Review:

Hybrid-augmented intelligence: collaboration and cognition[*]

Nan-ning ZHENG[†‡1,2], Zi-yi LIU[1,2], Peng-ju REN[1,2], Yong-qiang MA[1,2], Shi-tao CHEN[1,2], Si-yu YU[1,2], Jian-ru XUE[1,2], Ba-dong CHEN[1,2], Fei-yue WANG[3]

(*[1]Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China*)

(*[2]National Engineering Laboratory of Visual Information Processing Applications,*

Xi'an Jiaotong University, Xi'an 710049, China)

(*[3]State Key Laboratory of Management and Control for Complex Systems, Institute of Automation,*

Chinese Academy of Sciences, Beijing 100190, China)

[†]E-mail: nnzheng@mail.xjtu.edu.cn

Received Jan. 16, 2017; Revision accepted Jan. 22, 2017; Crosschecked Jan. 22, 2017

Abstract: The long-term goal of artificial intelligence (AI) is to make machines learn and think like human beings. Due to the high levels of uncertainty and vulnerability in human life and the open-ended nature of problems that humans are facing, no matter how intelligent machines are, they are unable to completely replace humans. Therefore, it is necessary to introduce human cognitive capabilities or human-like cognitive models into AI systems to develop a new form of AI, that is, hybrid-augmented intelligence. This form of AI or machine intelligence is a feasible and important developing model. Hybrid-augmented intelligence can be divided into two basic models: one is human-in-the-loop augmented intelligence with human-computer collaboration, and the other is cognitive computing based augmented intelligence, in which a cognitive model is embedded in the machine learning system. This survey describes a basic framework for human-computer collaborative hybrid-augmented intelligence, and the basic elements of hybrid-augmented intelligence based on cognitive computing. These elements include intuitive reasoning, causal models, evolution of memory and knowledge, especially the role and basic principles of intuitive reasoning for complex problem solving, and the cognitive learning framework for visual scene understanding based on memory and reasoning. Several typical applications of hybrid-augmented intelligence in related fields are given.

Key words: Human-machine collaboration; Hybrid-augmented intelligence; Cognitive computing; Intuitive reasoning; Causal model; Cognitive mapping; Visual scene understanding; Self-driving cars

http://dx.doi.org/10.1631/FITEE.1700053   **CLC number:** TP18

1 Introduction

The unprecedented development of artificial intelligence (AI) technology (Marr, 1977; Russell and Norvig, 1995) is profoundly changing the relationships and interactive modes between humans and between humans and their physical environments and society (McCarthy and Hayes, 1987; Holland, 1992). With the help of AI, solving various problems of high complexity, uncertainty, and vulnerability in every field of engineering technology, scientific research, and human social activities (Eakin and Luers, 2006; Martin, 2007; Gil *et al.*, 2014; Ledford, 2015) and continuously promoting the development of society and socioeconomics have become the cherished goals

of science and technology. AI is an enabling technology leading to numerous disruptive changes in many fields (Minsky, 1961; Stone *et al.*, 2016). Using AI technology reasonably and effectively can greatly promote valuable creativity and enhance the competitiveness in both humans and machines. Thus, AI is no longer an independent, isolated, and self-cycling academic system, but a part of the human evolutionary process.

In recent years, deep learning methods have gained rapid development with the boost in computer data acquisition, storage, and calculation capabilities (Hagan *et al.*, 2002; Sun *et al.*, 2014). A new boom in AI has been triggered, especially in high demand fields such as cloud computing (Youseff *et al.*, 2008), big data (O'Leary, 2013), wearable devices (Son *et al.*, 2014), and intelligent robots (Thrun *et al.*, 1998), which all promote the development of AI theory and technology.

The development of AI can be described using a three-dimensional (3D) space, which includes strength, extension, and capability. Strength refers to the intelligence level of AI systems, extension refers to the scope of the problems that can be solved by AI systems, and capacity refers to the average solution quality that AI systems can provide. General AI systems can do unsupervised learning deftly based on experience and knowledge accumulation. However, general AI cannot be realized with a simple combination of computing models and algorithms from AI methods. DeepBlue (Campbell *et al.*, 2002), Watson (?Rachlin, 2012; Shader, 2016), and AlphaGo (Silver *et al.*, 2016) are AI systems that have achieved great success in challenging human intelligence in some fields by relying on the powerful processing ability of computers. However, these systems cannot evolve to a higher intelligence level by virtue of their own thought processes yet. There is a gap between those systems and general AI in regards to a high self-learning ability (Simon, 1969; Newell and Simon, 1972; Selfridge, 1988).

Intelligent machines have become the intimate companions of humans, where the interaction and cooperation between a human and an intelligent machine will become integral in the formation of our future society. However, many problems that humans face tend to be of high complexity, uncertainty, and open-ended. Because the human is the service object and arbiter in the ultimate 'value judgment'

of an intelligent machine, human intervention in the machine has been consistent throughout the evolution of these systems. In addition, even if sufficient or infinite data resources are provided for AI systems, human intervention cannot be ruled out of intelligent systems. There are many problems to be solved in AI, for example, how to understand the nuances and fuzziness of human language in the face of the human-computer interaction system, and especially how to avoid the risks or even harms caused by the limitations of AI technology in some important applications, such as industrial risk control (de Rocquigny *et al.*, 2008), medical diagnosis (Szolovits *et al.*, 1988), and the criminal justice system. To solve these problems, human supervision, interaction, and participation must be introduced for verification purposes. Hence, on the one hand, the confidence level in intelligent systems will be improved, and human-in-the-loop hybrid-augmented intelligence will be constructed; on the other hand, human knowledge will be optimally used. Therefore, in this paper, we highlight the concept of hybrid-augmented intelligence, which skillfully combines human cognitive ability and the capabilities of computers in fast operations and mass storage. Particularly, the definitions are as follows:

Definition 1 (Human-in-the-loop hybrid-augmented intelligence) Human-in-the-loop (HITL) hybrid-augmented intelligence is defined as an intelligent model that requires human interaction. In this type of intelligent system, human is always part of the system and consequently influences the outcome in such a way that human gives further judgment if a low confident result is given by a computer. HITL hybrid-augmented intelligence also readily allows for addressing problems and requirements that may not be easily trained or classified by machine learning.

Definition 2 (Cognitive computing based hybrid-augmented intelligence) In general, cognitive computing (CC) based hybrid-augmented intelligence refers to new software and/or hardware that mimics the function of the human brain and improves computer's capabilities of perception, reasoning, and decision-making. In that sense, CC based hybrid-augmented intelligence is a new framework of computing with the goal of more accurate models of how the human brain/mind senses, reasons, and responds to stimulus, especially how to build causal models,

intuitive reasoning models, and associative memories in an intelligent system.

In addition, because of issues with qualification (Thielscher, 2001) and ramification (Thielscher, 1997), not all problems can be modeled; i.e., it is impossible to enumerate all the prerequisites of an action, or to enumerate all the branches following an action. Machine learning cannot understand real world environments, nor can it process incomplete information and complex spatial and temporal correlation tasks better than the human brain does. It is impossible for a formal system of machine learning to describe the interaction of the human brain across the spectrum of non-cognitive factors and cognitive functions or to emulate the high plasticity of the brain's nervous system. The brain's understanding of non-cognitive factors is derived from intuition and influenced by empirical and long-term knowledge accumulation (Pylyshyn, 1984). All these biological characteristics of the brain contribute to enhancing the adaptability of machines in complex dynamic environments or on the scene, promoting machine abilities in non-integrity and unstructured information processing and self-learning, and inspiring the building of CC hybrid-augmented intelligence.

CC frameworks can combine the modules for complex planning, problem solving, and perception as well as actions. These frameworks can possibly provide an explanation for some human or animal behaviors and study their actions in new environments, and they could build AI systems that require much less calculation than existing systems.

2 Human-computer collaborative hybrid-augmented intelligence

2.1 Human intelligence vs. artificial intelligence

Humans can learn, speak, think, and interact with the environment to perform actions and study. The capacity for human movement also depends on such learning mechanism. The most ingenious and important ability of human beings is to learn new things. The human brain has the ability for self-adaptation and knowledge inference, which transcends experience. In addition, human is gregarious, a quality where cooperation and dynamic optimization show that collective intelligence is much better

than that of any individual. In one word, human intelligence is creative, complex, and dynamic (Guilford, 1967; Sternberg, 1984). The creativity of human beings means that human intelligence is skillful in abstract thinking, reasoning, and innovation, creating new knowledge and making associations. The complexity of human intelligence implies the structural complexity and connective plasticity of the neural system inside the human brain, and the complexity inherent to a series of intuitive, conscious, and thinking mechanisms. At present, there is no common conclusion regarding the mechanism of human intelligence, but it is precisely because of the complex structure of the human brain that human intelligence can better specialize in dealing with non-integrity and unstructured information. The dynamic nature of human knowledge evolution and learning ability makes humans more adept at learning, reasoning, collaborating, and other advanced intelligence activities.

In considering an analogy with human intelligence, AI has the features of normalization, repeatability, and logicality. Normalization refers to the fact that AI can deal only with structural information at present; i.e., the input of programs must conform to certain norms. Repeatability refers to the mechanical nature of AI. Repetitive work does not degrade the efficiency or accuracy of the machine because of the powerful computing ability and abiotic characteristic of a computer. Logicality means that AI has advantage in dealing with the symbolized problem, which means that AI is better at processing some discrete tasks, instead of discovering or breaking the rules by itself (Poole *et al.*, 1997). Fig. 1 shows a comparison between human intelligence and AI. It can be seen that although AI and human intelligence each have distinctive advantages, they are highly complementary.

2.2 Limitations of existing machine learning methods

Machine learning (Nilsson, 1965; Michalski *et al.*, 1984; Samuel, 1988) makes it possible to predict the future through the patterns of past data. In short, machine learning can be considered to be the automation of predictive analyses and it generates models based on computer data. When dealing with a new task, the system makes a corresponding judgment according to a data-based model, which is a

'training & test' learning mode. This learning mode depends entirely on the machine's performance and learning algorithms (Bradley, 1997). In fact, the process of using machine learning to deal with complex, dynamic, and unstructured information (Wang *et al.*, 2017) is much more complex than that of the human process because a machine has to make choices between data sources and options, while a human can quickly make a decision according to slight differences in the tasks and the complex relationships among the data.

Machine learning relies excessively on the rules, which results in poor portability and scalability. Thus, it can work only in an environment where there are tight constraints and limited objectives, and it cannot process dynamic, non-complete, and unstructured information. Although hybrid-augmented computational intelligent systems can be constructed by artificial neural networks (Yegnanarayana, 1994), fuzzy reasoning (Mizumoto, 1982), rough sets, approximate reasoning (Zadeh, 1996), and optimization methods, such as evolutionary computation (Fogel, 1995) and group intelligence (Williams and Sternberg, 1988), to overcome individual limitations and achieve synergies to some degree with the integration of different machine learning methods and adaptive techniques, these systems are still incapable of exercising common sense, to solve time-varying complex problems, and to use experience for future decisions (Jennings, 2000). Indeed, no matter how much development happens in machine learning, it is impossible for a machine to complete all the tasks in human society individually. In other words, a human cannot rely completely on machine learning to carry out all work, such as economic decision making, medical problem solving, and mail processing.
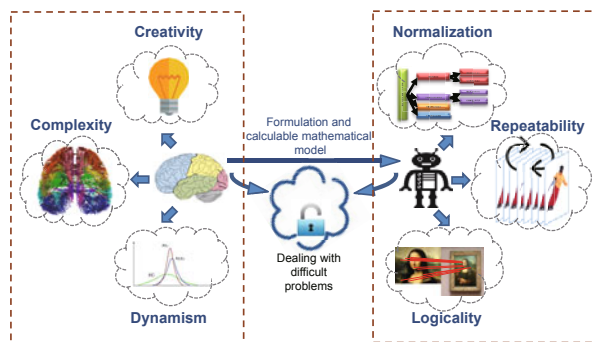
Humans are capable of extracting abstract concepts from a small number of samples. However, even though deep neural network (DNN) has made great progress in recent years, it is still difficult to make a machine do such things like a human. However, Lake *et al.* (2015) used Bayesian learning methods so that a machine can learn how to write letters like a human through a small amount of training data. Compared with traditional machine learning methods, which require a great deal of training data, this method requires only a rough model, and then uses a reasoning algorithm to analyze the case and update the details of the model.

The growth in the amount of data is a source of 'complexity' that must be tamed via algorithms or hardware, whereas in statistics, the growth in the amount of data brings 'simplicity' in a statistical sense, which often provides more support for reasoning, leading to stronger, asymptotic results. At a formal level, the gap is made evident by the lack of a role for computational concepts (such as 'runtime' in core statistical theory) and the lack of a role for statistical concepts (such as 'risk' in core computational theory). Therefore, machine learning with a stronger reasoning capacity requires more integration of computational and inferential aspects at the foundational level (Jordan, 2016).

2.3 Human-in-the-loop hybrid-augmented intelligence

Introducing human intelligence to the loop of intelligence systems can realize a close coupling between the analysis-response advanced cognitive mechanisms in fuzzy and uncertain problems and the intelligent systems of a machine (Fig. 2). Hence, the two adapt to and collaborate with each other, forming a two-way information exchange and control. Such a '1 + 1 > 2' hybrid-augmented intelligence can be achieved by integrating human perception, cognitive ability, machine computing, and storage capacities (Pan, 2016). Ultimately, information from a large-scale, non-complete, and unstructured knowledge base can be processed, and the risks of out of control brought by AI technologies can be avoided.

