Co-evolution in Hawk-Dove Games in a Mobile Environment

Yukikazu Murakami, Hiroshi Sato, Akira Namatame
Dept. of Computer Science,
National Defense Academy
Yokosuka, 239-8686, JAPAN
E-mail: {g40046, hsato, nama}@nda.ac.jp

Abstract

This paper is about social evolution ,and shows how the society as a whole co-evolves when individuals composing it learn from each other. We consider the situation where a society of agents is repeatedly matched to play hawk-dove games. Agents are completely naive and do not perform optimization calculations. Instead they observe the current performance of other agents, and cross-over with the most successful strategy of his neighbors.

We also consider co-evolution in a mobile society. It is known that the only mixed strategy will result in equilibrium in the hawk-dove games and both hawks and doves coexist in the long-rum. In this paper, we consider evolutionary dynamics in a mobile environment where each agent moves randomly in a network. We show that most agents learn to behave as doves which realizes a social efficiency.

Keyword: *co-evolution, evolutionary games, hawk-dove games, replicator dynamics, mobile society*

1. Introduction

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 technique of modeling the evolutionary process as a system of a deterministic difference or differential equations may tell us little about equilibrium concepts other than that strict Nash equilibrium are good. We can attempt to probe deeper into these issues by modeling the choices made by the agents with their own internal models.

While the concept and techniques of game theory have been used extensively in many diverse contexts, game theory has been unsuccessful in explaining on how agents realize the mixed strategy if a game has the unique equilibrium of the mixed strategy [4]. Introspective or educative theories that attempt to explain equilibrium play directly at the individual decision-making level impose very strong informational assumptions. As a consequence, attention has shifted to evolutionary explanations motivated by the work of biologists in evolutionary game theory [9]. Two features of this approach distinguish it from the introspective approach. First, agents are not assumed to be so rational or knowledgeable as to correctly guess or anticipate the other agent's strategies. Second, an explicit dynamic process is specified describing how agents adapt their strategies over time as they repeat the games many times. The standard interpretation of game theory is that the game is played exactly once between fully rational individuals who know all details of the game, including each other's preferences over outcomes. Evolutionary game theory, instead, assumes that the game is repeated many times by individuals who are randomly drawn from large populations. An evolutionary selection process operates over time on the population distribution of behaviors [3][5]. An interesting problem, which has been widely investigated, is under what circumstances will agents, by learning, converge to some particular equilibrium? We want to endow our agents with some simple way of learning and describe the evolutionary dynamics that magnifies tendencies toward better situation. By incorporating a consideration of how agents interact into models we not only make them more realistic but we also enrich the types of aggregate behavior that can emerge.

An important aspect of social evolution is the learning strategy adapted by individuals [2]. Evolution in the hawk-dove game drives the population to an equilibrium polymorphism state. But this symmetrical mixed equilibrium of hawk-dove is so inefficient that is far from optimal. How do we interpret these results?

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 mechanis ms. In this paper, we focus on evolutionary dynamics that may change in time according to certain local rules of individuals. Evolutionary models can be characterized both by the level at which the mechanisms are working and the dimensionality of the system.

We use the evolutionary models based on microscopic individuals who interact locally[1][7]. We provide a general class of adaptation models and relate their asymptotic behavior to new equilibrium concepts. We assume agents behave myopically, and they evolve their behavior over times. They learn from the most successful strategy of their neighbor. Hence their success depends in large part on how well they do in their interactions with their neighbors. If the neighbor is doing well, the behavior of the neighbor can be imitated, and in this way successful strategies can spread throughout a population, from neighbor to neighbor [10].

2. A Formulation of Hawk-Dove Games

A long-standing theme of ethnology is the prevalence of conventional fights. Conflicts animals are often settled by displays (dove) rather than all-out fighting (hawk). A whole gamut of threatening signals and harmless assessments of strength serves to settle contests for food, territory and mates, so that escaped fights leading to injury or death sure relatively rare. Such conventional fights have been observed and the restrained nature of animal aggression has often been stressed. It is obviously all to the good of the species, but need an explanation from an evolutionary point of view for why do escaped contests remain rare. The biologist J. Smith invented the hawk-dove games to explain this [9]. He showed that the proportion of animals who behave as hawks at equilibrium is given as the function of the prize of winning and the cost of the fighting, and if the loss from the fighting increases, the proportion of animals behaving as hawks decreases.

