Execution Model: Neuromorphic Computing

Armin Schaare

Arbeitsbereich Wissenschaftliches Rechnen
Fachbereich Informatik
Fakultät für Mathematik, Informatik und Naturwissenschaften
Universität Hamburg

2015-12-07

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

informatik die zukunft

Outline

1 Introduction

2 Benefits

3 State of the Art

4 Upcoming Technologies

5 Conclusion

6 Literature

Introduction

We all have a high performance computer in our bodies

Introduction

We all have a high performance computer in our bodies

The brain

Figure: Model of a brain [Nuf]

The Human Brain

- Consists of ~85 billion neurons
- Neurons connected to ~10000 synapses
  - All neurons connected via 3-4 synapses
- Neurons fire at up to ~1 kHz

- Performance of up to about 1 PFlop/s
- Energy consumption of a dim lightbulb

What & Why

- What is Neuromorphic Computing?
  - Hardware concept
  - Mimics nervous systems/brains

What & Why

- What is Neuromorphic Computing?
  - Hardware concept
  - Mimics nervous systems/brains
- Why do we need Neuromorphic Computing?
  - Engineering lessons to be learned
  - Better suited for special tasks
  - Lots of applications

Benefits

Efficiency for special tasks in terms of...

- ...speed
- ...energy
- ...space

Benefits

Efficiency for special tasks in terms of...

- ...speed
- ...energy
- ...space

Better than traditional computation
by orders of magnitude!

High performance computer 'K', Japan

Figure: Supercomputer 'K' by RIKEN [Ins13]

High performance computer ’K’, Japan

- Peak Performance: ~11.3 PFlop/s
- Power Consumption: ~12.7 MW
- Memory: ~1.5 PB

Brain simulation on 'K'

- Researchers simulated 1 second of 1% of brain activity
  - It took ~40 minutes
  - Consumed ~30.5 Gigajoule or ~8,500 kWh
  - Used ~1 PB of memory

Brain simulation on 'K'

- Researchers simulated 1 second of 1% of brain activity
  - It took ~40 minutes
  - Consumed ~30.5 Gigajoule or ~8,500 kWh
  - Used ~1 PB of memory

All this with a very simplified model of neurons and synapses

Speed efficiency

Assuming linear scaling

- It would take 66h 40min to simulate 1 second brain activity
- Even though performance of 'K' ∼10 times bigger than brain

Speed efficiency

Assuming linear scaling

- It would take 66h 40min to simulate 1 second brain activity
- Even though performance of 'K' ~10 times bigger than brain

A Human Brain is ~**240,000** times faster than 'K'

Energy efficiency

Again, assume linear scaling

- Would use up 850 MWh to simulate 1 second brain activity
- The brain would only consume 6 mWh

Energy efficiency

Again, assume linear scaling

- Would use up 850 MWh to simulate 1 second brain activity
- The brain would only consume 6 mWh

Human brains are ~**140 million** times more energy efficient than 'K'

Space efficiency

Assume every synapse represents one bit

- Brain has about 400 trillion synapses
- Results in a storage capacity of 50 Terabyte

Space efficiency

Assume every synapse represents one bit

- Brain has about 400 trillion synapses
- Results in a storage capacity of 50 Terabyte

It's not that easy to calculate brain storage capacity

(estimates at **2.5 Petabyte**)

Realization

Neuromorphic computation hardware is realized...

- Using digital or analog circuits
- Which mimic nervous systems/brains
- Very large scale integrated in microchips

Components of neuromorphic chips

- Analog or digital processor cores
- Chip interface
- Asynchronous package routing system
- Fault tolerance relaying
- Architecture specific parts

Asynchronous processing & fault tolerance

Asynchronous processing

- Event driven processing of packages
- Lessens energy consumption by orders of magnitude
- Weakens/eliminates Von-Neumann-Bottleneck

Asynchronous processing & fault tolerance

Asynchronous processing

- Event driven processing of packages
- Lessens energy consumption by orders of magnitude
- Weakens/eliminates Von-Neumann-Bottleneck

Fault tolerance:

- Rerouting around broken neurons
- Implemented on chip

Current teams and projects

- Human Brain Project (HBP)
  - SpiNNaker
  - Spikey
- Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE)
  - TrueNorth

Human Brain Project - SpiNNaker

- Short for "Spiking Neural Network Architecture"
- Entirely digital signal processing
- Chip utilises 18 ARM9 processors
- Die area of only 102 mm^2
- Functional SpiNNaker prototype chip in 2009
- First fully functional chips delivered in 2011

The SpiNNaker Chip

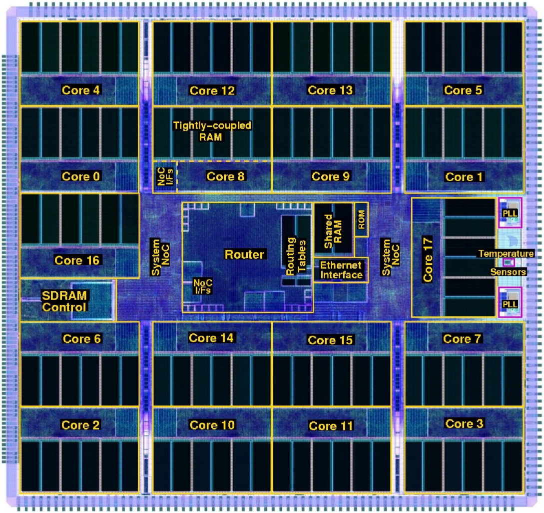


Figure: SpiNNaker Chip [Adv12b]

The SpiNNaker machine

- 1 million processor cores
- 1 billion neurons
- 1 trillion synapses
- Consumes less than 50KW on average
- Simulation takes place in realtime

The SpiNNaker machine

- 1 million processor cores
- 1 billion neurons
- 1 trillion synapses
- Consumes less than 50KW on average
- Simulation takes place in realtime

Considerably better than 'K' but still not quite human-brain-level

Human Brain Project - Spikey

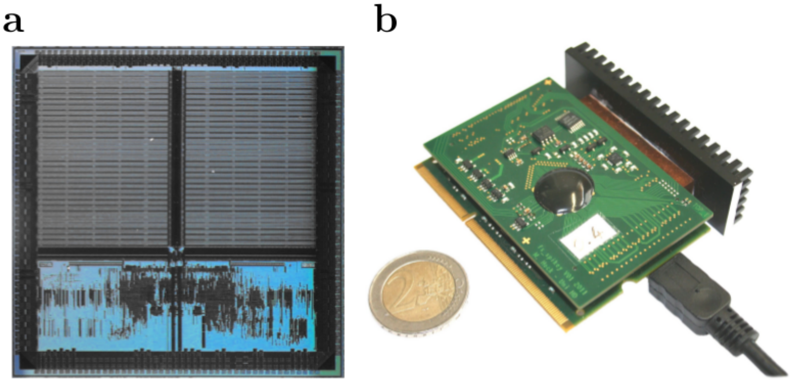


Figure: Spikey chip (a) and system with chip under sealing (b) [T+15]

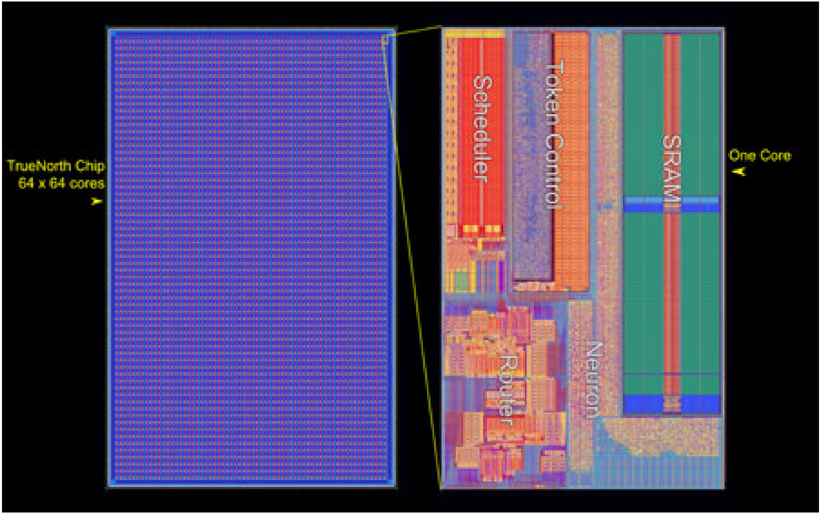
Human Brain Project - Spikey

- 25 mm^2 VLSI chip
- Analog hardware neuron and synapse realization
- Emulates 382 neurons with 256 synapses each
- Firing frequency of neurons 10^4 to 10^5 times higher than brain
- Even though analog, no memristors