The Internet provides an immense innovation space for HITL hybrid-augmented intelligence. Internet information processing is considered by some researchers as the processing of highly structured and standardized semantic information, a process they believe can be processed by computers as long as

Fig. 1 Human intelligence vs. artificial intelligence
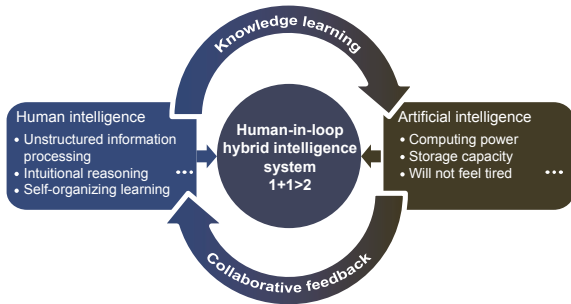
Fig. 2 Human-in-the-loop hybrid-augmented intelligence

human knowledge is properly marshaled. In fact, the Internet is full of disorganized, messy fragments of knowledge (Wang *et al.*, 2017), and much of it can be understood only by humans. Therefore, machines cannot complete all the tasks of Internet information processing. Human intervention is still needed on many occasions.
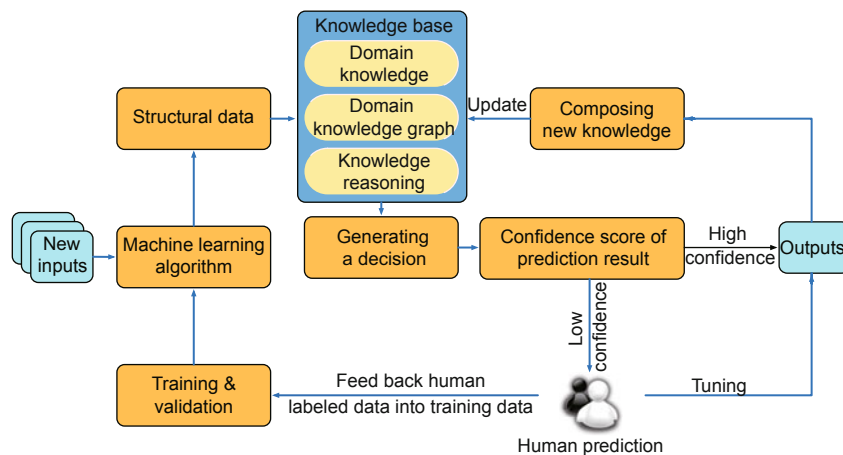
HITL hybrid-augmented intelligence needs to cover the basic functions of computable interaction models, including dynamic reconstruction and optimization, autonomy and adaptivity during interactive sharing, interactive cognitive reasoning, and methodologies for online evaluation. HITL hybrid-augmented intelligence can effectively realize the concept of human-computer communication, especially at the conceptual level of knowledge, where computers can not only provide intelligent-ware in different models, but also talk to human beings at the conceptual level of knowledge.

Different HITL hybrid-augmented intelligence systems should be constructed for different fields.

Fig. 3 shows the basic framework of HITL hybrid intelligence, which can be considered a hybrid learning model. The hybrid learning model integrates machine learning, knowledge bases, and human decision making. It uses machine learning (supervised and unsupervised) to learn a model from training data or a small number of samples, and predicts new data by using the model. When the predictive confidence score is low, humans will intervene to make judgments. In the hybrid learning framework shown in Fig. 3, when the system is abnormal, or when the computer is not confident in success, the confidence estimation or the state of the computer's cognitive load will determine whether the prediction needs to be adjusted by a human or whether human intervention is required, and the knowledge base of the system is automatically updated. In fact, human prediction and intervention in the algorithm increases (improves) the accuracy and credibility of the system. Of course, HITL hybrid-augmented intelligence needs to reduce human participation as much as possible, so that the computers can complete most of the work. The intelligence of a hybrid learning mode as shown in Fig. 3 is able to greatly expand the scale and efficiency of the tasks humans can complete.

The main research topics for HITL hybrid-augmented intelligence include:

1. how to break through the human-computer interaction barrier, so that machines can be trained in the intelligence circuit in a natural way,

2. how to combine human decision making and experience with the advantages of machine intelligence in logical reasoning, deductive inference, and

Fig. 3 Basic framework of human-in-the-loop hybrid-augmented intelligence (integrating supervised and unsupervised learning, knowledge bases, and human decision-making hybrid learning models)

so on, so that a man-machine collaboration of high efficiency can be realized,

3. how to build cross-task, cross-domain contextual relations, and

4. how to construct task- or concept-driven machine learning methods which allow machines to learn from both massive training samples and human knowledge, to accomplish highly intelligent tasks by using the learned knowledge.

HITL hybrid-augmented intelligence is able to process highly unstructured information, generating more accurate and more credible results than what can be derived from a single AI system.

3 Hybrid-augmented intelligence based on cognitive computing

In nature, human intelligence is undoubtedly the most robust. The construction of hybrid-augmented intelligence based on CC, which uses research on effective cooperation mechanisms between biologically inspired information processing systems and modern computers, could possibly provide a novel method to solve the long-term planning and reasoning problems in AI.

3.1 Computing architecture and computing process

The construction of CC hybrid-augmented intelligence should take into consideration computing architecture and computing processes. That is to say, the kind of computing architecture and the kind of computing process needed to complete the calculation must be decided.

Modern computers are based on the von Neumann architecture. The computing process is based on the fact that computing tasks can be formulated by a symbolic system. Running a modern computer is a process of calculation by a formal model (software) in the von Neumann architecture, which can achieve complete and undifferentiated copies of data. Different solutions (software programs) to different problems are required. Once the model is established, its computational capabilities and the tasks it faces are determined.

The computing architecture of a biological intelligence is based on the brain and nervous system. The calculation process of biological intelligence is the process of constantly adapting to an environment or a situation, that is, applying risk judgments and value judgments. The biological intelligence's information processing mechanism has two aspects. One is a natural evolutionary process, which requires the biological intelligence system to be able to model the status of the environment and of itself and then provide an 'interpretable model', which forms the measurement of risk and value. The other is 'selective attention' (Moran and Desimone, 1985), which provides an efficient mechanism for comprehensive judgments of risk or value and screening key factors in complex environments, such as children looking for a father's face in the crowd after school. In many cases, risk and value judgments are based on continuously cycling thinking activities of prediction and choice on the basis of cognitive models, and verification thinking activities evolve and improve the cognitive models, such as summing up an abstract or formulaic experience as a theorem. Biological intelligence is a process of evolution; in addition to the common characteristics, it presents individual differences such as individual experience (memory), value orientation (psychological factors), and even 'nerve expressions' of microscopic differentiation. For instance, the same face may be represented differently in different human brains.

Experience indicates that for different tasks, the computing process of a biological intelligence is possibly separated from its computing architecture. Yet, sometimes these two parts cannot be separated from each other (how to identify this separation is also worthy of study). For a computing architecture, the cognitive model of biological intelligence can be used to complete the 'modeling' progress and formalize its representation. Finally, taking advantages of modern computers (computing devices), an effective collaborative calculation can be realized. For the computing process, a neuromorphology model can be constructed to emulate the biological brain in structure and processing. Therefore, the critical step in forming an effective CC framework is to develop the hybrid-augmented intelligence inspired by biological intelligence.

3.2 Basic elements of cognitive computing

Fig. 4 shows a schematic diagram of the basic components in a framework for CC. A CC framework includes six interrelated cognitive components, which are understanding, verifying, planning,
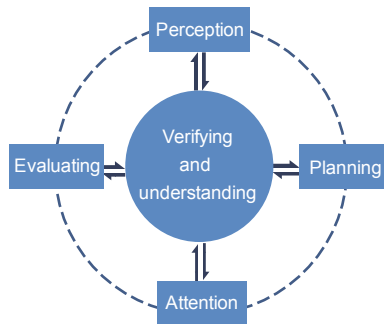
Fig. 4 Basic framework of cognitive computing

evaluating, attention, and perception. Any of them can serve as a starting point or an objective in a specific cognitive task. The system chooses a simple or complex interactive path (e.g., repeated iteration) to achieve the goal of cognition, according to the information required to interact with the outside world. Usually, the top-down selective attention is based on the planning attention, while the bottom-up selective attention is essentially based on the perception attention. Evaluation based on understanding or planning is the prior probability (performance as prediction), while evaluation based on perception is the posterior probability (performance as observation). In short, the process of CC is to constantly interact with the outside world based on the information required to meet objective tasks, and to gradually start a thinking activity, rather than be limited to knowledge-based processing. In the face of problems involving a lack of preparation, an intelligent system should have a cycle capability of 'do until ...', without traversing every possibility to achieve the goal of planning. This requires the CC process to contain verifying steps, including: What to do next? Did it produce the expected results? Whether to make further effort or try other methods? In such a process, the understanding and guidance of the environment is enriched based on reasoning and experience (long-term memory), and the ability to 'verify' is enhanced accordingly.

The above CC process requires construction of a causal model to explain and understand the world. Using the causal model to update the prior probability (the prediction) by the posterior probability (the observation), the association analysis is completed based on the probability analysis of given data, and the time/space-based imagination or prediction (such as spatial variation over time), provides understanding, supplement, and judgment of

the environment or situation. Planning action sequences are used to maximize future rewards, and prior knowledge is applied to enrich the reasoning of small-scale data to achieve good generalization ability and fast learning speed.

The main research topics for CC are as follows:

1. how to realize brain-inspired machine intuitive reasoning,

2. how to construct a causal model to explain and understand the world,

3. how to use the causal model to support and extend learned knowledge through intuitive reasoning, and

4. how to construct the knowledge evolution model, i.e., how to learn to learn and how to acquire and generate knowledge rapidly through the combination of knowledge.

3.3 Intuitive reasoning and casual model

3.3.1 Intuition and cognitive mapping

Intuition is a series of processes in the human brain including high-speed analysis, feedback, discrimination, and decisions (Fischbein, 2002). Studies have shown that the average accuracy of human intuitive judgment is higher than that of non-intuitive judgment (Salvi *et al.*, 2016). Humans make many decisions through intuition in their daily lives, such as judging the proximity of two objects, perceiving the unfriendliness of another's tone, and choosing one's partner or a book. Intuitive decision making is not just done by common sense. It involves additional sensors to perceive and become aware of information from outside.

Intuition can be divided into three processes, namely selective encoding, selective combination, and selective comparison (Sternberg and Davidson, 1983; Sternberg, 1984). Selective encoding involves sifting out relevant information from irrelevant information. However, selective encoding is still insufficient for a human to achieve accurate understanding. Selective combination is also needed to combine the encoded information in some way and form reasonable internal relations with other information as a whole. Thus, selective combination involves combining what might originally seem to be isolated pieces of information into a unified whole that may or may not resemble its parts. Selective comparison involves relating newly acquired information to old

information that one already has. When people realize the similarity to a certain degree between the old information and new information, people can use this similarity to achieve a better understanding of the new information.

Therefore, intuition helps humans make rapid decisions in complex and dynamic environments. Besides, it greatly reduces the search space in the process of solving problems and makes the human cognitive process more efficient.

Intelligence represents a model of characterization and facilitates a better ultimate cognition. One kind of cognitive 'pattern' that arises in the mind can be thought of as a world model constructed based on prior knowledge. This model contains three kinds of relationships: interaction, causality, and control. The world model can be considered a cognitive map of the human brain, which resembles an image of the environment. It is a comprehensive representation of the local environment, including not only a simple sequence of events but also directions, distance, and even time information. This concept of a cognitive map was first proposed by Tolman (1948). A cognitive map can also be represented by a semantic web (Navigli and Ponzetto, 2012). From the aspect of information processing theory, a cognitive map (or cognitive mapping) is a dynamic process with steps of data acquisition, encoding, storage, processing, decoding, and using external information (O'Keefe and Nadel, 1978).

People are able to model their own state and relationship within the environment, and then provide an interpretable model to form a basis and measure of evaluation and judgment of risk and value. Human cognitive activities are embodied in a series of decision-making activities based on the cognitive map, which is a process of pattern matching. The formation of a current cognitive map is related to the brain's perception and the understanding of external information. As shown in Fig. 5, through the human individual's growth and accumulation of learning, common sense, and experience, a human forms a 'decision-making space', and the brain searches decisions in the decision space randomly; once the selected decision matches the current cognitive map, where the match may be defined by a minimum cost, people will respond intuitively. In this process, the role that intuition plays can be considered guidance for a decision-making search as well as the construc-
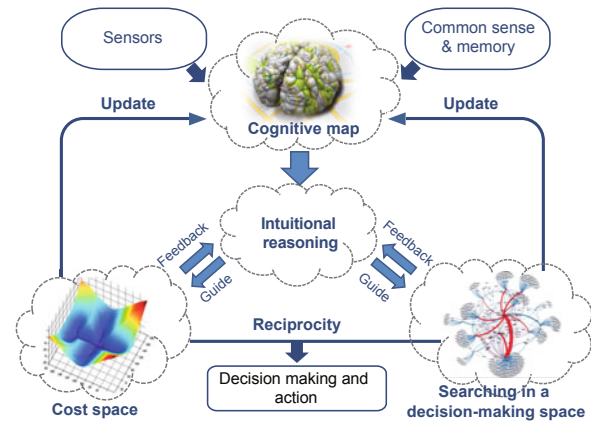
Fig. 5 Relation of intuitive reasoning and cognitive mapping

tion of a cost space in the computing process (Janis and Mann, 1977).

Humans' intuitive reasoning is closely related to the prior knowledge processing ability of the brain. This ability is about abstraction and generalization instead of the rote memory of prior knowledge. It is precise because of this ability that human intuition can make rapid risk mitigation decisions based on the world model in the human brain.

3.3.2 Machine implementation of intuitive reasoning

Although a machine has the power of symbolic computation and a storage capacity that the human brain cannot match, it is hard for existing machine learning algorithms to realize the concepts mentioned above, such as a cognitive map, decision space searching, and cost of space, like a human brain.

If the intuitive response can be considered as finding the global optimal solution in the search space, intuition can be regarded as the initial iteration position of the solution. This position is valid with large probabilities. This initial iteration position is not important when solving a simple problem. However, when solving complex problems, compared with traditional machine reasoning methods, the advantages of intuitive reasoning will be highlighted. In the latter case, traditional machine reasoning methods are likely to fall into local minima (Ioffe, 1979), while intuitive reasoning can provide a reasonable initial iteration position so that it can avoid the local minima problem to a great extent.

