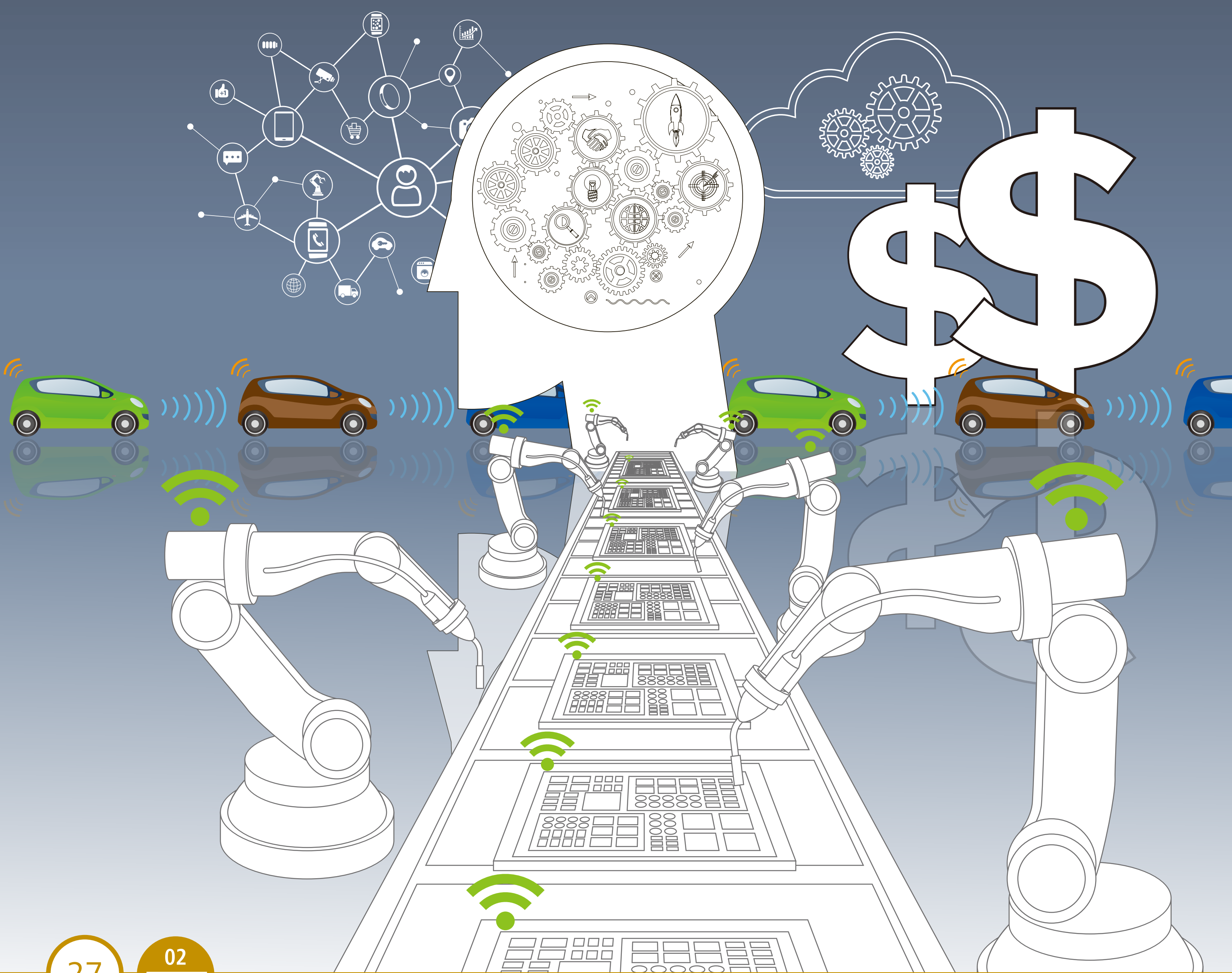
Thinking big with cognitive computing

McKinsey & Company predict that the six major technology sectors will be worth US$10 trillion by 2025. Notably, the industries framing the big six – mobile Internet, knowledge task automation, the Internet of Things, cloud technology, advanced robotics, and autonomous vehicles – all rely heavily on cognitive computing.

