Brain-like Computing – Technologies and Architectures

Prof. Pinaki Mazumder

University of Michigan

Ann Arbor, MI 48105

Acknowledgement: National Science Foundation (CCF & ECCS)

EVOLUTION OF NEUROMORPHIC COMPUTING:

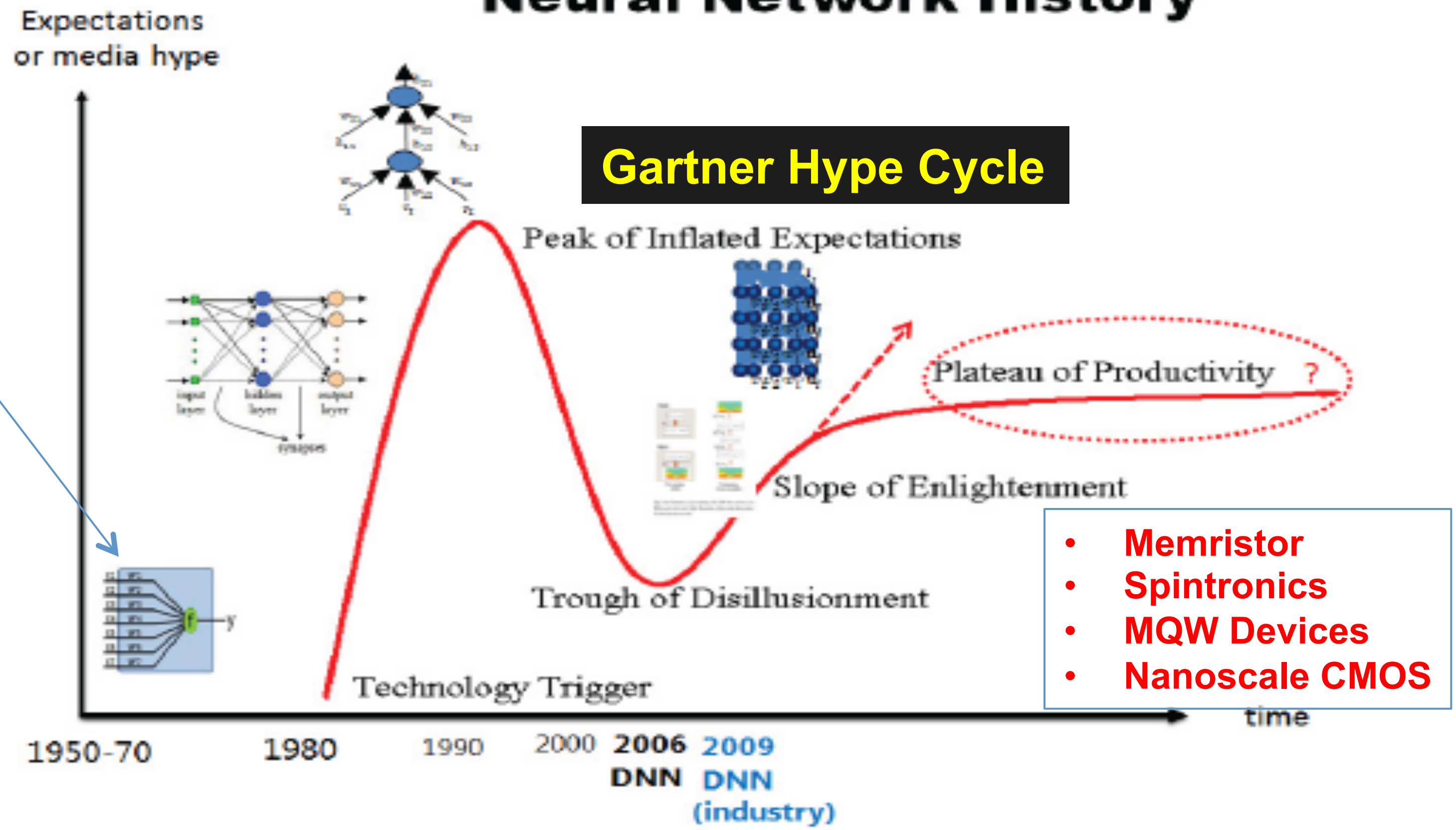
Perceptron ('60)
(10 E 2 neurons)

Neural Net ('90)
(10 E 3 Neurons)

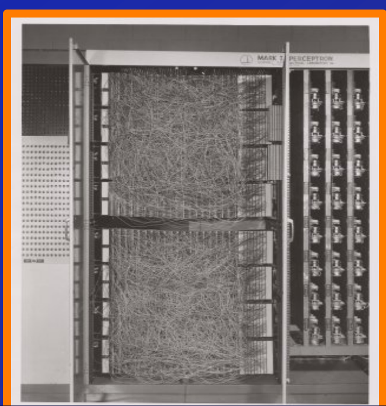
Neuromporphic Hardware ('00)
(10 E 5 neurons)

Multi-level Nanocrossbar
(10 E 7 neurons)

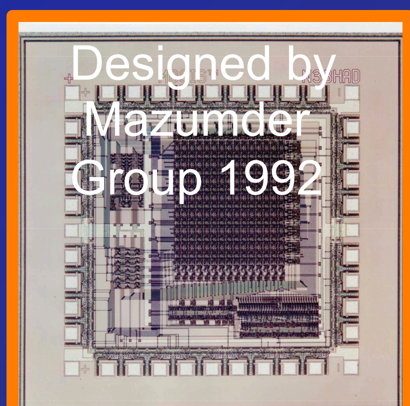
Neural Network History



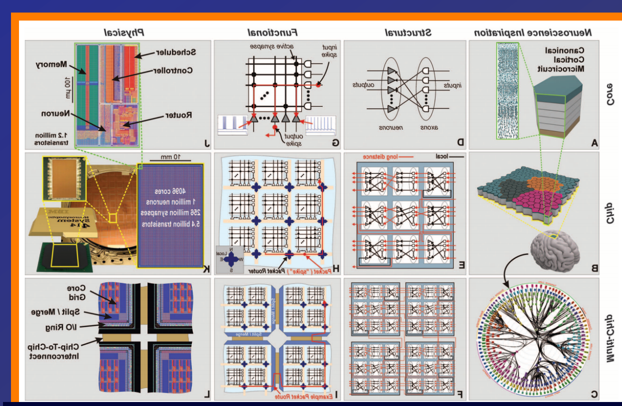
- Memristor
- Spintronics
- MQW Devices
- Nanoscale CMOS



Perceptron Mark 1



Hopfield Net Chip



TrueNorth (IBM) ~ 1 M Neurons

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

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 (2010-2016)

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

Q-Learning for Maze Search Algorithm on Memristor Array

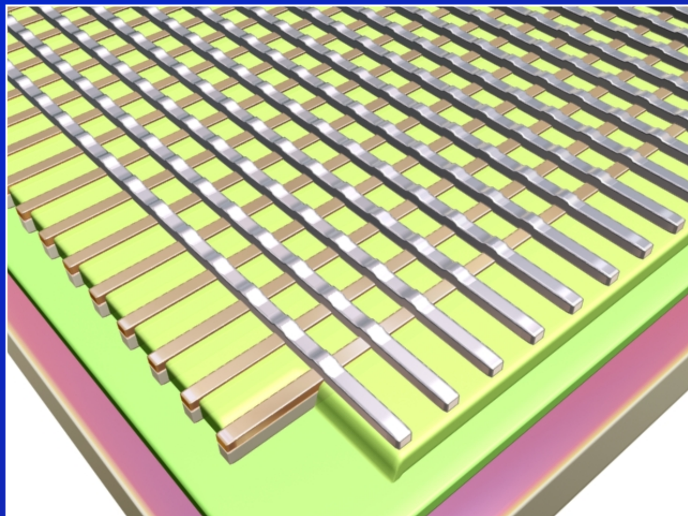
Reinforcement Learning/Actor-Critic NN for Optimal Nonlinear Control Applications

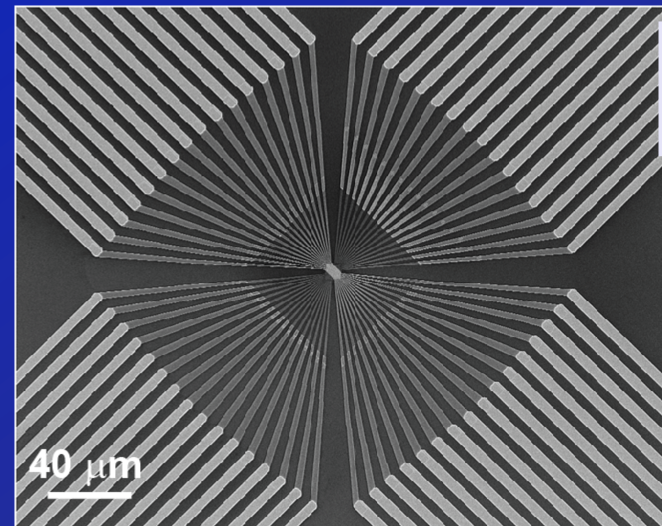
- Proceedings of the IEEE, 2012
- Nano Letters Journal, 2010
- IEEE Trans. on Computer, 2016
- IEEE Nanotechnology, 2011
- IEEE Nanotechnology, 2014
- IEEE Cellular Neural Networks, 2012

Neuromorphic Self-Healing Memory Design using Memristor Array

Product Code (SEC), Augmented PC (DEC) → Requires Muxes (~8%)
Hamming Code (SEC) → Higher Overhead
Projective Geometry Code (DEC/TED) → Galois Field Decoding, …
BCH & Reed Solomon (DEC) → Decoding complexity is high

High Density a-Si Based Nano-Crossbar

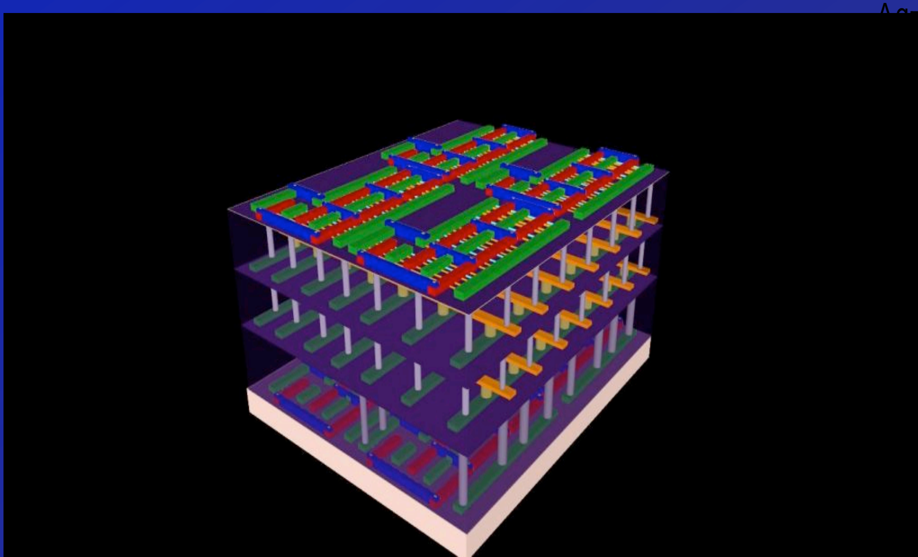


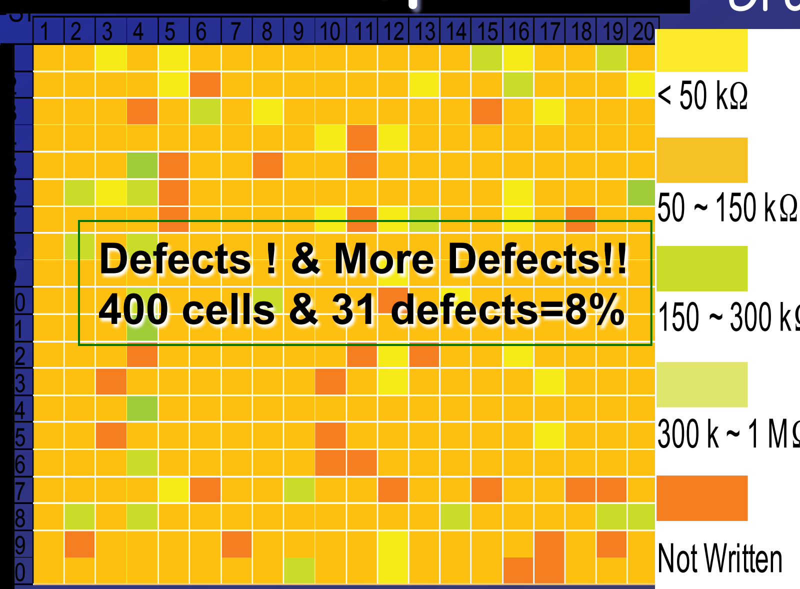


1kb crossbar array

Jo et al. *Nano Lett.*, 9, 870 (2009). From Prof. Wei Lu's Research Group

It is Scalable and CMOS Compatible.

