Why do neurons have thousands of synapses? A theory of sequence memory in the neocortex

Subutai Ahmad, Jeff Hawkins, and Yuwei Cui
sahmad@numenta.com, jhawkins@numenta.com, ycui@numenta.com

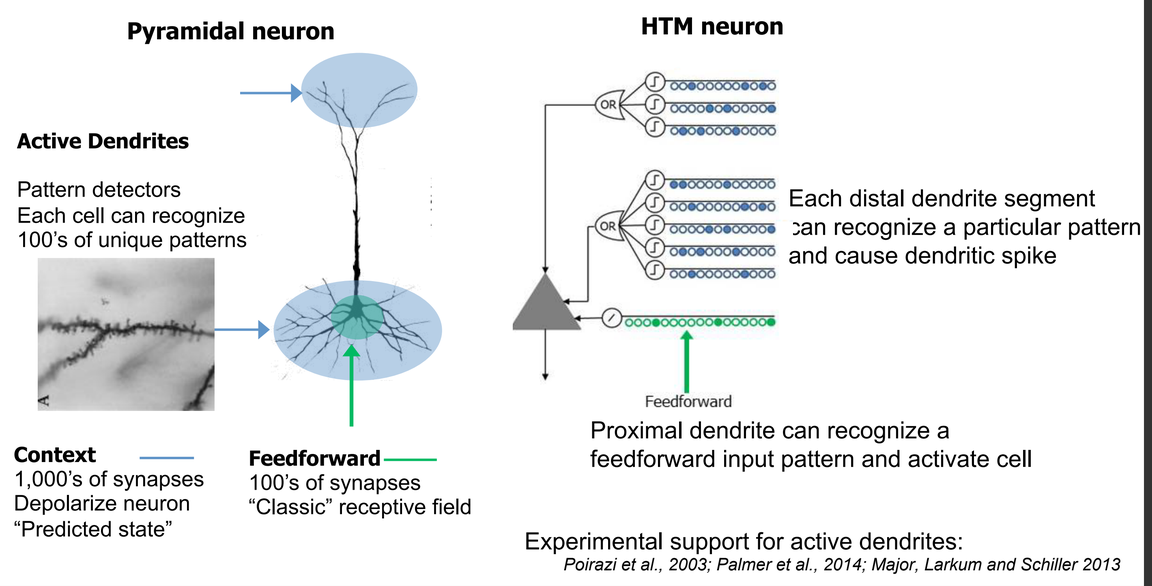
Sequence learning is ubiquitous in cortex

What is neural mechanism for sequence learning?

HTM sequence memory:

1. Neurons learn to recognize hundreds of patterns using active dendrites.
2. Recognition of patterns act as predictions by depolarizing the cell without generating an immediate action potential.
3. A network of neurons with active dendrites forms a powerful sequence memory.
4. Sparse representations lead to highly robust recognition.

