

Chapter 1

Introduction

1.1 Brief History of AI

Artificial Intelligence (AI) is usually defined as the science and engineering of imitating, extending and augmenting human intelligence through artificial means and techniques to make intelligent machines. In 2005, John McCarthy pointed out that the long-term goal of AI is human-level AI (McCarthy, 2005).

In the history of human development, it is a never-ending pursuit to free people from both manual and mental labor with machines. The industrial revolutions enable machines to perform heavy manual labor instead of people, and thus lead to a considerable economic and social progress. To make machines help relieve mental labor, a long cherished aspiration is to create and make use of intelligent machines like human beings.

In ancient China, many mechanical devices and tools have been invented to help accomplish mental tasks. The abacus was the most widely used classical calculator. The Water-powered Armillary Sphere and Celestial Globe Tower is used for astronomical observation and stellar analysis. The Houfeng Seismograph is an ancient seismometer to detect and record tremors and earthquakes. The traditional Chinese theory of Yin and Yang reveals the philosophy of opposition, interrelation and transformation, having an important impact on modern logics.

In the world, Aristotle (384-322, BC) proposed the first formal deductive reasoning system, syllogistic logic, in the *Organon*. Francis Bacon (1561-1626) established the inductive method in the *Novum Organum* (or “New Organon”). Gottfried Leibniz (1646-1716) constructed the first mechanical calculator capable of multiplication and division. He also enunciated the concepts of “*characteristica universalis*” and “*calculus ratiocinator*” to treat the operations of

formal logic in a symbolic or algebraic way, which can be viewed as the sprout of the “thinking machine”.

Since the 19th century, advancement of sciences and technologies such as Mathematical Logic, Automata Theory, Cybernetics, Information Theory, Computer Science and Psychology laid the ideological, theoretical and material foundation for the development of AI research. In the book “An Investigation of the Laws of Thought”, George Boole (1815-1864) developed the Boolean algebra, a form of symbolic logic to represent some basic rules for reasoning in the thinking activities. Kurt Gödel (1906-1978) proved the incompleteness theorems. Alan Turing (1912-1954) introduced the Turing Machine, a model of the ideal intelligent computer, and initiated the automata theory. In 1943, Warren McCulloch (1899-1969) and Walter Pitts (1923-1969) developed the MP neuron, a pioneer work of Artificial Neural Networks research. In 1946, John Mauchly (1907-1980). and John Eckert (1919-1995) invented the ENIAC (Electronic Numerical Integrator And Computer), the first electronic computer. In 1948, Norbert Wiener (1894-1964) published a popular book of “Cybernetics”, and Claude Shannon (1916-2001) proposed the Information Theory.

In a real world, quite a number of problems are complex ones, most of the times without any algorithm to adopt; or even if there are calculation methods, they are still NP problems. Researchers might introduce heuristic knowledge to solve such problem-solving to simplify complex problems and find solutions in the vast search space. Usually, the introduction of domain-specific empirical knowledge will produce satisfactory solutions, though they might not be the mathematically optimal solutions. This kind of problem solving with its own remarkable characteristics led to the birth of AI. In 1956, the term “Artificial Intelligence” was coined, and the Dartmouth Summer Research Project on Artificial Intelligence, proposed by John McCarthy, Marvin Minsky, etc., was carried on at Dartmouth College with several American scientists of psychology, mathematics, computer science and information theory. This well-known Dartmouth conference marked the beginning of the real sense of AI as a research field. Through dozens of years of research and development, great progress has been made in the discipline of AI. Many artificial intelligence expert systems have been developed and applied successfully. In domains such as Natural Language Processing, Machine Translation, Pattern Recognition, Robotics and Image Processing, a lot of achievements have been made, and the applications span various areas to promote their development.

In the 1950's, AI research mainly focused on game playing. In 1956, Arthur Samuel wrote the first heuristic game-playing program with learning ability. In the same year, Alan Newell, Herbert Simon etc. invented a heuristic program called the Logic Theorist, which proved correct 38 of the first 52 theorems from the "Principia Mathematica". Their work heralded the beginning of research on cognitive psychology with computers. Noam Chomsky proposed the Syntactics, the pioneer work of Formal Language research. In 1958, John McCarthy invented the Lisp language, an important tool for AI research which can process not only numerical values but also symbols.

In the early 1960's, AI research mainly focused on search algorithms and general problem solving (GPS). Allen Newell etc. published the General Problem Solver, a more powerful and universal heuristic program than other programs at that time. In 1961, Marvin Minsky published the seminal paper "Steps Towards Artificial Intelligence" established a fairly unified terminology for AI research and established the subject as a well- defined scientific enterprise. In 1965, Edward Feigenbaum etc. began work on the DENDRAL chemical-analysis expert system, a milestone for AI applications, and initiated the shift from computer algorithms to knowledge representation as the focus of AI research. In 1965, Alan Robinson proposed the Resolution Principle. In 1968, Ross Quillian introduced the Semantic Network for knowledge representation. In 1969, IJCAI (International Joint Conferences on Artificial Intelligence) was founded, and since then, the International Joint Conference on Artificial Intelligence (also shorted as IJCAI) was held biannually in odd-numbered years. Artificial Intelligence, an international journal edited by IJCAI, commenced publication in 1970.

In the early 1970's, AI research mainly focused on Natural Language Understanding and Knowledge Representation. In 1972, Terry Winograd published details of the SHRDLU program for understanding natural language. Alain Colmerauer developed Prolog language for AI programming at the University of Marseilles in France. In 1973, Roger Schank proposed the Conceptual Dependency Theory for Natural Language Understanding. In 1974, Marvin Minsky published the frame system theory, an important theory of Knowledge Representation. In 1977, Edward Feigenbaum published the well-known paper "The art of artificial intelligence: Themes and case studies in knowledge engineering" in the 5th IJCAI. He stated that Knowledge Engineering is the art of bringing the principles and tools of AI research to bear on difficult applications problems requiring expert knowledge for their solution. The

technical issues of acquiring this knowledge, representing it, and using it appropriately to construct and explain lines-of-reasoning, are important problems in the design of knowledge-based systems.

In the 1980's, AI research developed prosperously. Expert systems were more and more widely used, development tools for expert systems appeared, and industrial AI thrived. Especially in 1982, the Japan's Ministry of International Trade and Industry initiated the Fifth Generation Computer Systems project, which dramatically promoted the development of AI. Many countries also made similar plans for research in AI and intelligent computers. China also started the research of intelligent computer systems as an 863 National High-Tech Program.

During the past more than 50 years, great progress has been made of AI research. Theories of Heuristic Searching Strategies, Non-monotonic Reasoning, Machine Learning, etc. have been proposed. Applications of AI, especially Expert Systems, Intelligent Decision Making, Intelligent Robots, Natural Language Understandings, etc. also promoted the research of AI. Presently, Knowledge Engineering based on knowledge and information processing is a remarkable characteristic of AI.