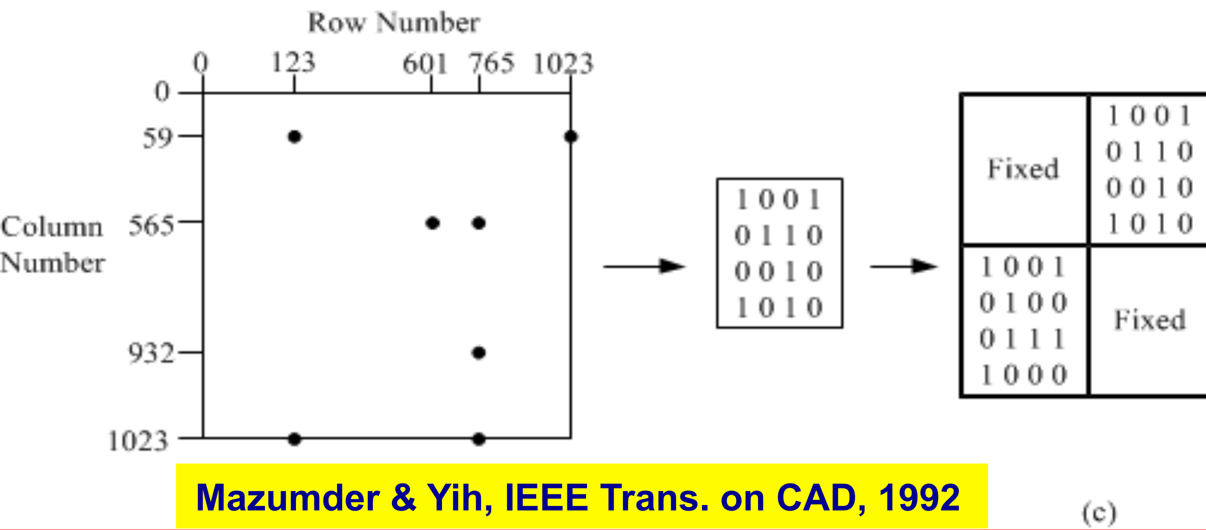


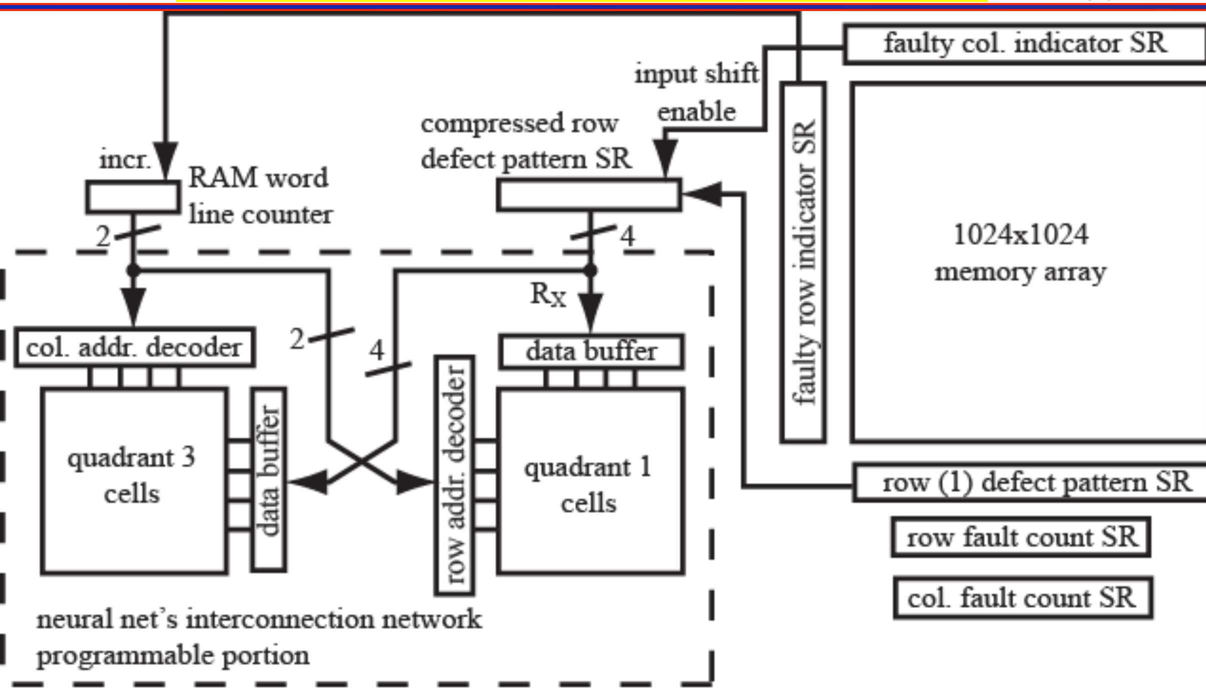


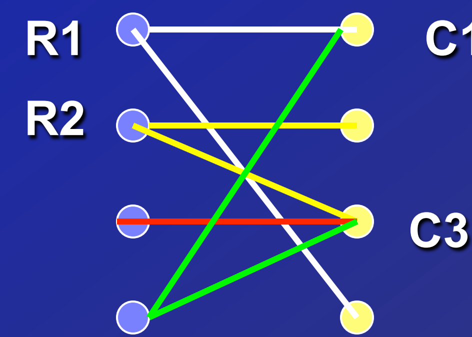
Defects ! & More Defects!!
400 cells & 31 defects=8%

Fabbed by Wei Lu's Group at Univ. of Michigan

Compaction of Faulty Array & Find Vertex Cover in Bipartite Graph







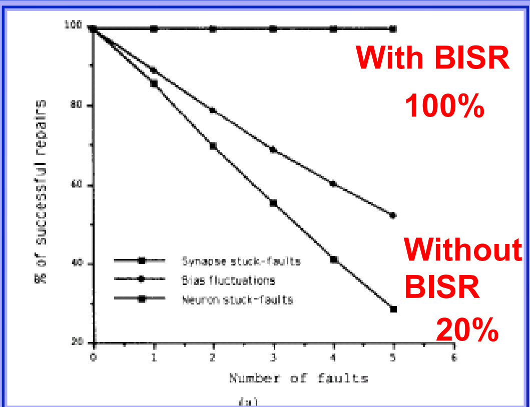
Find Vertex Cover:

[R1, R2; C1, C3]

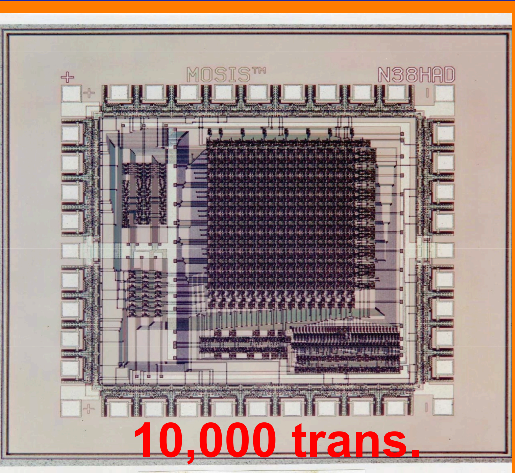
Defective Cells are Edges in Bipartite Graph

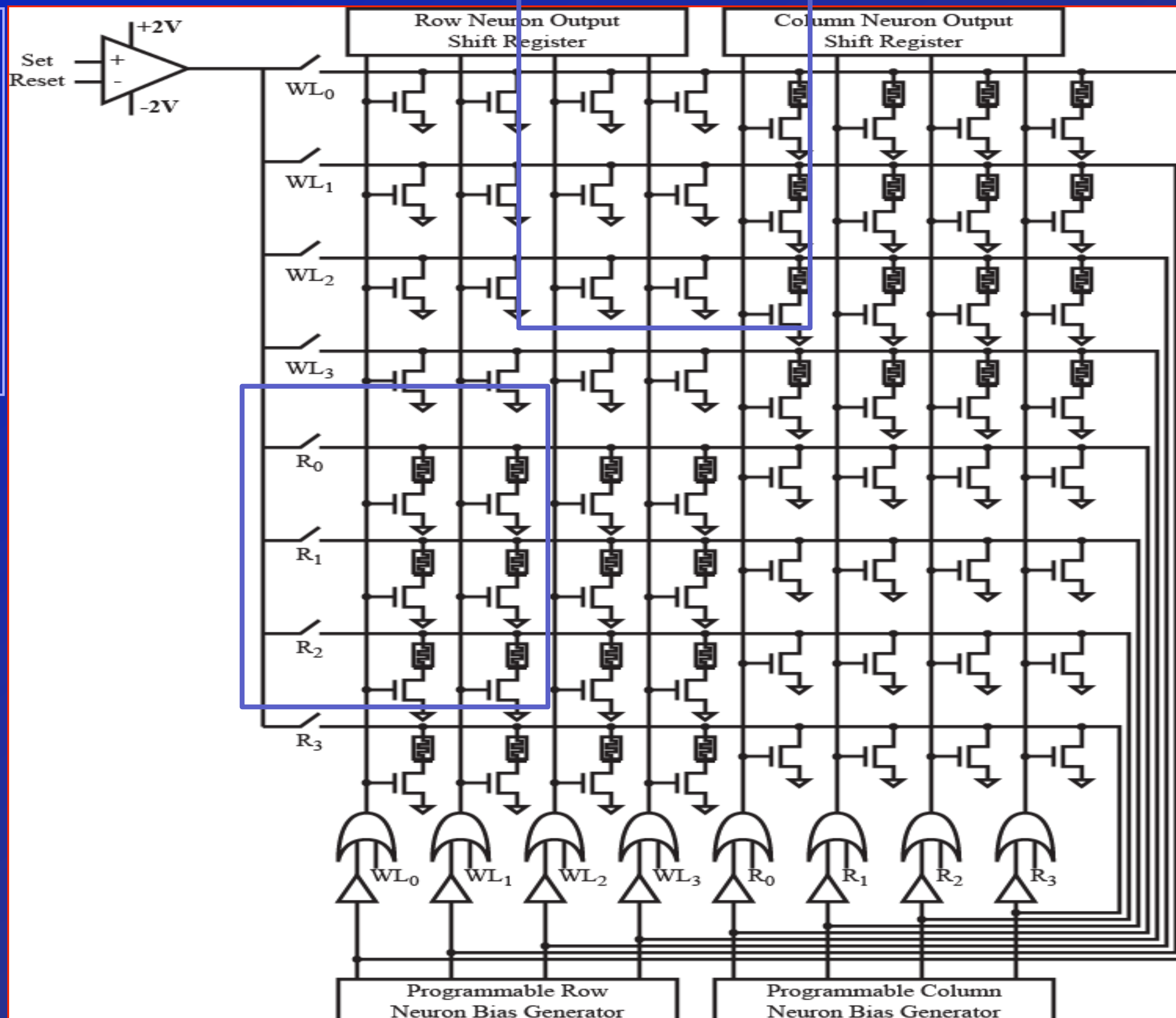
Unrestricted Vertex Cover Problem can Be solved in Polynomial Time by Bipartite Graph Matching Algorithm. However, Restricted Vertex Cover Problem is NP-complete.

