國 立 交 通 大 學

資訊管理研究所

博士論文

以認知學習修正 XCS
建構具知識教育與機械學習之雙模式學習機制
—以財務資料預測之知識學習為例

Applying Cognitive Learning to Enhance XCS to Construct a Dual-Mode Learning Mechanism of Knowledge-Education and Machine-Learning

— an Example of Knowledge Learning on Finance Prediction

研 究 生： 陳 怡 璋

指導教授： 陳 安 斌

中 華 民 國 九 十 四 年 七 月

以認知學習修正 XCS 建構具知識教育與機械學習之雙模式學習機制—以財務資料預測之知識學習為例

Applying Cognitive Learning to Enhance XCS to Construct a Dual-Mode Learning Mechanism of Knowledge-Education and Machine-Learning

— an Example of Knowledge Learning on Finance Prediction

研 究 生：陳怡璋　　　　Student：Yi-Chang Chen

指導教授：陳安斌 博士　　　　Advisor：Dr. An-Pin Chen

國 立 交 通 大 學

資 訊 管 理 研 究 所

博 士 論 文

A Dissertation

Submitted to Institute of Information Management

College of Management

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Information Management

July 2005

Taipei, Taiwan, the Republic of China

中華民國九十四年七月

以認知學習修正XCS建構具知識教育與機械學習之雙模式學習機制
—以財務資料預測之知識學習為例

學生：陳怡璋　　　　　　　　　　　　　　　指導教授：陳安斌 博士

國立交通大學資訊管理研究所

摘要

自 1956 年以來，人工智慧所定義的機器學習與長久以來研究人類心智行為的心理學所定義的學習，兩者明顯不同。由於電腦運算能力的提升，使得我們可以有能力再次重新檢視學習的定義，以此希望可以達成更高效率與準確率的智慧學習模型。

本研究企圖以認知心理學之認知結構來修正自 1956 年以來人工智慧之發展，由於人工智慧長期侷限於試誤學習之低效率學習模式，然而試誤學習於傳統心理學定義中僅限於刺激與反應之經驗行為而已，由此學習模式所建構之任何機器學習，均只能認定為經驗之適應模式而已，而較進階的種類，如演化式計算模型，也只是其能透過電腦強大的運算能力來達成所謂的動態環境下之演化式學習模式，其中演化之特色只是多考慮了外在環境的變化或內在參數的調整，而整個學習流程卻沒有進一步修正。這也可說明，當各人工智慧之原始模型發展針對封閉式環境問題，都會有很好的表現，但面對非封閉式問題時，卻只能經由大量實驗和透過參數的調整來片面獲取結果而無法自圓其說。

認知心理學中較完整的認知學習發展是在 1986 年以後，相關研究指出有效率的學習過程必須包含教育學習，而不再僅有透過試誤學習來達成。以此，本研究發展一套修正傳統機器學習之學習流程－雙模式智慧型學習機制。另外，由於 XCS 系統是試誤學習類之效果較佳及準確率較好的其中之一模型，透過以 XCS 為基礎加上本研究所提之學習流程架構，繼而發展出一個有效率之智慧型學習模型(E&R-R model)。

最後，本研究試圖以較複雜的問題來進行實驗模擬，而該問題為運用財務資料以建立財務預測知識模型，其模式為三種：XCS，R-R XCS 與 E&R-R XCS，透過三種模型的準確率與最後報酬率之比較來驗證本研究所提出之學習流程的效能。初步驗證，E&R-R XCS 均較 R-R XCS 和 XCS 之機制有顯著效能提升。

關鍵字：人工智慧，心理學，認知結構，試誤學習，教育學習，智慧型學習模型。

Applying Cognitive Learning to Enhance XCS to Construct a Dual-Mode Learning Mechanism of Knowledge-Education and Machine-Learning — an Example of Knowledge Learning on Finance Prediction

Student: Yi-Chang Chen Advisors: Dr. An-Pin Chen

Institute of Information Management

National Chiao Tung University

Abstract

From 1956, the learning definitions of Artificial Intelligence and Psychology to human mind/behavior are obviously different. Owing to the rapid development of the computing power, we have potential to enhance the learning mechanism.

This work tries to apply the learning process of the cognition structure defined in Cognitive Psychology to enhance or modify the development of AI, of which the learning models are almost based on trial and error style. However, this kind of learning style is definably given to the experience behavior of stimulus and response in Psychology. Thus, the relative AI models based on such style are design as an experience-adaptation system. For better ones, e.g. evolution-base algorithms, they belonged to the system with more powerful computing power to the dynamical environment. Even so, it was considered not only outside environment but also internal parameter tuning. As for the entire learning process, it has never been enhanced. That is, various original AI models are easily to be developed to their own close-form problem. To the unclose-form problems, their distinct results only come from huge amounts of experiments and tuning their model's parameters. As the result, it is not easy to make clear for the explanation to why or how.

The desirable cognitive learning of cognitive psychology is the development that has started since 1986. The relative literatures have pointed out that teaching-base education would increase the learning efficiency, but trial and error style is not sufficient to learning. That is the reason we enhance the AI learning process to develop a dual-perspective learning mechanism. Furthermore, since XCS is a better accuracy model of AI, we have applied it as a basement and involve the enhanced model proposed to develop an intelligence-learning model.

Finally, this work is designed a test of the more complex problem, which is constructing a finance prediction knowledge model. By comparing to the accuracy and accumulative profit of XCS, R-R XCS and E&R-R XCS respectively, the results obtain the obvious outcome. That is, the proposed learning framework has enhanced the original mechanism.

Keyword: Artificial Intelligence, Psychology, Cognition Structure, Trial and Error, Teaching-Base Education, Intelligence-Learning Model.

誌謝

博碩士生活終於告一段落，在這期間最要感謝的是指導教授　陳安斌老師，在這幾年間，陳老師對我的指導，不單只有在學業方面，在待人處事上也處處提醒我：不僅要時時感恩更要廣結善緣，這樣的過程，讓我不僅在學術研究上求得進步，更在生活觀上得到提升；同時特別感謝諸位口試委員：陳鴻基老師、楊千老師、劉敦仁老師、林妙聰老師以及鄭景俗老師，對於我的博士論文給予悉心指導；感謝您們提供寶貴的建議，使得論文能更臻完整。

此外，從小到大，爸爸和媽媽給我無止境的關懷，一直是我求學過程中最大的後盾，這是我最引以為傲的，對此，心中存有無限感激；大妹和小妹對我的鼓勵與支持，她倆更是代替我隨伺爸媽身邊盡孝順心，這是我無以回報的；另外，特別值得一提的是我那可愛的女朋友，總是會替我分憂解勞，她的聰明伶俐，讓我的研究生活充滿快樂的點滴。對於這些家人，我深深感謝他們，謹以此論文與他們分享。

最後感謝研究室裡，不管是和我一起打拼的小組或是 APC LAB 之其他成員，你們的存在，讓我的研究生活不會感到孤單；乃至於資管所的各位朋友，要感謝的人太多，無法一一陳述，這些種種我都點滴在心頭，你們都是我生命中不可或缺的一部分，我永遠感激你們！

陳怡璋 2005/7/31

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Motivation

Traditionally, Artificial Intelligence, according to the definition of Computer Science, works as helpful machines to find solutions to complex problems in a more human-like fashion [1]. This generally involves adopted characteristics from human intelligence, and it applies them as algorithms in a computer friendly way. A more or less flexible or efficient approach can be taken depending on the requirements established, which influences how artificial the intelligent behavior appears. Those researches, for example: Neural Network, Fuzzy Approach, Genetic Algorithm, and so on, all focus on Soft Computing. Of course, XCS (Extend Classifier System) is also a hybrid approach with high performance to the accuracy and the rule evolution on the prediction application. However, up to now, the Artificial Intelligence Techniques based on Soft Computing have all involved the concept, trial and error method or stimulus-response method even the series of evolution approaches [2,3], to construct their learning models. For this aspect, if possible, this example, a Chinese idiomatic phrase-"An Illusory Snake in a Goblet", is taken into consideration as an input-output pattern to training the learning model. The models are formed for sure. It is actually a wrong model trained by a bad experience. Besides, the parameters of those training models are exactly affected by the input dataset, especially the large difference of the training inputs and testing ones. Usually, in many researches it is chosen the high relation between the input and output datasets or given the strong assumption which is the inputs and outputs are relevant. Thus, a subjective black-box view and the tuning view are easily concluded [4].

The other sub-domain, Expert System, which's primary goal is to make expertise available to decision makers and technicians who need answers quickly. There is never enough expertise to go around -- certainly it is not always available at the right place in the right time. The same systems in-depth knowledge of specific subjects can assist supervisors and managers with situation assessment and long-range planning. These knowledge-based applications of artificial intelligence have enhanced productivity in business, science, engineering, and even the military. Although, the development of those expert systems is the view of anti-extreme to construct domain knowledge first but, for the reason, they are lack of the flexibility and the adaption. In fact, each new deployment of an expert system yields valuable data for what works in which context, thus fueling the AI research that provides even better applications.

Many researches, no matter Soft Computing techniques or Expert Systems try to consider into the human-like thinking way to make the simulation. But, from classic psychology, the human-mind researches are the researches to the human-behavior. Since Plato, Psychology is an unfathomable philosophy and those advanced AI researchers should concern this perfect development of Human Psychology, from simple to complex and from single factor to multiple ones. However, the traditional AI techniques are seldom focused on the high level of human-mind process and just paid attentions to the learning definition from the Empricalism Psychology. According to the development of Modern Psychology, the core of Psychology has been already transferred Empricalism-base into Information Process Theory of Human-Mind, Cognitive Psychology-base. As for the knowledge and the model construction, the teaching-base aspect has been involved as well to the learning process. Based on the aspect, this work tries to enhance the learning process of traditional AI techniques whose cognitive scotomas of learning definition, and it develops the novel learning model, involving the concept of Cognitive Psychology, which is utilized the high accuracy-prediction XCS model as the construction basement.

1.2 Purpose

Among learning artificial intelligence techniques, no matter neural network, fuzzy approaches, or any hybrid methods, all the models are formed by trial and error learning way, the traditional definition of learning [1]. It is practicable to be implemented that those models are utilized to a close-form problem. As for the others to unclose-form problems, however, it is critical the set of their relative input and output pairs needs to be modified. The datasets used to train or test should be all verified first as well, which is a boring work to the model designers. Besides, the relative problems of those evolution artificial intelligence techniques are also faced to my pre-statement. It is more significant to concern the proper datasets as inputs effects the model construction. By Darwin's Evolution Theory, Natural-Selection is easily to be concluded for the all organisms. The detailed steps could be realized that each obvious verified evolution result is always caused by the right things, the key factors, and the certain environment at the critical time. It is definitely not the random result. Take human evolution for instance, judged from the biotic evolution history of the earth – from the mitochondria, the cell, the microorganism, the multi-cell organism, …, the pithecanthrope, to Human, who dare to assure Human as the primate animal, still would own respectively two hands and two feet, each five fingers, if the history of the earth reshuffles?. That explains the reason, of which the dimension to solve problems could not be too complex, is that the training samples are not always sufficient to construct the model.

Nevertheless, much Knowledge discovery, Theory verification and Theorem definition are aggregated and not disregarded. They are all continually historical accumulated. That is also the reason that the civilization is enhanced, the culture is accumulated, and knowledge is transmitted. Either the voluntary learning or the passive learning through education is the key cores in each process. Following the previous concept, moreover, the hybrid approach, XCS [5], has already been verified its prediction accuracy and its ability to dynamical environment and it becomes the foundation of this work to construct the knowledge learning model. The above two assumptions/pre-statements are taken into consideration to develop the efficient knowledge learning model of the self-learning and the passive-learning. The methodologies are applying XCS with the reinforcement learning ability and involving the Human education [6] characteristic of Cognitive Psychology. Furthermore, it is the purpose to develop the high efficient learning model with the high accuracy knowledge accumulation is its purpose. The major contribution of this work is the proposed architecture. Once, the more accuracy ability of AI Techniques invented could be substituted for XCS and more performance would be more efficient.

1.3 Research Problem

The research issue will be arranged to develop the efficient knowledge learning model. First, the learning definition would be concluded from traditional AI, especially the classifier system. Second, in this work we would try to survey the psychology, thousands year of its development, as the basement to analyze the development of AI and the learning of human behavior. Moreover, this work focuses on Modern Psychology, Cognitive Psychology, to collect and induce and its learning concept to develop an enhanced model which increases the training process and the knowledge output. As for the design of the simulation, the traditional training/learning process of XCS model would be respectively compared to the proposed learning model and the education-learning proposed model. Finally, the performance would be verified.

1.4 Organization

The rest of this dissertation is organized as follows. In Chapter 2 we review the related work on Classifier System, Cognitive Psychology, and the Relationship of Cognitive Psychology and Classifier System. In Chapter 3, the cognitive learning from the evolved learning is distinguished and the definition of memory from Cognitive Psychology is described. In Chapter 4 it presents the dual-mode learning mechanism by education (E) learning and reinforcement-rehearsal (R-R) learning based on XCS, which contains the description of XCS, R-R XCS, and E&R-R XCS. Chapter 5 compares the experiments with the three learning model. Nevertheless, the design of finance prediction simulation would be detailed first. Conclusions and future work are made in the final Chapter 6.

Chapter 2. Literature Review

2.1 History of Classifier System

Learning classifier systems are a machine learning paradigm introduced by John H. Holland. They first appeared in 1978 in the paper “Cognitive Systems Based on Adaptive Algorithms” by Holland and Reitman [7]. However, before that, Holland [8] foreshadowed classifier systems in 1971. In learning classifier systems an agent learns through experiments to perform a certain task by interacting with a partially unknown environment, using rewards and other feedback to effect an internal evolutionary process which forms the rule-based model of the world. The agent senses the environment through its detectors; based on its current sensations and its past experience, the agent selects an action sent to the effectors in order to be performed in the environment. Depending on the effects of the agent's action, the agent occasionally receives a reward. The agent's general goal is to obtain as much reward as possible from the environment.

In his pioneer work, Holland combined two ideas which later became key topics of the research in machine learning. The first idea was that Darwinian Theory of the survival of the fittest could be used to trigger the adaptation to the artificial system to an unknown environment. This idea later became the basis for many important research areas, such as Evolutionary Computation. The second idea was that an agent could learn to perform a task just by trying to maximize the rewards it receives from an unknown environment. This model of learning through “trial and error” interactions has been formalized and developed in the area of reinforcement learning, which is now a major branch of machine learning research. Learning classifier systems have been wielded all-around through out more than twenty years. In these two decades they have receiving the more and more attention by many researchers from many areas.

In Holland's learning classifier system, there were a number of well-noted problems that prevented the system from achieving satisfactory performance in some cases. In 1987 Wilson [9] introduced a new type of “one-step” classifier system. Wilson showed that BOOLE could learn multiple disjunctive concepts faster than neural networks. Separately, Booker [10] introduced GOFER-1, a new type of classifier system that. In GOFER-1 the classifiers fitness is a function of both payoff and non-payoff information, and the genetic algorithm works in environmental niches instead of in the whole population.

Wilson [11] observed that the architecture of learning classifier systems is too complex to permit carefully, revealing studies of the learning capabilities of these systems. Accordingly he simplified the original framework and then introduced ZCS, a zeroth level classifier system. After that, optimal performance in different applications was finally reported in 1995 when Wilson [5] invented the XCS classifier system. While XCS maintains Holland's essential ideas about classifier systems, it differs pretty much from all the previous architectures. First in XCS Q-learning is used to distribute the reward to classifiers, instead of a bucket brigade algorithm. Second, in XCS the genetic algorithm acts in environmental niches instead of on the whole population, as it does in the work of Booker on GOFER-1 [10]. The most important of all, in XCS the fitness of classifiers is based on the accuracy of classifier predictions instead of the prediction itself, a solution partially anticipated in the works of Frey and Slate [12] and Booker [10]. Wilson [5, 13] showed that by using classifier accuracy as the fitness of the genetic algorithm, XCS is able to evolve classifiers that are (i) accurate. They give an accurate prediction of the expected reward, and (ii) maximally general. They match as many situations as possible without being overgeneral.

Anticipatory Classifier System (ACS), introduced by Stolzmann [14], differs greatly from other LCSs in that CS learns not only how to perform a certain task, but also learns an internal model of the dynamics of the environment or task. In ACS classifiers, there are not simple condition-action rules but they are extended by an effect part. The effect-part of a classifier is used to anticipate the environmental state which results from the execution of the classifier action. The model of the environment can be learned latently, that is it learned without any environmental reward, because the fitness of the classifiers depends on the accuracy of the anticipation. The classifier fitness is high if the next state is anticipated correctly while it is low if the anticipation is wrong. Besides genetic algorithms an Anticipatory Learning Process (ALP) is utilized for rule discovery which directly learns from the changes in the environment. ALP is a further development of a psychological learning theory, called anticipatory behavioral control [14]. ACS forms explicit condition-action-effect classifiers with a generalization capability in the classifier conditions. This leads to an internal model of the environment which consists of a minimal set of classifiers. The internal model can be used in many applications: (i) for mental acting and look ahead planning to improve learning, (ii) for action planning and goal directed planning in the absence of environmental reward, and many more.

LCSs have been applied in many domains [15]. However, most of the results reported fall into three main areas: autonomous robotics; knowledge discovery; and computational economics. For a good instance, Holmes' EpiCS [16] is an LCS specialized for classification and knowledge discovery tasks. It was developed from NEWBOOLE to meet the demands of epidemiologic data. EpiCS's distinctive features include: (i) techniques for controlling over- and under-generalization of data, (ii) the use of differential negative reinforcement of false positive and false negative errors in classification, and (iii) a methodology for determining risk as a measure of classification. All of these features have led to the successful usage of EpiCS in knowledge discovery applications to actual clinical databases of various sizes and levels of complexity. EpiCS was able to (i) derive models that identified features that were associated with outcomes such as appropriate child restraint in automobiles, (ii) occupational cancer (simulations), and (iii) head injury to children involved in automobile crashes (see [17] for an overview). Therefore, EpiCS appears to be a successful approach to apply evolutionary computation to the realm of knowledge discovery in databases.

The development of new LCS models [18], successful in many domains, has led to a resurgence of this area during recent years. Overall, the recent results represent probably the most significant advances in LCS research presented so far. However, most work still need to be done; there are many interesting research directions to be explored, and many open challenges. Besides, owing to the origin of LCS is described as a cognitive system by Holland, next will be discussed with cognitive psychology.

2.2 History of Cognitive Psychology

Cognitive Psychology is concerned with advances in the studies of memory, language processing, perception, problem solving, and thinking. However, to explore the beginning of Cognitive Psychology should be traced to the field of psychology whose history diagram was shown as Figure 1. The earliest roots of psychology would be divided into two different approaches to understand the human mind: philosophy and physiology. The pre-evidences are the two Greek philosophers Plato (ca. 428-348 B.C.) and his student Aristotle (384-322 B.C.) who has profoundly affected modern thinking in psychology and in many other fields. Both of them are the originators of rationalist and empiricist. A rationalist is one who believes that there is a route to knowledge is through logical analysis. In contrast, Aristotle's approach is that of an empiricist, the one who believes that we acquire knowledge via empirical evidence, obtained through experience and observation.

In Aristotle's view, then, it leads directly to empirical investigations of psychology, whereas Plato's view foreshadows the various uses of reasoning in theory development. But, most psychologists today seek a synthesis of the two: They all base empirical observations on theory but in turn of using these observations to revise their theories. To elaborate on Aristotle's ideas, Kemp (1996, 2000) [19] attempted to locate cognitive processes in the brain and to prove to have little to do with our current understanding of the brain. Furthermore, The German philosopher Immanuel Kant (1724-1804) [20] began the discussing empricalism versus rationalism. His impact on philosophy interacted with the nineteenth-century scientific exploration of the body and how it works to produce profound influences the eventual establishment of psychology as a discipline in the 1800s.

