1

Development of Artificial Intelligence

1.1 The Concept and Evolution of Artificial Intelligence

1.1.1 The Concept of Artificial Intelligence

Artificial intelligence (AI), also known as machine intelligence, refers to the intelligence represented by machines made by people. Generally, AI refers to human intelligence technology realized by means of various ordinary computer programs. The definition in the general textbook is "the research and design of intelligent agent," which refers to a system that can observe the surrounding environment and make actions to achieve goals [1, 2].

The definition of AI can be divided into two parts, namely, "artificial" and "intelligence." "Artificial" is easier to understand and less controversial. Sometimes we consider what humans can do and create, or whether a person's own level of intelligence is high enough to create AI, and so on. But to sum up an "artificial system" is an artificial system in the general sense. There are many questions about what "intelligence" is. This involves other issues such as consciousness, self and mind, including the unconscious mind. The only intelligence that people know is their own intelligence, which is a widely accepted view. But our understanding of our own intelligence is very limited, and our understanding of the necessary elements of human intelligence is also very limited, so it is difficult to define what "intelligence" is made by "artificial." Therefore, the research on AI often involves the research on human intelligence itself. Other intelligence with animals or other artificial systems is also generally considered as a research topic related to AI.

A popular definition of AI, as well as an earlier definition in this field, was put forward by John McCarthy of the Massachusetts Institute of Technology at the Dartmouth Conference in 1956: AI is to make the behavior of machines look like

Intelligent Satellite Design and Implementation, First Edition. Jianjun Zhang and Jing Li.

that of human beings. But this definition seems to ignore the possibility of strong AI. Another definition is that AI is the intelligence represented by artificial machines. In general, the current definition of AI can be divided into four categories, namely, machines "think like people," "move like people," "think rationally," and "act rationally." Here, "action" should be broadly understood as the decision to take action or specify action, rather than physical action.

Strong AI believes that it is possible to produce intelligent machines that can really reason and solve problems, and such machines will be considered as perceptual and self-conscious. There are two types of strong AI:

1) Human-like AI, that is, the thinking and reasoning of machines, is like human thinking.
2) Nonhuman AI, that is, machines produce perception and consciousness completely different from human beings and use reasoning methods completely different from human beings.

The term "strong artificial intelligence" was originally created by John Rogers Hiller for computers and other information-processing machines. Its definition is: strong AI holds that computers are not only a tool for studying human thinking. On the contrary, as long as it runs properly, the computer itself is thinking. The debate on strong AI is different from the debate on monism and dualism in a broader sense. The main point of the argument is: if the only working principle of a machine is to convert encoded data, then is the machine thinking? Hiller thought it was impossible. He gave an example of a Chinese room to illustrate that if the machine only converts data, and the data itself is a coding representation of some things, then without understanding the correspondence between this coding and the actual things, the machine cannot have any understanding of the data it processes. Based on this argument, Hiller believes that even if a machine passes the Turing test, it does not necessarily mean it is really thinking and conscious like a person. There are also philosophers who hold different views. Daniel Dennett believes in his book *Consciousness Explained* that man is just a machine with a soul. Why do we think: "Man can have intelligence, but ordinary machines can't?" He believes that it is possible to have thinking and consciousness when data is transferred to machines like the above.

The weak AI point of view believes that it is impossible to produce intelligent machines that can really reason and solve problems. These machines just look intelligent, but they do not really have intelligence, nor do they have autonomous consciousness. Weak AI came into being when compared with strong AI because the research on AI was at a standstill for a time, and it began to change and go far ahead until the artificial neural network (ANN) had a strong computing ability to simulate. In terms of the current research field of AI, researchers have created a large number of machines that look like intelligence, and obtained quite fruitful theoretical and substantive results, such as the Eureqa computer program developed by Cornell University Professor Hod Lipson and his doctoral student Michael Schmidt in 2009, as long as some data are given, This computer program can deduce the Newtonian mechanics formula that Newton spent years of research to discover in only a few hours, which is equivalent to rediscovering the Newtonian mechanics formula in only a few hours. This computer program can also be used to study many scientific problems in other fields. These so-called weak AI have made great progress under the development of neural networks, but there is no clear conclusion on how to integrate them into strong AI.

Some philosophers believe that if weak AI is realizable, then strong AI is also realizable. For example, Simon Blackburn said in his introductory philosophy textbook Think that a person's seemingly "intelligent" action does not really mean that the person is really intelligent. I can never know whether another person is really intelligent like me or whether they just look intelligent. Based on this argument, since weak AI believes that it can make the machine look intelligent, it cannot completely deny that the machine is really intelligent. Blackburn believes that this is a subjective issue. It should be pointed out that weak AI is not completely opposite to strong AI, that is, even if strong AI is possible, weak AI is still meaningful. At least, the things that computers can do today, such as arithmetic operations, were considered to be in great need of intelligence more than 100 years ago. And even if strong AI is proven to be possible, it does not mean that strong AI will be developed.

In a word, AI is an interdisciplinary subject, belonging to the intersection of natural science and social science. The disciplines involved include physics, philosophy and cognitive science, logic, mathematics and statistics, psychology, computer science, cybernetics, determinism, uncertainty principle, sociology, criminology, and intelligent crime. The research on AI is highly technical and professional, and each branch field is deep and different, so it covers a wide range. The core issues of AI include the ability to construct reasoning, knowledge, planning, learning, communication, perception, movement, and operation of objects that are similar to or even beyond human beings. Strong AI is still the long-term goal in this field.

There are different understandings about the definition of AI from different perspectives, and there is no unified definition at present. Here are some definitions with high acceptance:

1) In 1956, scientists such as John McCarthy of Stanford University, Marvin Minsky of Massachusetts Institute of Technology, Herbert Simon and Allen Newell of Carnegie Mellon University, Claude Shannon of Bell Laboratory, and other scientists first established the concept of "artificial intelligence" in Dartmouth College of the United States, that is to say, let machines recognize, think, and learn like people, that is, use computers to simulate human learning

and other aspects of intelligence. At the same time, seven typical task directions have been established: machine theorem proving, machine translation, expert system, game, pattern recognition, learning, robot, and intelligent control.