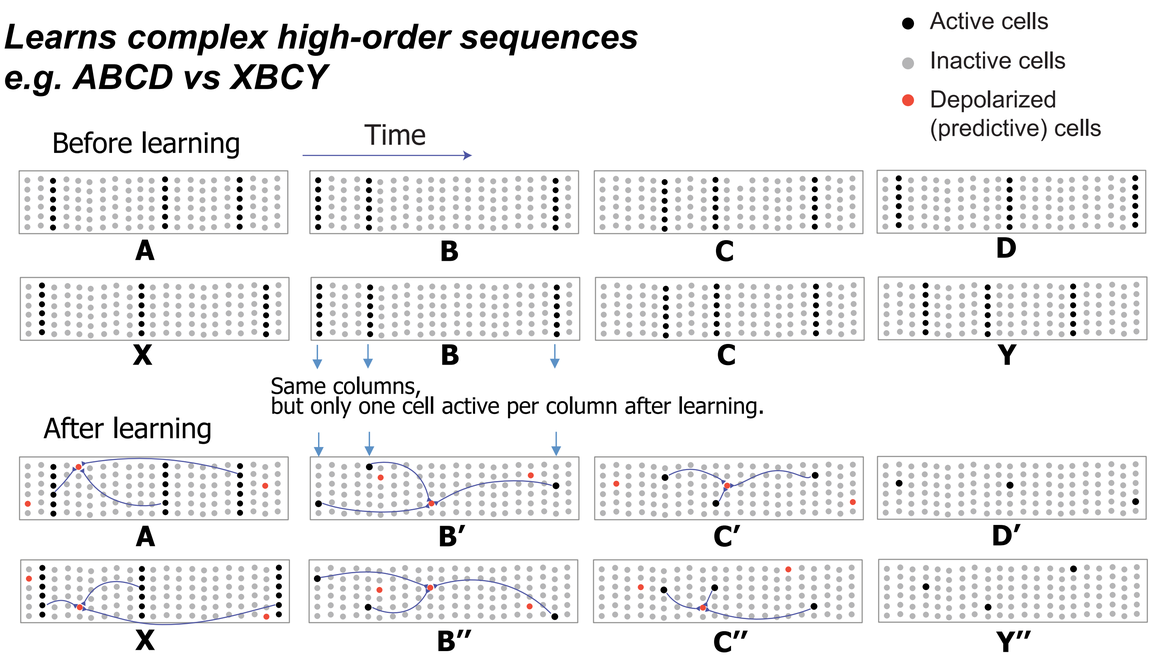
HTM neuron model:



Experimental support for active dendrites:
Poirazi et al., 2003; Palmer et al., 2014; Major, Larkum and Schiller 2013

HTM network model for sequence learning

Learns complex high-order sequences
e.g. ABCD vs XBCY



Learning and activation rules

Activation rules

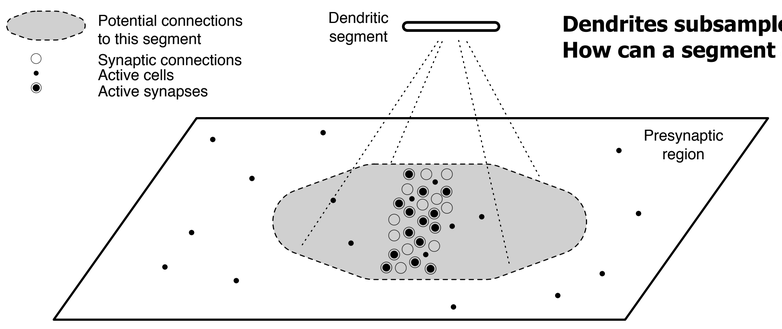
Select the top 2% of columns with strongest inputs on proximal dendrite as active columns
Detected pattern on distal dendrite causes cell to be depolarized (predicted)
If any cell in an active column is predicted, only the predicted cells fire
If no cell in an active column is predicted, all cells in the column fire

Unsupervised Hebbian-like learning rules:

If a depolarized cell becomes active subsequently, its active dendritic segment will be reinforced
If a depolarized cell does not become active, we apply a small decay to active segments of that cell
If no cell in an active column is predicted, the cell with the most activated segment gets reinforced

(Hawkins and Ahmad, 2016)

Sparsity, robustness, and active dendrites



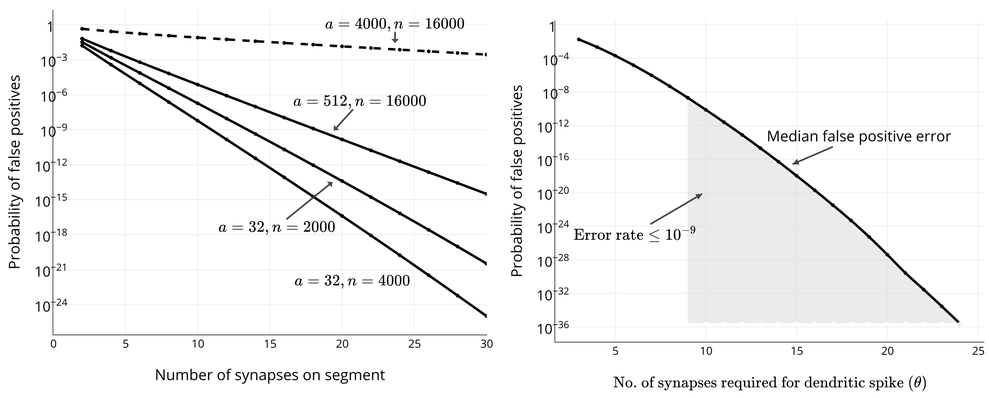
Dendrites subsample heavily from noisy population
How can a segment accurately detect patterns?