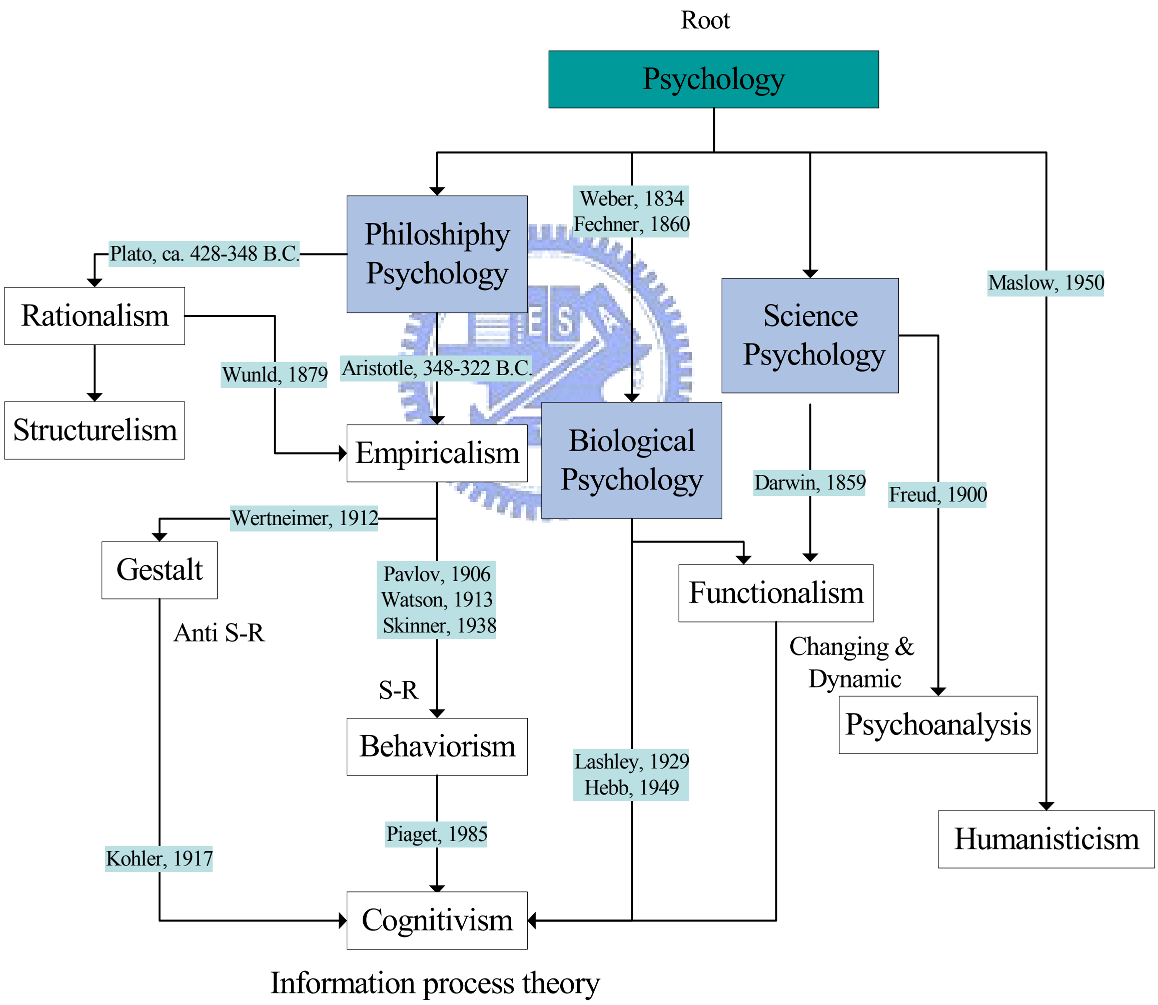

Figure 1. Evolution of Cognition Psychology

Wilhelm Maximilian Wundt was a German physiologist and Psychologist who made Psychology a field of its own. He was the first person in history to be called a “psychologist,” as well as the first person to teach a course in Physiological Psychology at Heidelberg in 1867. Wundt established psychology as a unique branch of science with its own questions and methods. Wundt was the first person to take all of the nineteenth century’s sprouting of the new psychology onto the old and creating his new science, and published a book on physiological psychology. The form of psychology Wundt called scientific metaphysics. This form of psychology would be used to integrate the empirical work in the lab with other scientific findings, reviewed by Piaget [6].

The philosophical and psychological developments lead to the emergence of cognitive psychology. Developments in other sub-fields also contributed to the development of cognitivism and modern psychology. Karl Spencer Lashley (1890-1958) [21, 22] studied topics not easily explained by simple conditioning, and to embrace methods other than the experimental manipulation of environment contingencies (Gardner, 1985). Lashley was deeply interested in neuroanatomy (the study of the structures of the brain) and in how the organization of the brain governs human activity. Lashley brashly challenged the behaviorist view that the human brain is a passive organ merely responding to environmental contingencies outside the individual; instead, he considered the brain an active and dynamic organizer of behavior. Donald Hebb (1949) [22, 23] was the first psychologist to provide a detailed, testable theory of how the brain could support cognitive processes. His influential work provides a strong foundation for some of the current trends in cognitive psychology. Behaviorists did not jump at the opportunity to agree with theorists like Lashley and Hebb. They thought that psychology should be the science of the behavior analysis but Human mind. From Watson [24] to Skinner [25], they applied their experimental analysis of behavior to almost everything, from learning to problem solving and even to the control of behavior in society. The other such as, Functionalism, was a major paradigm shift in the history of American psychology. As an outgrowth of Darwin’s evolutionary theory, the functionalist approach focused on the examination of the function and purpose of mind and behavior. Rather than the structures of the mind, functionalism was interested in mental processes and their relation to behavior. William James, a functionalism, became known to influence the psychology. The following, G. Stanley Hall, Mary Calkins, and Edward Thorndike are spreading functionalist psychology as well. As for Gestalt psychology, it is founder is Max Wertheimer. Those psychologists started to focus on "pattern" from "Gestalt", "Form", and "Configuration". They declared that Behavior is equal to the function of Human and Environment. Each pattern is sensitive to each case respectively. As this description, the definition of behavior is not purely only a set of "Stimulus-Response". For these instances above, they all were the emergences of cognitive psychology.

Generally, cognitive psychology is a science of the research of human cognitive process. The Switzerland philosopher Jean Piaget, originally a biologist, is now best remembered for his work on the development of cognition. Piaget (1985) [26] suggested that learning process is iterative, in which new information is shaped to fit with the learner's existing knowledge, and existing knowledge is itself modified to accommodate the new information.

Table 1. Arguments of Psychology to Soft Computing Techniques.

Psychology Branch	Specific Arguments	Soft Computing Branch
Behaviorism[24,25]	Stimulus-Response (S-R)	AI-based learning[1,2,27]
Biological Psychology[21,22,23]	Neurology, and Brain Theory	Neural Network[27]
Darwin-Science Psychology, [James]	Natural Selection, Theory of Evolution	Genetic Algorithm, Genetic Programming[32]
Gestalt[6]	Involve "Human" Factor and Anti S-R	None

To sum up the statements, the roots of the cognitive movement are extremely varied: It includes gestalt psychology, behaviorism, even humanism; it includes thinkers from linguistics, neuroscience, philosophy, and engineering; and it especially involves specialists in computer technology and the field of artificial intelligence. Cognitive psychology is far more sophisticated and philosophical than behaviorism. It does, of course, have the tremendous advantage of being tied to the most rapidly developing technology we have ever seen -- the computer. But more and more people saw AI as ultimately being a good model for human beings, and they are confused about cognitive psychology and other sub-psychology. For the reason, to develop the new human-thinking model to aggregate knowledge should be understood the psychology theory first even cognitive psychology. After all, the history of psychology is more continuous and complete for a long time than AI techniques. In Table 1, this work tries to summary some relationship about sub-psychology to soft computing.

2.3 Relationship of Cognitive Psychology and Classifier System

In 1956 John McCarthy regarded as the father of AI, organized a conference to draw the talent and expertise of others interested in machine intelligence for a month of brainstorming. He invited them to Vermont for "The Dartmouth summer research project on artificial intelligence." From that point on, because of McCarthy, the field would be known as Artificial intelligence. Artificial Intelligence (AI) belongs to the area of computer science focusing on creating machines that can engage on behaviors that humans consider intelligent. Today with the advent of the computer and 60 years of research into AI programming techniques, the dream of smart machines is becoming into reality, which is concluded from [27].

Learning Classifier Systems (LCSs) are a machine learning paradigm introduced by Holland (1986) [28], also the father of genetic algorithms. Before that, Holland and Reitman

(1978) [29] made their first appearance in the paper “Cognitive Systems Based on Adaptive Algorithms”. While there was still considerable research in the 1980s, the field began to wane at the end of the decade. In the early 1990s, learning classifier systems seemed too complicated to be studied, with few successful applications reported. In the mid 1990s the field appeared almost at a dead end. But, during the last five years, new enhanced models have been developed and new applications have been presented which caused a great resurgence of this area.

No matter the father of AI, McCarthy or the father of LCSs, Holland, both of them led the development of AI techniques and LCSs respectively. They caused the confusing definition of learning in various AI researches. Those researches all emphasized that the learning process of AI techniques is cognition. Therefore, AI researches all were developed the learning approaches that own the cognitive concept. Actually, their cognitive concept only presents the trial and error learning. It deserves to be mentioned that McCarthy [30, 31] has ever tried to stand at philosophy or psychology to redefine the learning and knowledge representation and taken seriously the idea of actually making an intelligent machine. Furthermore, McCarthy went on to the notions of metaphysically and epistemologically adequate representations of the world and then to an explanation of can, causes, and knows in terms of a representation of the world. Besides, he also reviewed the work in philosophical logic in relation to problems of artificial intelligence and a discussion of his previous efforts to program “general intelligence” from the point of view of this paper. Such as the above description, McCarthy at least knew that artificial intelligence should be enhanced by philosophy at that time.

Until Holland, cognitive system was mentioned the term, cognition, but not sufficient by the cognitive psychology. Maybe it is the reason that the concept of cognitive system was implemented and named to learning classifier systems (LCSs’s) by Holland [28]. However, cognitive psychology was brought into vogue after 1985 by Piaget [26]. And more cognitive models would be invented after Piaget. Nevertheless, few learning classifier systems focused on them and just enhanced the original Holland's one. In spite of the development of cognitive psychology, those following LCSs never refocused on the cognition definition, and the relative LCSs recognized their models owning the "cognitive" ability after Holland [28] (1978), for instance, ZCS or XCS and so on. Thereby, this work is given the strong suspicion that LCSs are not sufficient to the Cognition. That is, Table 2 additionally shows the description of the classifier systems for what aspects matching to the psychology by the suspicion.

Without surveying the relative cognitive studies and reconcentrating the cognition definition and cognitive model, the novel cognitive system is not easily to develop. Especially, when the cognitive psychology develops based on the thousands years of the psychology history, and it has already mimicked the human mind by several approaches to discovery the cognition process of human and knowledge aggregation.

Table 2. Arguments of Psychology to Classifier Systems.

Psychology Branch	Specific Arguments	Soft Computing Branch
Functionalism	LCSs don't have sufficient functions to the Cognition Mechanism, although it is based on solving the changeable issues in the dynamical environment. That is just satisfied to the Functionalism, but even Gestalt without "Human".	Classifier Systems[this work]

Chapter 3. "Evolved Learning" of XCS and Learning of Cognition

3.1 Introduction

Owing to the ambiguous of these two learning definitions, the evolved learning and the cognition learning are verified in different period. Besides, various kinds of researchers, such as biologists and philosophers gave the different definitions. Thus, identifying the learning definition would be the main job at the first job, and these two kinds of learning would be detailed next. Then, this work will combine their advantages to develop a high performance knowledge-learning framework that would be discussed first.

3.2 Evolved Learning

Traditional Evolved Learning is derived from Darwin's Evolution Theory which is the widely held notion that all life is related and has descended from a common ancestor: Complex creatures evolve from more simplistic ancestors naturally time over time. In a nutshell, as random genetic mutations occur within an organism's genetic code, the beneficial mutations are preserved because they aid survival -- a process known as "Natural Selection." These beneficial mutations are passed on to the next generation. Over time, beneficial mutations accumulate and the result is an entirely different organism. For this aspect, the evolved learning is easily described by Darwin. However, the learning result of this mechanism is sure, and optimal accuracy is verified. But the results would not be evolved the same by different times in insufficient samplings. In other words, the optimal results might be the certain case under the critical opportunity or effected by the certain factors in a specific environment.

3.2.1 Dynamical Evolved Learning

Genetic algorithms are based on a biological metaphor: They view learning as a competition among a population of evolving candidate problem solutions. A "fitness" function evaluates each solution to decide whether it will contribute to the next generation of solutions. Then, through operations analogous to gene transfer in sexual reproduction, the algorithm creates a new population of candidate solutions.

John Holland's pioneering Adaptation in Natural and Artificial Systems [32] (1975) described how an analog of the evolutionary process can be applied to solving mathematical problems and engineering optimization problems using what is now called the genetic algorithm (GA). Holland had two aims: to improve the understanding of natural adaptation process, and to design artificial systems having properties similar to natural systems. The basic idea is as follow: the genetic pool of a given population potentially contains the solution, or a better solution, to a given adaptive problem. This solution is not "active" because the genetic combination on which it relies is split between several subjects. Only the association of different genomes can lead to the solution. No subject has such a genome, but during reproduction and crossover, new genetic combination occurs and, finally, a subject can inherit a "good gene" from both parents. Holland method is especially effective because he not only considered the role of mutation (mutations seldom improve the algorithms), but he also utilized genetic recombination, (crossover): these recombinations, the crossovers of partial solutions greatly improve the capability of the algorithm to approach, and eventually find, the optimum.

3.2.2 Trial & Error Learning

Edward L. Thorndike (1943) [33] claimed that "A good simple definition or description of a man's mind is that it is his connection system, adapting the responses of thought, feeling, and action that he makes to the situation he meets." He worked on educational psychology and the psychology of animal learning. As a result of studying animal intelligence, he formulated his famous "law of effect", which states that a given behavior is learned by trial-and-error, and is more likely to occur if its consequences are satisfying. Thorndike's early experiments (1898 - 1911)[34] involved a hungry cat put in box that contains a concealed mechanism operated by a latch learning involves the goal of the cat manipulating the latch, opening the door, finding food, and eating initial random behavior is followed by the cat "catching on" and quickly opening the door.

Thorndike maintained that, in combination with the "law of exercise", the notion that associations are strengthen by use and weakened with disuse, and the concept of instinct, the law of effect could be explained to all of human behavior in terms of the development of myriads of stimulus-response associations. Briefly it is worth briefly comparing trial and error learning with classical conditioning. In classical conditioning a neutral stimulus becomes association with part of a reflex. In trial and error learning no reflex is involved. A reinforcing or punishing event (a type of stimulus) alters the strength of association between a neutral stimulus and quite arbitrary response. The response is not to any part of a reflex. The behaviorist points out that human behavior could be explained entirely in terms of reflexes, stimulus-response associations, and the effects of reinforcers upon them entirely excluding 'mental' terms like desires, goals and so on was taken up by John Broadhus Watson[35].

As for the reinforcement learning of XCS [36], its major thread concerns learning by trial and error and started in the psychology of animal learning. This thread runs through some of the earliest work in artificial intelligence and led to the revival of reinforcement learning in the early 1980s. This thread began in psychology, where "reinforcement" theories of learning were common. Perhaps the only first person to succinctly express the essence of trial-and-error learning was just Edward Thorndike [34]. This essence was taken to be the idea that actions followed by good or bad outcomes have their tendency to be re-selected altered accordingly. Additionally, in spite of the original development of each AI or the application of it, they more or less involved the trial & error method to invent their model. For instance, the most basic method of training a neural network is trial and error. Change the weighting of a random link by a random amount if the network isn't behaving the way it should. Undo the change and make a different one if the accuracy of the network declines. It takes time, but the trial and error method does produce results.

3.3 Cognitive Learning

Cognitive psychology is a theoretical perspective that focuses on the realms of human perception, thought, and memory. It portrays learners as active processors of information--a metaphor borrowed from the computer world--and assigns critical roles to the knowledge and perspective students bring to their learning. What learners do to enrich information, in the view of cognitive psychology, determines the level of understanding of that they ultimately achieve.

Cognition is defined as “the mental process or faculty of knowing.” To help the students reach a cognitive state about a certain subject should be one of the goals of both teaching and learning. Thus, the below discussions were the teaching learning and the rehearsal learning.

3.3.1 Teaching Learning

As articulated by Piaget (1969)[37], students learn better when they can discover knowledge through the way of inquiry and experimentation instead of acquiring facts presented by a teacher in class. Since the learner is portrayed as an active processor who explores, discovers, reflects, and constructs knowledge, the trend to teach from this perspective is known as the constructivist movement in education. As Bruning (1995)[38] explains, "The aim of teaching, from a constructivist perspective, is not so much to transmit information, but rather to encourage *knowledge formation* and development of metacognitive processes for judging, organizing, and acquiring new information." Several theorists have embellished this theme. Rumelhart (1981)[39], following Piaget, introduced the notion of *schemata*, which are mental frameworks for comprehension that function as *scaffolding* for organizing experience. At first, the teacher provides instructional scaffolding that helps the student construct knowledge. Gradually, the teacher provides less scaffolding until the student is able to construct knowledge independently.

Recently, there has been some interests in developing formal models of teaching [40, 41, 42, 43, and 44] through which we can develop a better understanding of how a teacher can most effectively speed up the training process. Although, the formal models of teaching that have been introduced in the learning theory community is that they place stringent restrictions on the learner to ensure that the teacher is not just providing the learner with an encoding of the target. In particular, the teaching models allow the teacher to present a set of examples for which only the target function is consistent. Thus, teaching under these models is made unnecessarily difficult since the problem reduces to teaching an obstinate learner that tries as hard as possible not to learn while always outputting a hypothesis consistent with all previous examples. In other words, teaching learning is necessary to a learner to reduce the complexity learn process.

3.3.2 Reinforcement-Rehearsal (R-R) Learning

Reinforcement Learning

There are several kinds of learning theories from behaviorists. You may be familiar with "conditioned response theory" developed by Pavlov 1903, whereby a response that already occurs in the presence of one stimulus can be "conditioned" to occur following a different stimulus. This learning theory is very important for emotional learning, but has little relevance to most learning of invariant tasks. Far more relevant is “reinforcement theory,” first developed by E. L. Thorndike (1913) [33] and further developed by B.F. Skinner (1956)[24] and others. In reinforcement theory, an invariant task is viewed as a “response” and is learned when it becomes “associated” with an appropriate stimulus. For example, “3.14” is a response that should become associated with “Pi”. This learning process occurs whenever “reinforcement” follows the response. For example, each time a learner responds with “3.14”, a reinforcer such as “Right!” or “Good!” or even just a smile with a nod will increase the probability of the learner responding the same way in the future. With sufficient repetition of these stimulus-response-reinforcement events, the response will come to occur automatically in the presence of the stimulus.

Also, the learning classifier system is a machine learning system with close links to reinforcement learning and genetic algorithms. LCS consists of a population of binary rules on which a genetic algorithm altered and selected the best rules. Instead of a using fitness function, rule utility is decided by a reinforcement learning technique.

Rehearsal Learning

Besides the reinforcement learning, rehearsal learning differs from it. A rehearsal strategy is used by the repeated practice of information to learn it. When a student receives the specific information that needs to be learned, such as a list, often he will attempt to memorize the information by repeating it over and over. He may read the words out loud, or he may sub vocalize the information (read it in his own mind). The repeated practice increases the student's familiarity with the information. For many people, the learning of our social security number, our telephone number, or the items we want to pick up at the grocery store prompts us to use a rehearsal strategy.

This strategy originally documented by Belmont and Butterfield (1971) [45] examines how regular review and recall techniques aid the transfer of information into LTM. Buzan

[46] goes on to propose a pattern that the rehearsal strategy should follow. By monitoring recall rates during, and immediately after learning has taken place and at timed intervals thereafter, Buzan concludes that “The first review should take place about 10 minutes after a one hour learning period and should itself take 5 minutes. This will keep recall high for approximately one day when the next review should take place, this time for a period of 2 to 4 minutes. After this, recall will probably be retained for approximately a week, when another 2 minutes review can be completed followed by a further review after about one month. After this time the knowledge will be lodged in LTM”.

Rehearsal strategies can be used to learn relatively brief amounts of information, and is good for learning “foundation information” or “correct information”. Foundation and correct information is necessary to be learned before more complex learning can take place. If you are using rehearsal to teach information that contributes to a larger concept or skill, keep in mind that lots of practice may be required for the students to learn the information to a level of automaticity. After initial learning takes place, you will need to review many times to ensure that the students have retained the information. We have all memorized information that we have promptly forgotten when we stopped rehearsing. For example, it is more concerning that “**3.14**” is a “**True**” response that should become associated with “Pi”. This learning process occurs whenever “Rehearsal” follows the response. Contrary to the “Reinforcement”, “**314**” is a “**False**” response that should become associated with “Pi”. This learning process occurs whenever “Reinforcement” follows the response. It is still practicable in the reinforcement learning process.

In spite of the mechanism of LCSs, it has the reward ability similar to the rehearsal learning as well. The truth of the rehearsal learning cognition is that teachers take the foundation or correct information to educate the students and students practice the information by themselves. The proper correct information or knowledge is worth to do the rehearsal. That is the difference of reinforcement and rehearsal. Furthermore, the fullness explanation of information process theory, the narrow terms of cognitive psychology would be detailed next.

