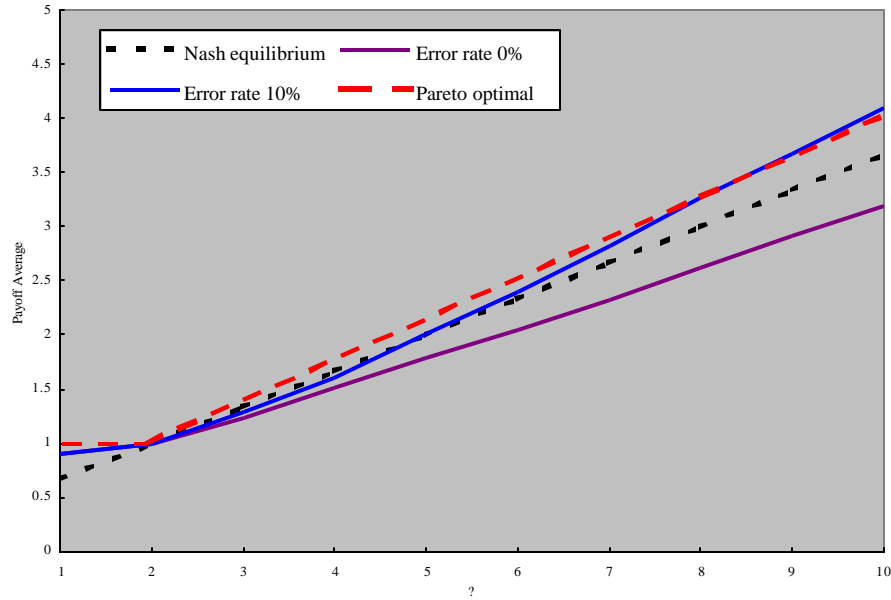


Fig 5. Progress of average payoff with changing the value of

At Fig.2(a), and Fig.2(b), all agents' payoff average is the same value 1. However, as payoff increases >2 , the cooperative relationship collapses when the gain increases, and the value of the average payoff decreases. 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. Therefore, payoff averages rise, on the other hand, social efficiency rises.

Fig.5 shows the progress of average payoff. In $=2$ the both value of Nash equilibrium and Pareto optimal are the same, and the simulation results achieved the value without no error or error 10%. As got higher, the simulation results showed difference between average payoff with error and it without error. Implementation errors became more effective. While the average payoff would not achieve Nash equilibrium without error, the average payoff could achieve Pareto optimal with error.

5. What Agents Co-evolved?

Next, we investigate the effect of implementation error. We show difference of the number of the same type agent who has the same strategy in his memory a certain generation in Fig.6 It shows that agents' memory is converged as four types that have common parts in 2000th generation, but there are about 400 types in primitive generation. According to this result, noise will work convergence of agents' memory.

Generation	Number of different rules	
	Error rate: 0%	Error rate: 10%
500	400	400
1000	400	250
1500	368	30
2000	238	4

Fig.6 The number of different learning rules among 400 agents

We analyze details of such a converged strategy decision rules by state diagrams. This simulation results in Fig.7 is the case of $\alpha=10$, error rate 10%, 2000th generation, and 1-history. 400 strategy decision rules in first generation are converged 4 type rules. Each rule has the same bits in common place without only 3 places. These 3 places are decision bits of the case that agents adopt the same strategy each other. Therefore, this means that agents adopt different strategy if agents adopt the same strategy each other. Each rule is called as rule type 1, 2, 3, or 4, and the number of agents who have type 1 is 202, type 2 is 114, type 3 is 58, and type 4 is 26.

Initial_Strategy	00	01	02	10	11	12	20	21	22	
Rule Type 1:	1	2	2	0	2	1	1	0	2	(202)
Rule Type 2:	1	1	2	2	0	2	1	1	0	(114)
Rule Type 3:	1	2	2	2	0	0	1	1	0	(58)
Rule Type 4:	1	1	2	2	0	0	1	1	0	(26)

Fig. 7 Converged strategy decision rules ($\alpha=10$, error rate 10%)
(): Some commonality of learning rules

6. The Effect of Co-existence of Heterogeneous Agents with Different Rules

(1) Homogeneous agents

We investigate a state diagram that an agent of rule type 1 plays with an agent who has the same rule in Fig. 8. Fig. 8 has broadly two cycles. Left cycle 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. All 400 agents have adopted S2 strategy at first of 2000th generation. Therefore, all agents rule 1 versus rule 1 start at state 11, then they go into state 22 and do not acquire payoff efficiently by acquiring 1 forever. However, if an implementation error happens in state 22, for example, an agent may go to state 02. Of course, he may come back to 02 from 22 in inefficient cycle.

