An Overview over Neuromorphic Computing

Ole Kröger
Applied computer science
Heidelberg, Germany
hn198@uni-heidelberg.de

***Abstract*—Neuromorphic computing is a multidisciplinary field which covers different approaches of building a computer system inspired by the structure of the human brain. This paper gives an introduction into the topic with a short introduction over the relevant part of neuroscience for this topic. Additionally various architectures are compared. SpiNNaker as a cluster for simulating the brain with conventional hardware as well as specialized neuromorphic chips like HiCANN, TrueNorth and Loihi. Main reasons for this research field is to gain a better understanding of neuroscience and be able to construct intelligent machines with a structure comparable to the human brain in contrast to Deep Neural Networks. An additional main research question is the reduction of the power consumption of such systems.**

***Index Terms*—neuromorphic computing, brain-inspired computing, HiCANN, TrueNorth, Loihi, SpiNNaker**

I. Introduction

The human brain or biological brains in general work vastly different than most available computer systems. Computer architecture is faster and more precise in calculation tasks but is not designed to learn from the ground up. Advances in computer science and especially machine learning are adding functionality to this architecture to be able to have learnable systems. Deep neural networks (DNNs) which are the main pillar of this branch in building artificial intelligent systems work conceptually very different than the brain but are easily implemented with conventional hardware systems.

Working with the current hardware like central processing units (CPUs) and graphics processing units (GPUs) the software side was able to teach systems like AlphaGo which is a Go engine developed by DeepMind to achieve super human "intelligence" in a very narrow field [1]. It was able to defeat one of the best players of Go in 2016 by teaching itself using self play only knowing the rules of the game. Before this various game engines for games like Chess used knowledge from professional players and opening and endgame tables to achieve super human ratings. Regardless of these achievements the implementation they used are not influencing current research in the field of neuroscience as they work differently than the brain. The current systems are only narrow intelligent systems which excel at specific tasks like playing Go but they can not solve IQ tests or drive cars and those that can are not able to play Go.

The only general intelligent system is our human brain and neuromorphic computing tries to reengineer the brain and help understanding it which might result in huge advances not only in computer science but also in neuroscience.

It is proven that neuromorphic computing is more powerful than the current state of art artificial neural networks (ANNs) with respect to the amount of neurons needed to perform a certain task [2]. A lot of different companies and research groups are investing in this field for different reasons. The main points are to get a better understanding of the only truly general intelligent system we know of, the human brain as well as to actually use it where artificial intelligence is used today and reducing the power consumption of current high performance clusters (HPC) that simulate parts of the brain. Three main projects are currently developed:

- The HiCANN chip at the university of Heidelberg is one project which tries to get more understanding about long-term learning and therefore the achievement is an about 1,000 times faster than biological real time emulation of the human brain [3].
- TrueNorth which is developed by IBM has the goal of a real time emulation with a very reduced power consumption [4].
- Loihi from Intel is the newest neuromorphic chip from a bigger company [5]. They announced it in 2017 and is the 5th neuromorphic chip from Intel.

and startups are interested in this field as well i.e. Rain Neuromorphics which try to run conventional ANNs on neuromorphic chips [6]. SpiNNaker is a project at the university of Manchester which uses conventional computer architecture with a many core system to build an artificial brain [7]. Their main goal is a biological real time simulation using software to have faster control of the system if new insights about the brain emerge due to research in the field of neuroscience.

Neuromorphic computing is not a new field but was delayed in development due to the invention of backpropagation by David E. Rummelhart in 1986 [8]. This was a necessary step to train DNNs and therefore no brain inspired technology was needed for the invention of artificial intelligent systems. Due to the usage of Graphics processing units (GPUs) for the matrix multiplications needed for this algorithms the performance of DNNs was sufficient enough for the industry. The ending of Moore's Law in the recent past is one reason why neuromorphic computing as a different pathway is a reasonable research topic again [9]. Moore's Law predicted the exponential growth of transistors in computer chips but ended due to physical limitations.

Additionally research is still determined to understand the human brain for various reasons like curing diseases such as

Alzheimer's disease and it can drastically reduce the power consumption when current HPCs are used for tasks in artificial intelligence. The human brain consumes only about 20 watts whereas a simulation of it might use around 10MW. This is one of two major goals in research the other one is to speed up the simulation significantly to study the change of plasticity over a human life time or even evolutionary changes. The rest of the paper is structured in the following manner. In the next part the difference between current ANNs and the brain structure is explained. The following one will explain the difficulties in simulating the brain with current hardware systems. In section IV various hardware approaches are explained. The penultimate section shortly mentions the software used on those neuromorphic questions and then the paper is concluded.

II. Comparison to Deep neural networks

Deep neural networks have a well structured design by having several layers of neurons. The input neurons are all densely connected to the hidden layer and those are connected with the next. This continues until the output layer is reached. For each step from one layer to the next a matrix multiplication is performed to determine the state of the next neurons. Those matrix multiplications scale with $\mathcal{O}(n^3)$ which is reason why those networks are deep but not very wide. This layered structure makes it simple and each matrix multiplication is done at one time stamp which means that the computation is moving from left to right which is not comparable with the structure of the brain. The brain has a very sparse small world connected graph similar to friendship connections on social networks in real life as well as platforms like Facebook. Additionally each neuron in the brain is independent with respect to time from its neighboring neurons.