3.4 Information Process Theory

There are at least two major kinds of cognitive theory relevant to learning invariant tasks: information-processing theory and schema theory. According to the information-processing model of learning (see Figure 2), there is a series of stages by which new information is learned (Gagne, 1985) [47]. Information is received by receptors (such as the eyes and ears), from which it is passed to the sensory register where all of it is held, but for only a few hundredths of a second. At this point of view, selective perception acts as a filter which causes some aspects of the information to be ignored and others to be attended to. For example, the ears (receptors) receive the sounds comprising "Pi equals 3.14," along with various other background sounds, and all those sounds are passed on to the sensory register in the brain. Then through the selective perception process, some of the information (hopefully the "Pi equals 3.14") is attended to the part.

That information which is attended to is transformed and passed on to short-term memory, which can only contain a few items of information at a time (depending on their complexity). For instance, if "Pi equals 3.14" is attended to, it is then passed on to short-term memory, where it might be said to "echo" for a few seconds, and the echoing can be prolonged through rehearsal." Items can persist in short-term memory for up to about 20 seconds without rehearsal, but with constant rehearsal they can be retained indefinitely.

Finally, the information may be passed on to long-term memory. This process is called encoding to memorize. For example, if appropriate encoding processes are exercised to link the "Pi equals 3.14" with prior knowledge, then the information is passed on to long-term memory. In the traditional model of human memory (Atkinson and Shiffrin, 1968 [48]; Waugh and D. A. Norman, 1968 [49]), immediate free recall yields items directly retrieved from a temporary short-term memory (STM) and items retrieved by retrieval cues from a more durable storage in long-term memory (LTM).

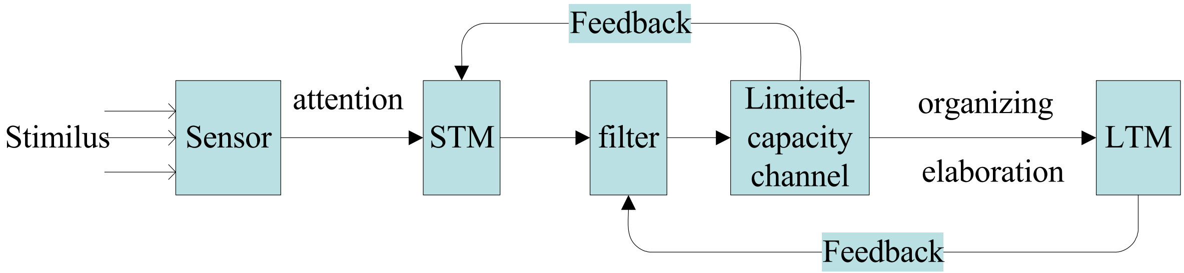

Figure 2. Information Process Theory proposed by Gagne [47], [50].

3.4.1 Short-Term Memory

Short-term memory (STM) lasts from a few seconds to a minute; the exact amount of time may vary somewhat. For instance, when you are trying to recall a telephone number that was heard a few seconds earlier, the name of a person who has just been introduced, or the substance of the remarks just made by a teacher in class, you are calling on short-term memory. STM is assumed to have a limited capacity (G. A. Miller, 1956)[51], and when attention is diverted to another demanding task, information originally stored in STM becomes unavailable.

3.4.2 Long-Term Memory

By contrast, long-term memory (LTM) lasts from a minute or so to weeks or even years. From long-term memory you can recall general information, which is valuable information, usually called to knowledge, about the world that you learned on previous occasions, memory for specific past experiences, specific lectures previously learned, and the like. The storage capacity of LTM is assumed to be vast and much more durable than that of STM. Storage in LTM is assumed to be primarily associative, relating different items to one another and relating items to attributes of the current situation (current context). The time required for storage of a new retrievable memory trace in LTM has been estimated to be relatively long--about ten seconds (Simon, 1973)[52].

3.4.3 Working Memory

In addition to LTM and STM, models of working memory (WM) have focused on the availability of information in STM which has limited to the capacity. No model of WM can reasonably allow greater working capacity during performance of a specific task than the maximal capacity of working memory measured in a pure memory task. That is, the capacity of WM must be much less than STM (G. A. Miller, 1956). Such a severe limit on WM might seem far too restrictive to allow for human performance levels.

Newell and Simon (1972) [53] proposed a production-system architecture for cognitive processes that has influenced most subsequent efforts to build models and theories. In this architecture the conditions of a large number of productions (condition-action pairs) are matched against the currently active elements (working memory). Such as Anderson's (1983) [54] ACT*, WM is the transiently activated portion of LTM. The limits on the number of elements in WM are not determined by a fixed number but rather by the amount of available activation. In his work on building ACT* models of cognitive processing Anderson found that WM can sometimes contain over 20 units at one time. To reconcile such a large capacity of WM with the much smaller capacity of STM, Anderson argued as follows: The activation of elements decays very rapidly. For this reason the number of units that can be actively maintained long enough to be included in immediate recall is much less than all of the information activated at the start of recall. Most investigators argue, however, that the capacity of WM must be far greater than the capacity of traditional STM (Newell, 1990 [55]).

In spite of the discussion of limited capacity of WM, the function of WM would be flexible to a branch from LTM or STM.

3.4.4 Accumulation of Knowledge

In the "Knowledge Society", there are two constants – continuous change and increasing volumes of information. Knowledge and skill currency can only be maintained in an era of such rapid change through active engagement in lifelong learning and the deployment of effective learning strategies. This subject draws upon relevant recent research and theory in the area of cognitive psychology to provide the knowledge and skills necessary to move beyond rhetoric to effective educational practice. Furthermore, cognitive psychology defined Cognition as the acquisition of knowledge [56]. In the others, cognitive psychology defined Knowledge as the storage and organization of information in memory [57].

The awareness of that knowledge more than heuristics or search strategies is at the core of much human cognitive functioning has also led applied researchers to pay much more attention to memory based knowledge structures. Accuracy of the model outcome is determined considerable by the amount and nature of the knowledge available and how effectively a model retrieves this knowledge from LTM. Upon the research of human information processing theory, understanding the mechanism of memory is necessary to how the knowledge is stored.

3.4.5 Summary

Using a computer as a metaphor for memory, the short-term phase is RAM (highly volatile and easily lost when some others else are entered), but long-term memory is such as a hard drive or diskette (the information is stored there even after the machine is turned off). This metaphor is especially helpful because a computer knows the address of each bit of information because of the manner information is entered. It is essential that information placed into a student's long-term memory be linked in a way that the student can retrieve it later. The teacher who should understand the relationship between memory and retrieval can lay out a lesson plan to assist the student in the process and enhance his learning.

As the pre-statement portrayed, while rehearsal is important to short-term memory, it can also be used to transfer information to long-term. Elaborating or making material memorable will also enhance the student's learning process. The effective teacher will elaborate and rehearse material so that the student can remember the information more easily. That is the reason the input material is high relevant to memorize to form valued-information, knowledge.

Organization of material into long-term memory involves sorting, relating, arranging, and grouping information so that it can be worth to been memorized. It is important to note that most application AI models have more trouble remembering/learning of what data they should remember/learn. Therefore, as the effective teacher will help the memory process by introducing the student to various organizational techniques. And if great teaching effects intended learning outcomes, learning is achieving those intended outcomes. However, the learning of teaching style is more various than traditional learning.

3.5 Conceptual Framework

During the Middle Period (mid 1900s), Knowledge is just thought of as the transformation of sensory inputs into associated thought, and the realization that sensory inputs are transformed prior to storage. In the early twentieth century, Knowledge is still considered as a framework of stimulus and response (S-R). The profound breakthrough of this period is that by studying S-R, one can gain insight into the working of cognitive knowledge. This research and its viewpoint of knowledge learning are largely based on narrow term of cognitive psychology, information processing theory. Besides, S-R of cognitive psychology research is historically analogous to the black box testing. Following these two aspects, this work applied the cognitive learning to modify the learning process of traditional soft techniques to increase the efficiency of forming knowledge storage. Furthermore, according to the accuracy ratio of LCSs model, we choose its best ones, XCS, as a kernel of that black box, we tested as a memorizing/learning model. We combine the information process theory and learning type to initial the concept of the dual learning mode framework, shown as Figure 3. It contains two parts: Knowledge Education learning and Reinforcement-Rehearsal(R-R) learning.

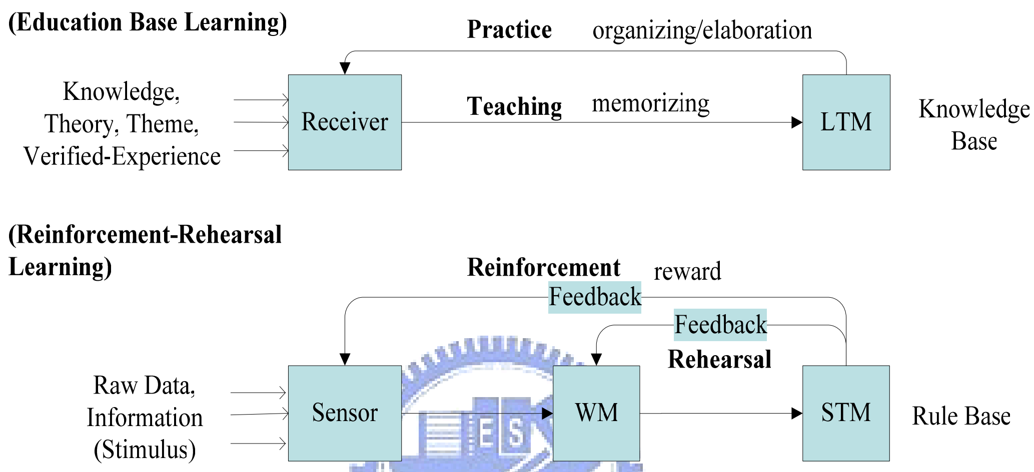

Figure 3. Education Learning Flow and Reinforcement-Rehearsal Learning Flow.

To construct an effective learning model, two aspects should be considered. Table 3 concludes some attributes of the proposed dual perspective model, Education and R-R perspective. The first is education base summarized by the teaching literatures. Knowledge is worth to be used as materials to teach students/train models. Nevertheless, using knowledge rule to build an expert system is not sufficient flexibility. That is, the learning/training model which utilized the knowledge materials as inputs is taken into consideration to form a learning model with education perspective. In this part, the model is as a learner, and the model operator as a teacher. The memory belongs to long-term memory (LTM) with permanent store. Knowledge transmission and the model that learns others' thinking are the two major purposes.

As for the superficial knowledge by the other learning model with reinforcement-rehearsal (R-R) perspective, it would be sum up the general experiences which come from stimulus-response actions. R-R perspective learning model is just like traditional soft computing techniques. Normally, utilizing huge amounts data as inputs to the "learning" model is the machine learning type which major method is trial and error style. In this part, it owns working memory (WM) and short-memory (STM). WM is a pre-storage of the stimulus-response. STM is a storage that maintains the short-term information in the model for rehearsal. If it is possible, the relative experiences would be concluded to rules which could be verified to form knowledge. Objectively, the entire process of this part is no efficiency because of trial and error method. That is because of the complexity of the learning, this work portrays the Education and R-R learning model to increase the efficiency of knowledge transmission and the accuracy of experience rules generation.

Table 3. Dual-Mode Learning Model of Education-Dominated and R-R Perspectives

Learning Attributes	Education Perspective	R-R Perspective
Subject	To Teach/Train the Model	Model as Learner Centered
Input source	Knowledge	Raw Data
Learning type	Teaching Style Learning	Trial and Error Learning
Process steps	Model Memorizes Knowledge	Model Summarizes the Experiences
Model type	Model as Memorizer	Model as Processor
Memory type	Rote Long Term Memory	Active Short Term Memory
Memory Capacity	Unlimited	Limited
Practice type	Repetitive	R-R
Output	Knowledge is stored	Experiences Rule is Created, but Need to be Verified
Instruction type	Sequential Instruction	Adaptive Learning
Training Flow	Operator Manages Model Learning	Model Self-Tuning by Pre-parameter
Thinking Type	Model Learn Others' Thinking	Model Develop and Reflect on its Self-Own Thinking
Knowledge	Knowledge Transmission	Knowledge Formation by Verified Experiences
Operator type	Operator as a teacher	Operator as a Data Inputter
Model type	Mechanistic/Training	Organismic/Evolution
Performance	High Efficiency	Low Efficiency
Flexibility	Low, Difficult to modify	High Flexibility but Need to Be Verified

Chapter 4. Education and R-R Model Based on XCS

According to the cognition theory, knowledge transmission by education is proven as a high efficiency mechanism of learning to human. As the human learning, the teaching-base learning style to form knowledge should be paid more attention on. As for machine learning, the machine is a software system running on a computer that could provide the ability to large continuous logical processes, while many kinds of learning algorithms are analogous to the human trial and error learning. Thus, this work combines the advantages of these two aspects to propose a dual perspective learning model which is implemented the conceptual framework that describes in chapter 3.

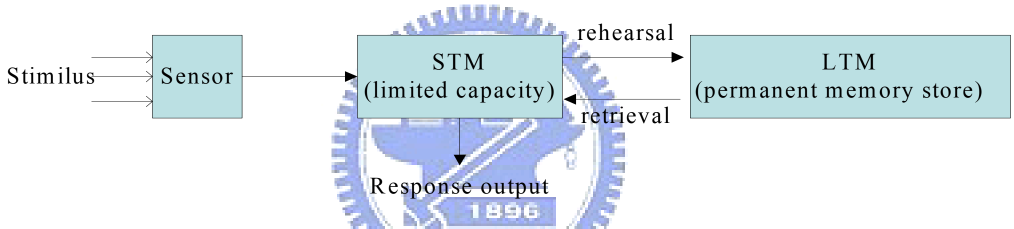

Figure 4. Richard Atkinson and Richard Shiffrin 1968 [48] proposed a theoretical model for the flow of information through the human information processor.

The dual perspective learning concept is knowledge education and reinforcement-rehearsal respectively. In addition to involve the cognition theory, information process theory (IPT) [47], [48], the relationship of LTM and STM are rehearsal and retrieval, shown as Figure 4. Although LTM has ever been mentioned by the original LCS, the definitions of LCS's LTM and IPT's LTM are different. In Fact, the function of LCS's STM is just equal to the function of IPT's WM, and the same aspect, LCS's LTM is equal to IPT's STM. As for LTM of IPT, it indeed owns an unlimited capacity to store information which is different from LCS's memory. While the inference of these memories is derived, the conceptual framework would enhanced by more considering the retrieval relation of LTM and STM, shown as Figure 5. LTM is presented to store knowledge base, and rules are collected in

STM. According to IPT, the relation of LTM and STM is existed. Although the availability and accessibility of retrieval memory from cognitive psychologists is still a troublesome problem, we won't discuss about it in this work, because LTM seems as a knowledge base with high priority to retrieve first.

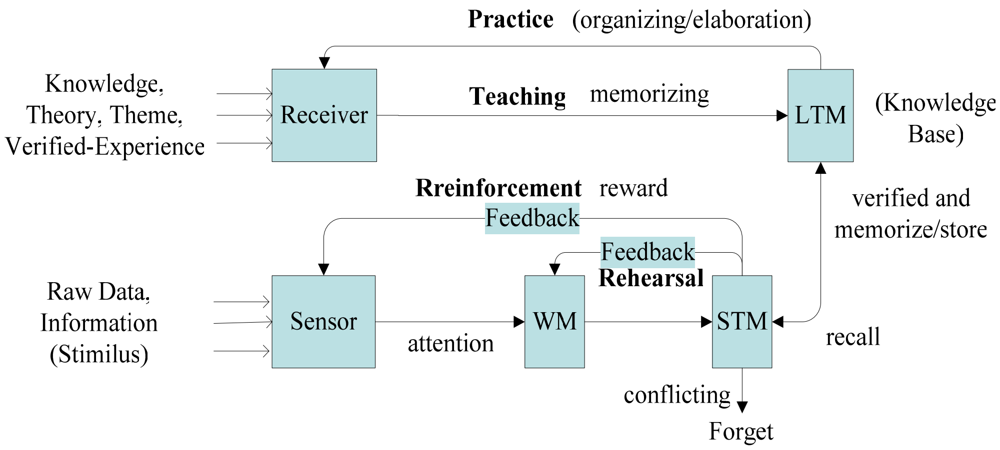

Figure 5. Dual perspective learning process of Education and R-R mechanism.

The following sections are described the development of the proposed dual-mode learning model. In 4.1, XCS is first detailed. After that, R-R XCS which involves rehearsal style is proposed. In final section 4.3, the dual-mode learning model referred to Figure 5 is proposed to enhance XCS.

4.1 XCS

Most machine learning techniques are developed by information process theory. No matter partial application of IPT concept or applying the entire flow of IPT, they all simulated various operations of memory. For example, those neural network types are applications of neuroanatomy. According to that, it is necessary to define the neural structures of the brain simulated as memory. The others would be evolution computing types, such as GA, GP, and LCSs. Among them, LCSs has flexible outcome on rule generation which represents information about the structure of the world in the form of rules and messages on an internal message list, such as its STM or LTM. The system can be used as the message list to store information about (a) the current state of the world (response), and (b) about previous states (stimulus). From now on, LCS has the ability to store rule according to the input information. Moreover, Wilson's XCS [36] is a recently developed learning classifier system (LCS) that differs in several ways from more traditional LCSs. In XCS, classifier fitness is based on the accuracy of a classifier's pay-off prediction instead of the prediction itself. As a whole, the genetic algorithm (GA) takes place in the action sets instead of the population. XCS's fitness definition and GA locus together result in a strong tendency for the system to evolve accurate, maximally general classifiers that efficiently cover the state-action space of the problem and allow the system's 'knowledge" to be readily seen. As a result of these properties, XCS has been considered and focused to the kernel of the proposed model in this work.

The detailed loop is shown in Figure 6, and the current situation is first sensed and the detector received the input from the environment. Second, the match set [M] is formed from all classifiers [N] that match the situation. Third, the prediction array [PA] is formed based on the classifiers in the match set [M]. [PA] predicts for each possible action a_i, the resulting pay-off. Based on [PA], one action is chosen for execution and the action set [A] is formed, which includes all classifiers of [M] that propose the chosen action. Next, the winning action is executed. Then the previous action set $[A]_{-1}$ (a previous action set) is modified by using the Q-learning-like payoff quantity P which is a combination of the previous reward p_{-1} and the largest action prediction in the prediction array [PA]. Moreover, the GA may be applied to $[A]_{-1}$. If a problem ends on the current, time-step (single-step problem or last step of a multi-step problem), [A] is modified according to the current reward, p, and the GA may be applied to [A]. The loop is executed as long as the termination criterion is not met. A termination criterion is a certain number of trials/inputs. As for the XCS detailed functions, they are listed in Appendix A.

Finally, XCS's architecture is much neater than that of previous models; accordingly

XCS is easier to study and analyze since it puts more the role of the various LCS components in evidence. As for the following section, the pre-proposed model and the entire proposed model, R-R learning model and E&R-R learning model, are both derived XCS.

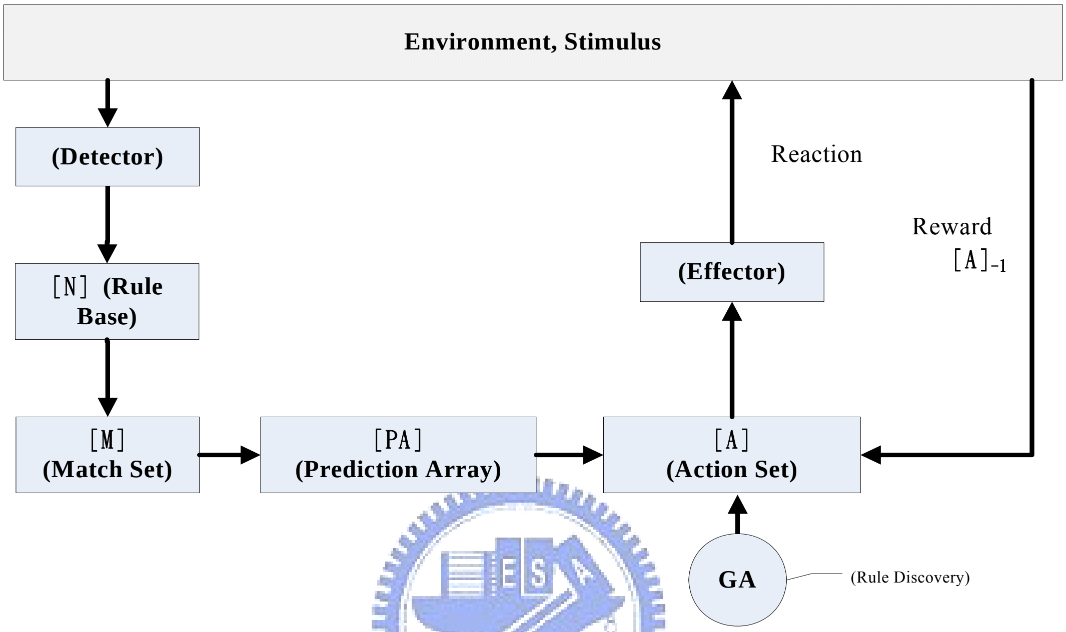

Figure 6. XCS Procedure.

