Brain-like Computing – Scalable Low-Power Chips for Learning & Optimal Control

Prof. Pinaki Mazumder
University of Michigan
Ann Arbor, MI 48105

Acknowledgement: National Science Foundation (CCF & ECCS)

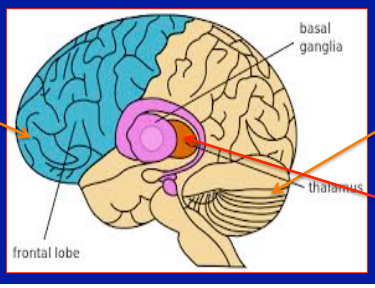
1

Brain-like Computing

Facets of Brain-like Computing:

1. Self-Healing
2. Associative Memory
3. Cognition
4. Learning & Plasticity

Transcribing Basal Ganglia on Silicon for Optimal Control Algorithms



Unsupervised Learning by Cerebral Cortex (Statistical Input)

Supervised Learning by Cerebellum (Error Vector)

Reinforcement Learning by Basal Ganglia (Reward Scalar)

Adaptive Hardware Platform for Nonlinear Optimal Control, Swarm Intelligence, Robot Control, and Markov Decision Process (MDP)

Mazumder Group's Neuromorphic Research

Self-Healing VLSI Design (1989-1996)

Hopfield Neural Net as Algorithmic Hardware for Spare Allocation by Node Cover over Bipartite Graph

- IEEE Trans. on CAS, 1993
- IEEE Trans. on CAS, 1993
- IEEE Trans. on Computer, 1996

Cognition Chips using Cellular Neural Networks (2008-2013)

- Color Image Processing
- Velocity Tuned Filter
- Memristor/RRAM based CNN
- RTD+HEMT based CNN

- IEEE Trans. on VLSI, 2009
- IEEE Trans. on Nanotechnology, 2008
- IEEE Trans. on Neural Nets, 2014
- IEEE Trans. on Nanotechnology, 2013
- ACM Journal on Emerging Technologies in Computing Systems, 2013

Learning based VLSI Chips (2010- Now)

STDP Learning for Position Detector
STDP Learning for Virtual Bug Navigation
STDP Learning for XOR/Edge Detection
Deep Learning for Pattern recognition

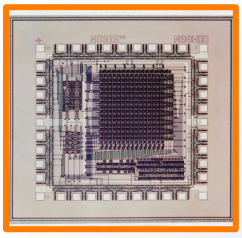
Q-Learning for Maze Search Algorithm on Memristor Array

- Proceedings of the IEEE, 2012
- Nano Letters Journal, 2010
- IEEE Nanotechnology, 2011
- IEEE Nanotechnology, 2014
- IEEE Cellular Neural Networks, 2012

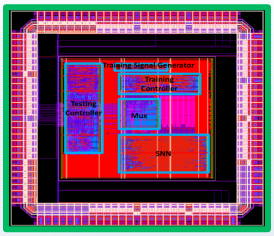
Reinforcement Learning/Actor-Critic NN (2016 – Now)

- IEEE Trans. on Computer, 2016
- IEEE Trans. on Neural Nets, 2018
- IEEE Trans. on Circuits & Systems, 2018

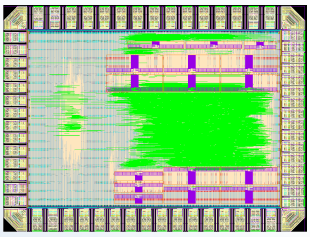
Neural-Inspired CMOS Chips Designed by Mazumder Group



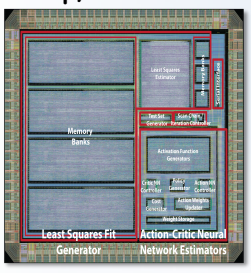
Self-Healing Chip, 1991



Synaptic Plasticity Chip, 2013



Deep Learning Chip, 2016



Actor-Critic Reinforcement Learning Chip, 2016

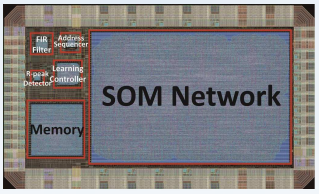
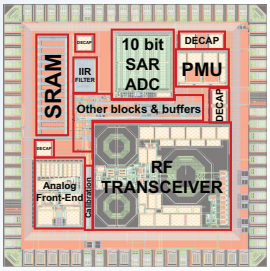