In practical terms, the solution space is often complex, non-convex, or even structurally indefinable (Hiskens and Davy, 2001). Therefore, the

selection of an initial iterative position is critical and can even decide whether the final result is the global optimal solution or not. In common machine learning methods, the initial iteration position is usually obtained at the sacrifice of the generalization abilities of the algorithm, such as the introduction of strong assumptions (Lippmann, 1987) and increasing human intervention (Muir, 1994). Constructing brain-inspired machine intuitive reasoning methods will avoid the problem of local minima and improve the generalization abilities of AI systems. Then, we can establish models for problems with uncertainty.

As seen from the above discussion, intuitive reasoning depends on the heuristics and reference points. The heuristic information is derived from experience, i.e., the prior information, which determines the direction of the problem solving. The choice of the reference point depends on other related factors, which determine the initial iteration position of the solution. Intuitive decision making does not seek to find the absolute solution of the target solution position, but to assess whether or not the deviation from the reference point is more conducive to the avoidance of loss. In reality, intuitive judgments tend to show the characteristics of the minimal cost (or minimal risk) based on the 'reward and punishment' rule. Therefore, intuitive reasoning can be simulated by machines. The hybrid-augmented intelligence based on CC requires optimally integrating the two reasoning mechanisms, i.e., intuitive reasoning (Tversky and Kahneman, 1983) and deductive reasoning (Dias and Harris, 1988), based on mathematical induction.

The success of AlphaGo can be seen as a successful example of the application of machine intuitive reasoning. The solution space for Go is nearly impossible to exhaust and the approaches based on rules or exhaustive searching cannot make Go programs reach the master level of a human. AlphaGo achieved intuitive reasoning to a certain extent. Its intuition is reflected in its simulation of the 'Go sense', which is realized by the policy network and value network (Fig. 6). The policy network is a quick judgment of where to move, i.e., which actions can be considered and which cannot. A value network evaluates overall positions. AlphaGo gains the 'Go sense' by training 30 million positions from the KGS Go server and the reinforcement learning process. The Go sense narrows the search space in the process of finding the
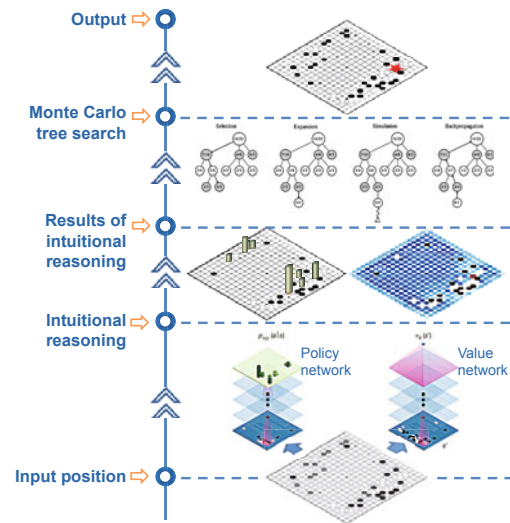

Fig. 6 Intuitive reasoning of AlphaGo

optimal solution, so that the computer can find the approximate optimal solution from the vast solution space through multithread iterations. The success of AlphaGo shows the importance of intuitive reasoning for problem solving.

Although AlphaGo has adopted a more general framework, it still involves a great deal of manually encoded knowledge. Designing specific encoding schemes for specific problems is the most common way to describe a problem to be solved in previous and present AI research. However, encoding methods are often manually designed for a particular purpose and do not ensure optimality. Besides, AlphaGo does not have the ability of associative memory. However, AlphaGo combines intuition (Go sense) with explicit knowledge (rules and chessboard) by the nonlinear mapping of deep learning and the jumping of Monte Carlo tree searching (Browne et al., 2012). It is of great value to the research into novel AI technologies.

In the research on hybrid-augmented intelligence, more attention must be paid to other methods of learning and reasoning, such as deep learning based reinforcement learning (Mnih et al., 2015), recurrent neuron network based methods (Mikolov et al., 2010), and differentiable neuron computers (Graves et al., 2016).

3.3.3 Casual model

Constructing an interpretable and understandable causal model is very important for the realization of the hybrid-augmented intelligence based on a

CC framework (Freyd, 1983). As shown in Fig. 7, the posture of a person riding downhill is clearly different from that when riding uphill. The angle between the postures and the slopes of the ground are bounded by a physical causal relationship. The relationship between the police and the thief can be seen as a kind of social causal relationship. Non-causal relationships manifest as no causal association between any two independent individuals. The causal model, which can be explained and understood in the framework of cognitive computation, should satisfy the physical constraints arising from the physical causal relationship in cognitive tasks, and regard the machine as 'itself' and understand itself and the causality involved in order to produce psychological reasoning judgments in the current cognitive task.
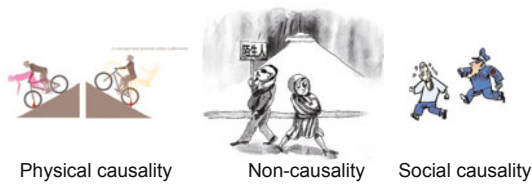

Fig. 7 Causal relationship: (a) physical causality; (b) non-causality; (c) social causality

Cognitive reasoning at the psychological level refers to the study and forecasting process of humans constrained or guided by their own mental state, such as in imitation activities (Premack and Premack, 1997; Johnson *et al.*, 1998; Tremoulet and Feldman, 2000; Schlottmann *et al.*, 2006). As shown in Fig. 8, a child remembers reward and punishment when he sees his friend playing a new game. On the next day, when he plays the same game, based on the memory of how his friend played the game, he can quickly find out how to deal with similar scenes. The child's behavior is guided by his psychological state when he is playing the same game. This is a kind of imitative learning and shows that people's perception of new things can be predicted based on their prior knowledge instead of complying with entirely new rules.

The causal model in CC can track the development spatiotemporally by cognitive inference at the physical level as well as at the psychological level, which means the learning procedure is guided by the mental state.

Fig. 9 shows the general framework of the causal model (Rehder and Hastie, 2001). Various objects
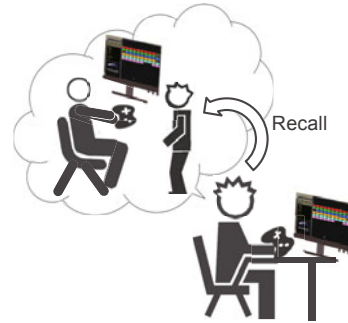

Fig. 8 Reasoning guided at the psychological level

that exist in the real world are represented by different class attributes in this model. A_1, A_2, A_3, and A_4 represent four different objects. In the common-cause schema (Fig. 9a), A_1 has causal effects on A_2, A_3, and A_4; i.e., in A_1/A_2, A_1/A_3, and A_1/A_4, A_1 is the cause. In the common-effect schema (Fig. 9b), A_4 is the effect, and the others are the cause. In the no-cause control schema (Fig. 9c), there exist no causal relationships between A_1, A_2, A_3, and A_4.
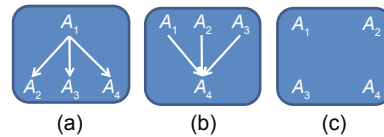

Fig. 9 General framework of the causal model: (a) common-cause schema; (b) common-effect schema; (c) no-cause control schema

Temporal and spatial causality widely exist in many AI tasks, especially in object recognition. For example, Chen D *et al.* (2016) proposed a tracking system for video applications. Because the spatial causality between the target and the surrounding samples changes rapidly, the 'support' in the proposed system is transient. Thus, a short-term regression is used to model the support. The short-term regression is related to support vector regression (SVR), which exploits the spatial causality between targets and context samples and uses spatial causality to help locate the targets based on temporal causality.

Perceptual causality is the perception of causal relationships from observation. Humans, even as infants, form such models from observation of the world around them (Saxe and Carey, 2006). Fire and Zhu (2016) proposed a framework for unsupervised learning of this perceptual causal structure from video. It takes action and object status detections as input and uses cognitive heuristics to produce the

causal links perceived between them. This method has the precision to select the correct action from a hierarchy. Similarly, a typical application of prediction or inference tasks based on a causal model is in the action recognition tasks in video sequences (Wei *et al.*, 2013; 2016). As shown in Fig. 10, a good action recognition system is expected to be able to deal with temporal and spatial correlation, discover causal relationships and constraints among samples, investigate the relationships between actions and environment at a sematic level, and then transform an action recognition task into a structured prediction problem.

An enhanced deep learning system with structured prediction modules (Honey *et al.*, 2010) has been applied to address issues including natural language processing (Fig. 11a) (Xiao *et al.*, 2016), pose detection (Fig. 11b) (Wang LM *et al.*, 2016), and semantic segmentation (Fig. 11c) (Noh *et al.*, 2015).

3.4 Memory and knowledge evolution

3.4.1 Implementation of memory by artificial neural networks

The human learning mechanism (Nissen and Bullemer, 1987; Norman and O'Reilly, 2003) is based on memory, which is the foundation of human intelligence. The human brain has an extraordinary memory capability and can identify individual samples and analyze the logical and dynamic characteristics of the input information sequence. The information sequence contains a great amount of information and complex temporal correlations. Therefore, these characteristics are especially important for developing models for memory. Memory is equally im-

portant, if not more than computation for cognitive functions. Effective memory is able to greatly reduce the cost of computation. Take the cognition of human faces as an example. The human brain can complete this cognitive process through a few photos or even a single photo of a face. This is because a common cognitive basis of human faces has been formed in the brain and the common features of the face have been kept in mind so that identifying the new or unique features of the faces is the only task for a human. However, machines require many training samples to achieve the same level as a human does. For example, in natural language processing tasks such as a question-answering system, a method is needed to temporarily store the separated fragments. Another example is to explain the events in a video and answer related questions about the events, where an abstract representation of the events in the video must be memorized. These tasks require the modeling of dynamic sequences in the time scale and forming long- and short-term memories of historical information properly.

However, information is converted into binary code and written into memory for computer devices. The capacity for memory is completely determined by the size of the storage devices. In addition, storage devices are hardly capable of processing data, which means storage and calculation are completely separated physically and logically. Therefore, it is difficult for existing AI systems to achieve an associative memory function. For future hybrid-augmented intelligence systems, a brain-like memory ability is in great demand (Graves *et al.*, 2013) so that the machine can imitate the human brain's long- and short-term memories. For example, we could form a
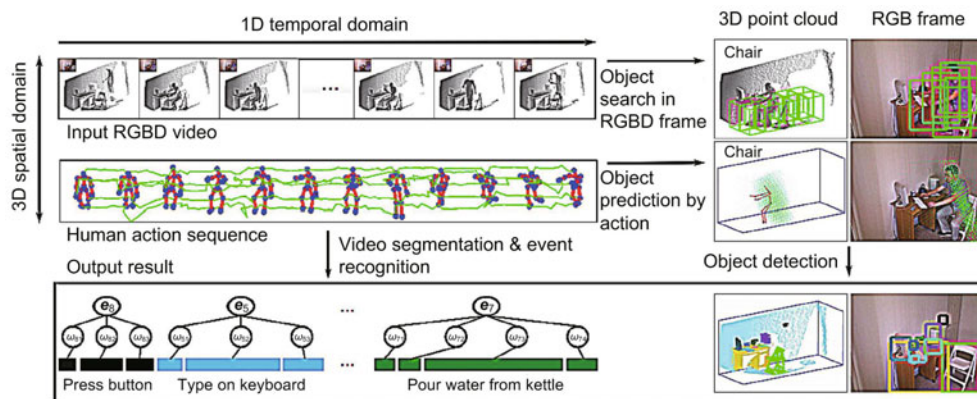

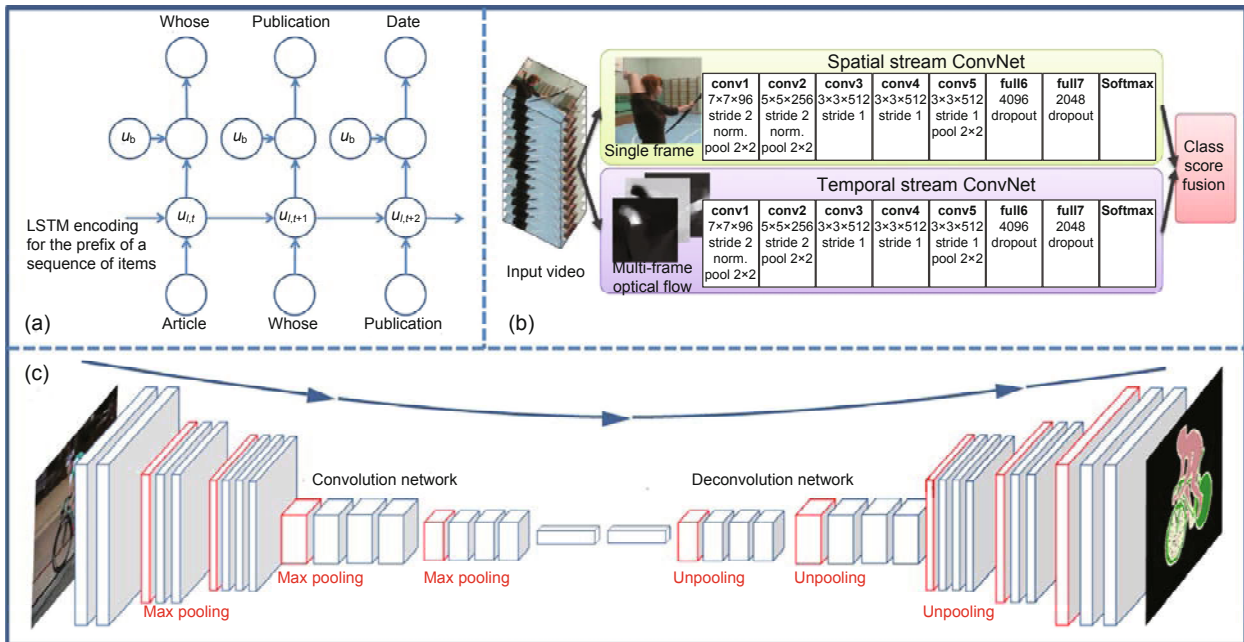**Fig. 10 Action recognition system integrating causal constraints**

Fig. 11 Combining deep learning with structured prediction to address issues of natural language processing (a), pose detection (b), and semantic segmentation (c)

stable loop that represents some basic feature, and maintain it for some period of time in a part of the artificial neural network (Williams and Zipser, 1989).

