國立交通大學

資訊科學系

碩 士 論 文

代理人自我覺察能力對於個體積效與合作行為之研究

A Study on Agent Self-awareness for Individual Performance and Collaborative Behavior

研 究 生：陳敬華

指導教授：孫春在 教授

中 華 民 國 九 十 四 年 六 月

代理人自我覺察能力對於個體績效與合作行為之研究

A Study on Agent Self-awareness for Individual Performance and Collaborative Behavior

研 究 生：陳敬華　　Student：Ching-Hua Chen

指導教授：孫春在　　Advisor：Chuen-Tsai Sun

國 立 交 通 大 學

資 訊 科 學 研 究 所

碩 士 論 文

A Thesis

Submitted to Department of Computer and Information Science

College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer and Information Science

June 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年六月

代理人自我覺察能力對於個體績效與合作行為之研究

學生：陳敬華　　　　　　　　　　　　　指導教授：孫春在　教授

國立交通大學資訊科學研究所

摘要

本研究從智慧型代理人的角度出發，回應傳統人工智慧領域中機器學習的核心問題：以世界模型為基礎的學習型代理人的設計不足之處。為了使代理人具備自我學習的能力，換言之，類似人類的自我覺察能力，我們提出一套新穎的代理人認知學習架構，包含外在學習與內在認知雙重概念，並相容於舊有的代理人架構。同時，為了有效驗證此架構是強固的、可靠的，可廣泛地應用在許多實際的狀況，如電子商務環境、社會科學模擬系統，我們將代理人的目標與生存環境之間所造成的衝突，以反覆囚犯困局來模擬與實驗，透過代理人的人格特質分析，我們提出一套以超我層次為自我覺察目標的代理人自覺模型，並以前述的反覆囚犯困局的實驗結果，來分析自我覺察能力對於學習型代理人的個體績效與合作行為的影響。

根據實驗結果，本研究證明以超我層次為自我覺察目標的代理人自覺模型，可有效幫助學習型代理人提升表現成效，並使合作行為提前浮現。更進一步的模擬與分析，我們發現只需要少數的代理人具備自覺能力，即可提升整體代理人社會公益，這些實驗結果也同時驗證本研究提出的代理人認知學習架構。最後，我們期望本研究所探討的方向與內容能讓大家重新思考與重視自我覺察對學習型代理人設計的重要性。

關鍵字：自我覺察、自我基模、學習型代理人、人工社會衝突、複雜行為浮現

A Study on Agent Self-awareness for Individual Performance and Collaborative Behavior

Student: Ching-Hua Chen Advisor: Dr. Chuen-Tsai Sun

Institute of Computer and Information Science
National Chiao Tung University

ABSTRACT

The approach of this research, how intelligent agents learn, is to deal with a core problem of Machine Learning. The problem of traditional artificial intelligence lies in the flaw that learning agents are designed on the basis of **World Model**. To endue agents with **Self-Learning** ability, in other words, the ability similar to self-awareness of human beings, we proposed a new cognitive learning model, which includes both external learning and internal cognition compatible with the former structure, for agents, called **A**gent **C**ognition **L**earning **M**odel **(ACLM)**. In order to prove this model is robust, reliable, and extensively applicable for real situations such as E-commerce or social science simulation systems, we will simulate and experiment **the conflict between the societal and self-interested goals** of agents with **Iterative Prisoner's Dilemma** on **Social Networks**. Through an analysis on personalities of agents, we proposed a **Self-Awareness Model** in **Superego Level** for agents. With the experimental results, we will analyze how individual performance and collaborative behavior of learning agents would be affected.

The results of the experiments would demonstrate that the self-awareness model aim for superego level could certainly improve the performance of learning agents and expedite the emergence of collaborative behavior. Further simulation and analysis would show that as few of the agents are capable of self-awareness, the whole social benefits of agents would be enriched. These results also very strong the agent cognitive learning model proposed by our research. Finally, we hope this research could make people reconsider the importance of self-awareness in the design of self-learning agents.

Keywords: Self-awareness, Self-Schema, Intelligent Agent, Artificial Society, Social Network, Emergent Behavior, Conflict

誌　謝

研究所這兩年感謝孫春在教授不斷給予啟發引導，不論是心靈的啟發、學術視野擴展以及對論文鉅細靡遺的指導，都讓我體會到真正做研究的方法與態度。論文口試時，很榮幸請到胡毓志教授、許永真教授、陳一平教授擔任口試委員，使我了解自己論文中的不足之處。感恩兩年來其他指導過我的老師，促使我學問與專業知識更進一步。

非常慶幸能加入孫老師的學習科技實驗室，在這裡遇到許多友善又學識豐富的學長姐們，在此特別感謝黃崇源學長，由於他的細心指導，使得我的研究受到口試委員們的認同與讚賞。感謝端木、其煜、冠傑、以及其它實驗室同學的建議以及學弟妹們熱心幫助，讓我深深體會實驗室的溫馨與熱情。

求學期間，最感謝的就是奶奶、父母與雅琪給我的支持與鼓勵，讓我無後顧之憂能夠對研究全力以赴。在這段甘苦交雜的日子裡，當我難過、悲傷、高興、快樂的時候，也有一群好室友世結、楀鴻、明祥陪伴著我，激勵我，我才有今天。

這段日子所得到的歡樂、幫助、知識等，用言語無法完全述說。這一切以感恩的心，對所有關心我的人，幫助我的人，致上最崇高的敬意，祝大家一切平安順利。

Contents

List of Tables

List of Figures

中文附錄目錄

附錄表目錄

附錄圖目錄

1. INTRODUCTION

Self-awareness is a kind of experience using the attention point to self (Duval & Wicklund, 1972). The eastern and western philosophers and the psychologists study the topic of self-concept year after year. Nowadays, self-awareness has become the central issue of Cognitive Science and Educational Psychology (Gardner, 1993; Carver & Scheier, 1981; Wickland & Frey, 1980). A Human being who has the correct ability of self-awareness outwardly can understand right or wrong, good or bad, and superior or inferior of self behavior and inwardly can realize self own emotions, motives, interests, and wishes to reach to self-identity and self-actualization further (Chen, 1996; Zhang, 1980). In other words, a man who lacks self-awareness is unfavorable to subsistence. For example, a man who lacks the emotional self-awareness and whose behavior is always influenced by his/her emotions cannot deliberate the consequence of his/her behavior. On the contrary, a man deficient in the comprehension of others' emotions usually over-reveals his/her advantage and is hard to understand that he/she should be more reticent to avoid offending others in some situations. According to these cases above, in the development of human-intelligence, self-awareness plays an extremely important role. Therefore, how to effectively assist human in raising the degree of the ability in self-awareness is always the central topic of research for scholars (Kondrat, 1999; Aronson, 1995).

Using computer simulation to analysis various complex systems has been a research trend. From simple activity to complex behavior, it provided a new approach to explain the social development and the economic system (Axelrod, 1997; Archer, 1995; Kontopoulos 1993). Computer Simulation and Artificial Society has been developed so far, mainly use the various learning models of Artificial Intelligence (e.g. Machine learning、Neural Network、Evolutionary Computing) to imitate every thinking methods and behavior models from humans, and to establish the intelligent agents (Wooldridge & Jennings, 1995). In these models, we conclude that the learning focus of traditional artificial intelligence only concerned with the outer environment, in other words, the outside world is just the whole thing of learning. It does not discuss the inner operation; the attention just put in the relation between outside incentives and behavioral responses. Point at the environment and the society; through constantly adjust and adapt to learn various skills and work strategies. Obviously, the learning concept putting attention on self, that is, human self-awareness is always ignored in the research territory of learning agents and intelligent agents. After all, whether self-awareness mechanism benefits intelligent agents or not is the main focus of this research.

If the learning agent has its own ability of self-awareness, it will be helped for 1) agents that can self debugging; 2) agents that help us to debug them; 3) new methods for machine learning; 4) fraud detection; 5) the interaction and communication between agents and humans can be more reality; 6) Computer Simulation and Artificial Society can be more closer to real world situations (McCarthy, 2004; Minsky, 2003; Mason & Sloman 2001). So, if the mechanism of self-awareness ability can be effectively incorporated into the structure of learning agents, we believed that a new research direction (approach) of artificial intelligent agents would be initiated and that the intelligent agent will get closer to the human high-level intelligence. Also the analysis and results of the simulation experiments appropriately using the technology of multi-agents will be more authentic and convincing.

This research proposed a new agent-cognition learning, called ACLM, to improve the deficiency of traditional AI learning approaches that only focus on World Model by the inner learning concept of self-schema. Furthermore, we instance the artificial societal conflict problem between public good and private interest that results from both the agents' environment and goal to proposed agent self-awareness model which is consistent with above cognition learning model. Also we expect to discuss how self-awareness resolves the collectively irrational behavior causes of individually rational behavior and prove the validity and stability from the analysis of individual performances and collaborative behaviors.

There are three main objectives in this paper; [1] to verify the ACLM is feasible; [2] to analyze how self-awareness improves agents' performance; [3] to prove the independence and compatibility of ACLM. First, according to the conflict between agents' self-interested and societal goals, we proposed an agent-learning model, which regarded superego as self-aware goal to prove the learning concept of agent cognition. Second, by analyzing the individual performance and collective behavior, we can understand how the agents' efficiency be improved and how the agents behavior performance be affected by the self-awareness. Third, from the conclusion of 1 and 2 above, we proved that the ACLM structure could be independent of and compatible with the former structure by agents from other designers.

2. BACKGROUND

2.1. Self-Awareness

Self-awareness is an attention which be focused either on the self or on external environment, when attention is self-focused we think more about trait, attitudes, and feelings and behave more consistently with them; self-awareness can be triggered by the scrutiny of others or by stimuli (such a mirrors or video cameras pointed at us) that trigger self-focused attention (Duval and Wicklund, 1972). From the consequences of self-focused attention, we compare our behavior to internal standards; when behavior doesn't match standards, we feel uncomfortable; this discomfort leads us either to change our behavior to match standards or to escape from self-directed attention. In 1995, Aronson suggested people the methods to self-recognize, including introspection, self-observation, self-schemas, and social interaction.

After collecting various scholars' literatures, we proposed a self-understanding circuit (Self → Self-schema → Self-awareness → Self) as shown in Figure 1. John, for example, supposed that he is a heterosexual. However, he is a homosexual in his subconscious. How could he understand and identify the real himself? Through our circuit, John could recognize that his real sex distinction is a homosexual by the awareness to self gender schema, and then would correct his own idea further. According to the examples above, we could clearly understand the relationship between these three concepts. Therefore, we regard the circuit as the basic process of self-awareness achievement for the learning agent (The details of self-awareness, please see appendix A.).

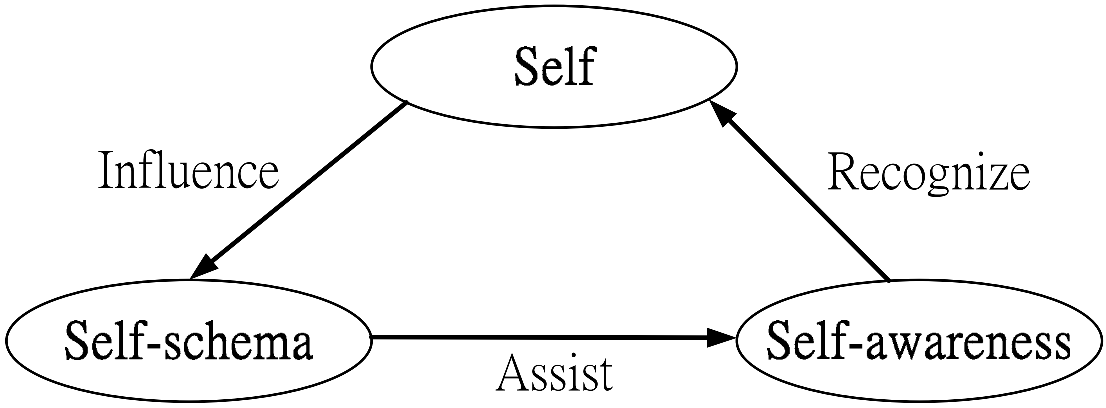


Figure 1: Relationships among Self , Self-schema, and Self-awareness

2.2. Classical Learning Agents

The agent intelligence will affect the capability to finish works. According to the formation of agent intelligence, it has three approaches which are Rule-based, Knowledge-based, and learning.

Because the rule-base and knowledge-based approaches needed to be given knowledge and experience from designers and experts, the deficiency is not easy to be modified and renewed (lack customization and learning flexibility), at the moment, using learning approach with built-in knowledge as the main topic of intelligent agents. This approach mainly use the way of sense-plan-act to proceed (Figure 2). Take the famous reinforcement learning model (Sutton & Barto, 1990, Figure 3) for example, the agent percept and collect outward signals, then planning a reasonable response through cost estimation and experiment accumulation, and at last, from proper rewards and punishments, try-and-error constantly to adjust the performance. Although different scholars have different opinions in agents learning structure, the interaction with environment has been identified and agreed.

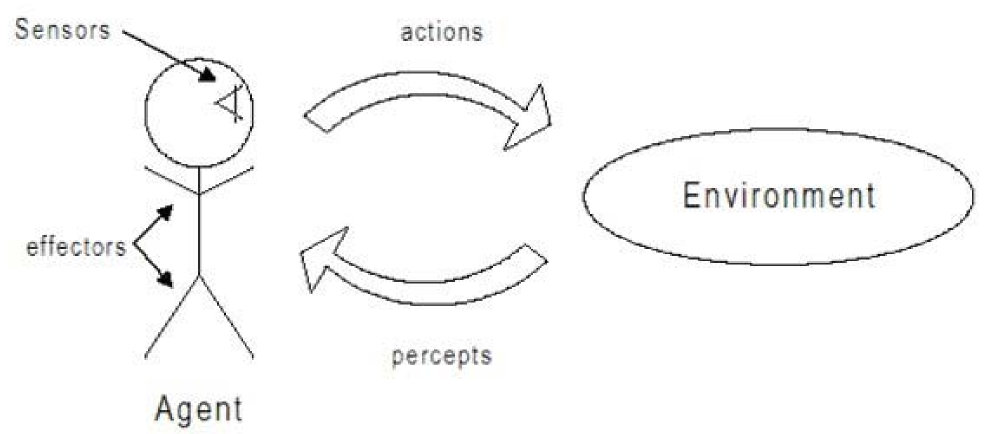


Figure 2: Classical AI Learning Approach

Anything that be viewed as perceiving its environment through sensors and Planning action based on knowledge or experience ,Finally acting upon that environment through effectors.

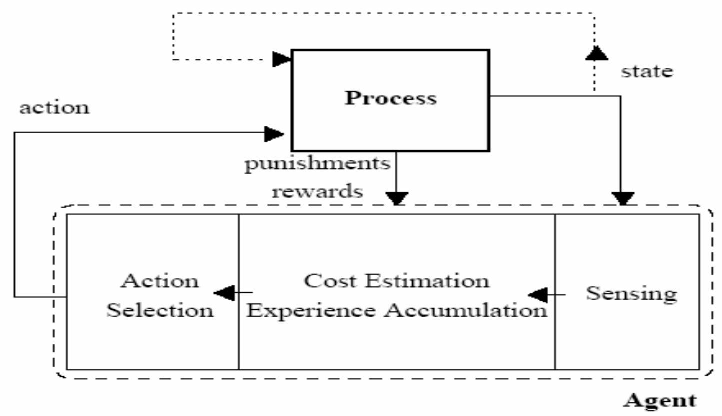


Figure 3: Reinforcement Learning Agent

Interaction between agent and dynamic process in Reinforcement Learning. The agent observes the states of the system through its sensors and chooses actions based on cost estimates which encode its cumulated experience. The only available performance signals are the reinforcements (rewards and punishments) provided by the process.

After Russell integrated the learning components of scholars, he proposed a general model of learning agents which can be divided into four conceptual components as shown in Figure 4. The **learning element** is responsible for making improvement, and **performance element**, which is responsible for selecting external actions. The **critic** is designed to tell the learning element how well the agent is doing. The last component of the learning agent is the **problem generator**, which is responsible for suggesting actions that will lead to new and informative experiences. So far, it presented completely the relation between learning components and environment for agent designers.