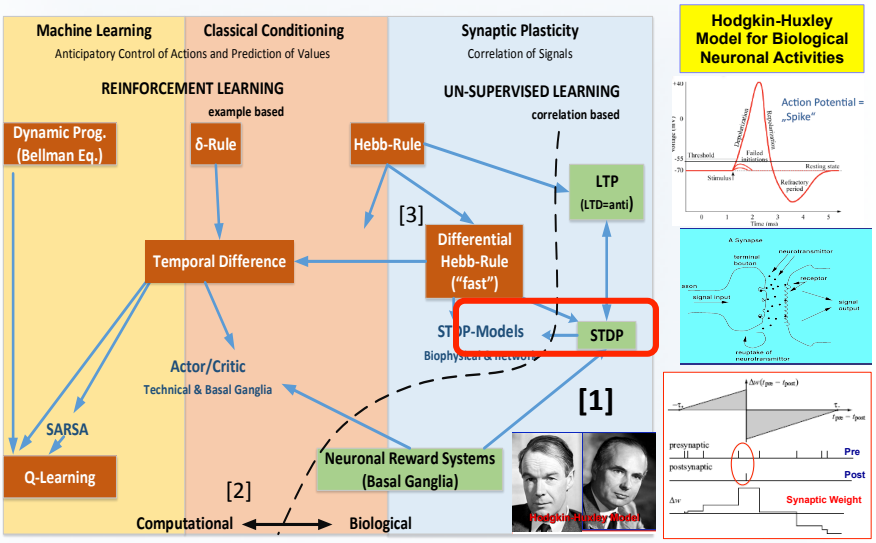
Figure 4.4: The proposed ECG clustering SOM chip layout.

Self-Organizing Map Chip for ECG Clustering 2016



Body-Sensing Network with Wireless Transceiver

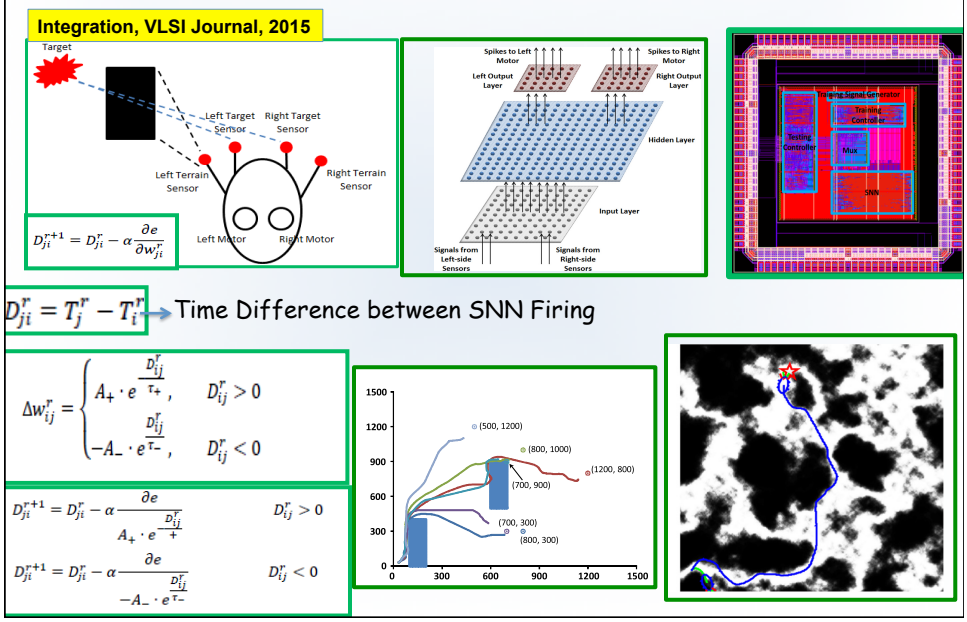
Gamut of Learning Models



Based on Woergoetter & Porr's Taxonomy

UNIVERSITY OF MICHIGAN

Plasticity (STDP) Based Learning Chip for Virtual Bug Navigation



$$D_{ji}^{r} = T_j^r - T_i^r$$

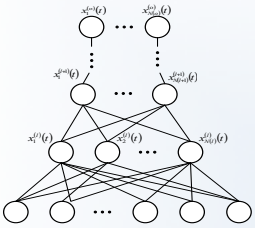
$$\Delta w_{ij}^{r} = \begin{cases} A_+ \cdot e^{\frac{D_{ij}^r}{\tau_+}}, & D_{ij}^r > 0 \\ -A_- \cdot e^{\frac{D_{ij}^r}{\tau_-}}, & D_{ij}^r < 0 \end{cases}$$