2) In 1981, A. Barr and E.A. Feigenbaum proposed from the perspective of computer science: “Artificial intelligence is a branch of computer science. It is concerned with the design of intelligent computer systems, which have intelligent characteristics associated with human behavior, such as understanding language, learning, reasoning, problem solving, etc.”
3) In 1983, Elaine Rich proposed that “Artificial intelligence is to study how to use computers to simulate human brain to engage in reasoning, planning, design, thinking, learning and other thinking activities, and to solve complex problems that are still considered to need to be handled by experts.”
4) In 1987, Michael R. Genesereth and Nils J. Nilsson pointed out that “AI is the science of studying intelligent behavior, and its ultimate purpose is to establish a theory on the behavior of natural intelligence entities and guide the creation of artificial products with intelligent behavior.”
5) Wikipedia defines “AI is the intelligence displayed by machines,” that is, as long as a certain machine has some or some “intelligence” characteristics or performance, it should be counted as “AI.” The Encyclopedia Britannica defines “AI is the ability of a digital computer or a robot controlled by a digital computer to perform some tasks that intelligent organisms only have.” Baidu Encyclopedia defines AI as “a new technological science that studies and develops theories, methods, technologies, and application systems for simulating, extending and extending human intelligence,” regards it as a branch of computer science, and points out that its research includes robots, language recognition, image recognition, natural language processing, and expert systems.
6) According to the White Paper on the Standardization of Artificial Intelligence (2018), “Artificial intelligence is a theory, method, technology and application system that uses digital computers or machines controlled by digital computers to simulate, extend and expand human intelligence, perceive the environment, acquire knowledge and use knowledge to obtain the best results. Artificial intelligence is the engineering of knowledge, which is the process that machines imitate human beings to use knowledge to complete certain behaviors.”
7) The US Defense Authorization Act of FY 2019 defines AI as all systems that can “act rationally” in a changing and unpredictable environment without sufficient human supervision or can learn from experience and use data to improve performance.

To sum up, AI allows computer/machine to simulate, extend, and expand human intelligence, so that the system can cope with changing and unpredictable environments without human supervision, deal with complex problems through learning and experience, and achieve performance improvement [2, 3].

1.1.2 Evolution of Artificial Intelligence

AI was born in the 1950s. Over the past 60 years, it has experienced three waves (Figure 1.1) and has formed such theoretical schools as semiotics, connectionism, and behaviorism. The first wave (1956–1976) was dominated by semiotics, and its main achievements were machine reasoning, expert system, and knowledge engineering; the second wave (1976–2006), dominated by the connectionist school, produced neural networks, machine learning, and other achievements; the third wave (from 2006 to now), led by the connecting school, achieved great success in deep neural network and deep reinforcement learning, and AI entered the stage of commercial development. The behavioral school theory has not played a leading role in the third wave so far. Its core is adaptive control, learning control, evolutionary computing, and distributed intelligence, which are the basis of modern control theory [3].

In the three waves of AI development, the key points in technology and application breakthrough are shown in Figure 1.2.

Notably, the behavioral school theory has not played a leading role in the third wave so far. Its core is adaptive control, learning control, evolutionary computing, and distributed intelligence, which is the basis of modern control theory. In recent years, different voices have emerged in the development of AI. They believe that AI technology based on the connectionist school theory has a huge gap in solving problems that cannot be reasoned and counted and cannot achieve strong AI with creative thinking. They believe that the symbolic school and behavioral school theory pay more attention to the abstraction of human higher intelligence and

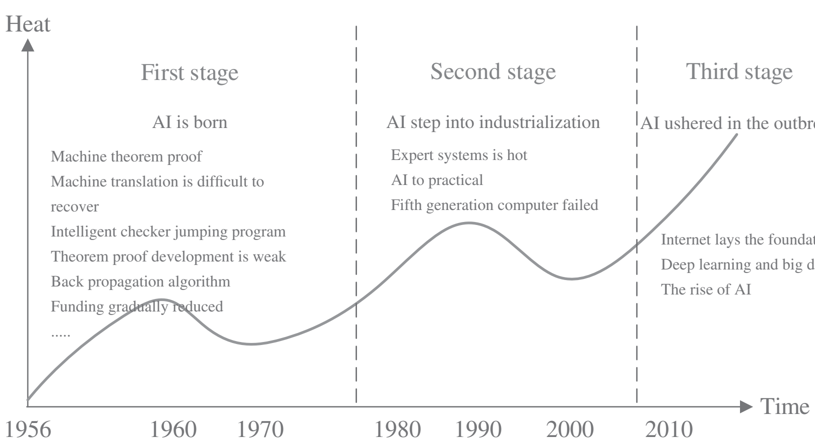

Figure 1.1 Three waves of AI.

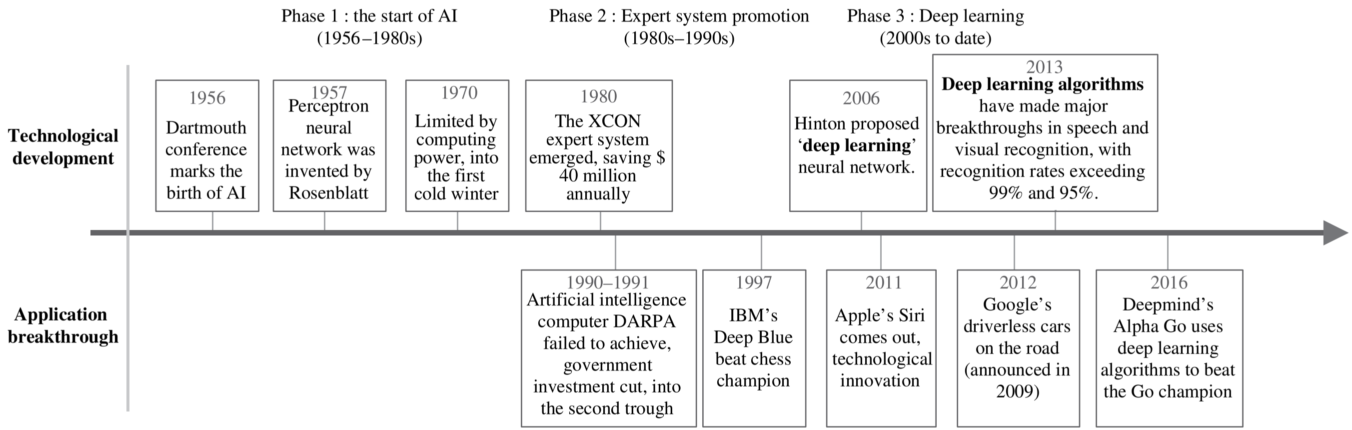

Figure 1.2 Key points of AI technology development.

more attention to adaptation and evolution; the realization of strong AI in the future may need to find another way.

1.2 The Current Scope and Technical Framework of Artificial Intelligence

1.2.1 Technical Scope

Compared with the early AI, the new generation AI is under the guidance of the new information environment, massive database, and continuously evolving and enriching strategic objectives, relying on the two basic platforms of cloud computing and big data and the three general technologies of machine learning, pattern recognition, and human–computer interaction and taking the new computing architecture, general AI, and open source ecosystem as the main guidance, and continues to build and improve the technical framework system. As a result of multi-disciplinary intersections and universal technology, AI technology has formed a complex network of technology systems together with related downstream technologies and applications. At present, the network is in its infancy, but it is still in a dynamic state of rapid renewal and drastic changes [4].