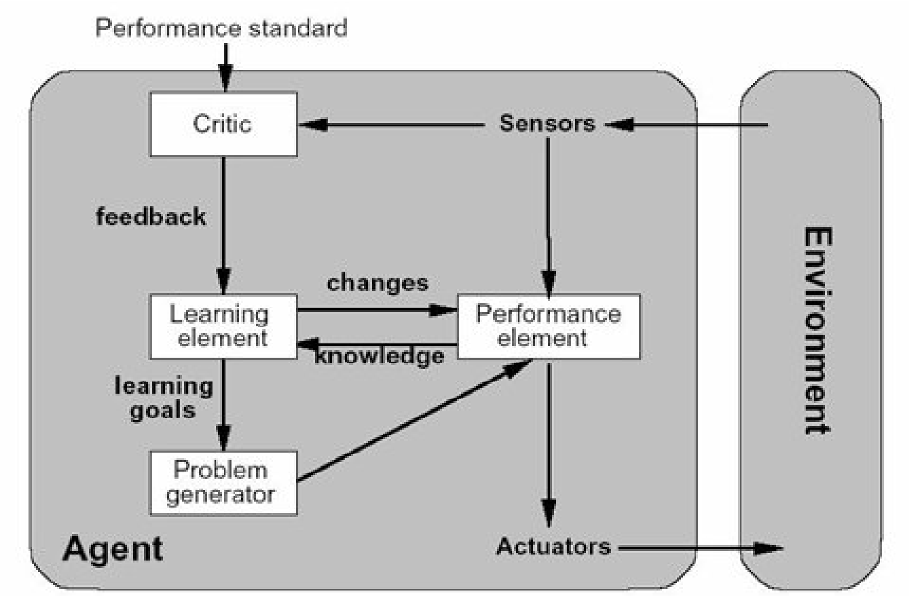


Figure 4: A General Model of Learning Agents

Russell's model had been studied completely, it explained clearly about all components of learning agents which be concerned by designers. Therefore, we use his model as a contrast paradigm. In next chapter, we will discuss the weakness of this model. (That is the part of all current agent designers didn't see.)

2.3. Social Games with Social Networks

Using computer simulation to account for the complex behavior of biological, social, and economic system has been the motivation of much interdisciplinary works in last decade (Haken, Poston, Stewart, 1978). In particular, the emergence of altruistic or cooperative behavior is a favorite problem of game theoretical approaches (Maynard

Smith, 1982). In this background, the Prisoner's Dilemma game (Axelrod & Hamilton, 1981) has been widely studied in different version, as a standard model for the confrontation between selfish and cooperative behaviors.

In a classic version of a PD game, two players decide whether each move they make should be one of cooperation or defection. A payoff is given to each player according to their combined moves. Table 1 shows a typical payoff matrix, including value constraints.

TABLE 1: PRISONER'S DILEMMA PAYOFF MATRIX AND CONSTRAINTS

Player B / **Player A**	**Cooperation**	**Defection**
Cooperation	**R**eward / **R**eward	**S**ucker's / **T**emptation
Defection	**T**emptation / **S**ucker's	**P**enalty / **P**enalty

Note: **T**emptation > **R**eward > **P**enalty > **S**ucker's, and

2x(**R**eward) > **T**emptation + **S**ucker's, or simplified:

T > **R** > **P** > **S**, and 2x**R** > **T**+**S**.

If both players know that they will play a PD game one time only, it is to their benefit to continually make defection moves in order to achieve a maximum outcome. If they know that they will play many games (a situation referred to as the Iterated Prisoner's Dilemma, or IPD), mutual cooperation is a better strategy for both players. Since most real-world dilemmas are iterated, the IPD is of great interest to researcher (Axelrod, 1984; Hoffmann, 2000). In the investigations of model behavior, there are several famous strategies (e.g. ALL-C: always cooperate、ALL-D: always defense、TFT: repeat what your opponent does in the previous round、PAVLOV: win-stay-lose-shift) that usually be discussed to explain the emergence of complex behavior and how to achieve the equilibrium of dynamic evolution (Bendor & Swistak, 2001).

IPD is usually implemented in zero dimensional systems, where every player can interact with any other. It has also been studied on a regular lattice (high clustering) called Cellular Automata, where a player can interact with its nearest neighbors in an array (Nowak & May, 1992). In a regular lattice the concept of a *k*-neighborhood is straightforward. It is composed of the *k* nearest individuals to a given one. However, social situations are rarely well described by such extreme networks. The topology of social communities is much better described by what has been called small-world networks (Watts & Strogatz, 1998). In the version of small worlds that we used in this thesis, the regular *k*-neighborhood of an individual is modified by breaking a fraction of its *k* original links. An equal amount of new links are created, adding to the neighborhood a set of individuals randomly selected from the whole system (high clustering & low separation). Figure 5 & 6 are key characterizations of small-world phenomenon.

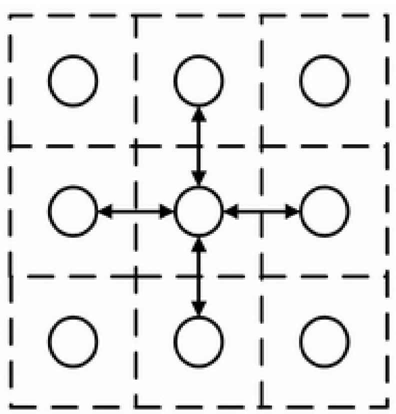


Figure 5: High clustering

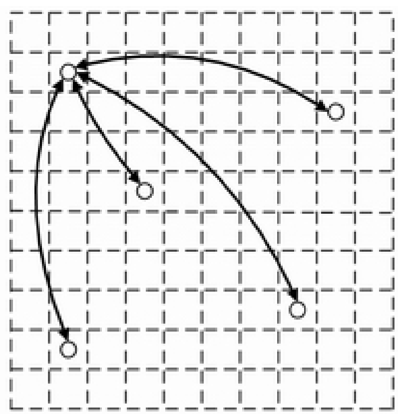


Figure 6: Low separation

3. AGENT COGNITION LEARNING MODEL

According to the viewpoint of World Model, we discuss the deficiency of A General Model of Learning Agents proposed by Russell. We have to deliberate the importance of self-learning ability in order to make agent get closer to human real thinking intelligence. After completely analyzing and collecting, we proposed a Cognition Learning Model which based on using the self-schema to be agent's internal learning focus and which could be compatible with the former agent system. By the model, agent pays learning attention to not only the World Model but also Self-Schema. Because of the achievement of inner learning by the awareness of self-schema, agents could be closer to human thinking intelligence. Also the model offers a brand-new designing concept to solve the high-level intelligent problem which the agent of traditional world model cannot solve.

3.1. The Proposed Model: ACLM

World Model Learning is a concept which put the learning focus (**Attention**) at the outer environment. Basically the outer world is the entire of learning. Actually it focuses on the association between outer incentives and behavioral responses but not discuss about the internal operations. After continuous adapting to environment and world, agent could acquire various skills and working strategies. Some common agent learning models (e.g. Reinforcement Learning、Neural Network、Evolutionary Computing、Classifier System、Decision Tree), for example, we could explore that all of them used World Model to be the learning focus (means entire) from these famous models of traditional artificial intelligence.

To discuss further into the example, Evolutionary Computing didn't include the viewpoint of internal learning. All its learning focus was put on the process for environment interaction. Through the pressure of existence given by evolution, it imitated other agents persistently, and through the process of crossover, mutation, and selection, it adapted to the environment, that is to say, **Learning equals Adaptation**. Although these World Model learning methods have good outcome, it's still not enough to simulate such human thinking intelligence on these models. In addition, more and more scholars started to research agent's emotion and mental state, agent's internal learning has to be regarded that is to pay attention on the self learning methods.

As the result of the analysis to World Model, the current designing concept has to be modified. Based on Russell's general agent model, we proposed a new agent cognition learning model (Figure 7). The model is composed of three elements (Performance element、World Model Learning element、Self-schema Cognition element). Performance element, such as Russell's design, is responsible for selecting external actions. The World Model Learning element is in charge of integrating the traditional learning components which only focused on outer environment to improve the learning efficiency. It takes some knowledge about the learning element and some feedback on how the agent doing, and determines how the performance element should be modified to do better in the future. Finally Self-Schema Cognition element, which is according to the past experience to get information into a knowledge structure, could help agent to understand, explain, and predict self-behavior. Because of the model design, World Model Learning and Self-Schema Cognition can coordinate each other, and present the most favorable method to improve performance. Eventually agent will be in possession of both external and internal learning concepts.

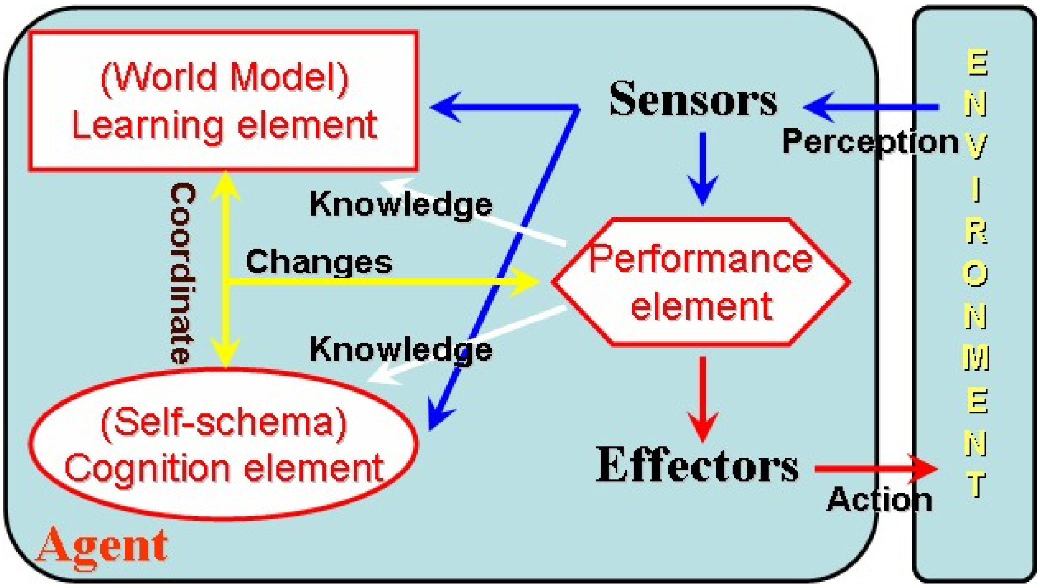


Figure 7: Agent Cognition Learning Model

The Agent cognition Learning Model we proposed could offer a new agent designing framework. Agent will possess external learning and internal cognition concepts of both World Model and Self-Schema. Therefore, the agent could self-discovery from self-awareness by adding various schemas that designers needed, and could effectively improve and promote the efficiency through the co-ordination between external learning and internal cognition, and then get closer to human thinking intelligent model.

3.2. Independence and Compatibility

Now the independence and compatibility of Agent Cognition Learning Model will be discussed in this section. If the model we proposed cannot be compatible with the former agent system, that is to say, we have to change the former agent design model before operation, it is unreasonable and uneconomical. Even the model is a perfect and complete one still cannot be accepted by the public.

Basically we separated the World Model and Self-schema completely based on the viewpoint of the external learning (learning theory) and internal cognition (cognition theory) in ACLM. Russell's model for example, we combined all World Model related components into the World Model Learning element, and embedded the self-schema cognition element which is independence of World Model Learning element in addition. Therefore, designers who hope to implement our ACLM in their own models don't have to re-modify the whole original learning structure but embed the part of self-schema cognition design. That proves, no matter what World Model learning approach the designers applied, the ACLM would be compatible with all these agent systems as well.

Through the internal cognition concept proposed by ACLM, agents will have the human self-awareness ability. Self-awareness can help agents inwardly self-understand and self-identification and outwardly learn various important skills or change the working strategies. In that event, during the research betweens the computer simulation using agent-skill and artificial society, agents will be able to present human high-level complex internal thinking behavior. Moreover, the result of simulation will be more realistic and persuasive. In next chapter, we will experiment with ACLM by using a general conflict problem in artificial societies, so that the experimental results can help us to verify ACLM, which can be implemented and is compatible with former agent system.

4. EXPERIMENTAL DESIGN

The environment that the agent exists still has a lot of other agents, thus if the designers occupy the most of the resources in the society or constantly get profit by defecting other agents in order to making his agent obtain the best performance, it will cause the harmful influence to the whole social benefits. According to the conflict between agents' basic societal and self-interested goal, we proposed an agent learning model, which regards superego as self-awareness achievement. This agent who owns self-awareness can make its life better (first getting private interest in the environment) and then make other agents live better (through public good to achieve the private interest) by the cooperative behavior at the same time. This is not only a reasonable ability for intelligent agents but also a rational behavior which conformed to the real world. It has been completely described as the strategic agents based on Iterative Prisoners' Dilemma. For these reasons, we adopt IPD with social networks which corresponding the conflict problem above as our research platform. By analyzing this experiment, we expected to recognize the learning agent with self-awareness, how will it effect the performance results? Does it help them to expedite the cooperative behavior early?

4.1. Simulation Model

The Simulation Model (Figure 8) uses the two-layer concept which is playing IPD game with social networks to be the research platform. The upper layer is IPD which adopted the Evolutionary Computing approach, and the lower layer is social networks which used two famous topologies, Cellular Automata and Small-World Network, to experiment respectively.

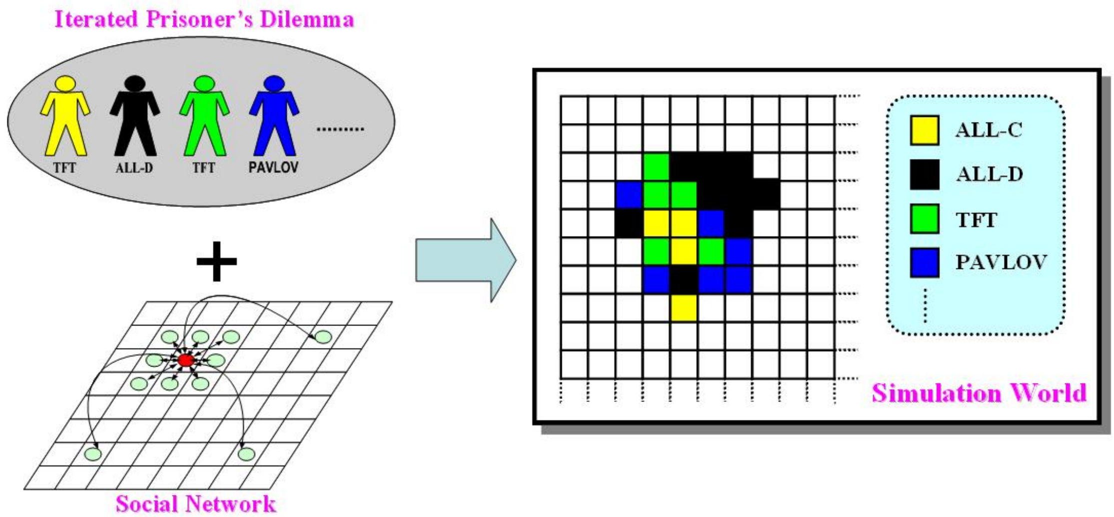


Figure 8: Simulation Model

Each upper agent adopts pure strategy, which means using the same policy to all coworkers. Besides, the Memory-1 deterministic strategy on its memory ability, there are 16 strategies can be chosen. Furthermore, in order to observe the emergent behavior of strategic agents, each agent has its own corresponding and unique color. At lower layer-Social Networks, the creation method of Cellular Automata adopted 2-D spatial relation, that is, each agent establishes links to contact with other surrounding cells. For example, if we extend outwardly k steps to establish the links of Cellular Automata, it is called radius-k neighborhood, which has $(k+2)^2-1$ nearest coworkers. After that, the version of small-worlds that we used in this experiment, radius-k neighborhood of an agent is modified by breaking a fraction of its $(k+2)^2-1$ original links. An equal amount of new links (shortcuts) are created, adding to the neighborhood a set of individuals randomly form the whole system.

The simulation world presents as 2-D lattice, each cell represents a strategic agent (or called an individual), and each agent just can occupy a specific cell. The colors of cells represent the corresponding strategy of agents. So, if there are two agents with the same strategy, that is, the colors in the simulation world will also be the same. Through this elementary relationship between space and color, we can recognize how the learning agents achieve their self-interested goal by constantly adapting to the environment. Besides, from observing the interactive process of cooperating or defecting, we can understand how to reach the evolutionary equilibrium and how to emerge the complex behavior clearly (The details of simulation model, please see appendix B.).

Evolutionary steps as follow; the corresponding flowchart as shown in Figure 9:

Step 1. Set up the environment parameters (including color mapping of strategies, social network parameters, interaction rules) and evolutionary parameter (population size, selection rules、mutation rate and rules、crossover rate and rules).

Step 2. Generate randomly the populations and establish two kinds of social networks.

Step 3. Select the coworkers.

Step 4. Calculate the scores (fitness) with coworkers.

Step 5. If all of the partners have been selected, then go Step 6; else go Step 3.

Step 6. Through the evolutionary pressure to select the individuals who are not suitable in the environment, and to execute the processes of Mutation and Crossover in order to generate a set of candidates for next generation using.

Step 7. Randomly select some candidates to evolve next generation.