DeepMind proposed a structure for a neural Turing machine (Graves *et al.*, 2014). Its structure consists of two basic components: the neural network controller and the memory pool. As shown in Fig. 12, each neuron in a neural Turing machine interacts with the outside via input and output vectors as in traditional neural networks. The difference is that the neural Turing machine interacts with a memory matrix having selective read and write operations. Therefore, it can achieve a simple memory based inference. Park and Kim (2015) developed a coherent recurrent convolutional network architecture based on the neural Turing machine and used it to create novel and smooth stories from a series of images. Based on the neural Turing machine, Deep-Mind has proposed a differentiable neural computer (DNC) (Graves *et al.*, 2016). As shown in Fig. 13, the DNC also interacts with the external storage unit to complete the memory function and it is similar to a differentiable function from a mathematical point of view. Thus, the structure is also used to solve the problem of vanishing gradient in the long short-term memory (LSTM) network for modeling longer time series.

Differentiable neural networks combine well the advantages of the structured memory of a traditional computer with the capabilities for learning and decision making in a neural network, so they can be used to solve complex structured tasks that a traditional neural network is quite unable to do. The core of the differentiable neural network is the controller, which is characterized by a deep neural network, where the controller can carry out intelligent read
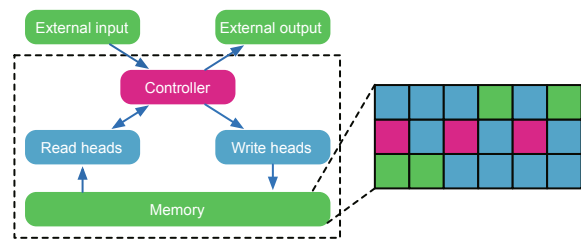

Fig. 12 Structure of the neural Turing machine
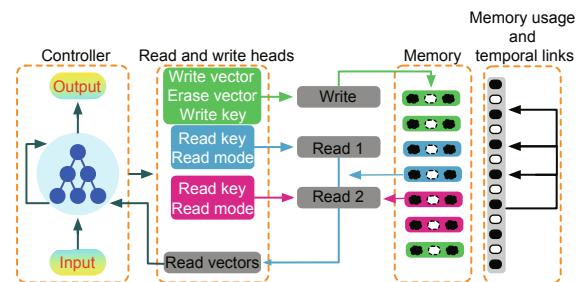

Fig. 13 Differentiable neural network

from and write in memory and reasoning decisions (Mnih *et al.*, 2013; 2015; Lillicrap *et al.*, 2016). For a given condition, the differentiable neural network makes inferences and autonomous decisions based on the relevant experience and knowledge in memory, and constantly makes itself remember, learn, and update strategies to complete the processing of complex data and tasks.

3.4.2 Knowledge evolution

The evolutionary process for human knowledge is the synergistic effect of the brain's memory mechanism and its knowledge transfer mechanism. Fig. 14a depicts the structure and hierarchy of the evolutionary process of human knowledge, which are alongside mental activities such as associative memory and association. This process is structural and hierarchical. Similarly, as shown in Fig. 14b, in the neocortex, neurons are not random, but have a certain structure and hierarchy. Moreover, knowledge is in a distributed representation in the human brain. The neural system stores information mainly by changing the strength of synaptic connections between neurons, and expresses different concepts through changes in multiple assemblies of neurons.

Human memory is associative memory (Ogura *et al.*, 1989), so the input information and the retrieved memory in the human brain are correlated at some level. For example, the former is part of the latter, or both of them are similar or related in content (such as the opposite), or they normally ap-

pear to exist simultaneously (spatial correlation) or sequentially (event-related), synchronously, or successively. Moreover, memory storage and retrieval is a well-structured sequence with rich dynamic features. This characteristic is the premise of knowledge evolution. In addition, memory and information processing are tightly coupled. The knowledge evolution model in the brain-inspired CC framework is expected to meet the following four requirements:

1. A probabilistic model is established for prior knowledge.

2. The evolution model is expected to achieve knowledge combinations and update them.

3. The evolution model is expected to determine whether to go further by using an existing strategy, or trying other methods, which is a self-proven process.

4. In the process of validation, the evolution model is expected to generate a rich understanding of the environment by taking advantage of causal models through intuitive reasoning and experience at the physical or psychological level (Lake *et al.*, 2016). Such understanding forms the basis of a verification capacity.

The general framework of the evolution model (Griffiths *et al.*, 2010; Tenenbaum *et al.*, 2011) is shown in Fig. 15. The first (bottom) layer represents the first-order logical expression of the abstract causal relationships, the weight status of the external intervention factors (x, y) that influence the development of the causal relationship, and the influence of

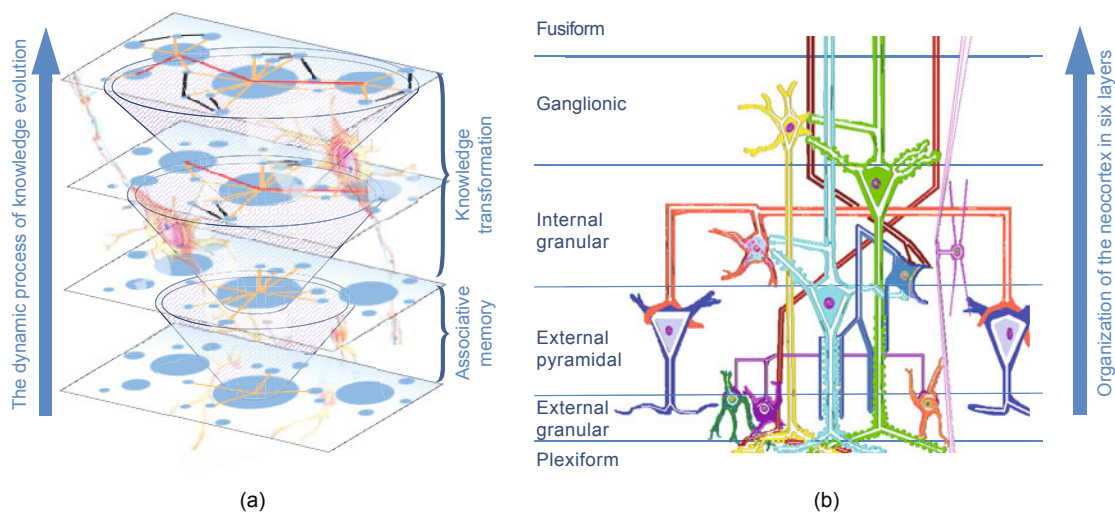(a)   (b)

Fig. 14 Process of the synergistic effect of brain's memory mechanism and knowledge transfer mechanism: (a) dynamic process of knowledge evolution; (b) organization of the neocortex in six layers
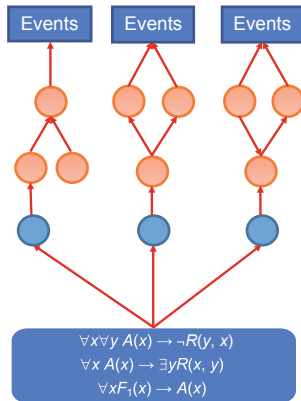
Fig. 15 General framework of the evolution model

the extrinsic intervention factors (F_1). Causality and external intervention factors are modeled as probability information (different data matrices). At the top level, the hypothesis space for event development is established, based on a probability computation, and the model can quickly converge to a certain event in the hypothesis space, that is, to predict the evolution of the results.

3.5 Visual scenes understanding based on memory and inference

The visual system plays a crucial role in understanding a scene through the visual center. The perception of the environment constructs a cognitive map in the human brain. Combining rules and knowledge stored in the memory with external information, such as map navigation information, a driver can make his/her driving decision and then control the vehicle following the decision.

Similarly, a brain-inspired automatic driving framework can be constructed and enlightened by the memory and reference mechanisms of the human brain. When someone is driving, the primary perception to create a basic description of the environment can be formed by the brain with a single glance. According to these environment perception results, integrating with the knowledge of the situation and related rules in the memory, the knowledge map of the traffic scene can be constructed.

Fig. 16 shows a hybrid learning network for an automatic driving vehicle using architecture with memory and inference. In this hybrid learning network, the road scene is first processed by multiple convolutional neural networks to simulate the function of human visual cortex and form a basic cogni-

tive map similar to human brain's, which is a structured description of the road sense and may contain explicit information and descriptions of hidden variables of the road sense. A more explicit cognitive map should be constructed based on the initial formation of the cognitive map and combined with prior elements of traffic and external traffic guidance information. Then, the cognitive map should contain both the description of the road sense and the driving strategy of the near future. Through the recurrent neural network (RNN) (Funahashi and Nakamura, 1993), the cognitive map formed in each frame is modeled to give the temporal dependency in motion control, as well as the long- and short-term memories of past motion states, imitating human motion (Kourtzi and Kanwisher, 2000). Finally, the control sequence of the automatic driving vehicle can be generated.

4 Competition-adversarial cognitive learning method

4.1 Generative model and adversarial network

The generative model and adversarial network (Xiao *et al.*, 2016) are combined by the competitive and adversarial cognitive learning methods, which can effectively represent the intrinsic nature of the data. This learning framework combines supervised learning with unsupervised learning to form an efficient cognitive learning. Adversarial training was first proposed by Szegedy *et al.* (2013) and Goodfellow *et al.* (2014a). The main idea of this learning framework is to train a neural network to correctly classify both normal examples and 'adversarial examples', which are bad examples intentionally designed to confuse the model.

Approaches to machine learning can be roughly divided into two categories: generative and discriminative methods. Models obtained by the two methods correspond to a generative model and a discriminative model, respectively. The generative model learns the joint probability distribution $P(X,Y)$ of samples and generates new data according to the learned distribution, and the discriminative model learns the conditional probability distribution $P(Y|X)$. A generative model can be used for unsupervised as well as supervised learning.
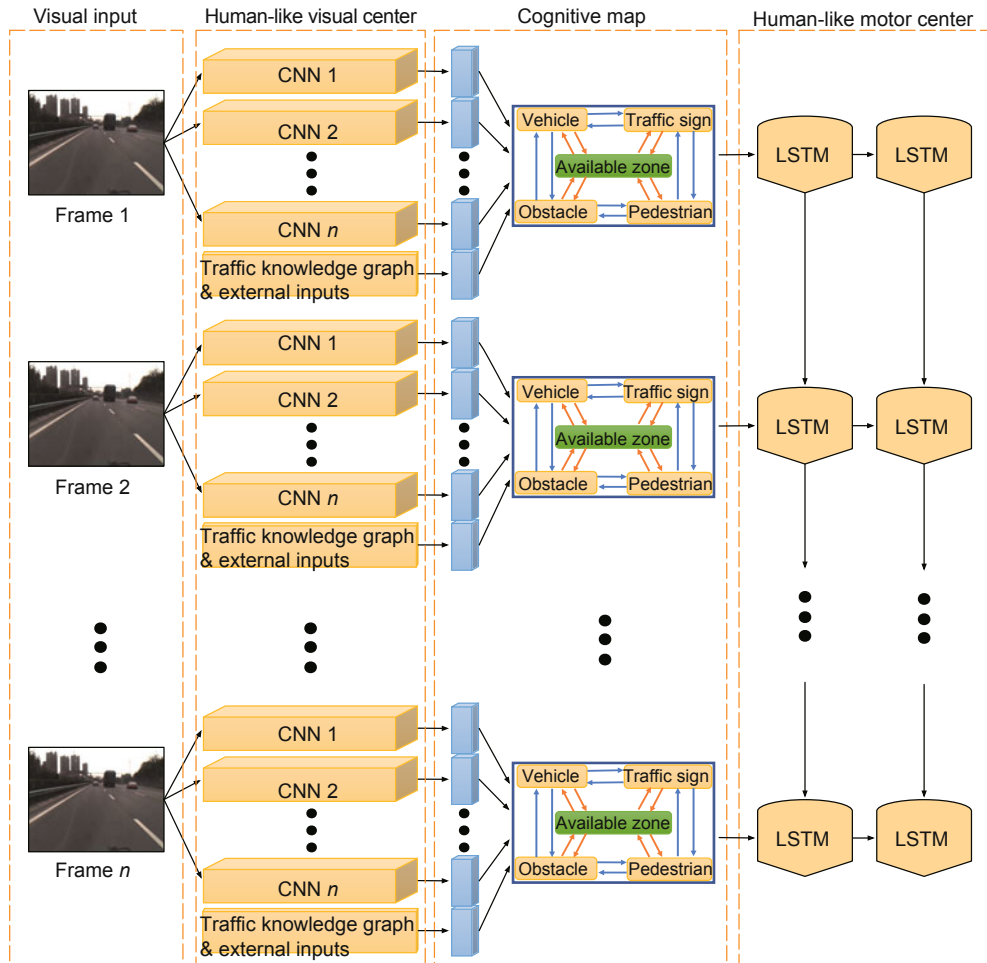
Fig. 16 Hybrid learning model for self-driving cars using architecture with memory and inference

In supervised learning, the conditional probability distribution $P(X|Y)$ is obtained from the joint probability distribution $P(X, Y)$ according to the Bayes formula; hence, many observation models have been constructed, such as the Naive Bayesian model (Lewis, 1998), mixed Gaussian model (Rasmussen, 2000), and Markov model. An unsupervised generative model is to learn the essential characteristics of real data, to give the distribution characteristics of samples, and to generate new data corresponding to the learned probability distribution. In general, the number of parameters of the generative model is far smaller than the size of the training dataset. Thus, the generative model can discover data interdependency and manifest the high-order correlation of the data without labeling information.

