

©Copyright JASSS



[Yutaka Nakai and Masayoshi Muto \(2008\)](#)

Emergence and Collapse of Peace with Friend Selection Strategies

Journal of Artificial Societies and Social Simulation vol. 11, no. 3 6
<<http://jasss.soc.surrey.ac.uk/11/3/6.html>>

For information about citing this article, click [here](#)

Received: 15-Oct-2007 Accepted: 27-May-2008 Published: 30-Jun-2008



Abstract

A society consisting of agents who can freely choose to attack or not to attack others inevitably evolves into a battling society (a 'war of all against all'). We investigated whether strategies based on C. Schmitt's concept of the political, the distinction of a friend and an enemy, lead to the emergence and collapse of social order. Especially, we propose 'friend selection strategies' (FSSs), one of which we called the 'us-TFT' (tit for tat) strategy, which requires an agent to regard one who did not attack him or his 'friends' as a 'friend'. We carried out evolutionary simulations on an artificial society consisting of FSS agents. As a result, we found that the us-TFT results in a peaceful society with the emergence of an us-TFT community. In addition, we found that the collapse of a peaceful society is triggered by another FSS strategy called a 'coward'.

Keywords:

Community, Carl Schmitt, a Friend and an Enemy, Tit for Tat, Coward, Evolutionary Simulation

Introduction

1.1

How does social order emerge among individuals acting freely? This fundamental question in sociology was named the "problem of order" by Parsons ([1937](#)). For example, if people, as selfish individuals, can freely choose to attack or not to attack others, the society will inevitably change into a "battling society". (Hobbesian state: a "war of all against all") However, real society is not in this state and seems to have order. For another example, if people could freely choose whether to pay taxes or not, everyone would want to receive public services without paying taxes (so called "free riders"), and as a result, nobody would receive public services. However, in real society, public services are maintained by tax revenues. As is well known, a variety of solutions to this problem have been proposed. For example, we can find such as "central authority and social contract" ([Hobbes 1651](#)), "institutionalization and

internalization of norms" ([Parsons 1937](#)), "system rationality and reduction of complexity" ([Luhmann 1968](#)), "mutual understanding based on communicative reason" ([Habermas 1981](#)) and "trust based on rational expectation" ([Coleman 1990](#)).

1.2

In this paper, we want to pay attentions to Schmitt's concept of the political. He states that the concept of the political is theoretically and practically independent from various areas such as religion, culture and economy, and he regards the distinction of a friend and an enemy, which becomes clear particularly in a wartime, as the fundamental concept causing all political actions ([Schmitt 1932](#)). Then, out of our sociological interest, we will investigate whether the social perception of "who is a friend and who is an enemy" leads to social order from the emergence of community. In addition we will also take into deep consideration the mechanism of collapse, because most social theories do not focus on the collapse of social order more than they do on the emergence of community.

1.3

There is a variety of social orders. In game theory, the Hobbesian state mentioned above corresponds to the 2-persons Prisoner's Dilemma (2PD) and the problem of the free riders does to the N-persons Prisoner's Dilemma (NPD). In our study, we focus our attention on the 2PD problem. To discuss the 2PD problem, we must make clear the structures of actions. We are more interested in generalized exchange (indirect reciprocity), which means an one-way interaction between two persons, than we are in restricted exchange (direct reciprocity), which means reciprocal interaction between two persons. The latter corresponds to an economic exchange of money and goods, and the former to a social exchange of kindness with no expectation that kindness will be returned, for example.

1.4

As an experimental platform, we prepare an artificial society with generalized exchange under a 2PD situation. We examine whether or not social order emerges in the case that agents follow strategies on the basis of the code of "a friend and an enemy". Here, we should note that a social perception of "who is a friend and who is an enemy" can be seen as quite neutral. It is impossible to decide intuitively whether this perception has positive or negative effects on the emergence of social order. We thus examine this using computational simulations.

1.5

In most studies on the 2PD problem, an action means either cooperation (help) or defection (no help). In our study, an action means either an attack or no attack because we want to focus on the distinction of a friend and an enemy. (Of course, the change in interpretation does not have an impact on the game structure.)



Research with Ideal Type

2.1

Most social theories do not place importance on "exact explanations" of a complex society, but do place importance on "deep understandings" with an ideal type's model of a society (the "idealtypus" by Weber). The ideal type is required to be simple rather than real, and to shed light on essential aspects behind a complex and real society. This approach remains an important guideline in artificial society studies, especially in a social scientific field. For example, the KISS principle asserts that simplicity of a model makes understanding the mechanism behind phenomena easy. Of course, the ideal type is not always applicable to every field. In the area of economic study such as e-commerce, there is interest in real applications, and importance is placed on objective performance. Thus, the reality of a model is usually checked using quantitative index such as efficiency.

2.2

Our study is a theoretical research for the purpose of examining whether the distinction of a friend and an enemy is a fundamental principle leading to social order. So, it seems to be meaningless to construct the complicated model with assumptions unrelated to the concept of a friend and an enemy. In this sense, we should make our model as simple as possible. We are also particularly interested in whether the neutrality of the distinction of a friend and an enemy leads to order. We should thus make our model as neutral as possible.



Battle Game Paradigm

3.1

Our model consists of an artificial society with equal N agents making generalized exchanges under a 2PD situation and strategies for evaluating who is a "friend" and who is an "enemy", which we call "friend selection strategies" (FSSs). As mentioned above, the artificial society and strategies are introduced to be as simple and neutral as possible. First, to formalize our artificial society describing a battling society, we introduce the battle game paradigm, as follows:

- (B1) N agents play the game.
- (B2) Each agent has a social perception of who is a "friend" and who is an "enemy".
- (B3) In one battle game, each agent (agent A) meets M other agents at random. (M stands for the "matching number" of one agent in one game.) Each agent (agent A) interacts with the other agent (agent B) one by one. If M equals $N - 1$, the game is a round robin game. Otherwise, the game is a random matching game.
- (B4) Agent A (as a performer) attacks or doesn't attack agent B (as a performed) on the basis of A 's social perception of B . A attacks B if B is A 's "enemy", and A doesn't attack B if B is A 's "friend".
- (B5) If A attacks B , A obtains a payoff of 0.5 and B loses a payoff of 1.0. Conversely, if A doesn't attack B , both obtain and lose nothing (Figure 1).
- (B6) In one battle game, there are $N \times M$ of these interactions.
-

3.2

A typical instance of the battle game is a burglary. The performer obtains a payoff of 0.5 by stealing the performed's property, while the performed loses the same amount. The performed also suffers a loss of 0.5 due to physical and/or mental damages.

3.3

To examine the ending state of the battle game, we assume two agents who interact reciprocally. The payoffs in Figure 1 result in the two-agent payoff matrix shown in Figure 2, which is just a typical payoff matrix for the 2PD problem. This means that a battle game society has to fall into a battling state.

3.4

It should be noted that the game's definition is incomplete—it doesn't include how to determine who is a "friend" and who is an "enemy". Without this piece of the definition, an agent cannot interact with anybody at all. In the next section, we complete the definition by introducing new strategies called "friend selection strategies" (FSSs), one of which can change a battling society into a peaceful society.

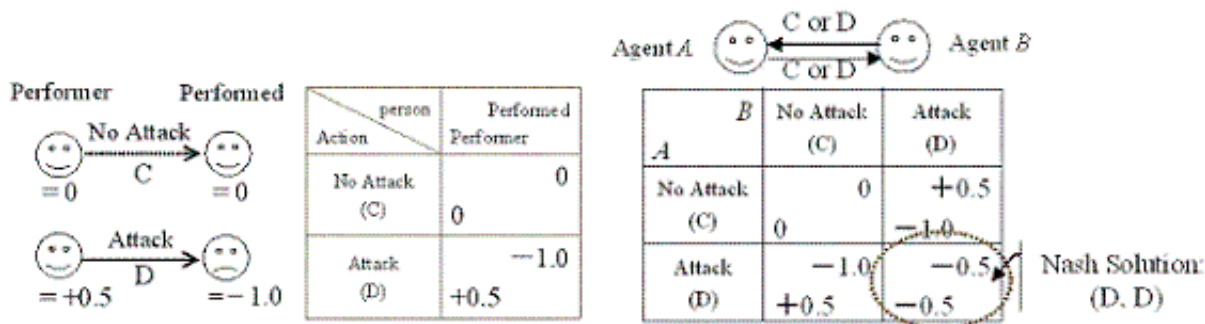


Figure 1 (left) Performer's and Performed's Payoffs. **Figure 2** (right) Battle Game's Payoff Matrix.

Friend Selection Strategies

4.1

As new strategies for determining who is a "friend" and who is an "enemy", it seems reasonable that an agent evaluates others on the basis of their direct actions toward the agent himself ("me"). This means that he doesn't need to know all actions of every other agent; he needs to know only actions taken directly toward him. We call these strategies "my-experience-based friend selection strategies" (MFSSs). We can expand the notion of MFSSs. It also seems reasonable that an agent can evaluate others on the basis of their actions toward the agent himself and his "friends" ("us" in place of "me"). The agent needs to know only actions taken toward his "us". We call these strategies "our-experience-based friend selection strategies" (OFSSs). Here, we define OFSSs more specifically. The agent needs to determine whether the other who attacked "us" should be regarded as a "friend" or an "enemy" and whether the other who did not attack "us" should be regarded as a "friend" or an "enemy". Thus, OFSS consists of 2×2 theoretical versions of strategies:

- (O1) us-ALL_D: the agent regards anybody who interacted with his "us" as an "enemy".
- (O2) us-TFT: the agent regards anybody who attacked his "us" as an "enemy" and anybody who did not as a "friend" (Figure 3).
- (O3) us-CWD: the agent regards anybody who attacked his "us" as a "friend" and anybody who did not as an "enemy". (CWD stands for a "coward".)
- (O4) us-ALL_C: the agent regards anybody who interacted with his "us" as a "friend".
- (O5) If other agents did not interact with his "us", he doesn't change his perceptions of them.

4.2

The us-ALL_D, us-TFT and us-ALL_C strategies are obviously variants of the strategies in the Iterated Prisoner's Dilemma game: ALL_D, TFT and ALL_C. The us-CWD strategy corresponds to Nash's solution in the chicken game. To obtain the MFSS definitions more specifically, we simply need to replace the "us" in the OFSS definitions with "me". Clearly, the me-TFT seems to be the most natural strategy among MFSSs.

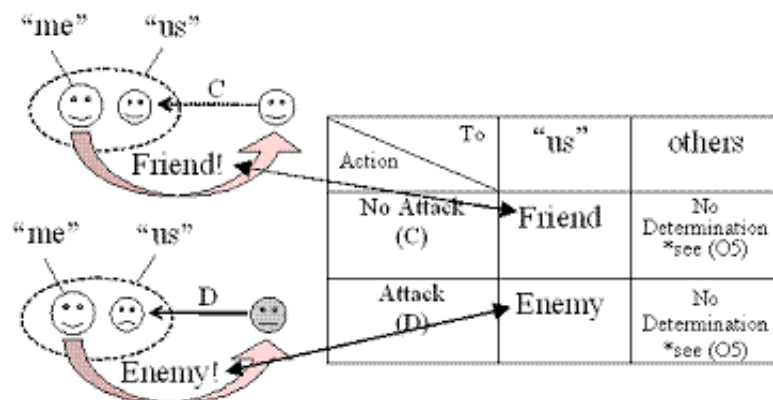


Figure 3. Concept of us-TFT Strategy

4.3

Here, it will be useful to consider more specific definitions of "friend", "enemy" and "us". In this study, the expression "friend" is defined as another agent who an agent decides arbitrarily not to attack. His decision is completely independent of the other's intention. In this meaning, the expression "friend" doesn't mean a close friend; it means simply his territory, which is set up without mutual agreement. Therefore, one agent's "us" doesn't necessarily correspond to another's "us". When one's "us" corresponds to another's "us", it means a friendly in-group, which we don't express by the term "us", but by the term us (without "").

4.4

Typical examples of people adopting FSS strategies are given in Table.1. First of all, the me-CWD requires an agent not to attack a militant agent against him and to attack a peaceful agent toward him. Figuratively, he obeys more powerful people than him and bullies less powerful people than him. Such a person reminds us of a chicken-hearted politician or a bully. An us-CWD agent obeys militant people against his "us" and bullies peaceful people toward his "us". And, his "friends" mean those who are more powerful than him. Such a person reminds us of a chicken-hearted bureaucrat or a hanger-on. He is located near powerful persons. He never fights with a stronger person, and he bullies a weaker person.

4.5

The me-ALL_D requires an agent to attack anyone interacting with him. Figuratively, he hates anyone who comes in contact with him, whether he is attacked or not. Such a person reminds us of a suspicious person. An us-ALL_D agent hates anyone who comes in contact with his territory ("us"). In particular, if a "friend" comes into contact with another "friend", the us-ALL_D agent comes to hate the former. Thus, he is an extremely suspicious person, like a lonely dictator.

4.6

The me-ALL_C requires an agent not to attack anyone interacting with him. Figuratively, he loves anyone who comes in contact with him, whether he is attacked or not. Such a person reminds us of a good natured person. An us-ALL_C agent loves anyone who comes into contact with his territory ("us"). If an "enemy" comes in contact with a "friend", the us-ALL_C agent comes to love the "enemy". He is thus an extremely good natured person, like a saint.