The hawk-dove games are formulated with the payoff matrix in Table 1. Let suppose there are two possible behavioral types: one escalates the conflict until injury or sticks to display and retreats if the opponent escalates. These two types of behavior are described as "hawk" and "dove". The prize corresponds to a gain in fitness V, while an injury reduces fitness by C, and we assume that $C > V$. If a hawk meets a hawk, they fight until one is seriously injured. The fitness of the winner increased by V, that of the loser reduced by C, so that the average increase in fitness is $(V-C)/2$, which is negative since the cost of the injury is assumed to exceed the prize of the fight. If a dove meets a dove, they engage in threatening display, but flee when confronted with real danger, and therefore, his expected fitness is given by $V/2$. It is well supported by observation that conflicts among animals especially within heavily armed species are often settled by displays rather than all-out fighting. If a hawk meets a dove, the dove runs away and the hawk wins the contested resource of value of V.

There is the unique symmetric Nash equilibrium in mixed strategies, both agents use the strategy S_1 ('hawk') with probability $q = V/C$ and the strategy S_2 ('dove') with the probability $1-q = 1-(V/C)$ [2]. Therefore if the cost of injury C is very large, the hawk frequency (V/C) will be small. At equilibrium of the mixed strategy, the expected fitness is given at the level of $(V/2)\{1-(V/C)\}$. If each agent chooses the strategy S_2 ('dove'), (however, the situation that both behave as doves are not equilibrium) he receives $V/2$. This implies that the mixed-strategy results in inefficient equilibrium.

When we consider the game in an evolutionary setting in order to investigate the long-run behavior of the population, the situation remains the same. The dynamic evolutionary process with the assumption of uniform matching can be also analyzed by using replicator dynamics [11]. The equilibria (S_1, S_2) and (S_1, S_2), or more precisely, the state in which all agents play S_1 (Hawk) and the state in which all agents play S_2 (Dove), are both asymptotically unstable with respect to the replicator dynamics. In a population of almost all hawks, the dove strategy of avoiding conflict does better than the hawk strategy. Then doves increase their proportion of the population. Similarly in a population of almost all doves, however, the hawk strategy does better than the dove strategy. Then hawks increase their proportion of the population, and only the mixed equilibrium remains, and the evolutionary dynamics drives the population to that equilibrium.

Own's strategy \ The other's strategy	S_1 (Hawk)	S_2 (Dove)
S_1 (Hawk)	(V-C)/2 / (V-C)/2	0 / V
S_2 (Dove)	V / 0	(V/2) / (V/2)

Table 1 The hawk-dove game $(0 < V < C)$

3. The Method lies of Interaction

Agents may interact with any other agents (global interaction) even though they may actually only do so with some randomly drawn sample. The approach of a random matching is modeled as follows: in each time period, every agent is assumed to match or interact with one agent drawn at random from a population [6]. With the assumption of random matching, each agent is modeled to match with one generic agent, and his decision become to choose an optimal strategy based on a sample of information about what others agents have done in the past. Agents are assumed to be able to calculate best replies and learn the strategy distribution of play in society. An important assumption of the random matching is that they receive knowledge of the current strategy distribution through random matching. Then agents gradually learn the strategy distribution in the society.

In many situations, agents may meet only with their neighbors (local interaction) Thus the consequences of their behaviors may take some time to have an effect on agents with whom they are not directly linked. Although the static equilibrium notion is worth examining, perhaps more interesting is the dynamic evolution. Agents are not assumed to be so rational or knowledgeable as to correctly guess or anticipate the other agent's strategies. They adapt other agents' successful strategies as guides for their own choices. Hence their success depends in large part on how well they do in their interactions with their neighbors. But neighbors can serve another function as well. If the neighbor is doing well, the behavior of the neighbor can be imitated. In this way successful strategies can spread throughout a population, from neighbor to neighbor.

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 spatiotemporal structures may allow for global optimality where the model with random matching would result in inefficiency. The presence of these various forms of spatiotemporal phenomena may also alter the evolutionary path compared with the global matching.