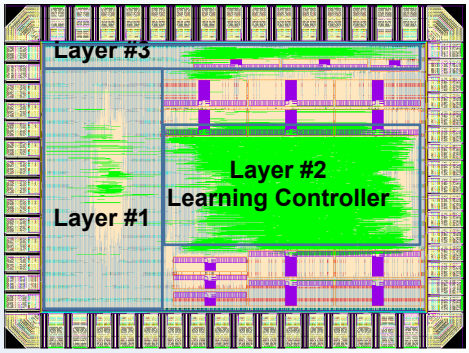
$$D_{ji}^{r+1} = D_{ji}^{r} - \alpha \frac{\partial e}{\partial w_{ji}^r}$$

$$D_{ji}^{r+1} = D_{ji}^{r} - \alpha \frac{\partial e}{A_+ \cdot e^{\frac{D_{ij}^r}{\tau_+}}} \qquad D_{ij}^r > 0$$

$$D_{ji}^{r+1} = D_{ji}^{r} - \alpha \frac{\partial e}{-A_- \cdot e^{\frac{D_{ij}^r}{\tau_-}}} \qquad D_{ij}^r < 0$$

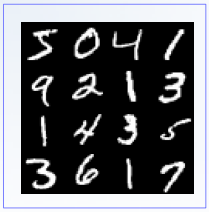
Discretized STDP Based Deep Neural Net



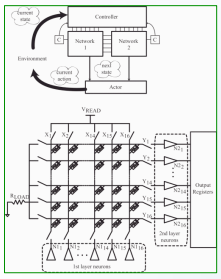


Performance Summary

- –Network size: 256-50-10
- –Technology: TSMC 65 nm
- –Bit width of synapse: 24 bits
- –Clock frequency: ~100 MHz
- –Power: ~100 mW
- –Time needed for one learning iteration: 30 µs
- –Speedup : ~700x v. digital computer
- –Energy: ~500 pJ/spike @ VDD = 1.2

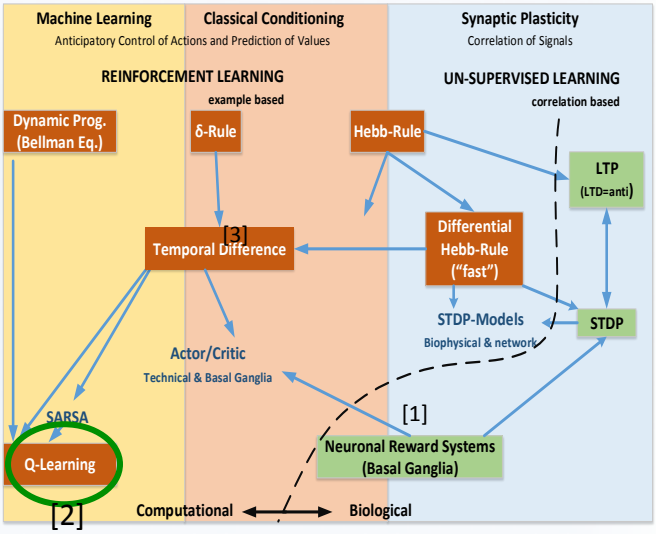


Gamut of Learning Models



Google's Deep Mind or Deep Q-Network (DQN) Played Atari 2600 Better than a Human

Google's Go Program uses SARSA and Monte Carlo Tree Search to beat World Champion (Kei Ji) in a Go contest.



Based on Woergoetter & Porr's Taxonomy

Q-Learning Hardware – Reinforcement Learning

SARSA: State-Action-Reward-State-Action

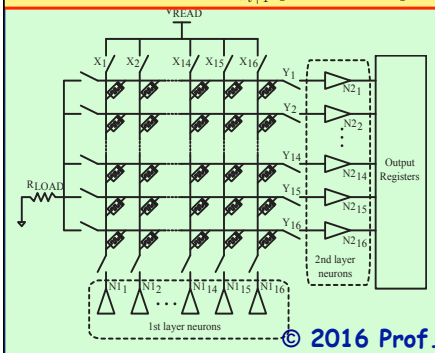
*NewEstimate ← OldEstimate + StepSize * (Target - OldEstimate)*

$$Q(s_t,a_t) \leftarrow \underbrace{Q(s_t,a_t)}_{\text{old value}} + \underbrace{\alpha_t}_{\text{learning rate}} \cdot \left(\underbrace{r_{t+1}}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \cdot \underbrace{\max_a Q(s_{t+1},a)}_{\text{learned value: estimate of optimal future value}} - \underbrace{Q(s_t,a_t)}_{\text{old value}} \right)$$