4.7

Of course, these examples do not completely correspond to the FSSs. Real social phenomena are too complicated to be described by only the FSSs. It should be remembered that the FSSs are ideal types. In order to make the FSSs simple and neutral, a variety of background factors, such as sense of value, power and role, are intentionally excluded, and only tangible elements, such as agents' actions, are included. The simplicity is expected to bring a "deep

understanding" of social phenomena in place of an "exact explanation".

Strategy	Description and Example People	Our Model
me-ALL_D	A person who hates anyone who contacts him ("me"). Ex. a suspicious person.	ALL_D
us-ALL_D	A person who hates anyone who contacts his "us" (his territory). When a "friend" contacts another "friend", the former changes into an "enemy". Ex. a lonely dictator.	
me-TFT	A person revenging for the attack against him. Ex. a lonely self defender.	me-TFT
us-TFT	A person revenging for the attack against his "us". Ex. a collective self defender.	us-TFT
me-CWD	Ex. A person obeying a militant person against him and submitting a peaceful person toward him. Ex. a chicken-hearted politician or a bully.	me-CWD
us-CWD	Ex. A person obeying a militant person against his "us" and submitting a peaceful person toward his "us". His "friends" are militant against him. Ex. a chicken-hearted bureaucrat or a hanger-on.	us-CWD
me-ALL_C	A person who loves anyone who contacts him ("me"). Ex. a good natured person.	ALL_C
us-ALL_C	A person who loves anyone who contacts his "us". When an "enemy" contacts any "friend", the former changes into a "friend". Ex. a saint.	

Table 1. Friend Selection Strategies (FSSs)

4.8

In sum, we have introduced eight FSS strategies. However, some may be too theoretical or a little bit unusual. For example, the us-ALL_C and us-ALL_D are rarely evident in real society. In addition, the large number of strategies complicates our analysis. Therefore, we integrated the me-ALL_C and us-ALL_C into the more common ALL_C, and the me-ALL_D and us-ALL_D into the ALL_D. We don't use the me-ALL_C and me-ALL_D to keep our study neutral between "me" and "us". In Section 7, we carry out agent-based simulations using the remaining six strategies. We also discuss whether a peaceful society can evolve from the FSSs and which strategy occupies a society. (Nakai and Muto (2005) previously conducted a study using all eight strategies [1].)



Related Studies

5.1

There are a lot of studies presenting solutions to the 2PD problem. These studies evaluate whether the other is good or bad on the basis of others' (third party) reputations about the other. They are sometimes called "reputation theories". They can be regarded as related to our study in the sense that our strategies also evaluate the other on the basis of others' reputations (friends' experiences). First, we can see typical related studies on economy, especially on e-commerce.

- (ER1) Janssen (2006) evaluated trustworthiness of the other using a variety of symbols in addition to the centralized (global) reputation.
- (ER2) Conte and Paolucci (2002) introduced two types of evaluations: one is "image" (own belief) and the other is "reputation" (socially circulating opinion).
- (ER3) Sabater, Paolucci and Conte (2006) used both image and reputation, and evaluated reputation using fuzzy functions.
- (ER4) Sabater and Sierra (2002) evaluated personalized (local) reputation on the basis of a variety of sources in a social network (witness, neighbourhood, system) and their credibility.

- (ER5) Hahn, Fley, Florian, Spresny and Fischer ([2007](#)) used personalized and centralized reputations (image, esteem, prestige) and evaluated reputation as a vector composed of trustworthiness, credibility and skill.
- (ER6) Ashri, Ramchurn, Sabater, Luck and Jennings ([2005](#)) evaluated personalized reputation on the basis of power relations (dependency, competition, cooperation).
- (ER7) Schlosser, Voss and Brückner ([2006](#)) evaluated centralized reputation on the basis of transaction values.

5.2

As we can see easily, all these economic studies deal with restricted exchange. Thus, our study, dealing with generalized exchange, essentially differs from these studies. Next, we will see non-economic studies, dealing with generalized exchange under a 2PD situation, as follows. They include common assumptions that an agent evaluates the other agent on the basis of own experiences and others' reputations about the other, which are neither a lie nor inaccurate information.

- (SG1) Younger ([2004](#)) introduced "communication", which means that two agents exchange their own reputations about another agent and update their own reputations by making an average operation of both reputations .
- (SG2) Castelfranchi, Conte and Paolucci ([1998](#)) introduced "communication", which means that two agents exchange their own reputations about another agent and update their own reputations by making Boolean OR operation of both reputations. Once a bad reputation has been established, changing a reputation is impossible.
- (SG3) Hales ([2002](#)) introduced "stereotyping", which means that a reputation about a member in a group is generalized to that of the whole group (a group reputation).

5.3

These studies include common assumptions unrelated to reputation. The artificial societies are made on a two dimensional lattice and have both economic and non-economic structures. Also, agents can move on the lattice and heterogeneous agents are introduced to the models, from a powerful agent to a powerless one. From our viewpoint of making a model simple and neutral, they are too complicated to be compared with our model.

5.4

In this sense, we consider the following studies as the previous studies for ours, because they can be regarded as having a simple model dealing with generalized exchange under a 2PD situation. They are "in-group altruistic strategy" ([Takagi 1996](#)), "imaging score strategy" ([Novak and Sigmund 1998a](#)), "standing strategy" ([Leimar and Hammerstein 2001](#)) and "strict discriminator strategy" ([Takahashi and Mashima 2003](#)).

5.5

These studies share the common strategy that an agent helps only agents who have helped a "good man". We can interpret the strategy as an altruistic strategy toward a selected good man, and so they are sometimes called "discriminator strategies" (DISC). In addition, they have the common characteristic that all agents with the strategy agree with the evaluation of good and bad. That is, FSSs give a political and relative evaluation like "my friend is your enemy", while DISC give a moral and absolute evaluation like "he is a good man for everyone". (The meaning of "good" in our study will be discussed at the end of the next section.)

5.6

Furthermore, DISC also share the common assumption that each agent knows every other agent's all actions (the "complete information assumption"), which is sometimes criticized as an unrealistic assumption. On the other hand, a FSS agent does not have to know all agents' all actions. He needs to know only actions toward himself and his "friends" ("incomplete information assumption"). With regard to this, we are also interested in whether social order

emerges even if we adopt the incomplete information assumption. In the next section, we examine in detail the similarities and differences between previous studies and ours.



Friend and Good Man

6.1

There are several similarities between previous studies and our study. First, the previous studies adopt a common platform called the "giving game" ([Takagi 1996](#); [Novak and Sigmund 1998b](#); [Leimar and Hammerstein 2001](#); [Takahashi and Mashima 2003](#)). The following is a typical definition of the game.

- (GG1) Two agents are assumed. One is a donor (Agent A), and the other is a recipient (Agent B).
- (GG2) The donor has a social perception about whether the recipient is a good man or a bad man.
- (GG3) The donor chooses whether or not to give to the recipient based on his own perception. That is, the donor gives to the recipient if the recipient is a good man and doesn't if the recipient is a bad man.
- (GG4) If the donor gives to the recipient, the recipient obtains a payoff of 0.5, and the donor loses a payoff of 1.0. Conversely, if the donor doesn't give to the recipient, both obtain and lose nothing.

6.2

If we suppose two agents interacting reciprocally, their payoff matrix turns out to be simply the matrix for the 2-persons Prisoners' Dilemma, the same as described in the analysis of the battle game. That is, the giving game leads inevitably to a selfish society. In addition, our "performer" corresponds to their "donor", our "performed" to their "recipient", and our "no attack" to their "help". Therefore, the giving game is similar to the battle game.

6.3

Second, to evaluate whether the other is good or bad, previous strategies consist of four rules, two of which are shared by them, as follows.

- (DS1) An agent regards anybody who helped a good man as a good man.
- (DS2) An agent regards anybody who did not help a good man as a bad man.

6.4

Concerning the remaining two, each strategy has different rules, as follows ([Takahashi and Mashima 2003](#)). In the "in-group altruistic strategy" ([Takagi 1996](#)):

- (IG1) An agent regards anybody who helped a bad man as a bad man.
- (IG2) An agent regards anybody who did not help a bad man as a good man.

In the "imaging score strategy" ([Novak and Sigmund 1998a](#)):

- (IS1) An agent regards anybody who helped a bad man as a good man.
- (IS2) An agent regards anybody who did not help a bad man as a bad man.

In the "standing strategy" ([Leimar and Hammerstein 2001](#)):

- (ST1) An agent regards anybody who helped a bad man as a good man.
- (ST2) An agent regards anybody who did not help a bad man as a good man.

In the "strict discriminator strategy" ([Takahashi and Mashima 2003](#)):

- (SD1) An agent regards anybody who helped a bad man as a bad man.
- (SD2) An agent regards anybody who did not help a bad man as a bad man.

6.5

As shown in (DS1) and (DS2), previous strategies require an agent to help only the other if the other helped a good man. So, they can be interpreted as altruistic strategies toward a selected "good man". Similarly, the me-TFT and us-TFT require an agent not to attack the other who did not attack him or his "friends". So, they can be seen as altruistic strategies toward a selected "friend". We can thus see the correspondence between a "good man" and a "friend".

6.6

There are also several differences between previous studies and our study. First, a key assumption in DISC is that each agent knows every other agent's all actions. In other words, a DISC agent is assumed to have complete information, which is sometimes criticized as an unrealistic assumption. The corresponding assumption in the FSSs is that each agent knows only actions taken toward him and his "friends". That is, a FSS agent is assumed to have incomplete information. Second, a DISC agent has the same social perception as others for the following reason. As mentioned earlier, in DISC, a good man is defined as one who helped a good man. However, this definition alone is insufficient for determining who is good because the latter "a good man" is not defined. We can improve the definition by defining a good man as one who helped someone who helped a good man. Unfortunately, this simply extends the problem further out. Each agent has to track the chain of others' actions into the past and, sooner or later, it is necessary to assign every agent an initial social perception about who is a good man and who is a bad man. In the previous studies, all agents are assumed to have the same initial social perception, i.e., all agents are good men. Furthermore, as mentioned above, a DISC agent knows every other agent's all actions. This means that each agent has the same experience as others. Therefore, all DISC agents continue to update their social perceptions on the basis of the common initial perception and their common experience^[2]. As a result, all DISC agents turn out to have a common social perception about who is good and who is bad^[3]. In contrast, each FSS agent is assumed to know only actions toward him and his "friends". That is, FSS agents have different social experiences, and each agent updates his social perception based on his own experience^[4]. Therefore, each agent has a unique social perception.

6.7

Let us now explain the meaning of "good" in our study. When DISC agents are seen as a group, as mentioned above, all group members have a common perception and take the same actions toward others, and so it is possible to interpret that a standard of actions works within the group. We call this aspect "good". So, the "good" is not just a rebel but it is a normative concept. On the other hand, the "friend" is not a normative concept at all because an agent's "friend" is not always other agent's "friend". Each agent can have a different perception and take a different action from others.



Evolutionary Simulation of Peace

7.1

To consider whether MFSSs and/or OFSSs can change a battling society, where all agents have no "friends", into a peaceful society, where all agents have no "enemies", we constructed an artificial society and carried out evolutionary simulations, as seen in Figure 4.

7.2

The society is composed of a number of agents. Each agent has his own strategy and social perception. Each simulation run is composed of a sequence of iterated turns, and each turn

consists of four phases: perception, action, selection, and mutation. In the perception phase, each agent updates his social perception on the basis of his own strategy with an occasional error called the "perception error". In the action phase, each agent plays the battle game in accordance with his updated perception. In the selection phase, agents whose results were inferior abandon their strategies and adopt that of agents whose results were superior. At the 0th turn, all agents are assumed to have the ALL_D strategy and to perceive that all other agents are "enemies", which expresses the state of "war of all against all". After many iterated turns, the superior strategies survive and the inferior ones fade away. Our interest is in finding out what strategy survives and whether it leads to a peaceful society.

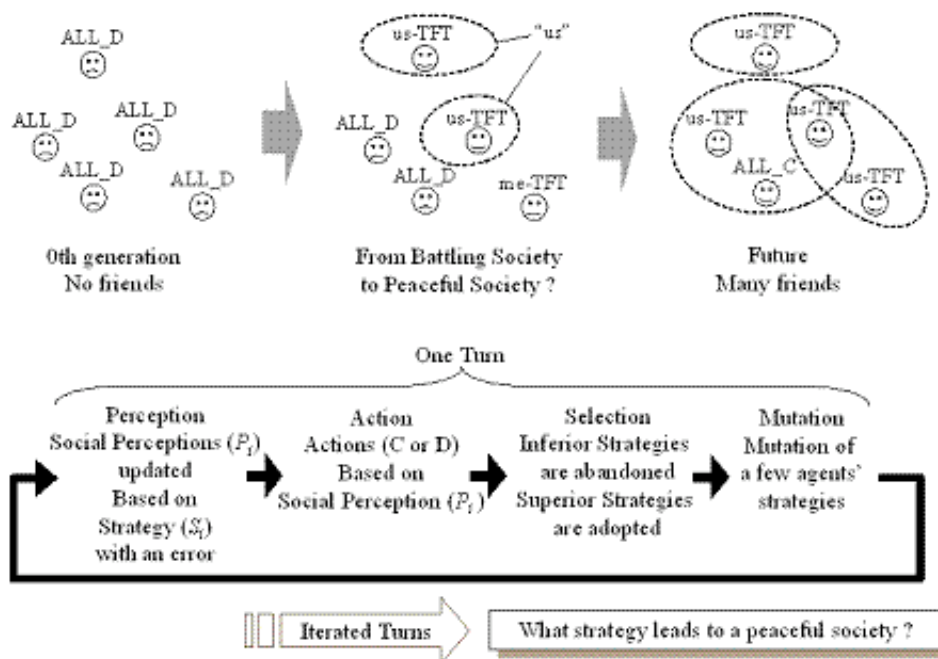


Figure 4. Evolutionary Simulation of Peace

7.3

The perception phase appears at the beginning of every turn. It is illustrated in Figure 5, and the algorithm is as follows.