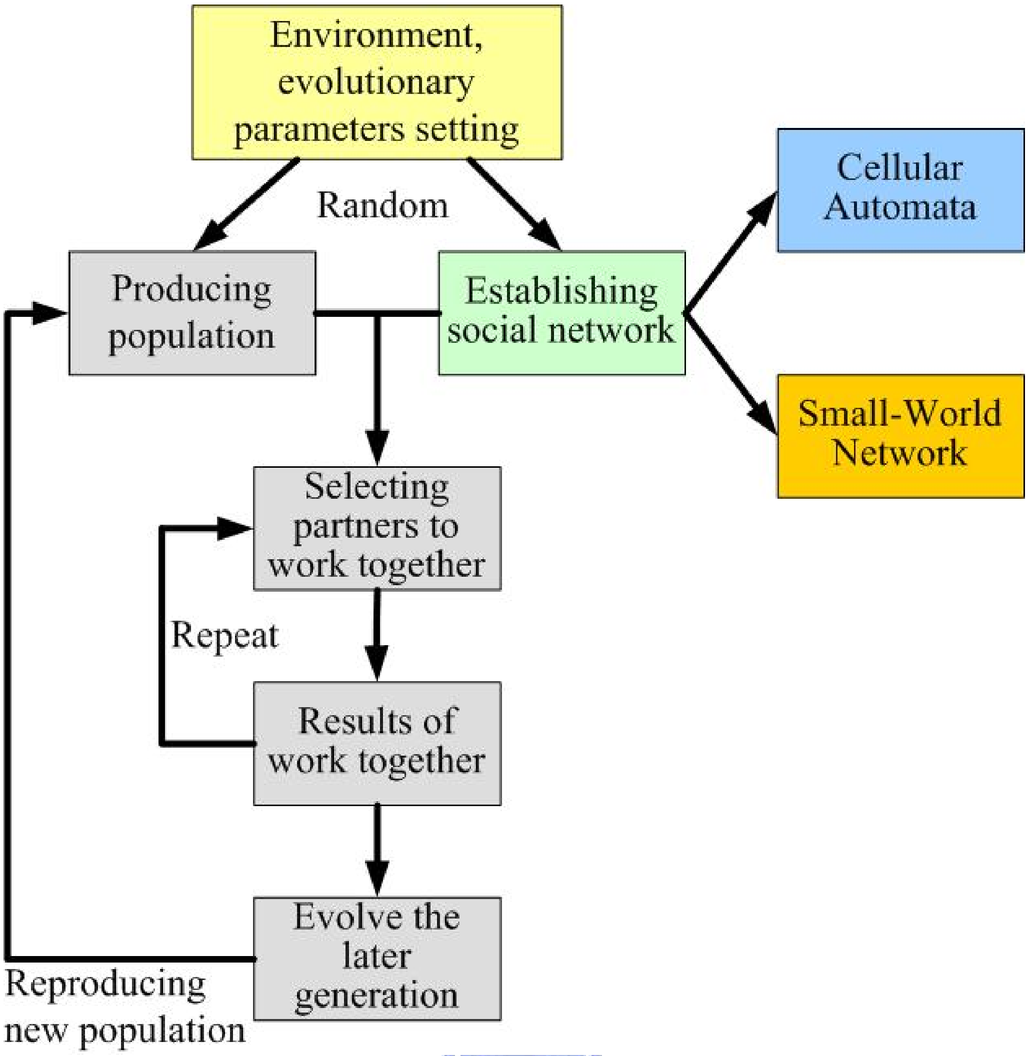


Figure 9: Evolutionary Procedure

4.2. Agent Self-Awareness Model using Superego Idea

After analyzing the personalities of intelligent agents and learning agents in the traditional AI by Sigmund Freud's Three Components of Personality, we found out that the agents have no concepts of id and superego (Figure 10). To discuss furthermore, if the agents have the idea of superego, it will be helped to understand the expectation in society and to make the collaborative behavior to emerge earlier. Thus, using superego as the aware goal, it will be a resolution in the conflict between public good and private interest in artificial societies. By the analysis above, we correct the personality model of agent (Figure 11) and propose an agent-learning model which regarded superego as self-aware goal according to the concept with external learning and internal cognition in ACLM. It means this model is put into Self-Schema Cognition element in ACLM. At last, we expect the agents could concern about the rational behavior both in public good and private interest to solve the conflict in artificial societies to prove that the ACLM we proposed can be achieved. (The detail of agent personality model, please see appendix B.)

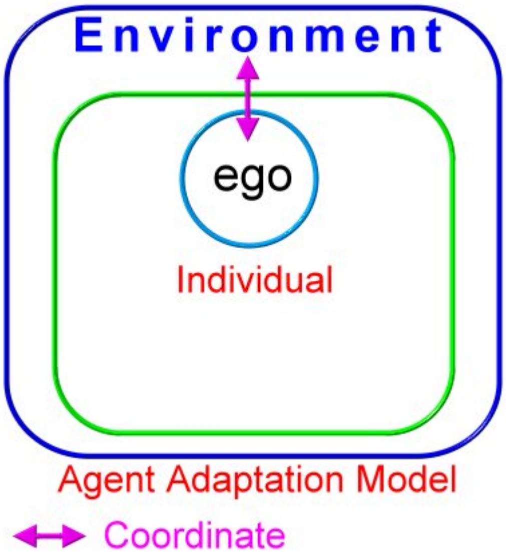


Figure 10: Agent Adaptation Model

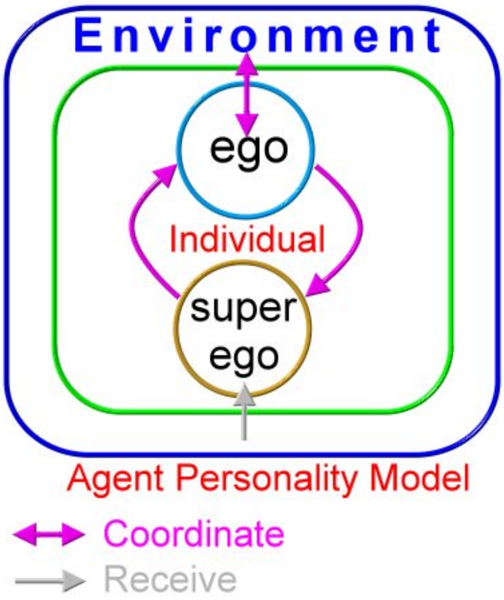


Figure 11: Agent Personality Model

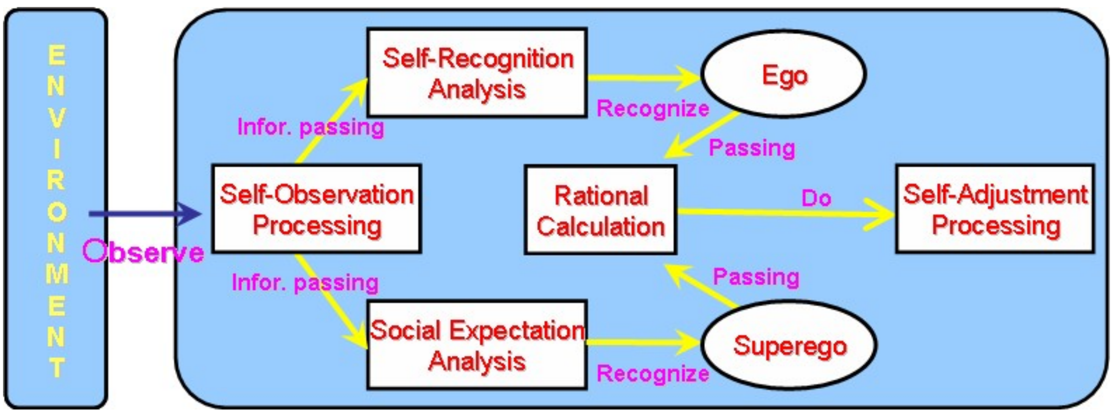


Figure 12: Agent Self-Awareness Model

According to the Figure 12 above, our superego awareness consisted of four steps: self-observation → self-recognition and social expectation analysis → rational calculation → self-adjustment. First, the self-aware agent collected information by observing the social interaction and self-status. Second, the agent recognized the ego and superego by using the recognition and expectation to self from coworkers. Third, through the rational calculation to decide the self-adjustment will be needed or not. Forth, correct the differences between ego and superego. (The detail of implementation, please see appendix C.)

Figure 13 is the flowchart of experiment group, the left part is elementary evolutionary procedure (control group), and the right one is the superego awareness model (experiment group) which we proposed.

<u>Experiment steps as follows; the corresponding flowchart as shown in Figure 13;</u>

Step 1. Set up the environment parameters (strategy color mapping, social network parameters, interaction rules), and evolutionary parameters (population size, selection rules、mutation rate、crossover rate).

Step 2. Generate randomly the populations and establish two kinds of social networks.

Step 3. Select the coworkers.

Step 4. Calculate the scores (fitness) with coworkers.

Step 5. According to the evaluating rules, then give the reputation to coworkers.

Step 6. Go Step 3 (repeat the game, till all of the coworkers have been selected, then Go Step 7).

Step 7. Collect the recognition to the agent from the coworkers. (**Reputation**)

Step 8. From the social expectation analysis to find out the expectation to the agent form the coworkers. (**Social expectation**)

Step 9. Through the rational calculation to compare with the matching degree between reputation and social expectation on agents, if it under the threshold then do nothing; else do Self-Adjustment. (The detail of Self-Adjustment, please see appendix C)

Step 10. Use the Self-Adjustment to select a suitable social expected strategy.

Step 11. Select the candidate agents as the next generation, and return to zero from the reputation and expectation.

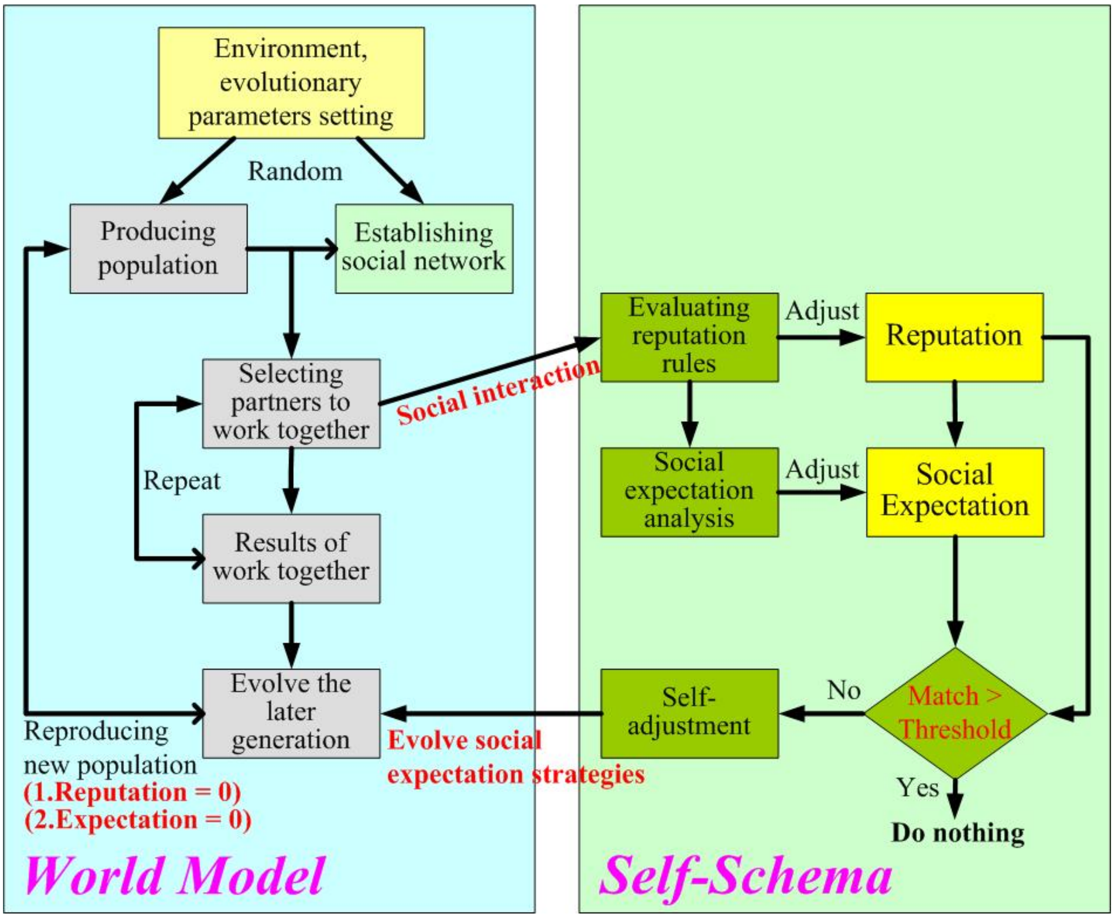


Figure 13: Experimental Procedure (World Model plus Self-Schema)

5. RESULTS

There are two kinds of social networks, cellular automata and small-world network, in our experiments. The control group of both social networks is the elementary evolutionary IPD model that is no self-aware agents in the environment. In the experiment group, we add self-aware agents with the ratio 1.0, 0.5, 0.3, and 0.1 into the simulation world. The ratio 1.0 means there are overall self-aware agents in the environment, and others so on.

5.1. A few agents with self- awareness that can improve whole interest

Figure 14 presents the experimental results of first class social network topology. The five squares in the right side represent the ratios of mixing self-aware agents in the environment. By observing the black curve, we can find some interesting phenomena. In the beginning of evolution, individuals (strategic agents) choose their own strategies randomly to co-work with their partners. After several generations, the individuals want to maximize their own fitness, thus these individuals tend to betray their partners. When most of the agents turn their strategies to defect their coworkers, the society will sink into a vicious circle. And cause the whole social benefits decrease rapidly. The decreasing of social benefits will cause the private interest to decrease either. After more few generations, the agents turn to accept the strategies of cooperation. This phenomenon fits with “the mutual cooperation is a better strategy for agents in the iterative games.” in game theory. Finally, because of the mutual cooperation, the social benefits increases, and the society tends to an evolutionary equilibrium.

By observing the evolutionary dynamics of the strategic agents in the control group, we can figure out that the simulation model fits the results of rational analysis in the game theory. And the effectiveness of the simulation model is thus being verified. In next step, we will observe the experiment group to determine whether the whole social benefits can be improved by self-awareness or not.

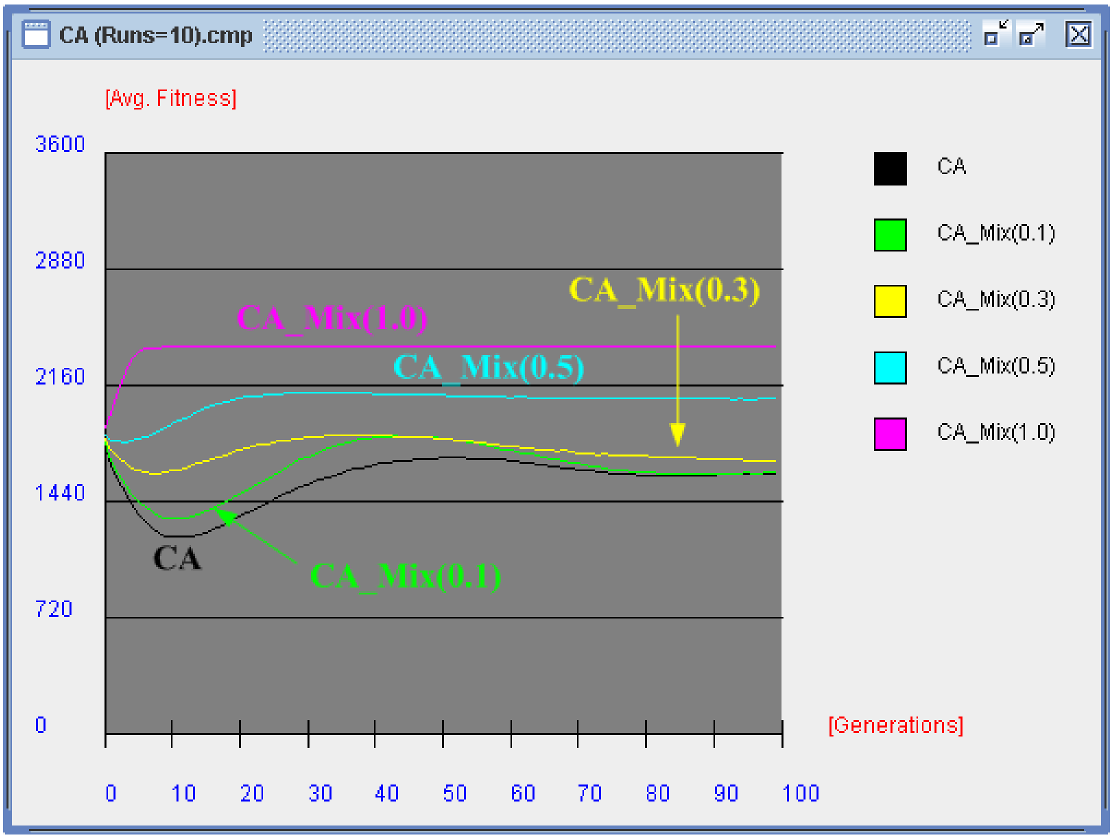


Figure 14: Comparsion with mixing partial self-awareness agents in CA

By observing the curve CA_Mix (1.0), we can figure out that if all the agents have the abilities of self-awareness, the whole social benefits will be improved and will not sink into a vicious circle, which causes an immediate evolutionary equilibrium. The vicious circle means the agents do not trust each other and the whole social benefits decrease rapidly. This experimental result seems to be perfect but actually not practical. Because we want to add the self-aware agents into existing agent system, in other words, we want to add some self-aware agents into an agent society without any self-aware agents. And expect the few self-aware agents to make the behavior of cooperation emerge earlier so that the whole social benefits and the private interests can both be improved. In this analysis, we will focus on the effects caused by a few self-aware agents in a given environment. Although CA_Mix (0.5), CA_Mix (0.3) and CA_Mix (0.1) may not reach a stable state as fast as CA_Mix (1.0) does, self-awareness do help the society to get rids of the chaos phenomenon earlier. It means the proposed self-awareness model do improve the whole social benefits. Take the curve CA_Mix (0.1) as an example, among ten agents, if there is only one agent has the ability of self-awareness, then the social benefits will be improved and so do the private interest of agents. It proves that, although there are few self-aware agents in cellular automata, they still provide an important influence.