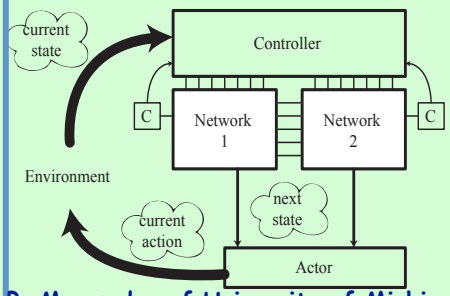
$$\tilde{Q}(s_t,a_t) \leftarrow \tilde{Q}(s_t,a_t) \times (1-\alpha_t(s_t,a_t)) + \alpha_t(s_t,a_t) \times r_t + \alpha_t(s_t,a_t) \times \max_{a_{t+1}} \left[\tilde{Q}(s_{t+1},a_{t+1})\right]$$

$$\max_{a_{t+1}} \left[\tilde{Q}(s_{t+1},a_{t+1})\right] = \tilde{Q}(s_t,a_t) + \delta_t.$$

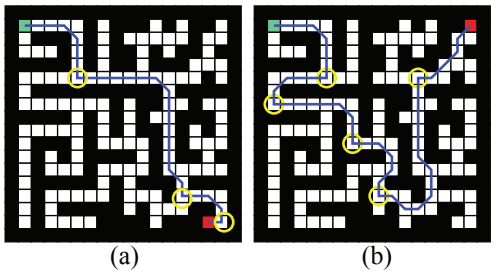
$$\tilde{Q}(s_t,a_t) \leftarrow \tilde{Q}(s_t,a_t) + \alpha_t(s_t,a_t) \times (r_t + \delta_t)$$

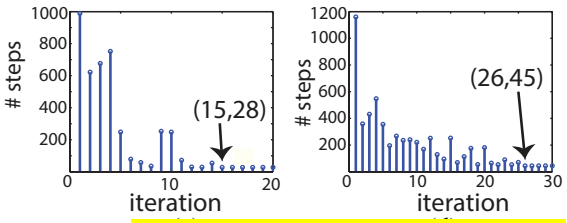


Evaluative Feedback (Rewards)

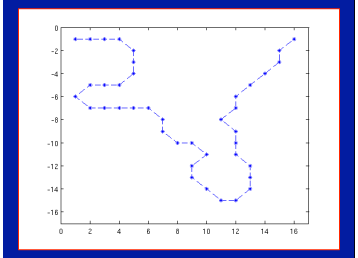


Performance of Memristor Q-Learning Hardware



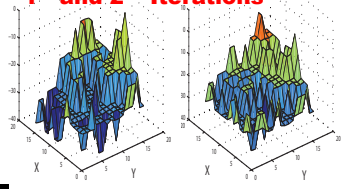


Ebong & Mazumder, IEEE Nano 2014



Memristor Model Used in Matlab Simulation

Synapse States after 1st and 2nd Iterations

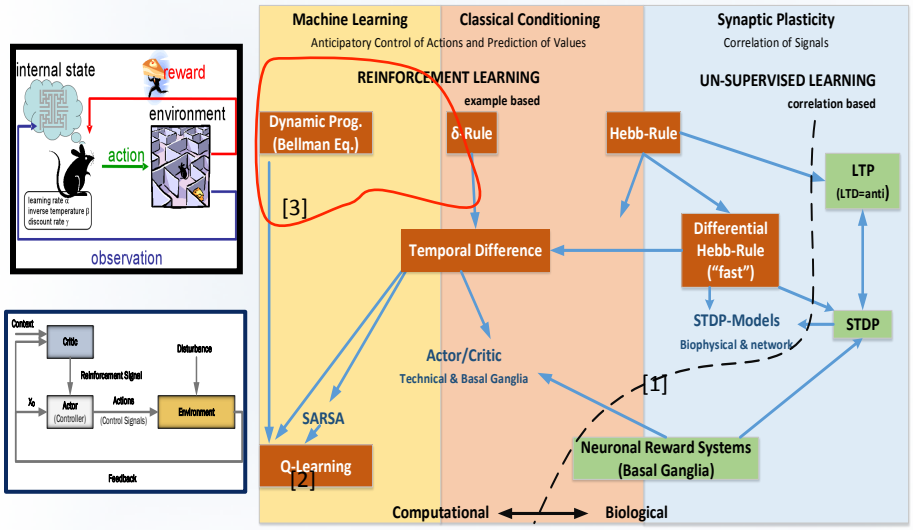


Superior Application of Q Learning by Google, UK

Human-level control through deep reinforcement learning
Nature, v. 518, pp. 529-533, Feb 2015.