4.2 R-R Learning Based XCS Model

After the reinforcement and rehearsal literatures, they would be easily distinguished. The original XCS model owns the reinforcement learning ability but rehearsal. Considering the significance of rehearsal, the value and correct information should be paid more attention and thereby the XCS model would be enhanced to R-R Learning Based XCS (R-R XCS).

For the rehearsal of "foundation information" or "correct information", a repeater should be added after the final step of XCS. In Figure 7, the detailed flow is that the current situation is still first sensed the raw data and the enhanced detector with working memory (WM) received the stimulus from the environment. That is, WM stores the current situation.

Second, the match set [M] is still formed from all classifiers [N] that match the situation. Third, the prediction array [PA] is formed based on the classifiers in the match set [M]. [PA] predicts for each possible action a_i, the resulted pay-off. Based on [PA], one action is chosen for execution and the action set [A] is formed, which includes all classifiers of [M] that propose the chosen action. Next, the winning action is executed. Then the previous action set $[A]_{-1}$ (a previous action set) is modified by using the Q-learning-like payoff quantity P which is a combination of the previous reward p_{-1} and the largest action prediction in the prediction array [PA]. Moreover, the GA may be applied to $[A]_{-1}$. If a problem ends on the current, time-step (single-step problem or last step of a multi-step problem), [A] is modified according to the current reward, p, and the GA may be applied to [A]. From now on, the above steps are the same as XCS flow, but before the effector, the repeater has the function of the judgment on the current loop is finished or the pattern is worthy to repeat/rehearsal to return to the detector. The significance is just a valued information or information, which is necessary to been verified, seems to be repeated again. In other words, for the inputs, stimulus-response pairs, they are tagged different learning weight onto the R-R learning model. As for the entire loop, it is also executed as long as the termination criterion is not met. A termination criterion is a certain number of trials. Furthermore, some R-R XCS functions, the same as XCS ones, are detailed in appendix A.

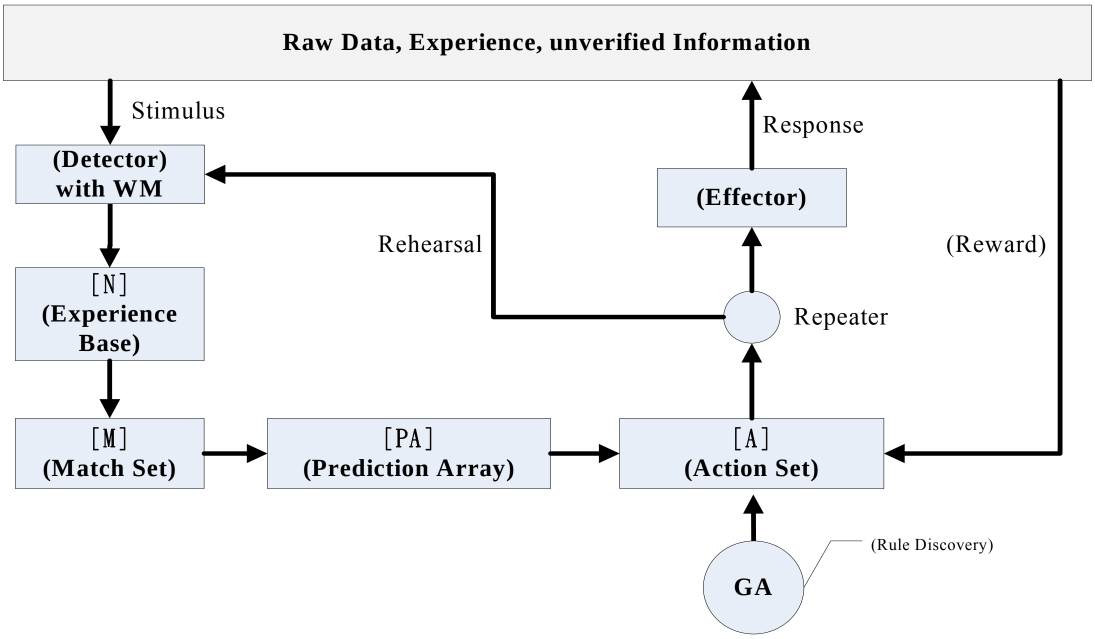

Figure 7. R-R XCS Procedure.

Due to adding the rehearsal mechanism to XCS, the degree of the result should be increased obviously, but the accuracy of the model may not be. As well as, the major purpose for this R-R XCS is that we theoretically pay more attention to the value (correct) information/pattern. In opposition, more system load of R-R XCS actually occurs at the end.

4.3 Education & R-R Based XCS Model

R-R XCS model is an enhanced version from XCS by adding a rehearsal mechanism. Owing to them both adapting GA as an evolution methodology of classifiers and based on XCS, their working accuracy rate should be equivalent by the same training data and testing data. The leverage of R-R XCS to XCS deserves to be mentioned. The rational assumption is that R-R XCS has higher leverage to XCS. This reason originates from R-R XCS considering more value information. But its performance would be decreased and its accuracy ratio might not be better than XCS.

Education & R-R XCS referred to Figure 5 is proposed to increase the accuracy ratio by concerning the education efficiency of learning. In Figure 8, there are two starting points in E & R-R XCS. E & R-R XCS includes R-R XCS discussed in pre-statement. In the additional education learning part, discovered knowledge, verified theory, and defined theorem are all considered as input patterns to the mechanism. Those data should be valued and worthy to "teach" the model or the model be "trained". Thus, we modify the practice route in the education part because this is the "model" not a student and the model is unnecessary to be practiced for more than twice times. For this, those input data would be easily memorized/ stored by the receiver and internalized to the knowledge rule base [N]. Population in knowledge rule base has higher weight or effectiveness than ones in experience rule base. Besides, the detector should consider more about the knowledge rule base [N] than about the experience rule base [N]. WM still stores the current situation in advance. Second, the match set [M] is formed from [N], which is either the knowledge rule base or experience rule base. The following steps are the same with R-R XCS ones. The difference is that the initial-picked population is more from knowledge rule base than experience one. In the mechanism, this kind population from knowledge rule base seems to be "principle". While the entire loop has finished, the new population should be generated from knowledge rule base to the experience one. Some experiences have possibility to produce from the real knowledge, if the knowledge really exists. The education knowledge should be increased to the rule base on the go, while the new knowledge or theory is discovered. Furthermore, while a rehearsal population from repeater to detector occurs, detector should verify the repeated population qualification that it may be transferred to receiver. The knowledge population, that is, does come not only from the outside environment but also from internal mechanism. As for the detailed functions, some of them, the same as XCS ones, are detailed in appendix A.

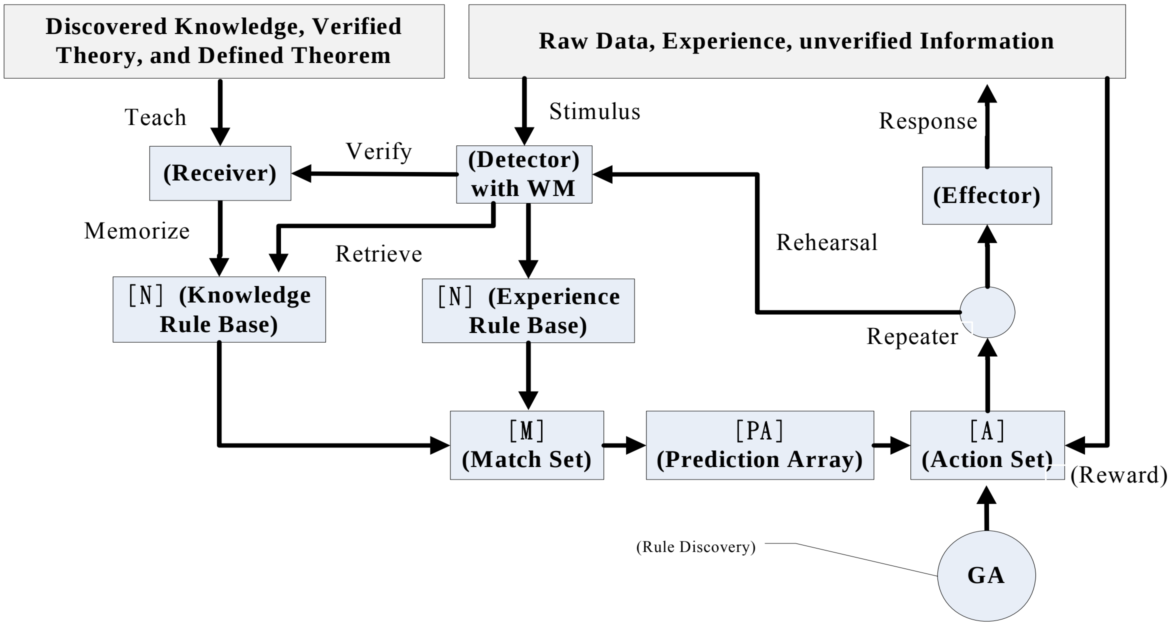

Figure 8. E&R-R XCS Procedure.

Actually, the case of adding new population from detector does not happen easily. E & R-R XCS exactly defines the stern discipline to the knowledge. In simulation, the percentage of knowledge from detector to receiver ones is low. That is, the population in knowledge rule base should be maintained spotless correctness.

4.4 Assumption to Education Materials

In 1987, George Lakoff said, "Meaning is not a thing; it involves what is meaningful to us. Nothing is meaningful in itself. Meaningfulness derives from the experience of a being of a certain sort…" [2, 58]. As an analogy, this paper does not propose a specific transformation process of information to knowledge. This process requires "the experience of a certain sort". Implying the knowledge does not manifest itself without purpose and meaning. Now that the sufficient derivation has been provided as to George Lakoff of this research paper, it is time to proceed to define the assumption of knowledge.

In simulation, this work only recognized the verified information as knowledge. As for the verified information, we first filter the conflict patterns in advanced. According to conflict patterns, one of them, even none of them, would be correct and others might be the wrong experience. In complex problem, the rational assumption is given to that the situation of the case without sufficient samples to analyze would occur. Now that the assumption indeed exists, knowledge would more easily been discovered by the human judgment of the conflict patterns. At least, the best result is that knowledge discovered from these conflict patterns is more efficient than knowledge discovered from the raw data. The worst result is spending narrow time on the judgment but no knowledge discovered. Therefore, the knowledge of education material is initialized by the analysis on the conflict patterns for the simulation.

4.5 Propositions

Following the descriptions of 4.2, 4.3, and 4.4, three propositions of these models are possibly deduced in this section. Their theoretical accuracy and accumulated performance would respectively be detailed as following. The x-axis, Time, in Figure 9, 10, and 11 might means time, or times which is the operating times of the model. The y-axis is just the theoretical accuracy or accumulated performance.

In Figure 9, γ is defined to the difference of the accuracy ratio of R-R XCS and XCS. λ is defined to the difference of the accuracy ratio of E&R-R XCS and XCS. It is sensible that $\lambda >> |\gamma| >= 0$. The reasonable explanation is R-R XCS with rehearsal learning focused on valuable information. When γ is approximate to zero, the two models are applied to the all original data. When $|\gamma|$ is large to zero, the two models are applied to identify the result for valuable information. As for λ, due to the education efficiency of learning, λ should be larger which means the accuracy ratio of E&R-R XCS is much better than XCS one.

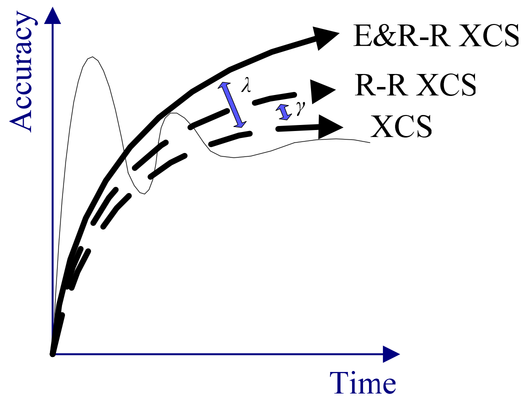

Figure 9. Theoretical Accuracy of XCS, R-R XCS, and E&R-R XCS.

In Figure 10, μ is defined to the difference of the accumulative output of R-R XCS and XCS. It is sensible that $|\mu| >= 0$. The reasonable explanation is just R-R XCS with rehearsal learning focused on valuable information, but its accuracy rate is not absolutely better than XCS one. Indeed, the leverage effect of R-R XCS originates from it focused on more valuable information. If the output is correct and positive to the result, the accumulative output should be increased more. Contrary to the wrong one, the accumulative output should be decreased more as well.

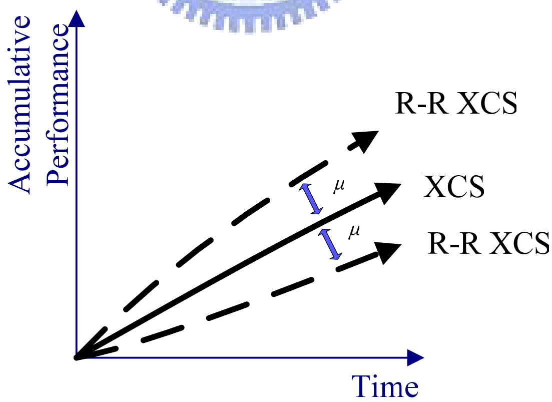

Figure 10. Theoretical-Accumulative Performance of XCS and R-R XCS

In Figure 11, μ_1 and μ_2 are defined to the difference of the accumulative output of E&R-R XCS and XCS. It is sensible that $|\mu_1| >> |\mu_2| >= 0$. The reasonable explanation is that E&R-R XCS has not only the ability with rehearsal learning focused on value information but also involves the education efficiency of learning. Therefore, its accuracy rate is absolutely better than XCS one. Indeed, E&R-R XCS still owns the leverage effect, which originates the same to R-R XCS. Owing to the accuracy ratio increased, the output is usually positive to the result, and the accumulative output should be increased much more.

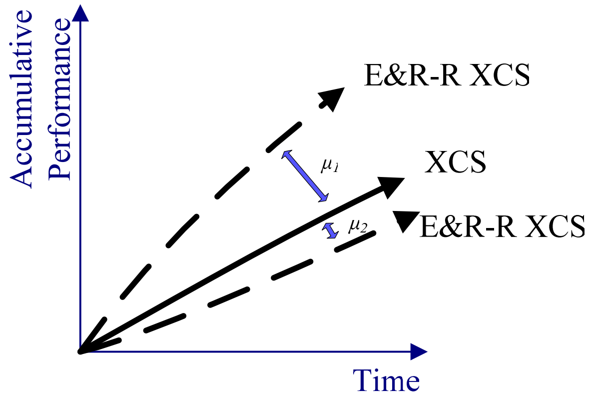

Figure 11. Theoretical-Accumulative Performance of XCS and E&R-R XCS

In a word, the learning accuracy of the proposed E&R-R XCS is much better than XCS. R-R XCS comparing with XCS has the leverage effect to the accuracy and the accumulated performance at least.

Chapter 5. Simulation and Comparison of XCS, R-R XCS, and E&R-R XCS

We applied the proposed dual-mode learning model to financial prediction on global overnight effect to verify the model performance. Therefore, the phenomenon, global overnight effect, would be discussed first. And the prediction model based on the phenomenon would be the application of the proposed learning model. Finally, the simulation results of XCS, R-R XCS and E&R-R XCS would be shown. Besides, their comparison and discussion are figured out.

5.1 Simulation on Finance Prediction

Global overnight effect is a well-known problem originated, which from trading time restriction each economic market [59], [60]. This designed simulation is just for the financial phenomenon that applied each prediction model, which is XCS, R-R XCS, or E&R-R XCS, to predict the stock trend of advance-decline ratio. Based on the complex finance issue, we developed three knowledge models to forecast the local stock market respectively. In the experiments, Dow Jones index (DJi) and Taiwan weight index (Twi) are chosen as reference and predicted markets respectively, and all models are trained by the historical data from markets.

5.1.1 Prediction on Global Overnight Effect

For development of the prediction model, some researches were almost successfully proven that their models were practicable for predicting the trend, but few of them were considered to the relative factors that was exactly affected the results. According to that, we concerned the relative factors from global economic market much more. In the following sections, the relative factors would be firstly mentioned, and then the foundation of prediction flow would be detailed.

5.1.2 Input Factors and Overnight Effect Theory

For the model development and simulation, we choose the Taiwan weighted index (Twi) as an observed market and Dow-Jones index (DJi) as the overnight information reference market. First, the overnight effect theory is figured out in part I, shown as Figure 12. As the concern of the phenomenon, the time series data of stock becomes non-continuous and some trading behaviors maybe exist [61], [62]. Owing to the restriction on trading time, the opening price of the observed market would be affected not only by the overnight effect and the closing price of the previous day, but also by the opening or the closing price of the referent market. As regards the time θ, it is the proposed model starting to predict the trend of the next trading day by the current stock price and the following overnight information (roi_t) of non-trading period. As for part II, it is an inference from part I by DJi substituting for roi_t, and the following equations are the deduced steps.

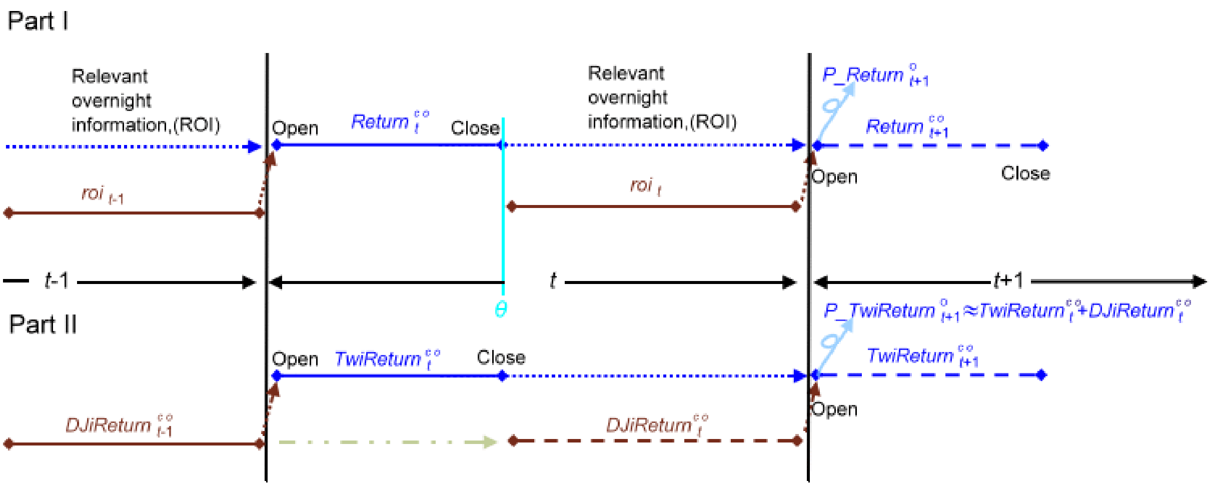

Figure 12. Overnight Effect Phenomenon

1. As part I shown and only one observed market considered, the overnight effect could

affect the next day opening price, which is exhibited by eq.1, while $P\_Return_{t+1}^{o}$ and $Return_t^{co}$ are respectively denoted as the opening price of the next trading day and the trading return of the current day at the closing time.

$$P\_Return_{t+1}^{o} = Return_t^{co} + roi_t \tag{1}$$

2. In this step, the referent market, DJi, is additionally involved. Because the trading time of DJi is just during the time between the closing time and the opening time of Twi next day, the trading return of DJi ($DJi\_Return_t^{co}$) is utilized to substitute the local overnight information (roi_t). In the meanwhile, eq. 1 would represent to eq. 2. $P\_TwiReturn_{t+1}^{o}$ is denoted the prediction return of the t+1 day opening of Twi. $TwiReturn_t^{co}$ is the Twi return of the current day.

$$P\_TwiReturn_{t+1}^{o} = TwiReturn_t^{co} + DJiReturn_t^{co} \tag{2}$$

5.1.3 Prediction Model

Due to the overnight theory and the different time zone between Taiwan and US, the trading time of the DJi between PM 22:30 and AM 05:00 (Taiwan time), which crosses two days, the previous model we proposed contains two stages[63]. The first stage is to generate the predicted return of DJi. Second, the prediction trend of Twi is the final output. Based on this research, we simplify the prediction issue, which been only applied well-known DJi first to predict the trend of Twi.