SyNAPSE - TrueNorth

- Chip simulates 1 million neurons and 256 million synapses
- Consists of 4096 cores
- All digital approach
- Consumes less than 70 mW while simulating neural networks
- Already built systems of 16m neurons and 4b synapses
- Goal of 4b neurons and 1t synapses system, consuming 4KW

SyNAPSE - TrueNorth



The Memory Resistor

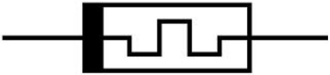


Figure: Symbol of a memristor [Jos14]

- Two terminal fundamental circuit Element
- The more intensely it is used, the lesser its resistance
- Raising resistance again by reversing current
- Remembers resistance when voltage turned off (non-volatile)

Memristor half adder

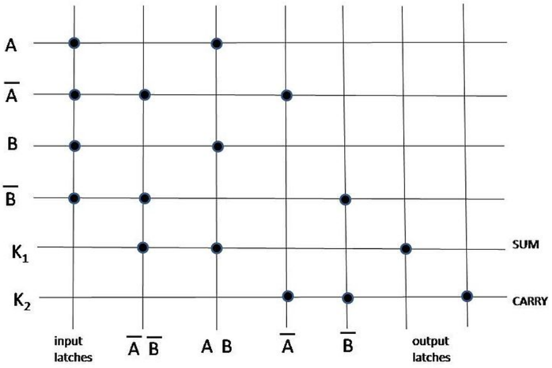


Figure: Crossbar latch architecture for half-adder [Blm08]

"The Machine" - Hewlett Packard

- Revolutionary computer architecture with usage of memristors
- 'K' uses 12,600 KW with 28.8 GUPS
- 'The Machine' should only consume 160 KW for 160 GUPS
- HP suggests 'The Machine' for exascale computing

The Memory Resistor

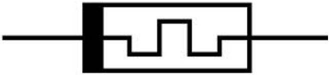


Figure: Symbol of a memristor [Jos14]

- Two terminal fundamental circuit Element
- The more intensely it is used, the lesser its resistance
- Raising resistance again by reversing current
- Remembers resistance when voltage turned off (non-volatile)

The Memory Resistor

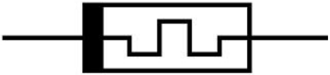


Figure: Symbol of a memristor [Jos14]

- Two terminal fundamental circuit Element
- The more intensely it is used, the lesser its resistance
- Raising resistance again by reversing current
- Remembers resistance when voltage turned off (non-volatile)

Essentially a model of a synapse

The Neuristor

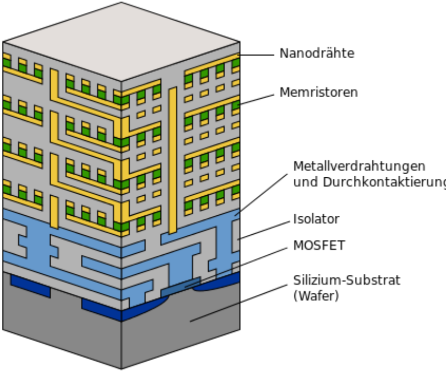


Figure: Neuristor concept [Mov14]

Memristors in Neuromorphic Computing

- One memristor is used for one synapse
- Memristors are fundamental circuit elements, therefore
  - small (cubes of 3nm edge length)
  - energy efficient
- Much faster than traditional approach of many transistors

Memristors in Neuromorphic Computing

- One memristor is used for one synapse
- Memristors are fundamental circuit elements, therefore
  - small (cubes of 3nm edge length)
  - energy efficient
- Much faster than traditional approach of many transistors