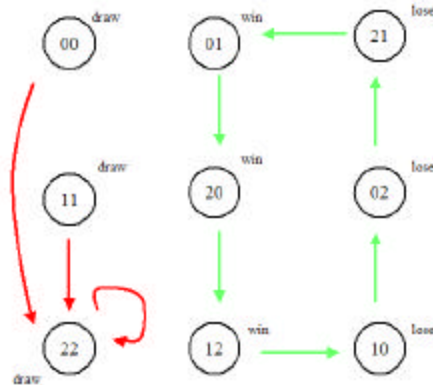


Fig. 8 A state diagram of plays among agents with the same rule (rule type 1 versus rule 1)

The truth is that 2 cycles (one is a draw cycle, another is a winning 3 times and losing 3 times cycle) appear as a state diagram if the same agent type play, for example type 2 versus type 2, and type 3 versus type 3. This state diagram means that agents can not acquire the payoff efficiently by staying draw state if agents with the same rule play without implementation errors from initial state. However, agents escape from the inefficient cycle as they would mistake to adopt strategies with probability p , and then they stay in the efficient cycle.

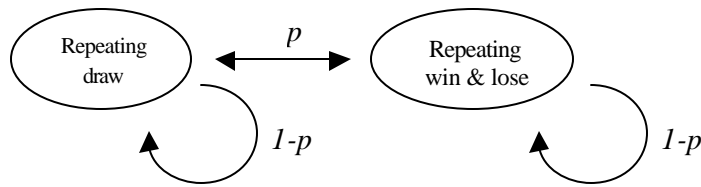


Fig. 9 Interaction with the same type of agents

(2) Heterogeneous agents

Next, we investigate state diagrams of plays among agents with the different rules in Fig. 10. Fig. 10 shows a state diagram of type 1 versus type 3. As all 400 agents have adopted S2 strategy at first the 2000 generation, starting state is 10. Fig. 10 is different from previous state diagram in the point that agents can move to efficient cycle without implementation errors.



Fig. 10 An example of state diagram of agents with the same rule (rule type 1 versus rule type 3)

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.

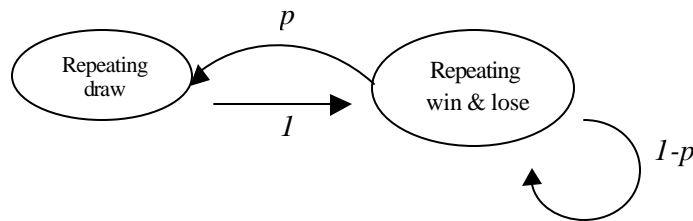


Fig. 11 Interaction with the different type of agents

7. Conclusion

We analyzed the competitive interactions in a finite population of agents in which agents are repeatedly matched within a period to play a stage game. We only imposed a weak monotonous condition reflecting the inertia and myopia hypotheses on the dynamics, which

describe the inter-temporal changes in the number of agents playing each strategy. The hypotheses we employed here reflect limited ability (on the agent's part) to receive, decide, and act upon information they get in the course of interactions.

We examined how efficiency and the fairness of the society changed when the gain that was able to be acquired increased. Moreover, what influence the act of making a mistake exerted on them was verified. The mistakes of deciding strategy in the society bring an important result, besides there is a possibility that the influence not intended conceals oneself.

We examined how the agents achieved efficient equilibrium and they have acquired strategy decision rules by learning and mistake. If agents mistook, agents acquired ideal and efficient cycle of strategic decision. Depending on first strategies, agents start at inefficient cycle but the mistakes lead agents to efficient one. We cannot pass that mistakes re-lead them inefficient one from efficient one.

References

- [1]Axelrod, R., *The Complexity of Cooperation*, Princeton Univ. Press, 1997.
- [2]Fogel, B. and Chellapia, K., *Inductive Reasoning and Bounded Rationality Reconsidered*, IEEE Trans. of Evolutionary Computation 3: pp.142-146, 1999.
- [3]Fudenberg, D. and Levine, D., *The Theory of Learning in Games*, The MIT Press, 2000.
- [4]Hansaryi, J. and Selten, R., *A Game Theory of Equilibrium Selection in Games*, MIT Press, 1988.
- [5]Namatame, A and Sato, N, *Co-evolutionary Learning in Modified Rock-Scissor -Paper Games*, Workshop on Economics and Heterogeneous Interacting Agents, Kyoto, Japan, 2004.
- [6]Namatame, A., Sato, N., and Murakami, K., *Co-evolutionary learning in Strategic Environments*, Advances in Natural Computation, ed Chen, T. K. to appear in 2004 World Scientific.
- [7]Samuelson, L., *Evolutionary Games and Equilibrium Selection*, The MIT Press, 1998.
- [8]Sipper, M., *Evolution of Parallel Cellular Machines: The Cellular Programming Approach*, Springer, 1998.
- [9]Uno, K. and Namatame, A., *An Evolutionary Design of the Networks of Mutual Reliability*, CEC'99 Workshop on Agent systems and Co-evolution, Washington D.C., pp.1717-1723, 1999.
- [10]Uno, K. and Namatame, A., *Evolutionary Behaviors Emerged through Strategic Interactions in the Large*, GECCO'99 Workshop on Artificial life, Adaptive behavior and Agents, Orlando, Florida, pp.1414-1421, 1999.
- [11]Weibull, J., *Evolutionary Game Theory*, The MIT Press, 1996.
- [12]Yao, X. and Darwen, P., *The experimental study of N-player iterated prisoner's dilemma*, Informatica, Vol.18: pp.435-450, 1994.