ANNs are working with weights on the connections between two neurons in consecutive layers which is similar to weights in neurons in brains but in ANNs those weights will be multiplied and summed up with the other incoming connections to form the weight of the current neuron. This weight is the input for the outgoing connections. This is fundamentally different from the biological system where the neurons only have a firing and a non firing state. The continuos component in the brain is the timing of the firing neurons and the differences between fire timing of previous and next neurons which are called presynaptic and postsynaptic neurons respectively. In a learning system the weight of a neuron would depend on the timing of the spikes in a way that if the postsynaptic neuron fires shortly after the presynaptic neuron the weight will be strengthen and if the postsynaptic neuron fires even though one presynaptic neuron did not, the weight of this particular connection will be reduced. This learning structure is called spike-timing-dependent plasticity (STDP). A synaptic influence can be negative which is referred to as inhibitory whereas a positive connection is excitatory. The biological structure of a neuron and its components is shown in Figure 1 which is a graphic from Wikipedia [10]. The dendrites are the input connection of the neuron which consist of the the main part called soma which is comparable to the processing unit of a neuron and the axon which is the outgoing connection to the next neurons. The synapse is between an axon terminal of the presynaptic neuron and one dendrite of a postsynaptic neuron. The relatively slow connection between neurons is the reason why silicon based technology can give in principal faster than real time simulation of the brain. The connection speed depends whether the axon has a myelin sheath and is ranging from 1m/s without this sheath to about 100m/s with it. The information encoding in the brain is done with timing the spikes which are the firing signals.

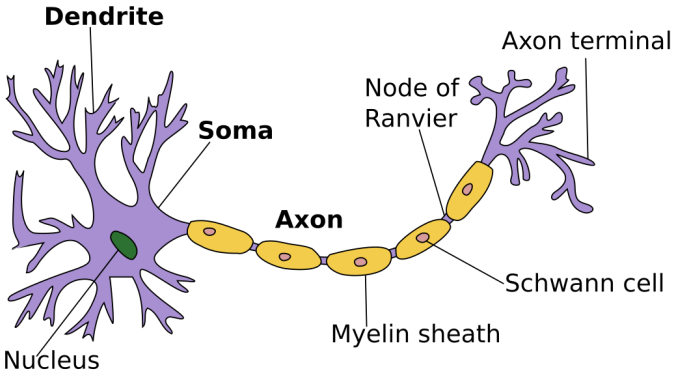

Fig. 1. Diagram of a biological neuron and its components from Wikipedia [10]

In general a neuron has a potential and a threshold if the potential is higher than the threshold a firing occurs and the threshold falls back to a resting potential. Additionally it has a resting time so that a spiking train (a sequence of spikes) can be differentiated. A spike has a current and a duration but these are not used to encode information only the firing start time and the the amount of spikes contain information.

In the following subsection the most basic neuron firing model namely the Leaky integrate and fire model (LIF) is explained.

A. Leaky integrate and fire model

The potential of a neuron i is time dependent and is represented as $u_i(t)$. The postsynaptic potential (PSP) ϵ_{ij} at the postsynaptic neuron j after neuron i fires at $t = 0$ is given by:

$$u_i(t) - u_{\text{rest}} =: \epsilon_{ij}(t)$$

The leaky integrate and fire model (LIF) is based on a physical model of a capacitor C and a resistor R connected in parallel. In this model the current $I(t)$ is:

$$I(t) = \frac{u(t) - u_{\text{rest}}}{R} + C\frac{\mathrm{d}u}{\mathrm{d}t}$$

Defining $\tau = RC$ gives the differential equation:

$$\tau_m \frac{\mathrm{d}u}{\mathrm{d}t} = -\left[u(t) - u_{\text{rest}}\right] + RI(t)$$

The voltage of the neuron rises if the presynaptic neuron fires and then declines to the resting potential. The following

equation assumes we have a voltage: $u = u_{\text{rest}} + \Delta u$ at $t = 0$ and set $I(t) = 0$ for $t > 0$.

$$u(t) - u_{\text{rest}} = \Delta u \exp\left(-\frac{t - t_0}{\tau_m}\right)$$

this describes the decaying part of a spike. The rising part can be described with a constant current in the time $0 < t < \Delta$ with:

$$u(t) = u_{\text{rest}} + RI_0\left[1 - \exp\left(-\frac{t}{\tau_m}\right)\right]$$

A firing signal of a presynaptic neuron would then result in the orange curve of Figure 2. If several of those spikes enter the dendrites of the neuron in an overlapping way in time they will be added up and if the resulting potential has a peak higher than the neuron threshold the postsynaptic neuron itself fires as visualized. For simplification the resting potential is set to 0 in this case.

