On-line Learning Architectures

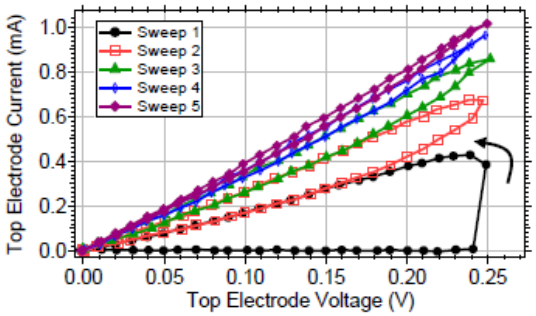
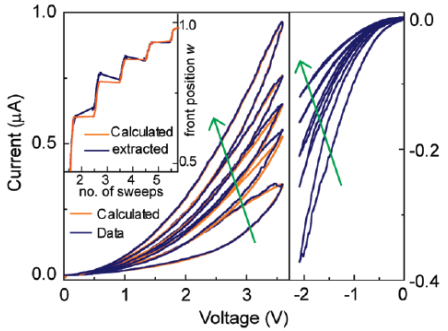
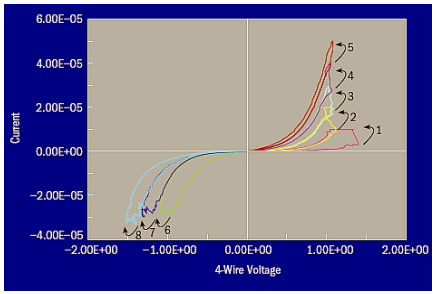
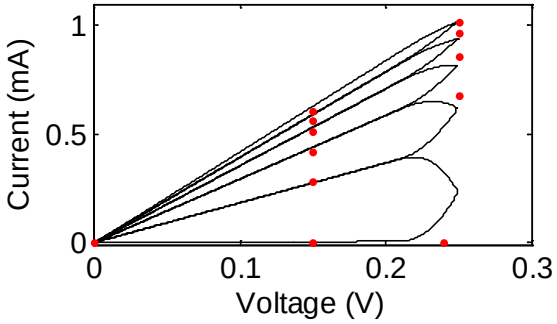
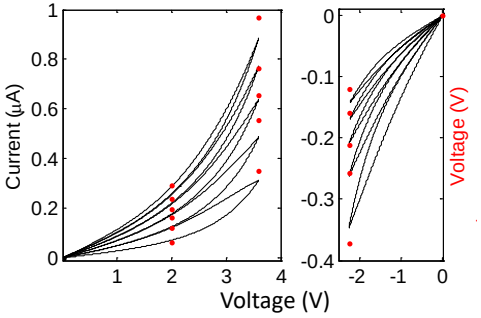
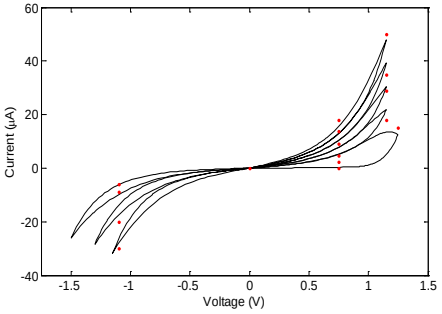
March 2017

Tarek M. Taha

Electrical and Computer Engineering Department

University of Dayton

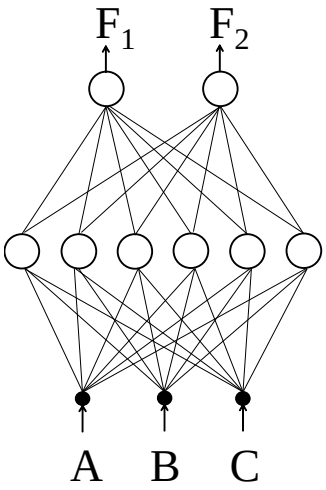
Device SPICE Model

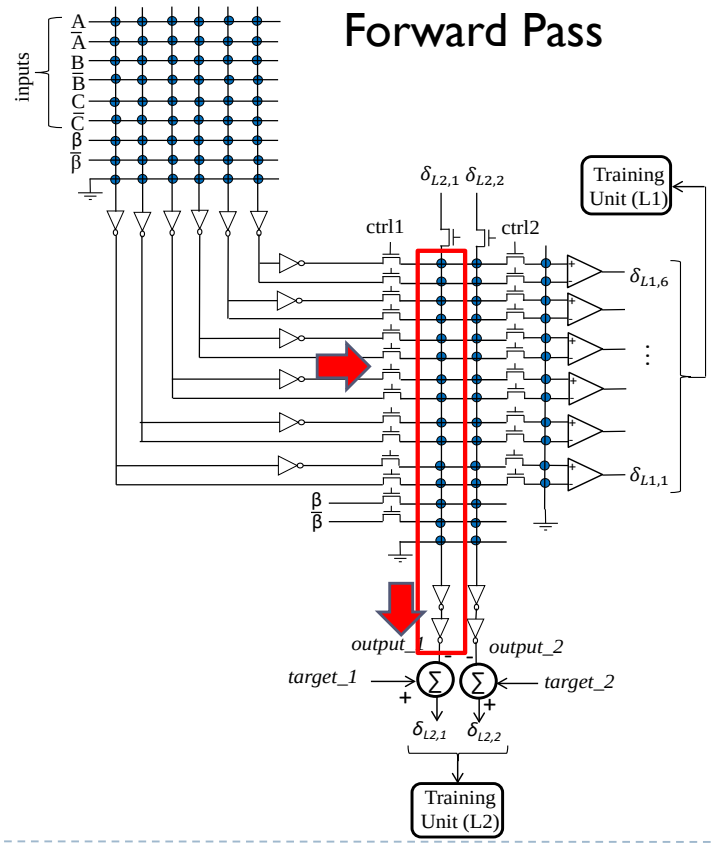
	Boise State	Univ of Michigan	HP Labs
Actual Device			
SPICE Model			

