Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

# Mutually trustworthy human-machine knowledge automation and hybrid augmented intelligence: mechanisms and applications of cognition, management, and control for complex systems*

Fei-Yue WANG[1], Jianbo GUO[2], Guangquan BU[3], Jun Jason ZHANG[†‡4]

[1]*The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China*

[2]*State Grid Cooperation of China, Beijing 100031, China*

[3]*China Electric Power Research Institute, Beijing 100192, China*

[4]*School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China*

[†]E-mail: jun.zhang.ee@whu.edu.cn

Received Sept. 2, 2021; Revision accepted Oct. 3, 2021; Crosschecked Jan. 5, 2022; Published online June 28, 2022

**Abstract:** In this paper, we aim to illustrate the concept of mutually trustworthy human-machine knowledge automation (HM-KA) as the technical mechanism of hybrid augmented intelligence (HAI) based complex system cognition, management, and control (CMC). We describe the historical development of complex system science and analyze the limitations of human intelligence and machine intelligence. The need for using human-machine HAI in complex systems is then explained in detail. The concept of "mutually trustworthy HM-KA" mechanism is proposed to tackle the CMC challenge, and its technical procedure and pathway are demonstrated using an example of corrective control in bulk power grid dispatch. It is expected that the proposed mutually trustworthy HM-KA concept can provide a novel and canonical mechanism and benefit real-world practices of complex system CMC.

**Key words:** Complex systems; Human-machine knowledge automation; Parallel systems; Bulk power grid dispatch; Artificial intelligence; Internet of Minds (IoM)

https://doi.org/10.1631/FITEE.2100418 **CLC number:** TP11

## 1 Introduction

### 1.1 Review of complex system science

The science of complex systems is the interdisciplinary fusion of complexity science and system science. In general, a complex system has the following characteristics: openness (the system continuously exchanges matter, energy, and information with its ambient environment); component complexity (the system has a complex, hierarchical architecture and operational process, and the system components are inter-connected, interactive, and heterogeneous); emergence (the system may develop new and unforeseen characteristics in the process of its operation); self-organization and adaptivity (when exchanges occur between the system and its environment, the system can adjust its own architecture, functionalities, and behaviors to actively form the capability of self-learning and self-adaptivity). As a result, the research objects of complex systems include system functionalities, behaviors, and their relations, and system characteristics of emergence, self-organization, autonomy, and

‡ Corresponding author

* Project supported by the National Key R&D Program of China (No. 2018AAA0101504) and the Science and Technology Project of the State Grid Corporation of China: Fundamental Theory of Human in-the-Loop Hybrid-Augmented Intelligence for Power Grid Dispatch and Control

ORCID: Fei-Yue WANG, https://orcid.org/0000-0001-9185-3989; Jun Jason ZHANG, https://orcid.org/0000-0001-6908-2671

self-evolution. The theory, approaches, and techniques of complex systems are applied with profound impact in the areas of managing and controlling real-world complex systems, such as living, aerospace, Earth systems, social organization, economic, military, and complex engineering systems.

The development of complex system science has a history going back over 100 years and roughly experienced six eras. It was represented by the evolutionary theory and statistical physics from the 1930s to the 1950s, by the system theory (von Bertalanffy, 1968), cybernetics (Wiener, 1948), and informatics (Shannon, 1948) from the 1950s to the 1960s, by the dissipative structure theory (Prigogine and Lefever, 1973), catastrophe theory (Thom, 1975), and synergetics (Haken, 1977) from the 1970s to the 1980s, by the chaos (Lorenz, 1963), fractal genometry (Mandelbrot, 1989), and critical phenomena theories (Binney et al., 1992) from the 1980s to the 1990s, and by the complex adaptive system theory (Holland, 1995), multi-agent simulation theory, and artificial life and society theory from the 1990s to the 2000s, providing the main contributions to the science. After 2000, with the rapid development of data science and artificial intelligence (AI), research of complex systems has entered the era of "data science," represented by the complex system analysis by big data and AI (Hey et al., 2009), complex network theory (Manoj et al., 2018), parallel system theory, and the Artificial systems, Computational experiments, and Parallel execution (ACP) approach as shown in Fig. 1 (Wang, 2004).

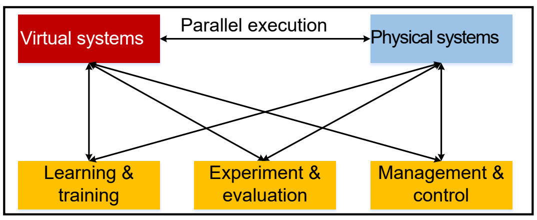


**Fig. 1 The parallel system theory and the Artificial systems, Computational experiments, and Parallel execution (ACP) approach**

A large number of complex system modeling, simulation, management, and control approaches have been developed (Thurner et al., 2018). These include mechanism, statistical and regression analysis, analytic hierarchical decomposition methods, Monte-Carlo analysis, grey system analysis, fuzzy sets, and artificial neural networks recently. Popular approaches include composite modeling, hybrid modeling, agent-based modeling, Petri-net-based modeling, Markov process based models, bootstrap-based models, cellular automata, complex adaptive systems, complex self-organizing systems, and fractal theory. The control approaches in complex systems also experienced multiple phases of development, including synthetic multi-modal linear control (which uses multiple system modes and the corresponding linear controller to cope with complex scenarios) and nonlinear control (which suffers from modeling limitations). Recently, intelligent control becomes a mainstream approach for complex systems, including fuzzy control, adaptive dynamic programming, expert systems, and knowledge-aided intelligent control (Miller and Page, 2007; Liu et al., 2008; Sayama, 2015; Manoj et al., 2018; Thurner et al., 2018; Wang and Chen, 2020).

### 1.2 Complex systems and human-machine hybrid augmented intelligence (HM-HAI)

The transition of complex system cognition, management, and control (CMC) theory and approaches into big data and AI-based technology is in progress. The new generation of information technology brings new opportunities for complex system research and application. However, its limitations in applicability are also fully revealed:

1. Machine learning models of the new generation of AI, represented by deep neural networks, rely mostly on training by big data. As a result, a large volume of unbiased numerical or semantical training data is essential for AI-based complex system modeling and analysis. Most machine learning models rely on labeled data for supervised or semi-supervised learning. As a result, a huge amount of work on data labeling is necessary to train usable AI models. This creates a major barrier for AI applicability. In fact, for real-world complex system applications, it is usually difficult to acquire a massive amount of unbiased, labeled data for model training.

2. The lack of explainability and interpretability also brings severe issues for AI applicability to critical management and control problems. On one hand, AI usually relies on deep neural networks for decision-making. However, these models are usually large-scale

models and opaque to human minds. On the other hand, most real-world critical decisions in complex systems need expandability and interpretability to ensure their safety and completeness.

3. Although transfer learning techniques exist and are under development, their applicability is very limited. Usually, trained AI models from certain datasets and with dedicated purposes are not universal or transferrable. As a result, it is difficult to transfer a model from one problem or scenario to another.

In a nutshell, current AI techniques are normally applicable to decision problems with sufficient training data, clear objectives, and distinct boundary conditions of problem definition. However, complex system CMC problems are usually highly abstract and ambiguous, and cannot be solved solely by AI tools.