- (PF1) At the T th turn, agent i has its own strategy S^T_i , which is one of the six strategies mentioned earlier.
- (PF2) Agent i has social perception P^T_i . It is described as a vector. If agent j is agent i 's "friend", element j of P^T_i is set to 1. Conversely, if agent j is agent i 's "enemy", element j of P^T_i is set to 0. Even if agent j is agent i 's "friend", agent i is not necessarily agent j 's "friend".
- (PF3) Each agent updates his social perception. Specifically, each agent updates his preceding perception P^{T-1}_i to the present one P^T_i , using his present strategy S^T_i and information on the $T-1$ th others' actions toward "me" or "us". The "friends" in "us" are those indicated by the $T-1$ th social perception P^{T-1}_i .
- (PF4) An ALL_C agent always identifies every other agent as a "friend".
- (PF5) A me-TFT agent identifies another agent who did not attack "me" as a "friend" and another agent who attacked "me" as an "enemy".
- (PF6) An us-TFT agent identifies another agent who did not attack "us" as a "friend" and another agent who attacked "us" as an "enemy". Specifically, an us-TFT agent assigns a +1.0 to a peaceful agent toward "us" and a -1.0 to a militant agent against "us". (See PF11 and PF12.)

- (PF7) A me-CWD agent identifies another agent who did not attack "me" as an "enemy" and another agent who attacked "me" as a "friend".
- (PF8) An us-CWD agent identifies another agent who did not attack "us" as an "enemy" and another agent who attacked "us" as a "friend". Specifically, an us-CWD agent assigns a -1.0 to a peaceful agent toward "us" and a $+1.0$ to a militant agent against "us".
- (PF9) An ALL_D agent always identifies every other agent as an "enemy".
- (PF10) If another agent did not directly interact with a me-TFT (CWD) agent, the latter agent retains his preceding perception of the former. Likewise, if another agent did not directly interact with an us-TFT (CWD) agent and his "friends", the latter agent retains his preceding perception of the former.
- (PF11) When an us-TFT or us-CWD agent evaluates another agent, he has to take into account all the other's actions toward "us". Specifically, he sums up all the scores assigned to that agent (PF6 and PF8) and identifies that agent as a "friend" if the total score is positive and as an "enemy" if it is negative. (See Figure 6.)
- (PF12) If the total score equals zero, the agent doesn't identify that agent. He retains his preceding perception of that agent.
- (PF13) An agent sometimes makes a mistake and updates his perception incorrectly. We call the mistake the "perception error". In the above updating, every bit in each perception vector is reversed by the "perception error rate" μ_p .
- (PF14) As the initial state of each simulation run, the 0th social perception P^0_i is set to a vector of zeros, meaning a battling state. The T th social perception P^T_i continues to be updated and is not reset at any turn.

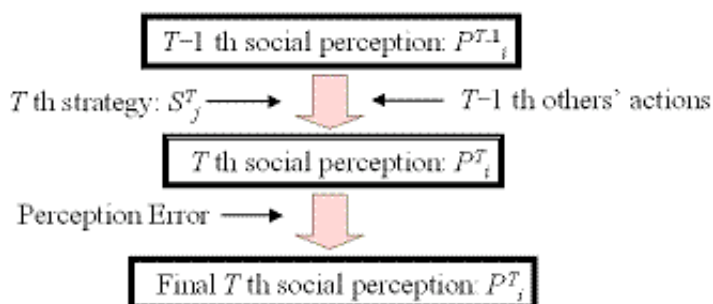


Figure 5. Process of Perception Phase

7.4

Let us explain PF11 in more detail. There are assumed to be four agents; i , f_1 , f_2 , and j . Agent i is an us-TFT agent, and f_1 and f_2 are i 's "friends". Suppose that agent j attacked f_1 and did not attack f_2 and i one turn ago. Does agent i identify j as a "friend" or as an "enemy"? The us-TFT agent i regards j as an "enemy" because of j 's attacking i 's "friend" f_1 , but regards j as a "friend" because of j 's not attacking i and i 's "friend" f_2 . Clearly, these identifications are contradictory. Therefore, we have to take all the identifications into consideration. In this example, given PF6, the corresponding scores would be -1.0 , 1.0 , and 1.0 . Since the total score equals 1.0 , agent i identifies j as a "friend" on the basis of PF11.

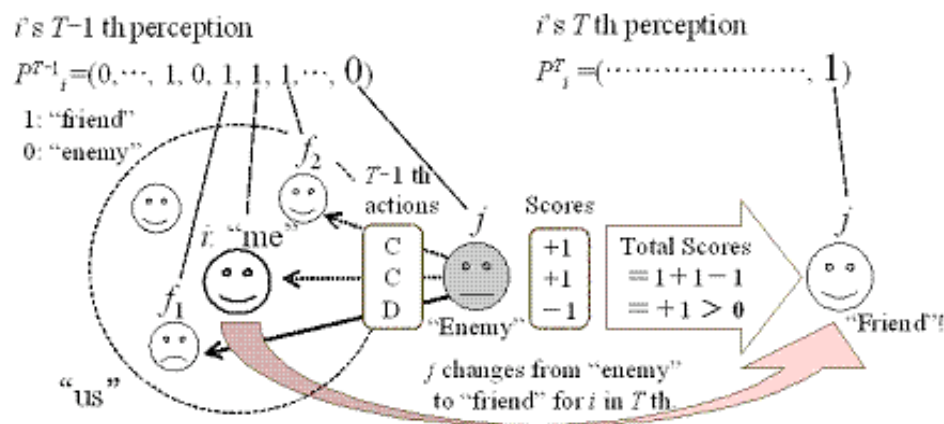


Figure 6. Us-TFT Agent's Perception: "Friend" or "Enemy"?

7.5

The action phase follows the perception phase. The algorithm is as follows.

- (AF1) Each agent interacts with M other agents randomly in every turn, and plays the battle game with each one.
- (AF2) Each agent decides how to act toward the other agent based on his updated social perception. That is, he doesn't attack a "friend", and he attacks an "enemy".
- (AF3) Both agents earn the payoffs shown in Figure 1.
- (AF4) After all agents finish playing one game, each agent calculates his payoff for that turn. That is, he sums up all the payoffs he earned as a performer and as a performed during the turn.
-

7.6

Following the action phase, there are agents with high payoffs and agents with low payoffs. In the subsequent selection phase, agents with low payoffs are assumed to be disappointed with their own failures and to replace their inferior strategies with superior ones. The algorithm is as follows.

- (SF1) The action phase results in agents with either high or low payoffs.
- (SF2) The lower $R\%$ agents in payoffs abandon their strategies. (We call R the "reflection ratio".)
- (SF3) These agents adopt the strategy of the agent with the highest payoff.

The final phase in a turn is mutation, where the strategies of a few agents spontaneously mutate.

- (MF1) A few agents are selected at random based on the "strategy's mutation rate" μ_S .
- (MF2) They abandon their strategies and adopt one of the six strategies at random.

Emergence of Peace due to us-TFT's Misunderstanding

8.1

We carried out two simulations to determine the effect of the notion of "us" on the emergence of a peaceful society. One simulates what society evolves from MFSSs (ALL_D, me-TFT, me-CWD, and ALL_C). The other simulates what society evolves from MFSSs and OFSSs (ALL_D, me-TFT, me-CWD, us-TFT, us-CWD, and ALL_C). The conditions for the two simulations are exactly the same, except for the strategies. The number of agents is 20 ($N=20$), and the perception error rate is 5% ($\mu_p = 5\%$). The number of agents each agent encountered in one

battle game (the matching number) is 19 ($M=19$), which means a round robin battle game corresponding to the highly frequent generalized exchange. The reflection ratio is 10% ($R=10\%$). The strategy's mutation rate is 0.3% ($\mu_s=0.3\%$).

8.2

Typical results are shown in Figure 7, which shows the friend ratio over turns. The friend ratio is calculated by dividing the mean number of "friends" of each agent by the total number of agents excluding the agent himself ($N-1$). Hence, a friend ratio of 1.0 corresponds to a peaceful society because all agents are friends mutually. The upper diagram shows the result of the MFSSs simulation and the lower shows that of the MFSSs and OFSSs simulation. Comparing the lower one with the upper one, we can find the following.

- (R1-1) With only the MFSSs, the friend ratio remains near zero for most of the turns, which means that the battling society continues.
- (R1-2) With both the MFSSs and OFSSs, the friend ratio reaches to 1.0 and remains there, which means that a peaceful society emerges and continues.

8.3

A peaceful society thus evolves from OFSSs but not from only MFSSs. This finding leads to the conclusion that the notion of "us" is essential to establishing social order. And, it is not necessary to know all actions of all agents; it is sufficient to know only actions taken toward "us".

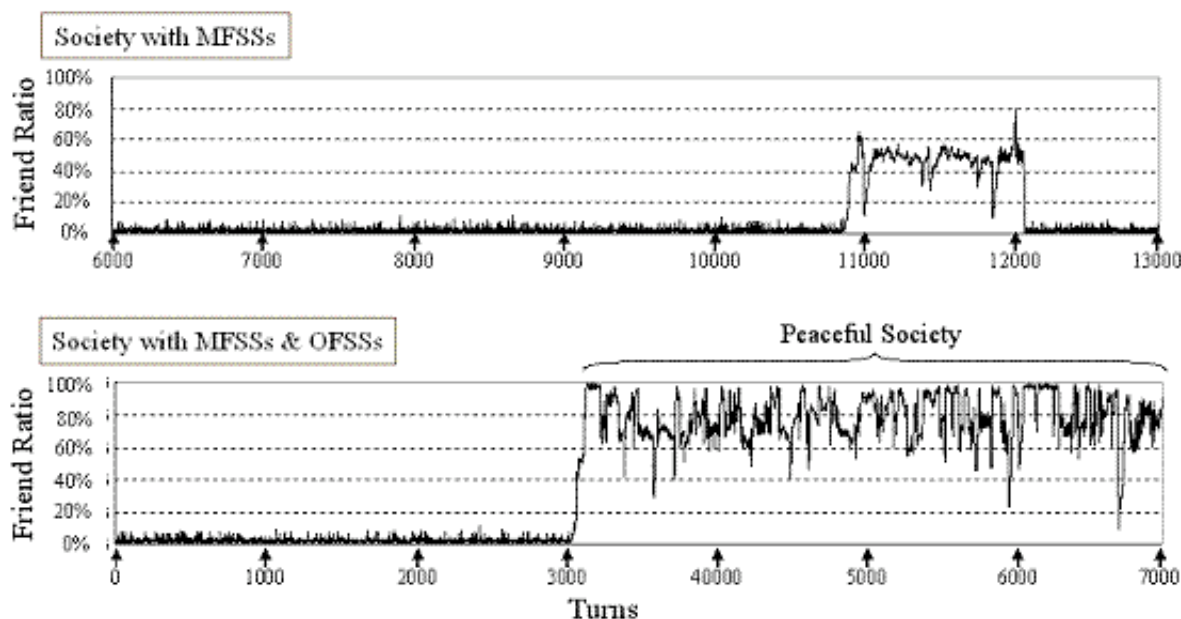


Figure 7. Friend Ratio vs. Turns (Upper: MFSSs, Lower: MFSSs and OFSSs)

*Number of Agents: $N=20$ agents, Matching Number of Agent in One Battle Game: $M=19$ agents, Reflection Ratio: $R=10$, Perception Error Rate: $\mu_p=5\%$, Strategy's Mutation Rate: $\mu_s=0.3\%$

8.4

The above parameters are one example leading to results, such as those shown in Figure 7, and so it is not fruitful to discuss why the parameters' values were selected. For example, consider the well-known "Self-Forming Neighborhood Model" by Schelling (1969). He examined what the distribution of races would be under the assumption that people make a move if the ratio of neighbours being the same race falls below the threshold (related to generosity). With this model, he found that clusters of races emerge in a local community. The reason why this study is regarded as meaningful is the fact that a very simple assumption

results in order. Because his model does not have to be so real, it isn't so important to examine why values of grid number or population density should be just the values and not other values. Since our model also does not have to be so real, it is not so fruitful to consider the reason for the parameters' values in our simulations. In the research using ideal type, we first have to make efforts to interpret simulation results sociologically and then verify whether only the parameters' values lead to the results or not. (Regarding the parametric survey, see Section 10.)