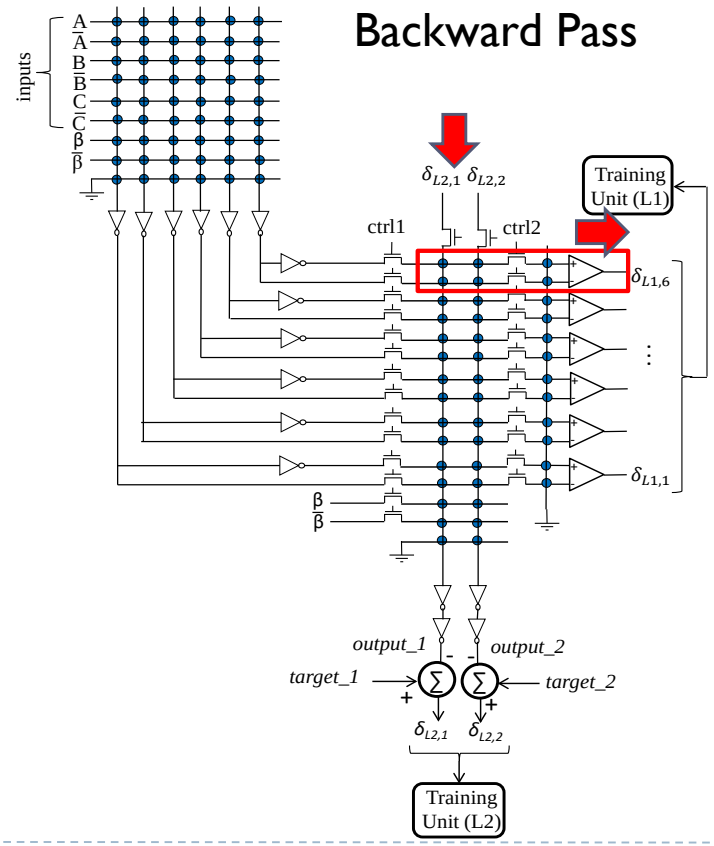
C. Yakopcic, T. M. Taha, G. Subramanyam, and R. E. Pino, "Memristor SPICE Model and Crossbar Simulation Based on Devices with Nanosecond Switching Time," IEEE International Joint Conference on Neural Networks (IJCNN), August 2013. [**BEST PAPER AWARD**]

Memristor Backpropagation Circuit

- Online training via backpropagation circuits
- The same memristor crossbar implements both forward and backward passes

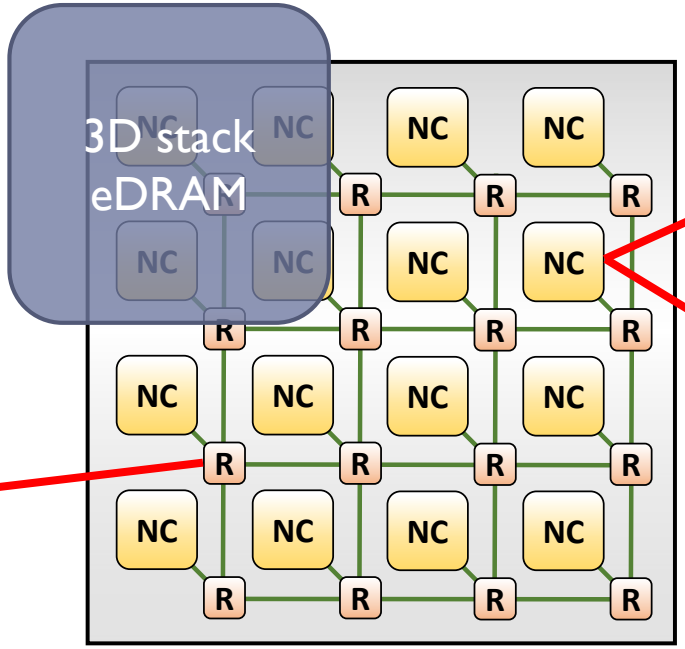




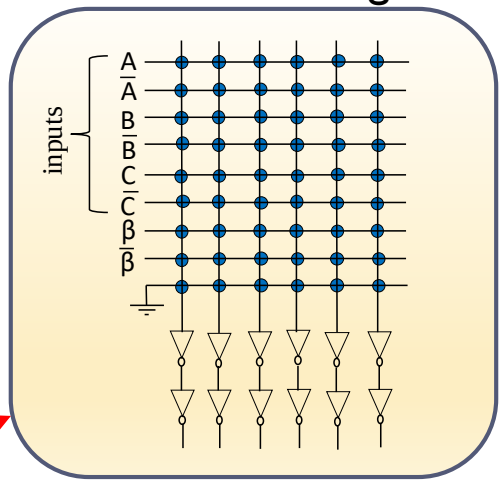


Socrates: Online Training Architecture

- Multicore processing system
- **Implements backpropagation based online training**
- Two versions: 1) memristor core and 2) fully digital core
- Has three 3D stack memory for storing training data