In 1999, *Science* published the special issue on complex systems, with the tag line "beyond reductionism" (Gallagher and Appenzeller, 1999). However, this does not mean that reductionism should be abandoned. The analysis of complex systems needs to follow the scientific principle of integrating holism and reductionism (Liu et al., 2008). Therefore, as complex systems usually possess a large number of hierarchical components and organizations, their analysis at different hierarchies also needs different methodologies and approaches. Some of the complex system CMC tasks are beyond human cognition capability, while some CMC tasks are highly abstract and cannot be solved by current AI techniques. As a result, human intelligence or machine intelligence alone cannot solve all the complex system problems given the limitation of the current state-of-the-art technologies.

There is an obvious and huge barrier between the current needs of complex system analysis and the limitations of machine intelligence. Complex system problems have the characteristics of complexity, uncertainty, relevance, and openness, all of which rely on human intelligence for high-level, abstract thinking. Human cognition capability is limited, information processing bandwidth is low, processing speed is low, and the ability of accurate calculation is limited. Therefore, complex system CMC based solely on human intelligence usually cannot support secure, efficient, and accurate decisions and operations. These need assistive tools such as computing, simulation, and AI for broadening cognition bandwidth and enhancing decision efficiency. As a result, in the process of complex system management and control, especially on critical decisions and execution, human minds need to divide problems into the ones with clear boundaries for machine intelligence based problem solvers. Also, it is human responsibility for how to design appropriate mechanisms for acquiring a vast amount of unbiased and high-quality data for robust training and modeling processes.

To tackle the above challenges brought by CMC, human-machine hybrid augmented intelligence (HM-HAI) has recently emerged as a new form of AI; i.e., human intelligence and machine intelligence are mixed aiming to enhance each other, and they are coordinated, integrated, and used throughout the process of system CMC (Zheng et al., 2017). In Zheng et al. (2017), two sub-forms of HM-HAI were proposed. One was "human-in-the-loop hybrid augmented intelligence," and the other was "cognitive computing-based hybrid augmented intelligence." The conclusion was that "hybrid-augmented intelligence is one of the important directions for the growth of AI." The proposed human-machine knowledge automation (HM-KA) mechanism focuses on the "human-in-the-loop hybrid augmented intelligence."

In this paper, we aim to study the mechanism of HM-HAI on complex system CMC, namely HM-KA, which is achieved by integrating and coordinating human and machine knowledge and intelligence in the task process of complex systems.

## 2 Overall problem of HM-HAI for complex system CMC

The core issue of HM-HAI for complex system CMC is to study the interaction among three groups of entities and design their coordinative operation schemes, i.e., solve the CMC game problem among the group of human intelligent agents, the group of machine intelligent agents, and the group of complex system component agents.

Neither human intelligence nor machine intelligence is able to completely understand and analyze a complex system, and neither can thoroughly understand each other. As a result, these two groups of different agents do need to coordinate and collaborate to the largest extent to achieve the goal of CMC for the complex system. Therefore, this CMC problem belongs to the category of complex mixed games of cooperative and non-cooperative multiple parties as shown in Fig. 2. In this problem, the task of HM-HAI is to minimize the information incompleteness among human intelligence, machine intelligence, and the systems, while improving and optimizing the coordination in the CMC game.

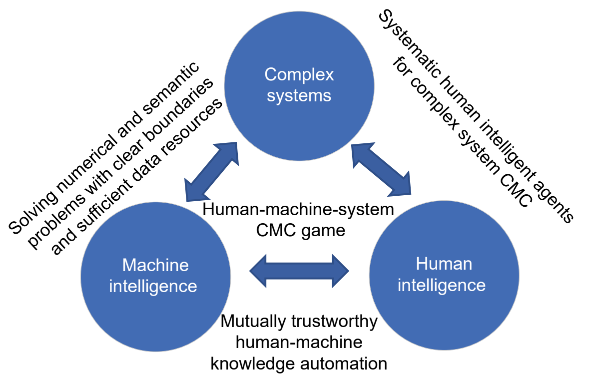


**Fig. 2 Cognition, management, and control (CMC) game of human intelligence, machine intelligence, and complex systems**

In the complex system CMC, the complexity, relevance, and criticality of the systems are three important evaluation criteria:

1. Complexity refers to the numbers of components, relevant feedback paths, and predictable and non-predictable events in the systems.

2. Relevance refers to the degree of closeness among components in the system. In a system with high relevance, one component can affect many other components, and the state changes of one part of the system can promptly affect other parts of the system.

3. Criticality refers to the complex system CMC decisions' real-time properties that must be satisfied, and also refers to the decision degree of significance in terms of the impact on society, economy, security, and other areas, when the decisions are being made and executed.

Both human intelligence and machine intelligence are driven by these three system characteristics in the CMC operation of complex systems.

## 2.1 Human intelligence for complex system CMC

CMC of a complex system requires enormous and various resources. CMC based on human intelligence is usually based on organized, systematic human intelligence groups. Through information exchange and knowledge collaboration, human intelligence groups may accomplish the CMC task of a complex system. Currently, most real-world complex system CMC processes are performed in such a way, including examples of social organization, economic management, and military combat systems.

With the development of the human society, requirements on the complexity, relevance, and criticality of the CMC of various complex systems continue to increase. The organized and systematic human intelligence agent based CMC begins to face great challenges. Specifically, the increasing system complexity requires decentralized CMC schemes with an increasing number of teams and team members, and also demands flexibility in the CMC structure for the system, so that it can authorize CMC staff to process local problems in a timely manner. Therefore, in a system with high complexity, it is normal for a large number of people to be working together with different roles and in different positions to analyze the system and implement decisions. High relevance of the system determines that one decision made by an operator often relies on the aggregation of information and knowledge from multiple relevant system components. At the same time, an operator's decision may affect many other sub-systems. Therefore, high-quality and centralized processing is required by system CMC.

The distributed CMC structure required by the high degree of system complexity and the centralized CMC structure required by the high degree of system relevance form a major contradiction in the complex system CMC. Furthermore, criticality adds a new dimension of difficulty to this contradiction.

Complex systems with high criticality require decisions to be accurate, error-free, and made and executed within specified real-time requirements. This makes very strict demands on the CMC mechanism. On some occasions, such as situations where the system information and knowledge exceed the cognitive bandwidth of human beings or the response time requirements are far below the human response time limitation, the criticality requirements become very difficult to achieve for human agent groups. Therefore, in general, human intelligence based CMC of complex systems is performed in a certain distributed and centralized form, and operated through a certain CMC process to organize the team and team members to conduct cognition tasks at different dimensions and levels of sub-systems and sub-operation processes. The cognition results are then aggregated accordingly, CMC decisions are made subsequently, and the executive management, control, and orders are delivered to each sub-group, sub-system, and operator. The increasing requirements on complexity, relevance, and criticality of various complex systems make information aggregation increasingly difficult and also put forward increasingly stringent real-time requirements and high accuracy for the CMC decisions.

With the continuous emergence and development of new complex systems, the human intelligent agent group based complex system CMC mode is facing severe challenges. Although this mode is being continuously supported by advanced tools from communication, computation, and simulation technologies, the cost and efficiency of operating a system with more human agents and larger organized groups in response to complexity have become unrealistic because of the prohibitive personnel number and cost, as well as inefficiency and impossibility of coordination. In a system with high degree of relevance, the amount of information aggregated from and shared by all subsystems and sub-components is too large and far exceeds the cognitive bandwidth of human minds. For criticality, the response time of human agents is severely limited. Moreover, human cognition and decision-making performance vary dramatically with different psychological, physiological, and social factors, and most of the time, they are not entirely controllable. Therefore, it is necessary to introduce machine intelligence into the current human agent group-centered complex system CMC paradigm to build essential hybrid augmented intelligence and knowledge automation with human-machine mutual trust.