As regards the input variables, the historical moving average data of price and volume are required based on stock prediction. Therefore, the price-volume moving average data and the moving average convergence-divergence (MACD) extended by the price moving average are both utilized as the input variables which would be translated into bit-string type. Predicting the Twi, about the price, the daily, weekly, monthly, and quarterly moving averages (1, 5, 20 and 60 days) are separately adopted as input variables (4 bits), down-trend (0) or up-trend (1), and their trend-permutation (4!=24 kinds) which would be encoded by 5 bits. As for the volume, its daily, weekly and monthly (1, 5 and 20 days) moving averages are utilized as the input variables (3 bits), down-trend (0) or up-trend (1), and their trend-permutation (3!=6 kinds) which would be encoded by 3 bits. We have already considered the quarterly moving averages (60 days) of volume as input, but it is insensitive. Also, 8 kinds of statuses that encoded by 3 bits are presented by 12 and 26 MACD patterns. Additionally, DJi_Returni (3 bits) should be considered, and total input bits becomes to 21 bits. After the prediction flow, $P\_TwiReturn^{o}_{i+1}$ (3 bits) would be obtained.

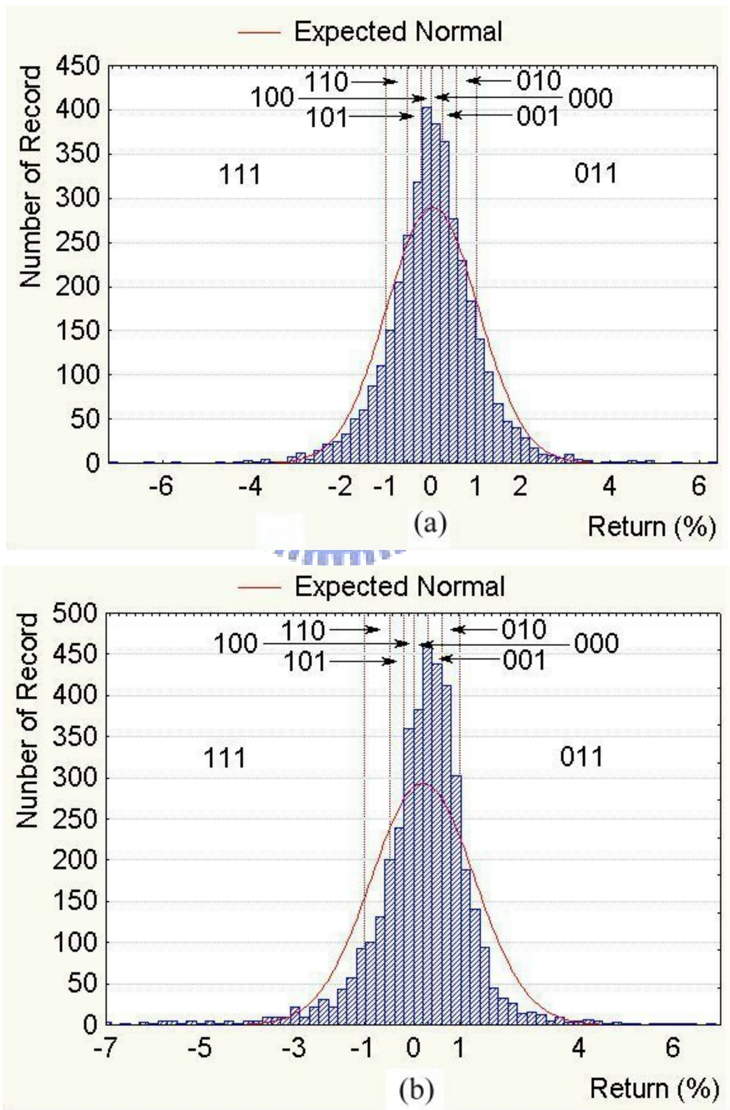

Figure 13. Distribution of Historical Return of (a) DJi and (b) Twi.

In Figure 13, it shows the distribution of the historical return data of DJi and Twi. In order to analyze, this research separates the data into 8 groups equally, 4 up-trends and 4 down-trends. These 8 groups are encoded by c_1, c_2, and c_3, detailed as Table 4

In this simulation, we assume that the critical time for making prediction is several minutes before the day t+1 opening of Twi, when depends on the model performance. Predicting the day t+1 opening return of Twi is the purpose of proposed model. Once day t+1 opening, the accuracy of predication would be calculated that means investors will balance its investments at opening time to win the profit from the gap price because of overnight.

Table 4. Encoding Rule to the Fluctuation of DJi and Twi

Upswing (%)		$c_1 \sim c_3$	Downswing (%)		$c_1 \sim c_3$
DJi	Twi		DJi	Twi	
(0, 0.26]	(0, 0.33]	000	(-0.23, 0]	(-0.19, 0]	100
(0.26, 0.57]	(0.33, 0.62]	001	(-0.53, -0.23]	(-0.5, -0.19]	101
(0.57,1.03]	(0.62, 0.97]	010	(-1.01, 0.53]	(-1.06, -0.5]	110
(1.03, ∞]	(0.97, ∞]	011	(-∞, -1.01]	(-∞, -1.06]	111

5.2 Experiments

5.2.1 Experiments

For each prediction model, all data in these simulations is separated into the training period and the testing period, shown in Figure 14. First of all, the data in the training period is applied to establish classifier rule populations. Second, the testing period data is applied to the opening price accuracy test.

Because of the trading day consistence of Twi and DJi, DJi daily data should be processed for synchronizing to predict Twi, which preprocesses are listed as follows:

1. In case of the day of Twi open but DJi non-open, we assume that DJi advance-decline ratio is zero, which means $DJiReturn_t^{co} = 0$.
2. In case of the day of Twi non-open but DJi open, we delete the DJi day (t to s-1) data and accumulate the day (t to s-1) Twi advance-decline ratio to next trading days, which denotes $TwiReturn_s^{co} = TwiReturn_t^{co} + \cdots + TwiReturn_{s-1}^{co}$.

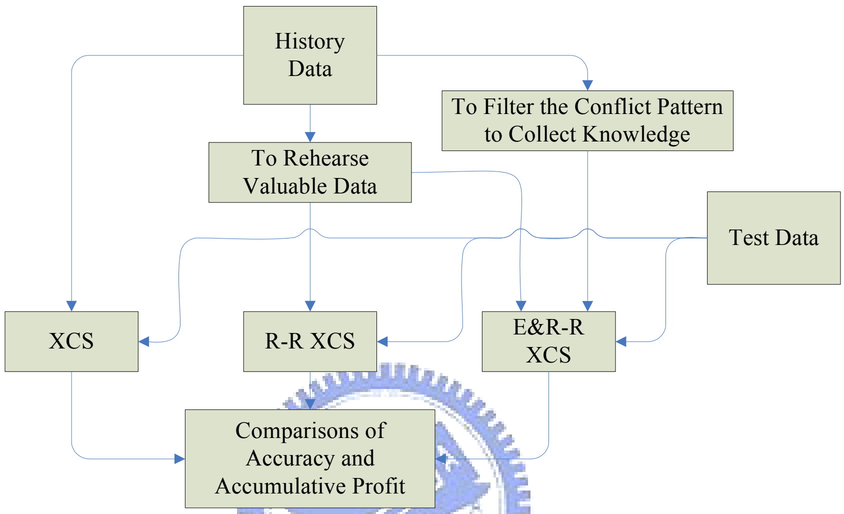

Figure 14. Flow of XCS, R-R XCS and E&R-R XCS Experiments.

- **Evaluation approaches:**

This work evaluates the three models, XCS, R-R XCS, and E&R-R XCS, by calculating their accuracy ratios and analyzing their accumulative profits. Before that, we first define three more investment strategies, listed as Table 5, to calculate each accumulative profit. For instance, the first strategy means when the output code of the model is {000,001,010,011}, the investment strategy is long future. The code set, {000,001,010,011} means the predicted advance-decline ratio is positive. In the contrary, when the output of the model is {100,101,110,111}, the investment strategy is short future. The code set, {100,101,110,111} means the predicted advance-decline ratio is negative. As for decoding the code set, the predicated fluctuation is translated and decoded by the Table 6.

Calculating accuracy ratio:

The accuracy ratio is basically utilized to evaluate the prediction performance of the model. The accuracy ratio is calculated by the judgment shown as Table 5. For instance, if the output code is {010}, which the predicted advance-decline ratio falls in the interval (0.62%, 0.97%), and the real advance-decline ratio is upswing, the predicted result is true. That is, the accuracy ratio is increased. The others are false and the accuracy ratio is decreased.

Table 5. Listing of Investment Strategies

Strategy	output codes{$c_1c_2c_3$}	predicated ratio (r %)	Investment strategy
1	{000,001,010,011}	r % is positive.	long future
	{100,101,110,111}	r % is not positive.	short future
2	{001,010,011}	r % > 0.33 %	long future
	{101,110,111}	r% < -0.19 %	short future
3	{010,011}	r% > 0.62 %	long future
	{110,111}	r% < -0.5 %	short future

Table 6. Table of Predicted Advance-Decline Ratio of Twi Return and its Accuracy Indicator

predicted upswing			predicted downswing		
$c_1 \sim c_3$	advance-decline ratio (%)	accuracy indicator	$c_1 \sim c_3$	advance-decline ratio (%)	accuracy indicator
000	(0, 0.32]	(-0.5, ∞]	100	(-0.19, 0]	(-∞, 0.61]
001	(0.32, 0.61]	(-0.19, ∞]	101	(-0.5, 0.19]	(-∞, 0.32]
010	(0.61, 0.97]	(0, ∞]	110	(-1.04, -0.5]	(-∞, 0]
011	(0.97, ∞]	(0, ∞]	111	(-∞, -1.04]	(-∞, 0]

Calculating accumulative profit:

The accumulative profit, calculated by making the investment, is to evaluate the leverage effect of the models. Once the model responses the output, the investment strategy is working to be made. As for the investment, because the problem is the model to simulate the global overnight effect to the market stock and to input those market indexes, the future should be chosen as the proper financial commodity to invest. The accumulative profit is calculated by summing up each investment profit as well.

- **Others**

The testing period is from 2004/01 to 2004/09 including 183 trading days, shown as Figure 15. The investment finance commodity is Taiwan weight index future. For each model, we finished the experiments of three strategies. For those analyses to the result of different investment strategies, we could verify the performance of the rehearsal function. Furthermore, for those comparisons of accuracy ratios, we could conclude that E&R-R XCS has better performance than the other two models. Next, the following subsections detail the experiments of Figure 14 indicated.

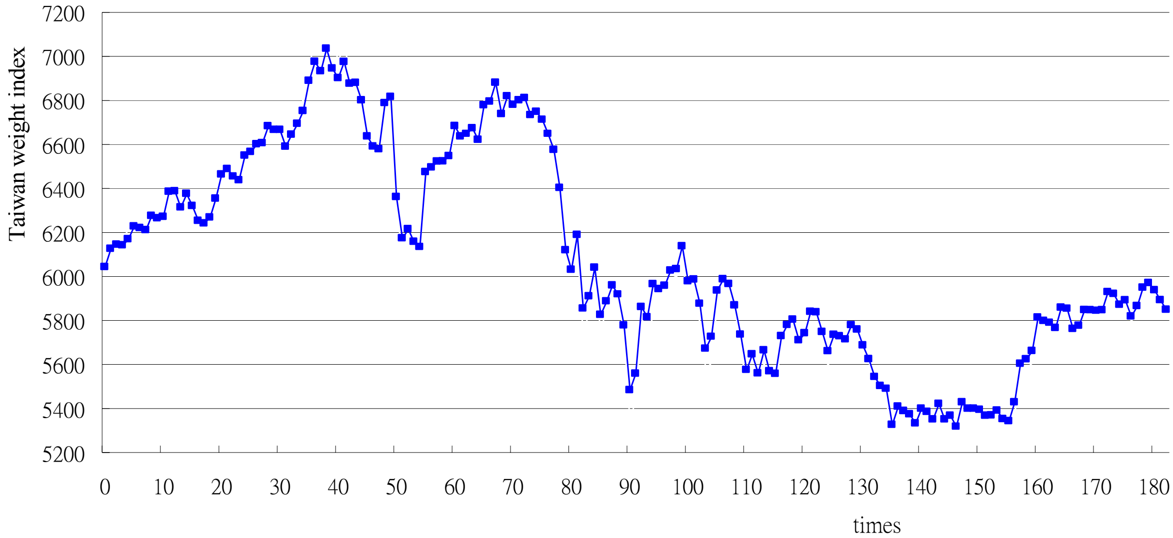

Figure 15. Testing Data of Taiwan Weight Index from 2004/01 to 2004/09

5.2.2 XCS Experiments

In XCS experiments, four rounds were finished respectively. Three kinds of investment strategies were predefined to evaluate the accuracy ratio and the accumulative profit, listed as Table 5. For the following graphs in 5.2.2, its x-axis means the trading day sequences, and its y-axis means the accuracy ratio or accumulative profits.

- **XCS accuracy ratio**

Figure 16, Figure 17, and Figure 18 show the accuracy ratio of three investment strategies respectively. In those figures, XCS model has a good predicted accuracy and a stable result. XCS model seems more proper to the first strategy. In Figure 16, the best result about 70% is in the first experiment and the worse is roughly 60 % in the forth experiment. The average accuracy ratio of first strategy is 67.2%, shown as Table 7.

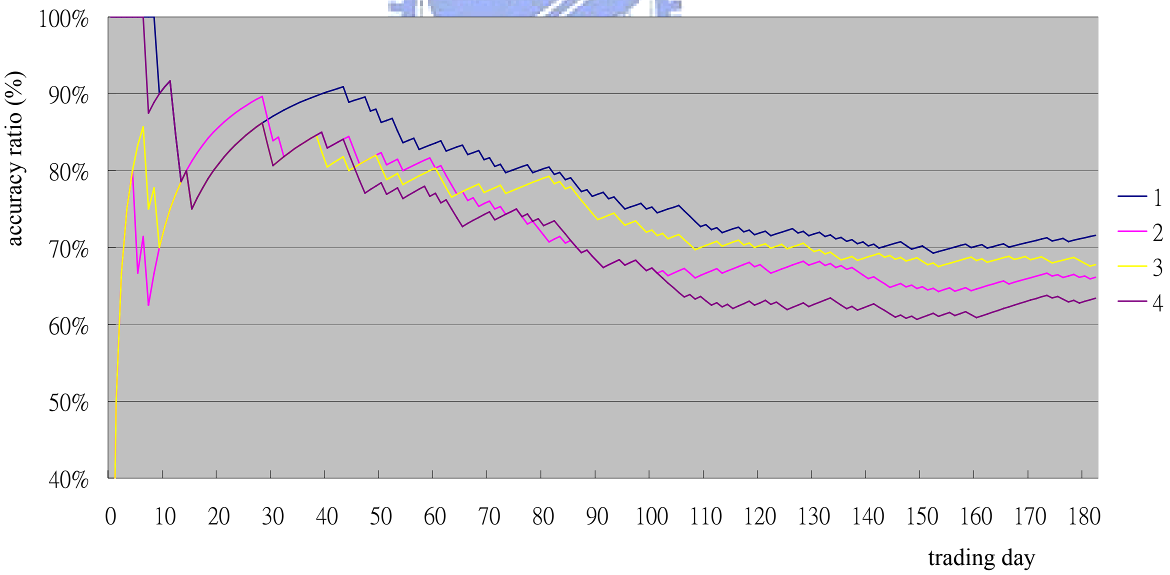

Figure 16. Strategy 1: Accuracy Ratio of XCS

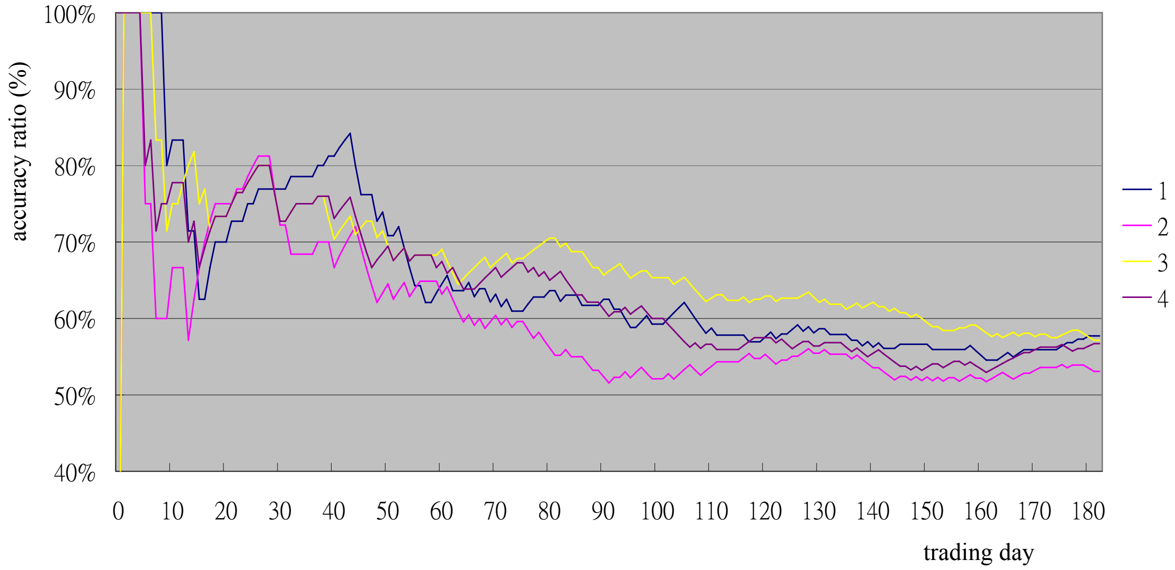

Figure 17. Strategy 2: Accuracy Ratio of XCS

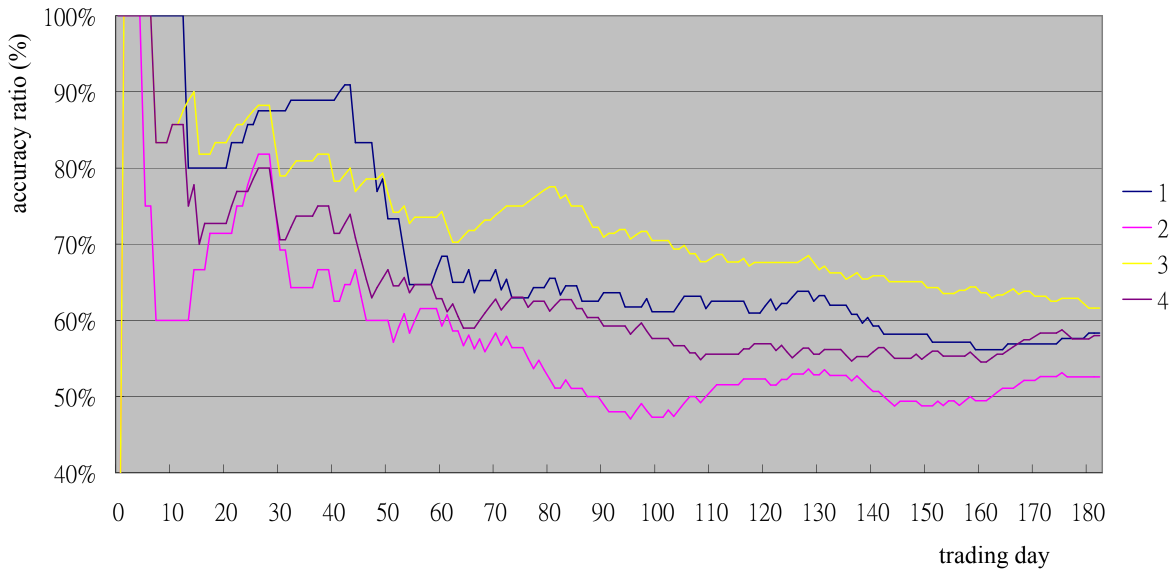

Figure 18. Strategy 3: Accuracy Ratio of XCS

- **XCS accumulative profits**

Figure 19, Figure 20, and Figure 21 present the accumulative profits according to XCS model. Besides, the third investment is not all positive. The reason just originates its accuracy ratio in figure 15 which shows the worse predicted result in the second experiment. The other two strategies are both positive. For 1st strategy, the highest accumulative profit is the 3rd experiment in Figure 19, but the highest accuracy ratio is not. Therefore, the accumulative profit does not depend on the accuracy ratio. In addition, a result deserves to be mentioned that the accumulative profits for all of the XCS experiments usually do not increase stably.