However, just as the development of any other discipline, there are also obstacles in the history of AI research. Even from the beginning, AI researchers had been criticized for their being too optimistic. In the early years of AI research, Herbert Simon and Allen Newell, two of the AI pioneers, optimistically predicted that:

- Within ten years, a digital computer will be the world's chess champion, unless the rules bar it from competition.
- Within ten years, a digital computer will discover and prove an important new mathematical theorem.
- Within ten years, a digital computer will write music that will be accepted by critics as possessing considerable aesthetic value.
- Within ten years, most theories in psychology will take the form of computer programs, or of qualitative statements about the characteristics of computer programs.

These expectations haven't been completely realized even till today. 3 year old little child can easily figure out a tree in a picture, while a most powerful super computer only reaches middle level as children in tree recognition. It is

also very difficult to automatically understand even stories written for little children.

Some essential theories of AI still need improvements. No breakthrough progresses have been made for some key technologies such as Machine Learning, Non-monotonic Reasoning, Common Sense Knowledge Representation and Uncertain Reasoning. It is also very difficult for global judgment, fuzzy information processing, multi-granular visual information processing, etc

Conclusively, AI research is still in the first stage of Intelligence Science, an indispensable cross discipline which dedicates to joint research on basic theories and technologies of intelligence by Brain Science, Cognitive Science, Artificial Intelligence and others. Brain Science explores the essence of brain and investigates the principles and models of natural intelligence in molecular, cellular and behavioral level. Cognitive Science studies human mental activities, such as perception, learning, memory, thinking and consciousness. AI research aims at imitating, extending and augmenting human intelligence through artificial means and techniques, and finally achieving machine intelligence. These three disciplines work together to explore new concepts, new theories and new methodologies for Intelligence Science, opening up prospects for a successful and brilliant future in the 21st century (Shi, 2006a).

1.2 Cognitive Issues of AI

Cognition is generally referred to as the process of knowing or understanding relative to affection, motivation or volition. Definitions of cognition can be briefly summarized into 5 main categories according to American psychologist Houston etc:

- (1) Cognition is the process of information processing;
- (2) Cognition involves symbol processing in psychology;
- (3) Cognition deals with problem solving;
- (4) Cognition studies mind and intelligence;
- (5) Cognition consists of a series of activities, such as perception, memory, thinking, judgment, reasoning, problem solving, learning, imagination, concept forming, language using, etc.

Cognitive psychologist David H. Dodd etc. held that cognition involves three aspects of adaptation, structure and process, i.e., cognition is the process of information processing in certain mental structures for certain objectives.

Cognitive Science is the science about human perceptions and mental information processing, spanning from perceptual input to complex problem solving, including intellectual activities from individuals to the whole society, and investigating characteristics of both human intelligence and machine intelligence (Shi, 1990). As an important theoretical foundation for AI, Cognitive Science is an interdisciplinary field developed from Modern Psychology, Information Science, Neuroscience, Mathematics, Scientific Linguistics, Anthropology, Natural Philosophy, etc.

The blooming and development of Cognitive Science marked a new stage of research on human-centered cognitive and intelligent activities. Research on Cognitive Science will enable self understanding and self control, and lift human knowledge and intelligence to an unprecedented level. Moreover, it will lay theoretical foundations for the intelligence revolution, knowledge revolution and information revolution, as well as provide new concepts, new ideas and new methodologies for the development of intelligent computer systems.

Promoted by works of Allen Newell and Herbert Simon, research related to cognitive science originated in the late 1950's (Simon, 1986). Cognitive scientists proposed better models for mind and thinking than the simplified model about human developed by behaviorism scientists. Cognitive Science research aims at illustrating and explaining how information is processed during cognitive activities. It involves varieties of problems including perception, language, learning, memory, thinking, problem solving, creativity, attention, as well as the impact of environment and social culture on cognition.

In 1991, the representative journal "Artificial Intelligence" published a special issue on the foundation of AI in its 47th volume, in which trends about AI research are discussed. In this special issue, David Kirsh discussed five foundational questions for AI research (Kirsh, 1991):

- (1) Pre-eminence of knowledge and conceptualization: Intelligence that transcends insect-level intelligence requires declarative knowledge and some form of reasoning-like computation-call this cognition. Core AI is the study of the conceptualizations of the world presupposed and used by intelligent systems during cognition.
- (2) Disembodiment: Cognition and the knowledge it presupposes can be studied largely in abstraction from the details of perception and motor control.

- (3) Kinematics of cognition are language-like: It is possible to describe the trajectory of knowledge states or informational states created during cognition using a vocabulary very much like English or some regimented logic-mathematical version of English.
- (4) Learning can be added later: The kinematics of cognition and the domain knowledge needed for cognition can be studied separately from the study of concept learning, psychological development, and evolutionary change.
- (5) Uniform architecture: There is a single architecture underlying virtually all cognition.

All these questions are cognitive problems critical to AI research, which should be discussed from the perspective of fundamental theories of Cognitive Science. These questions have become the watershed for different academic schools of AI research, as different academic schools usually have different answers to them.

1.3 Hierarchical Model of Thought

Thought is the reflection of the objective realities, i.e. the conscious, indirect and general reflection in a conscious human brain on the essential attributes and internal laws about the objective realities. Currently, we are in a stage emphasizing self knowledge and self recognition with the development of the Cognitive Science. In 1984, Professor Xuesen Qian advocated the Noetic Science research (Qian, 1986).

Human thought mainly involves perceptual thought, imagery thought, abstract thought and inspirational thought. Perceptual thought is the primary level of thought. When people begin to understand the world, perceptual materials are simply organized to form self-consistent information, thus only phenomena are understood. The form of thought based on this process is perceptual thought. Perceptual thought about the surface phenomena of all kinds of things can be obtained in practice via direct contact with the objective environment through sensories such as eyes, ears, noses, tongues and bodies, thus its sources and contents are objective and substantial.

Imagery thought mainly relies on generalization through methods of typification and the introduction of imagery materials in thinking. It is common to all higher organisms. Imagery thought corresponds to the connection theories of neural mechanisms. AI topics related to imagery thought include Pattern Recognition, Image Processing, Visual Information Processing, etc.

Abstract thought is a form of thought based on abstract concepts, through thinking with symbol information processing. Only with the emergence of language is abstract thought possible: language and thought boost each other and promote each other. Thus, physical symbol system can be viewed as the basis of abstract thought.

