INFORMATION, COMMUNICATION & COMPUTING

Fields of Expertise TU Graz

Source: istockphoto.com

Kay Uwe Römer,
Information, Communication & Computing
Source: Lunghammer – TU Graz

In 2017 the Bavarian state government announced the establishment of a new Technical University in Nuremberg – about 20 kilometers away from the University of Erlangen-Nuremberg, which has traditionally had a strong technical faculty, some say the strongest among all its faculties. The enthusiasm in Erlangen about this new Technical University and future competitor right in front of the door was very limited as reported by different media. Skillfully, the Bavarian state government later also decided to spend a large additional amount of money to also strengthen the University of Erlangen and the University of Applied Sciences Erlangen, such that criticism from the latter is rarely heard – at least not in public. The German Science Council – which usually analyses demand and makes recommendations before a new university is established – was not involved in the decision to establish this new university, but it was only after the political decision was made that a concept for the new Technical University was developed under the leadership of Wolfgang A. Herrmann (retired president of TU Munich) and the Science Council was asked to comment on that concept. History seems to be repeating now in Upper Austria (hopefully also the part with the additional money for the existing universities...)

In this issue of TU Graz research, Robert Legenstein – who was recently promoted to full professor and head of the Institute of Theoretical Computer Science (congratulations!) – writes about his research. Enjoy reading!

Robert Legenstein

The Future of Computing: Learning-Based, Energy-Efficient and Brain-Inspired

Computer science is at a turning point. Novel Machine Learning (ML) methods are revolutionizing how we think about computation and build computers. ML is believed to provide a path to artificial intelligence. Current ML systems however are energy hungry, which renders them unsuitable for edge applications and contributes to environmental problems. Hence, in the forthcoming transformation, energy-efficiency will play a major role, in which respect there is a lot to learn from the brain.

Computers are like empty boxes. They provide possibilities but need to be filled with content to be useful. The way we have done this filling since the advent of computing has only changed recently. An expert writes a computer program which defines the computation performed by the machine. It turned out that this way of filling the box has a severe drawback. It is extremely hard to design programs for many computational problems, in particular those that can be easily solved by humans.

COMPUTER SCIENCE IS AT A TURNING POINT

Within the last decade, tremendous progress has been made in this respect by using the Machine Learning (ML) approach. Instead of telling the computer how to exactly process each input in order to determine the desired output, one lets the computer figure it out by itself. For example, data sets exist with millions of images accompanied by labels that define which objects can be seen in each image. An ML algorithm can then figure out how to process input images in order to recognize objects in it. ML has been used to recognize objects or speech, to understand text, to play video games, or to control robots. Virtually all that progress was achieved by deep learning, an ML method based on deep neural networks. These successes have convinced many experts that deep learning provides a path to artificial intelligence.