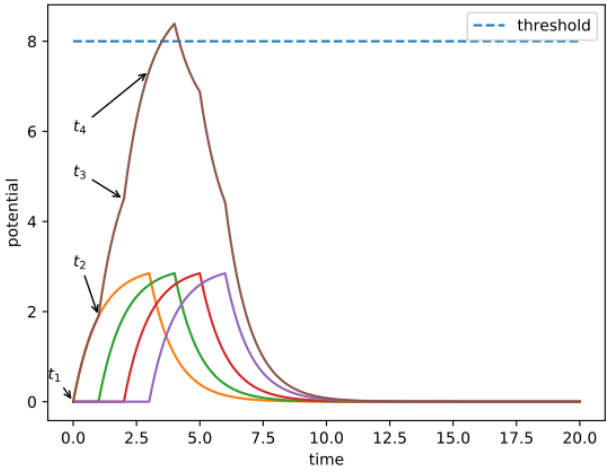

Fig. 2. Four incoming spikes that lead to a spike

This model is a simplified model of the current understanding of the biological neuron but is the base model. Different projects use different models which contain this model with a parameter setting as their base case.

Biological neurons are adapting to presynaptic spikes such that the time between two spikes increases with overstimulation. That is referred to as frequency adaptation.

III. Simulating on current hardware

There are projects which implement the above structure using conventional hardware as CPUs and GPUs. The most famous example is the SpiNNaker project developed at the university of Manchester [7]. In that project ARM chips with 18 cores and 200MHz clock speed each are used and connected in a toroidal structure. One of the SpiNNaker boards with 48 of those chips is represented in Figure 3 which is Figure 1 of [11].

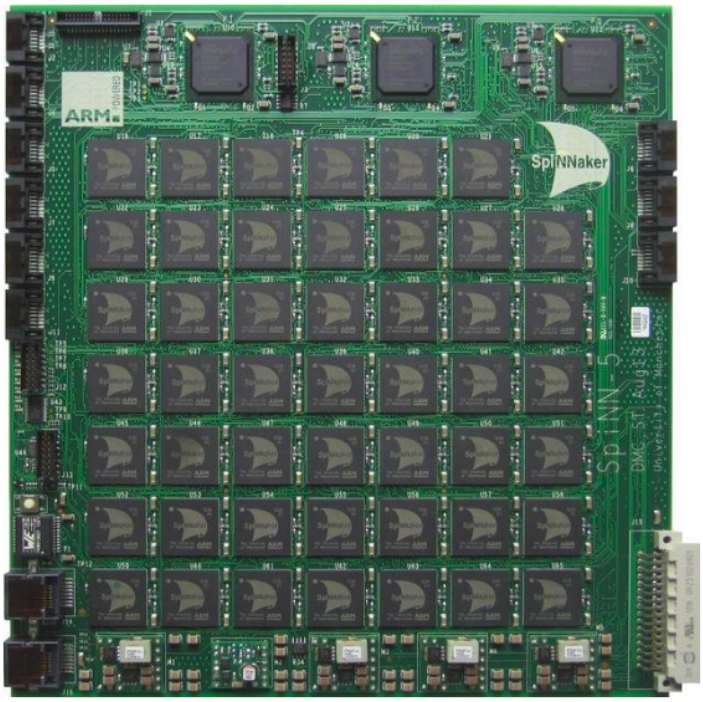

Fig. 3. One SpiNNaker board with 48 chips each with 18 cores. Figure 1 of [11]

Each core is able to simulate 1,000 neurons in biological real time (BRT). In 2018 the project team reached their goal of one million processor cores which is in theory able to simulate 1% of the human brain which has approximately 100 billion neurons [12]. An actual simulation of the human brain or part of it is still not achieved as the necessary full understanding of the human brain is not yet reached. In comparison to other hardware approaches this one uses standard cores and is more flexible as they are programmable and not hardware limited as other approaches. As a downside it consumes more power as they are not highly specialized for the task of simulating the human brain. According to a paper from 2018 each core uses about 56mW which results in $\approx 56kW$ for the whole system [13]. In comparison the human brain uses roughly 20W for 100 times more neurons. It can be still considered as a neuromorphic system if the definition is not too strict as the main goal is to simulate the human brain as compared to general purpose high performance clusters. The cores used are small integer cores in comparison to the bigger general purpose chips used in conventional hardware. Their structure minimizes the distance and therefore speed of accessing frequently needed data. Besides the computing cores the structure of the network is as well specialized for transferring a lot but small packets which represent spikes instead of bigger packets which are used in HPC.