1.2.2 Technical Framework

According to the description in the White Paper on the Development of New Generation AI (2017), the current AI technical framework is mainly composed of three parts: basic layer, technical layer, and application layer, as shown in Figure 1.3:

1) Basic layer
   The basic layer mainly includes big data, smart sensors, smart chips, and algorithm models. Among them, smart sensors and smart chips belong to basic hardware, and algorithm models belong to core software [5].
2) Technical level
   The technical level mainly includes pattern recognition, autonomous planning, intelligent decision-making, autonomous control, human–machine cooperation, group intelligence, etc.
3) Application layer
   The application layer mainly includes intelligent agents (robots, unmanned driving, intelligent search, and unmanned aerial vehicles) and industrial applications (finance, medical, security, education, human settlements, etc.).

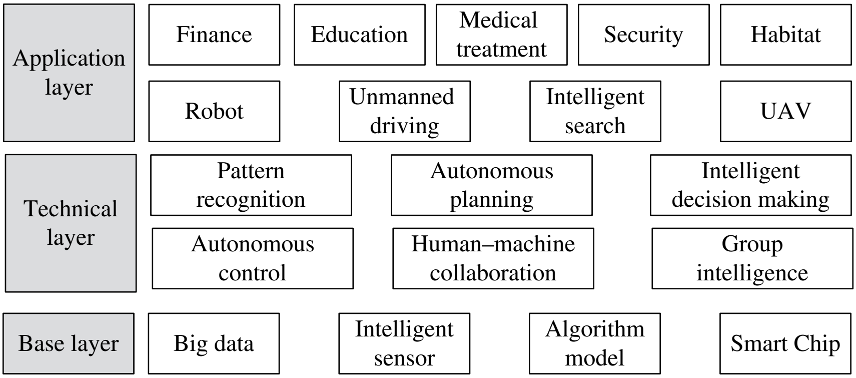

Figure 1.3 General technical system of AI technology.

1.2.3 Technical Features

At present, the development of AI is based on the theory of connectivity school. The main driving factors are explosive growth of data, continuous improvement of computing power, continuous optimization of algorithm models, and deep coupling of capital and technology. The main technologies include computer vision, machine learning, natural language processing, robot technology, speech recognition, etc. At present, intelligent medical, finance, education, transportation, security, home manufacturing, and other fields have been widely used and developed rapidly in unmanned driving, intelligent robots, and others. Among them, deep learning is a typical representative of the development of the new generation of AI, but "deep learning" is not a synonym for AI [6, 7].

Driven by data, computing power, algorithm model, and multiple applications, AI is evolving from auxiliary equipment and tools to assist and partner for collaborative interaction and is more closely integrated into human production and life (Figure 1.4):

1) Big data has become the cornerstone of the sustained and rapid development of AI.
2) Text, image, voice, and other information realize cross-media interaction.
3) Network-based swarm intelligence technology has begun to sprout.
4) Autonomous intelligent system has become an emerging development direction.
5) Human–machine collaboration is giving birth to a new type of hybrid intelligence.

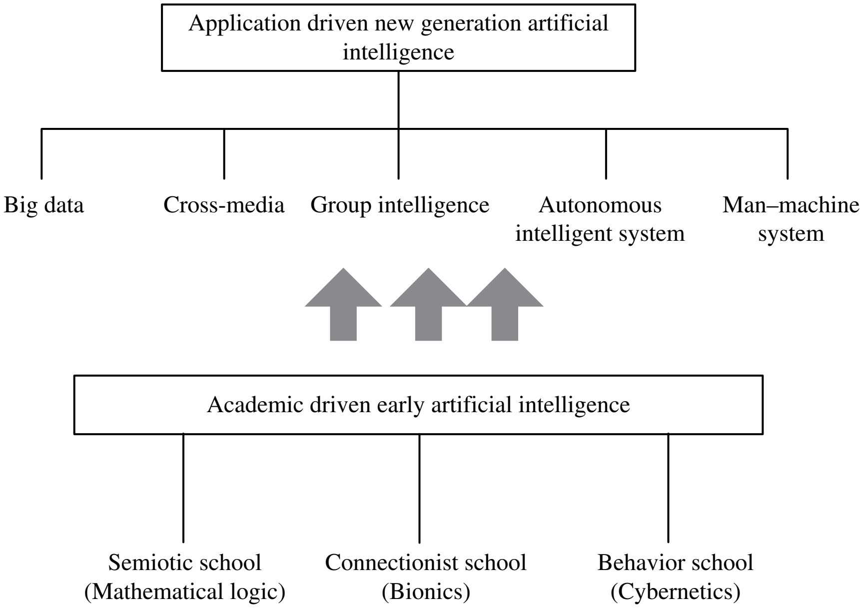

Figure 1.4 Technical characteristics of new generation AI technology.

1.3 The Overall Development Trend of Artificial Intelligence

1.3.1 Current Development Trend

1) **Artificial intelligence technology, as an important factor leading industry disruption and technological change, has achieved explosive growth in all fields around the world.**
   In the past decade, intelligent algorithms represented by deep learning and intensive learning have promoted AI to continue to exert its power in cutting-edge technologies and applications, and giant companies such as Google and IBM have made major breakthroughs in search engines, human–computer games, intelligent medical care, unmanned driving, and other fields. By December 2018, more than 5000 AI startups had been born and flourished in reading and writing assistant, financial industry, enterprise management, advanced manufacturing, and other industries, which had a profound impact on various fields [8–10].
2) **The major developed countries in the world regard the development of AI as a major strategy to enhance national competitiveness and maintain national security.**

The United States was the first country to release a national strategy for AI. In 2016, it released the National Strategic Plan for AI Research and Development. In the next two years, a total of 15 countries and regions released their AI strategies and accelerated their planning and layout in the field of AI (Figure 1.5). Among them, thanks to decades of federal research funds, industrial production, academic research, and the inflow of foreign talents, the United States has been leading the global wave of AI development in basic theory, software and hardware, talents, enterprises, and other aspects.