Played Atari 2600
Better than a Human

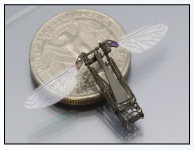
Gamut of Learning Models



Based on Woergoetter & Porr's Taxonomy

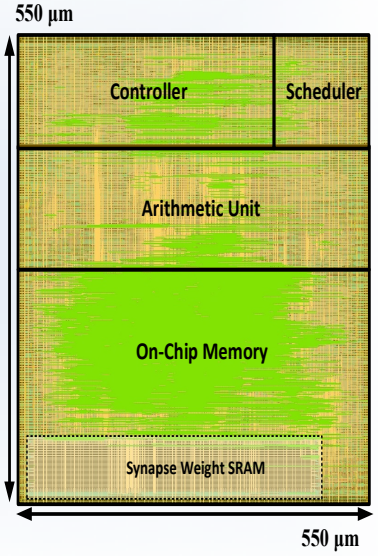
Actor-Critic NN CMOS Chip

- **Post-Layout Metrics**
  - Technology: 65 nm
  - Area: 550 μm × 550 μm
  - Arithmetic precision: 24-bit fixed-point
  - Supply voltage: 1.2 V
  - Clock frequency: 175 MHz
  - Power Consumption: 25 mW
  - 270x speed up over software ADP



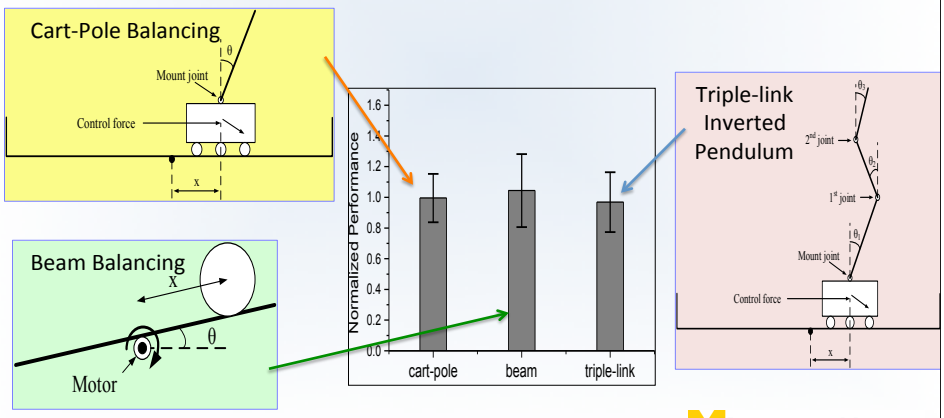
Controlling balance, rotation, movement of microrobots that operate on batteries

Provide smart sensing strategies for energy-constrained sensors



Benchmark Results

- **Benchmark Tests**
  - Cart-Pole Balancing
  - Beam Balancing
  - Triple-link Inverted Pendulum
  - Results obtained from the accelerator are normalized w.r.t software results
  - Hardware solution consumes 25 mW and is faster than software by 225 times



UNIVERSITY OF MICHIGAN

Problem Statement

- **Optimal Control/Decision-Making Problem**
  - Looking for the optimal policy/strategy
  - Target is to maximize reward &minimize cost
  - Solve using Bellman equation: The optimal equality.
  - Continuous-space problem is handled by Hamilton-Jacobi-Bellman equation

System Model: $\mathbf{x}(t+1) = f[\mathbf{x}(t), \mathbf{a}(t)]$

Reward to be maximized: $J[\mathbf{x}(t)] = \sum_{k=1}^{\infty} \gamma^{k-1} r[\mathbf{x}(t+k)]$

Bellman equation: $J^*[\mathbf{x}(t)] = \max_{\mathbf{a}(t)} (r[\mathbf{x}(t+1)] + \gamma J^*[\mathbf{x}(t+1)])$

$\mathbf{x}(t)$-state of the system at time t

$\mathbf{a}(t)$-action performed at time t

$J[\mathbf{x}(t)]$-state value function at state time$\mathbf{x}(t)$

$J^*[\mathbf{x}(t)]$-optimal state value function

$r[\mathbf{x}(t+1)]$-reward received at time t

γ-discount factor used to encourage near-term reward

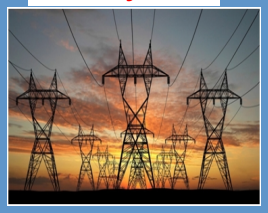
UNIVERSITY OF MICHIGAN

Traditional Applications of ADP in Software

Autopilot

- Used to control flight of a plane
- Cope with unexpected conditions

Power Systems



- Used as static compensators to regulate voltage in power systems

Communication Systems

- Self-learning call admission control scheme that can adapt for new communication environment