Neuromorphic Self-Healing for Any Type of Memory Array



M.D. Smith & P. Mazumder, ***IEEE Trans. on Computers***, Vol. 45, No. 1, Jan. 1996, pp. 109-115.

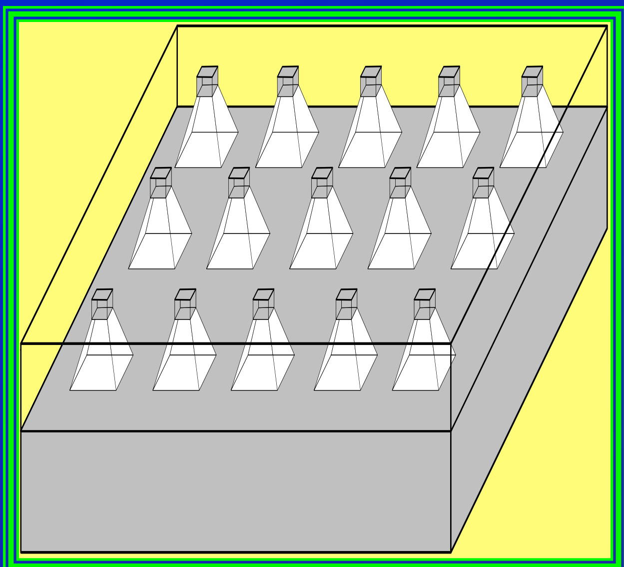


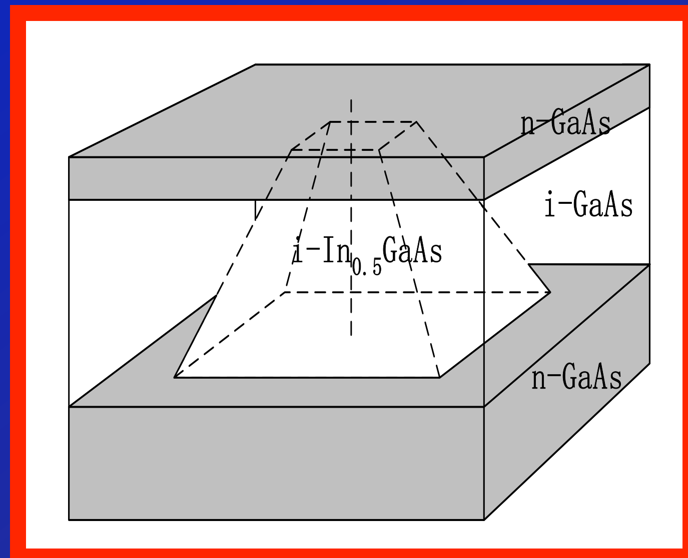


Fusion of Sensing & Processing

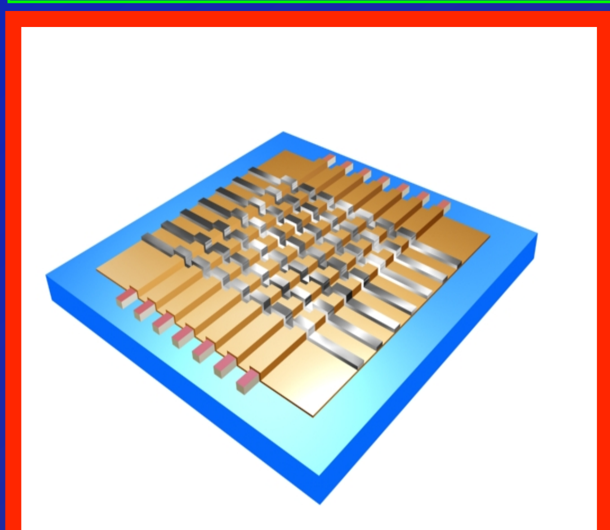
Nanoscale
Cellular Neural Network (CNN)
by Using
Quantum Tunneling Devices &
Memristor Array

Fusion of Sensing & Processing

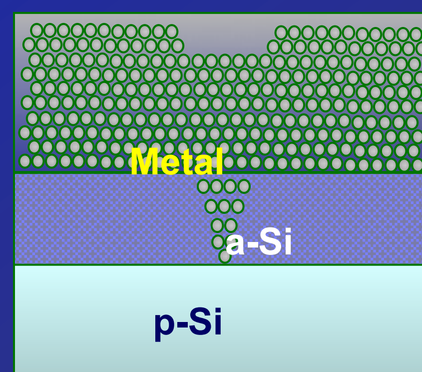




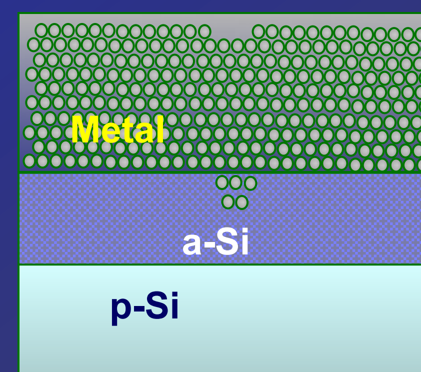
3-D Confined Quantum Box Array



Memristor Array



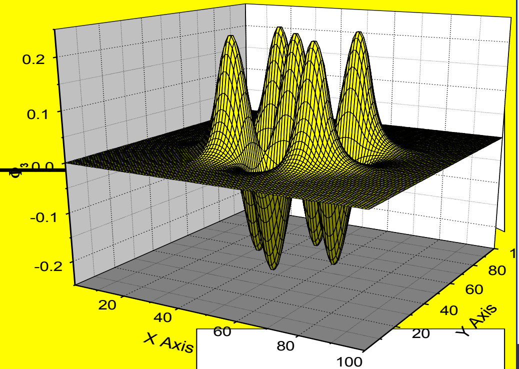
Low Resistance State (LRS)



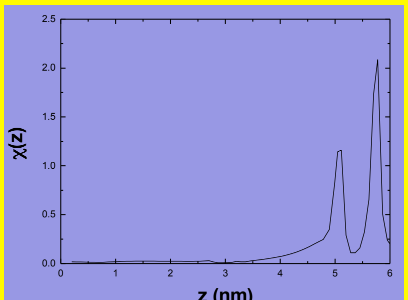
High Resistance State (HRS)

3D-SBEM-TIS

start → Define quantum dot structure → Calculate 2D wave function in xy plane → Solve 1D Schrodinger equation (in z direction) → Using boundary condition to get 3D S-Matrix → Integrate T(E) over all possible energy → end



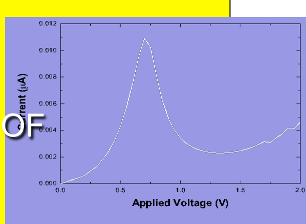
- Calculate 2D wave function in xy plane → 2D wave function



- Using boundary condition to get 3D S-Matrix → Transmission rate T(E)

$$T = T^{(N)} \cdot T^{(N-1)} \cdots T^{(2)} \cdot T^{(1)}$$

- Integrate T(E) over all possible energy → Tunneling current

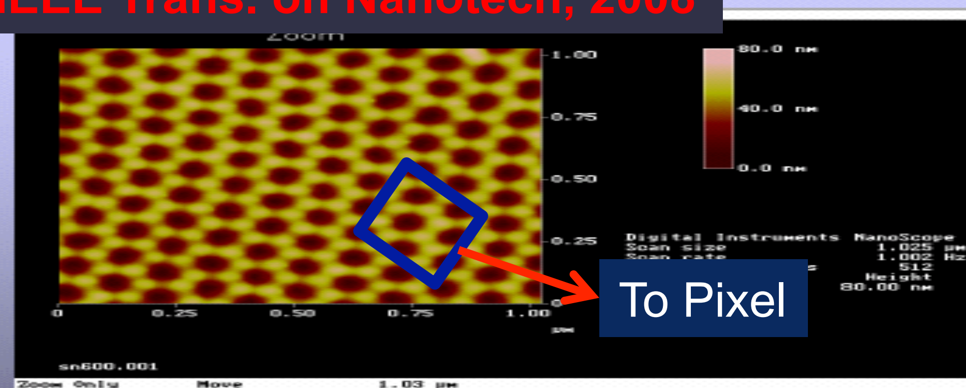


VCU Circuit parameters based on measured values

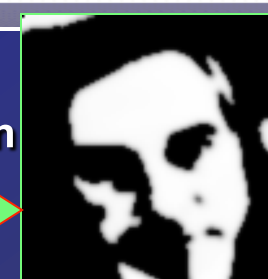
- Single dot : $R_{inter\text{-}dot}$ = 640 MΩ
- Single dot: $C_{inter\text{-}dot}$ = 5 aF
- Single dot: $C_{substrate}$ = 0.5 aF
- Single dot: Peak current = 1.5 nA

- Superdot (1 pixel = 6400 dots): $R_{inter\text{-}superdot}$ = 8 MΩ
- Superdot: $C_{inter\text{-}superdot}$ = 4 fF
- Superdot: $C_{substrate}$ = 3.2 fF
- Superdot: Peak current = 0.1 μA