Generative adversarial networks (GANs) (Goodfellow *et al.*, 2014b; Denton *et al.*, 2015; Radford *et al.*, 2015) were proposed to promote the training efficiency of the generative model, and to solve the problems that the generative model fails to process. GAN consists mainly of two parts, a generative network used to generate samples and a discriminator used to identify the source of the samples. When new samples generated by the generative network and real-world samples are fed into the discriminator, the discriminator will distinguish the two kinds of samples as accurately as possible. The generative network tries to generate new samples that cannot be discriminated by the discriminator (Mirza and Osindero, 2014; Salimans *et al.*, 2016; van den Oord *et al.*, 2016). Actually, generative adversarial learning is inspired by the zero-sum game from the game theory (Nash, 1950). During the training, the parameters of the generative model and the discriminative model are alternately updated (update one when the other is fixed) to maximize each other's error rate. Thus, the two

parts compete with each other in an unsupervised manner, and ultimately a nearly perfect generative model can be obtained.

4.2 Generative adversarial networks in self-driving cars

Self-driving car is a hotspot of recent AI research. Fig. 17 shows a framework of the generative adversarial model used in unmanned vehicles. There are two critical problems in self-driving technology. One is how to acquire enough training samples, especially negative samples; the other is how to build a vivid off-line test system to verify the performance of unmanned vehicles. Generative adversarial models can be used to generate abundant and more natural scenes for solving these two problems.
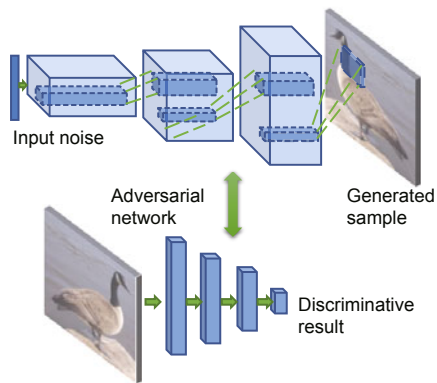

Fig. 17   Generative adversarial networks

First, the construction of a reliable and safe unmanned vehicle system requires to learn a variety of complex road scenes and extreme situations. However, in reality, the collected data cannot cover all of the road conditions. Therefore, complex and vivid scenes with more traffic elements need to be constructed to train a more robust unmanned system online by combining the generative adversarial network with traffic knowledge base and structure of the cognitive map.

Second, off-line test and evaluation require a real-time simulation system that combines the test requirements with the real vehicle status skillfully. As illustrated in Fig. 18, a small number of samples can be used to train a real-time system in unsupervised manner to simulate a variety of road environments by taking advantage of the generative adversarial model. This system can evaluate the performance of the unmanned vehicle through generating

virtual traffic scenes according to the requirements of a real-time simulation environment and the constraints of a road scene.
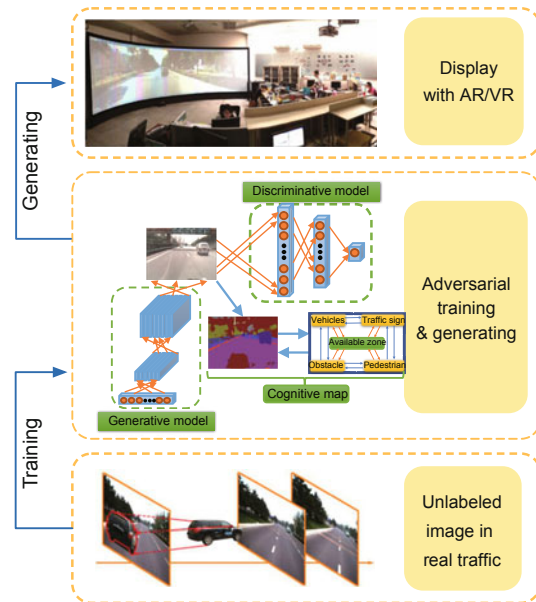

Fig. 18   Generative adversarial networks in self-driving cars

Machine learning systems will no longer rely too much on manually labeled data with models that can generate similar data according to the limited labeled data. Hence, an unsupervised computing architecture, which relies only on a small amount of manually labeled data, can be constructed for efficiently competitive adversarial learning.

5 Typical applications of hybrid-augmented intelligence

AI technology is creating numerous new products and changing the way of people's work, study, and life in almost every aspect. It has become a powerful driving force to promote sustained growth and innovative development of social economy. In this section we introduce some typical applications of hybrid-augmented intelligence.

5.1 Managing industrial complexities and risks

Managing industrial complexities and risks is a typical application of hybrid-augmented intelligence. In the networked era, how to manage the complexity and inherent risks of industry in a modern economic

environment has become a daunting task for many sectors (Shrivastava, 1995). Due to the dynamic nature of the business environment, various industrial environments are facing extensive risks and uncertainties. In addition, the importance of enterprise risk control, business process socialization, business social networks, and the configuration of social technology are promoted extensively because of the advances in our information society and sociocultural environment. The socialization of business is a process of socialization defined, specified, and implemented by an organization for the purpose of achieving implicit or explicit business benefits. The socialization and business social networks require establishing a specific business-driven social structure, or a specific business configuration. They facilitate the flow of information and knowledge (primarily through advanced technologies such as the Internet and AI) and contribute to business intelligence. In particular, the external networks of enterprises not only directly affect the competitiveness of enterprises, but also indirectly affect the competitiveness of enterprises by influencing the internal resources such as total assets and levels of technical expertise (Hu *et al.*, 2010; 2013).

To manage the inherent complexity brought about by society-economy (Wang, 2004), society-technology, and society-politics relationships, a modern process of business social networks and socialization is formed (Fig. 19). In this context, enterprises need innovative solutions to reconstruct different internal organizational functions and operational models, as well as to optimize the scheduling of resources and technology. Innovative solutions depend not only on the ability of decision makers and cognitive conditions (how much information is possessed), but also on the social capabilities based on technology. These capabilities are provided by hybrid-augmented intelligence (Liyanage, 2012; Wang FY *et al.*, 2016), including advanced AI, information and communication technologies (ICTs), social networks, and business networks. This hybrid-augmented intelligence integrates organizational events, technological components, and society to create a human-computer interaction environment where learning, understanding, reasoning, and decision making are supported and core technologies are available. The applications of hybrid-augmented intelligence can greatly improve the risk management capability of modern enterprises, enhance their value creation, and promote competitiveness.

5.2 Collaborative decision-making in enterprises

Collaborative decision-making is critical to almost all businesses (Hoffman, 1998; Ball *et al.*, 2001). The free exchange of ideas in an enterprise is likely to create more innovative products, strategic solutions, and lucrative business decisions (Fjellheim *et al.*, 2008).

Human-computer collaborative hybrid-augmented intelligence can provide application solutions for large-scale workflow coordination, which has great potential in value creation. Fig. 20
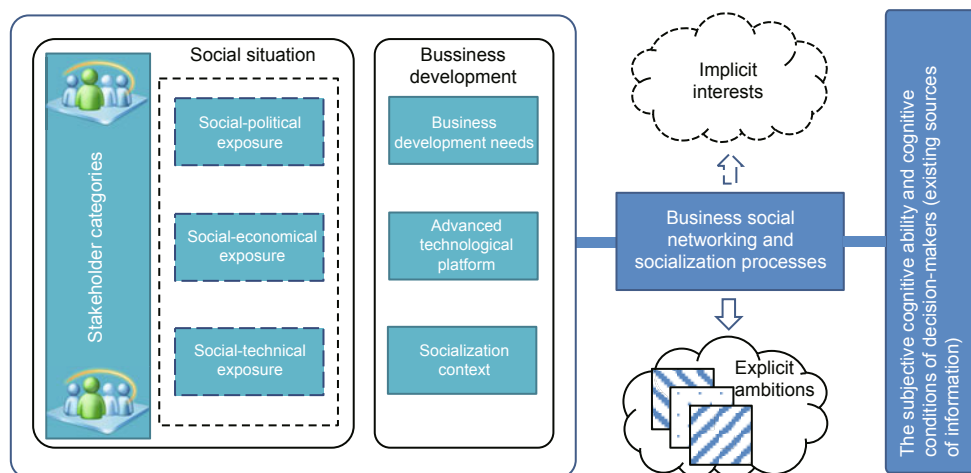

Fig. 19 Modern process of business social networks and socialization (by using hybrid-augmented intelligence to handle complex data and tasks in this process)
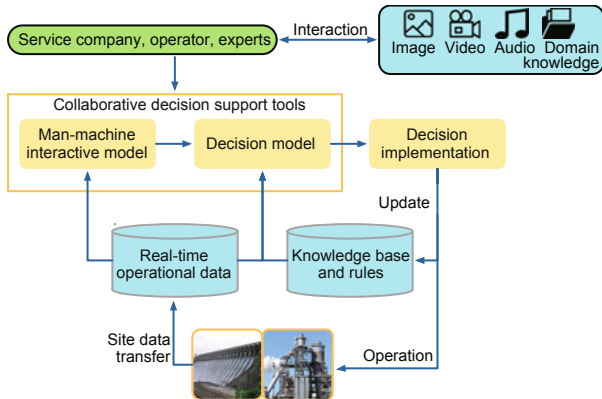
Fig. 20 General framework of hybrid-augmented intelligence for enterprise collaborative decision-making

shows an example of hybrid-augmented intelligence for enterprise collaborative decision-making that supports coordination and communication among participants in the process. The hybrid-augmented intelligence systems of enterprise collaboration decision-making must be accessible to all CEO partners to provide transparency and make it easy to follow workflows at any time. The integration of multiple machine learning methods, decision models, and domain knowledge is critical for hybrid-augmented intelligence systems. That is to say, the integration process is very complicated. In addition, a collaborative application is considered to include an expert system that provides recommendations for an optimal solution through a combination of existing explicit knowledge in the knowledge base,

rule reasoning, and experts' implicit knowledge. Such a collaborative application demands smooth interfaces (decision support, communication, work process compliance, etc.) among different modules. For example, in an application, members are able to discuss and solve problems by communicating and sharing pictures, videos, audio, and other language contexts. During the process of solving the problem, different solutions need to be combined into the decision-making model to obtain the recommended best solution.

5.3 Online intelligence learning

AI makes education traceable and visible. Online learning is another important application of hybrid-augmented intelligence (Yau *et al.*, 2003; Atif and Mathew, 2015). Future education must be personalized, and students will benefit from interacting with an online learning system. As shown in Fig. 21, such an online learning system is based on a hybrid-augmented intelligence system under the framework of CC. The human-computer interaction in online learning is not a simple interface interaction, but the continuous impartation and update of knowledge (Marchiori and Warglien, 2008) between students and machines during the learning process. Online hybrid-augmented intelligence learning system will be designed to provide personalized tutorial according to each student's knowledge structure, intelligence, and proficiency.
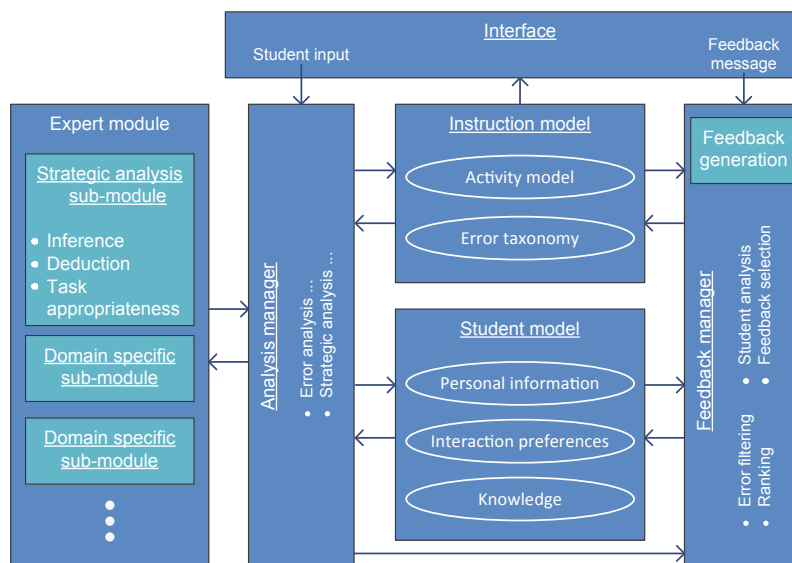

Fig. 21 Online hybrid-augmented intelligent learning system

To provide personalized tutorial, the online hybrid-augmented intelligence learning system can construct a sense-making model dynamically, and plan different learning schemes according to different abilities and responses of learners. The core of the system is to transform traditional education into a customized and personalized learning system, which will profoundly change the formation and dissemination of knowledge.

5.4 Medical and healthcare

In the medical field, a large amount of knowledge and rules need to be memorized, of which most are empirical, complex, and unstructured and have been changing over time. Furthermore, there are complex causal relationships between medical knowledge and rules (Lake *et al.*, 2016). Fig. 22 shows a schematic of various medical relationships among patients, precision medicine, healthcare, diagnosis, and clinical practice. In addition, the 'human disease space' cannot be exhaustively searched. Therefore, it is necessary to establish a medical care oriented hybrid-augmented intelligence system.
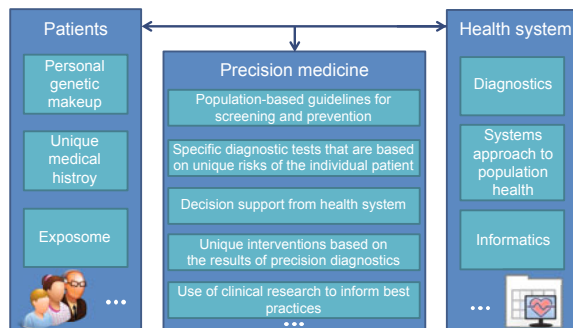

Fig. 22 Precision medical schematic

The medical field is closely related to human life, and a wrong decision is intolerable. So, completely replacing doctors by AI is impossible and unacceptable. At present, the most successful application in medical field is IBM's Watson health system, which is still in rapid development and improvement (Ando *et al.*, 2005; Ando, 2007; Chen Y *et al.*, 2016). For a doctor, the necessary preconditions to become an expert are formal training, reading a large amount of medical literature, rigorous clinical practice, and knowledge accumulation through cases. However, the knowledge and experience accumulated in the whole life of a doctor are still very limited. Mean-

while, the knowledge in each academic field is rapidly increasing; it is impossible for any expert to understand and master all the latest information and knowledge. In contrast to humans, the Watson system can accumulate knowledge easily by memorizing the literature, cases, and rules and by translating a number of doctors' diagnosis about diseases into an improvement in system capability. The Watson system is able to understand natural language, answer questions, and mine patient data and other available data systematically to obtain hypotheses and present them by a confidence score. Then, a doctor can give the ultimate diagnosis according to the information offered by the system. To some extent, AI systems can diagnose individually (Dounias, 2003), but it is difficult to exhaust human diseases by the rules. So, the involvement of doctors is required (Fig. 23a). Integrating doctors' clinical diagnostic process into a medical AI system with powerful storage, searching, and reasoning capabilities (Fig. 23a) can make a better and faster diagnosis. Fig. 23b shows the basic framework of a medical hybrid-augmented intelligence system.