3) **In many aspects, AI is far from mature application, and there are still technical bottlenecks.**
   At present, it is generally believed that AI has surpassed human beings in large-scale image recognition. Although some progress has been made in machine translation, it is not close to the ideal level, while there is a big gap in chat conversation. In terms of driverless driving, the current commercial autonomous vehicle is auxiliary driving, and the real autonomous vehicle is still in the development and testing stage, in which dealing with emergency and abnormal traffic conditions is the difficulty encountered by AI. Similarly, in speech recognition, although AI recognition is close to humans in the experimental environment, in reality, especially in the presence of environmental interference, the recognition rate of AI is actually unsatisfactory, and it often makes some common sense mistakes that humans cannot make. Some experts even believe that AI technology will not make a major breakthrough in 20 years, because there are not many new topics in the field of AI research done by the current academic community. Even if scientists work hard, it will take about 20 years to accumulate a theoretical basis that makes people feel very excited and surprised.
4) **AI will accelerate the intelligent transformation and upgrading of traditional industries.**
   In the future, with the continuous improvement of AI capabilities, the integration of AI and traditional industries will continue to deepen, driving the transformation and upgrading of traditional industries to intelligence and at the same time creating huge benefits for society. According to the research report, by 2025, the market scale of AI in agricultural applications will reach 20 billion dollars, and the financial industry will save and generate 34–43 billion dollars in costs and new market opportunities, and the annual expenditure in the medical field will be reduced by 54 billion dollars.
5) **Human–computer cooperation will become the main direction of AI commercialization.**
   With the complementary nature between humans and AI systems, the collaborative interaction between humans and AI systems will become the main direction of AI commercialization. Although fully autonomous AI systems

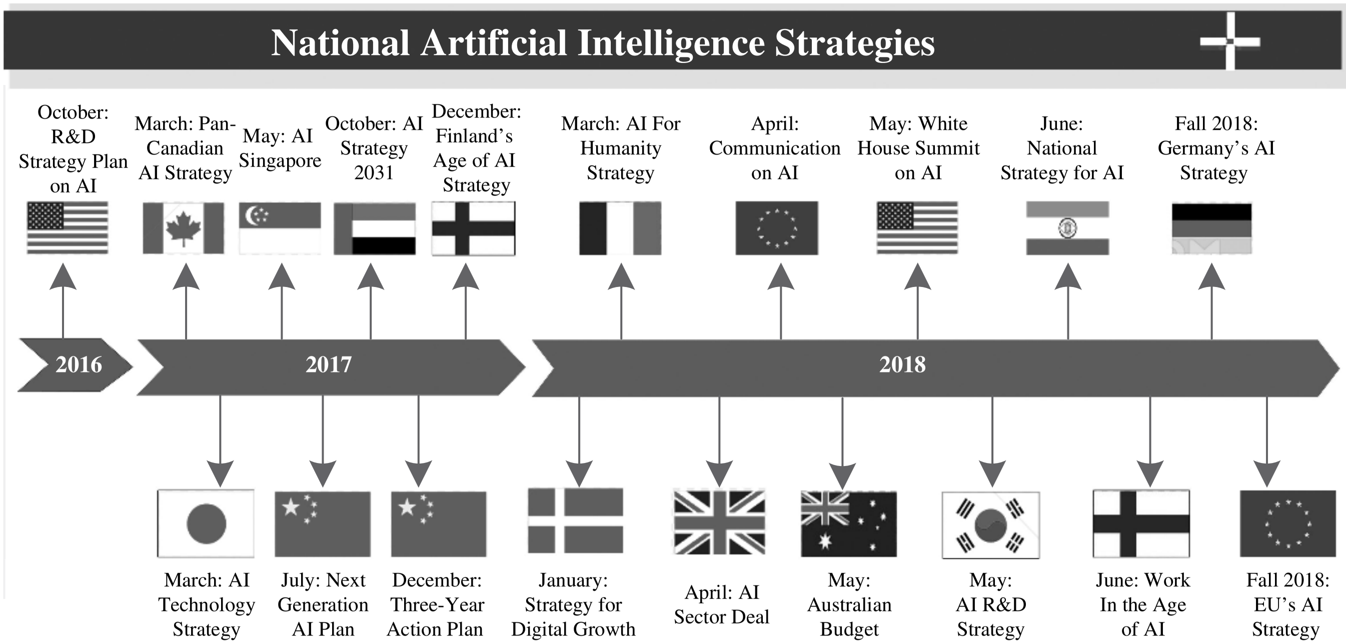

Figure 1.5 National AI strategies.

will play an important role in underwater or deep space exploration and other applications, AI systems cannot completely replace humans in the short term in disaster recovery, medical diagnosis, and other applications.

Human–machine cooperation can be divided into three ways, namely, joint execution, auxiliary execution, and alternative execution:

- Joint execution

  AI system positioning: perform peripheral tasks supporting human decision makers.

  Typical applications: short-term or long-term memory retrieval and prediction tasks.
- Auxiliary execution

  AI system positioning: when human beings need help, the artificial intelligence system performs complex monitoring functions.

  Typical applications: ground proximity alarm system, decision-making, and automatic medical diagnosis in aircraft.
- Alternative execution

  AI system positioning: AI systems perform tasks with very limited capabilities for humans.

  Typical applications: complex mathematical operations, dynamic system control and guidance in controversial operating environment, automatic system control in hazardous or toxic environment, nuclear reactor control room, and other rapid response systems.

6) **AI will present a competitive pattern of leading platform plus scenario application.**

Under the trend of AI platformization, AI will present a competitive pattern of several leading platforms and extensive scene applications in the future, and ecological builders will become the most important model among them. As shown in Figure 1.6, the future AI competition pattern will mainly present five modes.

Mode 1: Ecological construction – take the whole industrial chain ecology and scenario application as a breakthrough. Take Internet companies as the main body, mainly invest in infrastructure and technology for a long time. At the same time, scenario applications will be used as traffic entry, accumulate applications, and become the leading application platform. It will become the builder of AI ecosystem.

Mode 2: Technical algorithm-driven – take technical layer and scene application as a breakthrough. Take software companies as the main body, deeply cultivate algorithm platform and general technology platform, and gradually establish application platform with scene application as the traffic entrance.

Mode 3: Application focus – scene application. Based on the scenario or industry data, many segmented scenario applications are developed mainly for startups and traditional industry companies.

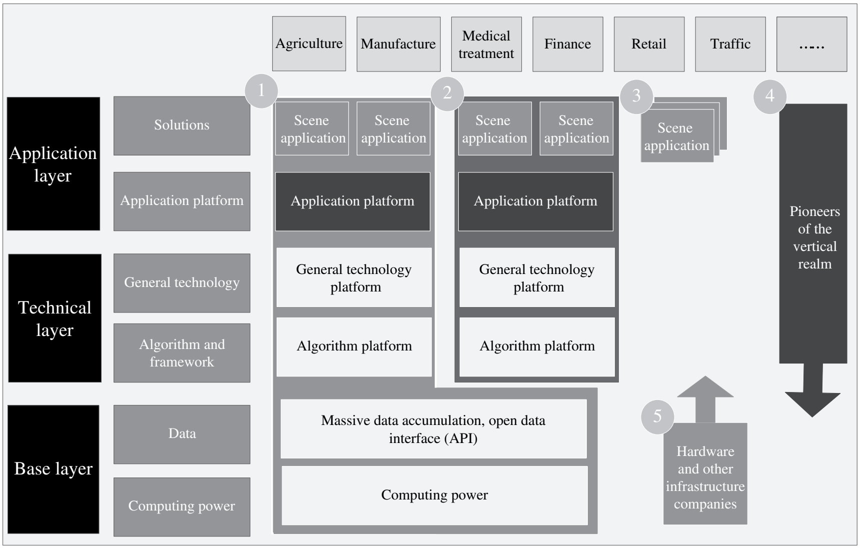

Figure 1.6 Competition pattern of AI in the future.