Additionally a real time simulation of the brain is only part of what the research community wants to achieve. The goal is to have a system which is 1,000 times faster than

BRT which would help to understand learning at a much faster pace. If we would be able to start with an equivalent of a newborn brain a simulation of the first 20 years of life would be able to be simulated in a week with such a system. The reason why the hardware simulation in this is only in BRT is that the asynchronous spike timings are not an intended use case by the von Neumann architecture the current hardware is working with. In principle it is possible to achieve the 1,000 fold speed up due to the fact that biological neurons are relatively slow compared to the transfer speeds of electronics. The von Neumann architecture works with the principle of having input and output devices as well as distinction between memory units and processing units. Brains work different in this regard as there is no specific region in the brain which stores memories. Memories are stored in the different connections between neurons and one event in a life time is stored at different regions i.e visual information is stored in different connections than the smell attached to that situation. The transfer between processing and memory takes significant time which can be reduced if a non-von Neumann architecture is used which is the idea of neuromorphic chips explained in the next section. They can reduce the power consumption and/or increase the speed up compared to the SpiNNaker model.

IV. Neuromorphic Chips

A. *HiCANN chip*

The High Input Count Analog Neural Network (HiCANN) chip developed at the university of Heidelberg is as well as SpiNNaker part of the European Human Brain Project. It is also a part of the Brain-inspired multiscale computation in neuromorphic hybrid systems (BrainScaleS) project which aims at understanding and emulating the functionality of the human brain at various spatial and temporal scales.

Their research group takes a different approach than the SpiNNaker project as it uses specialized hardware and is able to run 1,000 times faster than BRT which enables the system to understand different biological aspects of the brain. The hardware used is more inspired by the LIF model described before. A neuron is represented by a system of a capacitor and resistor. In more detail the neuron circuit has two input connections for synapses one for inhibitory and the other for excitatory inputs.

It has a capacitor, a current input, a leaky unit for the LIF model as well as two units for the adaptation which closer resembles the functioning of a biological neuron. One part controls the resetting of the voltage after a spike occurs and then a connection to the network exists which transfers the spikes and handles the spike-timing-dependent plasticity. A schematic diagram of this is shown in Figure 4 from the original paper (Fig 1.) [3].

The synapses and neurons are together implemented in the Analog Network Core (ANC) which enables neurons to have a variable amount of synapses connected to them. Each neuron circuit has a connection to 224 synapses and up to 64 neuron circuits can be combined to one logical neuron which results

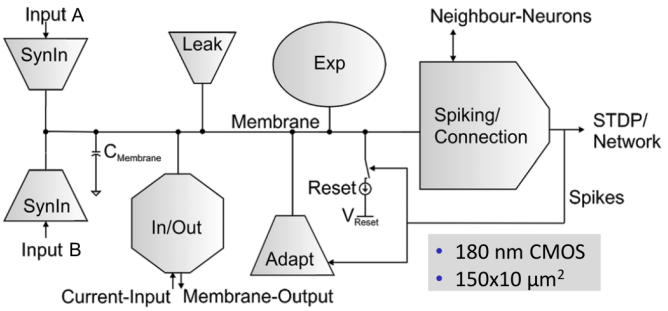

Fig. 4. Schematic diagram of the AdExp neuron circuit from [3]

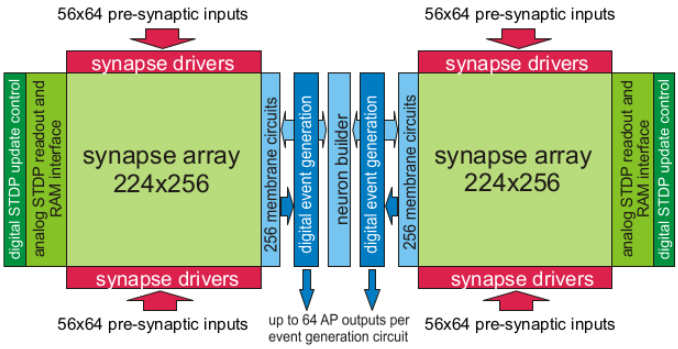

Fig. 5. Schematic diagram of the Analog network core (ANC) from [3]

in up to 14,336 presynaptic inputs per logical neuron which is close to the maximum of roughly 15,000 synapses a human neuron is connected to. A synapse in the system has a four bit address and a 2 bit enable signal such that a synapse can have 64 presynaptic inputs. The saved four bit address is compared to the input signal if the synapse is activated over the enable signal. If the synapse address is used, an analog input current will be multiplied with the 4 bit weight of the synapse using a digital analog converter (DAC) which then gets transferred to one of the inputs of the corresponding neuron. Those synapses are part of a matrix structure and each

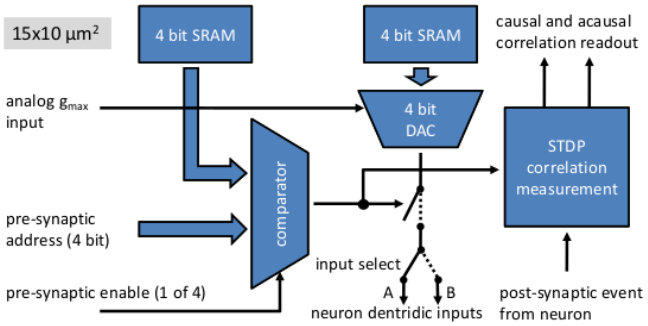

Fig. 6. Schematic diagram of a synapse from [3]