A very promising concept for Neuromorphic Computing

Prospect of Memristors in NC

- Current development of CrossNets in hybrid NC
- Faster processing rate than human brain
- Possibility of higher neuron density
- Higher but still manageble energy consumption

Prospect of Memristors in NC

- Current development of CrossNets in hybrid NC
- Faster processing rate than human brain
- Possibility of higher neuron density
- Higher but still manageble energy consumption

University of California, Santa Barbara active research:

- 100 neuron memristive NC system
- Able to do simple image recognition tasks

Applications

- Further extension of Moore's Law
- Understanding the human brain
- Brain prosthetics for neurodegenerativ diseases
- Face, Speech, Object recognition
- Language interpretation
- Robotic terrain manuevering
- Virtually any tasks where humans are better than computers

Conclusion

- Neuromorphic computing is very young
- Has much potential to be revolutionary
- Breakthroughs expected in less than 15 years

Conclusion

- Neuromorphic computing is very young
- Has much potential to be revolutionary
- Breakthroughs expected in less than 15 years

"As engineers, we would be foolish to ignore the lessons of a billion years of evolution"

— Carver Mead, 1993

Questions?

Literature I

[Adv12a] Advanced Processor Technologies Research Group, Manchester. SpiNNaker Home Page. http://apt.cs.manchester.ac.uk/projects/SpiNNaker/, 2012. [Online; accessed 30-November-2015].

[Adv12b] Advanced Processor Technologies Research Group, Manchester. spinn_labeled_bw.png. http://apt.cs.manchester.ac.uk/projects/SpiNNaker/SpiNNchip/, 2012. [Online; accessed 30-November-2015].

[And07] Andreas Grüebl. VLSI Implementation of a Spiking Neural Network. http://www.kip.uni-heidelberg.de/Veroeffentlichungen/download.php/4630/ps/agruebl_diss_kip.pdf, 2007. [Online; accessed 02-December-2015].

[Blm08] Blm19732008. Crossbar latch architecture for half-adder. https://en.wikipedia.org/wiki/Crossbar_latch, 2008. [Online; accessed 05-December-2015].

[Chr99] Chris Diorio. Why Neuromorphic Engineering? http://homes.cs.washington.edu/~diorio/Talks/InvitedTalks/Telluride99/index.htm, 1999. [Online; accessed 16-November-2015].

[Den89] Dennis Coon. Neurons. http://www.noteaccess.com/APPROACHES/ArtEd/ChildDev/1cNeurons.htm, 1989. [Online; accessed 17-November-2015].

[Dha14] Dharmendra S. Modha. Introducing a Brain-inspired Computer TrueNorth's neurons to revolutionize system architecture. http://www.research.ibm.com/articles/brain-chip.shtml, 2014. [Online; accessed 04-December-2015].

Literature II

[Har14] Harald Weiss. HP will Computerwelt mit "The Machine" revolutionieren. http://www.heise.de/ix/meldung/HP-will-Computerwelt-mit-The-Machine-revolutionieren-2220581.html, 2014. [Online; accessed 04-December-2015].

[Ins13] Institute of Neuroscience and Medicine Computational and Systems Neuroscience, Jülich. Largest neuronal network simulation to date achieved using Japanese supercomputer. http://www.fz-juelich.de/SharedDocs/Pressemitteilungen/UK/EN/2013/13-08-02LargestSimulation.html, 2013. [Online; accessed 19-November-2015].

[J+01] Jacqueline Ling et al. Power of a Human Brain. http://hypertextbook.com/facts/2001/JacquelineLing.shtml, 2001. [Online; accessed 17-November-2015].

[Jam12] James Randerson. How many neurons make a human brain? Billions fewer than we thought. http://www.theguardian.com/science/blog/2012/feb/28/how-many-neurons-human-brain, 2012. [Online; accessed 17-November-2015].

[Jim10] Jim Walker. Memristors and the future. http://www.nobeliefs.com/memristor.htm, 2010. [Online; accessed 05-December-2015].

[Jos14] Jose Fritz. Maybe the Memristor. http://tenwatts.blogspot.de/2014_08_01_archive.html, 2014. [Online; accessed 28-November-2015].