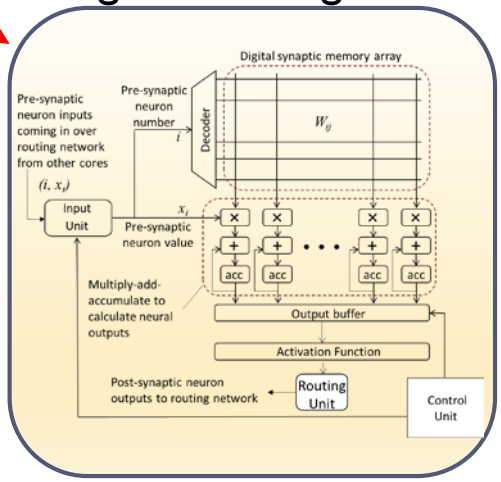


Memristor learning core

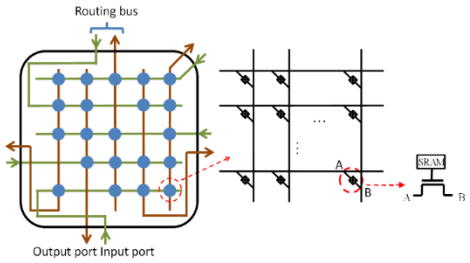


--OR--

Digital learning core



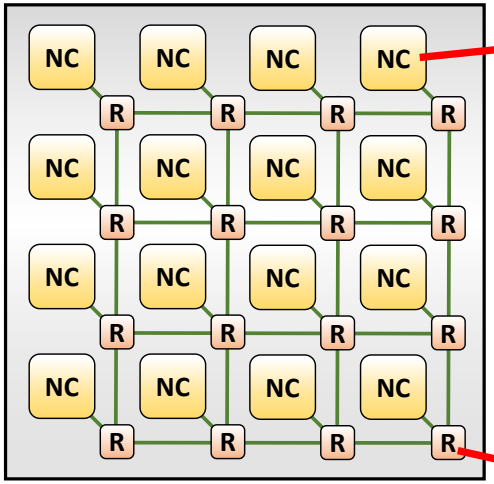
Router



FPGA Implementation

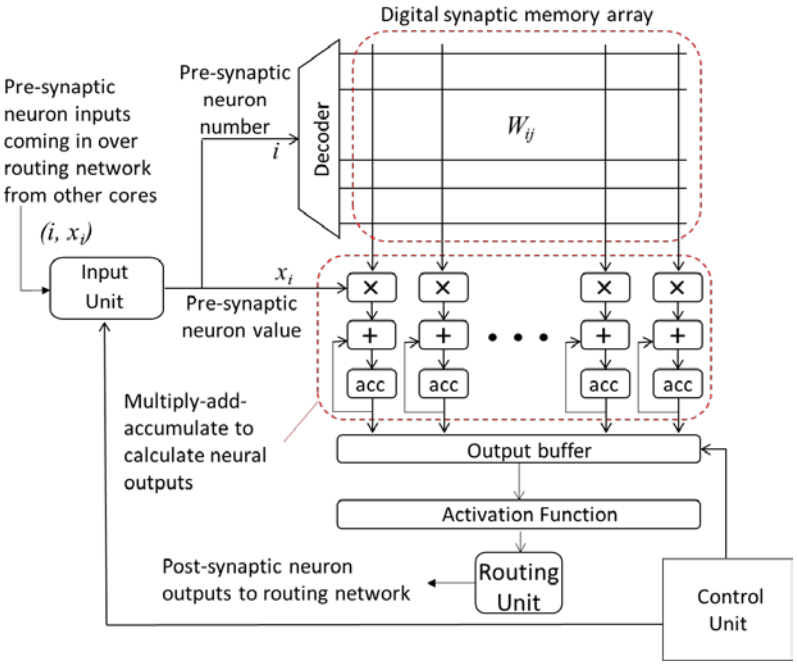
- We verified the system functionality by implementing the digital system on an FPGA.

NC = Neural Core
R = Router

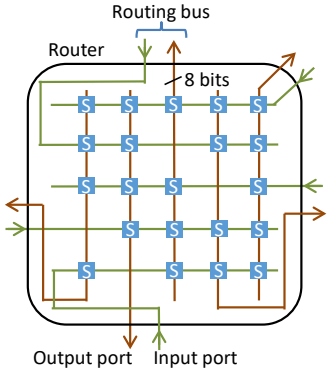


Actual images to and out of FPGA based multicore neural processor:

Neural Core



Router



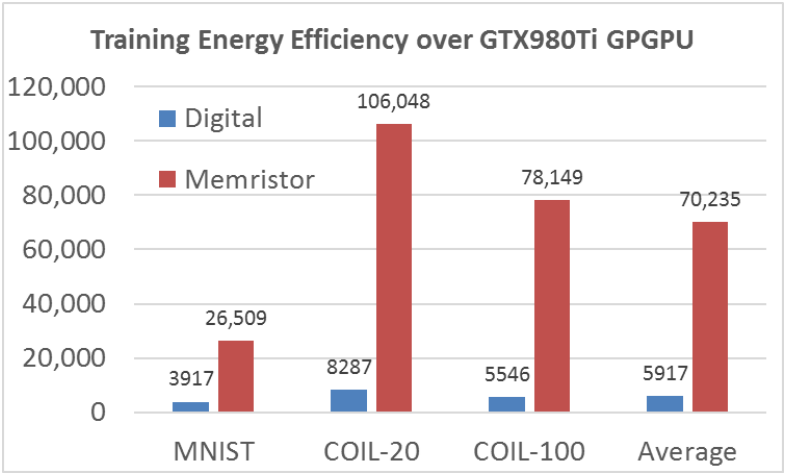
Training Efficiency

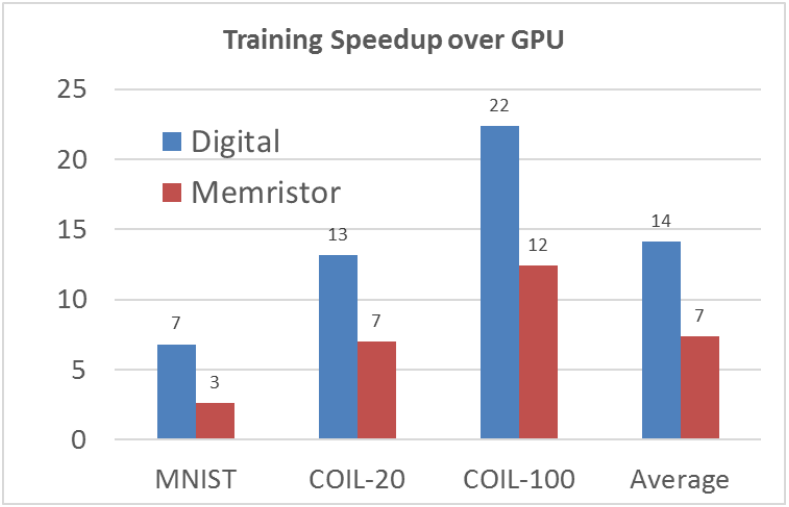
Training compared to GTX980Ti GPU

Energy Efficiency:
- Digital: 5900x
- Memristor: 70000x

Speedup:
- Digital: 14x
- Memristor: 7x
- Batch processing on GPU but not on specialized processors

Accuracy(%)	GPU	Digital	Memristor
MNIST	99%	97%	97%
COIL-20	100%	98%	97%
COIL-100	99%	99%	99%