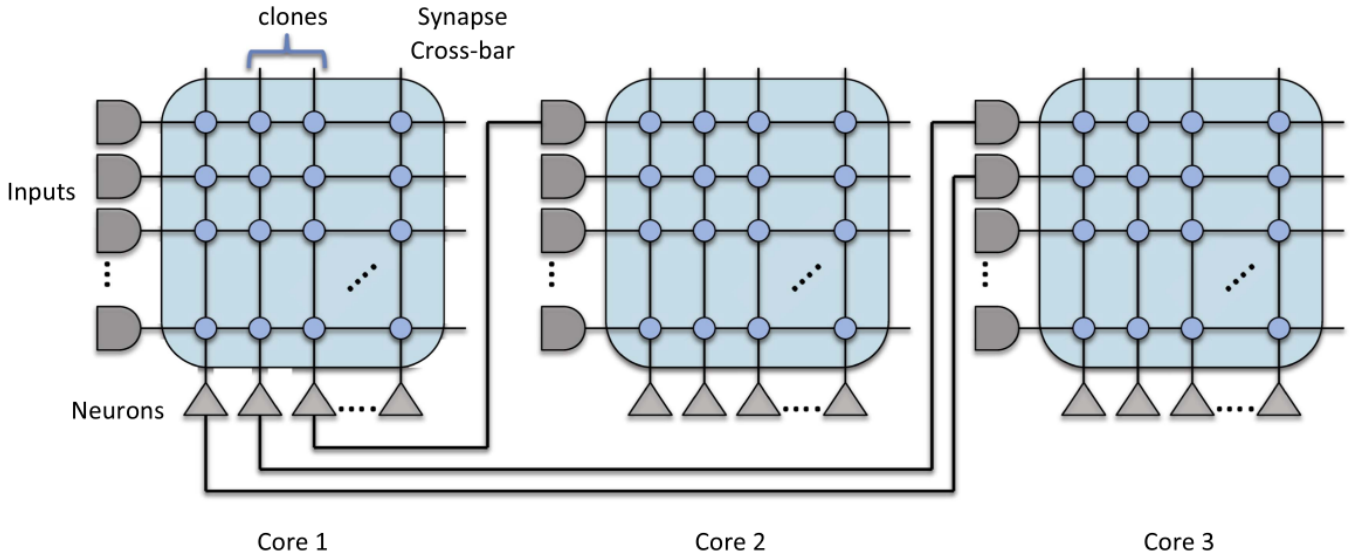

Fig. 7. Communication diagram in the TrueNorth architecture from [11] (Figure 1)

column is either connected to the inhibitory or the excitatory input of the neuron. The overall structure of the synapse array and the ANC is showed in Figure 5 which is copied from [3] Fig. 3. A single synapse is visualized in Figure 6 from the same original paper Fig 4.

The main goal of this project is to speed up the emulation of the human brain whereas the two following systems namely TrueNorth and Loihi specialize in reducing the power consumption of the emulation.

B. TrueNorth

The hardware company IBM is working on the neuromorphic chip TrueNorth since 2014 [14]. It has a different approach than the research group in Heidelberg and the goal is to be able to simulate the brain in real time instead of 1,000 times faster but drastically reduce the power consumption.

Each chip has 4096 cores and each core simulates 256 neurons which results in a bit over 1 million neurons per chip. Those chips can be connected via a network. Each core has a 256x256 crossbar where the inputs are axons and the output is a neuron. This connections are fully programmable with 256×256 synaptic connections which can be switched on or off and have weights associated to them. Each neuron-axon connection has also a parameter for an incoming delay. This delay can be used for coincidence detection which is used in the brain for the sound localisation for example [15]. In the BrainScaleS project in Heidelberg the coincidence detection is done by using a single neuron by saving previous spikes as described in [16]. The two ears of an animal hear sound at slightly different times which is then send over many different neurons with different axon length such that for sounds from different directions different neurons get the same input signal from both ears at different points in time. The neuron where this time difference is the smallest can be used to specify the angle from which the sound came from. This works without saving past spike events such that the simple overlap of the same spike train can be used instead of cross-referencing with previous spike trains.

The communication between neurons is done by a mesh network of size 64×64 on the 4096 core chip. This is a different technique than the brain uses as there the connections are one to one from one neuron to the end of an axon and then splits into single synapses. Compared to this the choice of IBM is reasonable as it reduces the amount of wires drastically and the same spike can be transferred more efficiently using less bandwidth until the spike packet is copied to the several ANCs at the last possible moment. The mesh network has five different directions, the four normal ones: north, east, south and west and additional local one as each core itself has 256 neurons. The network uses a deadlock-free dimension order routing which sends the packet over the x direction first and then in y direction if it is not a local package. Each spike has a package containing a 4 bit axonal delay, a 8 bit local address for the 256 target axons as well as two 9 bit addresses for the hops in the network. Those chips are connected in a network as well with different chips which is easily scalable. Their system has a specified power consumption of $26\,\mathrm{pJ}$ per synaptic event which is 3,250 times higher than the $8\,\mathrm{fJ}$ in the brain but nearly 176,000 times more efficient than a comparable simulation on a general-purpose microprocessor according to their paper [14]. In their paper they describe a real-time multi-object recognition system which uses a 400x240 pixel input. It can recognize five different classes using two input systems one for high resolution object recognition and one lower resolution system for detection. The network was trained offline therefore only the inference was tested on the neuromorphic chip itself. This test run on a single chip with a bit more than 1 million neurons and 256 million synapses. The architecture is shown in Figure 7 from [11] where it is Figure 1.