Scan for mobile reading

By Xue Xijun and Yan Zheng



> Cognitive computing is the CNS of these revolutionary technologies and products, forming the key component to perceive people, things, events, and the environment.

The current state of play

Cognitive computing is the CNS of the nascent crop of six revolutionary technologies and products, forming the key component to perceive people, things, events, and the environment.

Imbued with the ability to simulate human intelligence and efficiently process fuzzy data and problems, cognitive computing is also essential for data analysis and decision-making. Specifically, cognitive systems are information systems that contain the cognitive hallmarks of learning, inferring, deducing, and computing.

Research in cognitive computing has gone through several waves of ups and downs in both momentum and progress. Enormous achievements in fields such as databases and expert systems, for example, have been counterbalanced by stutters in machine translation and natural language processing. Since entering the 21st century, big data and high-performance computing have driven cognitive computing into an exciting period marked by the commercial use of Deep Learning in speech-recognition, image, and video.

Many companies, including Google, IBM, Facebook, Microsoft, and Baidu, have been active in the cognitive computing sector, leading a commercial upsurge in the field and the rapid emergence of related technologies. Based on Deep Learning, machines' cognitive abilities have grown to be on par with humans in areas such as facial recognition and image classification, and are in fact more accurate than the human eye. Equally, the Watson system was designed under IBM's DeepQA project to answer questions from the quiz show Jeopardy, and defeated two former winners in 2011.

Brain-inspired computing that uses a mass system of integrated circuits to simulate the behavior of neurons is marching towards commercialization, with Qualcomm already planning to use the Zeroth neural network chip to improve its mobile phone processors.

A series of exciting advancements in current research and applications for cognitive computing has been accompanied by a gradual upward trend in investment in R&D by ICT companies. Huawei stands out as one of the few companies to develop cognitive computing plans based on both the Von Neumann and brain-inspired architectures. Huawei plans to coordinate its resources in various areas including communications, network topology, computing, and storage to underpin the public Huawei cognitive cloud platform, Huawei REN.

Huawei REN

Huawei believes that sustained technical evolution will create new types of business models, which will see the influence of cognitive computing on the commercial sector continue to grow. Sustained commercial development is also expanding in application scope and technical improvements, and cultivating an open and progressive ecosystem. Huawei REN is committed to promoting the evolution and development of cognitive computing in four areas: cloud-pipe-device synergy with layered intelligence, cognitive computing as a service, heterogeneous coordination, and brain-inspired intelligence.

Cloud-pipe-device with layered intelligence

Layered intelligence is displayed on end devices through perception and action, on network transmission through planning, and on the cloud through learning and evolution. It is a naturally selected state of intelligence that humans have attained after 200,000 years of

> High refinement and granularity denote the ability to perceive spatial and temporal differences in client needs. Personalization provides layered services that target client needs and problem-areas by embedding cognitive computing as a service.

evolution, with the central nervous system providing a particularly good example.

Cognitive learning on a coordinated cloud-pipe-device platform may be unlike human learning in form, but it's similar in spirit. Memory and cognition are obtained from the cloud, while perception, interaction, and action control are coordinated on the device. For example, in video analysis-type applications, the end device is responsible for needs perception, initial cognitive analysis, and transmitting the initial analytical results to the cloud. The network transmits data, develops an understanding of the type of service being handled, and provides an analytical forecast. The cloud uses a cognitive model and algorithms to analyze the video and return the results to the end device for visual display. In this example, the device, pipe and cloud perform cognition and processing on video analysis applications from different angles and layers to achieve coordinated, layered intelligence.

Layered intelligence will bring several benefits. Elements of service control and decision making will shift down to the end device and intelligent network gateway. Both will accelerate decisions and, respectively, reduce the sudden shock that perceptual information creates on access networks and the shock that traffic causes on operators' backbone networks.