Figure 15 is the experimental result of class 2 social network. This experiment is close to the real topology of interpersonal network. By the characteristics of high clustering and low separation in Small-World Network, we can realize how the proposed self-awareness model affects the real world situations. This experiment is essential because the existing agent-based societies are all combined with human nature (e.g. network economics and on-line games) and have the small-world properties. Thus, we need to understand the results of this experiment.

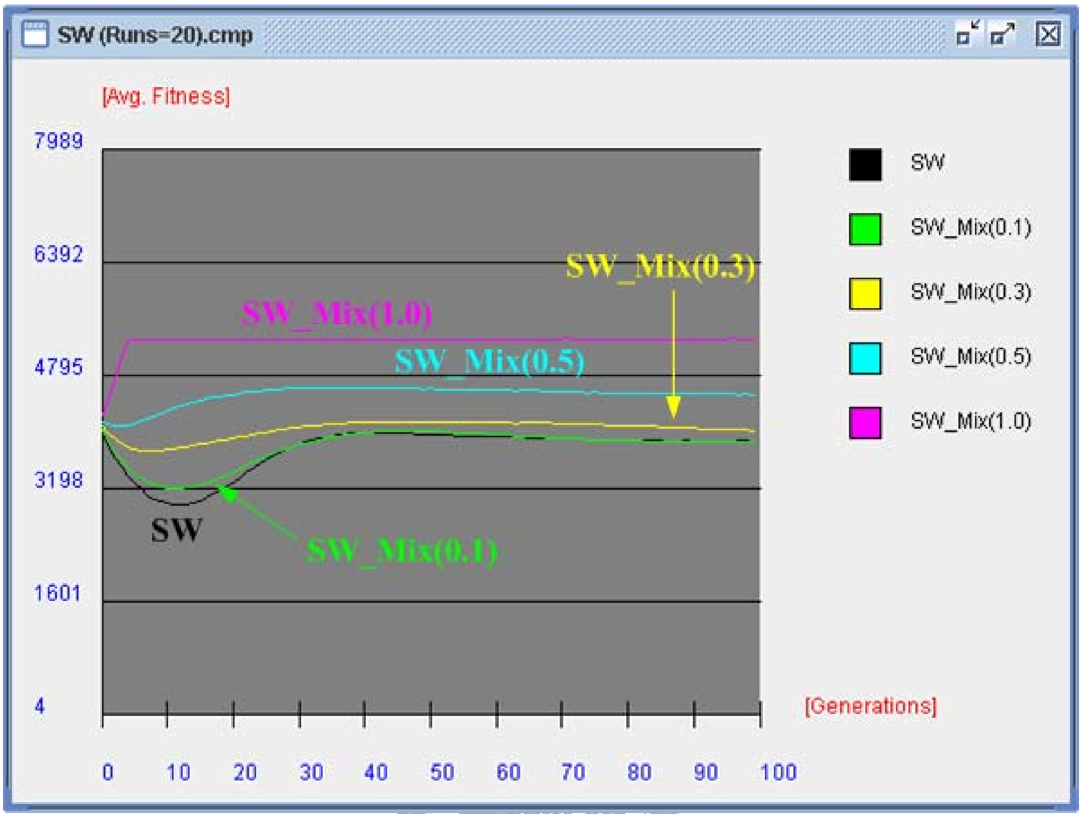


Figure 15: Comparison with mixing partial self-awareness agents in SW

By the figure above, we can realize that if only a few agents have the ability of self-awareness, the whole social benefits also can be improved and the evolutionary equilibrium can be achieved earlier. This fits the result of Cellular Automata and proves the effectiveness of the proposed self-awareness model. The experimental result also presents that the proposed model is practical to the existing agent societies. We even need not to modify the learning mechanism of agents. By adding the proposed self-aware agents, the behavior of cooperation can be emerging earlier.

5.2. Emergence of Social Behavior

There were 16 1-memory strategies agents in the experiment. In model behavior investigation of IPD, the representative strategies analyzed and discussed could be classified by 4 kinds, ALL-C, ALL-D, TFT, and PAVLOV. Among them, the ALL-C (always cooperate) refers to the ones which would always cooperate with their partners, no matter their partners cooperate with or betray them. On the contrary, ALL-D would always betray and take advantage of their partners. TFT refers to the "repeats the opponent's previous move" pattern, which cooperates with ALL-C but betrays ALL-D. TFT is the well-known good strategy in the IPD research. However, its flaw lies in that with the fail to synchronize its memory would cause its despiteful breach against STFT. The last one, PAVLOV, refers to the "vicar of bray" pattern. The rest of the other 12 strategies could be classified to these 4 patterns. Thus the relations among these well-known strategies are usually used by researchers to discuss complex social dynamics and equilibrium phenomenon. Thus our experiment analysis would take these four patterns as our samples.

- **Cellular Automata**

Figure 16 is an illustration of the reaction among the 4 well-known strategies in Cellular Automata. In the beginning of evolution, there was no significant difference in quantity. About 3 generations later, we find the quantities of ALL-D agents increased dramatically, while the amount of ALL-C and PAVLOV decreased gradually as a result of being invaded by ALL-D. TFT started to emerge when the quantities of ALL-D reached certain level. TFT would check and balance the growth of ALL-D and coexist with PAVLOV and ALL-C. After approximately 20^{th} generations, TFT would exceed

ALL-D in quantity. Under the pressure of TFT, ALL-D would decrease rapidly. During about the 30th generation, when TFT has grown to certain amount, the memory asynchronous problem of TFT against STFT started to emerge (thus began the vicious circle of despiteful breach). Then PAVLOV would start to increase, because it does not has the problem of failure in memory synchronization. In generation 60, amount of TFT had been less than ALL-D, so ALL-D started growing again. Meanwhile, PAVLOV would decrease as ALL-D increased. At last, in generation 80, TFT would exceed ALL-D again, and the artificial society would reach a dynamic equilibrium, in which amount of PAVLOV and ALL-C were kept stable (evolutionary-stable-strategy, ESS), while ALL-D and TFT checked and balanced each other.

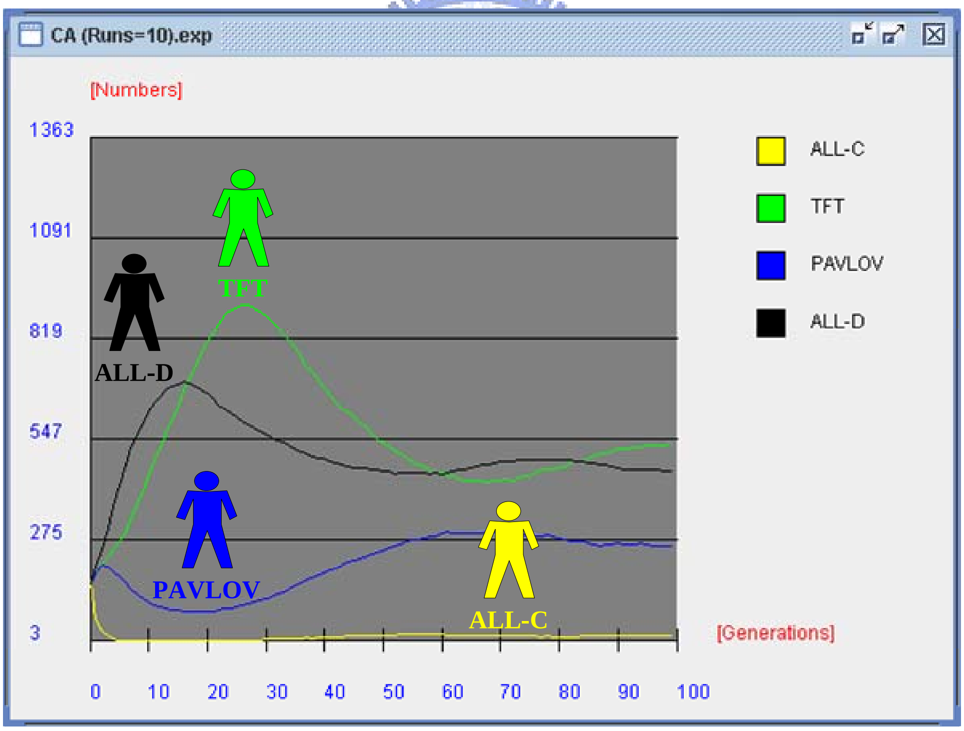


Figure 16: Four well-known strategies in Cellular Automata

Next, we will discuss our experiment group in Cellular Automata. As shown in Figure 17 and Figure 18 are four well-known strategies with mixing ratio 0.1 and 1.0 in Cellular Automata.

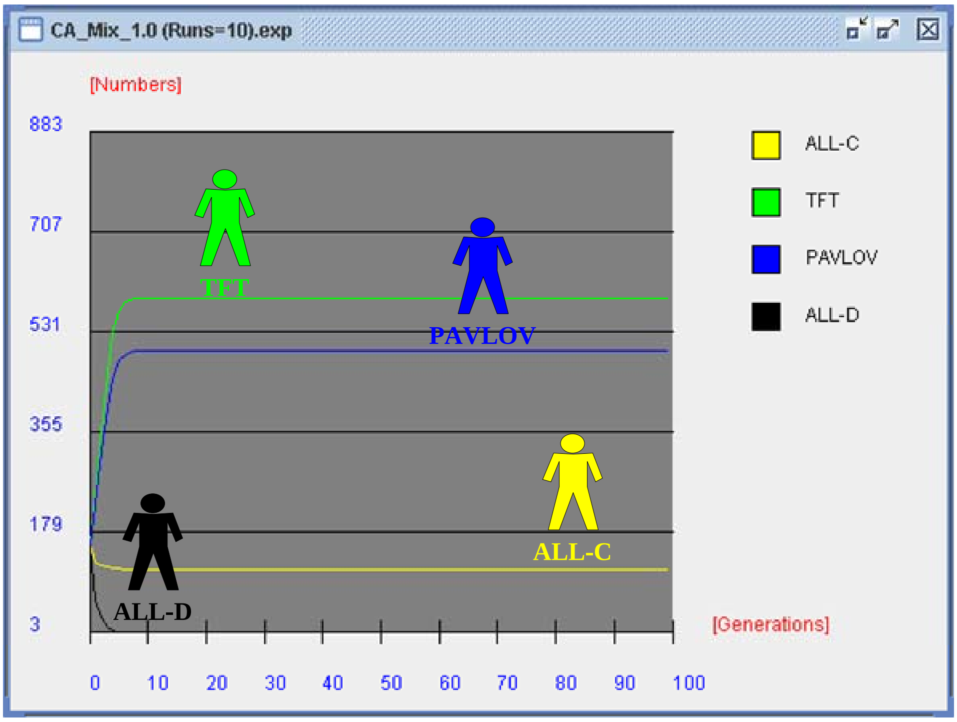


Figure 17: Four well-known strategies in CA
(Mixing self-aware agents with ratio 1.0)

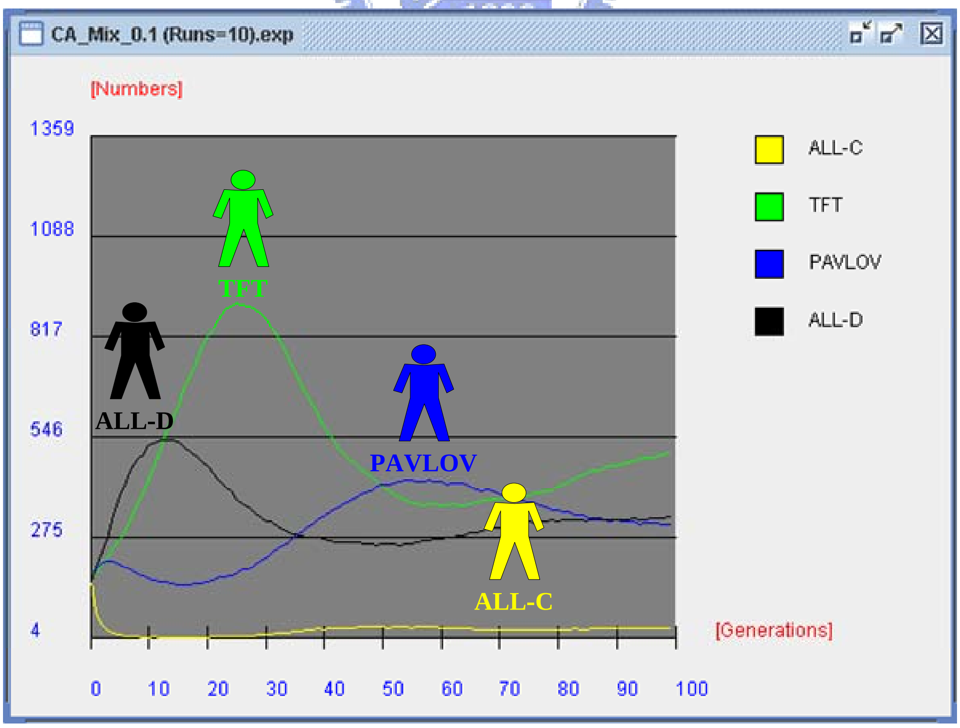


Figure 18: Four well-known strategies in CA
(Mixing self-aware agents with ratio 0.1)

According to Figure 17, we find that ALL-D disappeared at the beginning of evolution when Cellular Automata was filled with self-aware agents. Since ALL-D does not match social good expected strategy, our self-aware agent would discover that the existence of ALL-D is not permitted by superego. Thus the self-adjustment mechanism of self-awareness model would start. In order to meet the social expectation, strategy modification would begin. Therefore, in about 3rd or 4th generation, evolutionary equilibrium would be accomplished.

Figure 18 is the main point of this experiment, in which we put self-aware agents into Cellular Automata in the proportion of 0.1. Through our observation, we find that in the beginning of evolution, strategy ALL-D were not as vigorous as what we see in Figure 16 (control group). It has been controlled by self-aware agents. Furthermore, comparing the quantities of high peak (in about 15th generation) of ALL-D in these two figures, we find there were 700 ALL-D in the control group (There are 2500 strategic agents in the simulation world), the group without self-aware agent, while there were only 550 ALL-D in the experiment group, the group mixing self-aware agents with ratio 0.1. There existed an obvious gap of 150 in quantities. This is the key to accelerate the progress in getting rid of social vicious circle. Besides, another special phenomenon is that PAVLOV would exceed ALL-D and TFT in certain period of time and then decrease. That is, PAVLOV could not rival against ALL-D, with the increase of its quantities, it would be more easily invaded by few ALL-D.

- **Small-World Network**

Figure 19 were the reaction among the 4 well-known strategies in Small-World Network. In the beginning of evolution, the reaction would look like the one in Cellular

Automata. The key difference would not emerge until the 30th generation. ALL-D would thus reach an evolutionary stable, significantly lower than the quantities in Cellular Automata. The Shortcuts in Small-World Network would decrease the world separation, so the reaction among agents would get complicated, and the effectiveness of strategies would get stronger. Thus the evolutionary dynamics would be more vigorous and faster than the one in Cellular Automata. Under Small-World, ALL-D would reach its equilibrium in 30th generation, so the other 3 strategies would reach their evolutionary stable gradually.