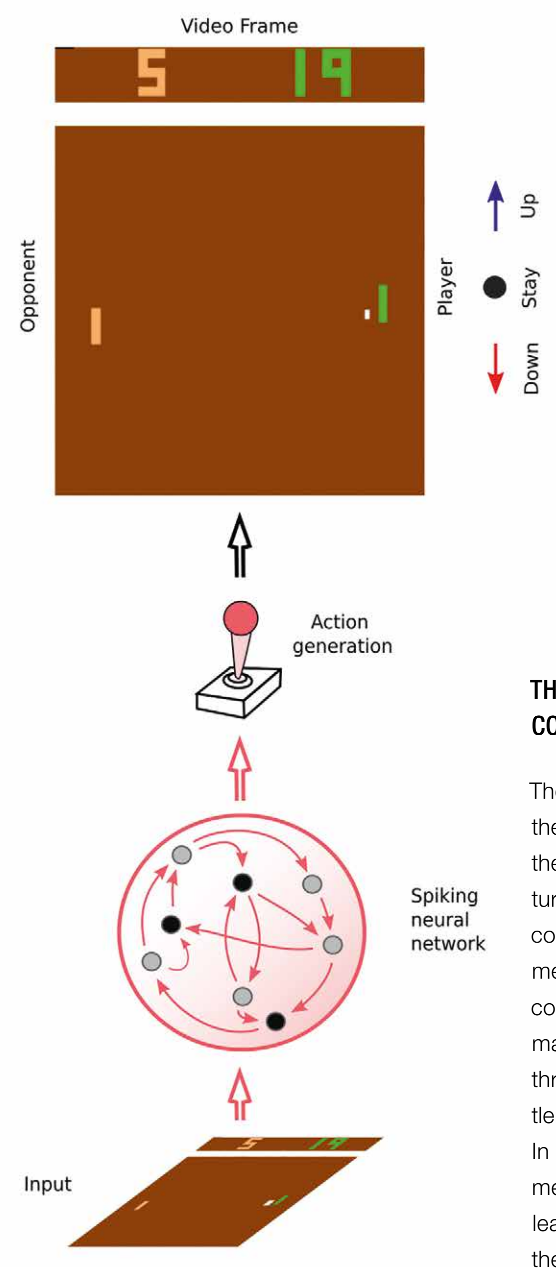

Figure 1: A spiking neural network learns to play the Atari game Pong.
Source: Franz Scherr

Note that ML does not just provide another tool in the toolbox of the computer scientist, it potentially preludes and fuels three fundamental paradigm shifts in computer science. First, a shift from the era of programming to the era of training. Second, a shift from computers as useful tools to computers as intelligent systems. Third, as discussed below, it brings about a complete reworking of the computing hardware we use.

THE ENVIRONMENTAL COSTS OF AI

The von-Neumann architecture has been the dominant computing architecture since the early days of computer science. It features a central processing unit (CPU) that communicates with a random access memory. However, the fact that during a computation, data has to be shuffled permanently between the CPU and memory through a tiny bus (the von-Neumann bottleneck), renders this architecture inefficient. In particular, it is not suited to the implementation of neural networks used for deep learning. Neural networks are inspired by the architecture of the brain, where simple computational units (neurons) form a complex network. Computation in this network is extremely parallel and memory is not separated from computation. Hence, there is no von-Neumann bottleneck. The currently used better option involves graphical processing units (GPUs). While they do provide significant speedups, they are very energy hungry. This is not only a problem for low-energy AI systems in edge devices, it is also becoming an environmental problem. For example, the training of a deep neural network for the GPT-2 model was estimated to emit about five times as much CO_2 as an average American car during its lifetime.

ENERGY-EFFICIENT BRAIN-INSPIRED COMPUTATION

To build AI systems with reasonable power budgets, novel technology is needed. Here, the brain can serve as a source of inspiration. While having the computing capabilities of a super computer, it consumes only 20 Watts. Many universities and big IT players such as Intel and IBM have thus developd so-called neuromorphic hardware which implements neural networks in a more brain-like manner. Like the brain, it uses spikes as the main communication unit between neurons. Spikes are binary pulses that are communicated only if necessary, hence making the computation much more power efficient.

The Institute of Theoretical Computer Science is at the forefront of this research. It has worked on the foundations for spiking neural networks (SNNs) for more than 20 years, and is now utilizing its expertise in several projects where energy-efficient brain-inspired hardware is developed.

NEUROMORPHIC COMPUTING AND BEYOND

Wolfgang Maass is leading a research team at the Institute of Theoretical Computer Science within the European flagship project – the Human Brain Project. The team has developed the machine learning algorithm e-prop (short for eligibility-propagation) for training SNNs. Previous methods achieved too little learning success or required enormous storage space. E-prop now solves this problem by means of a decentralized method copied from the brain. The method approaches the performance of the best-known learning methods for artificial neural networks [1]. For example, we used e-prop to train SNNs to play video games (Fig. 1).