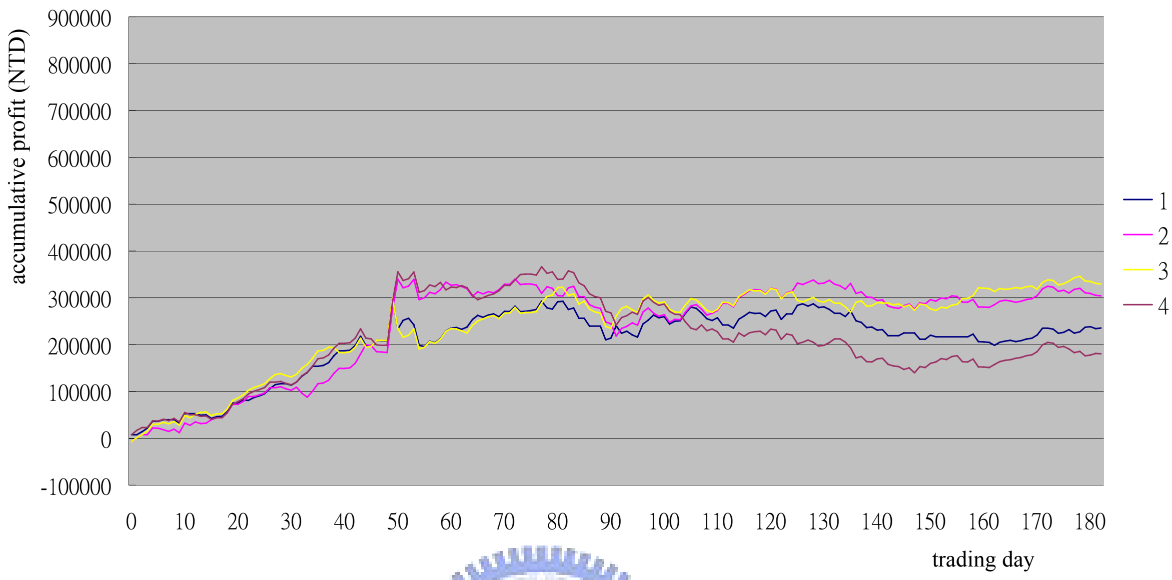

Figure 19. Strategy 1: Accumulative Profit of XCS

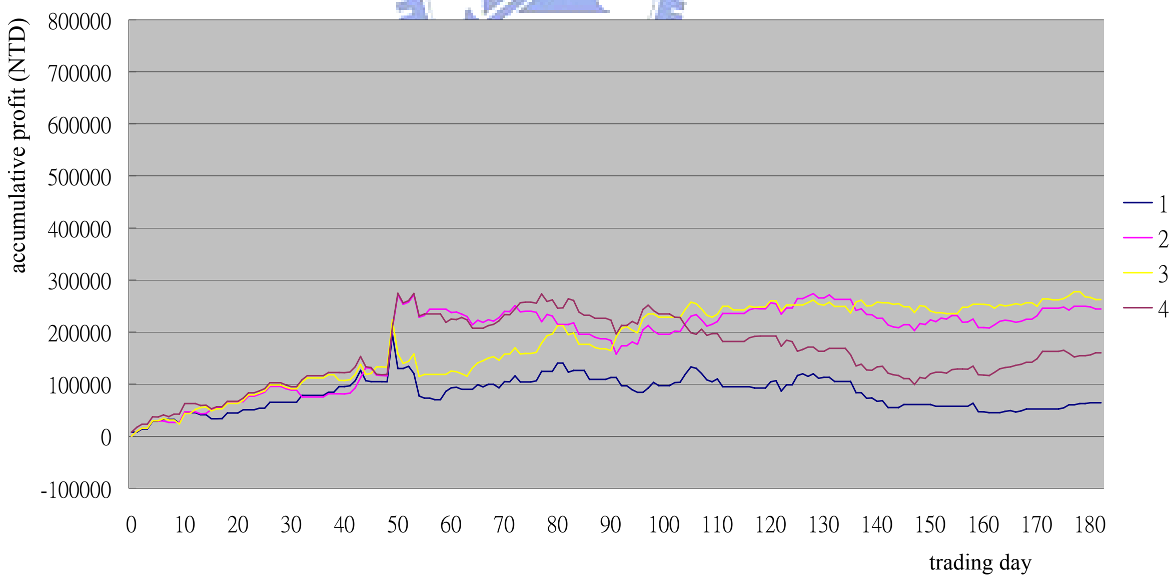

Figure 20. Strategy 2: Accumulative Profit of XCS

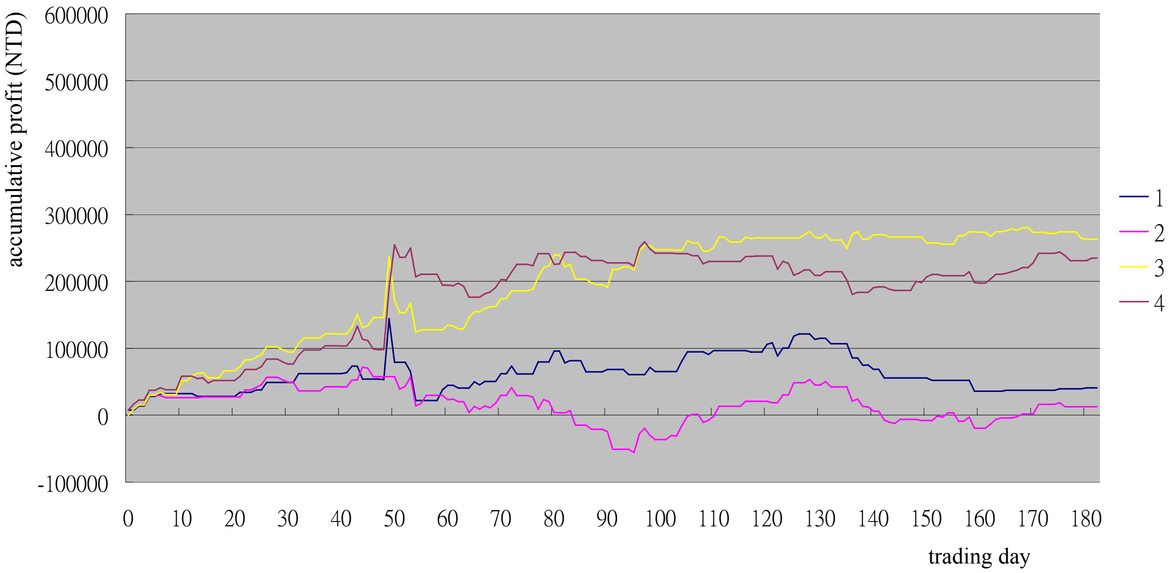

Figure 21. Strategy 3: Accumulative Profit of XCS

5.2.3 R-R XCS Experiments

In R-R XCS experiments, four rounds were finished respectively. Three kinds of investment strategies were predefined to evaluate the accuracy ratio and the accumulative profit, listed as Table 5. For the following graphs in 5.2.3, its x-axis means the trading day sequences. The difference between XCS and R-R XCS is rehearsal part. In advance, the valued rehearsal data is defined as the higher vibration of the stock fluctuation, which are the code set {010,011} and {110, 111}. Besides, if any conflict experience happens, the mutation mechanism starts or the system holds and waits the system operator to verify the condition.

- **R-R XCS accuracy ratio**

Figure 22, Figure 23, and Figure 24 show the accuracy ratio of three investment strategies respectively. In those figures, R-R XCS model has a better predicted accuracy and a stable result. In Figure 22, the best result about 72% is in the first experiment and the worse is roughly 65 % in the forth experiment. The average accuracy ratio of first strategy is 70.6%.

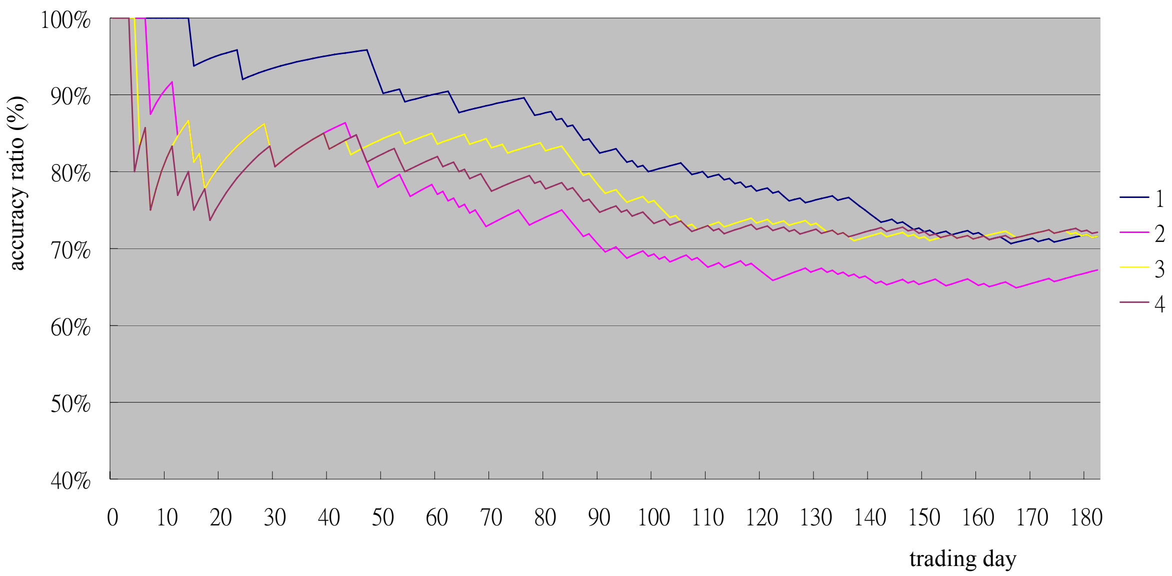

Figure 22. Strategy 1: Accuracy Ratio of R-R XCS

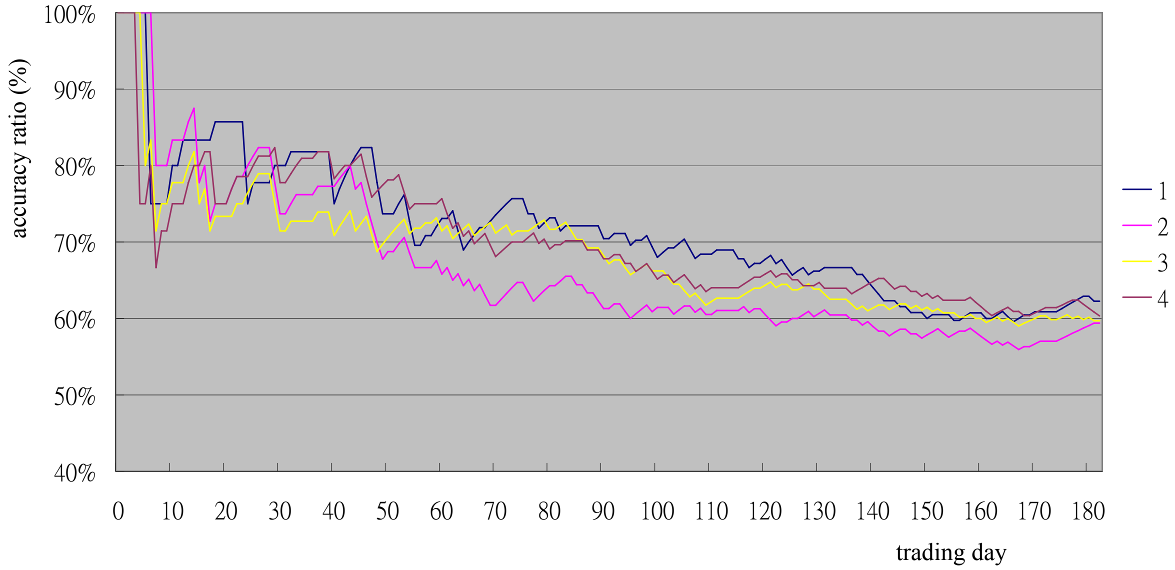

Figure 23. Strategy 2: Accuracy Ratio of R-R XCS

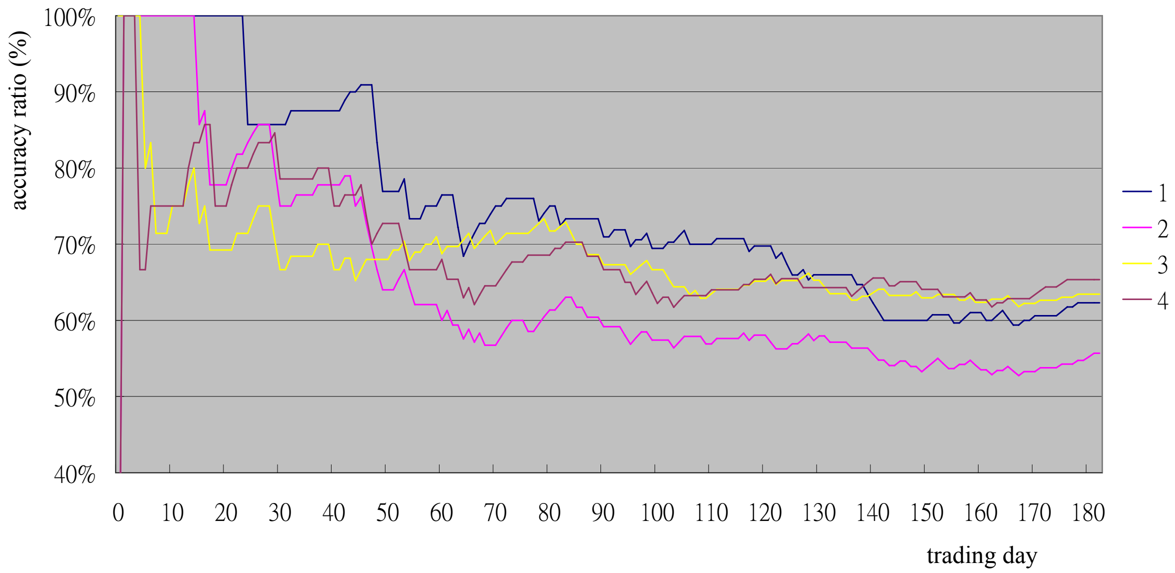

Figure 24. Strategy 3: Accuracy Ratio of R-R XCS

- **R-R XCS accumulative profits:**

Figure 25, Figure 26, and Figure 27 show the accumulative profits according to R-R XCS model. Besides, the third investment is not all positive. The other two strategies are both positive. In the first strategy, the highest accumulative profit is the 3rd experiment in Figure 19, but the highest accuracy ratio is not. Therefore, the accumulative profit does not depend on the accuracy ratio. In addition, a result deserves to be mentioned that the accumulative profits for the R-R XCS all experiments increase with a slow gradual degree.

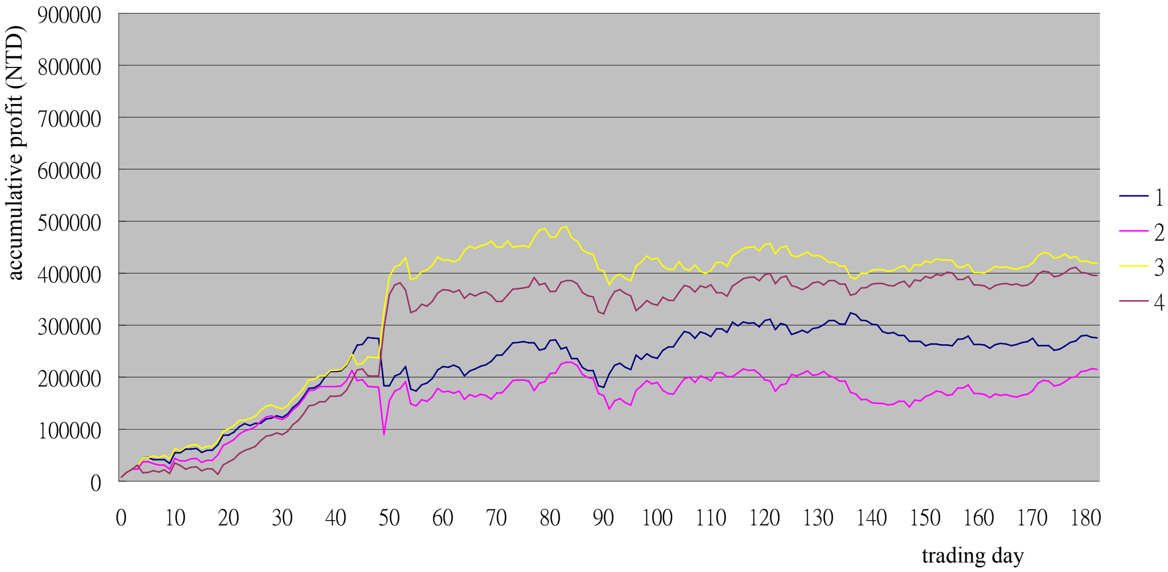

Figure 25. Strategy 1: Accumulative Profit of R-R XCS

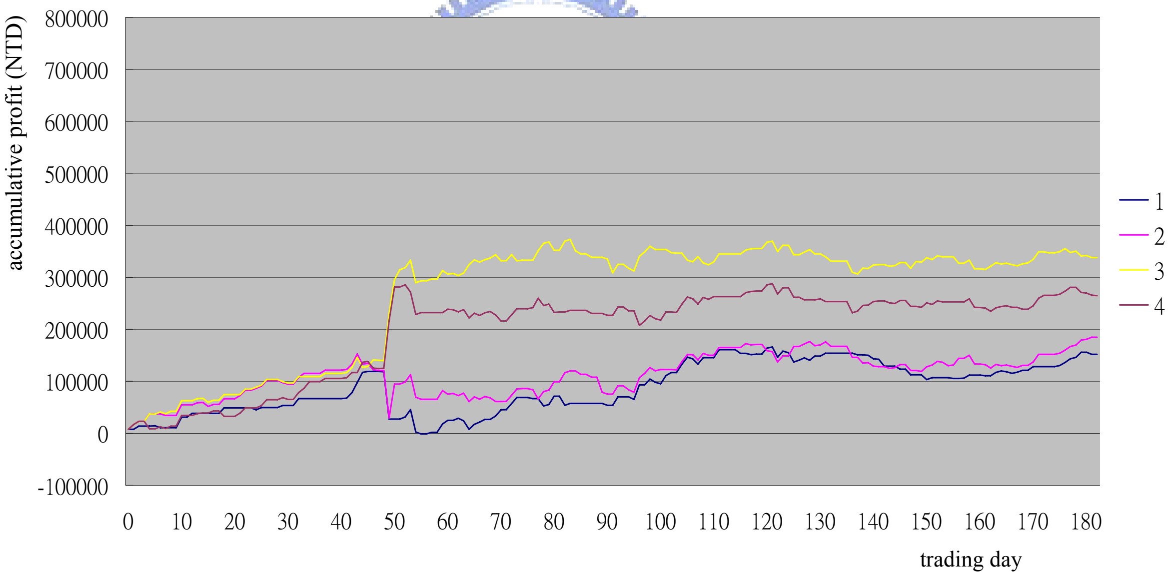

Figure 26. Strategy 2: Accumulative Profit of R-R XCS

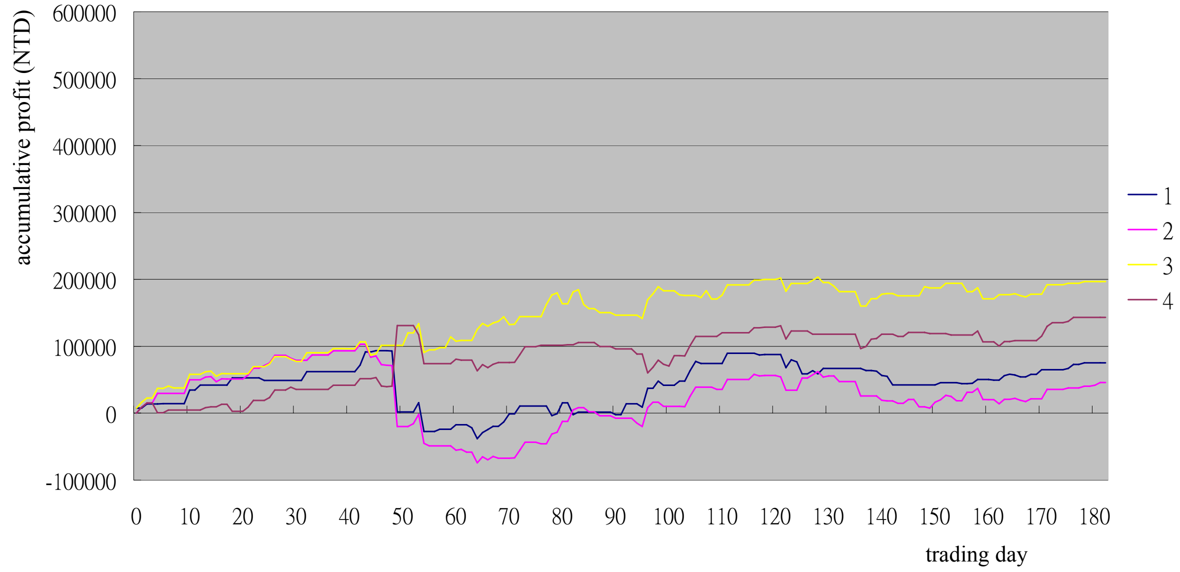

Figure 27. Strategy 3: Accumulative Profit of R-R XCS

5.2.4 E&R-R XCS Experiments

In E&R-R XCS experiments, four rounds were finished respectively. Three kinds of investment strategies were predefined to evaluate the accuracy ratio and the accumulative profit, listed as Table 5. For the following graphs in 5.2.4, its x-axis means the trading day sequences. In the simulation, R-R part in E&R-R XCS works the same as R-R XCS. Besides, for knowledge experiments and the description in section 4.4, we filter conflict patterns from the experience rule base of R-R XCS. Then the real and verified knowledge population with 29 classifiers is formed, listed in Appendix B. This knowledge population is utilized to build the knowledge rule base in advance. Actually, if any knowledge classifier increases, it would be possible to start the memorizing mechanism of E&R-R model and memorize it into knowledge rule base.