Little research has been done on inspirational thought. Some researchers hold that inspirational thought is the extension of imagery thought to sub-consciousness, during which a person does not realize that part of his brain is processing information. While some others argue that inspirational thought is sudden enlightenment. Despite all these disagreements, inspirational thought is very important to creative thinking, and need further research.

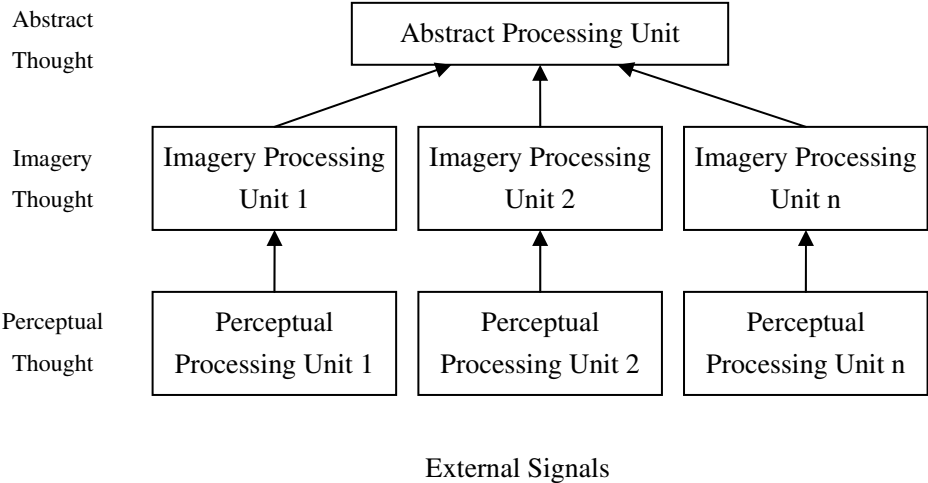


Fig. 1.1. Hierarchical model of thought

In the process of human thinking, attention plays an important role. Attention sets certain orientation and concentration for noetic activities to ensure that one can promptly respond to the changes of the objective realities and be better accustomed to the environment. Attention limits the number of parallel thinking. Thus for most conscious activities, the brain works serially, with an exception of parallel looking and listening.

Based on the above analysis, we propose a hierarchical model of human thought, as shown in Fig. 1.1 (Shi, 1990a; Shi 1992; Shi 1994). In the figure, perceptual thought is the simplest form of thought, which is constructed from the

surface phenomena through sensories such as eyes, ears, noses, tongues and bodies. Imagery thought is based on the connection theories of neural networks for highly parallel processing. Abstract thought is based on the theory of physical symbol system in which abstract concepts are represented with languages. With the effect of attention, different forms of thought are processed serially most of the time.

The model of thought studies the interrelationships among these three forms of thought, as well as the micro processes of transformation from one form to be other. Presently, much progress has been made. For example, attractors of neural networks can be used to represent problems such as associative memory and image recognition. Yet there is still a long way to go for a thorough understanding and application of the whole model. For example, further research is needed on the micro-process from imagery thought to logical thought.

1.4 Symbolic Intelligence

What is intelligence? Intelligence involves purposeful actions, reasonable thinking, as well as comprehensive capabilities to effectively adapt to the environment. Generally speaking, intelligence is one's capabilities to understand the objective world and apply knowledge to solve problems. Intelligence of an individual consists of comprehensive capabilities, such as: capability to perceive and understand objective things, the objective world and oneself; capability to gain experience and acquire knowledge through learning; capability to comprehend knowledge and apply knowledge and experience for problem analysis and problem solving; capabilities of association, reasoning, judgment and decision making; capability of linguistic abstraction and generalization; capabilities of discovery, invention, creativity and innovation; capability to timely, promptly and reasonably cope with the complex environments; capability for predictions of and insights into the development and changes of things. People live in the society, thus their intelligence is interrelated with the social environments. With the continuous development of human society, concepts of intelligence also evolve gradually.

AI (Artificial Intelligence), compared with natural intelligence of human, aims at imitating, extending and augmenting human intelligence through artificial means and techniques to achieve certain machine intelligence. The science of AI focuses on computational models of intelligent behaviors, develops

computer systems for noetic activities such as perception, reasoning, learning, association, decision making, etc., and solves complex problems that only human experts can solve.

In the history of AI research, different levels of thought are studied from different views of Symbolicism, Connectionism and Behaviorism.

Symbolicism is also known as traditional AI. It is based on the physical symbol system hypothesis proposed by Alan Newell and Herbert Simon, which states that a physical symbol system has the necessary and sufficient means for general intelligent action. A physical symbol system consists of a set of entities, called symbols, which are physical patterns that can occur as components of another type of entity called an expression (or symbol structure). The system also contains a collection of processes that operate on expressions to produce other expressions: processes of creation, modification, reproduction and destruction.

Connectionism, also known as neural computing, focuses on the essentials and capabilities for non-programmatical, adaptative and brain-like information processing. The research field is rapidly developing in recent years, with a great number of neural network mechanisms, models and algorithms emerged continuously. Neural network systems are open neural network environments providing typical and practically valuable neural network models. The open system enables convenient adding of new network models to the existing system, so that new network algorithms can be debugged and modified with the friendly user interfaces and varieties of tools provided by the system. Moreover, it is also convenient to improve existing network models, thus the system provides excellent environment to develop new algorithms.

Neural computing investigates the brain functionalities based on the nervous system of human brains, and studies the dynamic actions and collaborative information processing capabilities of large numbers of simple neurons. The research focuses on the simulation and imitation of human cognition, including processes of perception and consciousness, imagery thought, distributed memory, self-learning and self-organization. Neural computing is particularly competent in parallel search, associative memory, self-organization of spatio-temporal data statistical descriptions, and automatic knowledge acquisition through interrelated activities. It is generally considered that neural networks better fitted low level pattern processing.

Basic characteristics of neural networks include: a. distributed information storage, b. parallel information processing, c. capabilities of self-organization and self-learning (Shi, 1993). Owing to these characteristics, neural networks provide

a new means for information processing with computers. With more and more applications and in depth research of artificial neural networks, the researchers have found many problems of existing models and algorithms, and even met with some difficulties of the nonlinear theories or approximation theory. Despite these problems and difficulties, we believe that with the in-depth and extensive applications, neural networks will continue to develop and promote current techniques. The theory of neural field we proposed is such a new kind of attempt.

Currently, integration of symbol processing systems and neural network models is an important research direction. Fuzzy neural networks integrate fuzzy logic and neural networks, taking each other's advantages in theory, methodology and application, to develop systems with certain learning and dynamic knowledge acquisition capabilities.

Behaviorism, also known as behavior-based AI, in many respects reflects the behavior physiological views in AI. Rodney Brooks brought forward theories of intelligence without representation (Brooks, 1991a) and intelligence without reasoning (Brooks, 1991b), and stated that intelligence is determined by the dynamics of interaction with the world.