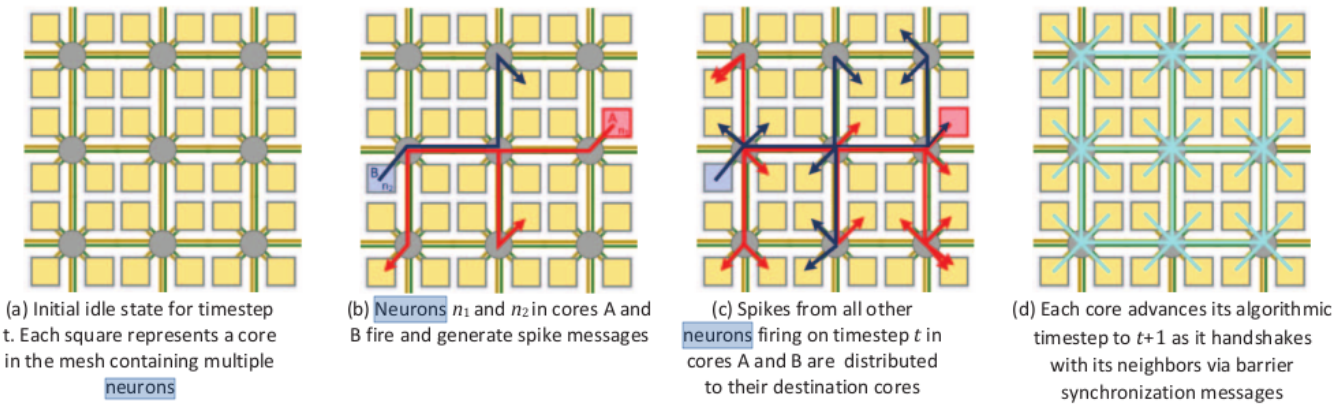

Fig. 8. Mesh operation in the Loihi chip from [5] (Figure 2)

C. Loihi

Loihi is the counterpart of TrueNorth from the chip manufacturer Intel. One Loihi chip contains 128 neuromorphic cores, each implements 1,024 biological neurons. It is able to simulate 130 million synapses using up to 1,000 synapses per neuron and can be combined with up to 16,384 chips in a mesh network. The research group managed to include on chip learning which is able use STDP but also more complicated learning rules. It is able to access the spiking rate to perform this learning which gives this chip more flexibility than for example TrueNorth where training has to be done offline.

In Loihi in each time step each core is checking each of the 1,024 neurons whether they generate a spike and then transfers this spike events over the network-on-chip (NoC). The messages are distributed using dimension-order routing. At the end of a time stamp they use a barrier synchronization algorithm to ensure that all cores receive their incoming spikes before they update to the next time stamp. This is visually shown in Figure 8 which appeared in [5] as Figure 2.

Table I shows the power consumption of Loihi for different events as stated in [5].

Loihi also has the possibility to have axon delays like TrueNorth but it is not specified how it is achieved exactly in their system.

TABLE I
POWER CONSUMPTION OF THE LOIHI CHIP FOR DIFFERENT EVENTS

Event	Energy consumption
Tile hop (E-W / N-S)	3.0 pJ / 4.0 pJ
Synaptic spike op (min)	23.6 pJ
Neuron update (active / inactive)	81 pJ / 52 pJ

In the TrueNorth paper only the power consumption per synaptic event is given with 26 pJ which might not contain the tile hops then Loihi and TrueNorth have a similar power consumption.

V. SOFTWARE

There are different software packages available to program the neuromorphic chips mentioned in the previous section. In this section only the two most often used softwares namely Nengo and PyNN are mentioned [17] [18].

- PyNN which is pronounced pine is a simulator-independent language which can run code on different platforms including the HiCANN chip as well as on SpiNNaker. It allows general as well as specific parameters for the overall structure of the network down to single neuron behavior. Besides running on neuromorphic hardware it is also supported by several simulators like Neuron, NEST and Brian.
- Nengo provides a graphical interface besides the traditional scripting interface and can be used on SpiNNaker as well as Loihi from Intel. The newest version 2.1 is written in Python and has its origin in the Computational Neuroscience Research Group of Prof. Chris Eliasmith at the university of Waterloo. Eliasmith is the author of the book “How to build a brain“ [19].