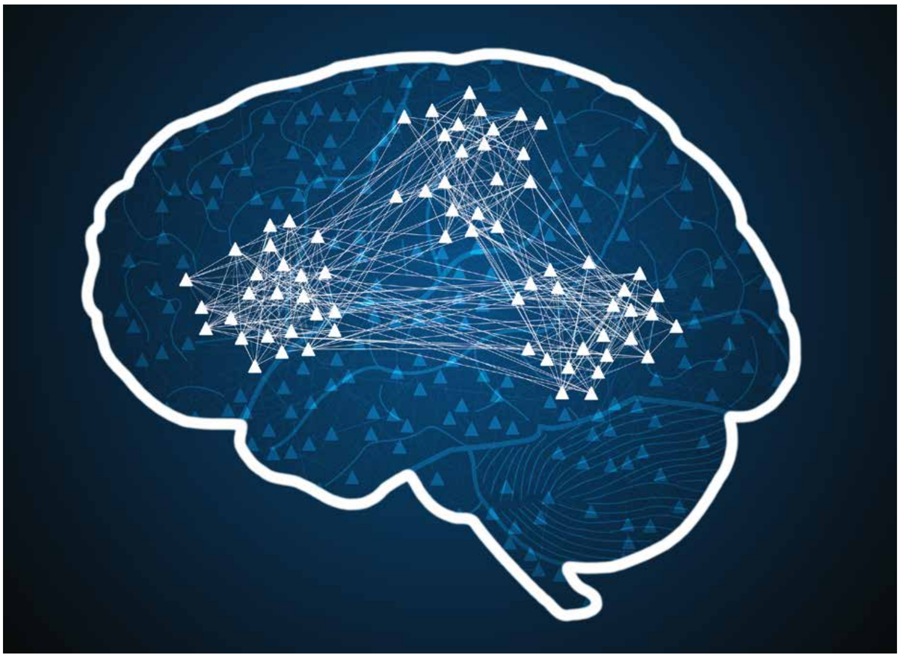

Figure 2: Neurons in the brain dynamically form interacting assemblies.
Source: Michael G. Müller

The energy-efficiency of neuromorphic systems can be further increased by using novel nano-scale circuit elements [2]. The potential of this approach is investigated in the Chist-era project SMALL (Spiking Memristive Architectures of Learning to Learn), launched this year and coordinated by the Institute of Theoretical Computer Science. Our team is developing computational paradigms that combine neuromorphic hardware developed by the University of Zurich with memristive devices from partners IBM Zurich research and University of Southampton.

Another computational substrate with high potential for ultra-fast computing are optical fibres. The Institute of Theoretical Computer Science is participating in the EU FET-Open project ADOPD (adaptive optical dendrites) that has just started. The project will develop ultra-fast computing units based on optical-fibre technologies that function according to the principles of information processing in dendritic branches of neurons in the brain.

Task 3: Three Supporting Facts
Sandra got the apple.
Sandra moved to the garden.
Sandra travelled to the bathroom.
Sandra dropped the apple.
John moved to the bathroom.
Where was the apple before the bathroom?
Answer: garden

Task 19: Path Finding
The kitchen is south of the office.
The bathroom is west of the garden.
The garden is west of the office.
The kitchen is east of the bedroom.
The hallway is north of the office.
How do you go from the kitchen to the garden?
Answer: north,west

Figure 3: Two question-answering tasks from the bAbI dataset. The neural network with a brain-inspired memory system observes a sequence of up to 320 sentences and has to provide the correct answer to a subsequent question.
Source: Thomas Limbacher

But there is much more to learn from the brain. It is known from neuroscientific experiments that neurons in the brain dynamically form assemblies to encode symbolic entities and relations between them. Recently, we have published work that shows how such assemblies could give rise to novel paradigms for symbolic processing in neural networks (Fig. 2) [3, 4]. In addition, the brain uses various memory systems to store information over many time scales. In an FWF project on stochastic assembly computations, we have recently shown how such memory systems can enable neural networks to solve demanding question-answering tasks (Fig. 3) [5].

In summary, the way we think about computation in computer science has remained largely decoupled from the way neuroscientists think about brain function. The ML revolution will change that. Future computing machines will benefit from our knowledge about how the brain is able to generate intelligent behaviour on a tiny energy budget.

[1] G Bellec et al., *Nature Communications*, 11:3625, 2020.
[2] R Legenstein, *Nature*, 521:37-38, 2015.
[3] CH Papadimitriou et al., *PNAS*, 117(25), 2020.
[4] MG Müller et al., *eNeuro*, 7(3), 2020.
[5] T Limbacher and R Legenstein, NeurIPS 2020, accepted.

Robert Legenstein is head of the Institute of Theoretical Computer Science.

Source: Lunghammer – TU Graz