- **E&R-R XCS accuracy ratio**

Figure 28, Figure 29, and Figure 30 show the accuracy ratio of three investment strategies respectively. In those figures, E&R-R XCS model has a good predicted accuracy and a stable result. In Figure 28, the best result about 81% is in the first experiment and the worse 75 % is in the forth experiment. The average accuracy ratio of 1^{st} strategy is 78.8%.

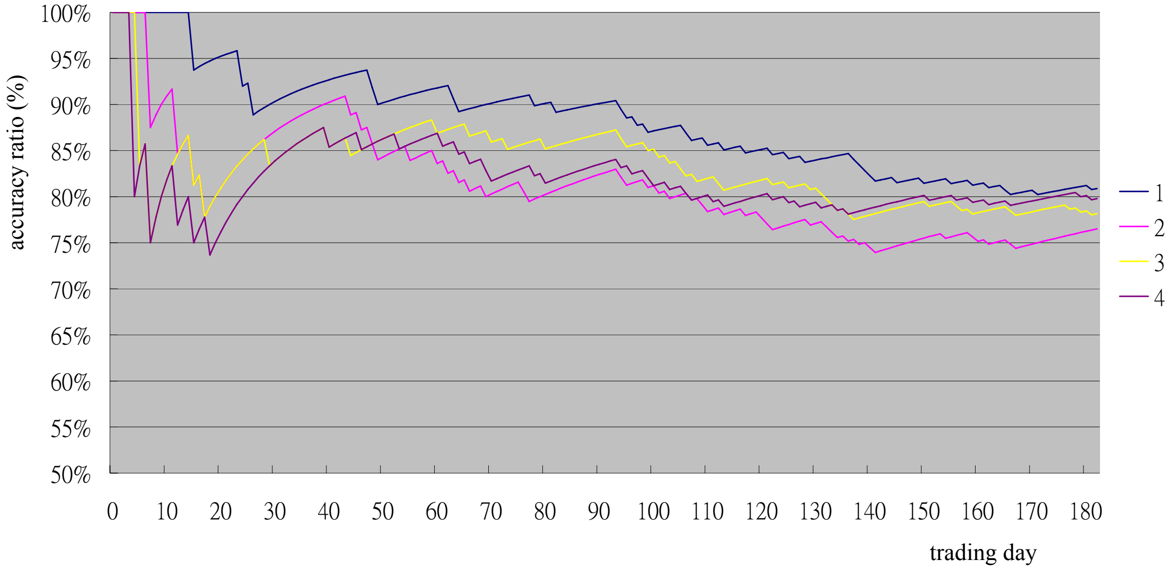

Figure 28. Strategy 1: Accuracy Ratio of E&R-R XCS

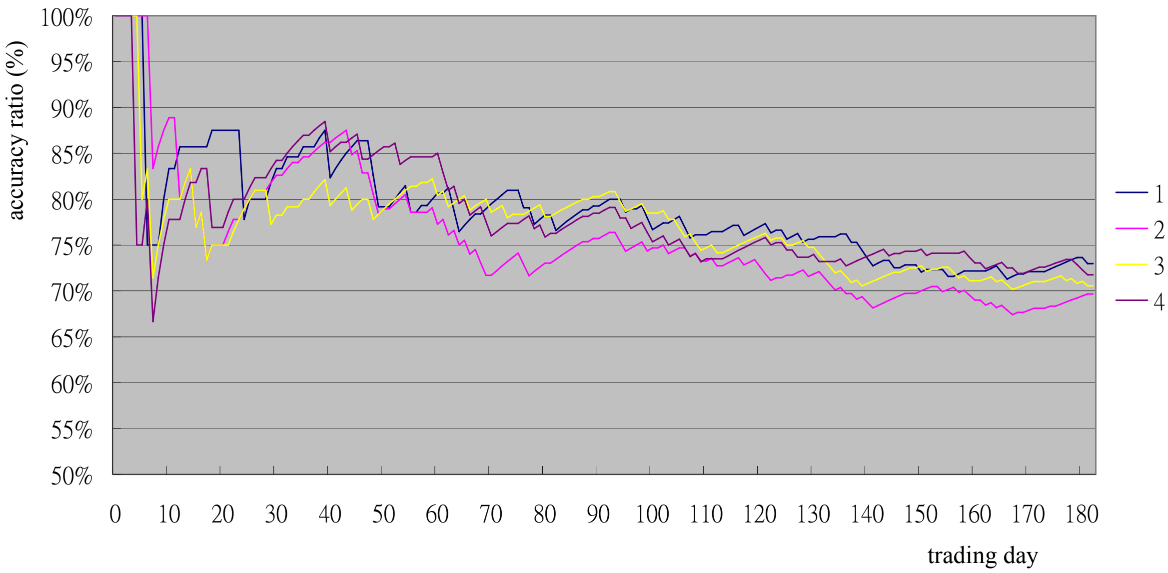

Figure 29. Strategy 2: Accuracy Ratio of E&R-R XCS

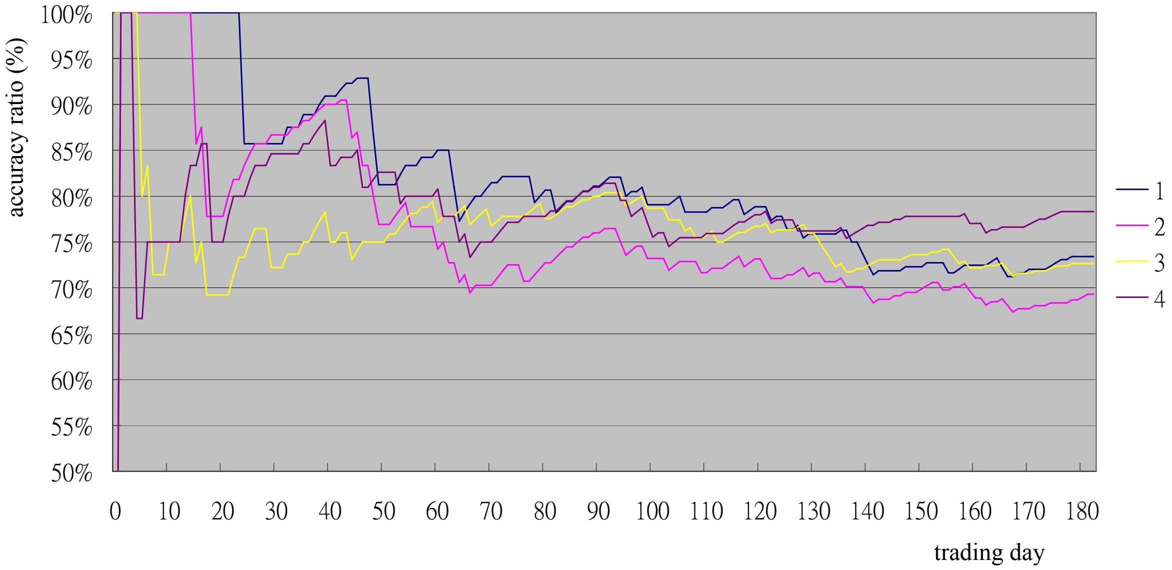

Figure 30. Strategy 3: Accuracy Ratio of E&R-R XCS

- **E&R-R XCS accumulative profits:**

Figure 31, Figure 32, Figure 33, and show the accumulative profits according to E&R-R XCS model. The three accumulative profits of the strategies are all positive. In the first strategy, the highest accumulative profit is the 3rd experiment in figure 28. In addition, a result deserves to be mentioned that the accumulative profits for all of the E&R-R XCS experiments increase stably.

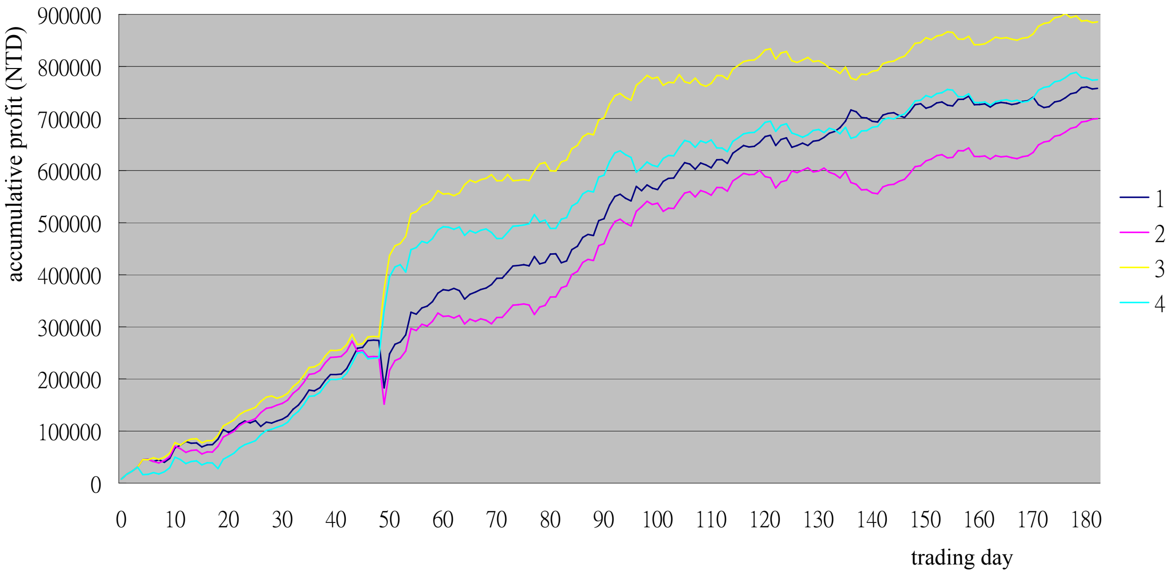

Figure 31. Strategy 1: Accumulative Profit of E&R-R XCS

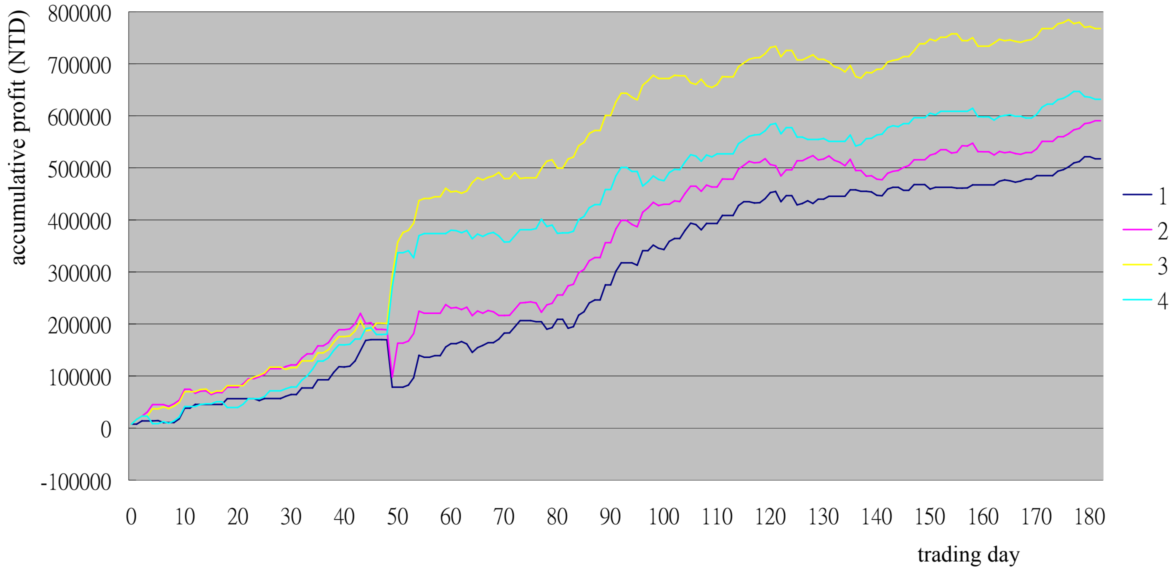

Figure 32. Strategy 2: Accumulative Profit of E&R-R XCS

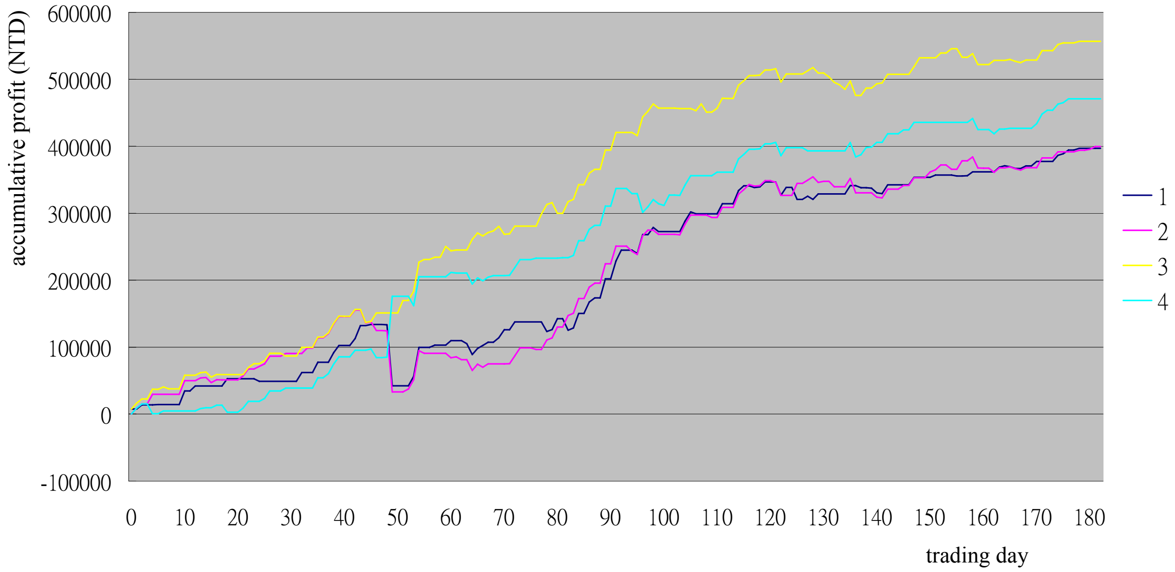

Figure 33. Strategy 3: Accumulative Profit of E&R-R XCS

5.3 Comparison and Discussions

Table 7, Table 8, and Table 9 respectively show the summary results of XCS, R-R XCS, and E&R-R XCS. In those tables, there are still three investment strategies according to each model.

Table 7. Summary of XCS Experiments

Strategy	Investing times	Correct times	Correct ratio	Average accumulative profit (NTD)	Accumulative profit (NTD)	Accumulative profit ratio	Annual profit ratio
1	183.00	123.00	67.2%	1,583.06	289,700.00	241.4%	321.9%
2	123.50	69.25	56.1%	1,478.95	182,650.00	152.2%	202.9%
3	89.00	51.25	57.6%	1,547.75	137,750.00	114.8%	153.1%

Table 8. Summary of R-R XCS Experiments

Strategy	Investing times	Correct times	Correct ratio	Average accumulative profit (NTD)	Accumulative profit (NTD)	Accumulative profit ratio	Annual profit ratio
1	183.00	129.25	70.6%	1,871.04	342,400.00	285.3%	380.4%
2	120.25	72.50	60.3%	1,951.77	234,700.00	195.6%	260.8%
3	83.50	51.25	61.4%	1,380.84	115,300.00	96.1%	128.1%

Table 9. Summary of E&R-R XCS Experiments

Strategy	Investing times	Correct times	Correct ratio	Average accumulative profit (NTD)	Accumulative profit (NTD)	Accumulative profit ratio	Annual profit ratio
1	183.00	144.25	78.8%	4,259.02	779,400.00	649.5%	866.0%
2	133.25	94.75	71.1%	4,704.69	626,900.00	522.4%	696.6%
3	92.25	67.50	73.2%	4,943.63	456,050.00	380.0%	506.7%

5.3.1 Models Self-Comparison

(1) In XCS model,

all the three accuracy ratios are all above 55%, and the average accumulative profit are roughly equal to NTD 1,500, shown as Table 7. Besides, the strategy 1 has the highest accumulative profit 241.1% because of its more investing times 183.

(2) In R-R XCS model,

all the three accuracy ratios are all above 60%, and the average accumulative profit are at least NTD 1,300, shown as Table 8. Besides, the strategy 3 has the lowest accumulative profit because the 2nd experiment, the worst one, causes the result to become worse. However, this result is just the leverage effect of R-R XCS, discussed in advance.

(3) In E&R-R XCS model,

all the three accuracy ratios are all above 71%, and the average accumulative profit are roughly above NTD 4,200, shown as Table 9. Besides, the strategy 1 has the highest accumulative profit 649.5% because of its more investing times, but strategy 3 has the highest average accumulative profit NTD 4,943.63 because its investment strategy just fitted the E&R-R model focused on valuable information.

Besides, it is deserved to be mentioned. While three trading intervals, 48-50, 77-79, and 88-90, some politic events happened in Taiwan. These unreasonable events caused the stock market crashed. In simulations, the prediction model accuracies all become worst. That is, we can easily find that the average accuracy before trading day 48 and after 150 should be better and increased continuously. In fact, according to these kinds of un-system risks, we can identify them as knowledge or not. If it is verified to knowledge, after that, the system performance would be increased. In the contrast, if it is not knowledge, we can filter it and avoid it to affect the model in advanced. The same is that the system performance would be increased as well.

5.3.2 Models Comparison

(1) Accuracy ratio:

Table 7, Table 8, and Table 9 show that the accuracy result of E&R-R XCS model is the best one, and the accuracy result of R-R XCS model is better than the result of XCS model. Actually, we induce the result as the discussions of 4.5 that the E&R-R XCS model is indeed the best model of all. Meanwhile, R-R XCS may be not a better t model than XCS one due to the leverage effect.

(2) Accumulative Profit:

Table 7, Table 8, and Table 9 show that the accumulative profit of E&R-R XCS model is really the best one, and the average accumulative profit of R-R XCS model is not absolutely better than the result of XCS model. The R-R XCS wins empirically ether the more profits or the worse profits to the XCS. Actually, we induce the result as the discussions of 4.5 that the E&R-R XCS model is indeed the best model of all. Meanwhile, R-R XCS may not be the better model than XCS one due to the leverage effect.

Chapter 6. Conclusions

6.1 Result and Summary

In this work, we propose a dual perspective learning model by adding cognitive learning mechanism through the review of Cognitive Psychology. Because the traditional AI researchers involuntarily restrict the learning type to the trial and error style only, the phenomenon causes the Soft Computing approaches developed to only flexible but not certainly correct-knowledge learning models. We successfully involve the education learning concept that permits the only knowledge as teaching materials to develop the model. From either the induction to construct the dual learning mechanism, E&R-R learning model, or the simulation of E&R-R XCS model to Taiwan-weight-index prediction on global overnight effect, this work has successfully enhanced the original AI learning procedure and finally implemented the E&R-R XCS model.

Among them, the prior proposed R-R XCS model has been verified that it could increase the leverage effect to XCS model first. Moreover, E&R-R XCS model implanted education concept that is the teaching materials from correcting conflict pattern in advance, and then the remarkable outcome obtained.

Finally, this work has also developed a global-overnight-effect knowledge analysis model based on E&R-R XCS, R-R XCS, and XCS model respectively. In the simulations, the investment strategy referred E&R-R XCS model wins the wonderful accumulative profit. For this, not only the proposed dual learning model based on knowledge-education and machine learning has swimmingly been verified, but also a novel prediction mechanism to financial investments has successfully been proposed.