Mode 4: Vertical field first – killer application and gradually builds vertical field ecology. It is mainly a pioneer in the vertical field. It relies on killer applications to accumulate a large number of users and data in the vertical field, and deeply cultivates the common technologies and algorithms in the field, becoming a subverter in the vertical field.

Mode 5: Infrastructure provision – starting from infrastructure and expanding to the downstream of the industrial chain. Take chip or hardware and other infrastructure companies as the main part, start from infrastructure, improve technical capabilities, and expand to the upstream of data, algorithms, and other industrial chains.

1.3.2 Technical Development Trend

At present, AI, represented by deep learning and big data, has made amazing progress in the fields of image classification, speech recognition, visual understanding, and machine translation. However, deep learning relies on annotated data and lacks the ability to express logical reasoning and causal relationships. It is difficult to deal with tasks with complex spatiotemporal correlations and cannot achieve the goal of strong AI. In the future, AI technology needs to continue to break through [10–12].

1) *From deep learning to brain intelligence algorithm*
   In the aspect of algorithms, the basis of AI technology is developed according to two main lines: deep learning improvement and new algorithms. Advanced in-depth learning focuses on breaking through the methods of adaptive learning, autonomous learning, zero-data learning, unsupervised learning, migration learning, etc., to achieve AI with high reliability, high interpretability, and strong generalization ability. Second, the academic community continues to explore new algorithms, and brain-like intelligent algorithms have become a frontier hotspot, focusing on breaking through brain-like information coding, processing, memory, learning, association, reasoning, and other theories, forming brain-like complex systems and brain-like control and other theories and methods, and establishing new models of large-scale brain-like intelligent computing and brain-inspired cognitive computing models [12, 13].
2) *From dedicated intelligence to general intelligence*
   With the continuous development of science and technology and the deep transformation of social structure, the problems facing human beings are highly complex. There is an urgent need for a wide range, high integration, and strong adaptability of general wisdom in professional fields such as game, identification, control, and prediction, which significantly improves the ability of human beings to read, manage, and reorganize knowledge. General AI has the characteristics of reducing dependence on domain knowledge, improving the applicability of processing tasks, and achieving the correction of machine-independent cognition. It has the ability to process multiple types of tasks and adapt to unexpected situations. Its substantive progress will truly start the prelude of intelligent revolution, highly integrate with the existing physical and information world, and profoundly affect all aspects of social and economic development.
3) *From machine intelligence to human–machine hybrid intelligence*
   Machine intelligence (AI) and human intelligence have their respective strengths and need to learn from each other and complement each other. Integrating multiple intelligence models, human–machine coexistence will become the new normal of the future society. It will focus on breaking through the theory of human–machine win–win situation understanding and decision learning, intuitive reasoning and causal model, memory and knowledge evolution, and realize hybrid enhanced intelligence that learning and thinking are close to or exceed the level of human intelligence.
4) *From single agent intelligence to group intelligence*
   Group intelligence originates from the research on the group behavior (including information transmission, collective decision-making, etc.) of social insects represented by ant colonies, bee colonies, etc. The group has self-organization, division of labor, and coordination. Group intelligence refers to the collective

intelligent behavior of a system composed of individual units through interaction between each other or with the environment, with the characteristics of intelligent emergence. Science believes that the era of group intelligence based on networks is coming. Group intelligence will focus on breaking through the theories and methods of organization, emergence and learning of group intelligence, establishing expressive and computable group intelligence incentive algorithms and models, and forming group intelligence capabilities based on a network environment [14, 15].

5) *Weak AI will gradually transform into strong AI*
   AI can be divided into three stages: weak AI, strong AI, and super AI. The level of AI in the three stages has been continuously improved. Weak AI refers to human ability in some fields; strong AI refers to having human capabilities in all fields, being able to compete with human beings in all aspects, and being unable to simply distinguish between human beings and machines; super AI means that it can surpass human beings in all fields, and can surpass human beings in the fields of innovation and creative creation to solve any problems that cannot be solved by human beings.

   From the current development level of AI, AI is still a weak AI based on specific application fields, such as image recognition, speech recognition, and other biometric analysis, such as intelligent search, intelligent recommendation, intelligent sorting, and other intelligent algorithms. When it comes to vertical industries, AI mostly assists human beings in their work in the role of assistant, such as the current intelligent investment advisor and autonomous vehicle, but the AI that completely gets rid of human beings and can reach or even surpass human beings in the true sense cannot be realized. In the future, with the substantial growth of computing power and data volume and the improvement of algorithms, weak AI will gradually transform into strong AI, and machine intelligence will advance from perception, memory, and storage to cognition, autonomous learning, decision-making, and execution.

6) *Data, algorithm, and computing power are the three carriages of AI*
   In recent years, with the increase of the data volume index, the algorithm theory continues to be updated iteratively, and the computing ability continues to be enhanced, which jointly promotes the accelerated development of AI. The relationship between the three and AI is shown in Figure 1.7.

With the rapid development of the Internet, not only the standard training set and test set data are increasing but also the massive and high-quality application scenario data are becoming increasingly rich. Real-time, massive, multi-sources, and multi-types of data can describe reality more closely from different angles, and machine learning algorithms can be used to mine multi-level associations between data, laying the foundation of data sources for AI applications.

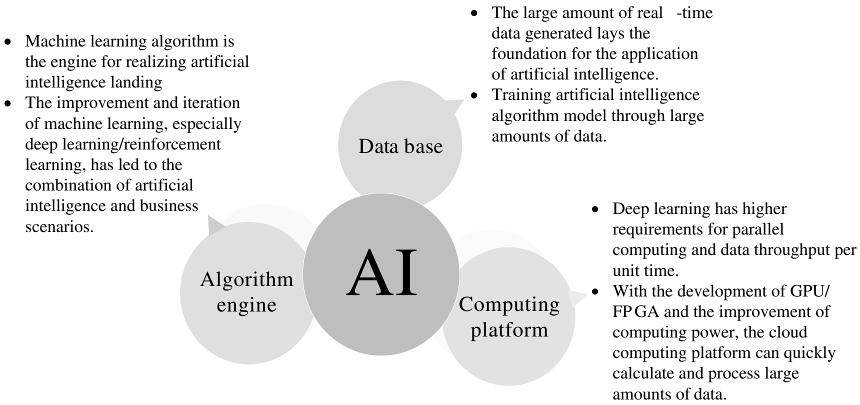

Figure 1.7 Relationship between data, algorithm, calculation, and artificial intelligence.