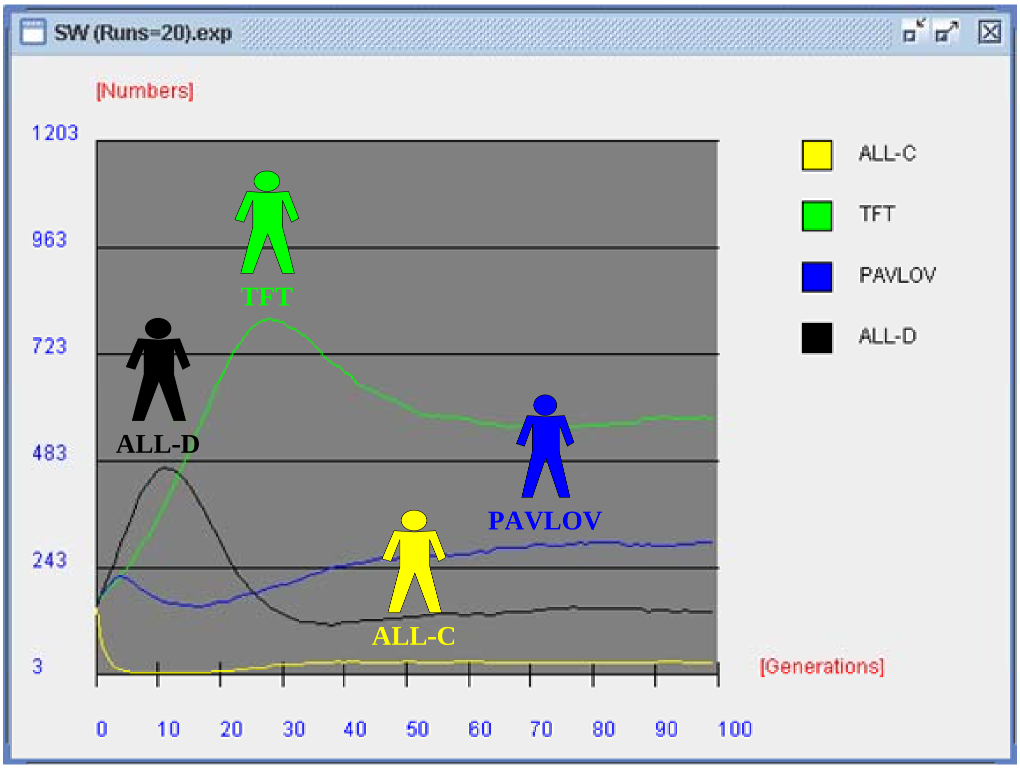


Figure 19: Four well-known strategies in Small-World Network

Figure 20 is the main point of Small-World experiment, in which we mixed self-aware agents with ratio 0.1 into the environment. After comparing the quantities of high peak of ALL-D in these two figures (Figure 19 and 20), we find there were 480 ALL-D strategies in control group and only 420 ALL-D in experiment group. There exists a gap of 60 for ALL-D. The gap is not as large as the one in Cellular Automata, in which there existed 150 for ALL-D; however, it is also helpful for getting rid of social vicious circle earlier.

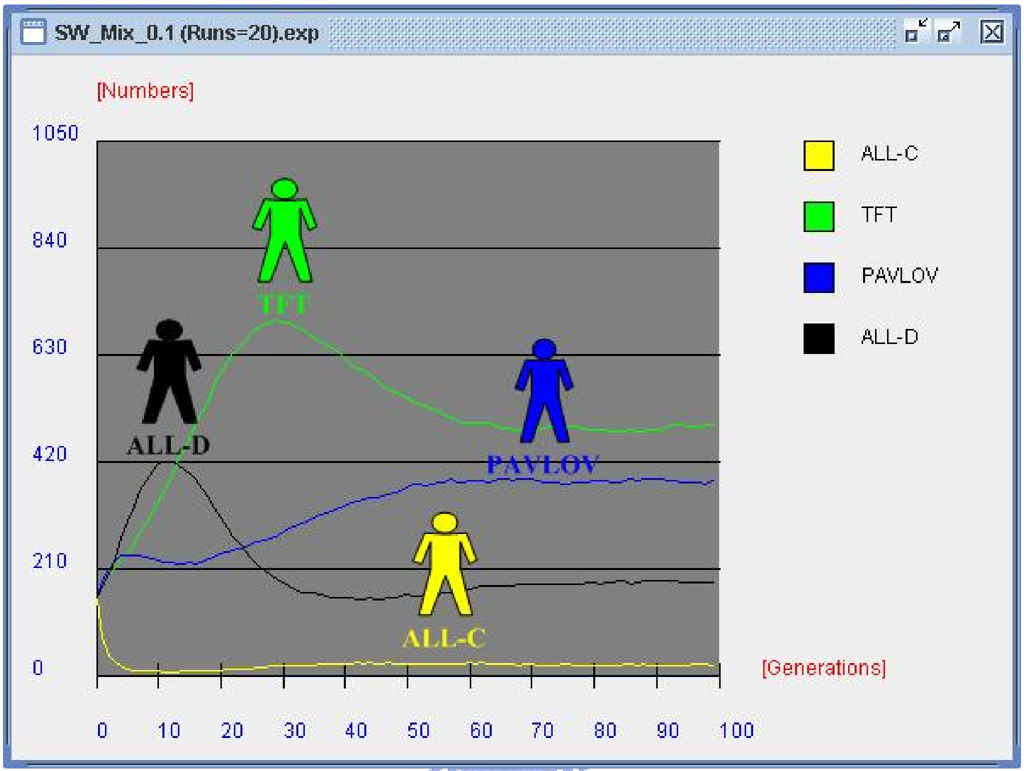


Figure 20: Four well-known strategies in SW
(Mixing self-aware agents with ratio 0.1)

After analyzing the emergent behavior of Cellular Automata and Small-World Network, we conclude that as long as a few self-aware agents, they can accelerate the progress in getting rid of social vicious circle and improved the whole social benefits certainly. This verified our self-awareness model in superego level would be a feasible and effective solution to the conflict between public goods and private interests. It also proved that our Agent Cognition Learning Model is feasible. (The details of all experimental results please see appendix D and E.)

6. CONCLUSION

The Agent Cognition Learning Model and Agent Self-Awareness Model in this thesis are related to 4 academic fields, including Artificial Intelligence, Cognitive Psychology, Economics, and Social Behavior. In Artificial Intelligence, we promoted the thinking level of agents to fix the flaw of both traditional intelligent agents and learning agents, whose learning focus were on the basis of World Model. In Cognitive Psychology, we established the personality model of agents, thus agents could achieve the goal of self-improvement through self-awareness. Our discussion through PD mathematics model on the conflict between public good and private interest in artificial society is involved with Economics. Our Superego Self-Awareness Model offered a good solution and verified the feasibility of ACLM. In Social Behavior, by observing the collaborative behavior of agents, we understood how they affect the operation pattern of former agent system.

Through the viewpoint of this thesis, we hope to clarify the importance of uniting internal cognition and external learning. We hope ACLM could offer a new approach for intelligent agents. We believe this approach is valuable and worthy striving for.

Appendix A　自我覺察相關文獻

自我概念（Self-concept）

自我的概念是人性的基礎，它會影響人類的一切行為(Tucker ladd)。根據存在主義的觀點，毛禮斯(Morris, 1954)曾經針對自我的意識指出：人不但能思想，而且能知其所思想(能批評、檢討、反省、糾正自己的思想)；人不但能感受，而且能知其所感受(感情中有理性，不為情感所迷惑)；人不但有意識，覺知到周圍的世界，而且更有自我意識(self-conscious)，覺知到自己在世界中的存在。

自我一經建立，將會引導我們的思維、感覺和行為(Smith &. Mackie,2001)。美國心理學創始者詹姆士(William James)針對人類的自我知覺兩元性(duality)曾指出：自我是由對自己的看法與信念所組成，或稱「已知」(known)，或更簡單的「我」(me)；自我也是個活躍的資訊處理器，叫做「知悉者」(knower)或「我」(I)。而詹姆士的「已知」也可當作自我概念或自我的定義，而自我的知悉者為察覺或意識狀態。也就是說，自我是一本書，也是這本書的讀者。前者充滿者長期累積下來的迷人內容，而後者則是一位在任何時刻都能自己讀取章節、任意增加章節的人。

自我覺察（Self-Awareness）

Duva and Wicklund (1972) 定義自我覺察是一種將注意力指向自我的體驗。國內學者張春興(民 79)認為自我覺察是個人對自己個性、能力、慾望等方面的了解。許玉佩(民 84)認為自我覺察的經驗包括了對知覺動作、感受、需求與價值觀等方面的了解，而此種了解與覺察的能力是由粗淺至深入、具體至抽象、外在至內在的一個發展過程。而陳金燕(民 85)的研究則將自我覺察界定為知道、了解、

反省、思考自己的情緒、行為、想法、人我關係及個人特質之狀況、變化、影響及發生的原因。而在陳金燕的定義中，指出自我覺察的焦點不再僅侷限於「自己」，也包含了「與自己互動的他人」；同時，覺察的內涵不再僅止於「靜態的」狀況，也關注過程中「動態的」變化及影響。

自我是知識的客體，但是要如何去深入的去瞭解自我？自我的印象是如何形成的？一直是心理學家及哲學家們研究探討的重要課題。Smith & Mackie 著的 Social Psychology 書中指出形成自我印象的方法與知覺他人的方法是很相似的。但是在對自我的感覺中，我們帶進了更多的偏誤。而人們往往從可自由選擇的行為中推論自己。自由選擇的行為是受內在動機所驅使的，即我們做自己想做而不是不得不做的事(Harackiewicz, 1979; Deci, 1970)。

歌德(Goethe) 曾說：「想了解自己，就去觀察你的周遭人之行為；想了解別人，就去探索你的內心」。因為我們觀察別人比觀察自己更客觀，了解他人可以增進自我了解，而自我了解有助於我們對別人的了解。而人們如何看待自己多受他人如何看待自己影響，因此別人對自己的看法可以幫助解釋個人對自我的錯誤理解，更可進一步幫助改正對自我錯誤的認知(Tucker ladd)。另外，Aronson(1995)建議人們認識自我可由以下幾種方法達成：經由內省(introspection)來認識自己；經由觀察自己(self-observation)的行為來認識自己；經由自我基模(self-schemas)來認識自己；經由社會互動(social interaction)來瞭解自己。

自我基模（Self-Schema）

自我基模是根據過去的經驗，將於自己認識之資訊彙整成知識結構，能幫助我們瞭解、解釋並預測自己的行為(Markus,1977;Markus & Sentis, 1982; Markus & Zajonc, 1985)。人們會透過記下一些他們認為既可反映其與眾不同之獨特性，又

具有跨情境一致性的核心屬性來建構一個統一而持久的自我感覺。而這些特性特質形成了一個基模(Markus, 1977)。人甚至在最平凡的行為上都尋找這些核心特質的證據，藉以增進穩固且統一的自我感覺(Cantor & Kihlstrom, 1987)。我們會利用自省與觀察自己的行為來了解自己，並將這些資訊組織起來置入自我基模中。我們也會利用那些指出我們的態度與行為可能如何改變的基模與內隱理論來理解我們的過去。

Appendix B　代理人之人格分析

佛洛依德之人格三我

佛洛依德(Sigmund Freud)認為一個人的人格結構(圖 1)有三個主要的部份，分別是本我(id)、自我(ego)與超我(super ego)，本我代表的是本能的驅欲，又叫「求爽的我」，趨樂避凶、心想事皆要成，都是這種心理在主導，例如食物與性等等，依循的是"享樂原則"；於此同時，超我則扮演著人格結構中偏向行為規範的那一面，又叫「道德的我」，禮義廉恥、忠孝仁愛、信義和平，全由它來提醒，依循的是"道德原則"，換句話說是外在規範的內化。而可以想見的，在一個人的內在，本我與超我經常是衝突的，此時便需要自我來居中協調。自我是在人格結構中與現實接觸的部份，又叫「理性的我」，依循的是"現實原則"，扮演調和的角色，換言之它會依遭遇的狀況來判斷何時可以滿足本我的需求，何時則需要接受超我的規範，此外它也會發展出許多的心理防衛機轉來平衡內在需求與外在現實之間的落差。舉個例子：假設走在街上，看到地上有一千元，本我最基本的想法就是將這一千元歸為已用，希望利用這不勞而獲的錢財來滿足自我的慾望，超我則會提醒我們，從小到大被教育「拾金不昧」的觀念，而自我則是撿到錢後，雖然不會交給警察但也不會立刻花掉，他可能會設定一個條件，就是萬一在三日內沒遇到失主，這筆錢的擁有權就屬於他，所以自我即是負責本我與超我衝突間的協調角色。

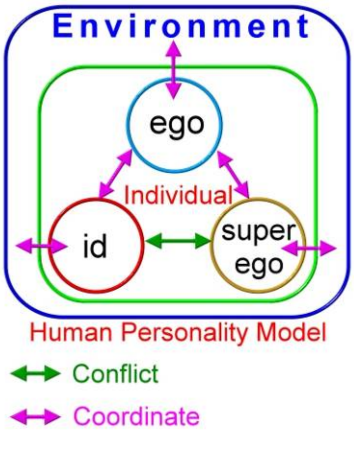
代理人調適模型

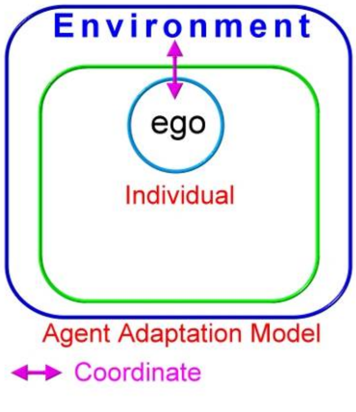
以佛洛依德的人格三我來分析傳統人工智慧中學習代理人與智慧型代理人的人格特質時，可以發現代理人的學習都是透過不斷地調適於環境，以獲得最理想的行為表現，換言之，學習等於調適。代理人本身並沒有本我(為了生存，代

理人並不會一直堅持同一個失敗策略)與超我(代理人並沒有道德規範)，它只擁有自我，也就是理性的我，用來幫助代理人達成生存目標而適應於環境中。因此我們將佛洛依德的人格三我稍做修正，以代理人調適模型(圖 2)來說明傳統人工智慧代理人的調適特質。

代理人人格模型

具自我特質的代理人(也就是現有代理人學習架構)，雖然能適存於環境，但卻不能為社會帶來最大利益。而具本我特質的代理人，可適用於個性化代理人(例如：電腦同伴、虛擬主播)，但缺點是個人私利行為或非理性決策會造成社會排擠的問題。最後探討的是超我，若代理人具備超我的道德觀念，可幫助了解社會期望，解決人工社會衝突，使代理人合作行為提早出現，並提升整體社會利益最有所提升。就解決代理人生存環境與自利目標的公益與私利衝突，可能是不錯的解決方案。因此我們在代理人調適模型當中，加入超我，形成代理人的人格模型(圖 3)，透過此模型，我們將人類社會的理性道德觀實作於代理人之中，期望代理人能對超我有所覺察，以反應並解決傳統人工社會中，個體非理性行為所造成的集體行為不理性問題。





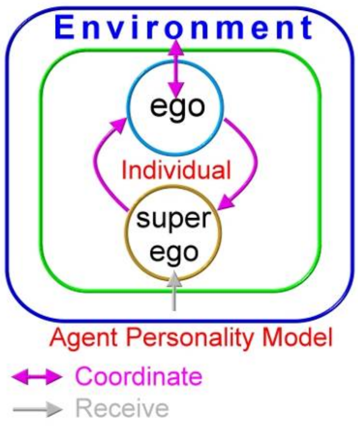


圖 1：佛洛依德人格三我　　圖 2：代理人調適模型　　圖 3：代理人人格模型

Appendix C　模擬系統

系統架構分為三層：第一層是社會網路拓樸；第二層是以囚犯困局為基礎的策略代理人模型；第三層是自我覺察模型。本論文主要以策略代理人加上社會網路(第一層與第二層)來模擬人工社會的公益與私利衝突問題(對照組)；同時，為解決此問題，我們以人格基模做為 ACLM 的對內學習焦點，提出一套以超我為自覺目標的代理人自覺模型(實驗組)。期望藉此自覺代理人，一方面能使自己活得更好(在環境中先顧到私利），一方面更要能透過促進合作，因而使大家和自己都活得更好(透過公益來實現私利)，來解決人工社會公益與私利的衝突問題。

社會網路拓樸

系統底層使用兩種著名的社會網路：細胞自動機與小世界網路。細胞自動機為一個具有高群聚度的網路結構，它易於我們觀察策略代理人的集體行為與群聚現象。而小世界網路是一同時具有高群聚度與低分隔度的網路結構，它符合真實的人際網路，而相關的代理人應用，也都直接與人類有所關聯(例如：電子商務、線上遊戲)。因此，在這兩類不同的社會網路結構下，超我自覺模型是否能改善社會整體利益，將是我們探討的焦點。

細胞自動機的建構方式是以每個二維空間的格子做為原點往外延伸，若延伸距離為 k 步，我們稱半徑-k 的鄰居關係(radius-k neighborhood)。以半徑-1 的鄰居關係(radius-1 neighborhood)進一步說明(圖 4)，每一個格子往其(上、下、左、右、左上、左下、右上、右下)延伸，建立代理人共事夥伴的連結，所以總共會有 8 個連結對象；同理，若以半徑-k 的鄰居關係做為細胞自動機的連結依據，將會有 $(k+2)^2-1$ 個連結。小世界網路建構方式同自胞自動機一樣，以半徑-k 的鄰居關係建立高群聚特性；另外，在建立周圍鄰居關係後，會破壞某些原有連結，並以

被破壞的數量，挑選遠方代理人加入捷徑(低分隔度)形成新的連結關係。(圖 5 為半徑-1 的鄰居關係之小世界網路示意圖)