Autonomous Robot

- Navigation in unknown environmen
- Control of a moving robotic arm

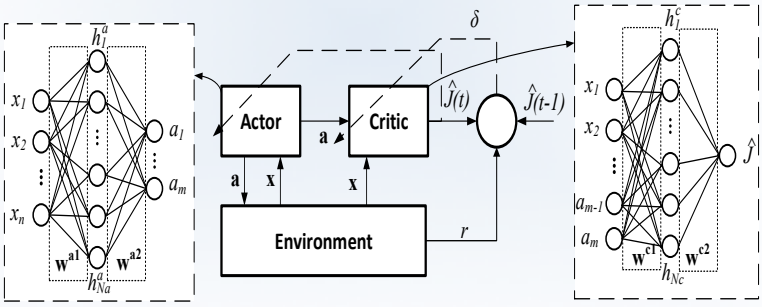
UNIVERSITY OF MICHIGAN

Action-Dependent Heuristic Dynamic Programming

- **Essentials**
  - Model-free learning
  - Learning by minimizing temporal difference error
  - Critic: estimate the reward-to-go
  - Actor: pick the right action to maximize future reward

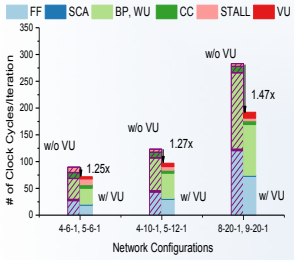
$$\delta(t) = \hat{J}[\mathbf{x}(t-1)] - \gamma\hat{J}[\mathbf{x}(t)] - r[\mathbf{x}(t)]$$

$$\hat{\mathbf{a}}[\mathbf{x}(t)] = \max_{\hat{\mathbf{a}}[\mathbf{x}(t)]}\left(r[\mathbf{x}(t+1)] + \gamma\hat{J}[\mathbf{x}(t+1)]\right)$$



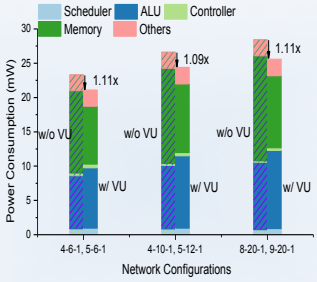
UNIVERSITY OF MICHIGAN

New Type of Updating Scheme



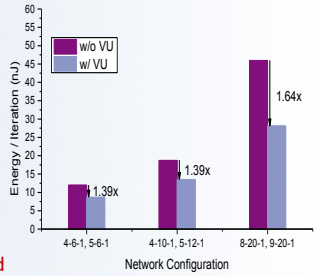
Throughput is increased

Fewer operations are needed when the virtual update algorithm is employed



Power consumption is reduced

Costly memory operations are greatly reduced



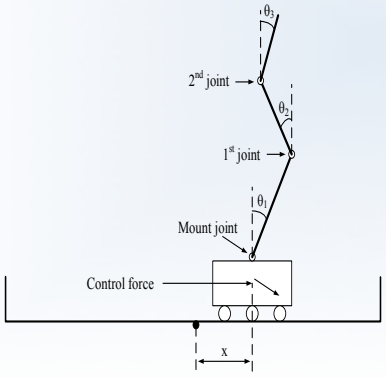
Energy efficiency is improved

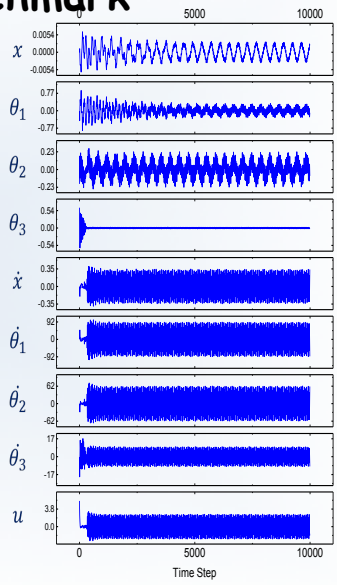
Consequence of improved throughput and power consumption