These three research genres investigate different aspects of human natural intelligence corresponding to different layers in the model of human thought. Roughly categorizing, it can be taken that Symbolicism focuses on abstract thought, Connectionism focuses on imagery thought, while Behaviorism focuses on perceptual thought. The comparisons of Symbolicism, Connectionism and Behaviorism is shown in Table 1.1

Table 1.1 Comparisons of Symbolicism, Connectionism and Behaviorism.

	Symbolicism	Connectionism	Behaviorism
Perceptual Level	Discrete	Continuous	Continuous
Representation Level	Symbolic	Continuous	Behavioral
Problem Solving Level	Top-down	Bottom-up	Bottom-up
Processing Level	Serial	Parallel	Parallel
Operational Level	Reasoning	Connection	Interaction
System Level	Local	Distributed	Distributed
Basic Level	Logic	Simulant	Intuitional Judgment

Some researchers classify AI research into two categories: symbolic intelligence and computational intelligence. Symbolic intelligence, also known as

traditional AI, solves problems through reasoning based on knowledge. Computational intelligence solves problems based on connections trained from example data. Artificial Neural Networks, Genetic Algorithms, Fuzzy Systems, Evolutionary Programming, Artificial Life, etc. are included in computational intelligence.

Presently, traditional AI mainly focuses on knowledge based problem solving. In the practical point of view, AI is the science of knowledge engineering: taking knowledge as the object and investigating knowledge representation, acquisition and application. This book mainly introduces and discusses traditional AI. For computational intelligence, please refer to the book “Neural Networks” by Zhongzhi Shi (Shi, 2009).

1.5 Research Approaches of Artificial Intelligence

During the development of AI since the 1950’s, many academic schools have been formed, each holding its specific research methodologies, academic views and research focuses. This section introduces some research methodologies of AI, focusing mainly on the cognitive school, logical school, and behavioral school.

1.5.1 Cognitive School

Cognitive school, with representative researchers such as Herbert Simon, Marvin Minsky and Allen Newell, focuses on functional simulation with computers based on human noetic activities. In the 1950’s, Newell and Simon advocated the “heuristic program” together, and worked out the “Logic Theorist” computer program to simulate the thinking process of mathematical theorem proving. Then in the early 1960’s, they developed the “General Problem Solver (GPS)”, which simulates the common principles of human problem solving with three steps: first, set the initial problem solving plan; then, apply axioms, theorems and rules to solve the problems according to the plan; continually proceed with the means-end analysis, and modify the problem solving until the goal is achieved. Thus the GPS possesses certain universality.

In 1976, Newell and Simon proposed physical symbol system premise, and stated that a physical symbol system has the necessary and sufficient means for general intelligent action. Thus, an information processing system can be viewed as a concrete physical system, such as human neural system, computer

construction system, etc. Each physical pattern is a symbol, as long as it can be distinguished from other patterns. For example, different English characters are different symbols. To operate on symbols relies on comparison among different symbols, i.e. distinguishing which symbols are the same and which ones are different. Thus fundamental task and functionality of a physical symbol system is to identify same characters and distinguish different ones.

In the 1980's, Newell etc. focused on the SOAR system, a symbolic cognitive architecture for general problem solving, based on the Chunking mechanism for learning and rule-based memory for representation of operators, search control, etc.

Minsky took the view of psychics, holding that in daily activities, people apply plenty of knowledge acquired and collected from previous experiences. Such knowledge is stored in the brain in a structure similar to frame. Thus, he proposed the frame knowledge representation structure in the 1970's. In the 1980's, Minsky believed that there is no unified theory for human intelligence. In the famous book "Society of Mind" he published in 1985, Minsky pointed out that the society of mind is a vast society of individually simple agents with certain thinking capabilities.

1.5.2 Logical School

Logical school, with representative researchers such as John McCarthy and Nils Nilsson, holds the logical perspective for AI research, i.e. describe the objective world through formalization. This academic school believes that:

- Intelligent machines will have knowledge of their environment.
- The most versatile intelligent machines will represent much of their knowledge about their environment declaratively.
- For the most versatile machines, the language in which declarative knowledge is represented must be at least as expressive as first order predicate calculus.

Logical school focuses on conceptual knowledge representation, model theoretic semantics, deductive reasoning, etc. in AI research. McCarthy claimed that everything can be represented with the unified frame of logics, and common sense reasoning will be difficult without some form of non-monotonic reasoning.

1.5.3 Behavioral School

Most AI research is based on too abstract and simple models for the real world. Rodney Brooks argued that it is necessary to go beyond this ivory tower of abstract models, and take the complex real world as the background instead, so that AI theories and technologies can be tested in real world problem solving, and improved in these tests.

In 1991 Brooks brought forward theories of intelligence without representation and intelligence without reason, and stated that intelligence is determined by the dynamics of interaction with the world. He simply called this work as “robots” or “behavior-based robots”. There are a number of key aspects characterizing this style of work as follows (Brooks, 1991b):

- **Situatedness:** The robots are situated in the world and the world directly influences the behavior of the system.
- **Embodiment:** The robots have bodies and experience the world directly.
- **Intelligence:** The source of intelligence is not limited to just the computational engine. It also come from the situation in the world.
- **Emergence:** The intelligence of the system emerges from the system’s interactions with the world and sometimes indirect interactions between its components.

Based on these ideas, Brooks programmed autonomous mobile robots, based on layered, asynchronous and distributed networks of augmented finite-state machines, each one being a comparatively independent unit for functionalities of advance, balance, prowl, etc. The robot walked successfully, and thus initiated a new approach to Robotics.

Different academic schools of AI research have different answers to the five foundational cognitive questions introduced in section 1.2. The logical school (represented by Nils Nilsson) holds positive answers to questions 1-4, and neutral answer to question 5; the cognitive school (represented by Allen Newell) holds positive answers to questions 1, 3 and 5; while the behavioral school (represented by Rodney Brooks) holds negative answers to all question of 1-5.

1.6 Automated Reasoning

Reasoning is the cognitive process of logically inferring a new judgment (conclusion) from one or more already known judgments (precondition). It is the reflection of the objective relationships in mind. People usually solve problems based on prior knowledge and make conclusions through reasoning. Theories and technologies of automated reasoning are important bases for research fields of program derivation, proof of program correctness, expert systems, intelligent robots, etc.

Early works of automated reasoning focused on automated theorem proving. Pioneer work includes the Logic Theorist developed by Herbert Simon and Allen Newell. In 1956, Alan Robinson proposed the Resolution Principle, making a great progress in research on automated reasoning. The resolution principle is easily applicable and logically complete, thus it becomes the computing model for the logic programming language Prolog. Though some methods outperforming the Resolution Principle in some aspects appeared later, e.g. natural deductive reasoning and term rewriting systems, yet they are limited due to the combination problem and the computational intractability essentially.