We consider two interaction model in the network formulated as a lattice as shown in Fig.1. We arranged agents for an area of 50×50 ($N = 2{,}500$ agents) of the lattice model shown in Fig. 1 with no gap, and four corners and end of an area connect it with an opposite side. At each time period t, each agent plays the hawk-dove game in Table 1 with his 8 neighbors. At the next time period, each agent cross-over with the strategy of the most successful neighbor who obtain the highest payoff. We also consider a mobile socially as show in Fig.2. On each time step, every individual on the lattice attempts to move at the same time (this is known as synchronous updating) and the movement is local in the sense that an individual can only move from hie current lattice site on location to an adjacent site and not to site a farther away. We want to avoid having individuals bumping into one another in the process of moving and since there is no traffic controller to coordinate the movements of people, two requirements are imposed on each individual that must be met in order for him to make a move, as opposed to remaining the adjacent site that the individual is attempting to move to must not already be occupied by another individual and the move must not cause a collision with another individual trying to move onto the same site.

We express the rules governing movement subject to these conditions as shown bellows in a set of rules, each having a simple form. The mobility of individuals a central future in many human activities.

<u>A movement rules of agents</u>

(1) Look out as far as permits in the four principal lattice directions and identify the unoccupied site(s)

(2) Move to this site;

(3) If they meet another agents, they play hawk-dove game with the payoff table1;

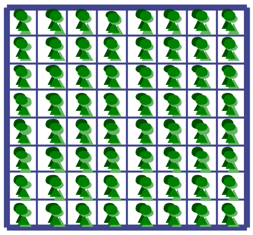


Fig.1 Hawk-Dove games in a fixed society

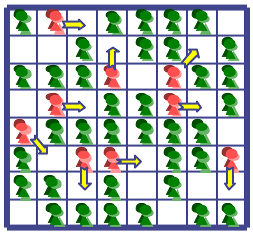


Fig.2 Hawk-Dove games in a mobile society

4. Learning Strategies of Agents

Here we assume that agents observe the current performance of their neighbors, and learn from the most successful agent. Agent are less sophisticated in that they do not know how to calculate best replies and are using other agent's successful strategies as guides for their own choices. Each agent interacts with the agents on all eight adjacent squares and imitates the strategy 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 converts to the strategy of the most successful of them, or cross-over with the strategy of the most successful neighbor.

An important aspect of social evolution is the learning strategy adapted by each individual. In the previous literatures, agents are viewed as being genetically coded with a strategy and selection pressure favors agents which are fitter, i.e., whose strategy yields a higher payoff against the population [3]. Each agent interacts with the agents on all eight adjacent squares and imitates the strategy 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 converts to the strategy of the most successful of them, or mutates with the strategy of the most successful neighbor.

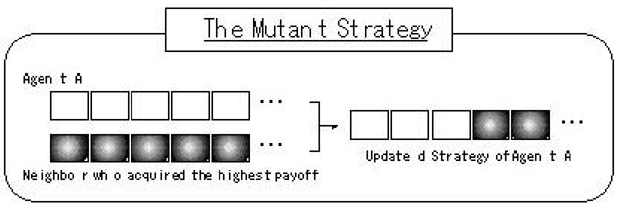


Fig.3 Illustration of learning strategies

Each agent's decision rules are represented by a population of N binary strings. At each generation gen ,gen ∈ [1…lastgen] agents repeatedly play the game for T iterations. An agent i, $i \in$ [1…N] uses a binary string i to make decision about his action at each iteration t ,$t \in$.[1…T]A binary string consists of 22 positions. Each position p_j , j∈ [1…22] represents as follows. The first and second position, p_1 and p_2, encodes the action that agent takes at iteration $t = 1$ and $t = 2$. A position p_j ,$j \in$ [3,6]encodes the memories that agent i takes at iteration $t - 1$ and $t - 2$ with his neighbor (opponent). A position p_j , j∈ [7,22] , encodes the action that agent i takes at iteration $t >$ 2, corresponding to the position p_j, $j \in$ [3,6]. An agent i compares the position p_j , $j \in$ [3,6] ,and decision table as shown in Table 2, 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:

An agent i represented with Table 2 behaves based on the strategy of TFT (Tit for Tat) [1]. This agent i at iteration t = 1 and t = 2, and then, the agent i mimic previous action of an opponent at t > 2.