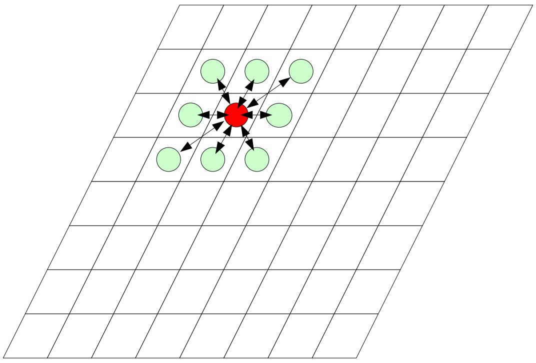


圖 4：半徑-1 的鄰居關係之細胞自動機

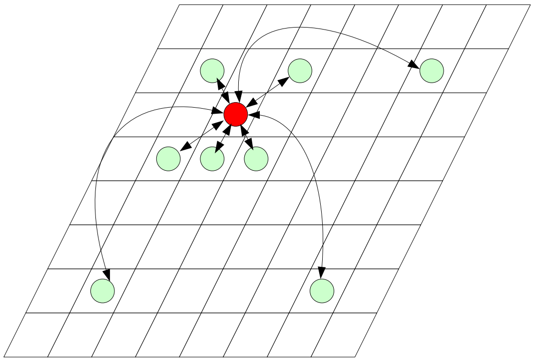


圖 5：半徑-1 的鄰居關係之小世界網路

策略代理人

代理人採取純粹策略，而純粹策略的定義是指：代理人在每個可能的決策節點上該如何抉擇的一種規則。它是一個完整的行動計劃，讓代理人知道在各種可能性之下要採取哪種行動。因此，首先需要釐清的問題，就是代理人到底有多少個決策節點？

我們的代理人是採用記憶為一的決定性策略(Memoery-1 Deterministric Strategy)，換言之，代理人可記錄上回合與夥伴間的共事情況；由於代理人共事情況只有四種可能，分別為雙方都履約(CC)、我履約而對方毀約(CD)、我毀約而對方履約(DC)、雙方都毀約(DD)。因此可將記憶為一的決定性策略記做(P_0, P_1, P_2, P_3)。舉例來說，若前世記憶是 CC，代表策略代理人就會以 P_0 回應之，其它依此類推；而每種回應只有兩種可能(履約與毀約)，所以總共會有 2^4 種策略組合。而這 16 種策略組合我們記作 $S_0, S_1, ..., S_{15}$，分別代表(C,C,C,C)、(C,C,C,D)、…、和(D,D,D,D)。

每個策略我們賦予一種相對應並且唯一的顏色，以便我們觀察策略代理人的演化行為浮現。代理人挑選共事夥伴是以底層的社會網路連結做為選擇的依據。共事報酬如表 1 所示。同時，為了方便計算共事結果，我們將報酬表轉換成圖 6 的報酬矩陣，並且透過下列的變數假設與算式推演，得到一組計算共事結果的一般化公式。

表 1：共事報酬表

Player B Player A	Cooperation	Defection
Cooperation	**3** **3**	**0** **5**
Defection	**5** **0**	**1** **1**

$$A = \begin{pmatrix} 3 & 1 \\ 5 & 0 \end{pmatrix}$$

圖 6：報酬矩陣

假設變數與符號說明：

$X_i = \begin{bmatrix} x_{i,0,0} & x_{i,0,1} \\ x_{i,1,0} & x_{i,1,1} \end{bmatrix}_{2x2}, x_{i,a,b} \in \{0,1\}, 0 \le a,b \le 1 \Leftrightarrow$ 代表代理人(i)的策略

$m_{ij,t} = (m_{ij,t,0}, m_{ij,t,1}) \Leftrightarrow$ 代表代理人(i)在時間點(t)面對共事夥伴(j)的前世記憶

註 1：$m_{ij,t,0}$ 與 $m_{ij,t,1}$ 分別代表自己與共事夥伴的前世決策(0 為合作，1 為背叛)

$d_{ij,t} \in \{(0,1),(1,0)\} \Leftrightarrow$ 代表代理人(i)在時間點(t)面對共事夥伴(j)的決策行動

註 2：(0,1)代表合作(履約)、(1,0)代表背叛(毀約)

$P_{ij}(g) \Rightarrow$ 代表代理人(i)在第(g)代面對共事夥伴(j)的總共事報酬

$P_i(g) \Rightarrow$ 代表代理人(i)在第(g)代面對所有共事夥伴的總報酬

$rounds \Rightarrow$ 代表每個代理人與其夥伴反覆共事的次數

$\Omega_i \Rightarrow$ 代表代理人(i)的所有共事夥伴集合

推演過程如下：

(1) $m_{ij,0} = \left(random_{one}\{0,1\}, random_{one}\{0,1\}\right)$

(2) $d_{ij,1} = \begin{cases} (1,0) & ,if \quad X_{i,m_{ij,0,0},m_{ij,0,1}} = 0 \\ (0,1) & ,if \quad X_{i,m_{ij,0,0},m_{ij,0,1}} = 1 \end{cases}$

(3) $m_{ij,1} = \left(X_{i,m_{ij,0,0},m_{ij,0,1}}, X_{j,m_{ji,0,0},m_{ji,0,1}}\right)$

(4) ....

(5) $d_{ij,t} = \begin{cases} (1,0) & ,if \quad X_{i,m_{ij,t-1,0},m_{ij,t-1,1}} = 0 \\ (0,1) & ,if \quad X_{i,m_{ij,t-1,0},m_{ij,t-1,1}} = 1 \end{cases}$

(6) $m_{ij,t} = \left(X_{i,m_{ij,t-1,0},m_{ij,t-1,1}}, X_{j,m_{ji,t-1,0},m_{ji,t-1,1}}\right)$

(7) $P_{ij}(g) = \sum_{1<t<rounds} d_{ij,t} \cdot A \cdot d_{ji,t}^{T}$ ，

(8) $P_i(g) = \sum_{j \in \Omega_i} P_{ij}(g)$

註 3：$random_{one}\{0,1\}$是指在$\{0,1\}$集合中隨機挑選一值，選到 0 代表代理人(i)對代理人(j)的前世記憶是履約；反之，若選到 1 則代表毀約。

挑選規則：

產生下一代策略代理人的挑選原則是以一個門檻值(T_s)做為基準，若代理人在某代的總報酬，比共事夥伴報酬排序後的倒數T_s名還高時，就直接延用於下一代；反之，則進行交配與突變以取得一組候選策略的集合，並以隨機的方式挑選某中一候選策略做為下一代演化之用(如公式 9 所示)。

$$(9)\quad X_i(g+1)=\begin{cases} X_i(g) & ,if\quad P_i(g)\geq \min_{T_s}(P\in\Omega_i) \\ random_{one}\{Crossover(X_i(g)), Mutation(X_i(g))\} & ,if\quad P_i(g)<\min_{T_s}(P\in\Omega_i) \end{cases}$$

註 4：$\min_{T_s}(P\in\Omega_i)$是指共事夥伴的總報酬中，第T_s小的報酬值。

註 5：$Crossover(X_i)$與$Mutation(X_i)$會各得一組交配與突變後的候選策略集合。

圖 7 是系統實作介面，右側部份是環境參數與演化參數，左側棋盤部份是策略代理人的演化動態情形，中間色塊部份是指代理人對應的策略顏色。透過模擬系統的顏色對應與空間關係，可用來觀察策略代理人在生存環境中，如何為了自利目標(生存壓力) 而導致人工社會走向集體毀約的不理性行為。並且，經由共事夥伴間(履約)合作與(毀約)背叛的互動過程，也可幫助我們了解人工社會的演化平衡現象與複雜行為的浮現。

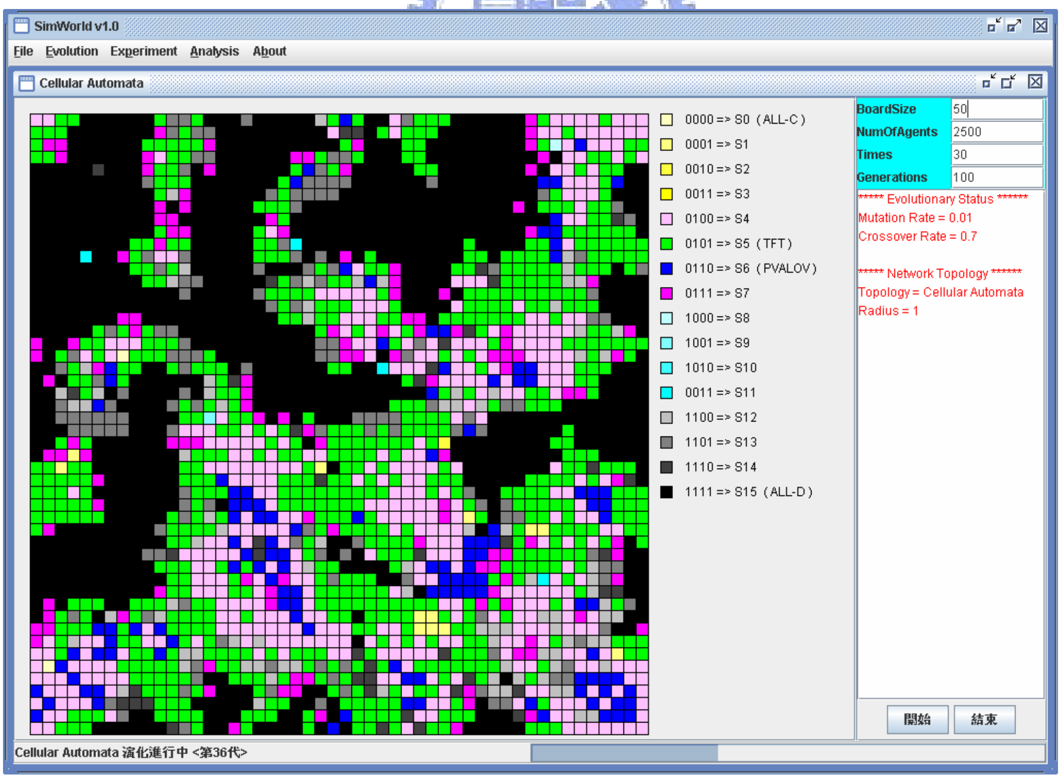


圖 7：模擬系統介面

自我覺察模型

經一系列的人格分析後，我們歸納出具備超我思考的學習代理人，將有助於提升整體社會利益，並使合作行為提早浮現。因此，代理人人格模型可應用於 ACLM 中的對內認知元件，換言之，我們藉用代理人對於自我與超我的認知，來實現代理人完成自我覺察的對內學習動作。

名聲機制

人工社會的名聲機制在許多策略代理人研究中都有提到，其主要用途是以名聲值高低來幫助代理人挑選有利的共事夥伴。但在人類社會中，除了挑選有利的共事夥伴外，個人名聲通常也可做為自我覺察與反省之用。因此，我們以此概念實作一組簡易的名聲機制於策略代理人之中。

評名聲規則

1. 假設每位代理人共有 n 個共事夥伴。
2. 每代終結時，代理人可得一組共事者對自己的合作次數集合 $\{C_0, C_1, ..., C_{n-1}\}$。
3. 建構 $C \sim \{C_0, C_1, ..., C_{n-1}\}$ 的標準常態分佈。
4. 根據圖 8 給予每位共事夥伴合適的名聲值。

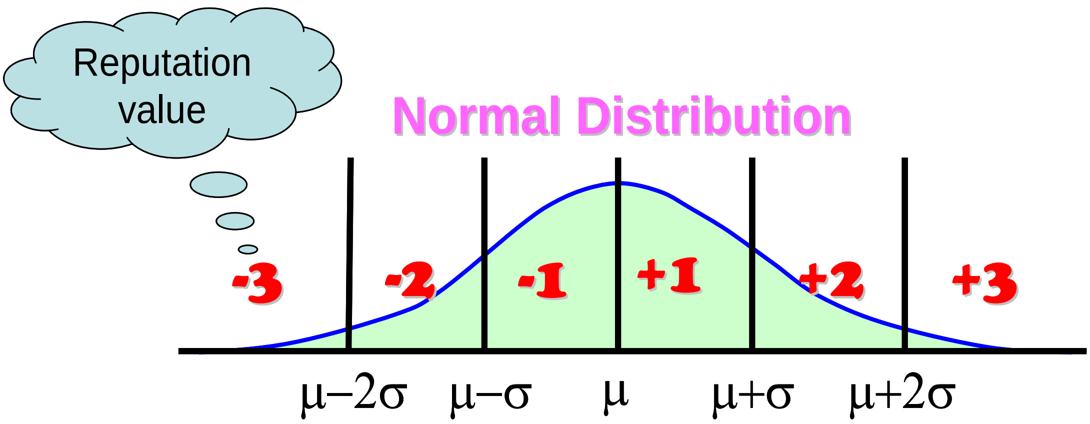


圖 8：代理人評斷共事夥伴名聲的對應圖

社會期望策略

當我們導入名聲機制於代理人社會後，代理人不僅知道自己的適存度，還可以透過名聲值來了解自己在所處群集的聲望。如此一來，由適存度與名聲值兩個變項的組合，將影響傳統只論適存度高低的演化挑選規則。表 2 將兩個變項均切分之三類：高、中、低，因此總共有 9 種可能的組合。而每一種組合，代表代理人與其共事夥伴集群的可能互動關係。

表 2：名聲值與適存度的對應集群互動關係

Self-Awareness (*Self → Neighbors*)		**Reputation** Low	**Reputation** Middle	**Reputation** High
Fitness	Low	(ALL-D) → (ALL-D) (ALL-D) → (TFT) 1 (TFT) → (ALL-D)		(ALL-C) → (ALL-D) 2
	Middle			
	High	(ALL-D) → (ALL-C) 3		(ALL-C) → (ALL-C) (ALL-C) → (TFT) (TFT) → (ALL-C)

Social Good Expected Strategy

經上表分析，可發現高適存度與高名聲是最符和社會期望的好策略。但是，傳統的演化挑選規則只認為高適存度就是好策略，這與我們的分析是有差異的。以上表的高適存度但低名聲的對應格子為例，代理人可能是採用 ALL-D 的惡性毀約策略來取得好的表現，這顯然不利於整體社會利益。因此，我們的超我自覺代理人必須先了解自己的適存度與名聲值在群體的位置，若所處的位置位於上表 1,2,3 格之中，則應進行自我調整的動作。而自我調整是指從共事夥伴中挑選出符合社會期望的好策略來加以模仿並進行交配與突變以產生新的候選策略之用。

自我覺察步驟與流程

步驟 1. 每位代理人收集其所有共事夥伴的名聲值與適存度。

- $Nei\_F \sim \{Nei\_F_0, Nei\_F_1, ..., Nei\_F_{n-1}\} \equiv$ 共事夥伴的適存度集合
- $Nei\_R \sim \{Nei\_R_0, Nei\_R_1, ..., Nei\_R_{n-1}\} \equiv$ 共事夥伴的名聲值集合

步驟 2. 分別建構 Nei_F 與 Nei_R 的標準常態分佈。

步驟 3. 根據圖 9 的兩個門檻值(T_L 與 T_H)來覺察代理人本身的名聲值與適存度在所處群體中的位置(高或低)。

步驟 4. 經理性計算，若代理人所處位置為表 2 的①、②、③其中之一，代表目前代理人的策略並不適合生存於共事夥伴之中，因此，自覺代理人將進行自我調整的動作。(虛擬碼 1 為實作流程)

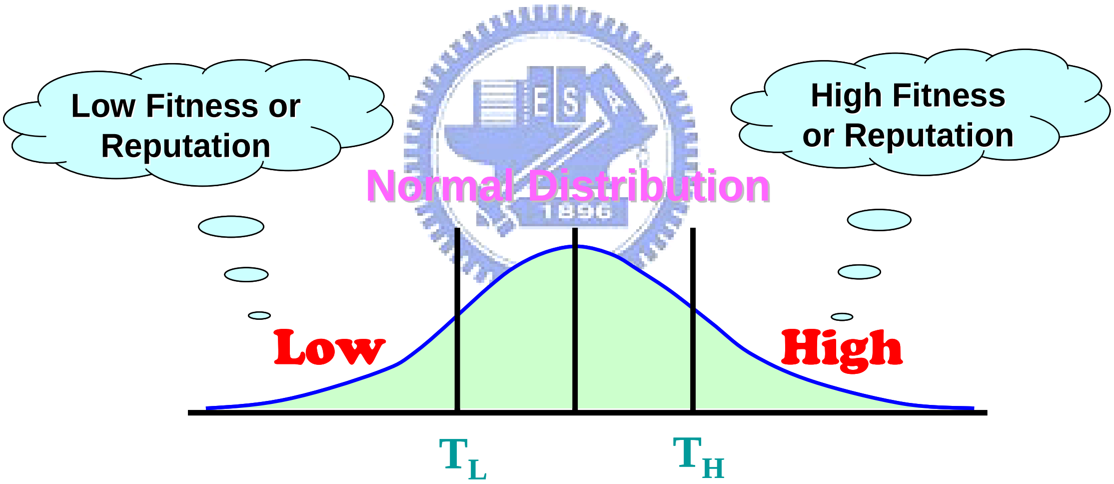


圖 9：名聲值與適存度的高低對應圖

虛擬碼 1：自覺代理人的理性計算與學習社會期望策略

If ($Fitness_{Me} = Low$) **or** ($Reputation_{Me} = Low$) **then**
    If (*My neighbors have social good strategies*) **then**
        Learning *social good expected strategy from neighbors*
    Else
        Do *original evolution procedure*
Else
    Do *nothing*