For a practical system, there always exist some non deductive cases. Thus, various reasoning algorithms have been proposed, which even weakens the attempt of finding a universal fundamental principle for AI. From the practical perspective of view, each reasoning algorithm conforms to its specific, domain related strategies based on different knowledge representation techniques. On the other hand, it is undoubtedly useful to find a universal reasoning theory. In fact, an important impetus for AI theoretical research is to find more general and universal reasoning algorithms.

An important achievement of automated reasoning research is nonmonotonic reasoning, a pseudo induction system. The so called nonmonotonic reasoning is the reasoning process in which adding new positive axioms to the system may invalidate some already proved theorems. Obviously, nonmonotonic reasoning is more complex than monotonic reasoning. In nonmonotonic reasoning, first hypotheses are made; then standard logical reasoning is carried out; if inconsistency appeared, then backtrack to eliminate inconsistency, and establish new hypothesis.

Raymond Reiter first set forth the closed world assumption (CWA) for nonmonotonic reasoning in 1978 (Reiter, 1978), and proposed the Default Reasoning (Reiter, 1980). In 1979, Jon Doyle developed the truth maintenance

system (TMS) (Doyle, 1979). In 1980, John McCarthy formalized the theory of Circumscription (McCarthy, 1980). Circumscription of a predict P means to exclude most models based on P, and select only a minimum set of models in which P is assigned to true. Different circumscription criteria will produce different minimizations of predicates.

Quantitative simulation with computers is commonly applied for scientific computing. Yet people often predict or explain system behaviors without detailed calculation data. Such problem solving can not be achieved simply through deduction, thus qualitative reasoning is proposed in AI for representation and reasoning without precise quantitative information. In qualitative reasoning, physical systems or procedures can be decomposed into subsystems or model fragments, each with structuralized specifications of the subsystem itself and its interrelationships with other subsystems. Through approaches such as causal ordering and compositional modeling, functionalities and behaviors of the real physical systems can be qualitatively represented. Typical qualitative reasoning techniques include: QDE (qualitative differential equation) based modeling and reasoning by Johan de Kleer, process-centered modeling and reasoning by Kenneth Forbus, and constraint-centered qualitative simulation by Benjamin Kuipers. Combined approaches of quantitative and qualitative reasoning will make great impact to scientific decision making of expert systems.

Uncertainty is ubiquitous to real world problems, which results from the deviation of people's subjective cognition from the objective realities. Various causes may reflect such deviation and bring about uncertainty, such as randomness of things, incompleteness, unreliability, imprecision and inconsistency of human knowledge, and vagueness and ambiguousness of natural language. With respect to different causes of uncertainty, different theories and reasoning methodologies have been proposed. In AI and knowledge engineering, representative approaches of uncertainty theories and reasoning methodologies are introduced in the following.

Probability theory is widely used to process randomness and uncertainty of human knowledge. Bayesian theory has been successfully applied in the PROSPECTOR expert system, yet it relies on assigned prior probabilities. The MYCIN model based certainty factors, adopting some assumptions and principles for conjunction of hypothesis, is a simple and effective method, though it lacks well established theoretical foundations.

Dempster-Shafer theory of evidence introduces the concept of belief function to extend classical probabilities, and defines that belief function satisfies a set of

axioms weaker than probability axioms, thus belief function can be viewed as a superset of existing probability functions. With belief function, even without precise probabilities, constraints on probability distributions can be set based on prior domain knowledge. The theory has well established theoretical foundations, yet its definition and computation is comparatively complex. In recent years, this theory of evidence has gained more and more research focuses, and many research achievements and application systems have been developed. For example, Lotfi Zadeh illustrated how the Dempster-Shafer theory can be viewed as an instance of inference from second-order relations, and applied in a relational database.

In 1965, Lotfi Zadeh proposed the Fuzzy Set, based on which a series of research have been made, including fuzzy logic, fuzzy decision making, probability theory, etc. For reasoning with natural language, Zadeh introduced fuzzy quantization to represent fuzzy propositions in natural language, defined concepts of linguistic variable, linguistic value and probability distribution, developed possibility theory and approximate reasoning. His work has attracted much research focuses. Fuzzy mathematics has been widely applied to expert systems and intelligent controllers, as well as for the research of fuzzy computer. Chinese researchers have done a lot in theoretical research and practical applications, drawing much attention from the international academics. However, many theoretical problems still remain to be solved in this domain. There are also some different views and disputes, such as, what is the basis for fuzzy logic? What about the problem of consistency and completeness of fuzzy logic? In the future, research focuses of uncertain reasoning may be centralized on the following three aspects: first, to solve existing problems of current uncertainty theories; second, to study the efficient and effective discrimination capabilities and judgment mechanisms of human beings for new theories and new methodologies to deal with uncertainties; and third, to explore methods and technologies to synthetically process varieties of uncertainties.

Theorem proving is a kind of specific intelligent behavior of human, which not only relies on logic deductions based on premises, but also requires certain intuitive skills. Automated theorem proving adopts a suit of symbol systems to formalize the process of human theorem proving into symbol calculation that can be automatically implemented by computers, i.e., to mechanize the intelligence process of reasoning and deduction. The mechanical theorem proving in elementary geometry and differential geometry proposed by Professor Wenjun Wu of Chinese Academy of Sciences is highly valued all over the world.

1.7 Machine Learning

Knowledge, knowledge representation and knowledge based reasoning algorithms are always considered at the heart of AI, while machine learning can be viewed as a most critical problem. For hundreds of years, the psychologists and philosophers held that the basic mechanism of learning is trying to transfer successful behaviors in one practice to other similar practices. Learning is the process of acquiring knowledge, gaining experience, improving performance, discovering rules and adapted to environments. Fig. 1.2 illustrates a simple model of learning with four basic elements of a learning system. The environment provides external information, similar to a supervisor. The learning unit processes information provided by the environment, corresponding to various learning algorithms. The knowledge base stores knowledge in certain knowledge representation formalisms. The performing unit accomplishes certain tasks based on the knowledge in the knowledge base, and sends the execution results to the learning unit through feedbacks. The system can be gradually improved through learning. Research on machine learning not only enables machines to automatically acquire knowledge and obtain intelligence, but also uncovers principles and secrets of human thinking and learning, and even helps to improve the efficiency of human learning. Research on machine learning also has a great impact on memory storage patterns, information input methods and computer architectures.