### 2.2 Machine intelligence for complex system CMC

The CMC of complex systems by machine intelligence is still at the frontier exploration stage. As mentioned above, under the current technological circumstances, machine intelligence is able to solve computational problems such as classification, reasoning, and decision-making with sufficient training samples and labeled data and well-defined numerical and semantic computing targets. However, in complex system CMC, the applications of machine intelligence are still influenced by the aforementioned complexity, relevance, and criticality requirements and criteria. As a result, only organized and systematic groups and teams of machine agents with coordinative intelligence can realize cognitive control of complex systems.

Recently, the concept of Internet of Minds (IoM) has been proposed and practiced in real-world applications (Wang and Zhang, 2017; Wang et al., 2018). The IoM, based on the technologies of Internet and Internet of Things, using knowledge automation systems as the core system, aims to accomplish critical knowledge engineering tasks including obtaining, expressing, exchanging, and coordinating knowledge. IoM also aims to establish connections of intelligent entities at the semantic level. The ultimate goal of IoM is to support and complete knowledge-related functionalities and services that require large-scale coordination and collaboration, especially in complex system CMC.

Technically, IoM aims to reach collaborative knowledge automation and collaborative cognitive intelligence of the group of intelligent agents, and provides foundations to implement knowledge service functionalities including reasoning, strategy, decision-making, planning, and control in a collaborative way. Therefore, the essence of IoM is a brand-new complex and collaborative knowledge automation system, which is directly oriented to coordinated intelligence.

The development of IoM is still at its early stage. A large number of intelligent entities form a complex system linked by knowledge in accordance with certain operating rules and mechanisms. The complex system will form a social network-like organization that is self-organized, self-operating, self-optimizing, self-adapting, and self-cooperating. Such IoM will have a revolutionary effect on solving the challenges brought by the complexity, relevance, and criticality requirements of complex systems. However, a summation of the solutions to each single problem often does not constitute the solution to a complex system CMC. As aforementioned, when we have highly abstract CMC problems, such as operational objectives, organizational structures, operational modes, component dimension, analysis methods, task processes, and technical pathways, human intelligence (which has abundant system CMC experience and knowledge and unique capability in solving open and highly abstract problems) will coexist for a long time with IoM and constitute the human-machine IoM.

### 2.3 CMC game between human intelligence and machine intelligence

For the highly abstract aspects of a complex system, such as system operational objectives, organization structures, operation modes, component interactions, analysis modes, task processes, and technology pathways, there is still a wide gap between the required solutions and the solutions which the current machine intelligence and the IoM technology can provide. The problem must be addressed by the fusion of human intelligence and machine intelligence.

Compared with the number of machine intelligence agents, the number of human agents is very small. When a small number of human agents and a large number of machine intelligence agents work together to complete CMC processes, the role of the human agents can generally be divided into the following three categories: (1) Human agents manage the crisis and mistakes within a required time frame, and should make right decisions in a timely manner; (2) Human agents need to make major decisions at a highly abstract level; (3) Human agents "encourage" and "motivate" machine agents under their control to achieve the desired group behavior (Wickens et al., 1998).

When facing system complexity, HM-HAI uses human intelligence to establish a reasonable system component structure, which is a highly abstract task, to schedule task objectives and processes and to specify the boundary of CMC problems. At the same time, a large number of machine intelligent agents are applied to solve the problems with clear boundaries in various system components. When addressing the issues of system relevance, machine agents are adopted to provide ultra-high cognitive bandwidth, massive knowledge sharing, and efficient collaboration mechanisms. When addressing the criticality issues, machine intelligence has high accuracy and real-time capability. These provide preemptive technical conditions for the realization of the critical performance of complex system CMC. At the same time, when making critical decisions for the system, the management and control decisions of human intelligence and machine intelligence must follow and obey certain decision-making procedures and rules, and must be compatible with each other.

The fusion of human-machine knowledge is generally achieved following the approaches below: (1) The behavioral data of human agents is collected through certain human-machine interfaces to form available input data for machine learning, and then integrated into the algorithms of machine intelligence and learning; (2) Machine intelligence provides knowledge representations which can be transformed into understandable knowledge representations by human intelligence; (3) Through linguistic expression and formal language processing, the knowledge of human agents forms symbolic knowledge representations, which can be integrated with symbolic knowledge representations of machine intelligence to eventually form human-machine cooperative knowledge representations; (4) In the process of human-machine coordination of CMC, when conflicts occur or decision fusion is needed, certain conflict-solving procedures should be adopted according to pre-defined fusion principles.

From the above consideration, the coordination mechanism of human-machine intelligence becomes the essential research topic in the field of HM-HAI based complex system CMC. We propose the concept of "human-machine knowledge automation" as the fundamental coordination mechanism in the complex system CMC process.

### 2.4 Mutually trustworthy knowledge automation in a complex system CMC process

For a complex system CMC process, especially the critical decision-making process in critical complex systems (such as energy, medicine, finance, and transportation) and in industrial systems, the human-machine hybrid system should meet all the needs of complexity, relevance, and criticality. Therefore, human intelligence and machine intelligence need to be able to achieve mutual understanding and trust. They also need to achieve automatic intelligence elevation and fusion of knowledge. Thus, on the basis of knowledge automation (KA), the concept of "mutually trustworthy HM-KA" is proposed below.

The key points of the HM-KA mechanism are shown in Fig. 3, including human-machine