Appendix D　集體行為與群聚現象

細胞自動機

從第一類細胞自動機實驗的對照組(未加入自覺代理人)，我們將社會群聚行為分成五個階段，圖 10~13 為對照參考圖，詳細演化過程，請下載我們的演化動態影片。

第一階段(大約 1~3 代)：16 種策略維持差不多的比例，這是因為系統預設以隨機方式來初始化策略代理人，而代理人間也都還在互相摸索中。

第二階段(大約 3~10 代)：由於代理人感受到黑色的 ALL-D 策略，可以在 16 種策略(記憶為 1 的決定性策略)分散平均的環境下取得大量利益，依循代理人的基本生存原則，大量地演化出 ALL-D 策略。由於 ALL-D 數量急速往上驟升，使得社會充滿不信任的混沌現象。此刻，由於細胞自動機的空間區域特性，代理人只能與附近的鄰居進行共事，若共事夥伴都是 ALL-D 策略，必定代理人本身也要是 ALL-D 策略才不會被其它人給侵佔，因此，代理人社會開始發生惡性毀約的 ALL-D 群聚現象。

第三階段(大約 10~20 代)：綠色的 TFT 策略開始挺身而出，它扮演制衡 ALL-D 的角色，阻止混沌現象惡化。由於大量代理人開始演化出 TFT 策略，此刻，我們發現 TFT 代理人開始聚集起來，並逐漸擴展領域至包圍 ALL-D 代理人，這種行為浮現，讓我們了解 TFT 如何斷絕 ALL-D 侵蝕能力，它將 ALL-D 困在一塊區域中，此區域對外得不到利益(因為 TFT 的制衡)，對內也得不到任何便宜(因為 ALL-D 的互相背叛)。

第四階段(大約 20~40 代)：此階段是 TFT 急速下降的時期，由於 TFT 前世記憶的不對稱(TFT 與 STFT)問題，造成大量群集的 TFT，開始產生互相背叛的決策，崩解了 TFT 代理人的群聚現象。不過，雖然大量的 TFT 消失了，但包圍在 ALL-D 旁的 TFT 代理人並不會減少，因為它一消失會使 ALL-D 再度擴大侵蝕的領域。TFT 群聚崩解後，S4、S6(PAVLOV)、S7 開始逐漸浮現，因為這幾個策略並不會有前世記憶不對稱的循環毀約問題，而且策略彼此又能互相合作，所以數量開始逐漸成長。

第五階段(大約 40~100 代)：進入演化動態平衡期，S4、S6(PAVLOV)、S7 群聚區域中間開始浮現 ALL-C 的策略，因為此區域中，大家都具合作共事的特性，使得 ALL-C 策略有機會浮現，這也就是人類社會中，當過度相信人時，所造就的好人國度。不過若有少數策略突變或外敵入侵，很有可能使得好人國度再次消失，如此不斷反覆的過程，形成代理人社會的動態平衡。

由此五個時期的觀察，我們可得知代理人社會在細胞自動機中，複雜行為浮現的詳細情形。而接下來要探討的是以分佈 0.1 比例自覺能力於環境的實驗組，圖 14 為對照組與實驗組兩者在演化世代 20 的策略分佈圖，左下角為對照組，右下角為實驗組，從觀察中可以發現具自覺能力的實驗組，對於破壞 ALL-D 的群聚現象，比對照組來得顯著，原因是在 ALL-D 群聚中的自覺代理人，並不會堅持 ALL-D 策略，因為我們的自覺能力是覺察社會中群體的期望，雖然 ALL-D 策略有利於自己生存，但對整體公益是不利的，因此具自覺的代理人，會演化出可以對抗 ALL-D 的策略，也就是 TFT，它生存在 ALL-D 的群集當中，試圖等待外界的援助，就像間諜在 ALL-D 區域中，它與外界包圍的 TFT 互相配合，一但能量累積足夠，就能瓦解 ALL-D 的惡性群聚現象，這對於提早脫離社會的混沌現象有非常大的助益。

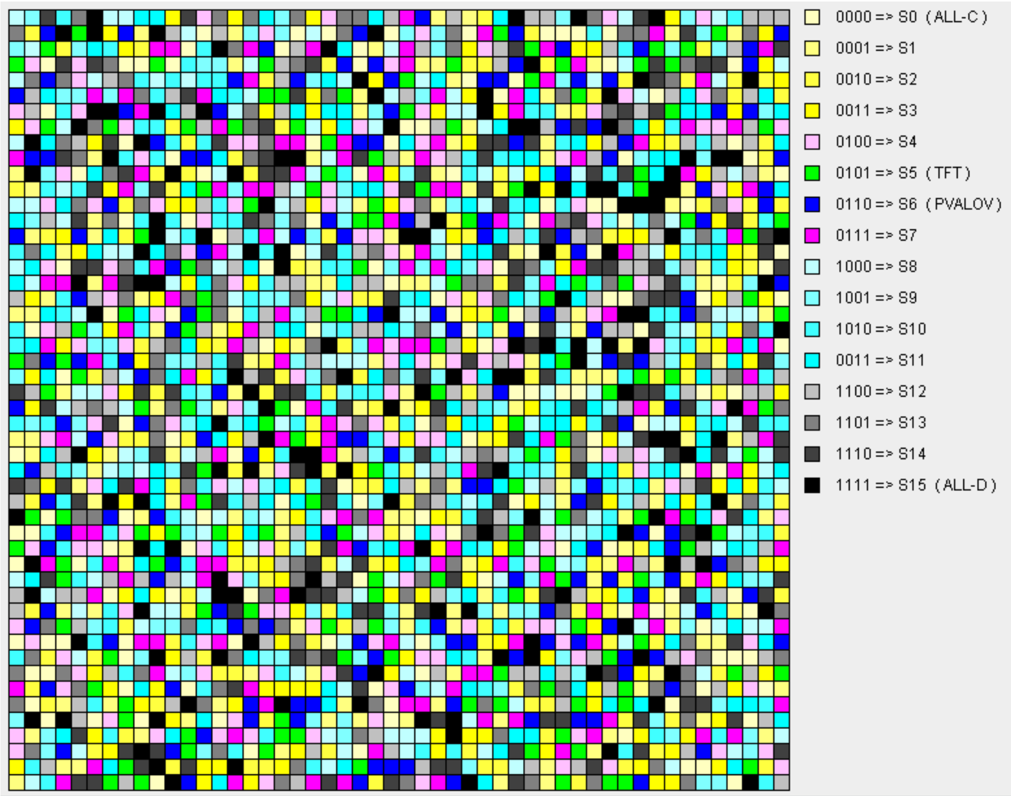


圖 **10**：記憶能力為 **1** 之策略代理人在細胞自動機的分佈情形 **(**演化世代 **0)**

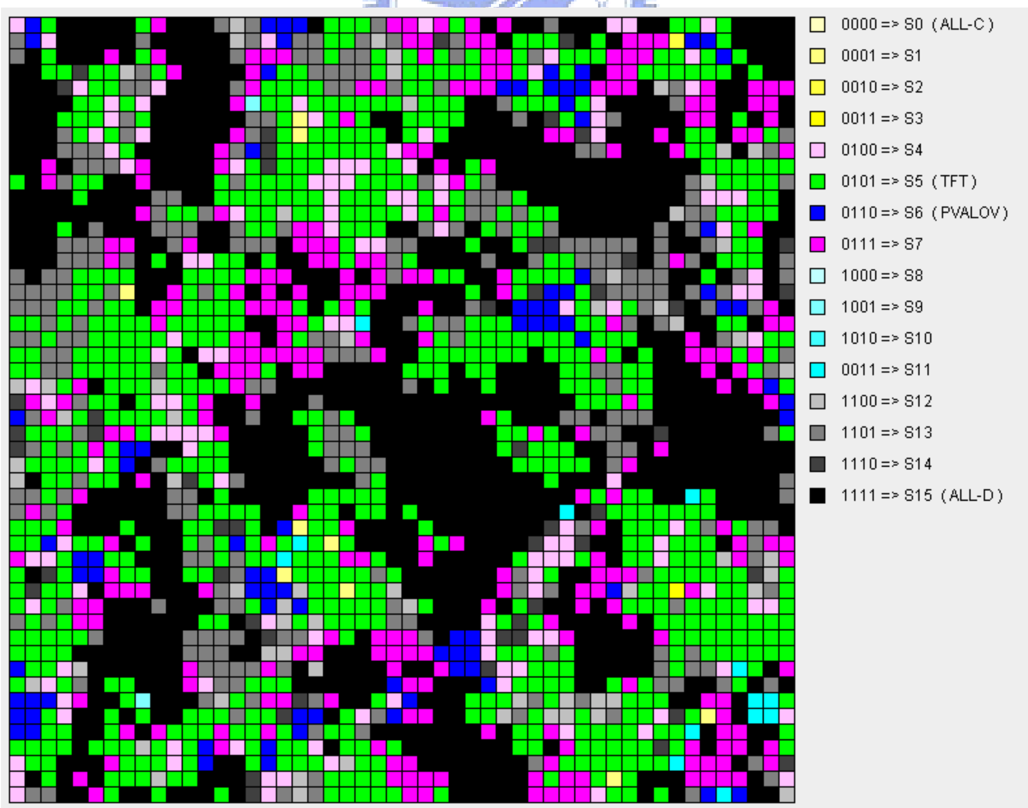


圖 **11**：記憶能力為 **1** 之策略代理人在細胞自動機的分佈情形 **(**演代世代 **20)**

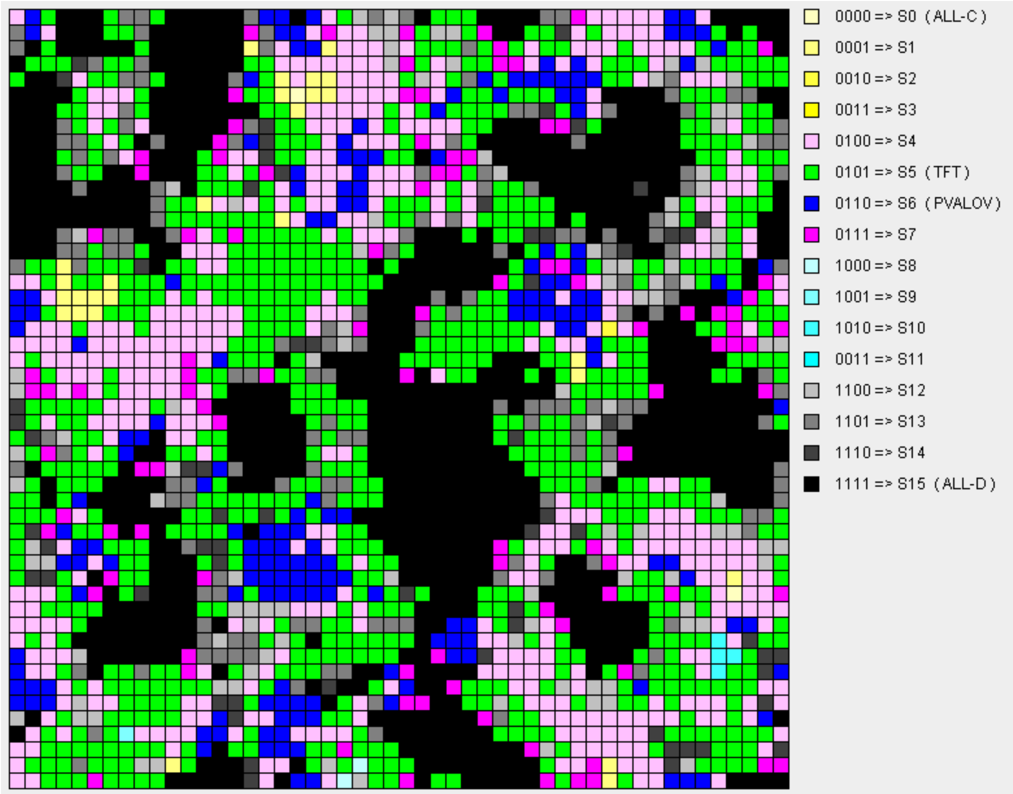


圖 12：記憶能力為 1 之策略代理人在細胞自動機的分佈情形 (演化世代 40)

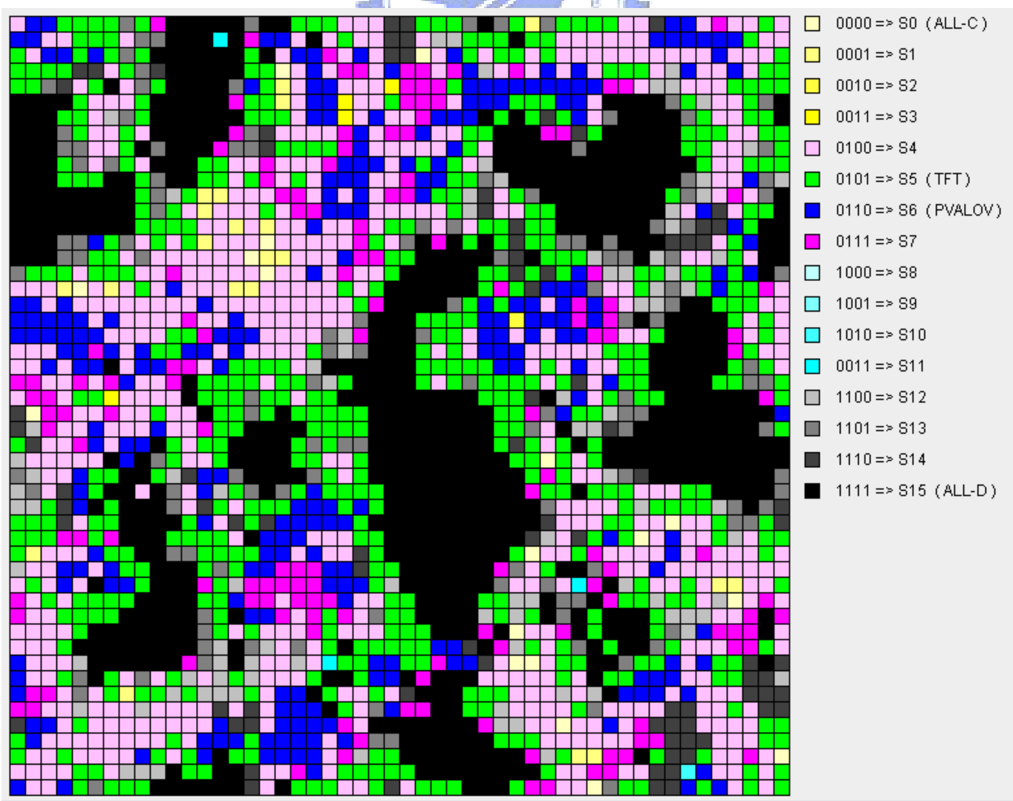


圖 13：記憶能力為 1 之策略代理人在細胞自動機的分佈情形 (演化世代 60)

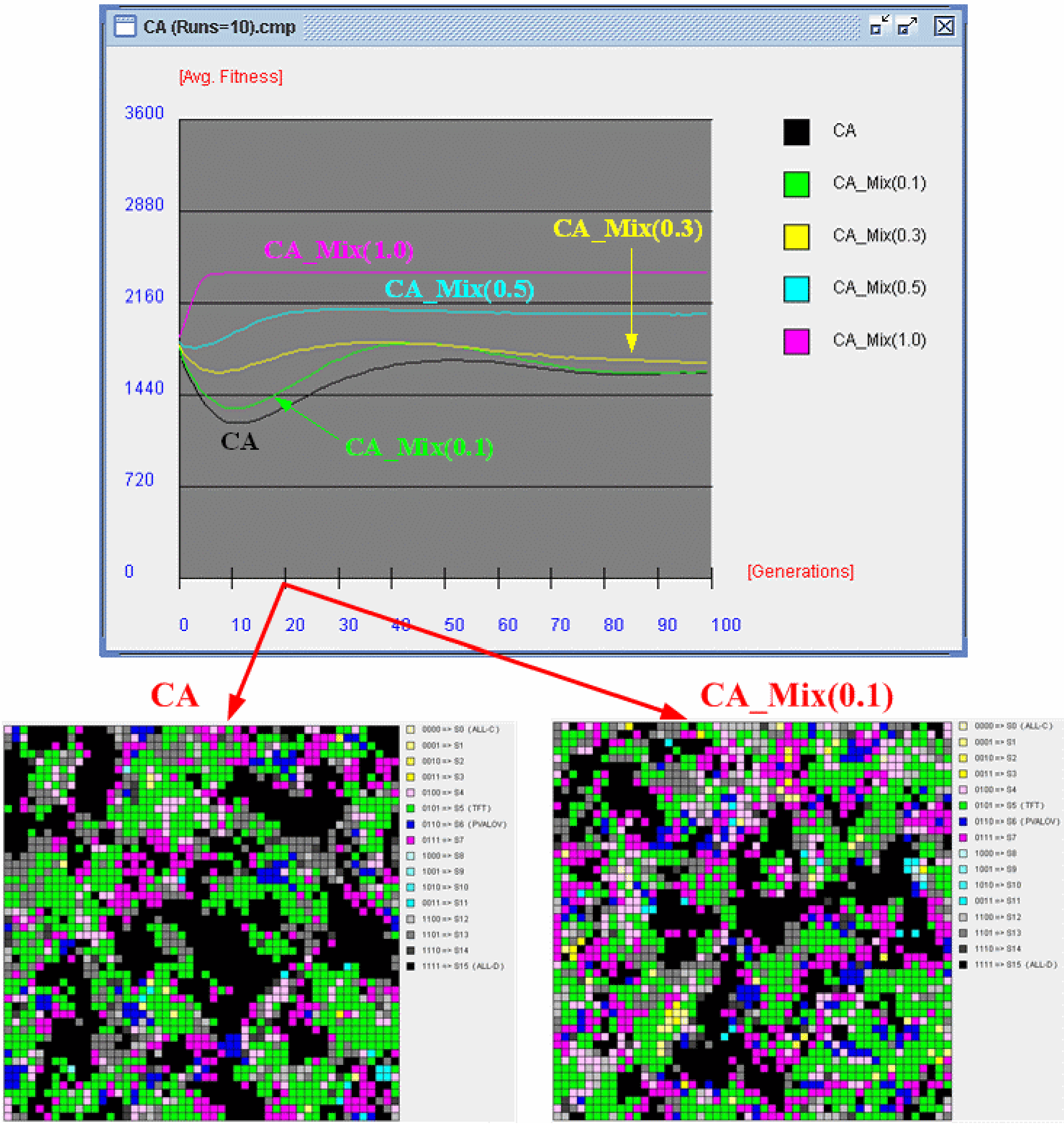


圖 14：細胞自動機對照組 vs.加入 0.1 比例自覺代理人實驗組 (演化世代 20)