Table 2: Strategy Table

	t=1, t=2		memory t-1 t-2				strategy decision part															
position Pj	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Agent i	H	H					H	D	H	D	H	D	H	D	H	D	H	D	H	D	H	D

The algorithm of the mutant strategy is implemented as follows:

(Step1) An agent i, i∈ , interacts with his neighbors at all eight adjacent squares as shown in Fig.3, and the game is repeated T iterations.

(Step2) An agent i chooses his action. Initial action, t = 1and t = 2, is determined by decoding bits in the position p_1 and. p_2 For t > 2, t ∈ [1…T], an agent i chooses his action at t by using the information about his memories (posit ion p_j , $j \in$ [3,6]). and strategy decision part (position p_j , j∈ [7,22]).

(Step3) The payoffs during T iterations of the games are accumulated.

(Step4) After T iterations, an agent i compares his payoff with his neighbors.

(Step5) An agent i updated his binary strings using the genetic operator: crossover with the neighbor who acquires the highest payoff (as shown in Fig.4(b)).

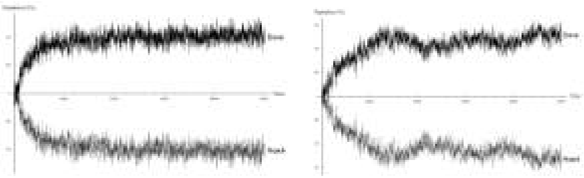
(Step6) Steps 1 through 5 are repeated for generations.

The operator of crossover operates on the pair of binary strings which is selected for an agent i and the neighbor who acquires the highest payoff. A number k, $k \in$ [1,2,7,…,22] , is randomly selected and two new binary strings are obtained by swapping the bit values to the right of the position k . Mutation randomly changes the value of a position within a chromosome. Each position has probability of being altered by mutation, if the bit of mutated position is H, this bit changed to D. And if the bit of mutated position is D, this bit changed to H. Given the updated set of binary strings, the algorithm proceeds to stage 1 for another generation of repeated games. Each simulation is conducted for a total of generations.

The algorithm for implementing the mimicry strategy is the same as the mutant strategy excluding the step 5. With the mimicry strategy, an agent i updated his binary strings by copying the whole binary strings of his neighbor who acquires the highest payoff at step 5.

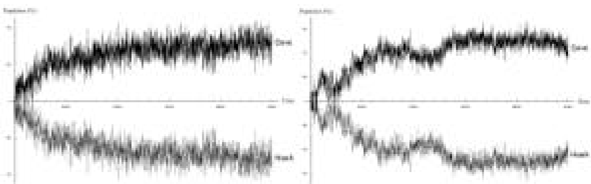
5. Simulation Results

We simulated several cases by changing the parameters V and C. In Fig.4, we show the cases when we set (a) (V, C)=(2,10), and then q = 0.2, and (b) (V,C)=(2,4), and then q = 0.5. The initial condition is set as follows: Only one agent at the center is set to choose $S2$ (dove) and the rest of the agents choose $S1$ (hawk). In that figure, the x-axis represents the update times of the strategies, and the y-axis represents the number of agents who choose $S1$ and $S2$ at each generation. The simulation results shown in Fig.4 indicate that the number of agents to choose $S2$ (dove) increases, and they gradually become to hate to fight each other, is close to efficient situation.



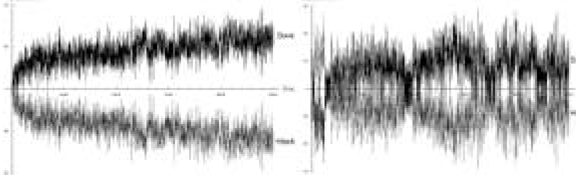
(a) A fixed society (b) A mobile society

Fig.4 The numbers of agents of choosing Hawks(S1) on Doves (S2) (V,C)=(2,10)



(a) A fixed society (b) A mobile society

Fig.5 The numbers of agents of choosing Hawk(S1) on Doves (S2) (V,C)=(2,4)



(a) A fixed society (b) A mobile society

Fig.6 The numbers of agents choosing Hawk(S1) and Dove (S2) (V,C)=(10,12)

In Fig. 6(a), we show the result when we set parameters (V, C)=(10,12), and then θ = 5/6 by increasing the benefit of winning the fighting. The initial condition is set as follows: the 75% of agents choose *S2* (dove) and the rest of the 25% of agents choose S_1 (hawk). The simulation result of this case indicates that all agents finally choose S_1 (hawk), and then all agents only receive -1 as the payoff. They loose the chance of receiving the payoff of 5 by avoiding the fight.