6.2 Future Works

Future work will be addressed three issues. First, this work still roughly derives cognitive learning from Cognition Psychology. As for the entire theory of cognition, lots of faultless Psychology models even Cognition Psychology models have been flooded. The better efficiency of learning mechanism by computing simulation has the possibility to been come true. Second, although this work is the first one to the aspect, we still expect more and more AI researchers would enhance their model considering this kind of philosophy thinking. Besides, from the pass to the future, the other following models with better accuracy, and the performance might be the substitution for XCS. Third, the model considering more complex factors to finance prediction issue would be declared. Actually, the ponderable model is important to apply the right factors to the right issue to obtain the remarkable outcome.

Reference

1. J. McCarthy. “Generality in Artificial Intelligence”, *Communications of the ACM*, 30(12), pp.1030-1035, Dec. 1987.
2. S. Ghirlanda, M. Enquist. “Artificial Neural Networks as Models of Stimulus Control”, *Animal Behaviour*, 56, pp.1383–1389, 1998.
3. S. Ghirlanda, M. Enquist. “The Geometry of Stimulus Control”, *Animal Behaviour*, 58(4), pp. 695-706, 1999.
4. V. Chiew. “A Software Engineering Cognitive Knowledge Discovery Framework”, *Proceedings of the First IEEE International Conference on Cognitive Informatic*s (ICCI’02), pp.163-172, Calgary, AB, Canada, Aug. 2002.
5. S. W. Wilson. “Classifier Fitness Based on Accuracy”, *Evolutionary Computation*, 3(2), pp.149-175, 1995.
6. J. Piaget. *Structuralism* (C. Maschler, Trans., original French edition published 1968). New York: Harper & Row, 1970.
7. J. H. Holland, J. S. Reitman. “Cognitive Systems Based on Adaptive Algorithms”, *Pattern Directed Inference Systems*, 7(2), pp.125-149, 1978.
8. J. H. Holland. “Processing and Processors for Schemata”, *Associative Information Processing*, New York, pp. American Elsevier, 127-146, 1971.
9. S. W. Wilson. “Classifier Systems and the Animat Problem”, *Machine Learning*, 2(3), pp.199-228, 1987.
10. L. B. Booker. “Triggered Rule Discovery in Classifier Systems”, *Proceedings of the 3rd International Conference on Genetic Algorithms* (ICGA89), Morgan Kaufmann Publishers Inc, pp. 265-274, June 1989.
11. S. W. Wilson. “ZCS: A Zeroth Level Classifier System”, *Evolutionary Computation*, 2(1), pp.1-18, 1994.
12. P. W. Frey, D. J. Slate. “Letter Recognition Using Holland-Style Adaptive Classifiers”, *Machine Learning*, 6, pp.161-182, 1991.
13. S. W. Wilson. “Generalization in the XCS Classifier System”, *Proceedings of the Third Genetic Programming Annual Conference*. Morgan Kaufmann: San Francisco, CA, 665-674, 1998.
14. W. Stolzmann. “An Introduction to Anticipatory Classifier Systems”, *Learning Classifier Systems: From Foundations to Applications, Lecture Notes in Artificial Intelligence*, 1813, Springer, pp.175–194, 2000.
15. P. L. Lanzi, et al. *Learning Classifier Systems: From Foundations to Applications, Lecture Notes in Artificial Intelligence*, 1813, Springer, Berlin, 2000.
16. J. H. Holmes. *Evolution-assisted Discovery of Sentinel Features in Epidemiologic Surveillance*.

Ph.D. thesis, Drexel University, Philadelphia, PA, 1996.
17. J. H. Holmes. “Learning Classifier Systems Applied to Knowledge Discovery in Clinical Research Databases”, *Learning Classifier Systems: From Foundations to Applications, Lecture Notes in Artificial Intelligence*, 1813, pp.243-264, 2000.
18. J. H. Holmes, P. L. Lanzi. “Wolfgang Stolzmann, and Stewart W. Wilson. Learning Classifier Systems: New models, Successful Applications”, *Information Processing Letters*, 82, pp.23-30, 2002.
19. S. Kemp. *Cognitive Psychology in the Middle Ages*, Westport, Connecticut and London: Greenwood Press, 1996.
20. I. Kant. “The Critique of Pure Reason”, *Great books of Western World*, 42. R.M.Hutchins (ed.), pp. ix-209, Chicago: Encyclopaedia Britannica, 1952.
21. K. S. Lashley. *Brain Mechanism and Intelligence: A Quantitative Study of Injuries to the Brain*, Chicago: Chicago University Press, 1929.
22. J. M. Hunt. “Psychological Development: Early Experience”, *Annual Review of Psychology*, 30, pp. 103-143, 1979.
23. D. O. Hebb. “A Neuropsychological Theory”, *Psychology: A Study of a Science*, 1, New York: McGraw-Hill, 1959.
24. M. W. Watson, K. W. Fische. “Development of Social Roles in Elicited and Spontaneous Behavior during the Preschool Years”, *Developmental Psychology*, 16, pp. 483-494, 1980.
25. B. F. Skinner. *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton-Century, Published in 1938 originally. Reprinted by the B. F. Skinner Foundation, 1991 and 1999.
26. J. Piaget. *The Equilibration of Cognitive Structures: The Central Problem of Intellectual Development*. Chicago: University of Chicago Press, 1985. Original work published 1975.
27. G. Luger, G. F. Luger. *Artificial Intelligence, Structures and Strategies for Complex Problem Solving*, 4th Edition, Harlow, England: Addison-Wesley, pp. 471, 2002.
28. J. H. Holland. “Escaping Brittleness: The Possibilities of General-Purpose Learning Algorithms Applied to Parallel Rule-Based Systems”, *Machine Learning, An Artificial Intelligence Approach,* 2(ch. 20), pp. 593-623, Morgan Kaufmann, 1986.
29. J. H. Holland, J. S. Reitman. “Cognitive Systems Based on Adaptive Algorithms”, *Pattern Directed Inference Systems*, New York: Academic Press, pp. 313-329, 1978.
30. J. McCarthy. “Generality in Artificial Intelligence”, *Communications of the ACM*, 30(12), pp.1030-1035, Dec.1987.
31. J. McCarthy, P. J. Hayes. “Some Philosophical Problems from the Standpoint of Artificial Intelligence”, *Machine Intelligence*, 4, pp. 463-502. Edinburgh University Press, 1969.
32. J. H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, 2nd Edition, MIT Press, 1992.
33. E. L. Thorndike. “Animal Intelligence: An Experimental Study of the Associative Processes in Animals”, *Psychological Review*, Monograph Supplement, 8, New York: Macmillan, 1898.

34. E. L. Thorndike. *Elements of Psychology*. New York: A. G. Seiler, 1905.
35. J. B. Watson. *Behavior: An Introduction to Comparative Psychology*, New York: Henry Holt & Co, 1914.
36. M. V. Butz, S. W. Wilson. "An algorithmic description of XCS", *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, Springer-Verlag GmbH, 6(3-4), pp.144-153, June 2002.
37. J. Piaget. *The Mechanisms of Perception*, New York: Basic Books, 1969.
38. R. H. Bruning, et al. *Cognitive Psychology and Instruction. Englewood Cliffs*, NJ: Prentice Hall, 1995.
39. D. E. Rumelhart. "Schemata: The Building Blocks of Cognition", In R. Spiro, B. Bruce, W. Brewer, *Theoretical Issues in Reading Comprehensio*n, pp. 33-58, Hillsdale, NJ: Lawrence Erlbaum, 1980.
40. M. Anthony, et al. "On Exact Specification by Examples", *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 311-318, ACM Press, July 1992.
41. R. Freivalds, et al. "On the Power of Inductive Inference from Good Examples", *Theorertical Computer Science*, 110(1), pp. 131-144, 1993.
42. S. A. Goldman, M. J. Kearns. "On the Complexity of Teaching", *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pp.303-314. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, August 1991.
43. J. Jackson, A. Tomkins. "A Computational Model of Teaching", *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, pp. 319-326. ACM Press, July 1992.
44. A. Shinohara, S. Miyano. "Teachability in Computational Learning", *New Generation Computing*, 8(4), pp. 337-347, 1991.
45. J. M. Belmont, E. C. Butterfield. "Learning Strategies as Determinants of Memory Deficiencies", *Cognitive Psychology*, 2, pp. 411-420, 1971.
46. T. Buzan, *Use Your Head*, (Millennium Ed), London, BBC, 2000.
47. R. M. Gagne. *The Conditions of Learning and Theory of Instruction*, (4th Ed.), New York, NY: Holt, Rinehart and Winston, 1985.
48. R. C. Atkinson, R. M. Shiffrin. "Human Memory: A Proposed System and Its Control Processes", *The Psychology of Learning and Motivation: Advances in Research and Theory*, 2, pp. 89-195, New York: Academic Press, 1968.
49. N. Waugh, D. A. Norman. "Primary Memory", *Psychological Review*, 72, pp. 89-104, 1965.
50. R. M. Gagne, K. L. Medsker. *The Conditions of Learning: Training Applications*, 1996.
51. G. A. Miller. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information", *Psychological Review*, 63, pp.81-97, 1956.
52. H. A. Simon. "The Structure of Ill Structured Problems", *Artificial Intelligence*, 4, pp. 181-201, 1973.
53. A. Newell, H. A. Simon. *Human Problem Solving*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1972.

54. J. R. Anderson. *The Architecture of Cognition*, Cambridge, MA: Harvard University Press, 1983.
55. A. Newell. *Unified Theories of Cognition*, Cambridge, MA: Harvard University Press, 1990.
56. S. K. Reed. *Cognition Theory and Applications*, 2nd Edition, Brooks/Cole Publishing Company, Pacific Grove, California, 1988.
57. R. L. Solso. *Cognitive Psychology*. 5th Edition, pp. 534, 1998.
58. G. Luger. *Cognitive Science: The Science of Intelligent Systems*. San Diego: Academic Press, 1994.
59. K. Chan, et al. "Intraday Volatility in the Stock Index and Stock Index Futures Markets", *Review of Financial Studies*, 4(4), pp. 657-684, 1991.
60. K. Chan, et al. "Overnight Information and Intraday Trading Behavior: Evidence from NYSE Cross-listed Stocks and Their Local Market Information", *Journal of Multinational Financial Management*, 10(3-4), pp. 495-509, 2000.
61. S. H. Chen. "Lecture 7: Rescale Range Analysis and the Hurst Exponent", *Financial Economics (I)*, Department of Economics, National Chengehi University, Taiwan, 2000.
62. M. J. Hinich, D. M. Patterson. "Evidence of nonlinearity in Daily Stock Returns", *Journal of Business & Economic Statistics*, 3(1), pp. 69-77, 1985.
63. A. P. Chen, et al. "Applying Two-Stage XCS Model on Global Overnight Effect for Local Stock Prediction" *Proceeding of (KES'2005) the 9th International Conference on Knowledge-Based & Intelligent Information & Engineering Systems*, Melbourne, Australia, 2005.

Appendix A. Relevant XCS Statements

In this appendix, all the following statements are reference from Wilson's XCS. The detailed descriptions about XCS should be looked it up in [36].

- A Classifier in XCS

XCS keeps a population of classifiers which represent its knowledge about the problem. Each classifier is a condition-action-prediction rule having the following parts:

- The condition $C \in \{0, 1, \#\}^L$ specifies the input states (sensory situations) in which the classifier can be applied (matches).
- The action $A \in \{a_1, ..., a_n,\}$ specifies the action (possibly a classification) that the classifier proposes.
- The prediction p estimates (keeps an average of) the payoff expected if the classifier matches and its action is taken by the system.

Moreover, each classifier keeps certain additional parameters:

- The prediction error ε estimates the errors made in the predictions.
- The fitness f denotes the classifier's fitness.
- The experience *exp* counts the number of times since its creation that the classifier has belonged to an action set.
- The time stamp *ts* denotes the time-step of the last occurrence of a GA in an action set to which this classifier belonged.
- The action set size *as* estimates the average size of the action sets this classifier has belonged to.
- The numerosity *num* reflects the number of micro-classifiers (ordinary classifiers) this classifier which is technically called a macroclassifier represents.

- The Different Sets

There are four different sets that need to be considered in XCS.

- The population $[P]$ consists of all classifiers that exist in XCS at any time t.
- The match set $[M]$ is formed out of the current $[P]$. It includes all classifiers that match the current situation $\sigma(t)$.
- The action set $[A]$ is formed out of the current $[M]$. It includes all classifiers of $[M]$ that propose the executed action.
- The previous action set $[A]_{-1}$ is the action set that was active in the last execution cycle.

- Learning Parameters in XCS

In order to control the learning process in XCS the following parameters are used:

- N specifies the maximum size of the population (in micro-classifiers, i.e., N is the sum of the classifier numerosities).
- β is the learning rate for p, ε, f, and *as*.
- α, ε_0, and υ are used in calculating the fitness of a classifier.
- γ is the discount factor used in multi-step problems in updating classifier predictions.
- θ_{GA} is the GA threshold. The GA is applied in a set when the average time since the last GA in the set is greater than θ_{GA}.
- χ is the probability of applying crossover in the GA.
- μ specifies the probability of mutating an allele in the offspring.
- θ_{del} is the deletion threshold. If the experience of a classifier is greater than θ_{del}, its fitness may be considered in its probability of deletion.
- δ specifies the fraction of the mean fitness in [*P*] below which the fitness of a classifier may be considered in its probability of deletion.
- θ_{sub} is the subsumption threshold. The experience of a classifier must begreater than 0,0 in order to be able to subsume another classifier.
- $P_{\#}$ is the probability of using a # in one attribute in C when covering.
- p_I, ε_j, and f_I are used as initial values in new classifiers.
- p_{explr}, specifies the probability during action selection of choosing the action uniform randomly.
- θ_{mna} specifies the minimal number of actions that must be present in a match set [*M*], or else covering will occur.
- *doGASubsumption* is a Boolean parameter that specifies if offspring are to be tested for possible logical subsumption by parents.
- *doActionSetSubsumption* is a Boolean parameter that specifies if action sets are to be tested for subsuming classifiers.

- An Algorithmic Description of XCS

This section presents the algorithms used in XCS. When XCS is started, the modules must first of all be initialized. The parameters in the environment must be set. After the initialization, the main loop is called. *RUN EXPERIMENT* is the main loop. Besides, *GENERATE MATCH SET*, *DOES MATCH*, *GENERATE COVERING CLASSIFIER*, *GENERATE PREDICTION ARRAY*, *SELECT ACTION*, *GENERATE ACTION SET*, *UPDATE SET*, *UPDATE FITNESS* are the detailed sub-functions, shown as following.

RUN EXPERIMENT ( ):

$\rho_{-1} \leftarrow 0$

do {

    $\sigma \leftarrow$ *env*: get situation

```
    GENERATE MATCH SET [M] out of [P] using σ
    GENERATE PREDICTION ARRAY PA out of [M]
    act ← SELECT ACTION according to PA
    GENERATE ACTION SET [A] out of [M] according to act
    env: execute action act
    ρ ← rp: get reward
    if ([A]-1 is not empty)
        P ← ρ-1 + γ * max (PA)
        UPDATE SET [A]-1 using P possibly deleting in [P]
        RUN GA in [A]-1 considering σ-1 inserting and possibly deleting in [P]
    if (rp: eop)
        P ← ρ
        UPDATE SET [A] using P possibly deleting in [P]
        RUN GA in [A] considering v inserting and possibly deleting in [P]
        empty [A]-1
    else
        [A]-1 ← [A]
        ρ-1 ← ρ
        σ-1 ← σ
} while (termination criteria are not met)
```

GENERATE MATCH SET ([*P*], σ):

```
initialize empty set [M]
while ([M] is empty)
    for each classifier cl in [P]
        if (DOES MATCH classifier cl in situation σ)
            add classifier cl to set [M]
    if (the number of different actions in [M] < θ_mna)
        GENERATE COVERING CLASSIFIER cl_c, considering [M] and σ
        add classifier cl_c to set [P]
        DELETE FROM POPULATION [P]
        empty [M]
return [M]
```

DOES MATCH (*cl*, σ):

```
for each attribute x in C_cl
    if(x <> # and x <> the corresponding attribute in σ)
    return false
return true
```

GENERATE COVERING CLASSIFIER ([M], σ):

initialize classifier cl
initialize condition C_{cl} with the length of σ
for each attribute x in C_{cl}
    if (RandomNumber (0, 1) < $P_{\#}$)
        $x \leftarrow$ #
    else
        $x \leftarrow$ the corresponding attribute in σ
  $A_{cl} \leftarrow$ random action not present in [M]
  $p_{cl} \leftarrow p_I$
$\varepsilon_{cl} \leftarrow \varepsilon_I$
$f_{cl} \leftarrow f_I$
$exp_{cl} \leftarrow 0$
$ts_{cl} \leftarrow$ actual time t
$as_{cl} \leftarrow 1$
$num_{cl} \leftarrow 1$

GENERATE PREDICTION ARRAY ([M]):

initialize prediction array PA to all null
initialize fitness sum array FSA to all 0.0
for each classifier cl in [M]
    if ($PA[A_{cl}]$ = null)
        $PA[A_{cl}] \leftarrow p_{cl} * f_{cl}$
    else
        $PA[A_{cl}] \leftarrow PA[A_{cl}] + p_{cl} * f_{cl}$
    $FSA[A_{cl}] \leftarrow FSA[A_{cl}] + f_{cl}$
for each possible action A
    if ($FSA[A]$ is not zero)
        $PA[A] \leftarrow PA[A] / FSA[A]$
return PA

SELECT ACTION (PA):

if (RandomNumber[0, 1) < p_{explr})
    //Do pure exploration here
    return a randomly chosen action from those not null in PA
else
    //Do pure exploitation here
    return the best action in PA

GENERATE ACTION SET ([*M*], *act*):
initialize empty set [*A*]
for each classifier *cl* in [*M*]
    if ($A_{cl} = act$)
    add classifier *cl* to set [*A*]

UPDATE SET ([*A*], P, [*P*]):
for each classifier *cl* in [*A*]
    exp_{cl}++
    //update prediction p_{cl}
    if ($exp_{cl} < 1 / \beta$)
        $p_{cl} \leftarrow p_{cl} + (\mathrm{P} - p_{cl}) / exp_{cl}$
    else
        $p_{cl} \leftarrow p_{cl} + \beta * (\mathrm{P} - p_{cl})$
    //update prediction error ε_{cl}
    if ($exp_{cl} < 1 / \beta$)
        $\varepsilon_{cl} \leftarrow \varepsilon_{cl} + (|\mathrm{P} - p_{cl}| - \varepsilon_{cl}) / exp_{cl}$
    else
        $\varepsilon_{cl} \leftarrow \varepsilon_{cl} + \beta * (|\mathrm{P} - p_{cl}| - \varepsilon_{cl})$
    //update action set size estimate as_{cl}
    if ($exp_{cl} < 1 / \beta$)
        $as_{cl} \leftarrow as_{cl} + (\sum_{C \in [A]} num_c - as_{cl}) / exp_{cl}$
    else
        $as_{cl} \leftarrow as_{cl} + \beta * (\sum_{C \in [A]} num_c - as_{cl})$
UPDATE FITNESS in set [*A*]
if (*doActionSctSubsumption*)
    DO ACTION SET SUBSUMPTION in [*A*] updating [*P*]

UPDATE FITNESS ([*A*]):
accuracySum $\leftarrow$ 0
initialize accuracy vector *k*
for each classifier *cl* in [A]
    if ($\varepsilon_{cl} < \varepsilon_0$)
        $k(cl) \leftarrow 1$
    else
        $k(cl) \leftarrow \alpha * (\varepsilon_{cl} / \varepsilon_0)^{-\upsilon}$
    $accuracySum \leftarrow accuracySum + k(cl) * num_{cl}$

for each classifier *cl* in [*A*]

$f_{cl} \leftarrow f_{cl} + \beta * (k(cl) * num_{cl} / accuracySum - f_{cl})$

Appendix B. Knowledge Population

Table: Knowledge Population

Knowledge	Condition part	Action part
1	111100000011000110000	001
2	011100000001010110100	101
3	111100000111000111010	011
4	111100000011000111100	011
5	011100000111000110000	010
6	000101001000111110011	100
7	000111001000111011100	011
8	100011100000111011111	111
9	000011100000111011111	001
10	000011100100111011001	011
11	110011101100110011000	110
12	000011100100111011000	111
13	000011100110111010111	000
14	100011100000111010010	011
15	000011111000111010010	001
16	000011100000111010001	011
17	100011100000111011101	010
18	010011101000111011100	110
19	000011111000111010100	000
20	000011100111001010001	011
21	000011100000111011000	011
22	110011101111001011000	101
23	110011101111001011111	110
24	010011101000010011100	000
25	100011100100110011000	100
26	111000001011011011000	100
27	111100000111000110000	100
28	011100101001111110110	110
29	001101100000111110110	100