小世界網路

第二類的小世界網路實驗，同細胞自動機的實驗結果，也就是具自覺能力的實驗組，一樣可以幫助代理人社會提早脫離混沌現象，可是相對於群聚行為的觀察，就比較難以了解細部的狀況。原因是小世界網路除了空間區域特性的高群聚度外，還有捷徑的低分隔特質，因此很難用肉眼觀察出複雜連結所造成的集體群聚現象。圖 15 為小世界網路對照組(無自覺代理人)的演化動態。

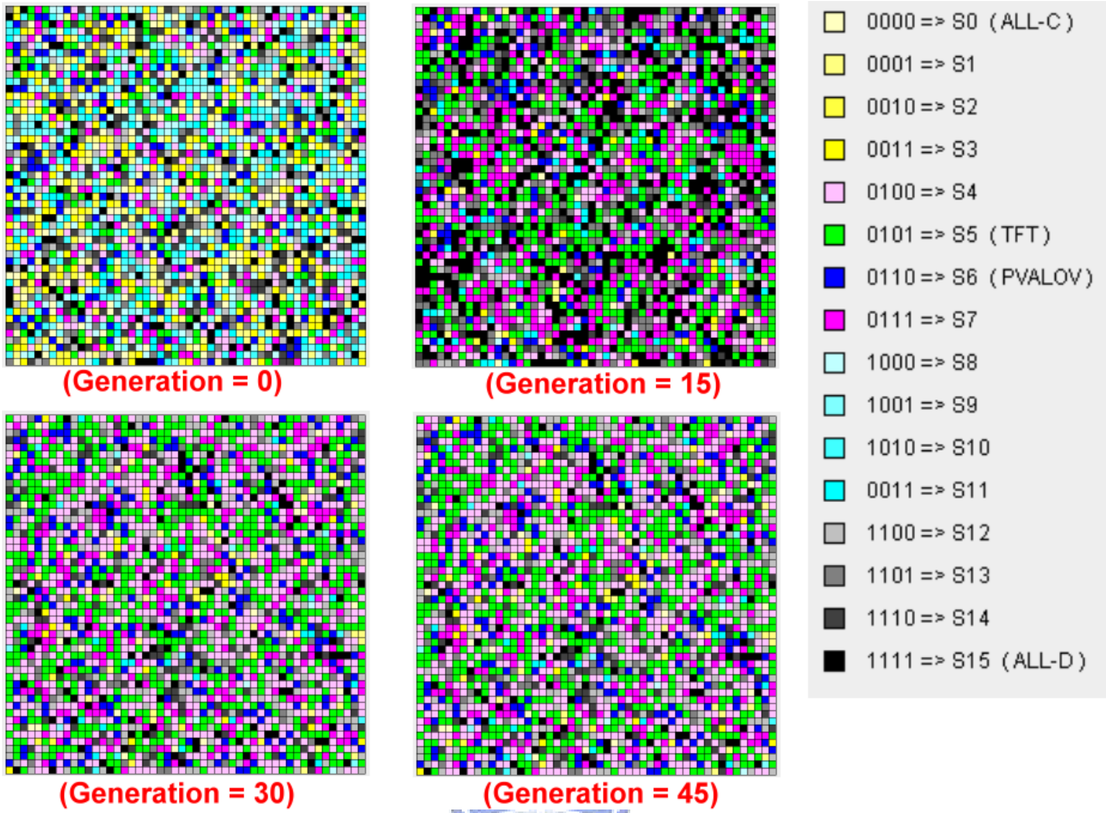


圖 15：小世界網路對照組之演化世代抽樣比較圖 (抽樣世代為 0、15、30、45)

上圖並沒有清楚而明顯的的集體群聚行為，這是因為小世界網路的捷徑使模擬世界的分隔度變低了，因此代理人之間的互動變的相當複雜，當然策略影響力也相對提高，所以社會的演化動態會比細胞自動機來的更加激烈、也更加快速。對照組大約到 30 代就快速達到整體的演化平衡狀態。由於沒有明顯的二維空間群聚現象，因此，我們無法進一步了解的複雜行為的浮現情形。

圖 16 為對照組與實驗組兩者在演化世代 15 的策略分佈圖，左下角為對照組，右下角為實驗組，從觀察中可以發現具自覺能力的實驗組，對於破壞 ALL-D 的群聚現象，並不如細胞自動機實驗來的如此顯著。雖然我們無法進一步了解複雜行為的浮現情形，但是，從 4 種名策略的消長分析，事實上還是可以發現超我自覺代理人，是有助於改善整體社會利益的。因此，沒有明顯的群聚現象並不會影響自覺代理人的有效性。

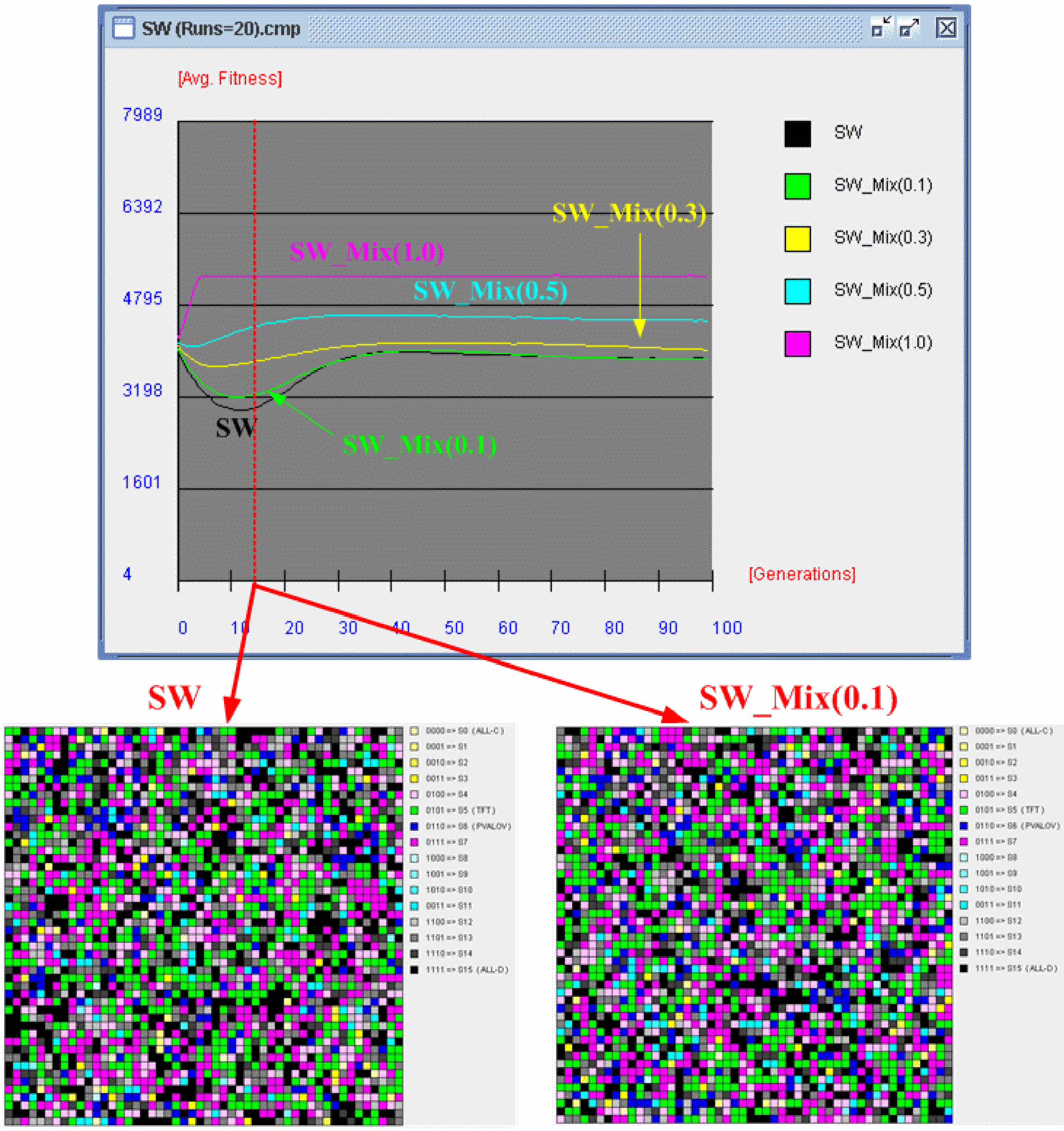


圖 16：小世界網路對照組 vs.加入 0.1 比例自覺代理人實驗組 (演化世代 15)

Appendix E　實驗設計與結果分析

實驗設計

實驗目的是要了解超我自覺代理人，有無能力解決人工社會的公益與私利衝突。因此，我們以模擬系統中的第一層與第二層做為實驗對照組。而實驗組的部份，分別以 1.0、0.5、0.3、0.1 的部份比例，加入第三層的自覺模型於策略代理人中。舉例來說，1.0 的實驗組是指整個環境中佈滿具自覺能力的代理人，其它比例請依此類推。下表是實驗的參數設定值。

表 3：實驗參數設定

	細胞自動機	小世界網路
反覆實驗次數	10	20
代理人總數	2500	2500
演化世代數	100	100
記憶能力	1	1
交配率	0.001	0.01
突變率	0.7	0.7
挑選壓力	0.25	0.25
建立連結方式	半徑-1 的鄰居關係	半徑-2 的鄰居關係

註 1：挑選壓力是指在共事夥伴之間的生存壓力。(例：假設有 10 共事夥伴而挑選壓力為 0.3，若代理人本身的適存度低於共事群體的倒數第三名(10 x 0.3 = 3)時，代表它不適應於環境中，而應進行演化的過程來產生較合適的策略(子代)。

註 2：半徑-k 的鄰居關係請參考附錄 C

結果分析

實驗結果使用三種分析方式。第一種：藉由觀察對照組與實驗組的整體平均報酬曲線，以了解各部份比例的自覺代理人對整體社會有何影響力存在，而觀察的結果證明了只要少數代理人具備自我覺察能力，即可有效提升整體社會利益。第二種：探討 4 種著名策略間的消長關係，來了解自覺代理人如何解決公益與私利間的衝突，而分析結果證明超我自覺代理人，可助社會提早脫離惡性循環的不信任毀約情形，這個結果使我們進一步了解自覺代理人的功能與作用。第三種：以模擬系統提供的策略顏色對應與二維空間關係，探討代理人間的集體行為與群聚現象。表 4 是我們比較兩類社會網路實驗結果的差異。

表 4：細胞自動機 vs. 小世界網路的實驗結果

細胞自動機	小世界網路
■ 資訊傳遞速度慢 ■ 易造成區域群聚現象 ■ 易於觀察策略之間相抗衡的關係 ■ 少量具自我覺察能力的代理人即可幫助提升整體社會利益	■ 資訊傳遞速度快 ■ 社會規範與制度形成速度快 ■ 社會平衡機制比細胞自動機快 ■ 少量具自我覺察能力的代理人即可幫助提升整體社會利益

經一系列的實驗與分析，我們最後歸納出超我自覺代理人具有以下特性：

- 可促進合作行為提早出現。
- 可使得代理人社會快速趨向穩定狀態。
- 少數代理人具自覺能力即可使整體利益提升。
- 經由超我道德分析可幫助代理人解決個別理性所帶來的集體不理性行為。

REFERENCE

Ana L. C. Bazzan, and Andrea P. Cavalheiro. 2003. "Influence of social attachment in a small-world network of agents playing the iterated Prisoner's Dilemma." *Third Workshop of Game Theoretic and Decision Theoretic Agents, Melbourne*.

Axelrod, Robert, and William Hamilton. 1981. "The Evolution of Cooperation." *Science* 211 (December 9), pp. 1390-96.

Axelrod, Robert, and Douglas Dion. 1987. "The Further Evolution of Cooperation." *Science* 242 (December 9), pp. 1385-90.

Ahmed E., and Elgazzar A. S. 2001. "On some applications of cellular automata." *Physica A*, 296, pp. 529-538.

Beom Jun Kim, Ala Trusina, Petter Holme, Petter Minnhagen, Jean S. Chung, and M.Y. Choi. 2002. "Dynamic instabilities induced by asymmetric influence: Prisoners' dilemma game on small-world networks." *Phys. Rev. E*, Vol. 66, 021907.

Bazzan, A. L. C., Bordini, R., Campell, and J. 2002. "Evolution of Agents with Moral Sentiments in an Iterated Prisoner's Dilemma Exercise." *Game Theory and Decision Theory in Agent-Based Systems*.

Bazzan, A. L. C., Bordini, R.H., Viccari, R., and Wahle, J. 2000. "Evolving Populations of Agents with Personalities in the Minority Game." *International Joint Conference.* Atibaia, SP, Brazil, November 11-22, LNAI , p. 166-175.

Boccaran, Goles E., Martinez S., and Picco P. 1993. "*Cellular Automata and Cooperative Phenomena*." Kluwer Academic Publisher, Boston.

Duval, S & Wicklund, R. A. 1972. "A theory of objective self-awareness." *New York: Academic Press.*

Guillermo Abramson, and Marcelo Kuperman. 2001. "Social games in Social network." *Phys. Rev. E*, Vol. 63, 030901.

Hines, and W. G. S. 1987. “Evolutionary Stable Strategies: A Review of Basic Theory.” *Theoretical Population Biology* 31:195-272.

Hsu-Chih Wu, and Chuen-Tsai Sun. 2004. ‘Comments on "A Computational Evolutionary Approach to Evolving Game Strategies and Cooperation",’ *IEEE Transaction on system, Man and Cybernetics, Part B: Man*.

Hsu-Chih Wu, Chuen-Tsai Sun and Sih-Shin Lee. 2005. “Visualization of evolutionary computation processes from a population perspective," *Intelligent Data Analysis*, Vol. 8.

Jonathan Bendor, and Piotr Swistak. 2001. “The Evolution of Norms.” *American Journal of Sociology*, Vol. 106, pp. 1493-1545.

Kihlstrom, J.F. 1987. “The cognitive unconscious.” Science, 237, 1445-1452.

Mayuko Nakamaru, and Simon A. Levin. 2004. “Spread of two linked social norms on complex interaction networks.” *Journal of Theoretical Biology*, Vol. 230, pp. 57-64.

Markus, H. 1977. “Self-schemata and processing information about self.” *Journal of Personality and Social Psychology*, Vol. 35, pp. 63-78.

Robert Behrman. (2004). “Social Network Influences on Strategic Choices,” *CASOS*.

Schoenfeld, Alan H. 1983. “Beyond the Purely Cognitive: Belief Systems, Social Cognitions, and Metacognitions As Driving Forces in Intellectual Performance” *Cognitive Science*, 7, 329-363.

WANG X. F., and Chen G. 2003. “Complex Networks: Small-World, Scale-Free and beyond.” *IEEE Circuits and Systems Magazine*, First Quarter, pp. 6-20.

Watts D. J and Strogatz S. H. 1998a. “Collective dynamics of 'small-world' networks.” *Nature*, 393(6684), pp. 440-442.

Yu, B. , and Singh, M.P. 2000. “A Social Mechanism of Reputation Management in Electronic Communities.” Proceedings of Fourth International Workshop on Cooperative Information Agents, pages 154-165.