8.5

To investigate the emergence mechanism of a peaceful society, we observed the change in friendships among agents. In particular, we observed the network structure in a matrix form, as shown in Figure 8. Each row in the matrix represents the social perception of one agent. In other words, each row shows the actions which an agent takes toward others. If agent i regards agent j as a "friend", the element (i, j) is expressed as a white dot. For example, agent 1 regards agents 3, 7, and 8 as his "friends", as shown in the figure. Conversely, each column shows the actions which an agent receives from others. For example, agent 1 is regarded as a "friend" by agents 6 and 10. It should be noted that a square on a diagonal indicates the emergence of a friendly in-group, where each member sees all other members as "friends".

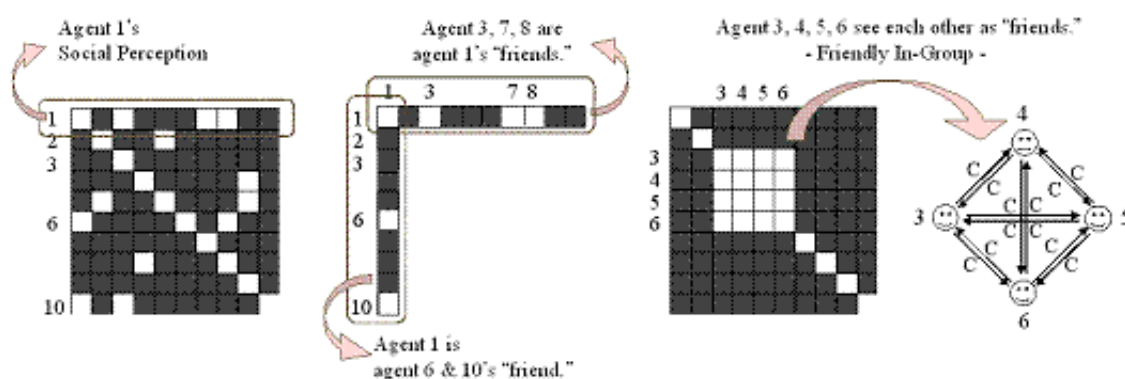


Figure 8. Observation of Relationships among Agents

8.6

The results are shown in Figure 9. It presents the change in friendships at around the 3000th turn (see the lower graph in Figure 7), when a peaceful society emerged. We see that the formation of a small us-TFT in-group triggered the emergence of peace as follows.

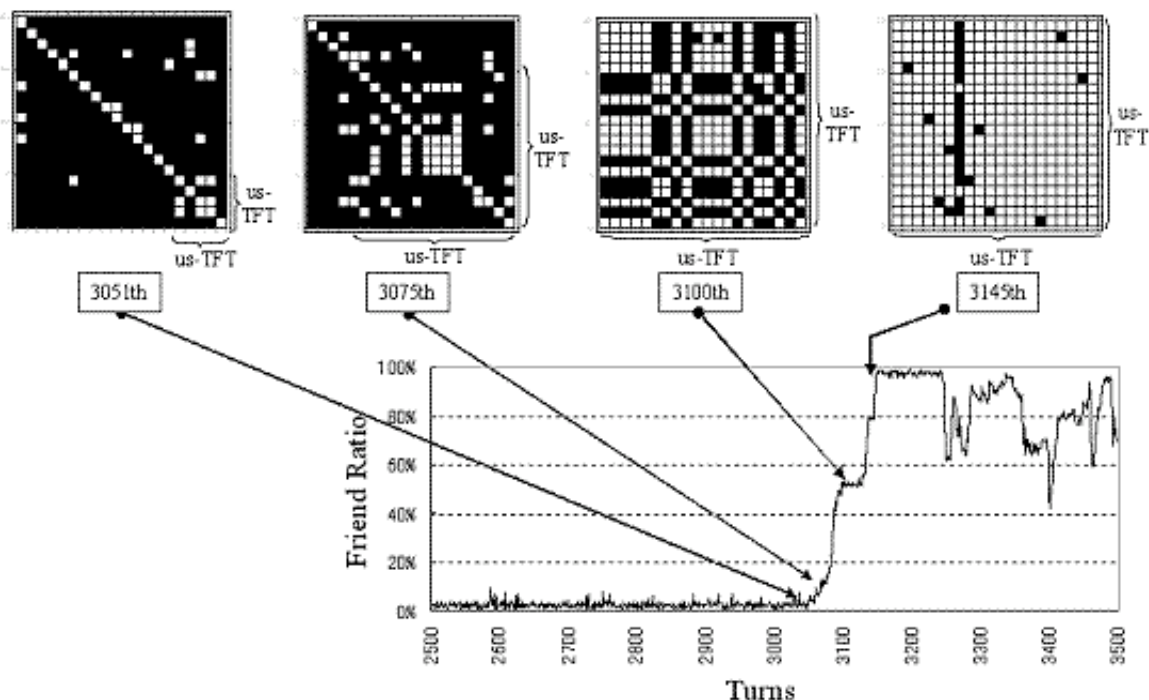


Figure 9. Typical Changes in Friendships

- (R2-1) A battling society continues up to about the 3000th turn.
- (R2-2) A few us-TFT agents with a few "friends" appear at around the 3050th turn.
- (R2-3) The number of such agents increases and they come to regard other us-TFT agents as "friends".
- (R2-4) Finally, a friendly us-TFT in-group appears, leading to the establishment of a peaceful society at around the 3150th turn.

8.7

We then examine how the us-TFT invades a battling society occupied solely by the ALL_D. To simplify our explanations, we first discuss the us-TFT's expansion mechanism and then discuss the initial state of the expansion.

8.8

To discuss the expansion mechanism, we assume that there are ALL_D agents and one us-TFT in-group whose members, of course, see each other as friends. The ALL_D agents and the us-TFT in-group see each other as enemies. Because the in-group members don't attack each other while the ALL_D agents attack each other, the members end up with higher payoffs than the ALL_D agents. This causes some ALL_D agents to change into us-TFT agents. Here, note that the new us-TFT agents see all others as "enemies". Because the new us-TFT agents fought against all others just one turn ago, all others are regarded as "enemies" according to the us-TFT definition. We call these agents "lonely us-TFT agents". Now, suppose that a lonely us-TFT agent happens to regard two members of the in-group as "friends" due to a misunderstanding (a perception error). The in-group members keep attacking the lonely agent, but don't attack the lonely agent's two new "friends" because the two are in-group members. From the lonely agent's viewpoint, all in-group members are peaceful toward his "friends", while they are hostile to himself. Following the algorithm PF6 and PF11, the total score of each in-group member assigned by the lonely agent becomes positive ($-1.0+1.0+1.0 = +1.0 > 0$)^[5]. As a result, the lonely agent changes his perception of the in-group members from "enemies" to "friends" and thus stops attacking the in-group members. In turn, the in-group members also change their perception of the lonely agent from an "enemy" to a "friend". In short, a perception error causes a lonely us-TFT agent to change into a member of an us-TFT in-group, which expands the size of the in-group

(Figure 10). In addition, note that the first step of the expansion isn't the in-group's peaceful action but the lonely agent's one. And note that the expansion proceeds like a snowball, once an us-TFT in-group begins to expand. The larger an in-group becomes, the more frequently a lonely us-TFT agent regards in-group members as "friends" due to a misunderstanding. This means that the larger group has higher probability to expand.

8.9

In the initial state of the expansion, a friendly us-TFT pair (or the smallest us-TFT in-group) happens to emerge. In particular, as the first step, mutation changes two ALL_D agents into us-TFT agents with no "friends" (lonely us-TFT agents), and then a perception error causes the two to regard each other as friends. In this manner, the smallest us-TFT in-group emerges, and it grows larger due to the mechanism described above. After all, the emergence of an us-TFT in-group leads to a peaceful society. (Hereinafter, we call a peaceful society, where all agents follow the us-TFT strategy and see each other as "friends", the "us-TFT society".)

8.10

This scenario reminds us of the following events sometimes seen in daily life. Consider the case of a lonely man and a friendly group. The man and some members of the group may become close friends by chance. This can lead to him eventually being welcomed into the group (being regarded as a member by the group).

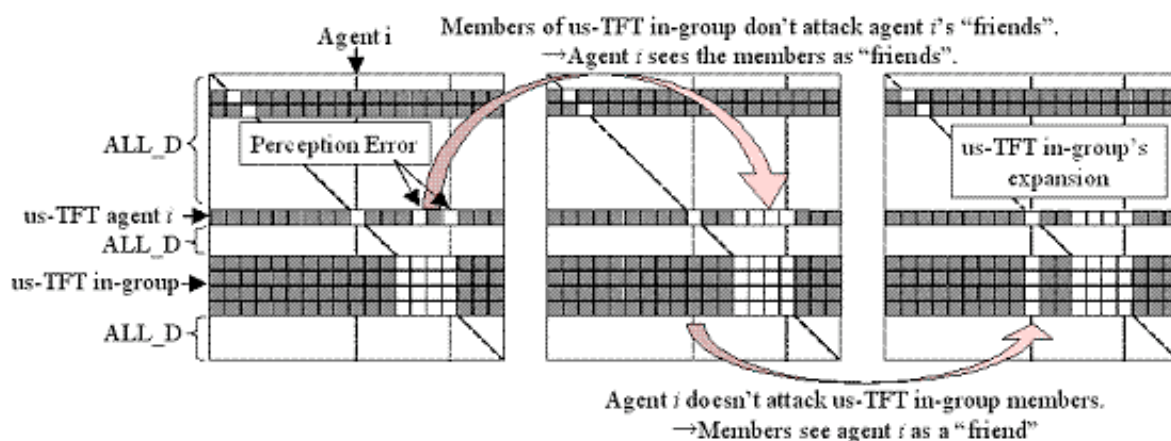


Figure 10. Expansion of us-TFT In-group due to Perception Error



Collapse of Peace due to us-CWD Cheaters

9.1

We are concerned about what effects other strategies except the us-TFT have on the emergence or collapse of peace. To investigate the question, we observed a friend ratio over turns and proportions of six strategies among all agents. Since it is complicated to analyze the six proportions at the same time, we simplified the task by introducing a prevailing strategy, which we defined as the strategy used by more than 50% of the agents. We observed the prevailing strategy for each turn in the simulations of Figure 7.

9.2

The results are shown in Figure 11. Graphs [1] and [3] show the friend ratio over turns, and graphs [2] and [4] show the corresponding prevailing strategy, which is shown as a colored zone and described together with a friend ratio.

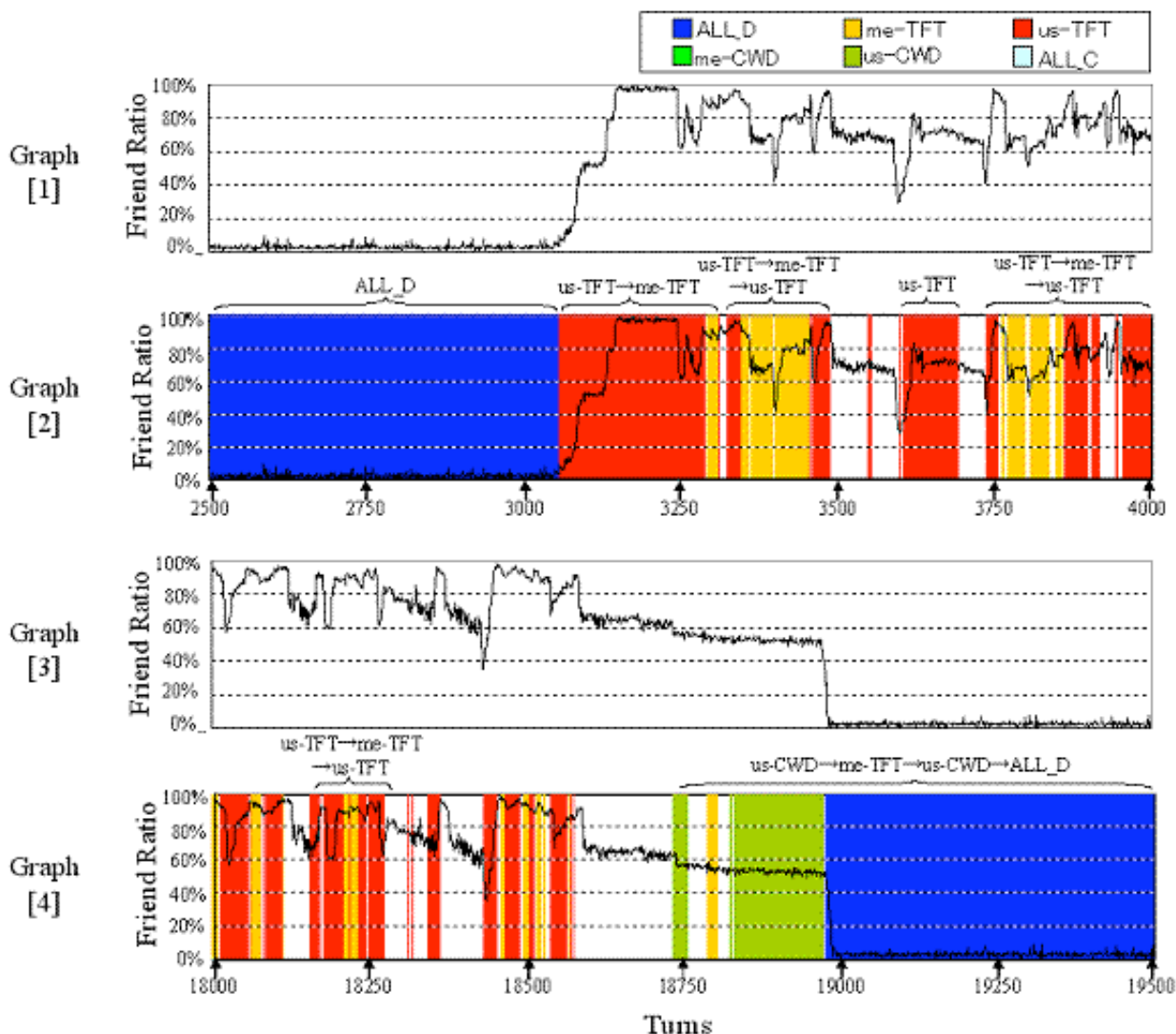


Figure 11. Prevailing Strategy vs. Turns (Round Robin Battle Game)

9.3

From these graphs, we can find the following.