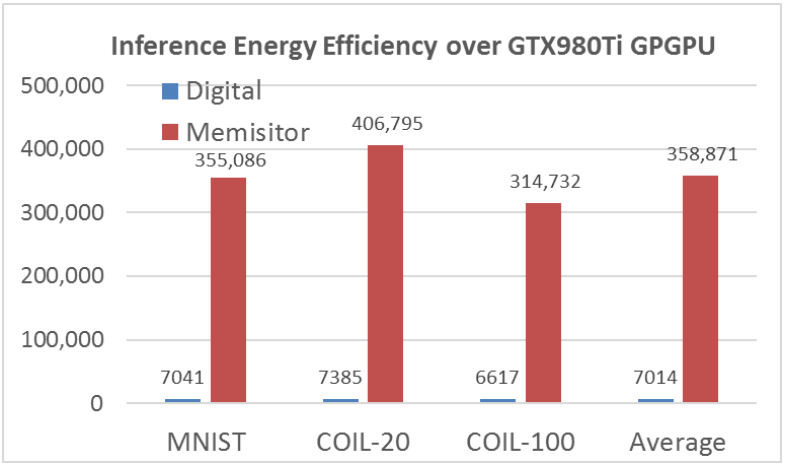
Inference Efficiency

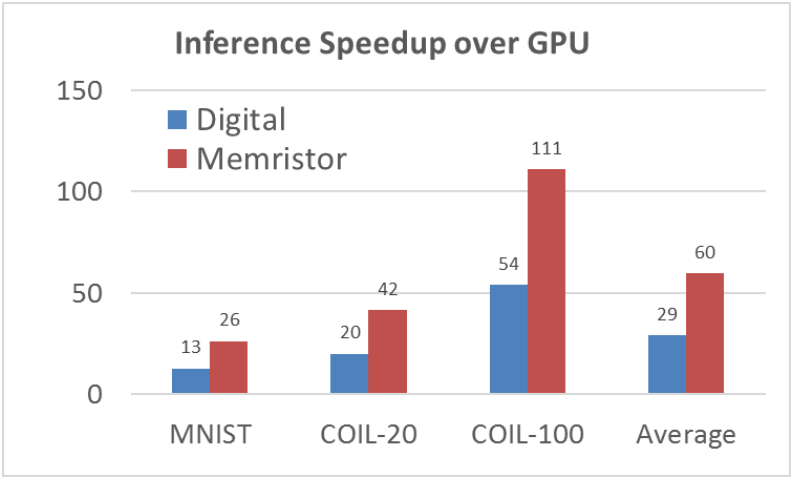
Inference compared to GTX980Ti GPU

Energy Efficiency:
- Digital: 7000x
- Memristor: 358000x

Speedup:
- Digital: 29x
- Memristor: 60x

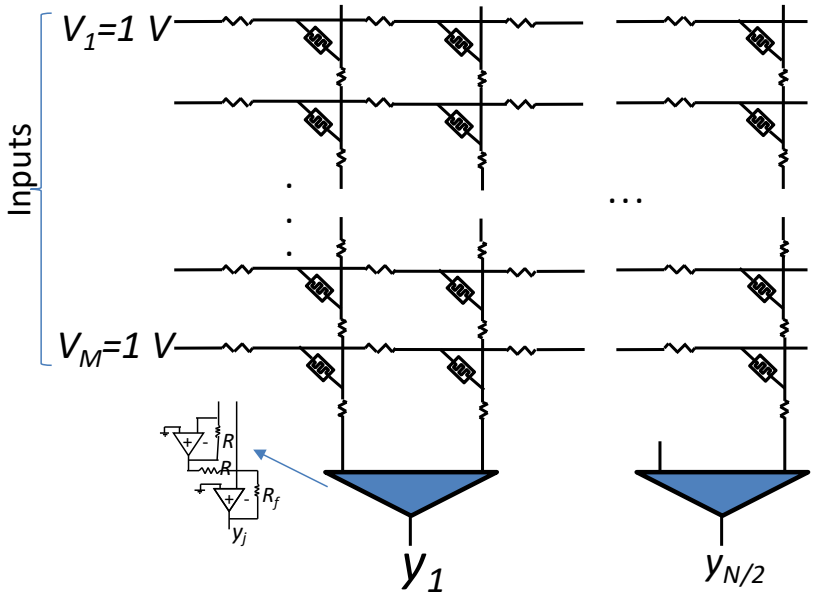
Memristor learning core is about 35x smaller in area:
- Digital: 0.52 mm^2
- Memristor: 0.014 mm^2