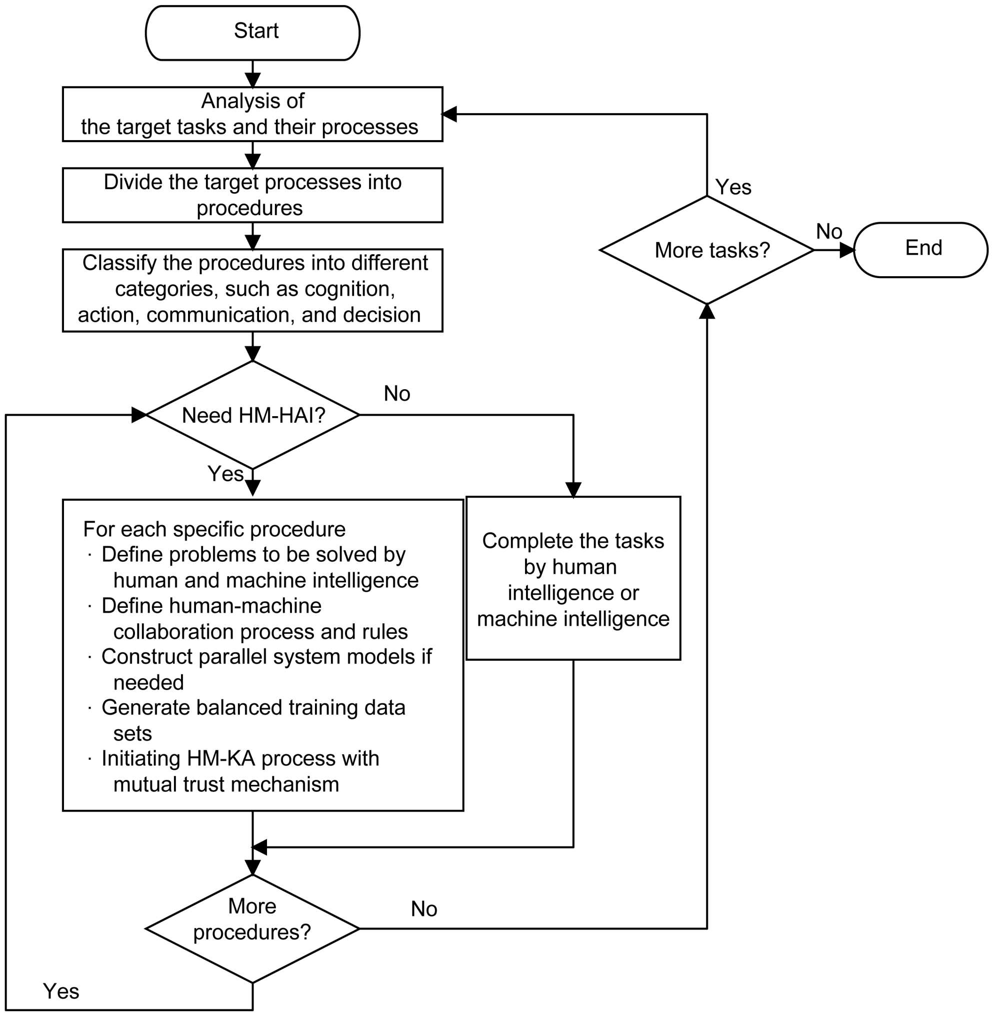


**Fig. 3 A general human-machine knowledge automation (HM-KA) process in complex system cognition, management, and control (CMC) tasks**

functionality deployment, human-machine behavior constraint, machine intelligence interpretation, and human-machine decision verification.

2.4.1 Task and process based human-machine functionality deployment

In the process of complex system CMC, human intelligence and machine intelligence should follow the principle of deployment based on operation and task processes and human-machine capability differences. The process of joint deployment of human-machine intelligence needs to be specific on the application scenario and task levels, and even on the human-machine behavior level. On one hand, human intelligence needs to define the problems with clear boundaries and sufficient data resources that are suitable for machine intelligence. On the other hand, high-level system abstract problems that need to be solved by human intelligence are summarized and assigned to human agents (Fig. 3). The IoM technology and human-machine intelligence connection technology are used to realize knowledge collaboration and intelligence fusion of multiple human-machine agents, and to realize knowledge automation for complex system CMC.

2.4.2 Rules and responsibilities based human-machine behavior regulations

The CMC mechanism of mutually trustworthy HM-KA needs to deploy human intelligence and machine intelligence with reasonable cooperation rules, norms, and standards. In particular, intelligent machine systems need to have enough common-sense knowledge through operational rules. Any machine intelligence that interacts with human intelligence must understand and abide by the rules when making and executing decisions in the system. It is necessary to establish standards that are suitable for evaluating the behaviors of machine intelligence agents. For example, in the field of power grids, it is necessary to establish hybrid human-machine intelligence deployment and scheduling rules.

It is very important to establish rules and mechanisms of assigning responsibilities in a human-machine knowledge automation system. Human intelligence based complex system CMC has a relatively mature responsibility assignment system. However, the responsibility assignment mechanism for machine and hybrid intelligence is basically non-existent. This novel responsibility mechanism needs to include the evaluation of algorithms, data, and design processes, and regular inspections of whether machine intelligence systems comply with required specifications and whether these specifications have produced the desired effect. The responsibility mechanism starts from the CMC rules, specifications, and standards, identifies and tracks the emerging problems, ensures the integrity of the CMC process system, and monitors the use of systems, the machine learning architecture, data sources and management, and the training and use of AI models.

#### 2.4.3 Mutual interpretability based human-machine intelligence fusion

For human-machine knowledge automation, the mutual interpretability of human-machine intelligence is a crucial link (Zhang TY et al., 2020; Zhang K et al., 2021b). To establish a robust human-machine hybrid system, we need to start with the establishment of a CMC mechanism that has an in-depth understanding of the target systems or tasks. For the human-machine hybrid system, "deep understanding" is more profound than statistical models and neural network models. Currently, the deep neural network technology is far from satisfying the requirement of being interpretable to human beings. As a result, it is important to explain the decision-making process and the results of deep neural network models using human understandable language, because the explanation ensures trust and transparency in the decision-making processes, especially in scenarios related to major safety concerns. Similarly, human intelligence can be understood and used by machine intelligence only in the form of knowledge representations. Its representation forms include symbolic knowledge representations, formal language, and teaching examples.

At the same time, the mutual interpretability of human-machine intelligence is a new and important means to human-machine interaction. In that process, interpretable machine intelligence can help human agents better understand the behavior of the machine intelligent agents and the states of the systems, and support the machine intelligent agents in acquiring valuable abstract thinking results and empirical knowledge.

#### 2.4.4 Empirical knowledge and parallel systems based human-machine decision verification

Both human intelligence and machine intelligence need to trust each other's decisions to ensure that the whole human-machine intelligence system meets the criticality requirements. In addition to the aforementioned rule- and interpretability-based principles, verification is one of the most crucial factors in achieving such mutual trust. There are two ways of verification in practice. The first is based on the knowledge extracted from existing practice cases and experience (Dai et al., 2021b). Most complex, open, and uncertain scenarios and tasks need to be completed in virtual systems. For example, parallel systems and related technologies can be used to predict and verify functionalities and the results of CMC in a large number of virtual scenarios by the ACP approach (Wang, 2004). Then the interpretability tools with deep understanding capabilities are used by human intelligence to infer, judge, and supervise machine intelligence for its correctness, reliability, and robustness. Also, through monitoring the system closely and fine-tuning it continuously, the verification principle aims to ensure the reliability of performance, detect and correct the deviations, and improve the transparency and inclusivity.

## 3 HM-HAI, HM-KA, and bulk power grid dispatch

The bulk power grid dispatch is essentially the CMC process for the complex system of a bulk power grid. The bulk grid has characteristics of open operational environments, complex system components, diverse operation modes, tightly coupled component behaviors, stringent demand on real-time response performance, high criticality, and so on (Song et al., 2006; He et al., 2008). Therefore, the HM-HAI technology needs to address the key scientific, technical, and engineering problems in this complex system CMC process, including the systematic modeling of bulk power grid dispatch, the construction of environment models shared by human agents and machine agents, and task and behavior models of human-machine intelligence. Based on the above models, complex CMC tasks are effectively completed, and the human-machine autonomous cooperative CMC mechanisms can be established.

Theoretically, a power grid should be observable and controllable as the physical operation mechanism of its circuitry and equipment is clear. However, in the real world, this is not true, and the following factors need to be taken into consideration: First, the contradiction between data integrity with real-time requirements and limitations on data sensing and collection in power grids affects the accuracy of the data and damages information integrity; Second, characteristics including time variation, high non-linearity, and uncertainty of power grid operation make it not completely knowable within the effective CMC process time. As a result, although the power grid operation follows physical laws and mechanisms, it is difficult to describe the power grid operation accurately in a required time frame. The coordination problems involving the power grid state variation time constant, the controller time constant, and the spatially distributed control measures make the power grid not entirely controllable. Therefore, timeliness is the core issue of bulk power grid dispatch, and timeliness, effectiveness, and strong adaptability (robustness) are the driving forces and developmental directions for this industrial complex system CMC.

### 3.1 Bulk power grid dispatch based on human intelligence

At present, the operational process of bulk power grid dispatch is organized and systematized by human agents, and the overall function is to realize the CMC of power generation, transmission, distribution, trading, load adjustment, and other businesses in the power industry. The CMC process needs to be in accordance with the processes of dispatch scheduling, mode operation, field dispatch, dispatch automation, system protection, system monitoring, and so on.