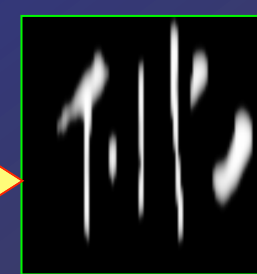
IEEE Trans. on Nanotech, 2008



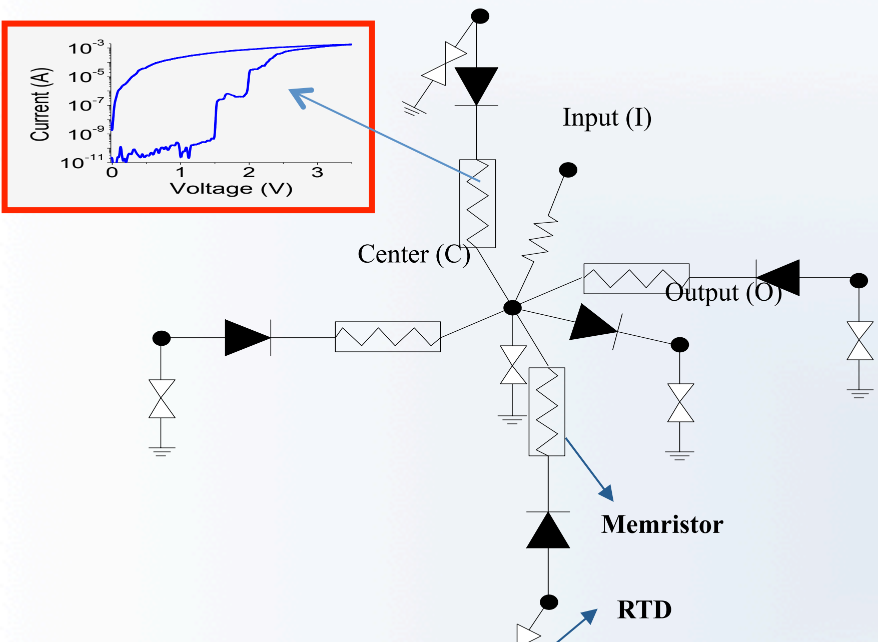
Edge Detection



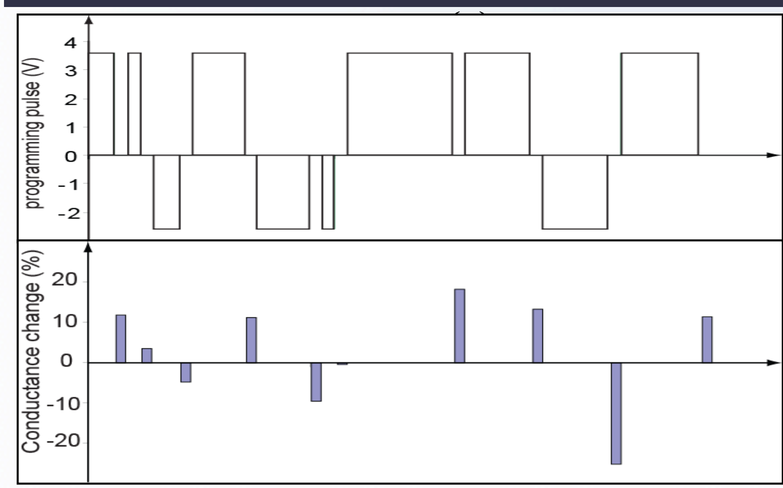
Vertical Line Detection

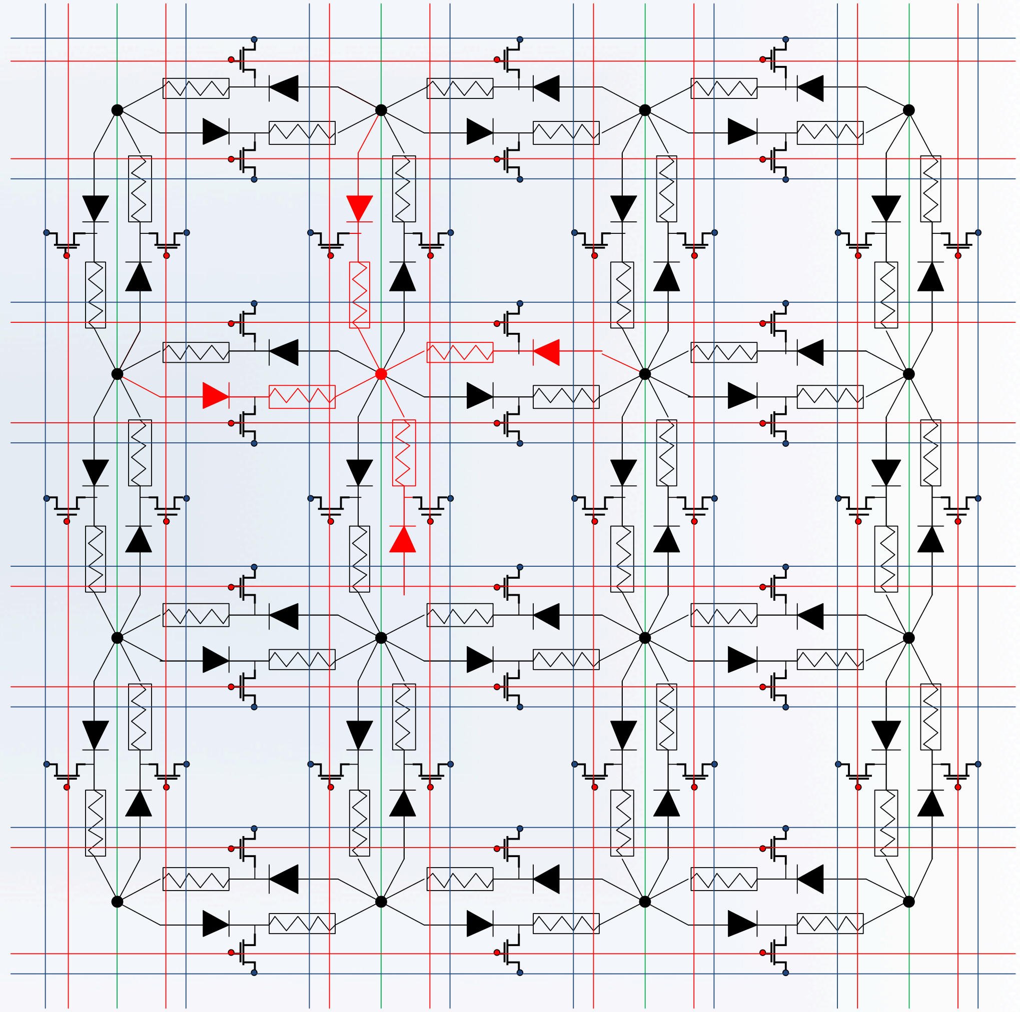


Analog Programmable Nano-Architecture for Static and Dynamic Image Processing



IEEE Trans. On Nanotechnology, 2013

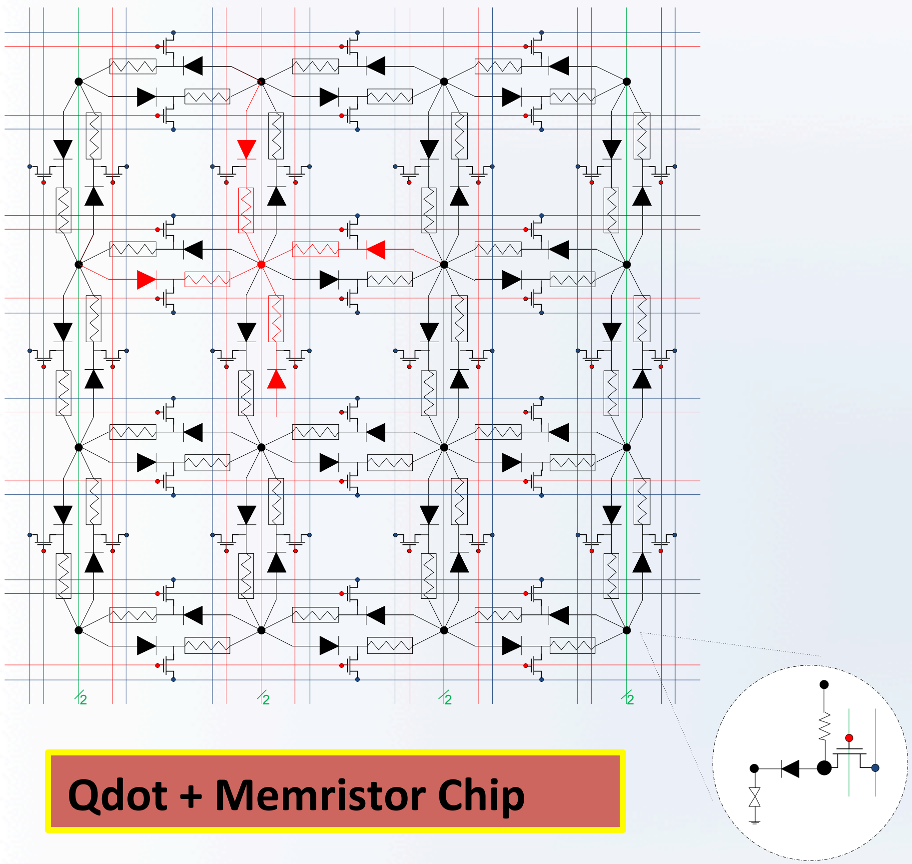




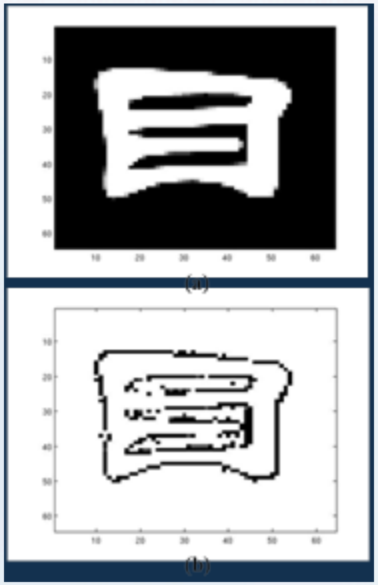
(b)

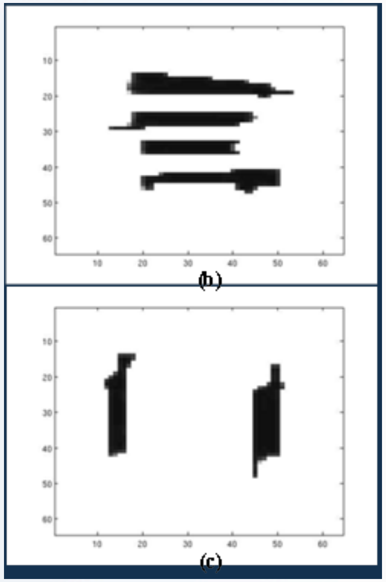
Qdot + Memristor Chip

Analog Programmable Unconventional Computing



IEEE Trans. On Nanotechnology, 2013





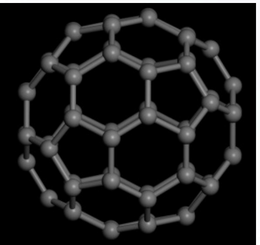
Edge Detection

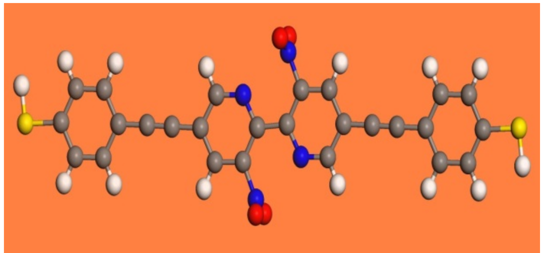
HLD & VLD

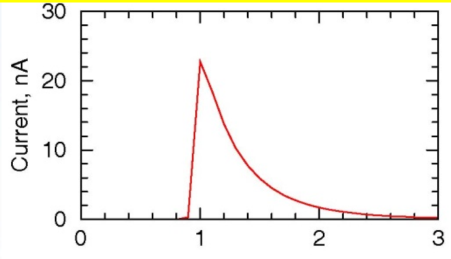
Qdot + Memristor Chip

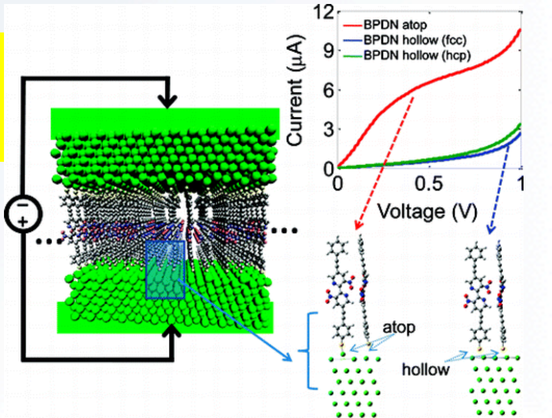
Molecular Computer?

$$H = \sum_i \varepsilon_0 (V_{tot}) c_i^+ c_i + h c_{i\pm1}^+ c_i$$



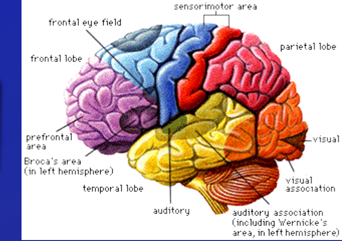




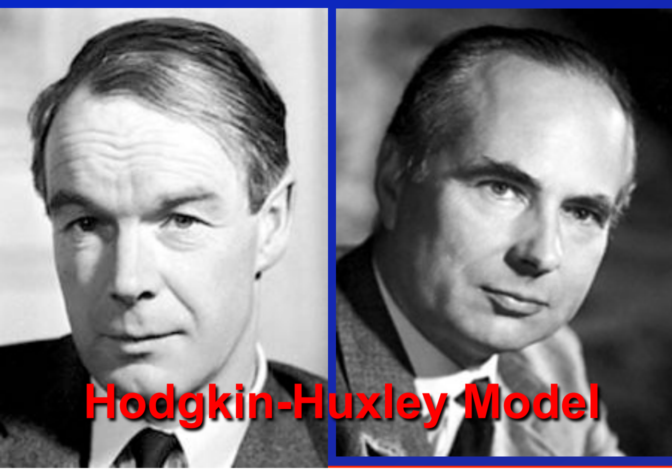


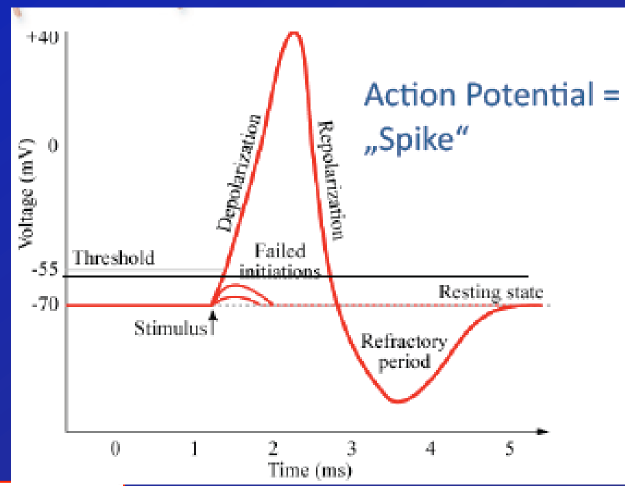
Spike Timing Dependent Plasticity (STDP) Learning Networks using Memristors