In general it is very valuable to be able to simulate it on conventional hardware as well as on as many neuromorphic hardware as possible to be able to use the tool in an efficient way.

VI. CONCLUSION

In this section the three different systems namely SpiNNaker, HiCANN and TrueNorth are compared. They all have upsides and downsides compared to the others such that there is no system which is clearly preferable.

- SpiNNaker is the most robust system with respect to new knowledge discovered in neuroscience as the system is programmable such that in theory every kind of neuron can be represented as long as the memory is big enough and it still uses a spiking system which is not likely to be wrong. Nevertheless the power consumption of this system is clearly the highest as it is not specialized enough to be able to reduce it as much as possible. It

TABLE II
COMPARISON BETWEEN BRAIN AND VARIOUS HARDWARE SOLUTIONS

	Human Brain	SpiNNaker	HiCANN/BrainScaleS	TrueNorth	Loihi
Neurons	100 billion	1 billion	$\approx$ 4 million	1 million	131 072
- using	1.4 kg	10 19-inch racks	20 wafers with 384 HiCANNs each	one chip [1]	one chip [1]
Mean number of synapses per neuron	7000	Programmable	224	256	1 000
Max number of synapses per neuron	$\approx$ 15 000	Programmable	14 336	256	1 000
Energy consumption per spike	8 fJ	4 nJ	0.1 nJ–10 nJ [2]	26 pJ	$>$ 23.6 pJ
- compared to brain	1	500 000	12 500–1 250 000[2]	3 250	$>$ 2 950
Speed up compared to brain	1	1	$10^3 - 10^5$	1	1
Run time plasticity	Yes	Programmable	STDP	No	STDP
Neuron model	Diverse	Programmable	Adaptive exponential leaky integrate and fire	Leaky integrate and fire	Leaky integrate and fire

[1] can be connected but numbers for how many could not be found
[2] depending on the resources and network activity [20]

will never be able to have a consumption similar to the 20W of the human brain.

- In comparison the HiCANN chip is the only one of those three which is able to simulate the brain in an accelerated pace with a speed of about 1,000 times faster than BRT. The specialized structure makes it harder to incorporate new insights from neuroscience if new neuron types and models emerge. In that case a similar system must be constructed and probably the costs are smaller than the initial project but it takes time to construct such a chip and it will have additional hardware costs compared to the reprogrammable SpiNNaker system.
- TrueNorth takes the route of being as power efficient but it lacks the possibility of run time plasticity. The maximum number of neurons per synapse could not be found in their papers where HiCANN has the ability to have a flexible number of synapses connecting to a neuron with a maximum comparable to the observed 15,000 in the human brain. TrueNorth has a mean number of 256 synapses per neuron.
- Loihi takes a similar route as TrueNorth in being power efficient and limits the number of synapses per neuron to 1,000 but it might as well be possible to combine two artificial neurons on the same core which is not mentioned in the paper. If this is not possible it is a downside compared to the HiCANN approach. Compared to TrueNorth it brings the functionality of on-chip learning.

In conclusion there is a diversity of systems available which are specialized for different approaches. The SpiNNaker system can be used if new knowledge in neuroscience emerges and tested. If it is successful it can be either implemented into HiCANN, TrueNorth and Loihi if possible or new chips might need to be realized but the knowledge of those two systems will be invaluable in that case. The HiCANN chip will be the chip which can bring new insights of long term learning and plasticity in the brain whereas the insights in the construction of TrueNorth and Loihi will be very helpful once a better understanding of the biology is achieved and then can construct a power efficient system.

In Table II the different hardware systems are compared as well to see the biggest differences as well as pros and cons for each of them.

All systems have drawbacks compared to the brain most notably in the understanding of learning and the power efficiency but all bring projects bring something special into the understanding of computer and neuroscience and can be used as a starting point for further improvements.

In the future the systems will be able to simulate a higher number of neurons and different neuron types to obtain a better understanding of the brain especially the human brain. Different systems will improve our understanding in neuroscience in different ways and once a new ideas in neuroscience arise the systems are able to simulate them to give feedback to neuroscientist. It is not possible at the moment as well as not ethical to manipulate the brain directly but using neuromorphic hardware we can try different ideas and can see how this would change the performance of the brain and during this testing get new ideas of how the brains works such that the approaches described have a relevance in the goal of understanding the currently only known general intelligent system we know of.

REFERENCES

[1] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, January 2016.

[2] Wolfgang Maass. Networks of spiking neurons: The third generation of neural network models. *Neural Networks*, 10(9):1659–1671, December 1997.

[3] J. Schemmel, D. Briiderle, A. Griibl, M. Hock, K. Meier, and S. Millner. A wafer-scale neuromorphic hardware system for large-scale neural modeling. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pages 1947–1950, May 2010.