The development of algorithms, especially the paper published by Professor Geoffrey Hinton in 2006, has started the wave of deep learning in academia and industry. The deep learning algorithm represented by ANN has become the core engine of AI applications.

AI has a high demand for computing power. With the rapid development of cloud computing technology and chip processing power, thousands of machines can be used for parallel computing. In particular, the development of tensor processing unit (TPU), graphics processing unit (GPU), field programmable gate array (FPGA), and special chips for AI has laid the foundation for implementing AI computing power, making the application of AI similar to human's deep neural network algorithm model a reality.

1.4 The Main Achievements of AI

1.4.1 Image Recognition

Image recognition is more troublesome than speech recognition because the object of speech recognition is always a variety of limited languages. However, the recognition of the face and cat in the image cannot be handled by a single model when it comes to the specific implementation. The current state is that if you select a point, such as face recognition, and hit dozens of PhDs and hundreds of GPUs, it will take one to two years. If you can still find the landing point and continue to obtain data, you can achieve very high accuracy (more than 99%). But this accuracy cannot cover other fields at once. You can only do it point by point. Because there are practical landing scenarios (banks, etc.) in such fields as face, they have developed rapidly. It will take some time for other fields to reach the same level.

In 2009, a Chinese scholar in the Department of Computer Science at Princeton University (the first author is Jia Deng) published the paper "ImageNet: A large scale hierarchical image database," announcing the establishment of the first super-large image database for use by computer vision researchers. In 2010, ImageNet Large Scale Visual Recognition Challenge 2010 (ILSVRC2010), a large-scale image recognition contest based on ImageNet, was held for the first time. The initial rule of the competition is to take 1.2 million images in the database as training samples, which belong to more than 1000 different categories and are manually marked. The trained program will be used to evaluate 50 000 test images to see whether it is accurate in image classification.

In 2012, Professor Hinton and his two graduate students, Alex Krizhevsky and Illya Sutkerver, applied the latest technology of deep learning to the problem of ImageNet. Their model is an eight-layer convolutional neural network with 650 000 neurons and 60 million free parameters. This neural network uses the discard algorithm and the excitation function of the rectified linear unit (RELU) introduced in the previous two articles. Professor Hinton's team used two GPUs to train the program with 1.2 million images, which took nearly six days. The trained model faces 150 000 test images, and the error rate of the first five categories predicted is only 15.3%. In the 2012 ImageNet competition with 30 groups, the test result ranked first. The second is from the model of the Japanese team, and the corresponding error rate is as high as 26.2%. This marks that neural network has greatly surpassed other technologies in the field of image recognition and has become a turning point in the breakthrough of AI technology.

In the ImageNet image recognition competition in December 2015, the team from Microsoft Research Asia (MSRA) won the championship. As the depth of the network increases, the efficiency of learning decreases. In order to solve the problem of attenuation of effective information in layer-by-layer transmission, the MSRA team tried an algorithm called "deep residual learning." MSRA hopes that in the deep residual learning model, using the 152 layers of neural network, the recognition error rate of the first five categories has reached a new low of 3.57%, which is lower than about 5% of the error rate of a normal person.

1.4.2 Speech Recognition

Speech recognition is a field that has been conquered by deep learning in recent years. If enough money is spent, the recognition accuracy can reach 99%. In speech recognition, improving the accuracy of the last few points is likely to be more difficult than the previous 90% accuracy. However, the accuracy of the last few points often formally crosses the key of being able to use and not being able to use. Before in-depth learning, people have tried to conquer speech recognition for many years. Generally speaking, the first speech recognition system that can recognize the pronunciation of 10 English numbers, studied by Bell Labs in 1952, is considered as the starting point of speech recognition. In this way, human beings have worked hard on this matter for more than 60 years. In the 1970s, speech recognition with a small vocabulary was completed, and in the 1980s, speech recognition with a large vocabulary was completed. Then the accuracy was stuck there, hovering around 85%, and one card was close to 30 years. Microsoft and IBM have tried to apply this technology in the past, but obviously, it has no effect. Many people do not even remember what they did. After the introduction of deep learning into speech recognition, things have changed fundamentally. Now, as long as there is enough data for training, most companies can train enough accurate speech recognition models themselves. This technology basically needs to be commercialized, and it can be more and more regarded as a technology that can be solved without too much investment.

RNN, also known as recurrent neural network or multilayer feedback neural network, is another very important type of neural network. In essence, the difference between RNN and feedforward network is that it can retain a memory state to process the input of a sequence, which is particularly important for handwritten character recognition, speech recognition, and natural language processing.

In October 2012, Geoffrev Hinton, Deng Li, and several other researchers representing four different institutions (University of Toronto, Microsoft, Google, and IBM) jointly published the paper "Application of Deep Neural Network in Acoustic Model of Speech Recognition: Common Viewpoint of Four Research Groups." The researchers borrowed the algorithm of "restricted Boltzmann machine" used by Hinton to "pre-train" the neural network. The depth neural network model is used to estimate the probability of recognizing characters. In a voice input benchmark test conducted by Google, the word error rate reached 12.3%.

In March 2013, Alex Graves of the University of Toronto led the publication of the paper "Deep Circulation Neural Network for Speech Recognition." The paper uses RNN/long short-term memory (LSTM) technology – a network with three hidden layers and 4.3 million free parameters. In a benchmark test called TIMIT, the "phoneme error rate" reached 17.7%, which is superior to the performance level of all other technologies in the same period.

In May 2015, Google announced that by relying on RNN/LSTM-related technologies, the word error rate of Google Voice decreased to 8% (about 4% for normal people).

In December 2015, Dario Amodei of Baidu AI Lab led the paper "End-to-End Speech Recognition in English and Chinese." The model in this paper uses a simplified variant of LSTM, called "closed cycle unit." Baidu's English speech recognition system has received nearly 12 000 hours of speech training. It takes three to five days to complete the training on 16 GPUs. In a benchmark test called WSJ Eval ’92, the word error rate was as low as 3.1%, which exceeded the recognition ability of normal people (5%). In another small Chinese test, the recognition error rate of the machine was only 3.7%, while the collective recognition error rate of a five-member group was 4%.