UNIVERSITY OF MICHIGAN

Triple-Link Inverted Pendulum Benchmark

- **Control Task**
  - Stabilize the system
  - 8 state variables as input: $x, \theta_1, \theta_2, \theta_3, \dot{x}, \dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3$ Control $\theta_1 - \theta_3$ to be within $[-20^\circ, 20^\circ]$
  - Control x to be within [-1m, 1m]
  - Magnitude of the control voltage has to be less than 30V

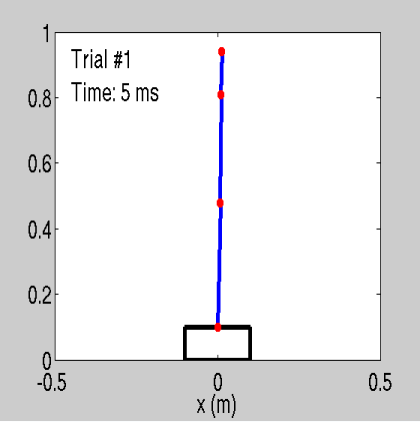




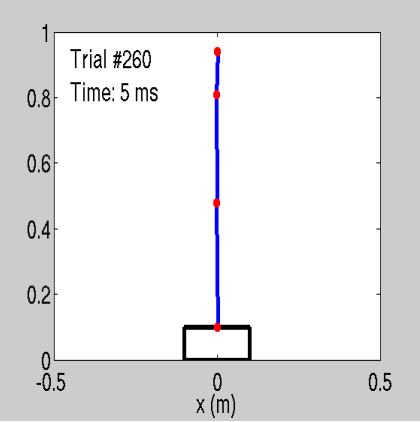
UNIVERSITY OF MICHIGAN

Triple-Link Inverted Pendulum Benchmark

Before Training



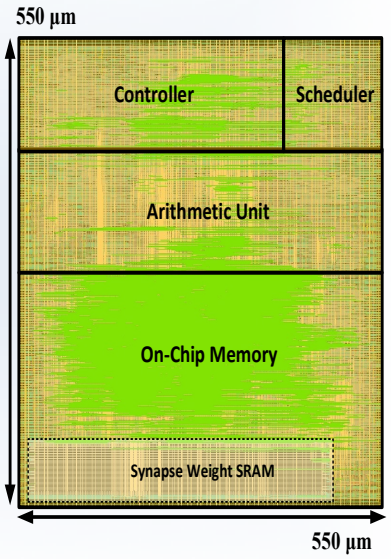
After Training



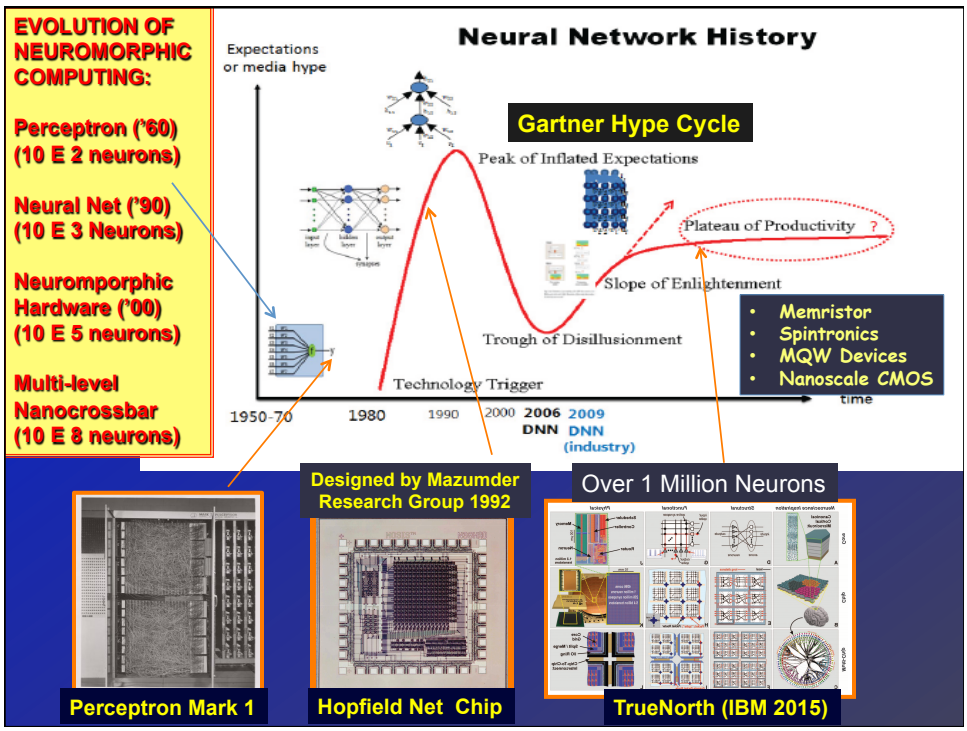
UNIVERSITY OF MICHIGAN

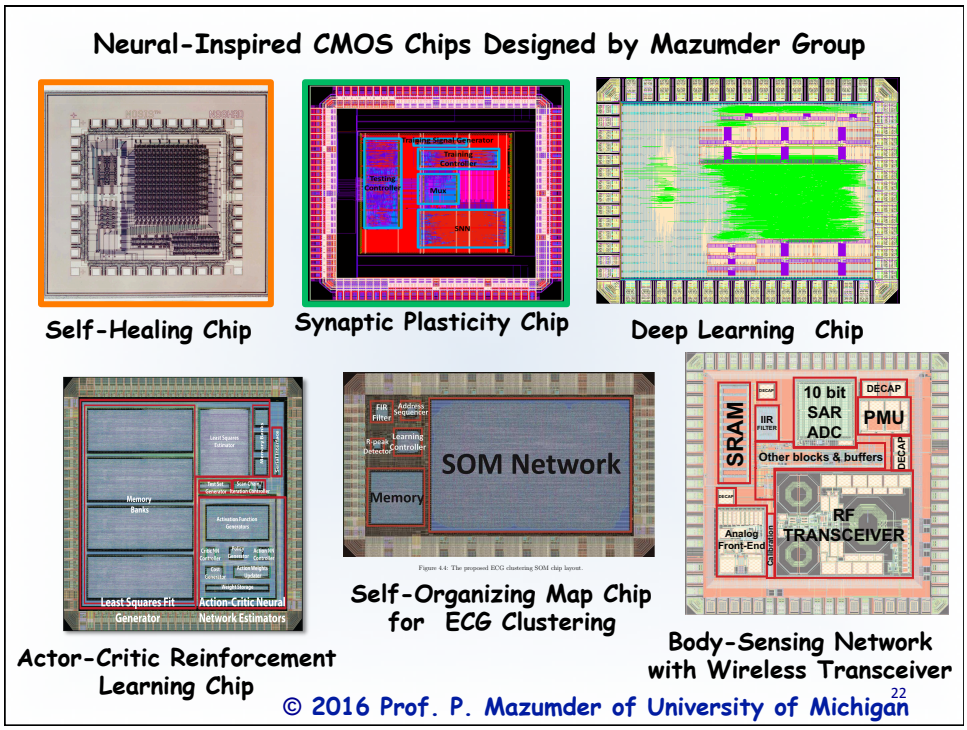
CMOS Implementation

- **Post-Layout Metrics**
  - Technology: 65 nm
  - Area: 550 μm × 550 μm
  - Number of datalanes: 4
  - Arithmetic precision: 24-bit fixed-point
  - Supply voltage: 1.2 V
  - Clock frequency: 175 MHz
  - Power Consumption: 25 mW
  - 270x speed up compared to the software approach



UNIVERSITY OF MICHIGAN