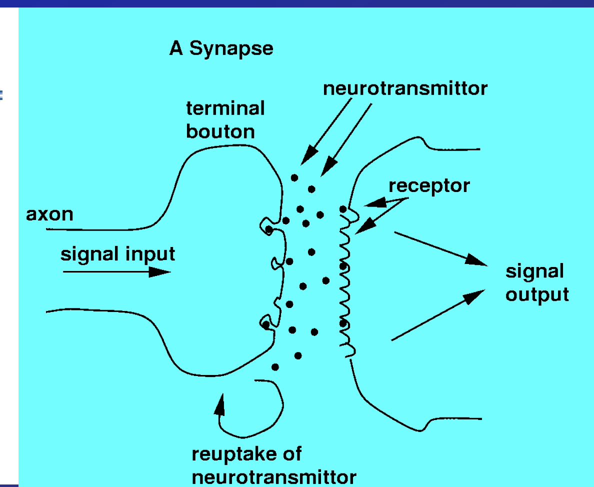
Biological Neuron Model



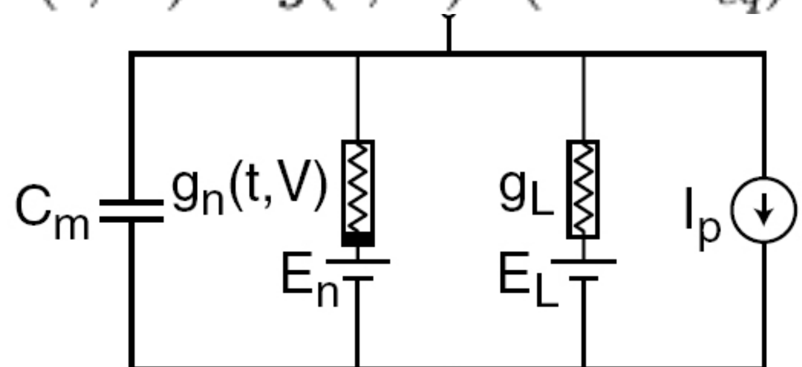
Ionic Transport in Biological Neuron & its Silicon Implementation



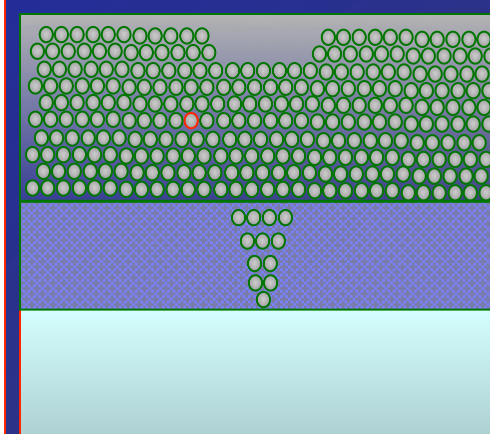


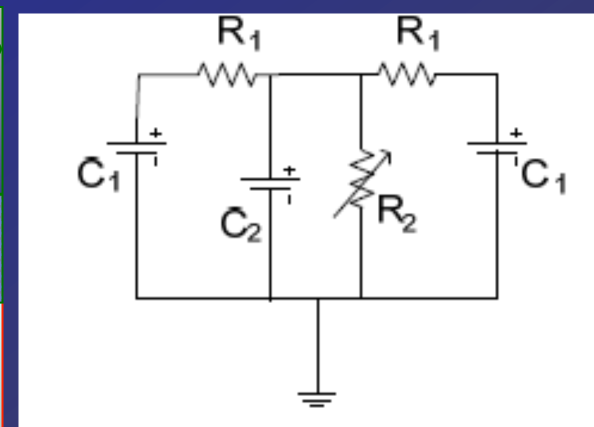


$$I(t,V) = g(t,V)\cdot(V - V_{eq})$$

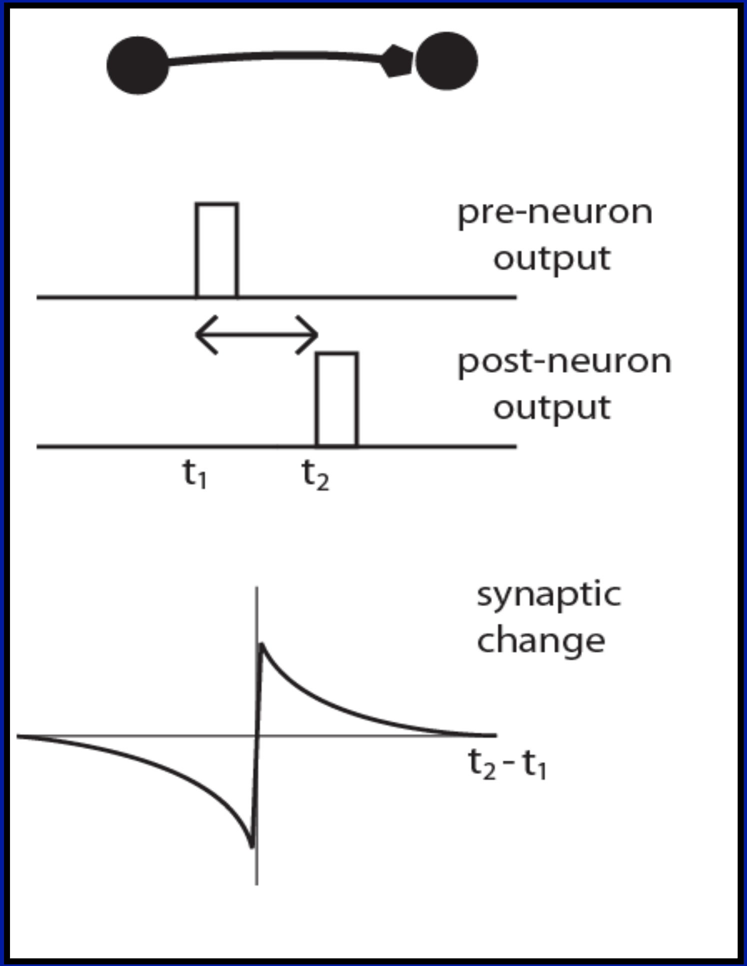


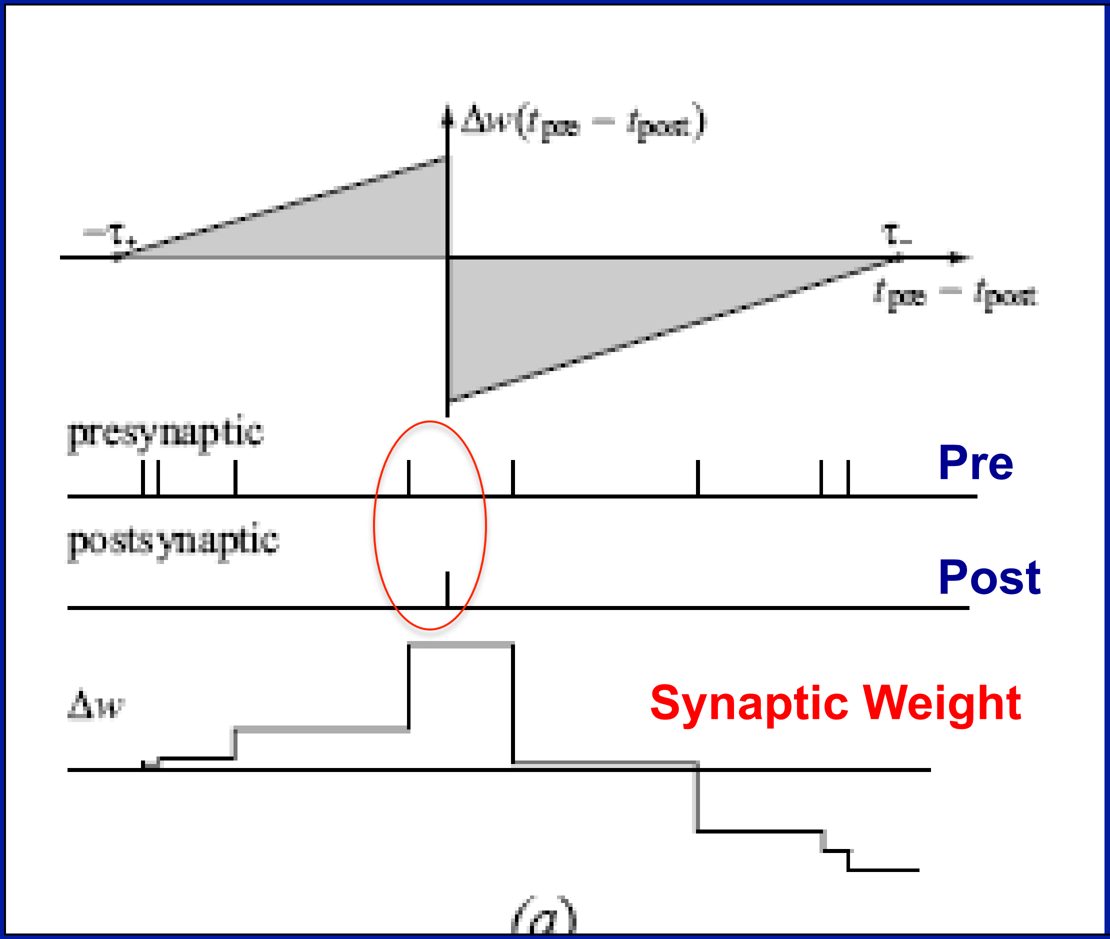
$$C_m \frac{dV(t)}{dt} = -\sum_i I_i(t,V)$$





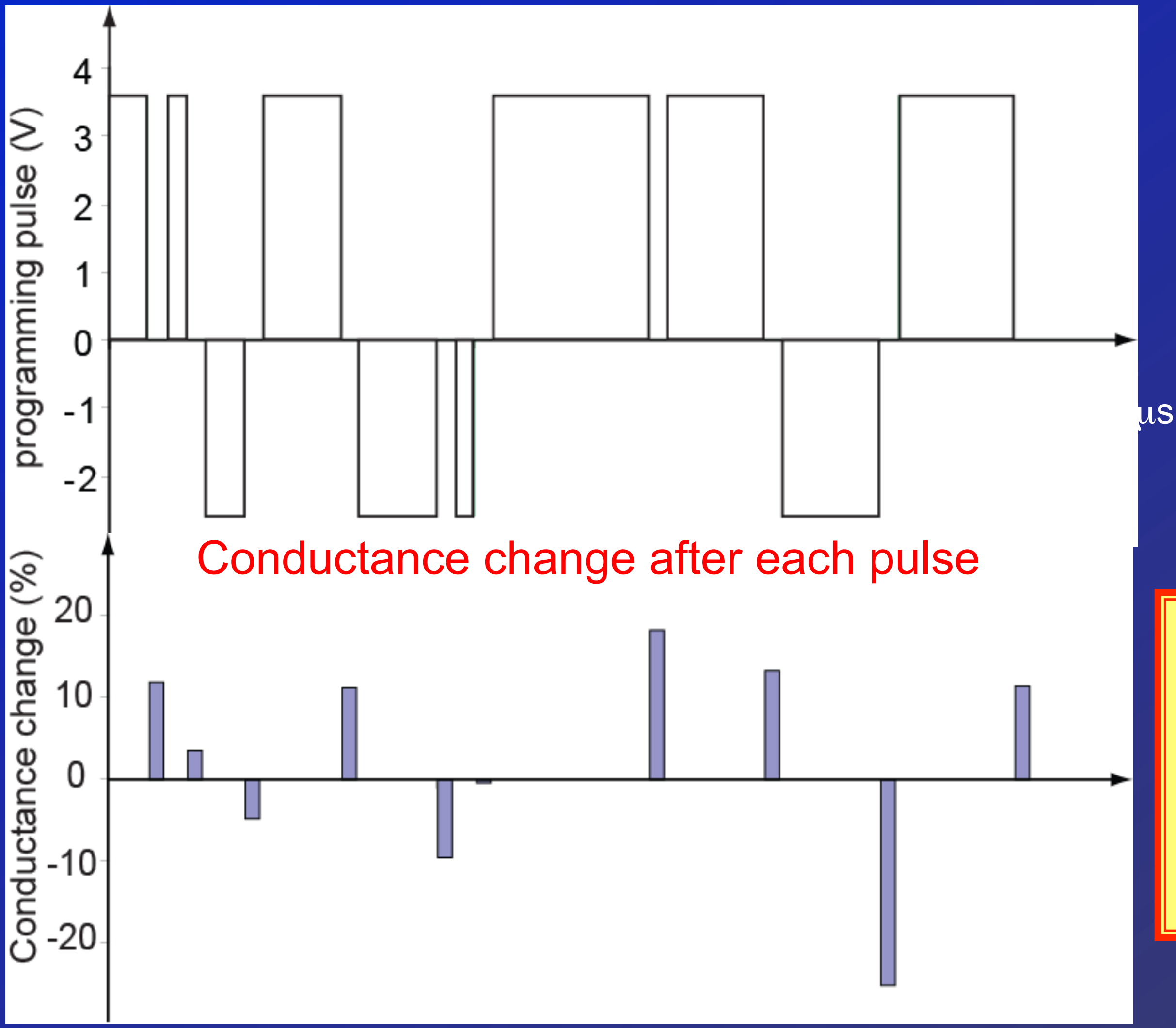
Spike Timing Dependent Plasticity (STDP) Learning Networks