- (R3-1) The us-TFT establishes a peaceful society as a prevailing strategy, and the friend ratio approaches 1.0.
- (R3-2) The me-TFT invades the society, and the us-TFT and me-TFT occupy the society alternately. As the me-TFT comes to prevail, the friend ratio decreases gradually. As the us-TFT comes to prevail, it increases gradually.
- (R3-3) The us-CWD invades and occupies the society in place of the us-TFT and me-TFT, and the friend ratio gradually decreases.
- (R3-4) Finally, the ALL_D invades and occupies the society in place of the us-CWD, and the friend ratio reaches 0. That is, the ALL_D eventually destroys the peaceful society.

9.4

We next investigate the rise and fall of these strategies. We first investigate whether the me-TFT can invade an us-TFT society. Suppose that one member of the in-group changes into a me-TFT agent by mutation. Such a me-TFT agent's payoff would be equal to that of each in-group member, because they all don't attack anyone and there is no difference between them. This means that the me-TFT cannot completely invade an us-TFT society through selection. Nevertheless, a perception error can change the situation. To examine this in more detail, let us suppose that a peaceful agent, who normally doesn't attack anyone, attacks a

me-TFT or an us-TFT agent due to a misunderstanding (a perception error). And let us compare the me-TFT's response and the us-TFT's one. In the me-TFT's case, the me-TFT agent attacks the peaceful agent, even though the agent is actually peaceful, because the peaceful agent directly attacked him. In the us-TFT's case, in contrast, the us-TFT agent doesn't attack the peaceful agent because he knows the agent's peaceful actions toward his "us" and doesn't mind the agent's direct attack against him. In short, both responses to the other's incorrect attack differ. An us-TFT agent tends to be generous and peaceful, while a me-TFT agent tends to be strict and militant (Figure 12). Therefore, a militant me-TFT agent gets a higher payoff than a peaceful us-TFT agent, and this leads to a me-TFT's invasion into an us-TFT society, like the ALL_D's invasion into an ALL_C society.

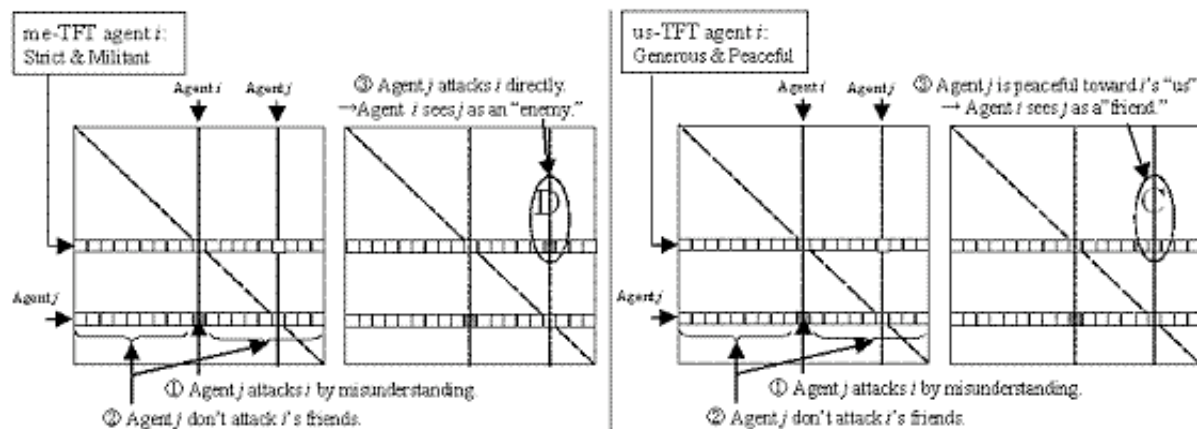


Figure 12. Strict me-TFT Agent vs. Generous us-TFT Agent

9.5

However, it should be noted that the me-TFT cannot completely exclude the us-TFT. This is because a me-TFT agent tends not to attack us-TFT agents since us-TFT agents generally don't attack him due to their generosity. On the other hand, a me-TFT agent often attacks me-TFT agents due to his strictness. Therefore, there are not me-TFT's attacks against us-TFT agents but frequent attacks among me-TFT agents. It reduces the me-TFT agents' payoffs, and these lower payoffs restrain the me-TFT's invasion.

9.6

Next, we examine whether the ALL_D can invade an us-TFT society. First, note that the ALL_D's payoff depends on whether or not the ALL_D agent and us-TFT society attacked each other. As the ALL_D's typical states corresponding to the relation with us-TFT society, we can see two types of the state for ALL_D agent, as follows.

- (AD1) ALL_D cheater: the agent attacks members of the us-TFT in-group, but the members don't attack him.
- (AD2) ALL_D fighter: the agent attacks members of the us-TFT in-group, and the members attack him.

9.7

We are interested in which state results in a higher payoff, an ALL_D cheater, an ALL_D fighter, or an us-TFT in-group member? Obviously, an ALL_D cheater's payoff is the highest, and an in-group member's payoff is higher than an ALL_D fighter's because the in-group members don't attack each other. Therefore, when we line them in order of ascending payoffs, the result is as follows: ALL_D fighter, in-group member, and ALL_D cheater.

9.8

Second, suppose that one member of the in-group changes into an ALL_D agent by mutation. Just after the mutation, he attacks the in-group members, while they don't attack him. That

is, he gets the highest payoff as an ALL_D cheater. However, the ALL_D's success doesn't last long because the in-group members identify who cheated them and they attack the cheater in the next turn. In other words, the ALL_D agent is successful for only one turn as a cheater. He is immediately enforced to fall into an ALL_D fighter against the in-group members. We call this process "us-TFT's revenge" (or "invalidation") of ALL_D cheaters. Once this revenge takes place, the ALL_D agents and the in-group members continue to attack each other. (The ALL_D agents remain fighters: Figure 13.)

9.9

Of course, the initial cheater produces many followers due to the high payoff. But the in-group immediately invalidates them one by one, and thus ALL_D fighters increase. Because the ALL_D fighter's payoff is lower than the in-group member's, only ALL_D fighters come to be selected and in-group members don't. Here note that the selected ALL_D fighters cannot change into ALL_D cheaters again. The selected ALL_D fighters imitate ALL_D cheaters due to the high payoff. But, because ALL_D fighters can not imitate the state (cheater) but imitate only the strategy (ALL_D), they remain ALL_D agents after imitating cheaters. As long as the in-group keeps attacking them, they keep attacking the in-group as ALL_D fighters. On the other hand, note that ALL_D cheaters appear only when the us-TFT in-group members are selected. That is, because the in-group members are not selected, ALL_D cheaters disappear after all. In this case, because an in-group member has the highest payoff, the selected ALL_D fighters change into us-TFT agents. That is, the ALL_D is excluded by the in-group us-TFT. In addition, the above scenario is applicable to the ALL_D's invasion into a me-TFT society in case of the round robin game.

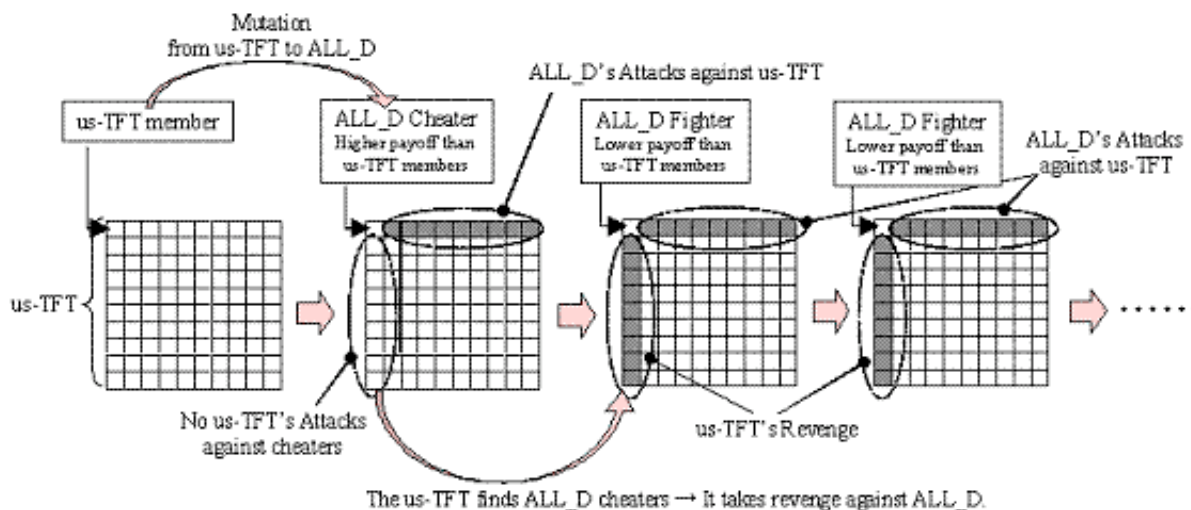


Figure 13. Changes in State of an ALL_D Agent

9.10

We then turn our attention to an us-CWD's invasion into an us-TFT society. As well as the ALL_D's case, we can see four types of the state for us-CWD agent, as typical states corresponding to the relation with us-TFT society, as follows.

- (CW1) us-CWD cheater: the agent attacks members of the us-TFT in-group, but the members don't attack him.
- (CW2) us-CWD fighter: the agent attacks members of the us-TFT in-group, and the members attack him.
- (CW3) us-CWD punished: the agent doesn't attack members of the us-TFT in-group, but the members attack him.
- (CW4) us-CWD pretender: the agent doesn't attack members of the us-TFT in-group, and the members don't attack him.

9.11

We are interested in which state results in a higher payoff, an us-CWD cheater, an us-CWD fighter, an us-CWD punished, an us-CWD pretender, or an us-TFT in-group member? The order of ascending payoffs for the us-CWD cheater, us-CWD fighter, and in-group member is the same as in the ALL_D case: us-CWD fighter, in-group member and us-CWD cheater. For the other states, the us-CWD punished's payoff is the lowest because it is attacked one-sidedly, and the us-CWD pretender's payoff is almost the same as that of the in-group members because the us-CWD pretender takes the same actions as in-group members. Therefore, when we line them up in order of ascending payoffs, the result is as follows: us-CWD punished, us-CWD fighter, us-CWD pretender/in-group member and us-CWD cheater. In addition, we also note that the state of an us-CWD agent keeps changing every turn as long as his strategy doesn't change: us-CWD cheater \Rightarrow us-CWD fighter \Rightarrow us-CWD punished \Rightarrow us-CWD pretender \Rightarrow us-CWD cheater \Rightarrow .

9.12

Now, suppose that one member of the in-group changes into an us-CWD agent by mutation. Just after the mutation, like the ALL_D's invasion, he earns the highest payoff as an us-CWD cheater, but the in-group members take revenge on him immediately. In response to this, as mentioned above, the us-CWD agent's responses are complicated (Figure 14). He fights against the in-group members, he is punished by the in-group members, he pretends to be an in-group member, and he cheats the in-group members again. (In contrast, an ALL_D agent keeps attacking the in-group members as a fighter after the in-group takes revenge on him.)