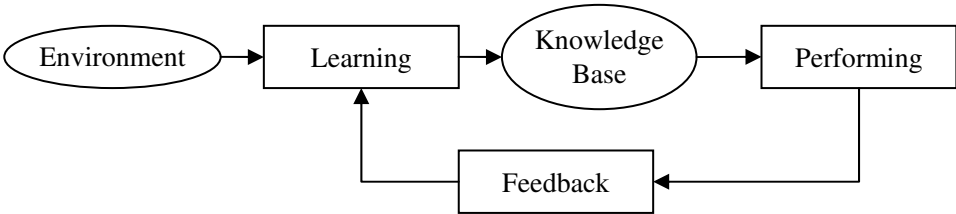


Fig. 1.2. Simple model of learning(Simon, 1983)

Research in machine learning roughly experienced four stages. The first and initial stage is learning without knowledge, focusing on neural models and self adaptative and self organization systems based on decision theories. However, as neural models and decision theories were fairly restricted and only achieved limited success, the research passion gradually depressed. The second stage in

the 1960's is the low tide, focusing mainly on symbolic concept acquisition. Then in the third stage, interest in machine learning rejuvenated and many distinctive algorithms appeared since Patrick Winston's important paper of "Learning Structural Descriptions from Examples" in 1975. More importantly, it was then popularly recognized that a learning system would not learn high level concepts without background knowledge. Thus, great amount of knowledge were introduced to learning systems as background knowledge, bringing about a new era and new prospects for machine learning research. Due to the mass applications of expert systems and problem solving systems, knowledge acquisition has become the key bottleneck, to solve which heavily relies on the advances of machine learning research. There comes the fourth stage and another climax of machine learning research.

Main paradigms of machine learning include inductive learning, analytical learning, discovery learning, genetic learning, connection learning, etc. (Shi 1992b). Inductive learning has been most extensively studied in the past, focused mainly on general concept description and concept clustering, and proposed algorithms such as the AQ algorithms, version space algorithm, and ID3 algorithm. Analogical learning analyzes similarities of the target problem with previously known source problems, and then applies the solutions from the source problems to the target problem. Analytical learning, e.g. explanation-based learning, chunking, etc., learns from training examples guided by domain knowledge. Explanation-based learning extracts general principles from a concrete problems solving process which can be applied to other similar problems. As learned knowledge is stored in the knowledge base, intermediate explanations can be skipped to improve the efficiency of future problem solving. Discovery learning is the method to discover new principles from existing experimental data or models. In recent years, knowledge discovery in databases (KDD, also known as data mining, DM) has attracted great research focuses, which is considered to be a very practically useful research discipline by AI and database researchers. KDD mainly discovers classification rules, characteristic rules, association rules, differentiation rules, evolution rules, exceptional rules, etc. through methods of statistical analysis, machine learning, neural networks, multidimensional database, etc. Genetic learning based on the classic genetic algorithm is designed to simulate biological evolution via reproduction and variation and Darwin's natural selection paradigm. It takes each variant of a concept as an individual of the species, and evaluates different mutations and recombinations based on objective fitness functions, so as to select the fittest

offsprings for survival. Connection learning recognizes different input patterns through training the neural networks with typical example instances.

Machine learning research is still in its primary stage, and needs extensive research efforts. Progress in machine learning research will enable breakthroughs in AI and knowledge engineering research. In the future, research focuses of machine learning will include cognitive models for the learning process, computational learning theories, new learning algorithms, machine learning systems integrating multiple learning strategies, etc.

1.8 Distributed Artificial Intelligence

Studies of human intellectual behaviors show that most human activities involve social groups consisting of multiple individuals, and large-scale complex problem solving also involves cooperation of several professionals or organizations. “Cooperation” is a major aspect of human intelligence pervasive in the human society, and thus the motivation for research in Distributed Artificial Intelligence (DAI).

With the development of computer network, computer communication and concurrent programming technologies since the 1980’s, DAI is gradually becoming a new research focus in the field of AI. DAI is a subfield of AI investigating how logically and physically distributed agents cooperate with each other to perform intelligent behaviors. It enables collaborated and coordinated knowledge, skills and planning, solves single-objective and multi-objective problems, and provides an effective means for the design and construction of large-scale complex intelligent systems or computers to support cooperation.

The term DAI was coined by American researchers, and the first International Workshop on Distributed Artificial Intelligence was held at MIT in Boston, U.S.A. in 1980. From then on, all kinds of conferences on DAI or DAI related topics have been held continually all over the world, which greatly promotes the development and popularization of DAI technologies, and gradually deepens and broadens the research and applications of the science of DAI. With the increase in scale, scope and complexity of new computer based information systems, decision support systems and knowledge based systems, as well as the requirement to encode more complex knowledge in these systems, applications and development of DAI technologies is becoming increasingly important to these systems.

Research of DAI can be generally categorized into two domains: Distributed Problem Solving (DPS) and Multi-Agent System (MAS), both sharing the same research paradigm yet adopting different problem solving means. The goal of DPS is to establish large-granularity cooperative clusters to accomplish the common problem solving objectives. In a pure DPS system, problems are resolved into sub tasks, specific task executors are designed to solve the corresponding sub tasks, and all interaction strategies are incorporated as an integral part of the system. Such systems feature top-down design, since the whole system is established to solve the predefined objectives at the top end.

On the opposite side, a pure MAS system generally comprises pre-existing autonomous and heterogeneous agents without a common objective. Research on MAS involves coordinations and cooperations in knowledge, plan and behavior among groups of autonomous intelligent agents, so that they can jointly take actions or solve problems. Though the agent here is also a task executor, it is “open” to other peer agents, and can deal with both single objective and multiple objectives.

Nowadays, applications of computers are becoming more and more extensive, and problems to be solved are becoming more and more complex, which makes centralized control of the problem solving process and centralized processing of data, information and knowledge more and more difficult. Such distributed and concurrent processing of data and knowledge hails great potentials along with many pending difficulties to the development of AI. The spatial distribution, temporal concurrency and logical dependant relationships of multiple agents make the problem solving more complex in multi-agent systems than in single-agent systems.

Despite such difficulties, research on DAI is feasible, desirable and important for the following reasons:

- (1) Technical foundations — Advances in technologies such as hardware architecture of the processors and communication between the processors make it possible to interconnect great amount of asynchronous processors. Such connection might be tightly coupled systems based on shared or distributed memory, or loosely coupled systems based on local networks, or even very loosely connected systems based on geographically distributed communication networks.
- (2) Distributed problem solving — Many AI applications are distributed in nature. They might be spatially distributed, such as the explanation and

integration of spatially distributed sensors, or the control of robots cooperated in a factory. They might also be functionally distributed, such as the integration of several professional medical diagnosis systems to solve complex cases. They might even be scheduling distributed, for example in a factory, the production line is composed of several working procedures, each scheduled by an expert system.