In terms of the required real-time requirements for task completion, the dispatch services are classified and divided into the following categories (Vadari, 2012):

1. Real-time service. This type of service aims to monitor the grid status, such as monitoring the voltage, load, frequency, the conditions of transmission lines, generator, and equipment, and to respond in real time, such as responding to various alerts as soon as possible. Future states of the power grid need to be estimated based on the current situation to take the necessary CMC measures in advance.

2. Management service for pre-scheduled events. For pre-scheduled power grid events, the service covers management work, including scheduling, authorization, implementation, and acceptance. Taking equipment maintenance as an example, the service needs to determine all the system protection settings and time sequences, and then to ensure the balance of the power of the designated area when the maintenance work begins.

3. Emergency management service. For unpredictable events, such as equipment failures and control system malfunctions, the service needs to make a real-time response to make the system return to a normal operational state as soon as possible. The specific responses include conducting effective and optimal recovery work on time, minimizing the impact on power network users, and ensuring the safety of users and employees in the process of re-energization.

4. Emergency response organization. If there is a large-scale power failure and blackout, the power grid system needs to enter an emergency status. The service needs to optimally organize resources to respond to emergencies.

5. Pre-scheduling service. The pre-scheduling service aims to analyze the current system, to make operational schedules for the next days, weeks, months, quarters, and years, and to make power grid operation arrangements in advance according to possible major events, such as significant changes in power grid load due to significant social events and possible power grid failures that may be caused by extreme weather conditions.

6. Power grid analysis. This service aims to effectively manage the grid assets and resources and to make short-, medium-, and long-term power grid operation decisions on dispatch, equipment replacement, grid maintenance, and investment.

7. Power grid performance analysis. This service aims to analyze the grid operational data to deliver reports on the performances of various system components and processes based on power industry standards.

In general, the power grid dispatch center is divided into many departments such as dispatching department, system operation department, automation department, planning department, monitoring department, and protection department. These departments carry out collaborative cognition, management, and control for the power grid operation following certain business processes and organizational architectures. The functionalities, responsibilities, and knowledge services of each department for supporting the power grid operation are shown in Fig. 4. Although this departmental setting supports the operation of the power grid, the knowledge exchange and intelligence collaboration are human-centric with very low efficiency, severely affecting a power grid's ability to respond to extreme events and fast-changing operational situations.

### 3.2 Bulk power grid dispatch based on machine intelligence

In this subsection, a case study is used to illustrate how machine intelligence completes a power grid CMC process under current technology conditions. The specific scenario is the "power correction control" of the power grid (Chen et al., 2021; Xu et al., 2021).

In this case study, the proposed machine intelligence for CMC is performed on a 36-bus power grid as shown in Fig. 5, which includes 59 transmission


**Fig. 4 Power grid dispatch operation departments and responsibilities**

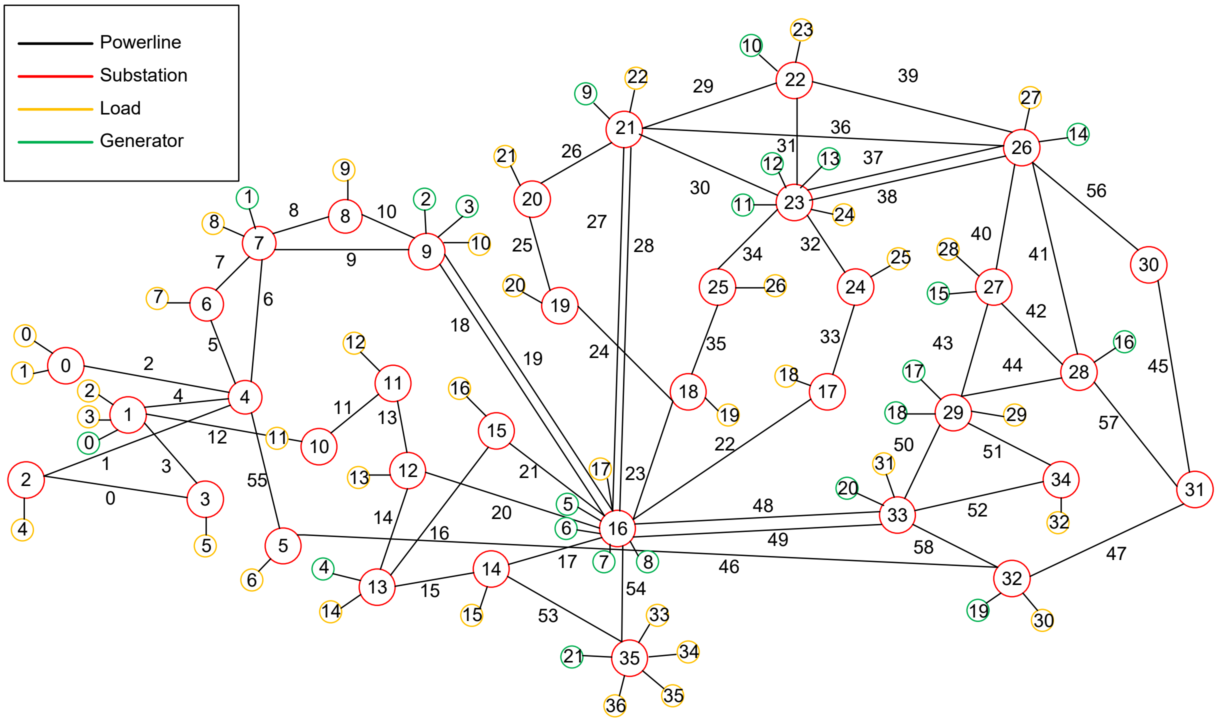


**Fig. 5 Topology of the 36-bus power grid**

lines, 22 buses with generators, and 37 buses with loads. The machine intelligent agents are trained on the platform Grid2Op (RTE-France, 2021), which provides data of 576 scenarios over a period of 48 years with 12 months each year. Each scenario contains the operational data at 5-min intervals over a 28-d period. In these scenarios, random transmission lines of the power grid are "attacked" at random time steps every day; that is, short-circuit faults with random occurrences and random durations may occur on every transmission line. In addition, the platform provides scenarios of 24-week duration, which are not included in the training data, for testing the trained agents. These test scenarios are generated according to the simulation of typical power grid operation scenarios. Because of the limitations of the algorithm and to prevent the system from unnecessary disturbance, each agent takes only a single control measure at each CMC time step. Even so, in the 36-bus power grid, there are still more than 60 000 available actions to choose at each time step, most of which are bus-bar switching actions. The recovery time impacted by the control actions is required to be within three time steps, that is 15 min. The machine intelligence agents are trained on a Linux server with 4 GPUs, each with 11 GB of memory.

The agents used in CMC, specifically, power correction control, of a power grid are trained by the algorithm of deep reinforcement learning. At each control time step, 1500 available actions are screened through the Monte-Carlo tree search. The whole action space of the agents is divided into four parts. Each part is managed and controlled by an agent with an attention mechanism based on a graph deep neural network. The structure of each agent is shown in Fig. 6. The four agents cooperatively exchange information and model parameters with each other to conduct power correction control of the power grid. The four agents are centrally trained and distributed in the execution. In the collaboration mechanism, the agents will share the global observations and reuse the model parameters. This will reduce the model complexity and improve the training efficiency.