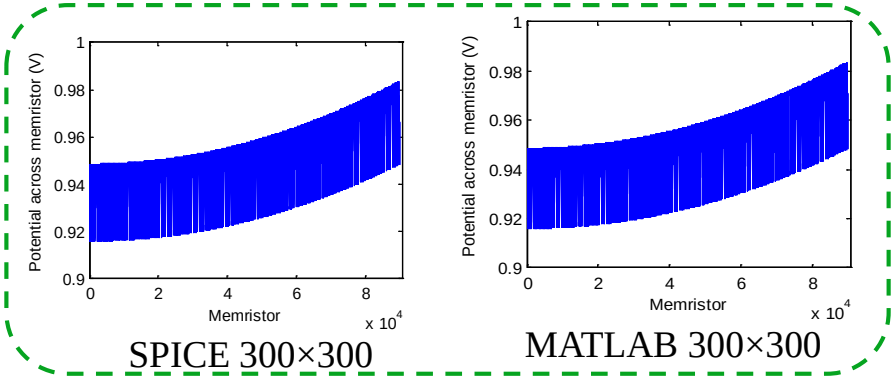




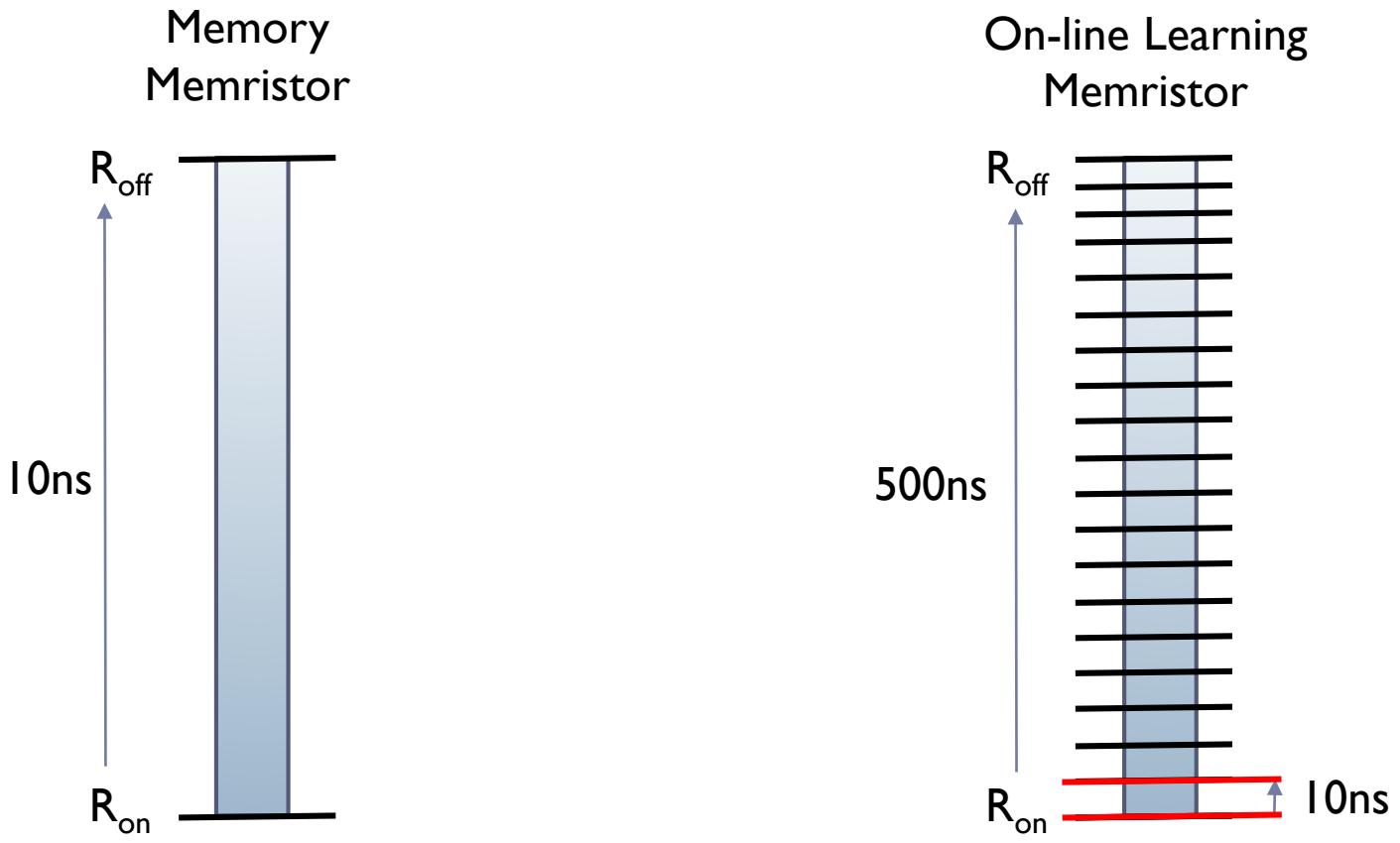
Large Crossbar Simulation

- Studying training using memristor crossbars needs to be done in SPICE to account for circuit parasitics.
- Simulating one cycle of a large crossbar in can take over 24 hours in SPICE. (training requires thousands of iterations and hence is nearly impossible in SPICE).
- Developed a fast (~30ms) algorithm to approximate SPICE output with over 99% accuracy.
- Enables us to examine learning on memristor crossbars.





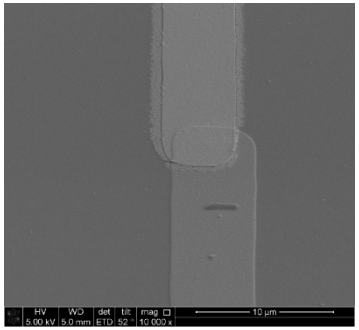
Memristor for On-Line Learning

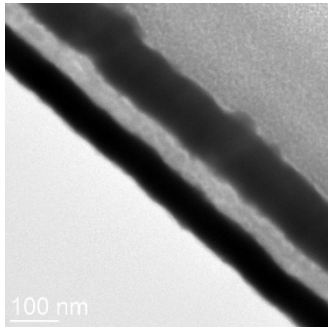


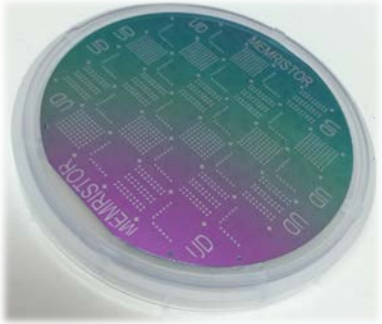
- On-line learning requires "highly tunable" or "slow" memristors
- This does not affect inference speed

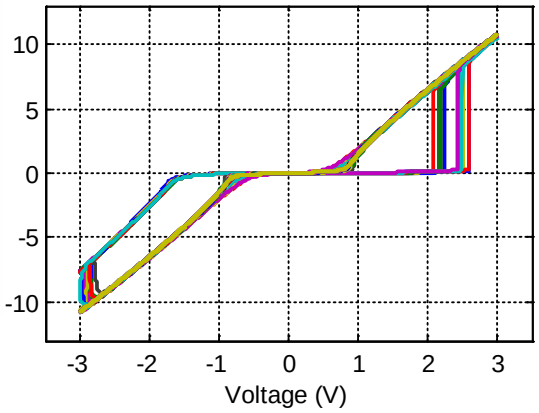
$LiNbO_3$ Memristor for Online Learning

- Device is quite slow: order of tens of millisecond switching time.
- Useful for backprogation training.



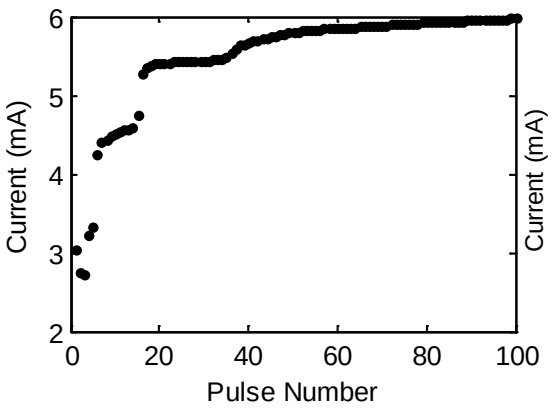


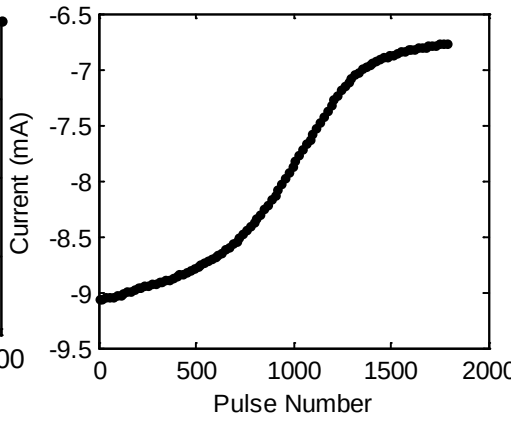




Switching speed of devices and minimum programming pulse needed in neuromorphic training circuits.

