Execution Model: Neuromorphic Computing

Armin Schaare

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

2015-12-07
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 \( \sim 40 \) minutes
  - Consumed \( \sim 30.5 \) Gigajoule or \( \sim 8,500 \) kWh
  - Used \( \sim 1 \) 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' \( \sim \) 10 times bigger than brain

A Human Brain is \( \sim 240,000 \) 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 \( \sim 140 \text{ million} \) 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/elimates 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 mm²
- 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 mm² 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

![TrueNorth Chip Core Array](image.png)
The Memory Resistor

- 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)

Figure: Symbol of a memristor [Jos14]
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

- 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)

Figure: Symbol of a memristor [Jos14]
The Memory Resistor

- 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 manageable 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 manageable 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 neurodegenerative diseases
- Face, Speech, Object recognition
- Language interpretation
- Robotic terrain maneuvering
- 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

"As engineers, we would be foolish to ignore the lessons of a billion years of evolution"
— Carver Mead, 1993
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


Literature II


Literature III


Literature IV


[Ste14] Steve Furber. Spikey school Introduction to the hardware system.


[TOP15] TOP500.org. K COMPUTER, SPARC64 VIIIIFX 2.0GHZ, TOFU INTERCONNECT.

Memristor - The missing element

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

Figure: Titanium Dioxide Memristor [Jim10]