EVOLUTION OF NEUROMORPHIC COMPUTING:
Perceptron ('60) (10 E 2 neurons)
Neural Net ('90) (10 E 3 Neurons)
Neuromorphic Hardware ('00) (10 E 5 neurons)
Multi-level Nanocrossbar (10 E 8 neurons)
Neural Network History
Expectations or media hype
Gartner Hype Cycle
Peak of Inflated Expectations
Plateau of Productivity ?
Slope of Enlightenment
Trough of Disillusionment
Technology Trigger
time
1950-70
1980
1990
2000
2006 DNN
2009 DNN (industry)
Memristor
Spintronics
MQW Devices
Nanoscale CMOS
Designed by Mazumder Research Group 1992
Over 1 Million Neurons
Perceptron Mark 1
Hopfield Net Chip
TrueNorth (IBM 2015)


Neural-Inspired CMOS Chips Designed by Mazumder Group
Self-Healing Chip
Synaptic Plasticity Chip
Deep Learning Chip
Actor-Critic Reinforcement Learning Chip
Self-Organizing Map Chip for ECG Clustering
Body-Sensing Network with Wireless Transceiver
SOM Network
Memory
SRAM
10 bit SAR ADC
PMU
Other blocks & buffers
RF TRANSCEIVER
Analog Front-End
Least Squares Fit Generator
Action-Critic Neural Network Estimators
© 2016 Prof. P. Mazumder of University of Michigan
22