The applications of cognitive medical hybrid-augmented intelligence systems with human-computer interaction, medical imaging, biosensors, and nano surgery will bring a revolutionary change to the medical field.

5.5 Public safety and security

The current public safety and security issues show a complex and diversified development trend, especially in security areas such as national security (Cimbala, 2012), financial security (Hilovska and Koncz, 2012), web security (Shuaibu *et al.*, 2015), public security (Ferreira *et al.*, 2010), and anti-terrorism. Hybrid-augmented intelligence can provide strong technical support and a basic infrastructure framework to meet the increasing challenges in those security areas. Generally, the processing of anomaly events can be divided into three parts: prediction, detection, and subsequent disposition. To make full use of human intelligence in complex problem judging and of AI in processing massive data, security systems should be a human-computer collaborative hybrid-augmented intelligence, that is, humans' participation in prediction, detection, and subsequent disposition. A general framework of the system is given in Fig. 24.
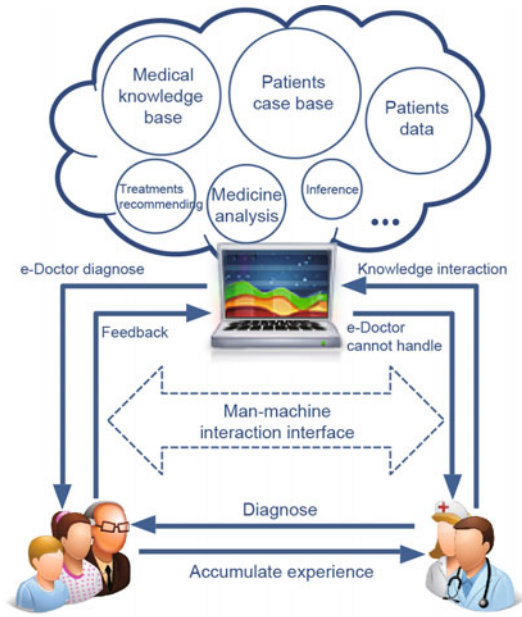
(a)
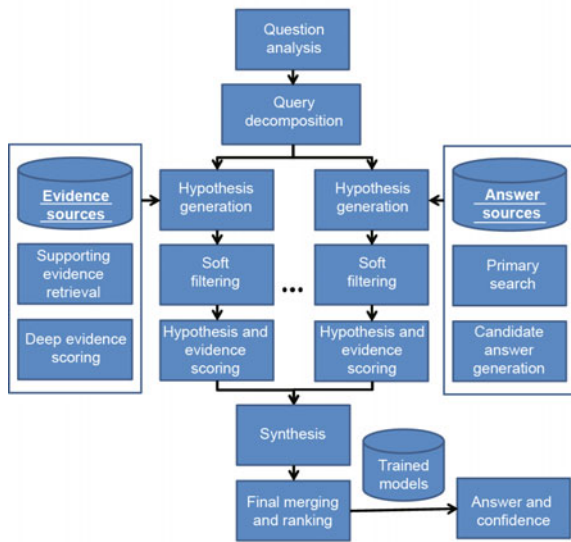
(b)

Fig. 23 Integrating doctors' clinical diagnostic process into a medical AI system (a) and the basic framework of a medical hybrid-augmented intelligence system (b)

A typical example of an anomaly prediction task is sentiment analysis (Zhao *et al.*, 2010). With the development of social networks, analyzing sentiment access by Internet data is possible. Sentiment analysis is an effective means to predict the occurrence of abnormal events (public safety events). Facing massive unstructured Internet data, it is, however, impossible for humans to predict abnormal events
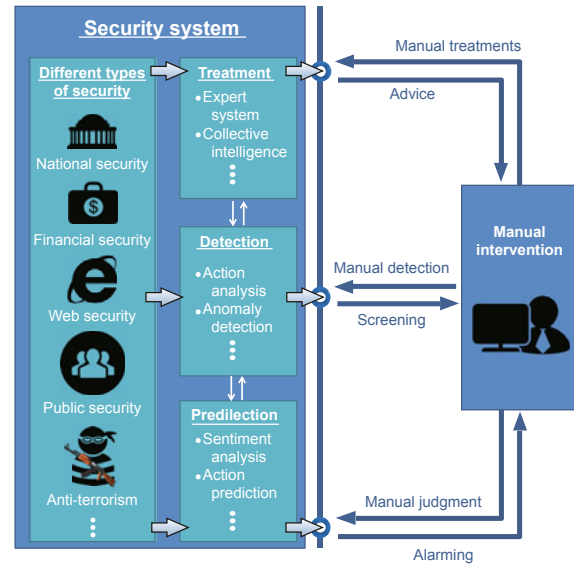
Fig. 24 General framework of hybrid-augmented intelligence for public safety and security

without the aid of AI. This requires the prediction module of the security system to process large-scale data automatically and hand the results to humans, who will make further judgment. In the process of anomaly detection and subsequent disposition, there is a similar interaction mechanism between a human and a security system. Thus, a security system based on human-computer collaborative hybrid-augmented intelligence is formed.

At present, surveillance cameras are deployed almost everywhere, which can provide massive video streams for monitoring public security. Due to the lack of manpower, those videos are not fully used. Hybrid-augmented intelligence based on CC can detect suspect events and characters from massive data (e.g., dangerous carry-on items, anomaly postures, and anomaly crowd behaviors). For results with low confidence or significant impact, experts will get involved and interact with the security system and make further judgments by their intuition and domain knowledge. Meanwhile, a cognition model can leverage experts' feedbacks to improve the analytical ability for video understanding and finally, a better and faster system can be achieved for prediction, detection, and subsequent disposition of anomaly events.

5.6 Human-computer collaborative driving

Automatic driving system (Varaiya, 1993; Waldrop, 2015) is a highly integrated AI system and also

a hotspot of research in recent years. Currently, fully automatic driving is still facing difficult technological challenges. A conception of human-computer collaborative driving was first put forward in the 1960s (Rashevsky, 1964). Along with the development of intelligent transportation systems, 5G communication technologies, and vehicle networking, human-computer collaborative driving has become more and more robust and advanced (Im *et al.*, 2009).

Human-computer collaborative driving refers to the sharing of vehicle control between a driver and the intelligent system. It means accomplishing the driving task cooperatively (Fig. 25). Obviously, this is an HITL human-computer collaborative hybrid-augmented intelligence system where there is a strong complementarity between a human driver and an assisted-driving machine. First of all, humans are of strong robustness and adaptability towards scene understanding, but humans' driving behaviors are easily affected by physical and psychological factors (such as fatigue) (Sharp *et al.*, 2001). Human-computer collaborative driving can reduce the risks of human error and free people from repetitive work. In addition, humans rely mainly on vision for environment perception, which is vulnerable to light, weather, and other factors. The machine assisted-driving system can take advantage of a variety of sensors to achieve continuous monitoring of the driving scenes with high precision, provide more driving information to make up for the lack of human manipulation, and broaden the perception domain. The system is also able to intervene humans' driving behaviors when humans fail to detect danger.
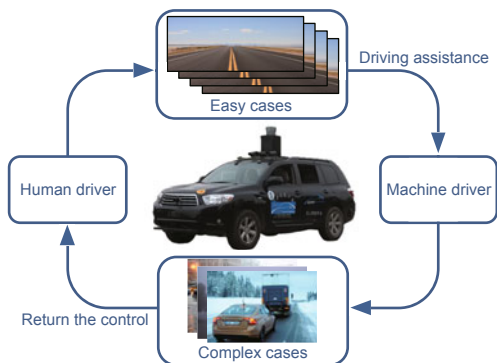

Fig. 25 Human-computer collaborative driving

The key problems of man-machine collaborative driving are how to realize machine perception and judgment, interaction of information in machine and humans' cognition, and decision-making (Saripalli *et al.*, 2003). Therefore, how to coordinate the two 'drivers' to realize safe and comfortable driving of the vehicle is a pressing fundamental problem faced by the hybrid-augmented intelligence man-machine collaborative driving system.

At present, automatic driving has been applied in specific situations, but technical difficulties still exist in public and natural traffic scenes. However, there are still more than 1 billion passenger cars on the road every day. Therefore, it is quite important to solve the current safety problems of passenger cars by man-machine collaborative driving (Zheng *et al.*, 2004). Fig. 26 shows a three-layer architecture for a driver assistance and safety warning system. The sensory layer completes data collection and communication with different types of in-vehicle sensors and roadside devices. The decision-making layer processes the data that the sensor layer collects, extracts valuable information, combines it with the GIS database for real-time decision, and recommends corresponding actions for the human driver. Simultaneously, the actions of the driver are compared with the driver's dangerous actions to make appropriate rational decision-making. The human interface layer displays a variety of guidance information, offers real-time presentations to the driver for the safety of high-level road information, and warns for unreasonable posture actions.
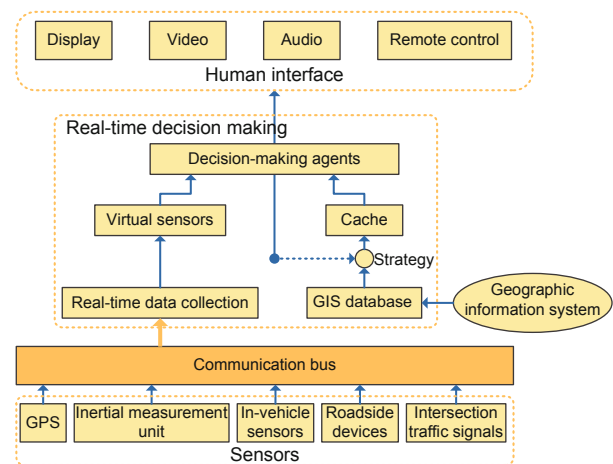

Fig. 26 Architecture for a driver assistance and safety warning system

Man-machine collaborative driving is also able to provide an approach to driving learning for the automatic driving intelligence system. The system

learns the actions of human drivers, including driving behavioral psychology from the process of man-machine collaborative driving.

5.7 Cloud robotics

In recent years, robots have been widely used in industrial manufacturing (Johnson *et al.*, 2014; Schwartz *et al.*, 2016), life services (Walters *et al.*, 2013; Boman and Bartfai, 2015; Hughes *et al.*, 2016), military defense (Gilbert and Beebe, 2010; Barnes *et al.*, 2013), and other fields. However, traditional robots have the problem of simplification of instructions, which makes them difficult to update the knowledge among the robots and hard to interact with humans; so, it is difficult to carry out complex tasks. Therefore, how to enhance the intelligence of an individual in a multi-robot collaborative system is a major challenge for multi-robot collaborative augmented intelligence

Cloud robot is one of the fastest fields of transforming hybrid-augmented intelligence research into commercial applications. An important application of mobile Internet is the Internet of Things (IoT). The concept of IoT is the support of millions of ordinary devices or all the items used in daily life, connected to a mobile Internet cloud. This is a long-term goal that people are pursuing, but this kind of interconnection has already been reflected in the cloud robot field. In these systems, different tasks can be optimized so that different robots can independently cope with specific tasks, and robots can share solutions with each other via the cloud. The robots can share data with each other via the cloud, enabling any robot or intelligent system connected to the same network to analyze the data. For example, if robot *A* sends some knowledge to robot *B*, robot *B* in turn can improve that knowledge and continue to transmit it in a cooperative way, and can realize multi-robot motion planning in the shared space and limited time. Thus, the learning potential and connectivity of the robots are significantly improved. Fig. 27 shows the hybrid-augmented intelligence framework for cloud robot interconnection.

In addition, entertainment is an important application of hybrid-augmented intelligence. In recent years, technologies such as augmented reality and virtual reality have been widely used in game industry, such as Pokemon Go, which enhances humans' participation by superimposing users' real scenes
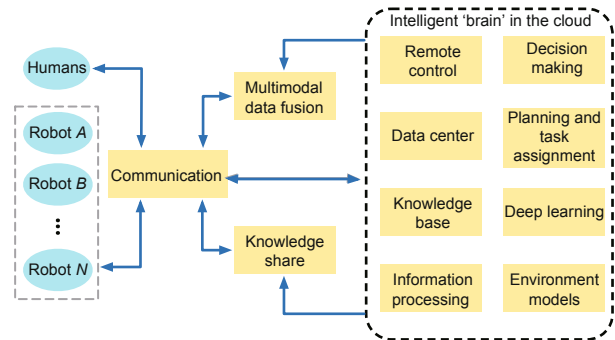

Fig. 27 Hybrid-augmented intelligent framework for cloud robot interconnection

and game virtual scenes, promotes the development of game industry, and contributes to the technological progress. Moreover, social platforms such as Facebook and WeChat, shopping websites, and other entertainment websites push related information to users by making personal preference analysis, which can become more effective and accurate by introducing human-computer collaborative hybrid-augmented intelligence.

6 Conclusions

Intelligence machines have become human companions, and AI is profoundly changing our lives and shaping the future. Ubiquitous computing and intelligence machines are driving people to seek new computational models and implementation forms of AI. Hybrid-augmented intelligence is one of the important directions for the growth of AI.