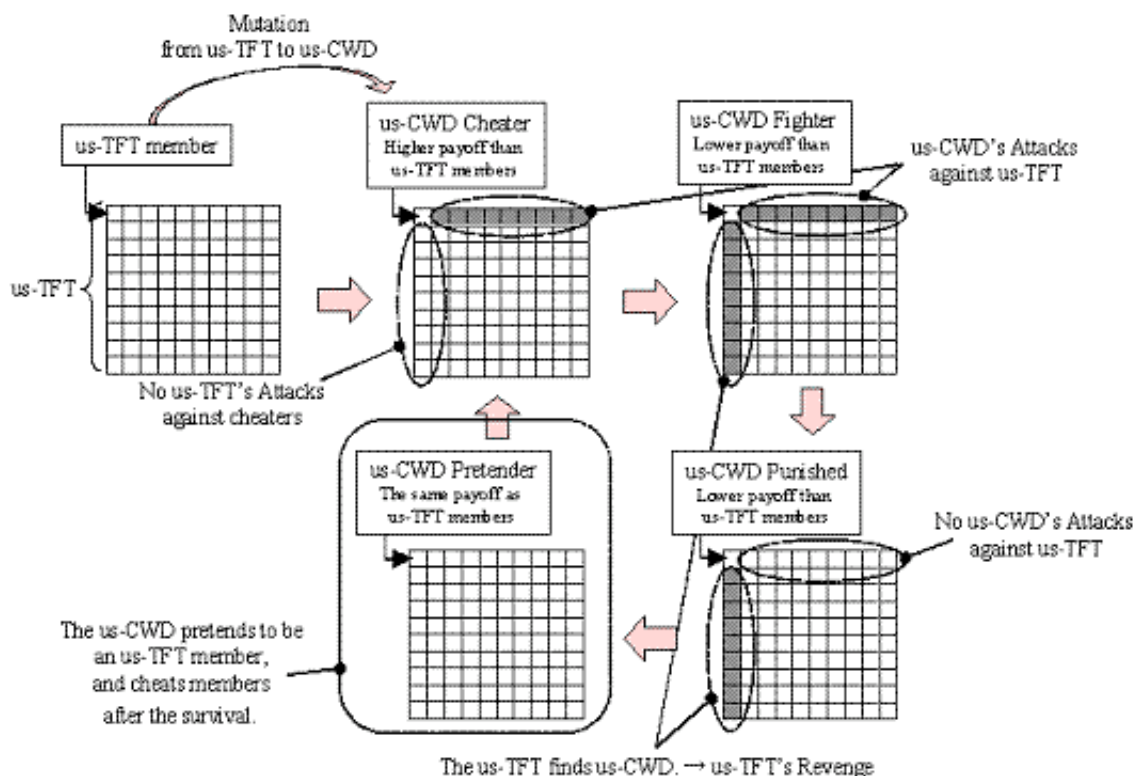


Figure 14. Changes in State of an us-CWD Agent

9.13

Here, for convenience, let us consider the case without us-CWD punisheds and pretenders. The mechanism would be same as that of the ALL_D's invasion. The initial cheater produces many followers due to the high payoff. But the in-group immediately takes revenge on them, and thus us-CWD fighters increase. Because the us-CWD fighter's payoffs are lower than the in-group member's, only us-CWD fighters come to be selected but in-group members don't.

When the selected us-CWD fighters imitate us-CWD cheaters due to the high payoff, they remain us-CWD fighters, because they can not imitate the state (cheater) but imitate only the strategy (us-CWD). On the other hand, because only selected in-group members can change into us-CWD cheaters and in-group members don't come to be selected, nobody change into us-CWD cheaters and then us-CWD cheaters disappear. In this case, because an in-group member has the highest payoff, the selected us-CWD fighters change into us-TFT agents. After all, the us-CWD is excluded by the in-group us-TFT. However, when we take us-CWD punisheds and pretenders into consideration, the situation changes. In the case, us-CWD fighters cannot remain us-CWD fighters and have to change into us-CWD punisheds and pretenders. Since an us-CWD pretender's payoff is almost as high as an in-group us-TFT's, the us-CWD pretenders are less likely to change their strategy. That is, they attack the in-group again as cheaters and produce many followers. Thus, the in-group us-TFT cannot stop an us-CWD's invasion because the in-group cannot identify who are the pretenders.

9.14

Furthermore, consider a society completely occupied by the us-CWD. If one of the us-CWD agents changes into an ALL_D agent, he keeps attacking the other us-CWD agents, while the us-CWD agents never attack him according to the definition of us-CWD. In other words, the ALL_D exploits the us-CWD all the time, and the ALL_D eventually completely excludes the us-CWD. We now can see the collapse mechanism of peace: after the us-CWD occupies a society in place of the us-TFT, the ALL_D excludes the us-CWD.

9.15

We also consider a me-CWD's invasion. The us-CWD's mechanism mentioned above is applicable to the me-CWD's invasion. However, if we take a perception error into account, both act differently. To better understand this, consider the situation in which a peaceful agent attacks a me-CWD or an us-CWD agent due to a misunderstanding. In the case of the me-CWD agent, he stops attacking the peaceful agent because he pays attention to only the agent's direct attack against him. In contrast, in the case of the us-CWD agent, he keeps attacking the peaceful agent because he pays attention to the agent's actions on his "us", which are peaceful. In sum, a me-CWD agent reacts sensitively to incorrect attacks as a fearful agent, while an us-CWD agent doesn't as a fearless agent. So, a me-CWD agent tends to be more peaceful than an us-CWD agent. So, the me-CWD cannot invade an us-TFT society easily, compared to the us-CWD.

9.16

We now summarize the considerations in this section. An ALL_D's invasion into an us-TFT society is almost impossible because of the us-TFT's revenge. Similarly, the me-TFT cannot invade an us-TFT society completely because of the me-TFT's cease-fire to the us-TFT. A me-CWD's invasion into an us-TFT society is not easy because me-CWD agents fear incorrect attacks very much. On the other hand, an us-CWD's invasion is more likely to succeed because us-CWD agents survive the us-TFT's revenge as pretenders and attack the in-group members as fearless cheaters^[6]. Then, with the round robin battle game, we can identify the typical transitions between prevailing strategies (Figure 15), and find the following story about the emergence and collapse of peace (Table 2).

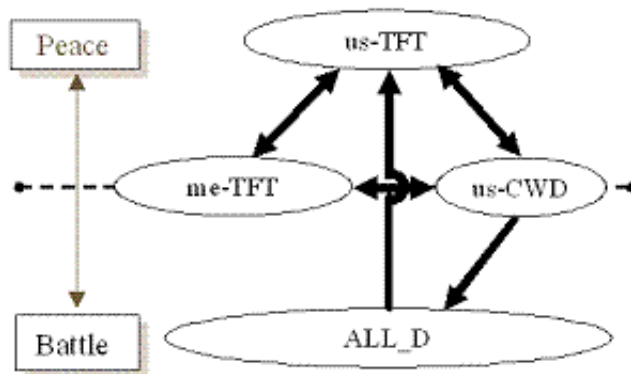


Figure 15. Typical Transitions between Prevailing Strategies (Round Robin Battle Game)

No	Description
(R4-1)	A small us-TFT in-group accidentally appears in a battling society. A perception error causes the us-TFT in-group to grow like a snowball. Eventually, the in-group establishes and maintains a peaceful society.
(R4-2)	Next, a variety of strategies appears. The me-TFT begins to invade the us-TFT society as strict and militant me-TFT agents. The me-TFT gradually makes the society hostile. But the invasion doesn't completely succeed due to the me-TFT's cease-fire to the us-TFT and mutual attacks among the me-TFT agents.
(R4-3)	The ALL_D invades the us-TFT society as cheaters. The in-group us-TFT immediately takes revenge (together with the me-TFT) and eventually drives them out.
(R4-4)	The us-CWD invades the us-TFT society as cheaters or pretenders. Because the us-TFT in-group (and the me-TFT) cannot identify the pretenders, the invasion is sometimes successful ⁶ .
(R4-5)	The ALL_D exploits the us-CWD as cheaters, which eventually leads to a battling society.

Table 2. The selected results of the simulations (Round Robin Game)

9.17

As is well known, the collapse of social order is sometimes triggered by the ALL_C. That is, the TFT establishes social order by defeating the ALL_D. Next, the ALL_C increases due to the TFT's umbrella protection against the ALL_D. The ALL_D exploits the ALL_C, which eventually destroys social order. In contrast, in our model, the ALL_C doesn't play an important role. This is because the ALL_C is always exploited by the us-CWD or me-CWD and cannot eventually increase. The ALL_D exploits the us-CWD which increases in place of the ALL_C, and destroys social order after all. Here, we remember such a metaphor that an us (me)-CWD agent corresponds to a person who selectively bullies unresisting weaker people. The above scenario on the us-CWD reminds us of a variety of sad examples, such as a boss dumping too many tasks on an uncomplaining staff member, a parent abusing a child, or a terrorist attacking unrelated and peaceful people.

 Collapse of Peace due to Careless me-TFT and Silent me-CWD

10.1

Do just the parameters' values used in Figure 7 result in the emergence of peace by chance?

To verify that our results are not a result of rare fortune, we carried out additional simulations with a variety of parameter values. Hereafter, we focus more on finding the sociological scenario rather than the quantitative discussion because our model does not significantly aim at reality. As well known, parametric surveys of Schelling's model showed that clusters of races emerges within a range of threshold related to peoples' moves to the same races. Because the model can be seen an ideal type, the most important finding is not the value but the existence itself of the range. The values of the range and parameters are not so often discussed in detail.

10.2

In parametric surveys, we found that a peaceful society emerged and collapsed in almost every case. Some notable results are as follows. The lower the perception error rate, the less frequently a peaceful society emerged. In particular, when it was set to zero, a peaceful society did not emerge, which is consistent with the explanation of the emergence of peace in Section 8. With respect to the strategy's mutation rate, there was a specific mutation rate at which the average friend ratio over turns was maximum. In other words, a peaceful society did not emerge when the mutation rate was equal to 0 or when it was sufficiently high. This seems reasonable. A mutation rate of zero causes an us-TFT agent not to appear, which leads to no emergence of peace. If the rate is high, the strategies are not selected in a selection process but at random. The lower the matching number in a battle game (M), the more frequently a peaceful society collapsed. In the case of a random matching battle game, corresponding to the low frequent generalized exchange, it seems fruitful to discuss in detail the emergence and collapse of peace, because results are different from the round robin's.

10.3

Typical results are shown in Figure 16. The simulation was run under the same conditions as that of Figure 11, except that the matching number (M) was set to be 4. As in Figure 11, we observed the friend ratio over turns and the corresponding prevailing strategy. From this figure, we can find that the collapses are triggered by me-TFT's or me-CWD's invasions.

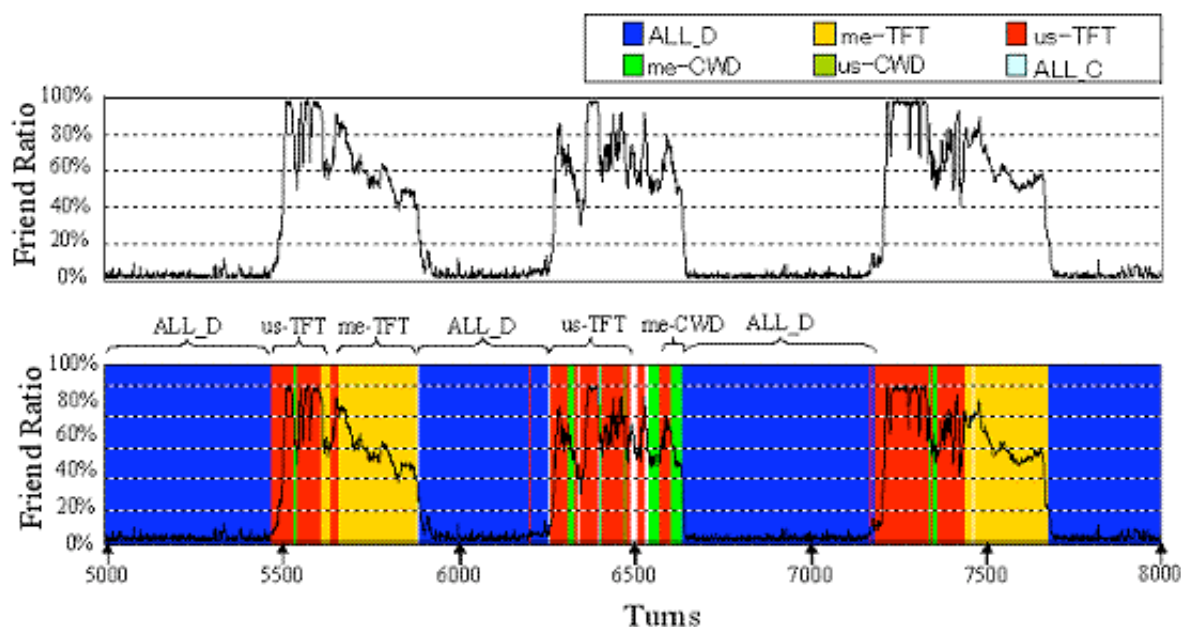


Figure 16. Prevailing Strategy vs. Turns (Random Matching Battle Game)

* Number of Agents: $N=20$ agents, Matching Number of Agent in One Battle Game: $M=4$ agents, Reflection Ratio: $R=10\%$, Perception Error Rate: $\mu_p=5\%$, Strategy's Mutation Rate:

$$\mu_s=0.3\%$$

- (R5-1) The us-TFT establishes a peaceful society as a prevailing strategy, and the friend ratio approaches 1.0.
- (R5-2) The me-TFT (me-CWD) invades and occupies the society in place of the us-TFT, and the friend ratio decreases gradually.
- (R5-3) The ALL_D then invades and occupies the society in place of the me-TFT (me-CWD), and the friend ratio reaches 0. That is, the ALL_D eventually destroys a peaceful society.

10.4

We are concerned about the reason why the me-TFT can invade an us-TFT society in the random matching battle game. Here, let us recall the mechanism already mentioned in the case of the round robin battle game. In this game, the me-TFT begins to invade an us-TFT society as militant me-TFT agents but the invasion is not completely successful. It is because the me-TFT agents stop attacking the us-TFT agents in response to the us-TFT's generosity, and because me-TFT agents keep attacking each other due to their strictness. That is, the me-TFT's peaceful attitude to the us-TFT and me-TFT's mutual attacks reduce the me-TFT's payoff.