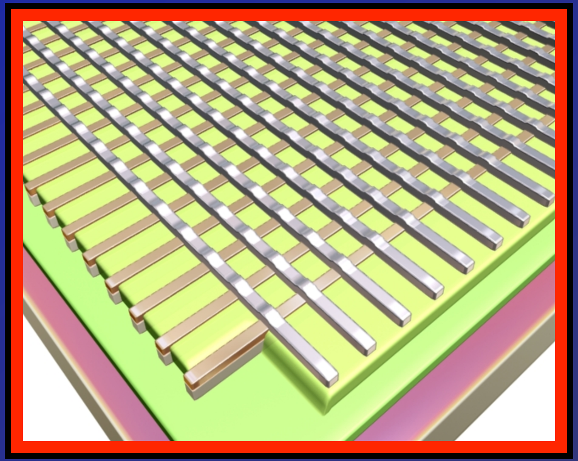




Previous Approach of STDP Implementation On Crossbar with Constant Amplitude

Programming pulses with different pulse widths

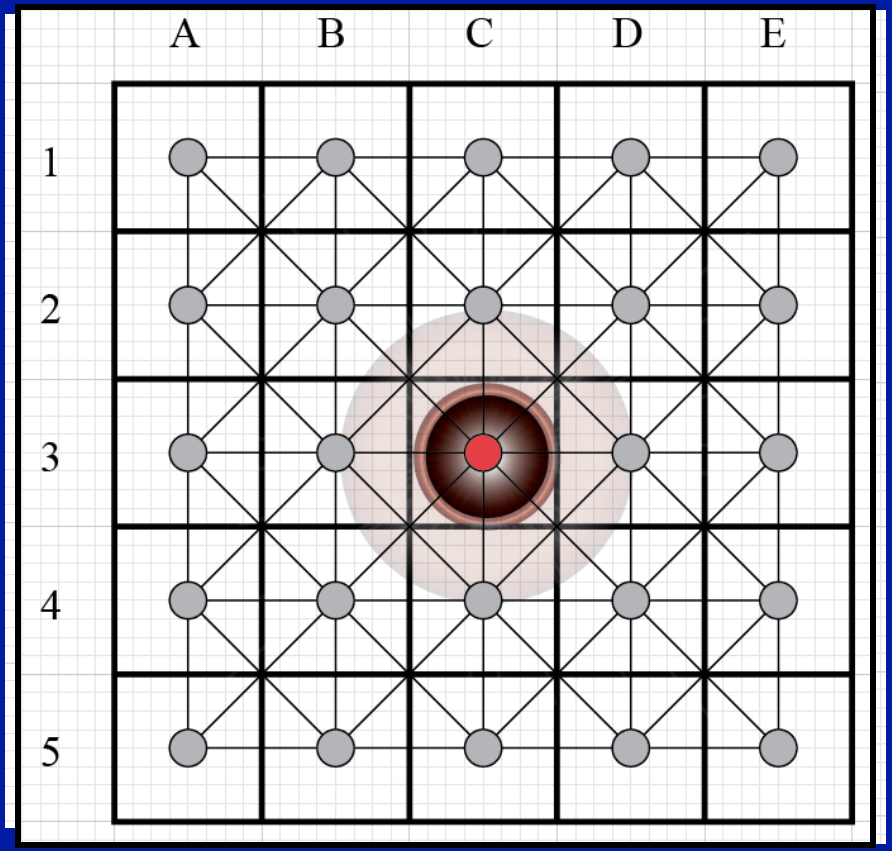




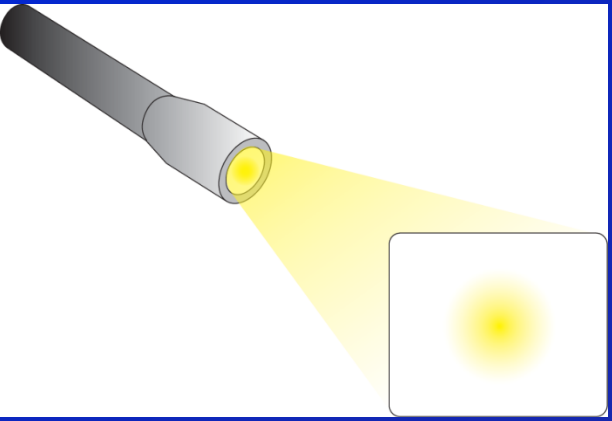
1. Digitally Controlled
2. Constant Amplitude
3. Temporal Correlation

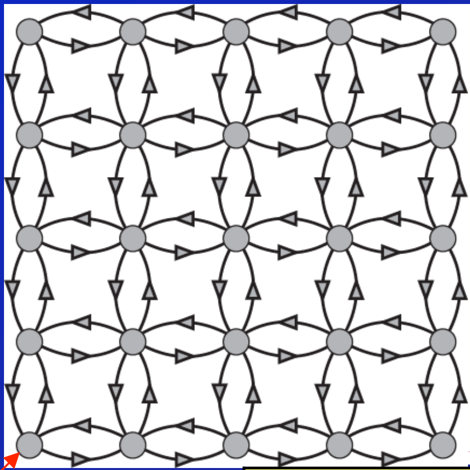
Position Detector Application

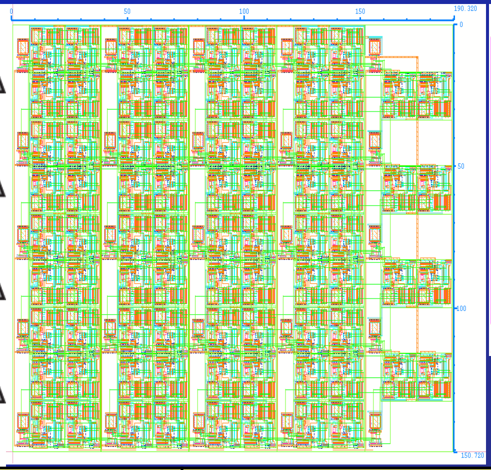
- **Split up area into a 5x5 grid**
- **Each grid has one neuron that is connected to an adjacent neuron through STDP synapse**
- **Detect a light source at any position on the grid**
- **No post-processing circuitry necessary**
- **k-Winner-Take-All (k-WTA) implementation**

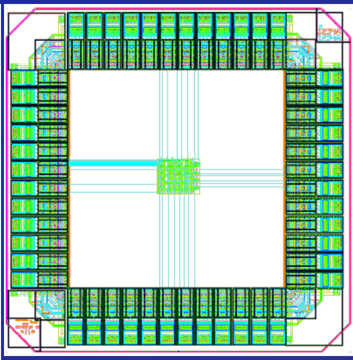


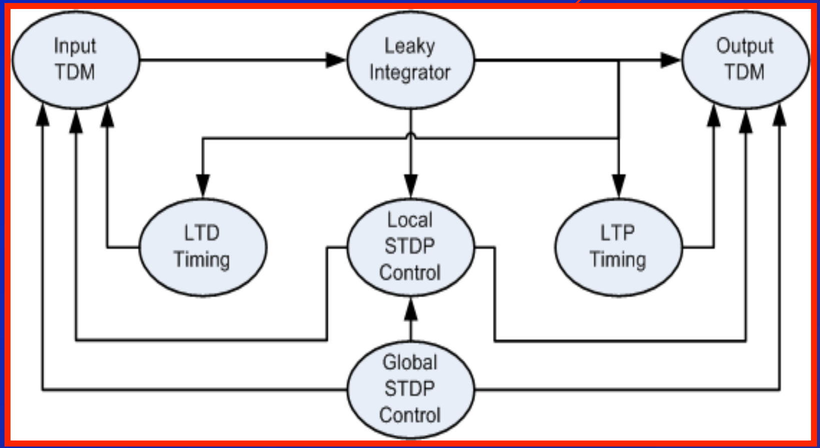
STDP Neural Circuit for Position Detector

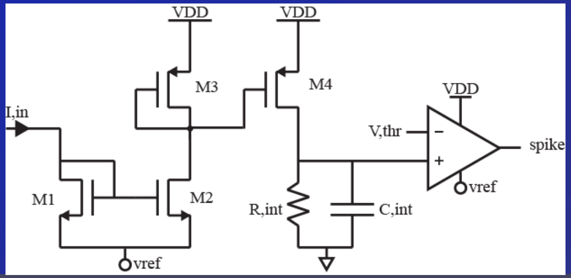








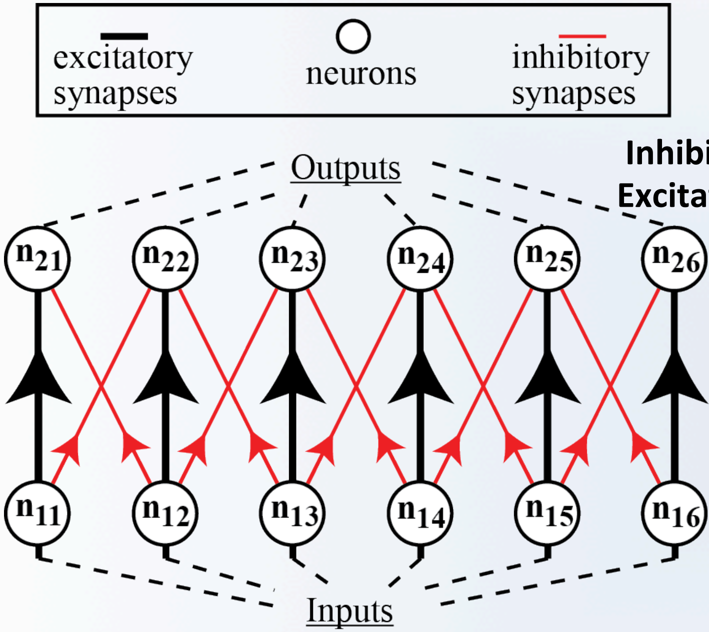




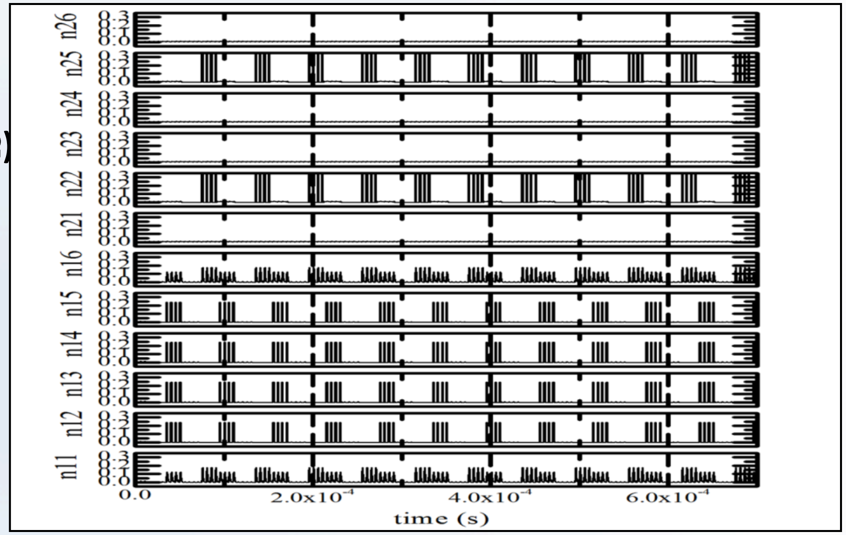
Bases	Memristor Design	CMOS Design
Synaptic area	< (0.5µm x 0.5µm)	17µm x 16µm
Synaptic Density	➢4 devices/μm^2 x1000	0.0037 devices/ μm^2
Neuron area	20µm x 10µm	8µm x 12µm
Neuron Density	0.005 devices/ μm^2 x2	0.0104 devices/ μm^2
Volatility	Nonvolatile	Volatile

Ebong, and Mazumder, *Proceedings of the IEEE*, Feb. 2012.