With an average decision time of 35 ms for each time step, the agents make continuous forward power correction control decisions, which can quickly eliminate the off-limit power flow at the next time steps and maintain the stability of the power grid in the new operational scenarios.

In the performance evaluation after deployment, deep reinforcement learning (DRL) agents have good

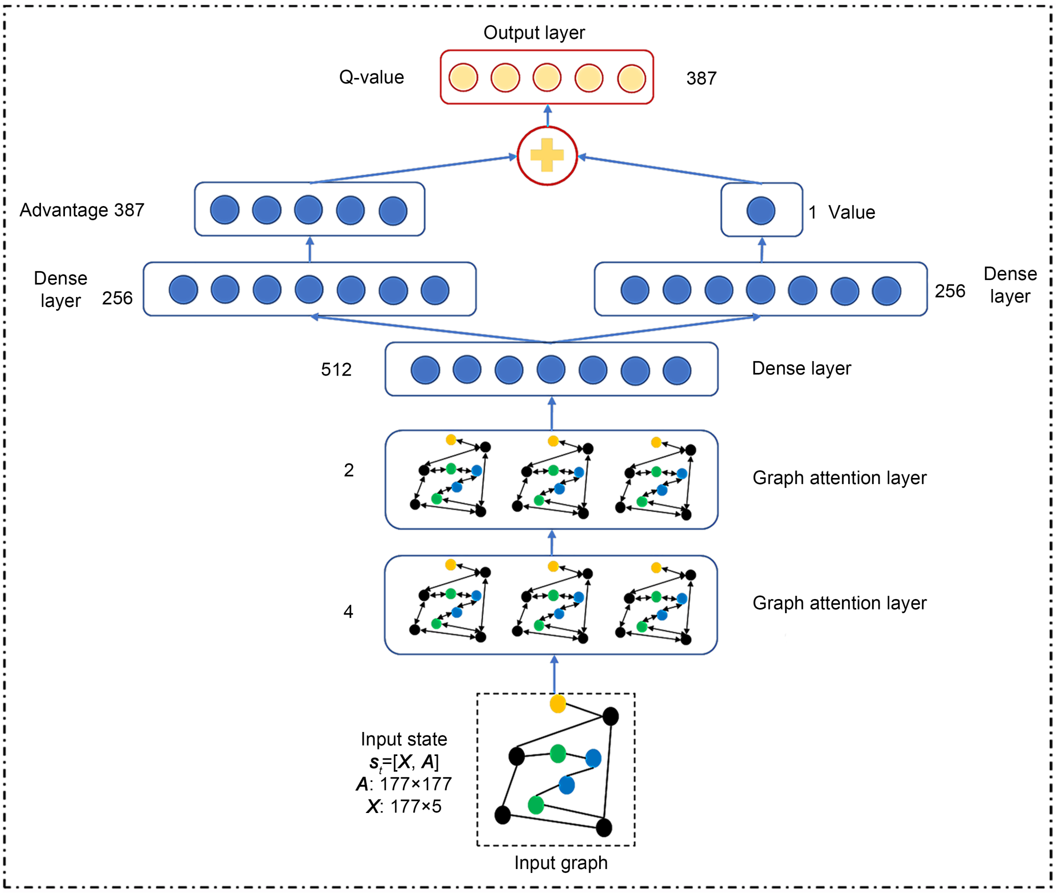


**Fig. 6 Architecture of graph deep Q-networks**

timeliness and provide effective control strategies in most scenarios. However, in extreme grid operational scenarios, the agent operations may not be quite effective. The possible reasons include complex task flow, large numbers of highly coupled coordinative operation components and processes, high dimension of power grid observation and control measure spaces, and unpredictable and unforeseen grid operation mode. This may also be the case for future normal operation scenarios with an extreme high rate of renewable energy generation. Also, high-level abstract problems in this application, such as area planning, typology definition, action definition, and action rules, are solved by human intelligence, and currently machine intelligence does not possess such capability. Therefore, DRL needs to cooperate with human knowledge in power correction control, as well as to avoid wrong decisions caused by lack of cognition when dealing with open and unknown scenarios.

### 3.3 Application of HM-KA in bulk power grid dispatch

The mechanism of "mutually trustworthy HM-KA" discussed above is supported by four aspects: functional deployment, collaborative rules, interpretability, and decision verification.

The technical support of the first two aspects, functional deployment and collaborative rules, was discussed in detail in Sections 2.4.1 and 2.4.2, respectively. In terms of the third aspect, mutual interpretability of human-machine intelligence, its supporting technologies include distributed reinforcement collaborative learning, semantic space fusion, assistive reinforcement learning, and interpretability of cognitive graphs. Some examples of such techniques (Zhang TY et al., 2020; Chen et al., 2021; Dai et al., 2021a; Xu et al., 2021; Zhang K et al., 2021a, 2021b) have been demonstrated. In terms of the fourth aspect, decision verification based on parallel system and knowledge engineering technology is a feasible and practical pathway for real-world applications. Its key supporting techniques include the construction of the empirical knowledge architecture, parallel system based computing architectures, deep reinforcement learning based automatic data generators, and semantical knowledge representation networks for complex system human-machine-material relations.

Based on the above key supporting technologies and guiding rules of human-machine knowledge automation, the human-in-loop intelligent CMC system for bulk grid dispatch can be expected to have the following characteristics: (1) It is a human-machine enhanced auxiliary decision mechanism superimposed upon the current dispatch system. The main functionalities of the auxiliary decision system are to provide adjustment decision suggestions in normal operational circumstances to keep the grid operating in a balanced manner with safety and efficiency, and to meet the requirements of fluctuation and uncertainty when incorporating new energy. (2) In abnormal working conditions, the auxiliary decision mechanism generates forward countermeasures and suggestions in real time and verifies the decisions automatically, aiding human operators in making the final decision. AI-assisted decision-making is expected to have better adaptability and more suitable security.

## 4 Example of HM-KA based power grid dispatch

The corrective control of power grids is used as a benchmark scenario to illustrate the application of HAI and HM-KA in power grid dispatch.

In this scenario, online corrective control is performed mainly to maintain the stability of the 62-bus regional power grid of a China 300-bus multi-region power grid. The topology of the power grid is shown in Fig. 7. The sending end of the power grid