According to this trend, the accuracy of the machine in various benchmark tests of speech recognition will soon catch up with and surpass that of ordinary people. This is another difficulty that AI will overcome after image recognition.

The essence of RNN is that it can handle the output and input (many to many) of a sequence with varying length. In a broad sense, if the traditional feedforward neural network does the optimization of a function (such as image recognition), then what the circular neural network does is the optimization of a program, and the application space is much wider.

Semantic understanding and speech recognition are different from image recognition. Semantic understanding is in a state that is basically unsettled. Many times, when we watch the demonstration, we can see a robot and an intelligent product communicate with people smoothly. There are two possibilities to achieve this. One is cheating, with individuals behind it, which belongs to AI. One is that the conversation is limited to specific scenarios, such as making a phone call in the car or allowing map navigation. The difficulty of semantic understanding is related to the number of concepts to be processed. When the number of concepts to be processed is less than a few thousand, it is still possible to solve the problem in a rule-based way for a specific scenario, which is relatively smooth. But once the scope is extended to the whole social life, it is at most like Google Now and Siri. One of the applications closely related to this is the ability of various intelligent voice assistants in conversation, and the other is translation.

1.4.3 Art Creation

For a long time, people tend to think that machines could understand humans’ logical thinking, but they cannot understand humans’ rich feelings, let alone humans’ aesthetic values. Of course, machines cannot produce works with aesthetic value. Facts speak louder than eloquence. The Alpha Dog took a ground-breaking step against Li Shishi. Mr. Nie Weiping, the chess master, saluted the Alpha Dog’s dismounting, which shows that the deep learning algorithm has been able to create aesthetic value. Many chess players spend their entire lives on the chessboard, and what they are looking for is a beautiful and wonderful player. It is so profound, beautiful, mysterious, and abstract that it becomes a dull neuron parameter overnight, which makes many people disillusioned.

In fact, in the field of visual art, the ANN has been able to separate the content and style of a work, at the same time learn the artistic style from the art masters, and transfer the artistic style to another work, rendering the same content with the style of different artists.

This means that the ANN can accurately quantify many vague concepts in the original humanities, such as the “artistic style” in specific fields and the “chess style” in games, and make these skills and styles that can only be understood and cannot be expressed in words simple, easy to copy, and promote.

1.4.4 Other Aspects

In terms of game games, the Deep Q-Network DQN developed by the Google DeepMind team has reached or exceeded the level of human professional players in 29 of the 49 Atari pixel games. Alpha Dog is a top player in Go that completely surpasses humans.

In May 2016, AI Lab from Google reported that the researchers used 2865 English romance novels to train the machine and let the machine learn the narrative and diction style of romance novels. From the perspective of the evolution process of the program, the machine model first comprehends the structure of the space between words and then slowly understands more words. From short to long, the rules of punctuation are slowly mastered, and some sentence structures with more long-term relevance are slowly mastered by the machine.

In May 2016, Google’s DeepMind team wrote an article and developed a “Neural Programming Interpreter.” This neural network can learn and edit simple programs by itself, which can replace the work of some junior programmers.

APP and mobile phones: Because of the emergence of intelligent photography technology, we have been able to do some new human–computer interactions. In some new fields, whether from the perspective of mobile internet companies, or from the perspective of live broadcasting, online education, mobile games, and other industries, there are several new products to do, representing some new growth in the industry. In the next step, there is new content to be done in the lens module of mobile phones and the industrial chain of mobile phones, from unlocking faces to intelligent photography and other new products.

Security field: At present, face recognition in the security field is changing from single-point camera recognition to the overall recognition of a camera cluster, which is a step in the progress of technology. There are still many technological processes to be completed before we can truly achieve product maturity and user experience maturity. In this process, it needs CPU supercomputing and other related products to drive new markets in smart cities, transportation, government affairs, and other aspects. In general, it is now possible to provide safer and more efficient urban management means and means to catch people in this field in the technical link.

Internet of Things: This is a trillion-level market with many scenarios. Although the needs are different, the technical commonalities are the same: VIP identification, identifying whether you are a VIP member; identity authentication, whether you are yourself; passenger flow analysis, through the data statistics and analysis of a large number of people to better optimize the offline sales, what kind of people to sell, how to lay out the goods and stores, and to carry out optimization and sales promotion.

The maturity of underlying technology drives the emergence of new retail. The corresponding market increment can be found in replacing access control, smart parks, buildings, hospitals, schools, and other scenarios. Accordingly, companies engaged in offline scene integration have relatively potential new market value increments.

Advertisement: The profit from advertising is very high. After the formation of better AI technology, user-generated content (UGC), live broadcast, augmented reality, and other technologies and industrial scale, some new channels have emerged, including the increment of advertising communication, advertising production, advertising services, and other links. In essence, the core is the emergence of new advertising forms; everyone can advertise the products they want to endorse. The live broadcast scenario can also optimize the overall advertising communication efficiency.

Google driverless car: Google Driverless Car is a fully autonomous vehicle developed by Google X Lab of Google, which can start, drive, and stop without the driver. It is currently being tested to drive 480,000 km. The project is led by Sebastian Thrun, the co-inventor of Google Street View. Google engineers used seven test vehicles, six of which were Toyota Prius and one was Audi TT. These cars have been tested on several roads in California, including the Jiuqu Flower Street in the San Francisco Bay Area. These vehicles use cameras, radar sensors, and laser rangefinders to "see" other traffic conditions and use detailed maps to navigate the road ahead. Google said that these vehicles are safer than manned vehicles because they can react more quickly and effectively.