10.5

On the other hand, in the case of the random matching game, a me-TFT agent's response changes. That is, a me-TFT agent sometimes keeps attacking an us-TFT agent who is his "friend". The reason is as follows. The random matching means that all agents don't necessarily meet all other agents. So, some us-TFT agents with friendly perceptions for me-TFT agents lose chances to take peaceful actions to them. A me-TFT agent may thus fail to receive peaceful signals from us-TFT agents and keeps attacking us-TFT agents. We call him a "careless me-TFT agent". Since a careless me-TFT agent attacks anybody and his payoff is always higher than us-TFT's, the me-TFT can invade an us-TFT society, like an ALL_D's invasion into the ALL_C.

10.6

Next, let us consider an ALL_D's invasion into a society occupied by the me-TFT. Suppose that one of me-TFT agents becomes an ALL_D agent by mutation and that he begins to invade as a cheater. In the round robin battle game, all the me-TFT agents are attacked directly by him, which causes all the me-TFT agents to take revenge against him. That is, the me-TFT can repel an ALL_D's invasion. In the random matching battle game, only a few me-TFT agents are directly attacked by the ALL_D cheater, and only these agents take revenge. Since the me-TFT's revenge is limited, the ALL_D's invasion is not repelled.

10.7

We also examine a me-CWD's invasion into an us-TFT society. In the round robin battle game, the me-CWD survives the us-TFT's revenge as pretenders, and attacks the in-group members again as cheaters. In the random matching battle game, the basic mechanism is the same, but there is a difference in terms of cheating intensity. Suppose that one member of the in-group becomes a me-CWD agent and begins to attack the in-group members as a cheater. According to the me-CWD's definition, he attacks the agents who did not attack him in the preceding turn. In the random matching battle game, a few agents interact with the me-CWD agent, so the me-CWD agent has only a few targets to attack. That is, a me-CWD agent in the random matching game is not so militant but rather peaceful. Hence, the in-group members cannot identify the me-CWD agent and thus continue to regard him as a "friend". This means that the me-CWD can secretly penetrate an us-TFT society, avoiding the us-TFT's revenge. We call him a "silent me-CWD agent". In addition, how about the us-CWD's invasion in the case of the random matching battle game? In the random matching battle game, a me-CWD agent has a few targets to attack, while an us-CWD agent has many targets due to much information from his "friends". So, the us-CWD agent is militant enough for the

in-group members to identify him. He is susceptible to the us-TFT's revenge. An us-CWD's invasion is thus less likely to succeed than a me-CWD's.

10.8

Finally, we consider an ALL_D's invasion into a society occupied by the me-CWD. As mentioned earlier, the ALL_D easily exploits the me-CWD. That is, once the me-CWD occupies a society, the ALL_D exploits and excludes the me-CWD completely.

10.9

We now summarize these considerations. A me-TFT's invasion is possible because a me-TFT agent attacks anyone due to its carelessness. A me-CWD's invasion has a better chance of success than an us-CWD's one, because the penetration is difficult to be detected, so the us-TFT's revenge is not induced. Concerning the random matching battle game, we can identify the typical transitions between prevailing strategies (Figure 17), and find the following story about the emergence and collapse of peace (Table 3).

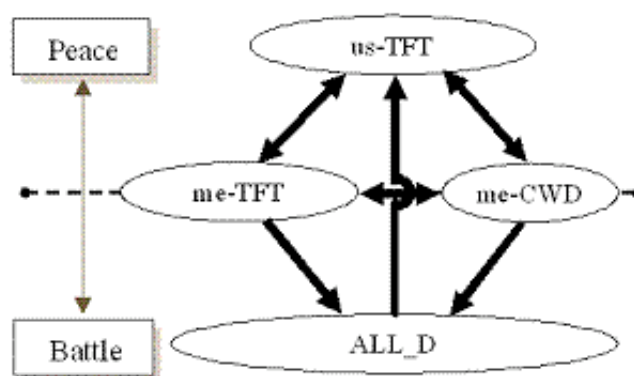


Figure 17. Typical Transitions between Prevailing Strategies (Random Matching Battle Game)

No	Description
(R6-1)	A small us-TFT in-group brings about a peaceful society (as well as the round robin battle game).
(R6-2)	The ALL_D begins to invade the us-TFT society as cheaters, but the invasion fails to invade due to the us-TFT's revenge (as well as the round robin battle game).
(R6-3)	Next, a variety of strategies appears. The me-TFT begins to invade the us-TFT society as careless agents to us-TFT's friendships, and the invasion sometimes succeeds due to the indiscriminate attacks against everyone.
(R6-4)	The me-CWD begins to invade the us-TFT society as cheaters or pretenders, and the invasion sometimes succeeds because the in-group members cannot identify them and take revenge ⁶ . In contrast, an us-CWD's invasion is less likely to succeed due to the us-TFT's revenge.
(R6-5)	The ALL_D invades the society occupied by the me-TFT or me-CWD as cheaters, and eventually establishes a battling society (as well as the round robin battle game).

Table 3. The selected results of the simulations (Random Matching Game)

10.10

Because the random matching game corresponds to the low frequent generalized exchange, we may be able to regard the game as a society with fewer interactions. Then, the scenario on the me-TFT reminds us that people become unconcerned with others and crime increases in such a society with the weaker ties between people. Also, if we recall that a me (us)-CWD

agent corresponds to a bully, the scenario on the me-CWD reminds us that bullying takes place easily in the weak-tie society.

Discussion

11.1

We have found new mechanisms driving the emergence and collapse of social order based on the distinction of a friend and an enemy by Schmitt^[7]. First, we embodied the distinction of a friend and an enemy in the us-TFT strategy, which takes revenge on the attack against "us", and we found that the us-TFT strategy brings about a peaceful society. This means that the emergence of social order does not require information on all agents' all actions, but only information on actions toward the agent himself and his "friends". This is interesting in the sense that the more realistic assumption about information results in social order. Second, we found that the me-TFT, me-CWD and us-CWD strategies invade an us-TFT society and the invasions lead to a battling society. This is also interesting, because collapses differ from the well-known collapse triggered by the ALL_C. We thus conclude that they are new scenarios.

11.2

Now, out of sociological interest, we will consider the meaning of an us-TFT society. Concerning the emerging phase of a peaceful society, we can interpret the society as is shown in Figure 18.

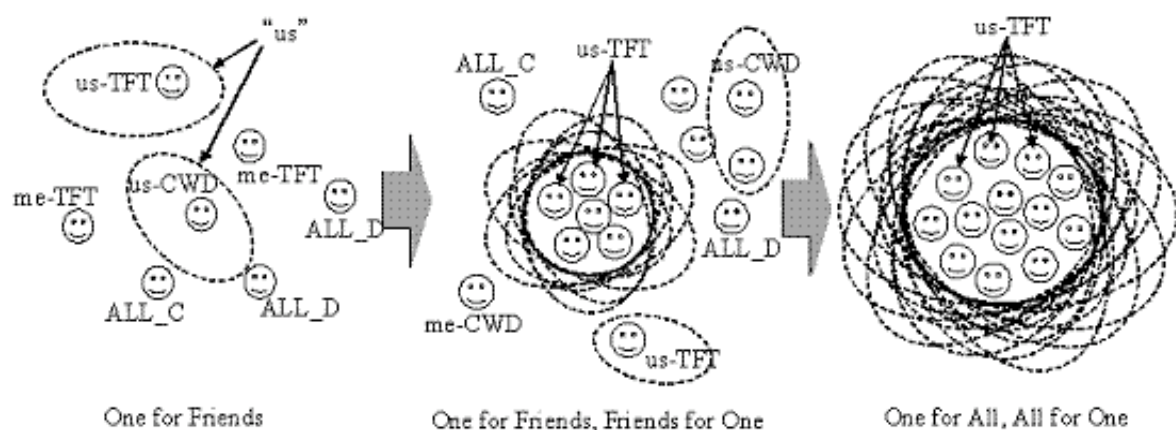


Figure 18. Emergence of Community

- (CC1) One for Friends: in a battling society full of militant people, some peaceful people appear by chance.
- (CC2) One for Friends, Friends for One: the formation of friendships among peaceful people results in a friendly in-group, whose members see each other as friends.
- (CC3) One for All, All for One: the in-group grows up to a peaceful society, where everyone regards everyone else as a friend.

11.3

Let us consider the last interpretation, One for All, All for One. In the us-TFT's definition, as described in Section 4, each us-TFT in-group member comes to have a common social perception and to take the same actions as the other members. This is because one in-group member's "us" corresponds to another's "us", and each member has common information on others' actions. For example, if a member attacks another member, all members come to regard the attacker as a common enemy. We can thus find that an us-TFT society works just as a "community". Here, let us note that the us-TFT's definition doesn't include assumptions leading to the emergence of a community directly. In the us-TFT's definition, an us-TFT agent can have a unique social perception and thus take actions different from those of other us-

TFT agents. When an agent attacks another agent, each us-TFT agent evaluates the attacker differently. This is because each agent's "us" is mutually different and each agent has unique information on others' actions. Therefore, we can conclude that the us-TFT's definition does not directly result in the emergence of community and that the community has emerged spontaneously^[8]. We can see some studies related to a community, such as on the in-group of sharing members and ostracism from it ([Younger 2005](#)), the powerful and powerless groups and sanctioning by the powerful ([Saam and Harrer 1999](#)) and so on. They introduced a group into their models *a priori*, and so they do not show the emergence of a group.

11.4

Next, we consider the morality of community. A lot of previous studies regarding a group's cooperation provide useful insights. In general, they pointed out that cooperation comes out from consideration of the group's payoff, commitment of the group's norm, and trust of the group members ([Messick and Brewer 1983](#); [Kerr 1995](#)), and that these mechanisms are activated by communication among members ([Dawes, McTavish and Shaklee 1977](#); [Dawes, Orbell and van de Kragt 1988](#); [Kerr and Kaufman-Gilliland 1994](#)). Studies on the consideration of group's payoff pointed out that people give weight to the group's payoff in individual decision making ([Brewer and Kramer 1986](#), [Messick and McClintock 1968](#)). It makes a contrast to the assumption that each us-TFT agent takes only his own payoff into account. Studies on the commitment of group's norm pointed out that group members make a commitment to the group's norm and are punished for violations ([Messick and Brewer 1983](#); [Brewer and Schneider 1990](#)). And other studies (e.g., [Axelrod 1986](#)) also explained that sanctions (punishments) are taken even if they impose a cost on the sanctioner (self-sacrifice). In contrast, the us-TFT agent doesn't attack the other as a sanction with self-sacrifice; he attacks the other simply as a revenge, which increases his own payoff. In sum, the "group" in previous studies consists of members who put importance on the group's payoff or commit to the group's norm, while the us-TFT's "community" consists of members who are socially aware and seek their own payoffs. In other words, the community doesn't result from morality. The term "friend" tends to remind us of morality. However, the reason why the us-TFT prevails is that inferior agents adopt it to obtain higher payoffs for themselves. In summary, we found that the code of a friend and an enemy, which is embodied in the us-TFT strategy, results in social order through the emergence of community. However, the us-TFT community can be regarded as a collective security system with calculating individuals seeking higher payoffs. It is not a moral community, but just a political community.

11.5

We next discuss our model's coverage. To directly see the effects of the distinction of a friend and an enemy, we made as simple and neutral a model as possible. Therefore, we cannot show a real society which our model describes, but we can show an ideal type of society to which our model corresponds. In other words, we show the limit of our model from the viewpoint of describing real societies.

11.6

First, as seen easily, the us-TFT is a solution to the 2PD problem but not to social dilemmas (NPD problem) in the public good's provision such as the environmental preservation and the pension system. In particular, the real society needs a sanctioning mechanism for suppressing free riders ([Coleman 1990](#)). However, since our model has no such mechanism, it cannot deal with social dilemmas. Second, the us-TFT doesn't work well in a large-scale society. The larger the us-TFT in-group, the less realistic becomes our assumption that an us-TFT agent acts based on others' actions toward him and his "friends". When a society becomes one large us-TFT in-group, each us-TFT agent has to watch the actions of all other agents. He cannot handle such a lot of information because of his limited capability. The model thus suffers from the same problem as in previous studies. Third, our model cannot discuss issues

regarding rumour and public opinion. As mentioned above, the larger the society, the more impossible it is to watch many other agents located in physically and socially long distances. The information exchange mechanism among a lot of agents seems necessary, for example, local information exchange or mass media^[9]. Fourth, our model cannot deal with ethnocentrism or nationalism. Previous studies pointed out that group identification increases cooperation under NPD situation (Messick and Brewer 1983; Kerr 1995; Brewer and Schneider 1990). Especially, minimal manipulation of group identification with abstract and vague similarities can increase cooperation rather than discussions among members (Brewer and Kramer 1986). These indicate that group identification using a common symbol can foster cooperation. Then, the community model using a common symbol may be able to be made. If we regard a race or a nationality as a symbol, for example, conflicts due to ethnocentrism or nationalism can be discussed as a system of opposing in-groups^[10]. However, the us-TFT doesn't have any symbols. It cannot describe several in-groups with different identities. In other words, our model results in a mixed state of only one in-group and independent individuals. Finally, our model cannot discuss inequality such as rich or poor classes. The agent in our model is concerned about whether or not the other acted peacefully, but not about whether or not the other has strong power politically or economically. In this sense, all agents are even^[11]. Our model cannot fundamentally deal with the problem of inequality.