Building human-computer interaction based HITL hybrid-augmented intelligence by combining perception and cognitive capabilities of humans with the computer's capabilities to calculate and store data can greatly enhance AI system's decision-making capability, the level of cognitive sophistication required to handle complex tasks, and adaptability to complex situations. Hybrid-augmented intelligence based on CC can solve the problems of planning and reasoning that AI research area has been facing for a long time through intuitive reasoning, experience learning, and other hybrid models.

In this survey, the importance of the development of human-computer cooperative hybrid-augmented intelligence and its basic framework are described on the basis of discussing the limitations of existing machine learning methods. The basic problems of hybrid-augmented intelligence based on

CC such as intuitive reasoning, causal modeling, memory, and knowledge evolution are discussed, and the important role and basic approach of intuitive reasoning in complex problem solving are described. The visual scene understanding method based on memory and reasoning is also presented. Finally, typical applications of hybrid-augmented intelligence in the fields of managing industrial complexities and risks, collaborative decision-making in enterprises, online intelligent learning, medical and healthcare, public safety and security, human-computer collaborative driving, and cloud robotics are introduced. We encourage both the industry and academia to investigate and enrich hybrid-augmented intelligence, in both theory and practice.

Acknowledgements

References

Ando, R.K., 2007. Biocreative II gene mention tagging system at IBM Watson. Proc. 2nd BioCreative Challenge Evaluation Workshop, p.101-103.

Ando, R.K., Dredze, M., Zhang, T., 2005. Trec 2005 genomics track experiments at IBM Watson. 14th Text REtrieval Conf., p.1-10.

Atif, Y., Mathew, S.S., 2015. Building a smart campus to support ubiquitous learning. *J. Amb. Intell. Human. Comput.*, **6**(2):1-16.
http://dx.doi.org/10.1007/s12652-014-0226-y

Ball, M.O., Chen, C.Y., Hoffman, R., *et al.*, 2001. Collaborative decision making in air traffic management: current and future research directions. *In*: Bianco, L., Dell'Olmo, P., Odoni, A.R. (Eds.), New Concepts and Methods in Air Traffic Management. Springer Berlin Heidelberg, Berlin, Germany, p.17-30.
http://dx.doi.org/10.1007/978-3-662-04632-6

Barnes, M.J., Chen, J.Y.C., Jentsch, F., *et al.*, 2013. An overview of humans and autonomy for military environments: safety, types of autonomy, agents, and user interfaces. Proc. 10th Int. Conf. on Engineering Psychology and Cognitive Ergonomics: Applications and Services, p.243-252.
https://dx.doi.org/10.1007/978-3-642-39354-9_27

Boman, I.L., Bartfai, A., 2015. The first step in using a robot in brain injury rehabilitation: patients' and health-care professionals' perspective. *Disab. Rehab. Assist. Technol.*, **10**(5):365-370.
http://dx.doi.org/10.3109/17483107.2014.913712

Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Patt. Recogn.*, **30**(7):1145-1159.
http://dx.doi.org/10.1016/S0031-3203(96)00142-2

Browne, C.B., Powley, E., Whitehouse, D., *et al.*, 2012. A survey of Monte Carlo tree search methods. *IEEE Trans. Comput. Intell. AI Games*, **4**(1):1-43.
http://dx.doi.org/10.1109/TCIAIG.2012.2186810

Campbell, M., Hoane, A.J.Jr., Hsu, F.H., 2002. Deep Blue. *Artif. Intell.*, **134**(1-2):57-83.
http://dx.doi.org/10.1016/S0004-3702(01)00129-1

Chen, D., Yuan, Z., Hua, G., *et al.*, 2016. Multi-timescale collaborative tracking. *IEEE Trans. Patt. Anal. Mach. Intell.*, **39**(1):141-155.
http://dx.doi.org/10.1109/TPAMI.2016.2539956

Chen, Y., Argentinis, J.D.E., Weber, G., 2016. IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research. *Clin. Therap.*, **38**(4):688-701.
http://dx.doi.org/10.1016/j.clinthera.2015.12.001

Cimbala, S.J., 2012. Artificial Intelligence and National Security. Lexington Books, Lanham, USA.

Denton, E.L., Chintala, S., Fergus, R., *et al.*, 2015. Deep generative image models using a Laplacian pyramid of adversarial networks. Proc. 28th Int. Conf. on Neural Information Processing Systems, p.1486-1494.

de Rocquigny, E., Nicolas, D., Stefano, T., 2008. Uncertainty in Industrial Practice: a Guide to Quantitative Uncertainty Management. John Wiley & Sons, Hoboken, USA.

Dias, M.G., Harris, P., 1988. The effect of make-believe play on deductive reasoning. *Br. J. Devel. Psychol.*, **6**(3):207-221.
http://dx.doi.org/10.1111/j.2044-835X.1988.tb01095.x

Dounias, G., 2003. Hybrid computational intelligence in medicine. Proc. Workshop on Intelligent and Adaptive Systems in Medicine.

Eakin, H., Luers, A.L., 2006. Assessing the vulnerability of social-environmental systems. *Ann. Rev. Environ. Resourc.*, **31**:1-477. http://dx.doi.org/10.1146/annurev.energy.30.050504.144352

Ferreira, F.J., Crispim, V.R., Silva, A.X., 2010. Detection of drugs and explosives using neutron computerized tomography and artificial intelligence techniques. *Appl. Rad. Isot.*, **68**(6):1012-1017.
http://dx.doi.org/10.1016/j.apradiso.2010.01.019

Fire, A., Zhu, S.C., 2016. Learning perceptual causality from video. *ACM Trans. Intell. Syst. Technol.*, **7**(2):1-22.
http://dx.doi.org/10.1145/2809782

Fischbein, H., 2002. Intuition in Science and Mathematics: an Educational Approach. Springer Science & Business Media, Berlin, Germany.

Fjellheim, R., Bratvold, R.R., Herbert, M.C., 2008. CODIO - collaborative decisionmaking in integrated operations. Intelligent Energy Conf. and Exhibition, p.1-7.
http://dx.doi.org/10.2118/111876-MS

Fogel, D.B., 1995. Evolutionary Computation: Toward a New Philosophy of Machine Intelligence. Wiley-IEEE Press.

Freyd, J.J., 1983. Representing the dynamics of a static form. *Memory Cogn.*, **11**(4):342-346.
http://dx.doi.org/10.3758/BF03202447

Funahashi, K.I., Nakamura, Y., 1993. Approximation of dynamic systems by continuous-time recurrent neural networks. *Neur. Netw.*, **6**(6):801-806.
http://dx.doi.org/10.1016/S0893-6080(05)80125-X

Gil, Y., Greaves, M., Hendler, J., *et al.*, 2014. Amplify scientific discovery with artificial intelligence. *Science*, **346**(6206):171-172.
http://dx.doi.org/10.1126/science.1259439

Gilbert, G.R., Beebe, M.K., 2010. United States Department of Defense Research in Robotic Unmanned Systems for Combat Casualty Care. Report No. RTO-MP-HFM-182, Fort Detrick, Frederick, USA.

Goodfellow, I.J., Shlens, J., Szegedy, C., 2014a. Explaining and harnessing adversarial examples. ePrint Archive, arXiv:1412.6572.

Goodfellow, I.J., Pougetabadie, J., Mirza, M., *et al.*, 2014b. Generative adversarial nets. Advances in Neural Information Processing Systems, p.2672-2680.

Graves, A., Mohamed, A.R., Hinton, G., 2013. Speech recognition with deep recurrent neural networks. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, p.6645-6649.
http://dx.doi.org/10.1109/ICASSP.2013.6638947

Graves, A., Wayne, G., Danihelka, I., 2014. Neural turing machines. ePrint Archive, arXiv:1410.5401.

Graves, A., Wayne, G., Reynolds, M., *et al.*, 2016. Hybrid computing using a neural network with dynamic external memory. *Nature*, **538**(7626):471-476.
http://dx.doi.org/10.1038/nature20101

Griffiths, T.L., Chater, N., Kemp, C., *et al.*, 2010. Probabilistic models of cognition: exploring representations and inductive biases. *Trends Cogn. Sci.*, **14**(8):357-364.
http://dx.doi.org/10.1016/j.tics.2010.05.004

Guilford, J.P., 1967. The Nature of Human Intelligence. McGraw-Hill, New York, USA.

Hagan, M.T., Demuth, H.B., Beale, M.H., *et al.*, 2002. Neural Network Design. PWS Publishing Co., Boston, USA.

Hilovska, K., Koncz, P., 2012. Application of artificial intelligence and data mining techniques to financial markets. *ACTA VSFS*, **6**:62-76.

Hiskens, I.A., Davy, R.J., 2001. Exploring the power flow solution space boundary. *IEEE Trans. Power Syst.*, **16**(3):389-395.
http://dx.doi.org/10.1109/59.932273

Hoffman, R., 1998. Integer Programming Models for Ground-Holding in Air Traffic Flow Management. PhD Thesis, University of Maryland, College Park, USA.

Holland, J.H., 1992. Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control and Artificial Intelligence. MIT Press.

Honey, C.J., Thivierge, J.P., Sporns, O., 2010. Can structure predict function in the human brain? *NeuroImage*, **52**(3):766-776.
http://dx.doi.org/10.1016/j.neuroimage.2010.01.071

Hu, P., Zhou, S., Ding, W.Z., *et al.*, 2010. The comprehensive measurement model of the member importance in social networks. Int. Conf. on Management and Service Science, p.1-4.
http://dx.doi.org/10.1109/ICMSS.2010.5577405

Hu, P., Wen, C.L., Pan, D., 2013. The mutual relationship among external network, internal resource, and competitiveness of enterprises. *Sci. Res. Manag.*, **V**(4):90-98 (in Chinese).

Hughes, D., Camp, C., O'Hara, J., *et al.*, 2016. Health resource use following robot-assisted surgery versus open and conventional laparoscopic techniques in oncology: analysis of English secondary care data for radical prostatectomy and partial nephrectomy. *BJU Int.*, **117**(6):940-947.

Im, D.Y., Ryoo, Y.J., Kim, D.Y., *et al.*, 2009. Unmanned driving of intelligent robotic vehicle. ISIS Proc. 10th Symp. on Advanced Intelligent Systems, p.213-216.

Ioffe, A.D., 1979. Necessary and sufficient conditions for a local minimum. 3: second order conditions and augmented duality. *SIAM J. Contr. Opt.*, **17**(2):266-288.
http://dx.doi.org/10.1137/0317021

Janis, I.L., Mann, L., 1977. Decision Making: a Psychological Analysis of Conflict, Choice, and Commitment. Free Press, New York, USA.

Jennings, N.R., 2000. On agent-based software engineering artificial intelligence. *Artif. Intell.*, **117**(2):277-296.

Johnson, M., Bradshaw, J.M., Feltovich, P.J., *et al.*, 2014. Coactive design: designing support for interdependence in joint activity. *Electr. Eng. Math. Comput. Sci.*, **3**(1):43-49.
http://dx.doi.org/10.5898/JHRI.3.1.Johnson

Johnson, S., Slaughter, V., Carey, S., 1998. Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Devel. Sci.*, **1**(2):233-238.
http://dx.doi.org/10.1111/1467-7687.00036

Jordan, M.I., 2016. On computational thinking, inferential thinking and data science. Proc. 28th ACM Symp. on Parallelism in Algorithms and Architectures, p.47.
http://dx.doi.org/10.1145/2935764.2935826

Kourtzi, Z., Kanwisher, N., 2000. Activation in human MT/MST by static images with implied motion. *J. Cogn. Neurosci.*, **12**(1):48-55.
http://dx.doi.org/10.1162/08989290051137594

Lake, B.M., Salakhutdinov, R., Tenenbaum, J.B., 2015. Human-level concept learning through probabilistic program induction. *Science*, **350**(6266):1332-1338.
http://dx.doi.org/10.1126/science.aab3050

Lake, B.M., Ullman, T.D., Tenenbaum, J.B., *et al.*, 2016. Building machines that learn and think like people. *Behav. Brain Sci.*, **22**:1-101.

Ledford, H., 2015. How to solve the world's biggest problems. *Nature*, **525**:308-311.
http://dx.doi.org/10.1038/525308a

Lewis, D.D., 1998. Naive (Bayes) at forty: the independence assumption in information retrieval. European Conf. on Machine Learning, p.4-15.
http://dx.doi.org/10.1007/BFb0026666

Lillicrap, T.P., Hunt, J.J., Pritzel, A., *et al.*, 2016. Continuous control with deep reinforcement learning. ePrint Archive, arXiv:1509.02971.

Lippmann, R.P., 1987. An introduction to computing with neural nets. *IEEE ASSP Mag.*, **4**(2):4-22.
http://dx.doi.org/10.1109/MASSP.1987.1165576

Liyanage, J.P., 2012. Hybrid Intelligence Through Business Socialization and Networking: Managing Complexities in the Digital Era. IGI Global, Hershey, USA.

Marchiori, D., Warglien, M., 2008. Predicting human interactive learning by regret-driven neural networks. *Science*,

319(5866):1111-1113.
http://dx.doi.org/10.1126/science.1151185

Marr, D., 1977. Artificial intelligence—a personal view. *Artif. Intell.*, **9**(1):37-48.
http://dx.doi.org/10.1016/0004-3702(77)90013-3

Martin, J., 2007. The Meaning of the 21st Century: a Vital Blueprint for Ensuring Our Future. Random House.

McCarthy, J., Hayes, P.J., 1987. Some Philosophical Problems from the Standpoint of Artificial Intelligence. Morgan Kaufmann Publishers Inc., Burlington, USA.

Michalski, R.S., Carbonell, J.G., Mitchell, T.M., 1984. Machine Learning: an Artificial Intelligence Approach. Springer Science & Business Media, Berlin, Germany.

Mikolov, T., Karafiát, M., Burget, L., *et al.*, 2010. Recurrent neural network based language model. Conf. of the Int. Speech Communication Association, p.1045-1048.

Minsky, M., 1961. Steps toward artificial intelligence. *Proc. IRE*, **49**(1):8-30.
http://dx.doi.org/10.1109/JRPROC.1961.287775

Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. ePrint Archive, arXiv:1411.1784.