[4] S. K. Esser, A. Andreopoulos, R. Appuswamy, P. Datta, D. Barch, A. Amir, J. Arthur, A. Cassidy, M. Flickner, P. Merolla, S. Chandra, N. Basilico, S. Carpin, T. Zimmerman, F. Zee, R. Alvarez-Icaza, J. A. Kusnitz, T. M. Wong, W. P. Risk, E. McQuinn, T. K. Nayak, R. Singh, and D. S. Modha. Cognitive computing systems: Algorithms and applications for networks of neurosynaptic cores. In *The 2013 International Joint Conference on Neural Networks (IJCNN)*, pages 1–10, August 2013.

[5] M. Davies, N. Srinivasa, T. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, Y. Liao, C. Lin, A. Lines, R. Liu, D. Mathaikutty, S. McCoy, A. Paul, J. Tse, G. Venkataramanan, Y. Weng, A. Wild, Y. Yang, and H. Wang. Loihi: A Neuromorphic Manycore Processor with On-Chip Learning. *IEEE Micro*, 38(1):82–99, January 2018.

[6] Rain Neuromorphics. http://rain-neuromorphics.com/.

[7] E. Painkras, L. A. Plana, J. Garside, S. Temple, S. Davidson, J. Pepper, D. Clark, C. Patterson, and S. Furber. SpiNNaker: A multi-core System-on-Chip for massively-parallel neural net simulation. In *Proceedings of the IEEE 2012 Custom Integrated Circuits Conference*, pages 1–4, September 2012.

[8] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533, October 1986.

[9] G. E. Moore. Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff. *IEEE Solid-State Circuits Society Newsletter*, 11(3):33–35, September 2006.

[10] Quasar Jarosz at English Wikipedia. English: Recreated File:Neuron-no labels2.png in Inkscape and hand-tuned to reduce filesize. Created by Quasar (talk) 19:59, 11 August 2009 (UTC). https://commons.wikimedia.org/wiki/File:Neuron_Hand-tuned.svg, August 2009.

[11] Steve Furber. Large-scale neuromorphic computing systems. *J. Neural Eng.*, 13(5):051001, August 2016.

[12] 'Human brain' supercomputer with 1 million processors switched on for first time. https://www.manchester.ac.uk/discover/news/human-brain-supercomputer-with-1million-processors-switched-on-for-first-time/.

[13] Chen Liu, Guillaume Bellec, Bernhard Vogginger, David Kappel, Johannes Partzsch, Felix Neumärker, Sebastian Höppner, Wolfgang Maass, Steve B. Furber, Robert Legenstein, and Christian G. Mayr. Memory-Efficient Deep Learning on a SpiNNaker 2 Prototype. *Front. Neurosci.*, 12, 2018.

[14] Paul A. Merolla, John V. Arthur, Rodrigo Alvarez-Icaza, Andrew S. Cassidy, Jun Sawada, Filipp Akopyan, Bryan L. Jackson, Nabil Imam, Chen Guo, Yutaka Nakamura, Bernard Brezzo, Ivan Vo, Steven K. Esser, Rathinakumar Appuswamy, Brian Taba, Arnon Amir, Myron D. Flickner, William P. Risk, Rajit Manohar, and Dharmendra S. Modha. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, 345(6197):668–673, August 2014.

[15] Lovorka Pantic, Joaqu n J Torres, and Hilbert J Kappen. Coincidence detection with dynamic synapses. *Network: Comput. Neural Syst.*, 14(1):17–33, February 2003.

[16] Johannes Schemmel, Laura Kriener, Paul Müller, and Karlheinz Meier. An accelerated analog neuromorphic hardware system emulating NMDA- and calcium-based non-linear dendrites. *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 2217–2226, 2017.

[17] Trevor Bekolay, James Bergstra, Eric Hunsberger, Travis DeWolf, Terrence C. Stewart, Daniel Rasmussen, Xuan Choo, Aaron Voelker, and Chris Eliasmith. Nengo: a Python tool for building large-scale functional brain models. *Front. Neuroinform.*, 7, 2014.

[18] Andrew P. Davison, Daniel Brüderle, Jochen M. Eppler, Jens Kremkow, Eilif Muller, Dejan Pecevski, Laurent Perrinet, and Pierre Yger. PyNN: a common interface for neuronal network simulators. *Front. Neuroinform.*, 2, 2009.

[19] Chris Eliasmith. *How to Build a Brain: A Neural Architecture for Biological Cognition*. OXFORD UNIV PR, Oxford, new edition, July 2013.

[20] Sebastian Schmitt, Johann Klaehn, Guillaume Bellec, Andreas Gruebl, Maurice Guettler, Andreas Hartel, Stephan Hartmann, Dan Husmann, Kai Husmann, Vitali Karasenko, Mitja Kleider, Christoph Koke, Christian Mauch, Eric Mueller, Paul Mueller, Johannes Partzsch, Mihai A. Petrovici, Stefan Schiefer, Stefan Scholze, Bernhard Vogginger, Robert Legenstein, Wolfgang Maass, Christian Mayr, Johannes Schemmel, and Karlheinz Meier. Neuromorphic Hardware In The Loop: Training a Deep Spiking Network on the BrainScaleS Wafer-Scale System. *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 2227–2234, May 2017. arXiv: 1703.01909.