11.7

Consequently, our model is more applicable to the society, a) which doesn't suffer from the NPD problem and has no sanctions, b) which has a small-scale population, c) which doesn't have rumor and mass media, d) which is unconscious of its group identity (which doesn't know other alien groups), e) which consists of equal people.



Acknowledgements

This work was supported by the Ministry of Education, Culture, Sports, Science and Technology, Japan, as a project of the 21st Century COE program: "Creation of Agent-Based Social Systems Sciences (ABSSS)", under the direction of Prof. Hiroshi Deguchi, the Tokyo Institute of Technology. We wish to thank the ABSSS project members, especially Prof. Takatoshi Imada and Prof. Hiroshi Deguchi, for their helpful discussions. This study was presented at NACCSOS 2005, Notre Dame, IN, June 2005. We thank the participants for their helpful comments and perceptive questions.



Notes

¹ The study showed that the emergence of peace is driven by a typical transition such as me-ALL_D→us-ALL_D→us-TFT→me-TFT. In this case, the us-CWD and me-CWD don't play an important role.

² In indirect reciprocity studies, the social perception that all others are friends is newly assigned to all agents at the beginning of every turn. And all agents' perceptions continue to be updated during the turn.

³ By introducing an error rate, indirect reciprocity studies assume a few agents who have different social perceptions from others. However, this can be seen as simply a perturbation because it doesn't change the basic logic that the same initial perceptions and experiences lead to the same perceptions. Moreover, even if the number of agents with different perceptions increases, all perceptions are abandoned and newly assigned at the beginning of

each turn.

⁴ In this study, the social perception that all others are enemies is assigned to all agents in the 0th turn. Additionally, all agents' perceptions are not updated during one turn but continue to be updated throughout the turns. That is, an initial perception in our study doesn't mean a perception at the beginning of a turn, but one in the 0th turn.

⁵ If a lonely us-TFT agent happens to regard only one member of the in-group as a "friend" due to a misunderstanding, nothing happens. From his viewpoint, the in-group members are hostile to himself but peaceful to his "friends". Following PF12 in the us-TFT definition, each member's total score equals to zero ($-1.0+1.0 = 0$), and so members remain as "enemies" for the lonely us-TFT agent. If we change PF11 and PF12 into a new rule that requires an agent to change his perception of another agent from an "enemy" to a "friend" if the total score of the other agent is positive or zero, the consequences change. In the case, members change from "enemies" to "friends" for the lonely us-TFT agent. That is, having one new "friend" due to a misunderstanding causes the lonely us-TFT agent to change into a member of an us-TFT in-group. Note that original PF11 and PF12 are regarded as neutral rules. A violation of the neutrality, the modification of PF11 and PF12 leads to more frequent emergences of a peaceful state.

⁶ The scenarios show that the us (me)-CWD destroys social order. In order to prevent it, it is necessary to identify us (me)-CWD agents. Previous studies pointed out that cheaters can be identified by a strategy using a signal, which is called "costly signal strategy" ([Smith and Bliege Bird 2000](#); [Bliege Bird, Smith and Bird 2001](#); [Bliege Bird and Smith 2005](#)) or "handicap theory" ([Zahavi 1975](#); [Zahavi 1977](#); [Boone 1998](#)). According them, a honest agent sends others a peculiar signal though it imposes a cost to him, while a cheater cannot. That is, a strategy based on a signal may evolve together with one based on others' actions (us-TFT), and then the signal-based strategy may exclude us (me)-CWD agents.

⁷ The findings of this study should be verified empirically. It is desirable to verify them quantitatively. (As a good example, Cederman ([2003](#)) quantitatively verified the power-law in international conflicts.) However, in general, models following the KISS principle don't aim at a precise prediction of social phenomena. Therefore, a quantitative comparison between simulation results and empirical facts is probably impossible, although it is an ideal approach. A more realistic approach would be to identify the similarities between simulation results and real phenomena, such as historical phenomena.

⁸ In previous studies, DISC agents hold a common perception (good or bad) and take a common actions (help or don't help) against the other. That is, there seem to be something like a norm. However, this norm doesn't emerge but comes directly from their assumptions of the common initial perception and the common information about others' actions.

⁹ Most related studies pay attentions to reputation and communication of reputations as a mechanism for sharing information ([Janssen 2006](#); [Conte and Paolucci 2002](#); [Sabater and Sierra 2002](#); [Sabater et al. 2006](#); [Sabater 2003](#); [Hahn et al. 2007](#); [Ashri et al. 2005](#); [Schlosser et al. 2006](#); [Younger 2004](#); [Castelfranchi et al. 1998](#); [Hales 2002](#)). In these studies, agents update their own reputations about the other by exchanging reputations among them. In contrast, our study does not assume the communication (exchange) of reputations about the other. Furthermore, in reputation theories related to e-commerce, a lie and credibility have been deeply considered ([Younger 2004](#); [Castelfranchi et al. 1998](#); [Hales 2002](#)).

¹⁰ Hales ([2002](#)) and Axelrod & Hammond ([2003](#)) showed the emergence of a group using a tag model.

¹¹ Most non-economic studies pay attentions to social classes. They assume that agent's actions depend on whether agents are strong or weak ([Younger 2004](#); [Castelfranchi et al. 1998](#); [Hales 2002](#)). Furthermore, many studies on e-commerce have introduced roles such as a seller and a buyer, for example ([Janssen 2006](#); [Conte and Paolucci 2002](#); [Sabater and Sierra 2002](#); [Sabater et al. 2006](#); [Sabater 2003](#); [Hahn et al. 2007](#); [Ashri et al. 2005](#); [Schlosser et al. 2006](#)).



References

- ASHRI, R., S. D. Ramchurn, J. Sabater, M. Luck and N. R. Jennings (2005), "Trust evaluation through relationship analysis", *International Conference on Autonomous Agents archive, Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*.
- AXELROD, R. (1986), "An evolutionary approach to norms", *American Political Science Review*, 80, 1095–1111.
- AXELROD, R. and R. A. Hammond (2003), "The Evolution of Ethnocentric Behavior", *Midwest Political Science Convention*, Chicago, IL
- BLIEGE BIRD, R., E. A. Smith and D. W. Bird (2001), "The Hunting Handicap: Costly Signaling in Human Foraging Strategies", *Behavioral Ecology and Sociobiology*, 50 (1), 9–19.
- BLIEGE Bird, R. and E. A. Smith (2005), "Signaling Theory, Strategic Interaction, and Symbolic Capital", *Current Anthropology*, 46 (2), 221–248.
- BOONE, J. L. (1998), "The Evolution of Magnanimity: When Is It Better to Give than to Receive?", *Human Nature*, 9 (1), 1–21.
- BREWER, M. B. and R. M. Kramer (1986), "Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing", *Journal of Personality and Social Psychology*, 50, 543–549.
- BREWER, M. B. and S. K. Schneider (1990), "Social identity and social dilemmas: A double-edged sword", in D. Abrams & M. Hogg (Eds.) *Social identity theory: Constructive and Critical advances*, Harvester/Wheatsheaf, NY.
- CASTELFRANCHI, C., R. Conte and M. Paolucci (1998), "Normative reputation and the costs of compliance", *Journal of Artificial Societies and Social Simulation*, 1, no.3.
- CEDERMAN, L. (2003), "Modeling the Size of Wars: From Billiard Balls to Sandpiles", *American Political Science Review*, 1 (97), 135–150.
- COLEMAN, J. S. (1990), *Foundations of Social Theory*, Cambridge University Press, Cambridge.
- CONTE, R. and M. Paolucci (2002), *Reputation in Artificial Societies: Social Beliefs for Social Order*, Kluwer Academic Publishers, Dordrecht.
- DAWES, R. M., J. McTavish and H. Shaklee (1977), "Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation", *Journal of Personality and Social Psychology*, 35, 1–11.
- DAWES, R. M., J. M. Orbell and A. J. van de Kragt (1988), "Not me or thee but me: The importance of group identity in eliciting cooperation in dilemma situations", *Acta*

Psychologica, 68, 83–97.

HABERMAS, J. (1981), *Theorie Des Kommunikativen Handelns*, Suhrkamp Verlag, Frankfurt/Main.

HAHN, C., B. Fley, M. Florian, D. Spresny and K. Fischer (2007), "Social Reputation: a Mechanism for Flexible Self-Regulation of Multiagent Systems", *Journal of Artificial Societies and Social Simulation*, 10, no.1.

HALES, D. (2002), "Group Reputation Supports Beneficent Norms", *Journal of Artificial Societies and Social Simulation*, 5, no.4.

HOBBS, T. (1651), *Leviathan*, printed for Andrew Crooke.

JANSSEN, M. (2006), "Evolution of Cooperation when Feedback to Reputation Scores is Voluntary", *Journal of Artificial Societies and Social Simulation*, 9, no.1.

KERR, N. L. and C. M. Kaufman-Gilliland (1994), "Communication, commitment, and cooperation in social dilemmas", *Journal of Personality and Social Psychology*, 66, 513–529.

KERR, N. L. (1995), "Norms in social dilemmas", in D. Schroeder (Eds.) *Social dilemmas: Social psychological perspectives*, NY: Pergamon Press, 31–47.

LEIMAR, O. and P. Hammerstein (2001), "Evolution of Cooperation through Indirect Reciprocity", *Proceedings of the Royal Society of London Series B: Biological Sciences*, 268, 743_753.

LUHMANN, N. (1968), *Zweckbegriff und Systemrationalität: über die Funktion von Zwecken in sozialen Systemen*, Suhrkamp, 1977.

MESSICK, D. M. and C. G. McClintock (1968), "Motivational basis of choice in experimental games", *Journal Experimental Social Psychology*, 4, 1–25.

MESSICK, D. M. and M. B. Brewer (1983), "Solving social dilemmas: A review", in L. Wheeler & P. Shaver (Eds.) *Annual review of personality and social psychology*, 3, 11–44.

NAKAI, Y. and M. Muto (2005), "Evolutionary Simulation of Peace with Altruistic Strategy for Selected Friends", *Socio-Information Studies*, 9(2), 59–71.

NOVAK, M. A. and K. Sigmund (1998a), "Evolution of Indirect Reciprocity by Image Scoring", *Nature*, 393, 573_577.

NOVAK, M. A. and K. Sigmund (1998b), "The Dynamics of Indirect Reciprocity", *Journal of Theoretical Biology*, 194, 561_574.

PARSONS, T. (1937), *The Structure of Social Action*, McGraw Hill.

SAAM, N. and A. Harrer (1999), "Simulating Norms, Social Inequality, and Functional Change in Artificial Societies", *Journal of Artificial Societies and Social Simulation*, 2, no.1.

SABATER, J. and C. Sierra (2002), "Reputation and social network analysis in multi-agent systems", In *Proceedings AAMAS-02*, Bologna, Italy, 475–482.

SABATER, J. (2003), "Trust and Reputation for agent societies", *PhD thesis. Artificial Intelligence Research Institute (IIIA-CSIC)*, Bellaterra, Catalonia, Spain.

SABATER, J., M. Paolucci and R. Conte (2006), "Repage: REputation and ImAGE Among Limited

Autonomous Partners", *Journal of Artificial Societies and Social Simulation*, 9, no.2.

SCHELLING, T. (1969), "Models of Segregation", *American Economic Review*, 59(2), 488–493.

SCHLOSSER, A., M. Voss and L. Brückner (2006), "On the Simulation of Global Reputation Systems", *Journal of Artificial Societies and Social Simulation*, 9, no.1.

SCHMITT, C. (1932), *Der Begriff des Politischen*, Dunker & Humblot

SMITH, E. A. and R. Bliege Bird (2000), "Turtle Hunting and Tombstone Opening: Public Generosity as Costly Signaling", *Evolution and Human Behavior*, 21, 245–261.

TAKAGI, E. (1996), "The generalized exchange perspective on the evolution of altruism", in W. Liebrand & D. Messick (Eds.) *Frontiers in Social Dilemmas Research*, Berlin: Springer–Verlag, 311_336.

TAKAHASHI, N. and R. Mashima (2003), "The emergence of indirect reciprocity: is the standing strategy the answer?", *Center for the Study of Cultural and Ecological Foundations of the Mind: Working Paper Series No.29*, Hokkaido University, Japan.

TAKAHASHI, N. and T. Yamagishi (1996), "Social Relational Foundation of Altruistic Behavior", *The Japanese Journal of Experimental Social Psychology*, 36(1), 1_11.

YOUNGER, S. (2004), "Reciprocity, Normative Reputation, and the Development of Mutual Obligation in Gift–Giving Societies", *Journal of Artificial Societies and Social Simulation*, 7, no.1.

YOUNGER, S. (2005), "Reciprocity, Sanctions, and the Development of Mutual Obligation in Egalitarian Societies", *Journal of Artificial Societies and Social Simulation*, 8, no.2.

ZAHAVI, A. (1975), "Mate Selection: A Selection for Handicap", *Journal of Theoretical Biology*, 53, 205–214.

ZAHAVI, A. (1977), "Reliability in Communication Systems and the Evolution of Altruism", in B. Stonehouse and C. Perrins (Eds.) *Evolutionary Ecology*, London: Macmillan Press, 253–259.

[Return to Contents of this issue](#)

© [Copyright Journal of Artificial Societies and Social Simulation, \[2008\]](#)