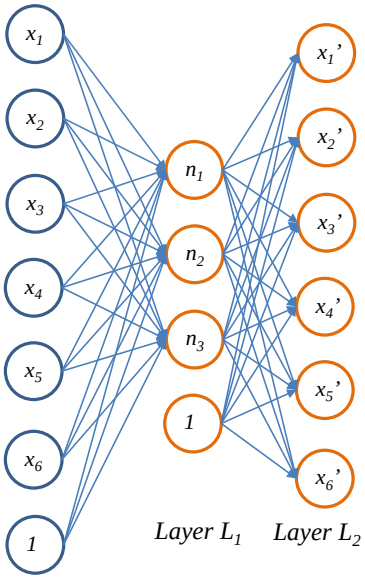
Approx. Switching Time	Switching Voltage	Minimum Pulse Width Required
10ns	6	1ps
10ns	4.5	10ps
100ns	4.5	100ps
1µs	4.5	1ns
10µs	4.5	10ns

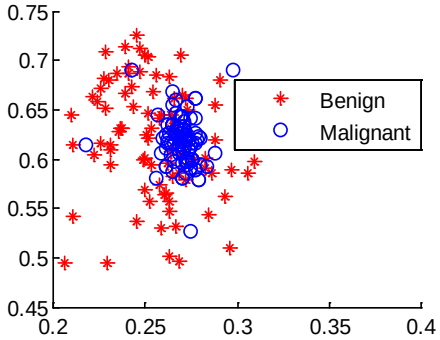
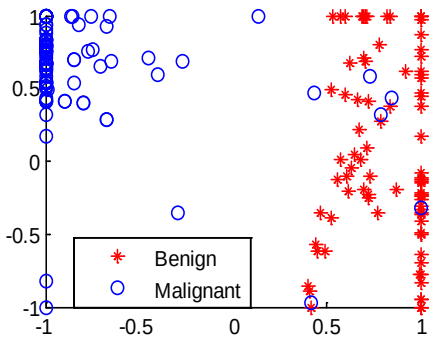
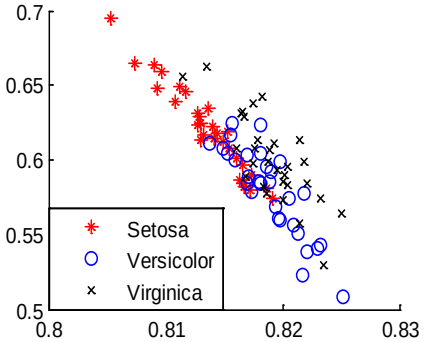
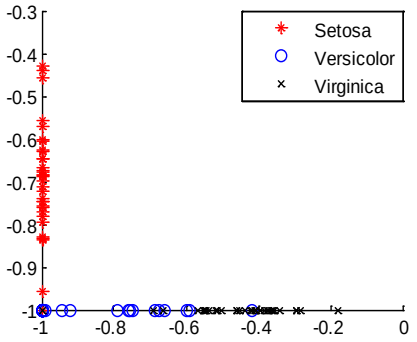
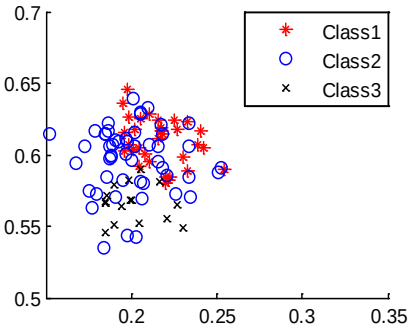
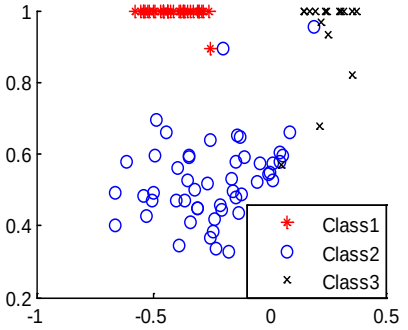




Unsupervised Clustering

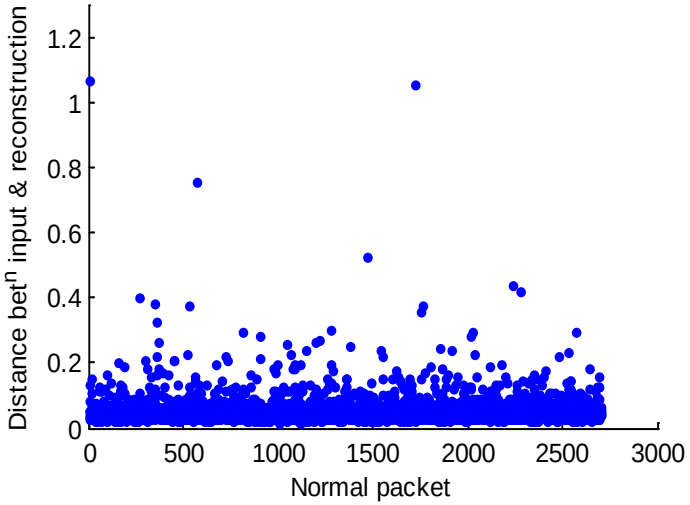
- We extended the backpropagation circuit to implement auto-encoder based unsupervised learning in memristor crossbars.
- These graphs show the circuit learning and clustering data in an unsupervised manner.