(3) System integration — DAI systems well support modular design and implementation. A complex system can be resolved into several comparatively simple and task specific sub-modules, in order that the system can be easily constructed, debugged and maintained. It is more flexible to handle errors of decomposed sub-modules than a single integral module. On the other side, great economic and social benefit will be gained if the many existing centralized AI application systems can be used to construct distributed AI systems with minor modifications. For example, it will be extremely time-saving and practically effective if existing independent systems of liver diagnosis system, stomach diagnosis system, intestines diagnose system, etc. can be slightly modified to construct a complex expert system to diagnose digestive tract diseases. The plug-in approach we proposed for agent construction is an effective means to integrate existing AI systems.

(4) New approach to intelligent behavior — Implement intelligent behavior with intelligent agents. To become societies of mind, AI systems should have functions for interaction with the environment, as well as capabilities to cooperate and coordinate with each other.

(5) Meanings in cognitive science — DAI can be used for research and verification of the problems and theories in sociology, psychology, management, etc. Cooperative MAS based on belief, knowledge, hope, intention, promise, attention, object, cooperation, etc. provide effective means to understand and simulate the cognitive problems.

Therefore, no matter technically or socially, the emergence and development of DAI systems is imperative. It is also natural to apply DAI technologies to solve large-scale martial problems. Presently, research in this domain has made certain achievements in China.

MAS is a branch of DAI research. In a multi-agent system, an agent is an autonomous entity which continuously interacts with the environment and co-exists with other peer agents in the same environment. In other words, agent is an entity whose mental states consist of components such as belief, desire and intention. In a multi-agent system, the agents can be either homogeneous or heterogeneous, and the relationships among them can be either cooperative or competitive. A common characteristic of DAI and MAS is distributed behaviors of entities or agents. Multi-agent systems feature bottom-up design, because impractical, the distributed automatic individual agents are defined first, and then problem solving is accomplished with one or more agents. Both single objective and multiple objectives can be achieved. Research on MAS is dedicated to analysis and design of large-scale complex cooperative intelligent systems such as large-scale knowledge and information systems and intelligent robots, based on theories of problem solving through concurrent computing and mutual cooperation among logically or physically distributed multiple agents.

At present, MAS is a very active research direction, which aims at simulation of human rational behaviors for applications in domains such as real world and society simulation, robotics, intelligent machines, etc. An agent is characterized with features of autonomy, interaction with the environment, cooperation, communication, longevity, adaptability, real-time, etc. In order to survive and work in the constantly changing environment of the real world, agents should not only react to emergencies promptly, but also make middle or short term plans based on certain tactics, and then predict the future state through modeling and analysis of the world and other agents, as well as and cooperate or negotiate with other agents using the communication language.

To achieve these features, agent architecture should be studied, because architectures and functions of agents are closely related to each other: improper architecture may greatly limit the functions, while appropriate architecture may well support high level intelligence of agents. We proposed a compound architecture for an agent, which systematically integrates multiple parallel and comparatively independent yet interactional forms of mind, including reaction, planning, modeling, communication, decision making, etc. A Multi-Agent Environment (MAGE) is implemented through the agent kernel based plug-in approach we proposed for agent construction (Shi, 2003). With MAGE and the plug-in approach, compound agents can be conveniently constructed and debugged.

1.9 Artificial Thought Model

Development of computers can be roughly divided into two stages. In the first stage, the Von Neumann architecture is applied for numerical computation, document processing, and database management and query. All these applications have specific algorithms, though somewhat difficult in programming. The second stage focuses on symbolic and logical processing, in which knowledge and information processing mainly bases on reasoning. How to choose effective algorithm is the key problem to this stage of research. All these applications are well defined and explicitly represented problems of the ideal world. However, many real-world problems are ill-structured, such as pattern recognition, problem solving and learning from incomplete information, etc. These problems are in the category of intuitive information processing.

For intuitive information processing, theories and technologies of flexible information processing should be studied. Flexibility in real world has the following characteristics:

- Integrate varieties of complex and intricately related information containing ambiguity or uncertainty information;
- Actively acquire necessary information and knowledge, and learn general knowledge inductively from examples;
- Automatically adapt to users and changing environment;
- Self-organization based on the object for processing;
- Error tolerant information processing.

Actually, human neural networks capable of large-scale parallel and distributed information processing inherently support flexible information processing. Thus, we proposed the artificial thought model in Fig. 1.3.

The artificial thought model in Fig. 1.3 clearly illustrates that artificial thought bases on open autonomous systems, takes fully advantages of varieties of information processing patterns to achieve collective intelligence, then proceeds with flexible information processing, and finally solves real-world problems.

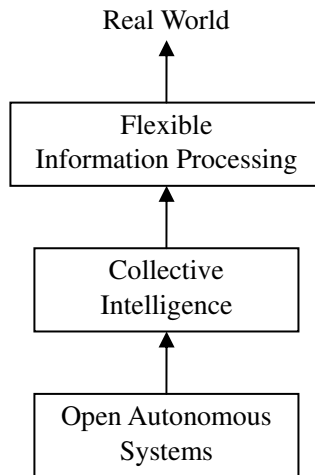


Fig. 1.3. Artificial thought model

1.10 Knowledge Based Systems

An important impetus for AI research is to construct knowledge based systems to automatically solve difficult problems. Ever since the 1980's, knowledge engineering has become the most remarkable characteristic of AI applications. Knowledge based systems (KBS) include expert system, knowledge base system, intelligent decision support system, etc. In 1965, DENDRAL, which was designed to illustrate organic chemistry structures, developed to a series of expert system programs. Such systems mainly include two parts: one is the knowledge base, which represents and stores the set of task-related specific domain knowledge, including not only facts about the related domain, but also heuristic knowledge in expert level; the other is the inference engine, which includes series of inference methodologies to retrieve the reasoning path, and thus to form premises, satisfy objectives, solve problems, etc. As different mechanisms and concepts can be adopted, the inference engines have multiple patterns.

In knowledge based systems knowledge will be stored in the computer in defined structure for knowledge management, problem solving and knowledge sharing. Projects and softwares of "Knowledge Based Management System (KBMS)" have been initiated and developed all over the world, such as in America, in Japan (the NTT Company), as well as in China. Remarkable

characteristic of KBMS is the integration of inference and query, which improves the maintenance of the knowledge base, and provides useful development environment for specific domain knowledge based systems.