STDP Circuits for XOR/Edge Detector

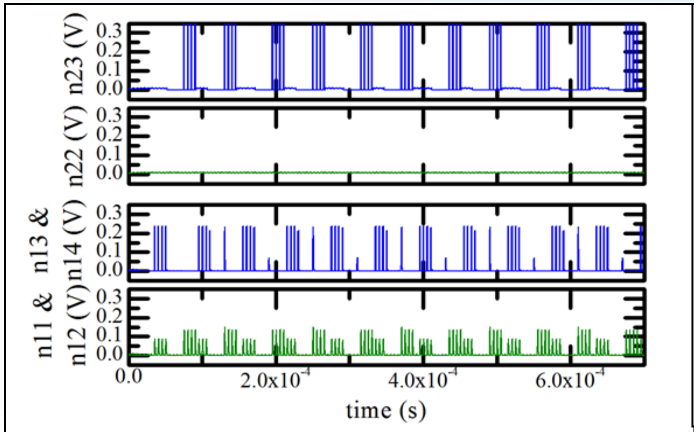


Inhibitory =20 MΩ
Excitatory = 5.6 MΩ)

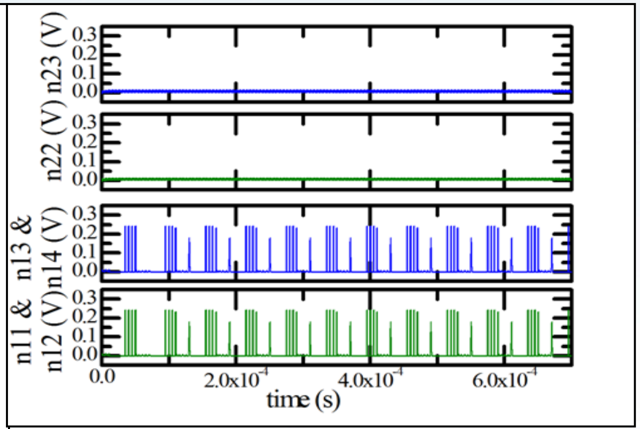


Input = 011110; Output = 010010

n23 = n12 XOR n14 = 1 when n12=0 & n14=1



n23 = n12 XOR n14 = 0 when n12=1 & n14=1



Basic Neuromorphic Functions: WTA, XOR, Edge Detector Were implemented Using STDP learning scheme

IEEE Nano 2011

Reinforcement Learning Networks

Memristor Based
Q-Learning Network

CMOS Digital
Actor-Critic Network

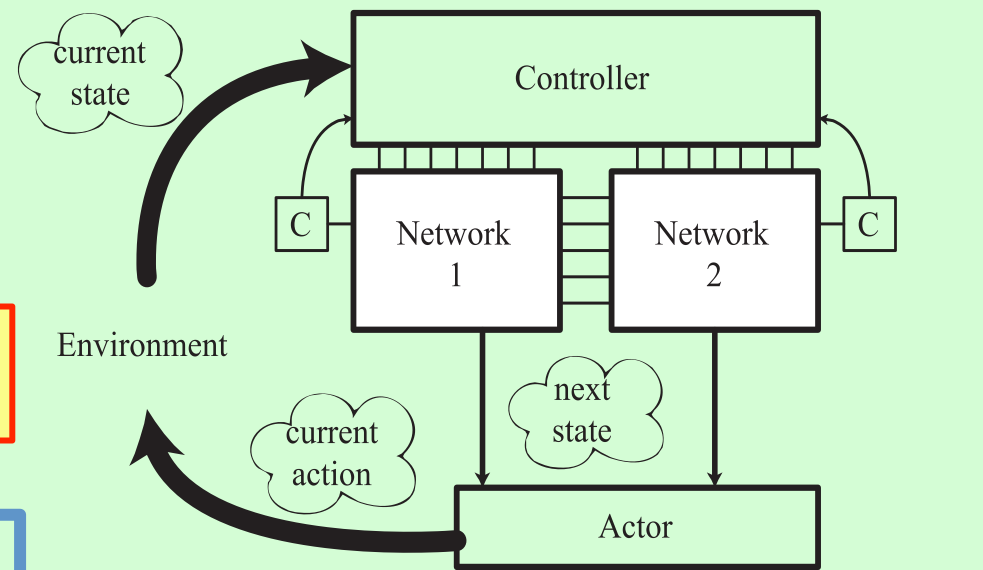
Q-Learning Hardware – Reinforcement Learning

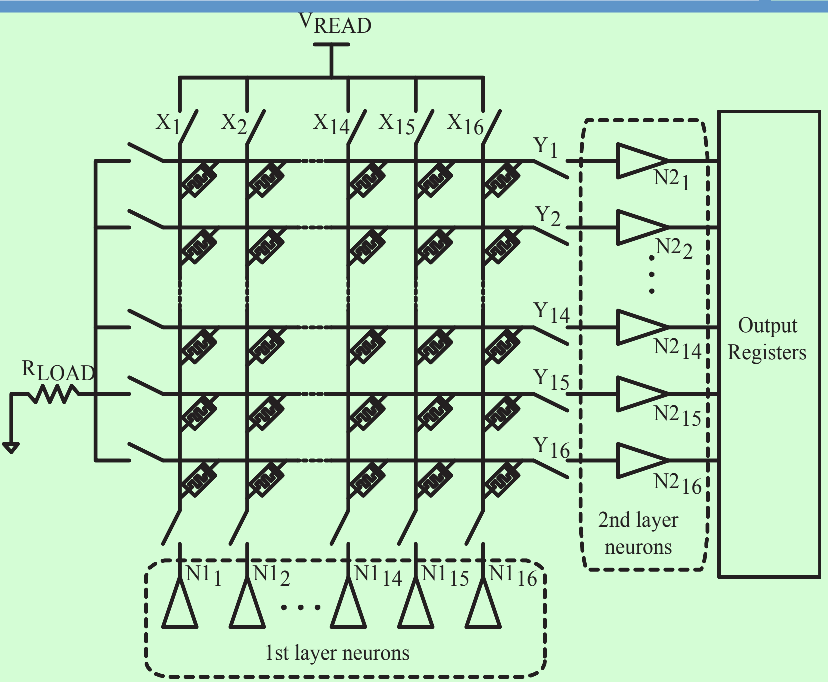
$$\tilde{Q}(s_t, a_t) \leftarrow \tilde{Q}(s_t, a_t) \times (1 - \alpha_t(s_t, a_t)) + \alpha_t(s_t, a_t) \times r_t + \alpha_t(s_t, a_t) \times \max_{a_{t+1}} \left[\tilde{Q}(s_{t+1}, a_{t+1}) \right]$$

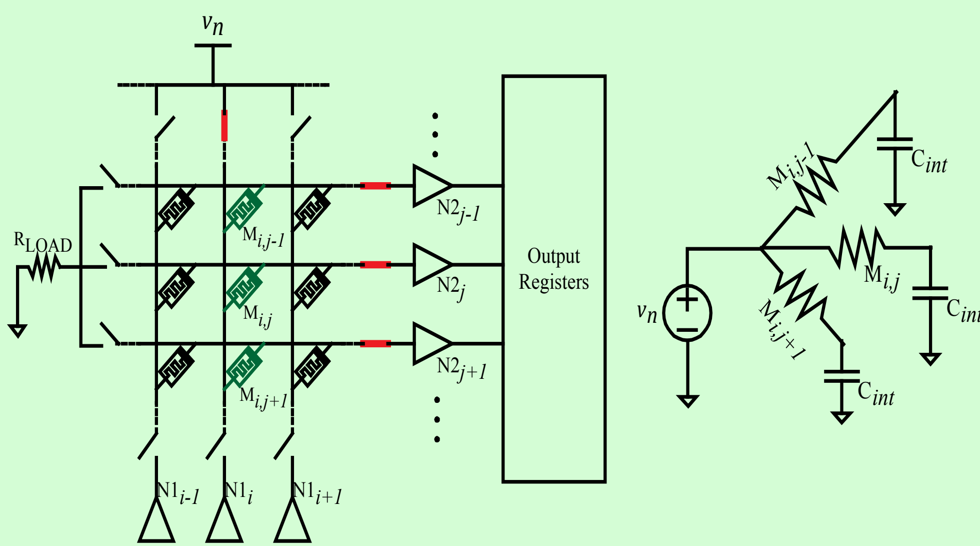
$$\max_{a_{t+1}} \left[\tilde{Q}(s_{t+1}, a_{t+1}) \right] = \tilde{Q}(s_t, a_t) + \delta_t.$$

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

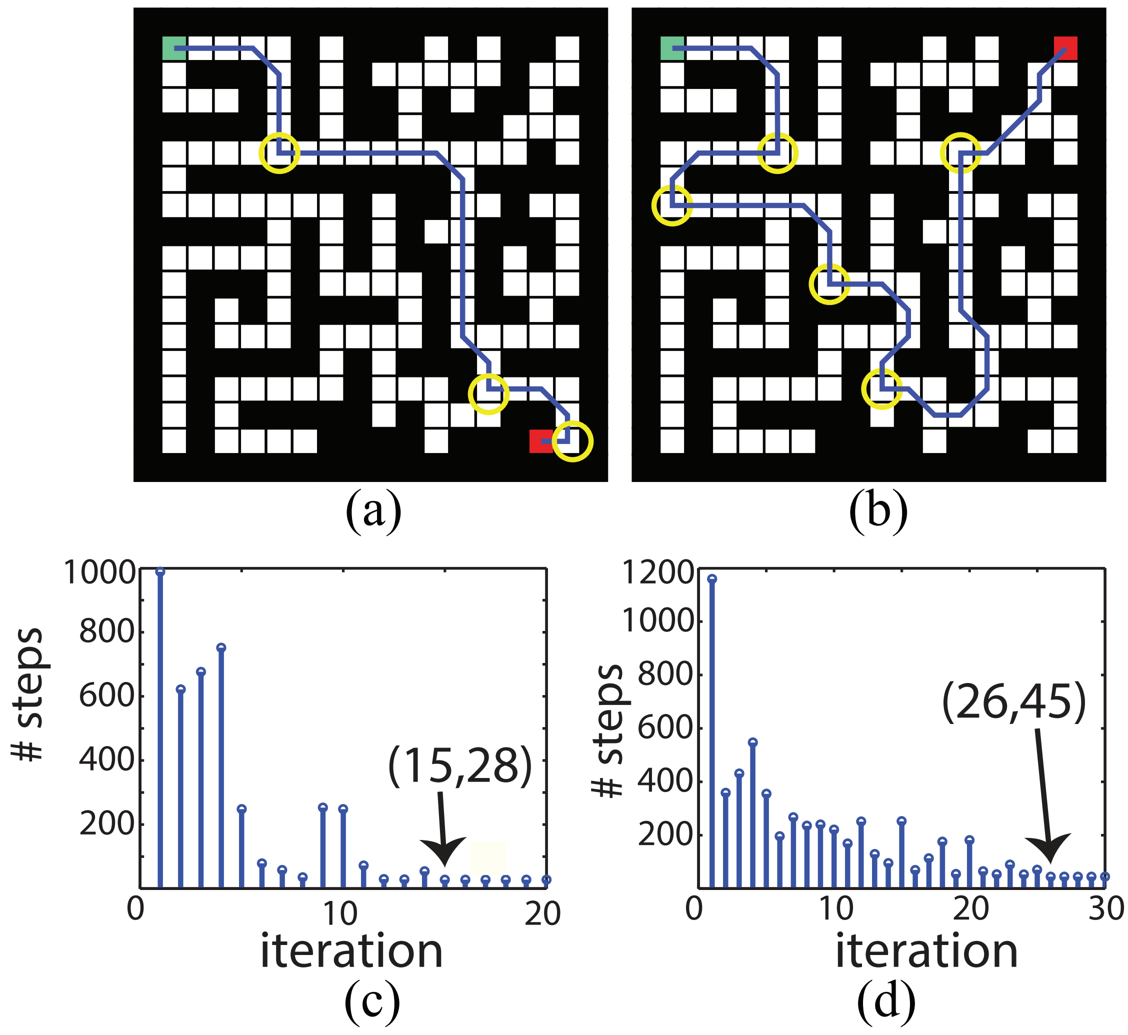
Evaluative Feedback (Rewards)