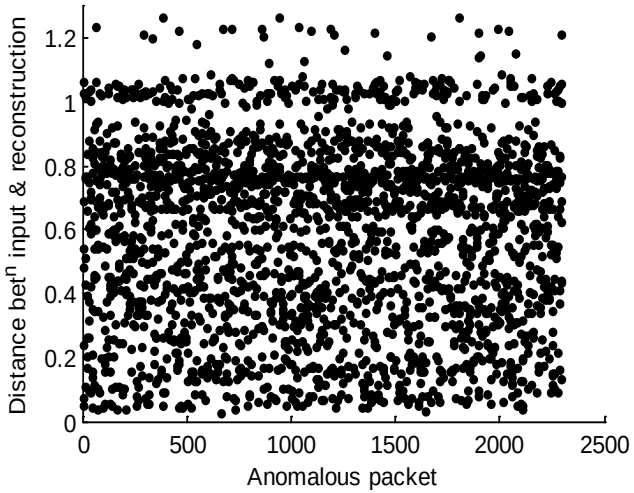


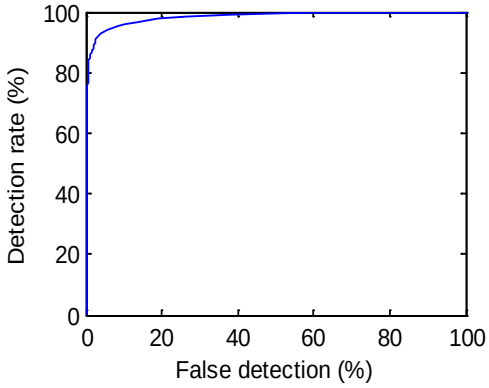
	Before Training	After Training
Wisconsin		
Iris		
Wine		

Application: Cybersecurity

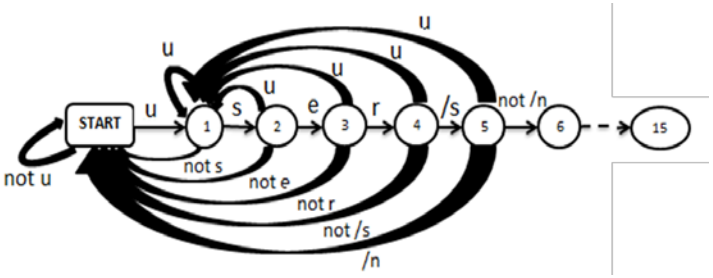
- The autoencoder based unsupervised training circuits were used to learn normal network packets.
- When the memristor circuits were presented with anomalous network packets, these were detected with a 96.6% accuracy and a 4% false positive rate.







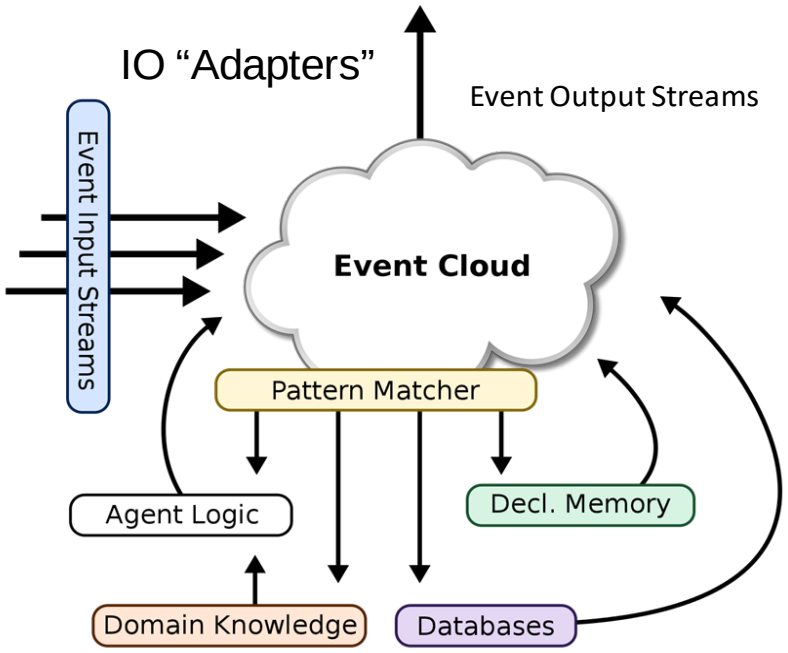
- Could also implement rule based malicious packet detection (similar to SNORT) in a very compact circuit



Regex : *user\s[^\n]{10}*

Application: Autonomous Agent for UAVs

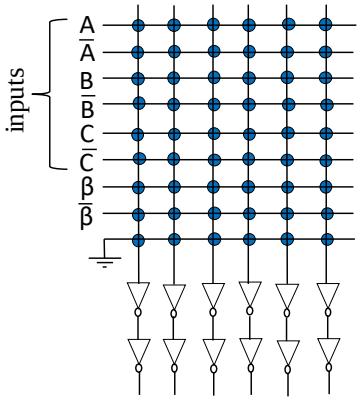
- Cognitively Enhanced Complex Event Processing (CECEP) Architecture
- Geared towards autonomous decision making in a variety of applications including **autonomous UAVs**



- Exploring implementation on:
  - IBM TrueNorth processor
  - Memristor circuits
  - High performance clusters: CPUs and GPUs

IBM TrueNorth

Memristor Systems



Parallel Cognitive Systems Laboratory

Research Engineers:

- Raqib Hasan, PhD
- Chris Yakopcic, PhD
- Wei Song, PhD
- Tanvir Atahary, PhD

Doctoral Students:

- Zahangir Alom
- Rasitha Fernando
- Ted Josue
- Will Mitchell
- Yangjie Qi
- Nayim Rahman
- Stefan Westberg
- Ayesha Zaman
- Shuo Zhang

Master's Students:

- Omar Faruq
- Tom Mealy

Sponsors:

LPS The Laboratory for Physical Sciences

DARPA

Air Force Research Laboratory

NSF

Air Force Office of Scientific Research, United States Air Force

http://taha-lab.org