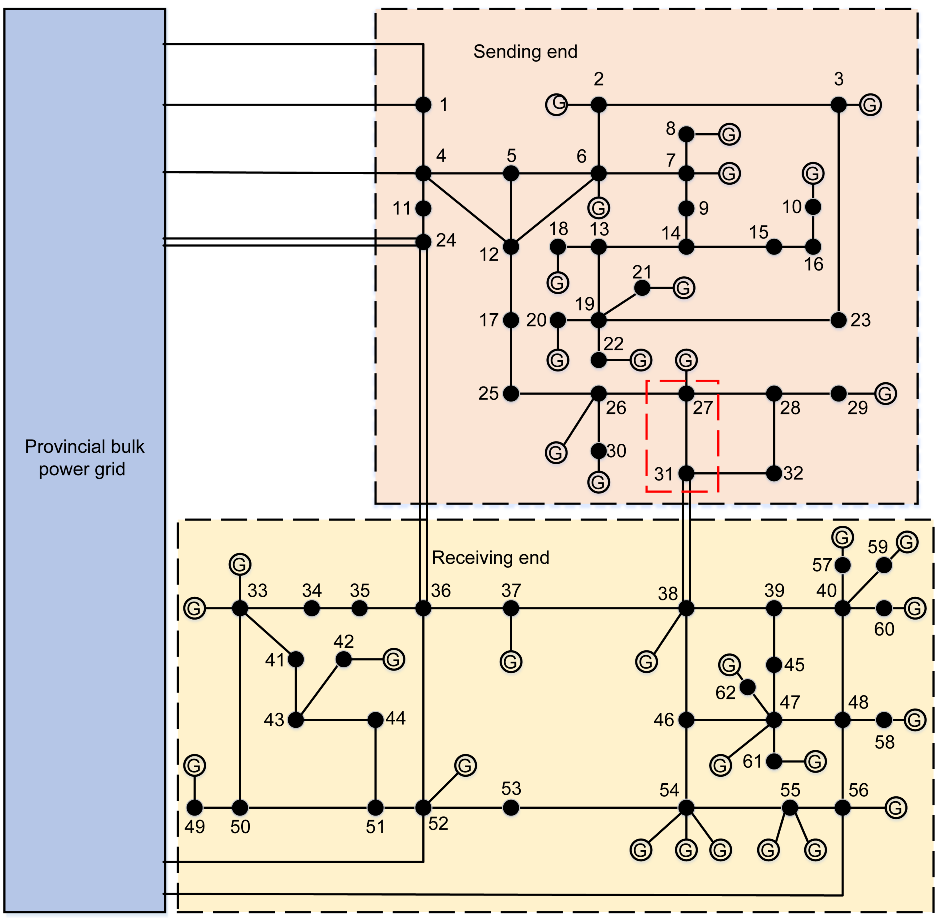


**Fig. 7 Topology of the regional power network**

transmits power to the receiving end through two alternating current transmission corridors. From the experience of the dispatch operators, the transmission power of lines 27–31 (that is, the origin of this line connects to bus 27, and the extremity connects to bus 31) in the sending end area is prone to exceed the thermal limit. Thus, lines 27–31 are defined as the critical transmission line. When the operational condition of the power grid constantly changes, the critical line is more likely to overrun the transmission capacity, and the other lines are also likely to suffer from overloads, which may lead to line tripping and succeeding cascading failures. In this situation, the dispatcher needs to adjust the line power flow through generator re-dispatch, transmission line switching, and other operations. By applying feasible corrective control strategies, the active power of lines 27–31 can be reduced within the thermal limit, and the stable operation of the whole region can be further maintained. Note that there are a large number of adjustable devices in this grid. Just in the 62-bus regional power network, the sending end and the receiving end contain 14 and 20 generators, respectively. Assuming that the maximum number of generators allowed to be dispatched is 3 in one control action, the number of candidate actions can be 4760.

The machine intelligent agent of corrective control is trained by the data from operator experience of transmission line switching to realize the fusion of human knowledge with AI. From the principles described in Section 3.3, the human-machine cooperation mode is further designed. Based on the current system line tripping situation and the duration of line overloads, the degree of power grid emergency is defined. When the grid is not at an emergency state, the machine intelligent agents generate the corrective control strategy. If the produced strategy accords with the prior rules and passes the simulation verification, the corrective control action is carried out. Otherwise, dispatch operators will adjust and control the power flow based on their own experience. In the case of emergency, dispatch operators will directly perform the corrective control operation according to their experience to ensure the timeliness and effectiveness of critical operations. The human-machine cooperation mode is shown in Fig. 8.

Given the above framework, the corrective control agent is trained based on the power grid simulation software. The machine intelligence and human-machine hybrid augmented intelligence are deployed in a typical scenario. The operating state and control process of the power grid are shown in Fig. 9.

As shown in Fig. 9, in the machine intelligence control mode, at the early stage, the DRL agent can

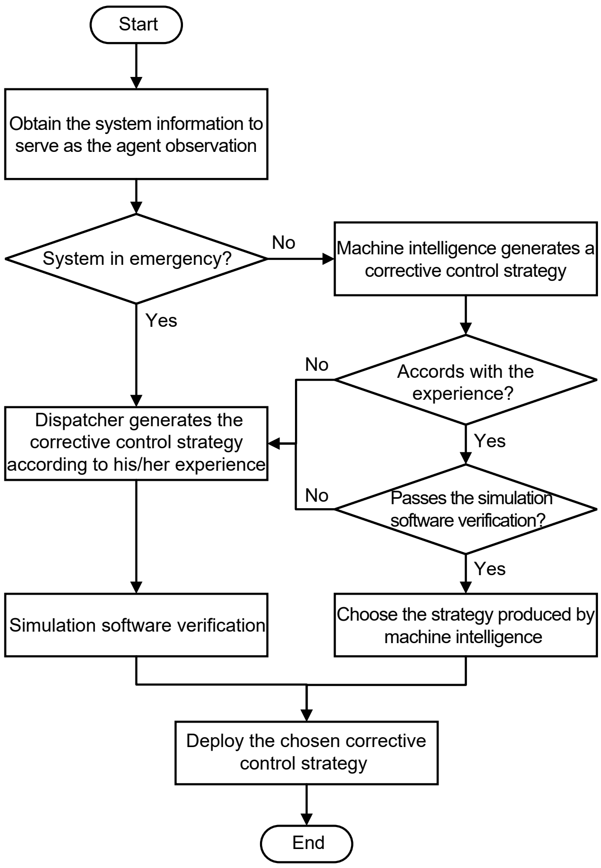


**Fig. 8 Schematic of the human-machine cooperation mode**

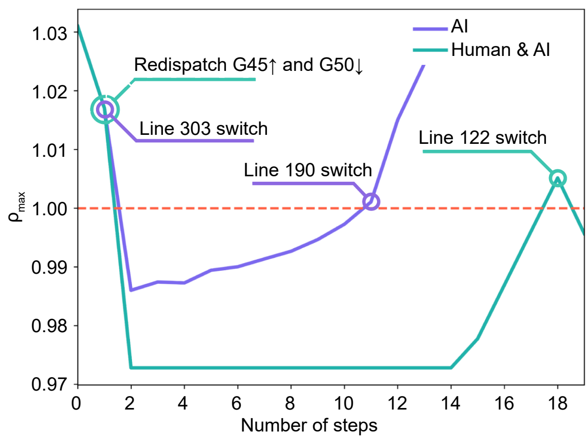


**Fig. 9 Operating state of the power grid under different control modes (every step takes 5 min)**

quickly eliminate the overload of critical lines 27–31 by adjusting the network topology. However, when the transmission power of another key line exceeds the limit of the later stage, machine intelligence fails to take effective measures and finally causes the system to collapse. In the human-machine hybrid augmented intelligence control mode, when confronted with the overload of critical lines 27–31 at the initial stage, the topology adjustment action taken by machine intelligence does not accord with the experience. Thus, the dispatcher carries out the corrective control action. The overload situation is effectively eliminated through the generator re-dispatch operation, and the load ratio of the transmission lines in the grid is improved. At the later stage, when another key line exceeds the limit, machine intelligence generates the strategy in line with the experience and passes the simulation software verification, the line overload is eliminated in time, and the continuous and stable operation of the power grid is realized. Based on this example, it can be inferred that with a reasonable human-machine cooperation mode, AI and the dispatcher can complement each other in CMC capability, and thus human-machine hybrid augmented intelligence is conducive to further improvement of power grid dispatch.