Boston Atlas robot is controlled by a robot. Atlas is a humanoid robot developed by Google's Boston Dynamics. The product iteration of Atlas focused on three major version updates. Atlas robot is based on the early PETMAN humanoid robot of Boston Dynamics, and PETMAN is a humanoid robot designed to detect chemical protective clothing, which can simulate how soldiers act with protective clothing under real conditions. Unlike the limited movement posture of the chemical protective clothing testing machine in the past, PETMAN can not only balance itself and walk freely, bend the body, but also do various calisthenics that have pressure on the chemical protective clothing in the chemical warfare agent operation workshop. PETMAN also simulates the actual test conditions by simulating the human physiology in the protective clothing to control the temperature, humidity, and sweating. Since PETMAN is a humanoid robot used to detect chemical protective clothing, starting from PETMAN, Boston Power Development Department has established Atlas humanoid robot to improve its movement ability. The first version of Atlas robot has four hydraulically driven limbs. Atlas robot is made of aviation grade aluminum and titanium, with a height of about 1.8 m and a weight of 150 kg, and is illuminated by a blue LED. The Atlas robot is equipped with two vision systems – a laser rangefinder and a stereo camera, which are controlled by an onboard computer. Its hands have the ability of fine motor skills. Its limbs have a total of 28 degrees of freedom. The main feature of the first version is that Atlas needs a long cable for power supply. It can walk on a road covered with stones in the laboratory environment and keep its body balanced without falling. Second, in the laboratory environment, let Atlas stand on one foot and introduce external impact. Atlas can still stand on one foot without falling when the impact force is not very large. When walking on the green belt, place obstacles on the path of one foot. When Atlas steps on the obstacles, it will automatically update its gait in the next step and plan the landing point of the next step to leave the obstacles. Finally, carry out the road test in the field environment, and Atlas can also maintain a good balance in the field environment. At the beginning of 2015, in order to participate in the DARPA Robot Challenge at the beginning of June, Atlas completed its own evolution. That is, Atlas has successfully evolved into the second edition. This time, 75% of Atlas's whole body has been redesigned. Only the lower legs and feet have followed the previous version of the design. The new design makes it stronger, faster, and quieter. Thanks to the large battery backpack behind it, it has been freed from the cable. The upgraded Atlas is 1.88 m high and weighs 156.5 kg. The newly improved Atlas is slimmer and smaller than the first version. The more effective airborne hydraulic pump also makes the Atlas robot move faster. The purpose of this improvement is to make Atlas robot easier to complete some challenges. For example, it is required to squeeze into a space specially designed for human beings. So the purpose of this Atlas is to put it in real disaster rescue scenarios. It can communicate with the upper computer, and the protective rope on its head can also be removed. The power supply system of the second version of Atlas is a 3.7-kwh lithium-ion battery pack, which can last for one hour when completing actions including walking, standing, and using tools. The powerful hydraulic pump gives Atlas the ability to stand up when it falls. The third version of Atlas can be operated indoors and outdoors. Similar to the power of the second version, Atlas uses a power supply and hydraulic drive. This also shows that in order to have great power, energy can only be compromised by hydraulic pressure at present. The sensors in Atlas's body and legs keep its body balanced by collecting posture data. The lidar locator and stereo camera on its head can enable Atlas to avoid obstacles, detect ground conditions, and complete cruise tasks. This version of Atlas is 1.75 m high and weighs 82 kg, both of which are lighter than the previous version. In the process of carrying boxes and pushing doors, it is also necessary to mark and complete the object recognition task, and the machine vision ability such as object recognition needs to be improved. The most striking feature of this version is the function that Atlas can climb up automatically after falling down. Compared with other robots, the performance of this function is excellent, and the efficiency of getting up is also very high.

AlphaGo of the game. AlphaGo is the first AI robot to defeat the human professional Go player and the world champion of Go. It is developed by a team led by Demis Hassabis, a Google's DeepMind company. Its main working principle is "deep learning." AlphaGo has used many new technologies, such as neural networks, deep learning, and Monte Carlo tree search, which has made a substantial leap in its strength. AlphaGo is mainly composed of four parts: strategy network, given the current situation, predict, and sample the next move; the goal of fast-moving is the same as that of the strategic network, but the speed is 1000 times faster than that of the strategic network block with an appropriate loss of the quality of moving; value network, given the current situation, it is estimated that the probability of white victory or black victory is high, and Monte Carlo tree search connects the above three parts to form a complete system. The previous version of AlphaGo combines the chess manual of millions of human go experts and the supervised learning of intensive learning for self-training. The capability of AlphaGoZero has been improved qualitatively on this basis. The biggest difference is that it no longer requires human data. In other words, it has not touched human chess at the beginning. The R&D team just let it play chess freely on the chessboard and then conducts a self-game. According to David Silva, the head of the AlphaGo team, AlphaGoZero uses new reinforcement learning methods to make itself a teacher. In the beginning, the system did not even know what is Go, but it started with a single neural network and played it by itself through the powerful search algorithm of the neural network. With the increase of self-game, the neural network is gradually adjusted to improve its ability to predict the next step and finally win the game. What is more, with the deepening of the training, the AlphaGo team found that AlphaGoZero also independently discovered the rules of the game and came up with new strategies, bringing new insights into the ancient game of Go.

References

1 Wang, J.J., Ma, Y.Q., Chen, S.T., et al. (2017). Fragmentation knowledge processing and networked artificial intelligence. *Scientia Sinica (Informationis)*, 47 (2): 171–192.

2 Hudyjaya Siswoyo, J. and Nazim, M.N. (2013). Development of minimalist bipedal walking robot with flexible ankle and split-mass balancing systems. *International Journal of Automation and Computing* 10 (5): 425–437.

3 Zhang, W.X., Ma, L., and Wang, X.D. (2017). Reinforcement learning for event-triggered multi-agent systems. *CAAI Transactions on Intelligent Systems* 12 (1): 82–87.

4 Lake, BM., Salakhutdinov, R., Tenenbaum, J.B. (2015). Human-level concept learning through probabilistic program induction. *Science* 350 (6266): 1332–1338.

5 Gao, X.S., Li, W., and Yuan, C.M. (2013). Intersection theory in differential algebraic geometry: generic intersections and the differential chow form. *Transactions of the American Mathematical Society* 365 (9): 4575–4632.

6 Luo, J.M. (2005). Intelligence and an intelligent model. In: *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics*, 5624–5629. Guangzhou: IEEE Press.

7 Luzeaux, D. and Dalgalarrondo, A. (2001). HARPIC, an hybrid architecture based on representations, percept ion and intelligent control: a way to provide autonomy to robots. In: *Proceedings of the International Conference on Computational Science*, 327–336. London: ACM Press.

8 Hintong, G., Deng, L., Yu, D. et al. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine* 29 (6): 82–97.

9 Deng, J., Dong, W., Socher, R. et al. (2009). Image Net: A large-scale hierarchical image database. In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on Miami*, 248–255. IEEE.

10 Dan, C.C., Meier, U., Gambardella, L.M. et al. (2010). Deep big simple neural nets excel on handwritten digit recognition. *Neural Computation* 22 (12): 3207–3220.

11 Hinton, G.E., Srivastava, N., Krizhevsky, A. et al. (2012). Improving neural networks by preventing co-adaptation of feature detectors. *Computer Science* 3 (4): 212–223.

12 Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. *Journal of Machine Learning Research* 15: 315–323.

13 Hinton, G.E., Osindero, S., and Teh, Y.W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation* 18 (7): 1527–1554.

14 Zhong, Y.X. (2006). A preliminary investigation on information knowledge intelligence transforms. *Frontiers of Electrical and Electronic Engineering in China* 1 (3): 60–66.

15 Socher, R., Lin, C., Ng, A. Parsing natural scenes and natural language with recursive neural Networks. *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, Bellevue, WA, USA. Germany (28 June–2 July 2011): International Machine Learning Society, 2011.