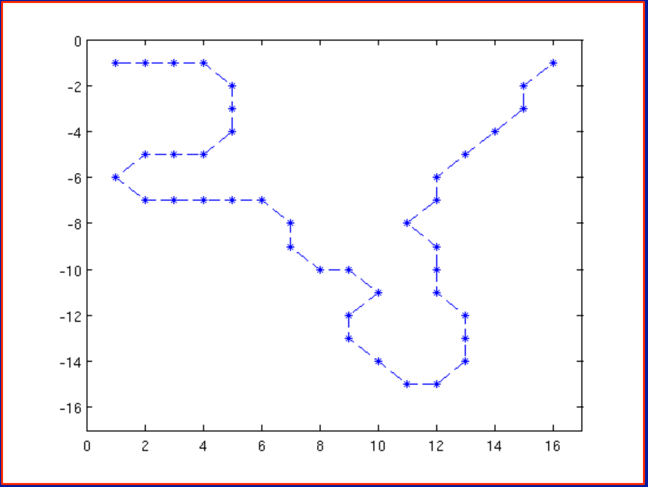




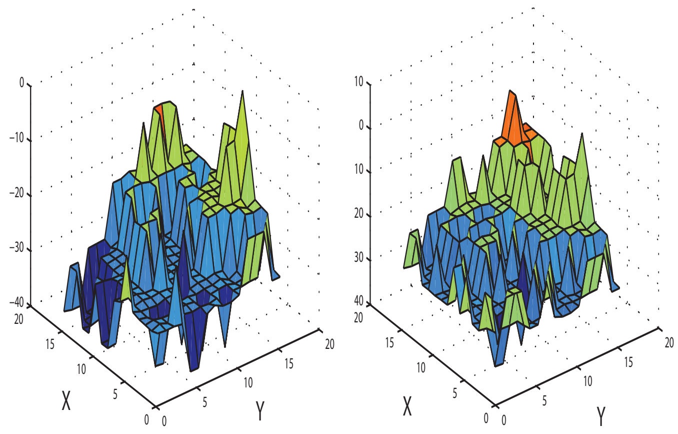


Performance of Memristor Q-Learning Hardware





Memristor Model Used in Matlab Simulation

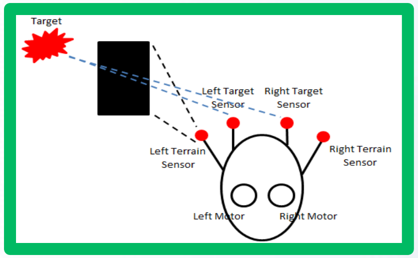


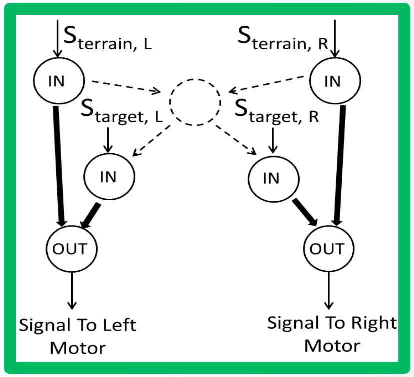
Idong and Mazumder, IIEEE Nano 2014

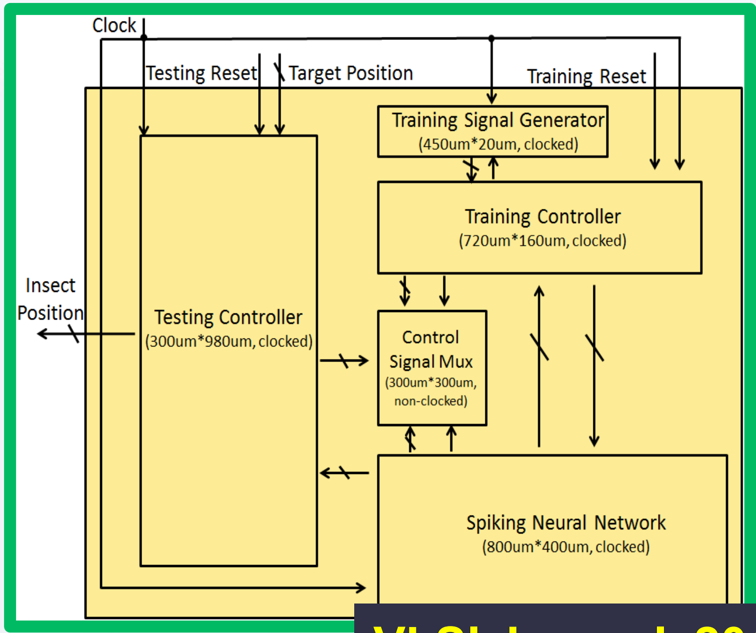
Synapse States after 1st and 2nd Iterations

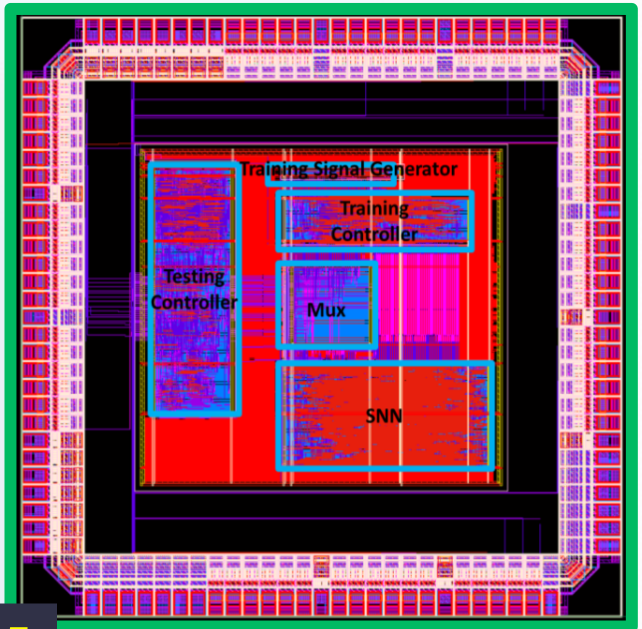
Plasticity (STDP) Based Learning Chip for Virtual Bug Navigation

Non-Evaluative Feedback (Correlation)









VLSI Journal, 2015

$$\begin{cases} V_x = V^* * \cos(\theta) \\ V_y = V * \sin(\theta) \\ V = \frac{V_L+V_R}{2} \\ \Delta\theta = \frac{V_R-V_L}{L} \\ \Delta V_L = -\frac{V_L}{\tau_{motor}} + \eta * (t == t_L^f) \\ \Delta V_R = -\frac{V_R}{\tau_{motor}} + \eta * (t == t_R^f) \end{cases}$$

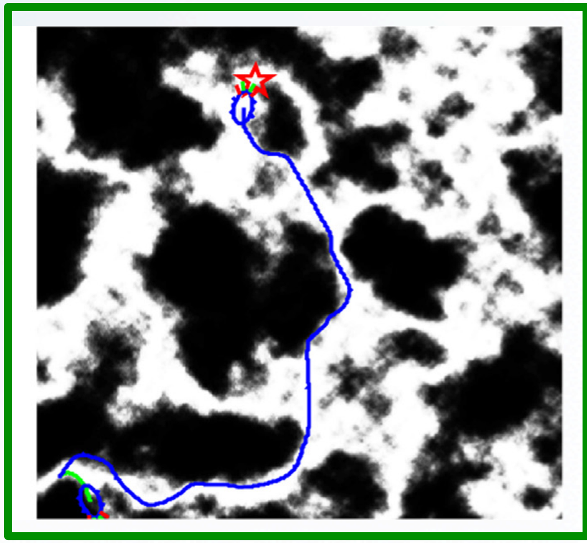
$$D_{ji}^r = T_j^r - T_i^r$$

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

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

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

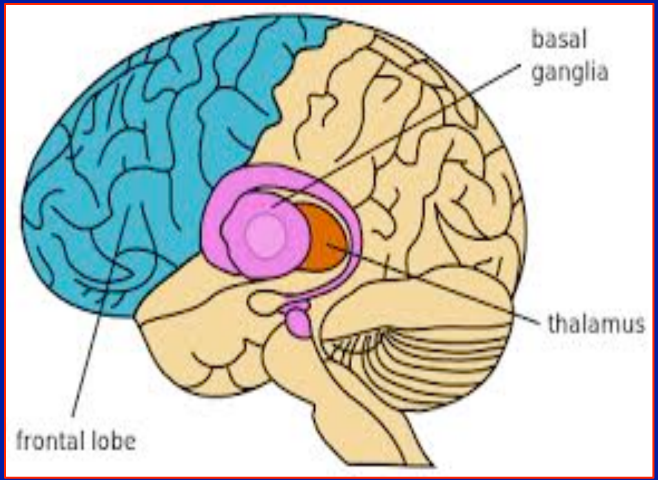
$$D_{ji}^{r+1} = D_{ji}^r - \alpha \frac{\partial e}{-A_- \cdot e^{\frac{D_{ij}^r}{\tau_-}}} \qquad D_{ij}^r < 0$$

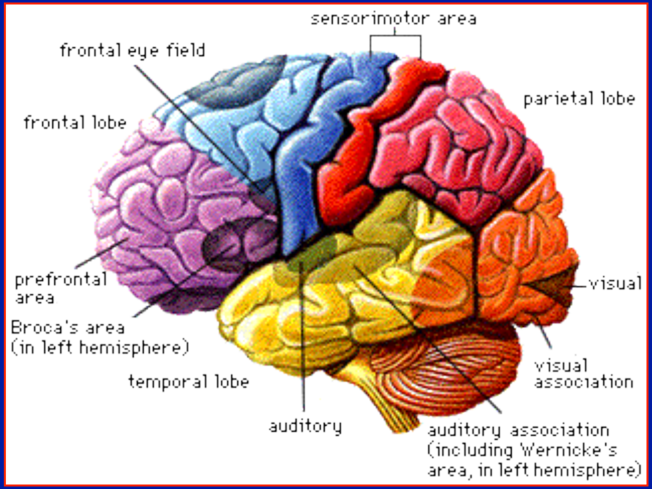


Conclusion

Facets of Neuromorphic or Brain-like Computing:

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





Adaptive Hardware Platform: Optimal control theory, multi-agent systems, swarm intelligence, robot control, computer games, telecommunications, smart grid for power distribution, and Markov decision process (MDP)