Moreover, a cloud with strong computing and storage abilities enables deep-level research and analysis on perceptual data and information from across large spans of time. It can also discover associations between information from different dimensions. Combined, these capabilities produce the most accurate results for decisions based on analysis and forecasts from a systemic perspective. It also allows information exchange and decisions to be adjusted based on the intelligent units distributed throughout the device and pipe.

Cognitive computing as a service

Although still in its infancy, the idea of cognitive computing as a service has already gained traction in the cloud computing market. For example, machine translation as a cloud service can translate cell phone apps into specified languages. In the future, service-based cognitive computing will form an important development direction for cognitive technology and business models, and evolve in a more refined, granular, and personalized direction.

High refinement and granularity denote the ability to perceive spatial and temporal differences in client needs. Personalization provides layered services that target client needs and problem-areas by embedding cognitive computing as a service in end equipment, devices, and sensors. These are interconnected, and can perceive client needs anywhere and anytime to provide the information and services clients may need.

A look at intelligent cloud e-assistant services for online businesses reveals clear differences between the requirements that different types of online businesses have on customer service e-assistants. For example, an online business for fresh produce might mention its need for a cloud e-assistant to Huawei REN. Huawei REN would first perceive the unique features of services corresponding to fresh produce, and then provide these features online via a cloud e-assistant with full knowledge about food, nutrition, and health. The e-assistant could then provide personalized advice and automated customer services for the online business client. Huawei REN is able to dynamically receive user feedback in line with the client's consumption records, and continue to improve the personalization and accuracy of the service.

Cognitive computing as a service allows applications

After almost 20 years, cognitive computing is finally growing – we hope that all major ICT companies and research organizations are passionate about jointly driving the cognitive computing era towards prosperity.

to offer a superior service experience based on cognitive computing technology.

Heterogeneous coordination

Cognitive systems form truly heterogeneous computing environments by virtue of components such as CPUs, GPUs, FPGAs, DSPs, and neuron processors, each with their own algorithms and tasks; for example, CPUs are adept at logical operations and GPUs are adept at numerical operations. Cognitive systems can automatically perceive the properties of higher-level applications based upon a priori knowledge and historical data. They can fully adapt applications and computing components, and use appropriate components for different use scenarios. Current heterogeneous cognitive systems focus on developing cognitive algorithms, meaning that software engineers use heterogeneous computing components to accelerate the speed of specified algorithms. The future will see an increasing number of cognitive algorithms based on heterogeneous system design.

Heterogeneous computing is highly efficient, requires little energy, and is especially useful for cognitive computing where demands on computing resources are high. Heterogeneous coordination can assist the technical evolution of converged software and hardware, and create new paths to realize device-pipe-cloud coordinated intelligence.

Brain-inspired intelligence

As the study of neuroscience has progressed, more researchers are beginning to understand that cognitive computing can and needs to obtain its inspiration from neuroscience to help develop new theories and methods, and improve machines' cognitive understanding of big data. An approach both domestic and international members of academia increasingly agree upon is the use of Turing machines to solve left brain problems, and brain-inspired computing to solve right brain problems. Cognitive computing requires the left and right sides of the brain to be integrated.

China's Brain Plan is active in three major areas: moving past Von Neumann's computing principle and model; making breakthroughs in the form of new components that evolve computing theory from mathematical models to physical models; and large scale, real-time systems that model biological brains. Brain-inspired intelligence has the potential to revolutionize cognitive computing and create a new paradigm of autonomous cognition.

A prosperity generator

Cognitive computing systems enjoy extremely broad business prospects, market value, and application scenarios, for example, smart devices, smart homes, smart robots, autonomous vehicles, SMEs, information centers, and intelligent office automation. They can also be used for financial analysis and gene analysis, and in the data centers of large enterprises to enable functions such as information retrieval, targeted ad pushes, mobile medicine, smart transportation, and smart power grids. After almost 20 years, cognitive computing is finally growing – we hope that all major ICT companies and research organizations are passionate about jointly driving the cognitive computing era towards prosperity.

Editor: Gary gary.marcus.maidment@huawei.com