Decision Support System (DSS) is evolved from the Management Information System (MIS), with its concept initiated in the early 1970's. It developed fast as an important tool to improve the competitiveness and productivity of companies, as well as to decide on the successfulness of a company. DSS has been adopted by various levels of decision makers in abroad, and attracted great focuses in China. Decision support techniques are critical to support scientific decision making. Early DSS is based on MIS and includes some standard models, such as the operational research model and the econometric model. In 1980, Ralph Sprague proposed a DSS structure based on data base, model base and dialog generation and management software, which has a great impact on later research and applications. In recent years, AI technologies have been gradually applied to DSS, and thus came in to being the intelligent decision support system (IDSS). In 1986, the author proposed the intelligent decision system composed of data base, model base, and knowledge base (Shi, 1988b), which improved the level of scientific management by providing an effective means to solve semi-structured and ill-structured decision problems. Characteristics of IDSS include the application of AI techniques to DSS, and the integration of database and information retrieval techniques with model based qualitative analysis techniques. In the 1990's, we developed the Group DSS (GDSS) based on MAS technologies, which attracted enormous research interests.

Building intelligent systems can imitate, extend and augment human intelligence to achieve certain "machine intelligence", which has great theoretical meanings and practical values. Intelligent systems can be roughly classified into four categories according to the knowledge contained and the paradigms processed: single-domain single-paradigm intelligent system, multi-domain single-paradigm intelligent system, single-domain multi-paradigm intelligent system, and multi-domain multi-paradigm intelligent system.

1. Single-domain single-paradigm intelligent system

Such systems contain knowledge about a single domain, and process only problems of a single paradigm. Examples of such systems include the first and second generation of expert systems, as well as the intelligent control system.

Expert systems apply domain-specific knowledge and reasoning methods to solve complex and specific problems usually settled only by human experts, so that to construct intelligent computer programs with similar problem solving capabilities as experts. They can make explanations about decision making procedure and learn to acquire related problem solving knowledge. The first generation of expert systems (such as DENDRAL, MACSYMA, etc.) had highly professional and specific problem solving capabilities, yet they lacked completeness and portability in architecture, and were weak in problem solving. The second generation of expert systems (such as MYCIN, CASNET, PROSPECTOR, HEARSAY, etc.) was subject-specific professional application system. They were complete in architecture with better portability, and were improved in aspects such as human-machine interface, explanation mechanisms, knowledge acquisition, uncertain reasoning, enhanced expert system knowledge representation, heuristics and generality of reasoning, etc.

2. Multi-domain single-paradigm intelligent system

Such systems contain knowledge about multiple domains, yet only process problems of a certain paradigm. Examples include most distributed problem solving system and multi-expert system. Generally, expert system development tools and environments are used to construct such large-scale synthetical intelligent systems.

Since intelligent systems are widely applied to various domains such as engineering technology, social economics, national defense affairs and ecological environment, several requirements are put forward for intelligent systems. To solve the many real-world problems such as medical diagnosis, economic planning, military commanding, financial projects, crop planting and environment protection, expert knowledge and experience of multiple domains might be involved. Many existing expert systems are single-subject, specific micro expert systems, which might not satisfy the users' practical demands. To construct multi-domain single-paradigm intelligent systems might be an approach to meet the users' requirements in certain degrees. Characteristics of such systems include:

- (1) solve the user's real-world complex problems;
- (2) adopt knowledge and experience of multiple domains, disciplines and professionals for cooperative problem solving;

- (3) based on distributed open software, hardware and network environment;
- (4) constructed with expert system development tools and environments;
- (5) achieve knowledge sharing and knowledge reuse.

3. Single-domain multi-paradigm intelligent system

Such systems contain knowledge of only a single domain, yet process problems of multiple paradigms. Examples include compound intelligent system. Generally, knowledge can be acquired through neural network training, and then transformed into production rules to be used in problem solving by inference engines.

Multiple mechanisms can be used to process a single problem in problem solving. Take an illness diagnosis system as an example, both symbolic reasoning and artificial neural networks can be used. Then, compare and integrate the results of different methods processing the same problem, through which correct results might be obtained and unilateralism can be avoided.

4. Multi-domain multi-paradigm intelligent system

Fig. 1.4 illustrates the sketch map of such systems, which contain knowledge of multiple domains and process problems of different paradigms. Collective intelligence in the figure means that when processing multiple paradigms, different processing mechanisms work separately, accomplish different duties, and cooperate with each other, so that to represent collective intelligent behaviors.

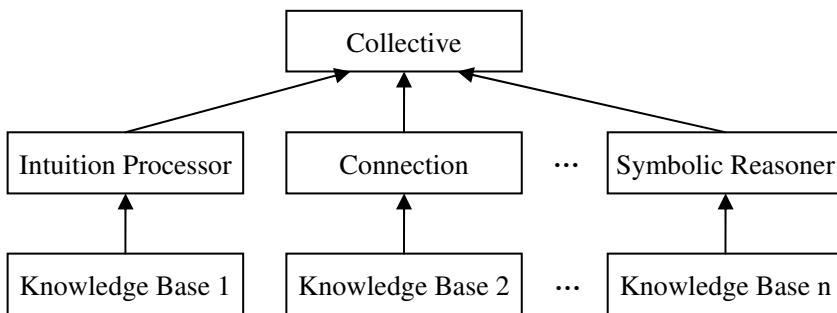


Fig. 1.4. Multi-domain multi-paradigm intelligent system

Synthetical DSS and KBS belong to this category of intelligent systems. In such systems, reasoning based abstract thought is based on symbolic processing,

while imagery thought such as pattern recognition and image processing applies neural network computing.

Most intelligence problems are ill-structured and continuously changing, thus they are difficult to solve with a single specific algorithm. A plausible approach to solve such intelligence problems is to construct human-machine united open systems which interact with the environment. An open system is one which may always run into unexpected results during the system processing, and can receive external new information at any time.

Based on summarization and analysis of the design methods and implementation technologies of existing KBS, intelligent agent technologies are studied to construct large scale synthetical KBS with functionalities of multiple knowledge representation, synthetical knowledge base, self-organization and cooperation, automatic knowledge acquisition, continually improved intelligent behaviors, etc. Such systems are the primary means to construct multi-domain multi-paradigm intelligent system.

Exercises

1. What is Artificial Intelligence (AI)? What is the research objective of AI?
2. Please briefly introduce the main stages of development in the history of AI.
3. What are the five fundamental questions for AI research?
4. What is the physical symbol system? What is the physical symbol system assumption?
5. What is symbolic intelligence? What is computational intelligence?
6. Please describe the simple model of machine learning and its basic elements.
7. What is Distributed Artificial Intelligence (DAI)? What are the main research domains of DAI?
8. Please refer to relevant literature and discuss whether the following tasks can be solved by current computers:
 - a) Defeat an international grandmaster in the world's chess competition;
 - b) Defeat a 9 dan professional in a game of Go;
 - c) Discover and prove a new mathematical theorem;
 - d) Find bugs in the programs automatically.
9. How to classify knowledge based systems? How to achieve collective intelligence behaviors?