Mizumoto, M., 1982. Comparison of fuzzy reasoning methods. *Fuzzy Sets Syst.*, **8**(3):253-283.
http://dx.doi.org/10.1016/S0165-0114(82)80004-3

Mnih, V., Kavukcuoglu, K., Silver, D., *et al.*, 2013. Playing Atari with deep reinforcement learning. ePrint Archive, arXiv:1312.5602.

Mnih, V., Kavukcuoglu, K., Silver, D., *et al.*, 2015. Human-level control through deep reinforcement learning. *Nature*, **518**(7540):529-533.
http://dx.doi.org/10.1038/nature14236

Moran, J., Desimone, R., 1985. Selective Attention Gates Visual Processing in the Extrastriate Cortex. MIT Press, Cambridge, USA.

Muir, B.M., 1994. Trust in automation: part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics*, **37**(11):1905-1922.
http://dx.doi.org/10.1080/00140139408964957

Nash, J.F., 1950. Equilibrium points in *n*-person games. *PNAS*, **36**(1):48-49.
http://dx.doi.org/10.1073/pnas.36.1.48

Navigli, R., Ponzetto, S.P., 2012. Babelnet: the automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artif. Intell.*, **193**(6):217-250.
http://dx.doi.org/10.1016/j.artint.2012.07.001

Newell, A., Simon, H.A., 1972. Human Problem Solving. Prentice-Hall, Englewood Cliffs, USA.

Nilsson, N.J., 1965. Learning Machines: Foundations of Trainable Pattern-Classifying Systems. McGraw-Hill, New York, USA.

Nissen, M.J., Bullemer, P., 1987. Attentional requirements of learning: evidence from performance measures. *Cogn. Psychol.*, **19**(1):1-32.
http://dx.doi.org/10.1016/0010-0285(87)90002-8

Noh, H., Hong, S., Han, B., 2015. Learning deconvolution network for semantic segmentation. IEEE Int. Conf. on Computer Vision, p.1520-1528.

Norman, K.A., O'Reilly, R.C., 2003. Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol. Rev.*, **110**(4):611-646.
http://dx.doi.org/10.1037/0033-295X.110.4.611

Ogura, T., Yamada, J., Yamada, S.I., *et al.*, 1989. A 20 kbit associative memory lSI for artificial intelligence machines. *IEEE J. Sol.-State Circ.*, **24**(4):1014-1020.
http://dx.doi.org/10.1109/4.34086

O'Keefe, J., Nadel, L., 1978. The Hippocampus as a Cognitive Map. Clarendon Press, Oxford.

O'Leary, D.E., 2013. Artificial intelligence and big data. *IEEE Intell. Syst.*, **28**(2):96-99.
http://dx.doi.org/10.1109/MIS.2013.39

Pan, Y.H., 2016. Heading toward artificial intelligence 2.0. *Engineering*, **2**(4):409-413.
http://dx.doi.org/10.1016/J.ENG.2016.04.018

Park, C.C., Kim, G., 2015. Expressing an image stream with a sequence of natural sentences. Advances in Neural Information Processing Systems, p.73-81.

Poole, D., Mackworth, A., Goebel, R., 1997. Computational Intelligence: a Logical Approach. Oxford University Press, Oxford, UK.

Premack, D., Premack, A.J., 1997. Infants attribute value to the goal-directed actions of self-propelled objects. *J. Cogn. Neurosci.*, **9**(6):848-856.
http://dx.doi.org/10.1162/jocn.1997.9.6.848

Pylyshyn, Z.W., 1984. Computation and Cognition: Toward a Foundation for Cognitive Science. The MIT Press, Cambridge, Massachusetts, USA.

Rachlin, H., 2012. Making IBM's computer, Watson, human. *Behav. Anal.*, **35**(1):1-16.

Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. ePrint Archive, arXiv:1511.06434.

Rashevsky, N., 1964. Man-machine interaction in automobile driving. *Prog. Biocybern.*, **42**:188-200.

Rasmussen, C.E., 2000. The infinite Gaussian mixture model. Advances in Neural Information Processing Systems, p.554-560.

Rehder, B., Hastie, R., 2001. Causal knowledge and categories: the effects of causal beliefs on categorization, induction, and similarity. *J. Exp. Psychol.*, **130**(3):323-360. http://dx.doi.org/10.1037/0096-3445.130.3.323

Russell, S.J., Norvig, P., 1995. Artificial Intelligence: a Modern Approach. Prentice Hall, Englewood Cliffs, USA.

Salimans, T., Goodfellow, I., Zaremba, W., *et al.*, 2016. Improved techniques for training gans. Advances in Neural Information Processing Systems, p.2226-2234.

Salvi, C., Bricolo, E., Kounios, J., *et al.*, 2016. Insight solutions are correct more often than analytic solutions. *Think. Reason.*, **22**(4):443-460.
http://dx.doi.org/10.1080/13546783.2016.1141798

Samuel, A.L., 1988. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.*, **44**(1-2):206-226. http://dx.doi.org/10.1147/rd.441.0206

Saripalli, S., Montgomery, J.F., Sukhatme, G., 2003. Visually guided landing of an unmanned aerial vehicle. *IEEE Trans. Robot. Autom.*, **19**(3):371-380.
http://dx.doi.org/10.1109/TRA.2003.810239

Saxe, R., Carey, S., 2006. The perception of causality in infancy. *ACTA Psychol.*, **123**(1-2):144-165. http://dx.doi.org/10.1016/j.actpsy.2006.05.005

Schlottmann, A., Ray, E.D., Mitchell, A., *et al.*, 2006. Perceived physical and social causality in animated motions: spontaneous reports and ratings. *ACTA Psychol.*, **123**(1-2):112-143. http://dx.doi.org/10.1016/j.actpsy.2006.05.006

Schwartz, T., Zinnikus, I., Krieger, H.U., *et al.*, 2016. Hybrid teams: flexible collaboration between humans, robots and virtual agents. German Conf. on Multiagent System Technologies, p.131-146.

Selfridge, O.G., 1988. Pandemonium: a paradigm for learning. National Physical Laboratory Conf., p.511-531.

Shader, R.I., 2016. Some reflections on IBM Watson and on women's health. *Clin. Therap.*, **38**(1):1-2. http://dx.doi.org/10.1016/j.clinthera.2015.12.008

Sharp, C.S., Shakernia, O., Sastry, S.S., 2001. A vision system for landing an unmanned aerial vehicle. IEEE Int. Conf. on Robotics & Automation, p.1720-1727. http://dx.doi.org/10.1109/ROBOT.2001.932859

Shrivastava, P., 1995. Ecocentric management for a risk society. *Acad. Manag. Rev.*, **20**(1):118-137. http://dx.doi.org/10.5465/AMR.1995.9503271996

Shuaibu, B.M., Norwawi, N.M., Selamat, M.H., *et al.*, 2015. Systematic review of Web application security development model. *Artif. Intell. Rev.*, **43**(2):259-276. http://dx.doi.org/10.1007/s10462-012-9375-6

Silver, D., Huang, A., Maddison, C.J., *et al.*, 2016. Mastering the game of Go with deep neural networks and tree search. *Nature*, **529**(7587):484-489. http://dx.doi.org/10.1038/nature16961

Simon, H.A., 1969. The Sciences of the Artificial. MIT Press, Cambridge, USA.

Son, D., Lee, J., Qiao, S., *et al.*, 2014. Multifunctional wearable devices for diagnosis and therapy of movement disorders. *Nat. Nanotechnol.*, **9**(5):397-404. http://dx.doi.org/10.1038/nnano.2014.38

Sternberg, R.J., 1984. Beyond IQ: a triarchic theory of human intelligence. *Br. J. Educat. Stud.*, **7**(2):269-287.

Sternberg, R.J., Davidson, J.E., 1983. Insight in the gifted. *Educat. Psychol.*, **18**(1):51-57. http://dx.doi.org/10.1080/00461528309529261

Stone, P., Brooks, R., Brynjolfsson, E., *et al.*, 2016. Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, USA.

Sun, Y., Wang, X.G., Tang, X.O., 2014. Deep learning face representation from predicting 10,000 classes. IEEE Conf. on Computer Vision and Pattern Recognition, p.1891-1898. http://dx.doi.org/10.1109/CVPR.2014.244

Szegedy, C., Zaremba, W., Sutskever, I., *et al.*, 2013. Intriguing properties of neural networks. ePrint Archive, arXiv:1312.6199.

Szolovits, P., Patil, R.S., Schwartz, W.B., 1988. Artificial intelligence in medical diagnosis. *Ann. Int. Med.*, **108**(1):80-87. http://dx.doi.org/10.1059/0003-4819-108-1-80

Tenenbaum, J.B., Kemp, C., Griffiths, T.L., *et al.*, 2011. How to grow a mind: statistics, structure, and abstraction. *Science*, **331**(6022):1279-1285. http://dx.doi.org/10.1126/science.1192788

Thielscher, M., 1997. Ramification and causality. *Artif. Intell.*, **89**(1-2):317-364. http://dx.doi.org/10.1016/S0004-3702(96)00033-1

Thielscher, M., 2001. The qualification problem: a solution to the problem of anomalous models. *Artif. Intell.*, **131**(1-2):1-37. http://dx.doi.org/10.1016/S0004-3702(01)00131-X

Thrun, S., Burgard, W., Fox, D., 1998. A probabilistic approach to concurrent mapping and localization for mobile robots. *Mach. Learn.*, **5**(3):253-271. http://dx.doi.org/10.1023/A:1008806205438

Tolman, E.C., 1948. Cognitive maps in rats and men. *Psychol. Rev.*, **55**(4):189-208. http://dx.doi.org/10.1037/h0061626

Tremoulet, P.D., Feldman, J., 2000. Perception of animacy from the motion of a single object. *Perception*, **29**(8):943-951. http://dx.doi.org/10.1068/p3101

Tversky, A., Kahneman, D., 1983. Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychol. Rev.*, **90**(4):293-315. http://dx.doi.org/10.1037/0033-295X.90.4.293

van den Oord, A., Kalchbrenner, N., Kavukcuoglu, K., 2016. Pixel recurrent neural networks. ePrint Archive, arXiv:1601.06759.

Varaiya, P., 1993. Smart car on smart roads: problems of control. *IEEE Trans. Autom. Contr.*, **38**(2):195-207. http://dx.doi.org/10.1109/9.250509

Waldrop, M.M., 2015. Autonomous vehicles: no drivers required. *Nature*, **518**(7537):20-23. http://dx.doi.org/10.1038/518020a

Walters, M.L., Koay, K.L., Syrdal, D.S., *et al.*, 2013. Companion robots for elderly people: using theatre to investigate potential users' views. IEEE Ro-Man, p.691-696. http://dx.doi.org/10.1109/ROMAN.2013.6628393

Wang, F.Y., 2004. Artificial societies, computational experiments, and parallel systems: a discussion on computational theory of complex social-economic systems. *Compl. Syst. Compl. Sci.*, **1**(4):25-35.

Wang, F.Y., Wang, X., Li, L.X., *et al.*, 2016. Steps toward parallel intelligence. *IEEE/CAA J. Autom. Sin.*, **3**(4):345-348. http://dx.doi.org/10.1109/JAS.2016.7510067

Wang, J.J., Ma, Y.Q., Chen, S.T., *et al.*, 2017. Fragmentation knowledge processing and networked artificial. *Seieat. Sin. Inform.*, **47**(1):1-22.

Wang, L.M., Xiong, Y.J., Wang, Z., *et al.*, 2016. Temporal segment networks: towards good practices for deep action recognition. *LNCS*, **9912**:20-36. http://dx.doi.org/10.1007/978-3-319-46484-8_2

Wei, P., Zheng, N.N., Zhao, Y.B., *et al.*, 2013. Concurrent action detection with structural prediction. IEEE Int. Conf. on Computer Vision, p.3136-3143. http://dx.doi.org/10.1109/ICCV.2013.389

Wei, P., Zhao, Y., Zheng, N., *et al.*, 2016. Modeling 4D human-object interactions for joint event segmentation, recognition, and object localization. *IEEE Trans.*

Softw. Eng.
http://dx.doi.org/10.1109/TPAMI.2016.2574712

Williams, R.J., Zipser, D., 1989.   A learning algorithm for continually running fully recurrent neural networks. *Neur. Comput.*, **1**(2):270-280.
http://dx.doi.org/10.1162/neco.1989.1.2.270

Williams, W.M., Sternberg, R.J., 1988. Group intelligence: why some groups are better than others. *Intelligence*, **12**(4):351-377.
http://dx.doi.org/10.1016/0160-2896(88)90002-5

Xiao, C.Y., Dymetman, M., Gardent, C., 2016. Sequence-based structured prediction for semantic parsing. Meeting of the Association for Computational Linguistics, p.1341-1350.

Yau, S.S., Gupta, S.K.S., Karim, F., *et al.*, 2003.   Smart classroom: enhancing collaborative learning using pervasive computing technology.   ASEE Annual Conf. and Exposition, p.13633-13642.

Yegnanarayana, B., 1994.   Artificial neural networks for pattern recognition. *Sadhana*, **19**(2):189-238.
http://dx.doi.org/10.1007/BF02811896

Youseff, L., Butrico, M., da Silva, D., 2008.   Toward a unified ontology of cloud computing. Grid Computing Environments Workshop, p.1-10.
http://dx.doi.org/10.1109/GCE.2008.4738443

Zadeh, L.A., 1996.   Fuzzy logic and approximate reasoning. *In*: Advances in Fuzzy Systems - Applications and Theory: Volume 6. Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems. World Scientific Publishing, Singapore, p.238-259. http://dx.doi.org/10.1142/9789814261302_0016

Zhao, Y.Y., Qin, B., Liu, T., 2010. Sentiment analysis. *J. Softw.*, **21**(8):1834-1848.

Zheng, N.N., Tang, S.M., Cheng, H., *et al.*, 2004. Toward intelligent driver-assistance and safety warning systems. *IEEE Intell. Syst.*, **19**(2):8-11.
http://dx.doi.org/10.1109/MIS.2004.1274904