[Kon11] Konstantin K. Likharev. CrossNets: Neuromorphic Hybrid CMOS/Nanoelectronic Networks. http://mysbfiles.stonybrook.edu/~klikharev/nano/SAM11.pdf, 2011. [Online; accessed 05-December-2015].

Literature III

[Luk] Luke Mastin. NEURONS & SYNAPSES. http://www.human-memory.net/brain_neurons.html. [Online; accessed 17-November-2015].

[Mic15] Michael Byrne. Neuromorphic Circuits Don't Just Simulate the Brain, They Outrun It. http://motherboard.vice.com/read/neuromorphic-circuits-dont-just-simulate-the-brain-they-outrun-it-2, 2015. [Online; accessed 05-December-2015].

[Mov14] MovGP. Hybrid CMOS Memristor Schaltung 3D. https://de.wikipedia.org/wiki/Neuristor, 2014. [Online; accessed 28-November-2015].

[Nuf] Nuffield Foundation. brain.jpg. http://www.nuffieldfoundation.org/science-society/watching-brain-working. [Online; accessed 16-November-2015].

[oTT10] California Institute of Telecommunications and Information Technology. Finding the Missing Memristor - R. Stanley Williams. https://www.youtube.com/watch?v=bKGhvKyjgLY, 2010. [Online; accessed 28-November-2015].

[P+14] Paul A. Merolla et al. A Million Spiking-Neuron Integrated Circuit with a Scalable Communication Network and Interface. http://paulmerolla.com/merolla_main_som.pdf, 2014. [Online; accessed 04-December-2015].

[Par13] Parcly Taxel. Relations between the four fundamental electronic variables and devices that implement these relations. https://en.wikipedia.org/wiki/Memristor, 2013. [Online; accessed 05-December-2015].

Literature IV

[Pau10] Paul Reber. What Is the Memory Capacity of the Human Brain? http://www.scientificamerican.com/article/what-is-the-memory-capacity/, 2010. [Online; accessed 20-November-2015].

[Ste14] Steve Furber. Spikey school Introduction to the hardware system. http://electronicvisions.github.io/hbp-sp9-guidebook/pm/spikey/spikey_school.html, 2014. [Online; accessed 02-December-2015].

[T+15] Thomas Pfeil et al. The effect of heterogeneity on decorrelation mechanisms in spiking neural networks: a neuromorphic-hardware study. http://arxiv.org/pdf/1411.7916.pdf, 2015. [Online; accessed 02-December-2015].

[The13] The Economist Newspaper Limited, London. The machine of a new soul. http://www.economist.com/news/science-and-technology/21582495-computers-will-help-people-understand-brains-better-and-understanding-brains, 2013. [Online; accessed 28-November-2015].

[TOP15] TOP500.org. K COMPUTER, SPARC64 VIIIFX 2.0GHZ, TOFU INTERCONNECT. http://www.top500.org/system/177232, 2015. [Online; accessed 19-November-2015].

[Wik15] Wikipedia. Neuromorphic engineering — wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Neuromorphic_engineering&oldid=690486231, 2015. [Online; accessed 16-November-2015].

Memristor - The missing element

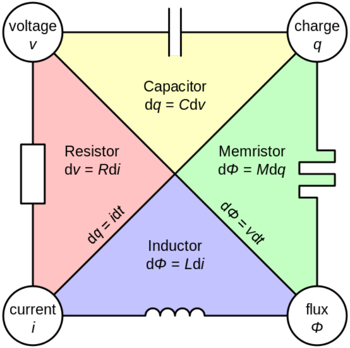


Figure: The four fundamental electronic variables and devices. [Par13]

Titanium Dioxide Memristor

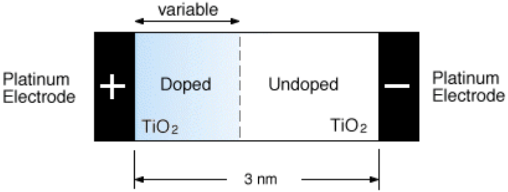


Figure: Titanium Dioxide Memristor [Jim10]