In Fig.6 (b), we showed the simulation result in a mobile society mimicry strategy, cross-over. In this case, each pair of agents interacts 20 times at one generation, and then they have many combinations of strategies *S1* and *S2*. The initial condition is set as follows: all agents randomly choose *S1* (hawk) or *S2*(dove) at the initial generation. The number of agents who choose *S2*(dove) increases over generation, and all agents finally choose *S2* (dove). This result indicates that they gradually co-evolve to hate of fighting, which is also the socially efficient situation.

6. 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 monotonically condition reflecting the inertia and myopia hypotheses on the dynamics, which describe the intertemporal 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 analyzed also about social learning and shown how the society as a whole learns even when the individuals composing it do not. Specifically, it is about the evolution of social norms. We especially examined how conventions evolve in a society that begins in an amorphous state where there is no established custom, and individuals rely on hearsay to determine what to do. With simulations, we also provided specific conditions as to which conventions are most likely to emerge.

We analyzed the role of learning in evolutionary games. The evolutionary model presented here is simple but presents a number of interesting phenomena in the form of evolutionary dynamics and spatiotemporal patterns, which cannot be observed by the previous models. Spatiotemporal phenomena may give rise to a stable coexistence between strategies that would otherwise be out competed. We also discussed the role of individual learning in social evolution. The hypotheses we employed here reflect limited ability of interaction with the individual learning capability. The learning strategy employed here is mimicry, and they mimic to the most successful neighbor. This methodology of learning is very slow to evolve, but it is essential especially in uncertain worlds. It is shown the co-evolution is essential for realizing social efficiency.

References

[1] Cohedent, P (eds): *The Economics of Networks*, Springer, 1998

[2] Fudenberg, D., and Levine, D., *The Theory of Learning in Games*, The MIT Pres, 1998

[3] Hammerstein, P. and Selten, R., "Game Theory and Evolutionary Biology", in *Handbook of Game Theory with Economic Applications*, Vol.2, Auman, R., Hart, S (Eds), Elsevier Science, pp.931-962, 1994.

[4] Hansaryi, J. and Selten, R., *A Game Theory of Equibrium Selection in Games*, MIT Press, 1988.

[5] Hofbauer, J. Sigmund, K., *Ecolutionary Games and Population Dynamics*, Cambridge Univ. Press, 1998.

[6] Kandori, M. and Mailath, G., "Learning, Mutation and Long Run Equilibria in Games", *Econometrica*, Vol.61, pp.29-56, 1993.

[7] Nowak, M. A., etc., "The Arithmetics of Mutual Help", *Scientific American*, June, 1995.

[8] Samuelson,,L., *Evolutionary Games and Equlibrium Selection*, The MIT Press, 1998.

[9] Smith ,J. M., *Evolution and the Theory of Games*, Cambridge University Press, 1982.

[10]Murakami.Y, Sato.H, and Namatame.A.,"Co-evolution in Negotiation Games "International Conference on Computational Intelligence and Multimedia Applications"

[11] Weibull J., *Evolutionary Game Theory*, The MIT Press, 1996.

[12] Lloyd A. L., "Computing Bouts of the Prisoner's Dilemma", *Scientific American*, June, 1995.

[13] Uno K. and Namatame A., "An Evolutionary Design of the Networks of Mutual Reliability", *CEC'99 Workshop on Agent systems and Coevolution*, Washington D.C., pp.1717-1723, 1999.

[14] 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.

[15] Uno K. and Namatame A., "N-Persons Prisoner's Dilemma (NIPD) Game with Bounded Rationality", *AJ'98 Workshop on Intelligent and Evolutionary Systems*, Kyoto, pp.83-90, 1998.

[16] Weibull J., *Evolutionary Game Theory*, The MIT Press, 1996.

[17] Yao X. and Darwen P., "The experimental study of N-player iterated prisoner's dilemma", *Informatica*, Vol.18:pp.435-450, 1994.