## 5 Conclusions

In this paper, we first reviewed the history of complex system science, characteristics of complex systems, the concept of HM-HAI, and the reason why HM-HAI is a necessary emerging technology for complex system cognition, management, and control (CMC). Then, we explained the mechanism of human intelligence and machine intelligence to complex system CMC and their advantages and limitations. The concept of “mutually trustworthy HM-KA” mechanism for complex system CMC was proposed. Finally, using a bulk power grid benchmark as the background of power grid dispatch, we analyzed how HM-HAI can be used in future power grid applications. Using human-machine enhanced corrective control as an example, we illustrated the technical path of mutually trustworthy HM-KA.

Through this work, we hope to provide a new mechanism, namely HM-KA, for the theory and methodology of complex system CMC based on HM-HAI. We discussed the core method, implementation process, and key supporting technologies for HM-KA. The development of this method will be a trans-disciplinary product, including results from system science, control science, management science, psychology, cognitive science, data science, and AI. We hope that in the process of digital transformation in typical complex social and industrial systems, such as aerospace, aviation, marine, energy, manufacturing, transportation, environmental protection, and other fields, the proposed theory, mechanisms, approaches, and techniques can play positive roles and make considerable contributions.

### Contributors

Fei-Yue WANG and Jianbo GUO designed the research. Guangquan BU and Jun Jason ZHANG conducted the analysis. Guangquan BU and Jun Jason ZHANG drafted the paper. Fei-Yue WANG and Jianbo GUO revised and finalized the paper.

### Compliance with ethics guidelines

Fei-Yue WANG, Jianbo GUO, Guangquan BU, and Jun Jason ZHANG declare that they have no conflict of interest.

### References

Binney JJ, Dowrick NJ, Fisher AJ, et al., 1992. The Theory of Critical Phenomena: an Introduction to the Renormalization Group. Clarendon Press, Oxford, UK.

Chen SY, Duan JJ, Bai YY, et al., 2021. Active power correction strategies based on deep reinforcement learning—Part II: a distributed solution for adaptability. *CSEE J Power Energy Syst*, early access. https://doi.org/10.17775/CSEEJPES.2020.07070

Dai YX, Chen QM, Gao TL, et al., 2021a. Deep reinforcement learning control policy extraction based on weighted oblique decision tree. *Electr Power Inform Commun Technol*, 19(11):17-23.

Dai YX, Zhang J, Ji ZX, et al., 2021b. Intelligent diagnosis and auxiliary decision of power system secondary equipment based on functional defect text. *Electr Power Autom Equip*, 41(6):184-191. https://doi.org/10.16081/j.epae.202106006

Gallagher R, Appenzeller T, 1999. Beyond reductionism. *Science*, 284(5411):79. https://doi.org/10.1126/science.284.5411.79

Haken H, 1977. Synergetics. *Phys Bull*, 28(9):412-414. https://doi.org/10.1088/0031-9112/28/9/027

He HL, Guo JB, Song YT, et al., 2008. Risk based probabilistic security evaluation algorithm of bulk power system. 3rd Int Conf on Electric Utility Deregulation and Restructuring and Power Technologies, p.1231-1236. https://doi.org/10.1109/DRPT.2008.4523595

Hey T, Tansley S, Tolle K, 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research, Redmond, USA.

Holland JH, 1995. Hidden Order: How Adaptation Builds Complexity. Addison-Wesley, New York, USA.

Liu XT, Liang BC, Liu L, et al., 2008. The Theory, Method & Technique for Complex System Modeling. Science Press, Beijing, China (in Chinese).

Lorenz EN, 1963. Deterministic nonperiodic flow. *J Atmos Sci*, 20(2):130-141.

Mandelbrot BB, 1989. Fractal geometry: what is it, and what does it do? *Proc Roy Soc Lond A Math Phys Sci*, 423(1864):3-16. https://doi.org/10.1098/rspa.1989.0038

Manoj BS, Chakraborty A, Singh R, 2018. Complex Networks: a Networking and Signal Processing Perspective. Prentice Hall, Boston, USA.

Miller JH, Page SE, 2007. Complex Adaptive Systems: an Introduction to Computational Models of Social Life. Princeton University Press, Princeton, UK.

Prigogine I, Lefever R, 1973. Theory of dissipative structures. In: Haken H (Ed.), Synergetics. Vieweg+Teubner Verlag, Wiesbaden, Germany, p.124-135. https://doi.org/10.1007/978-3-663-01511-6_10

RTE-France, 2021. Grid2Op. https://github.com/rte-france/Grid2Op [Accessed on July 13, 2021].

Sayama H, 2015. Introduction to the Modeling and Analysis of Complex Systems. Open SUNY Textbooks, New York, USA.

Shannon CE, 1948. A mathematical theory of communication. *Bell Syst Tech J*, 27(3):379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Song YT, He HL, Zhang DX, et al., 2006. Prsobabilistic security evaluation of bulk power system considering cascading outages. Int Conf on Power System Technology, p.1-6. https://doi.org/10.1109/ICPST.2006.321614

Thom R, 1975. Structural Stability and Morphogenesis. Addison-Wesley Benjamin, Inc., New York, USA.

Thurner S, Hanel R, Klimek P, 2018. Introduction to the Theory of Complex Systems. Oxford University Press, Oxford, UK.

Vadari S, 2012. Electric System Operations: Evolving to the Modern Grid. Artech House, Boston, USA.

von Bertalanffy L, 1968. General System Theory. George Braziller Inc., New York, USA.

Wang FY, 2004. Parallel system methods for management and control of complex systems. *Contr Dec*, 19(5):485-489, 514 (in Chinese). https://doi.org/10.3321/j.issn:1001-0920.2004.05.002

Wang FY, Chen JL, 2020. Intelligent Control Method and Application. China Science and Technology Press, Beijing, China (in Chinese).

Wang FY, Zhang J, 2017. Internet of Minds: the concept, issues and platforms. *Acta Autom Sin*, 43(12):2061-2070 (in Chinese). https://doi.org/10.16383/j.aas.2017.y000006

Wang FY, Zhang J, Zhang J, et al., 2018. Industrial Internet of Minds: concept, technology and application. *Acta Autom Sin*, 44(9):1606-1617 (in Chinese). https://doi.org/10.16383/j.aas.2018.y000004

Wickens CD, Gordon SE, Liu YL, 1998. An Introduction to Human Factors Engineering. Longman, New York, USA.

Wiener N, 1948. Cybernetics or Control and Communication in the Animal and the Machine. John Wiley & Sons Inc., New York, USA.

Xu PD, Duan JJ, Zhang J, et al., 2021. Active power correction strategies based on deep reinforcement learning—Part I: a simulation-driven solution for robustness. *CSEE J Power Energy Syst*, early access. https://doi.org/10.17775/CSEEJPES.2020.07090

Zhang K, Zhang J, Xu PD, et al., 2021a. Explainable AI in deep reinforcement learning models for power system emergency control. *IEEE Trans Comput Soc Syst*, 9(2):419-427. https://doi.org/10.1109/TCSS.2021.3096824

Zhang K, Xu PD, Gao TL, et al., 2021b. A trustworthy framework of artificial intelligence for power grid dispatching systems. IEEE 1st Int Conf on Digital Twins and Parallel Intelligence, p.418-421.

Zhang TY, Gao TL, Xu PD, et al., 2020. A review of AI and AI intelligence assessment. IEEE 4th Conf on Energy Internet and Energy System Integration, p.3039-3044.

Zheng NN, Liu ZY, Ren PJ, et al., 2017. Hybrid-augmented intelligence: collaboration and cognition. *Front Inform Technol Electron Eng*, 18(2):153-179. https://doi.org/10.1631/FITEE.1700053