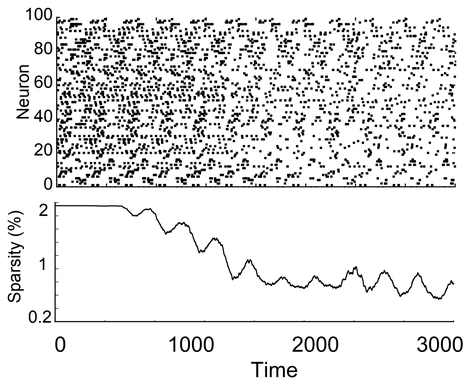
High-dimensional sparse representations = extremely accurate recognition

Number of synapses, s, between 20-40; typical threshold, θ, between 8-20
Presynaptic population, n, in the thousands; a = number of active presynaptic cells.

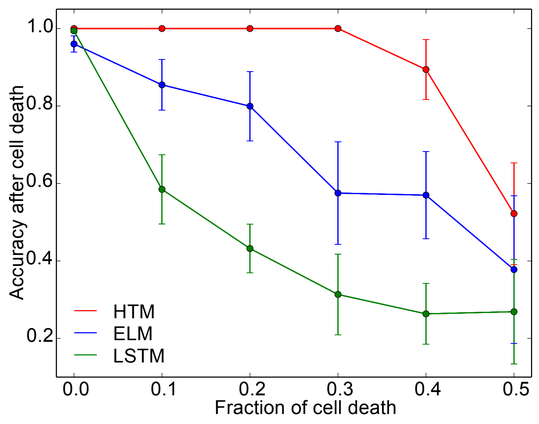
$$prob(false\ match) = \frac{\sum_{b=\theta}^{s} |\Omega(n,a,b,s)|}{\binom{n}{a}} \qquad |\Omega(n,a,b,s)| = \binom{s}{b} \times \binom{n-s}{a-b}$$

(Ahmad and Hawkins, 2016)





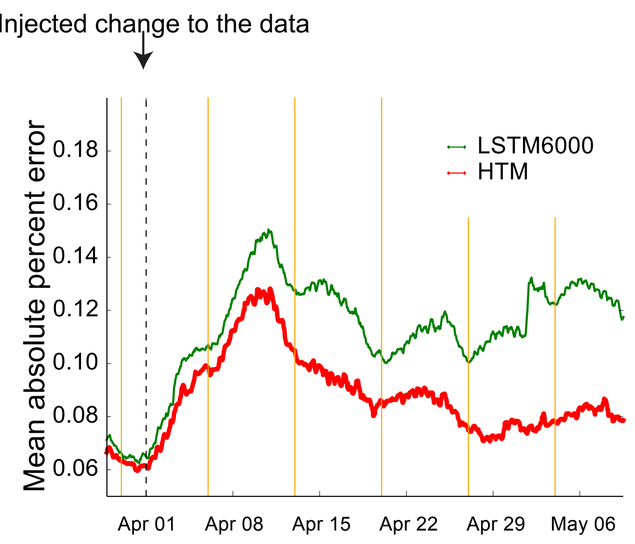
Highly sparse representations



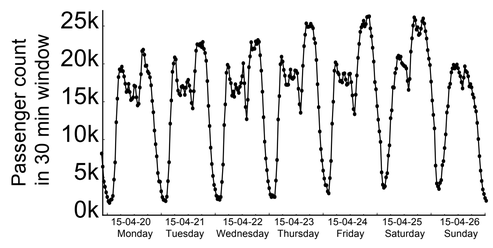
High tolerance to faults and noise

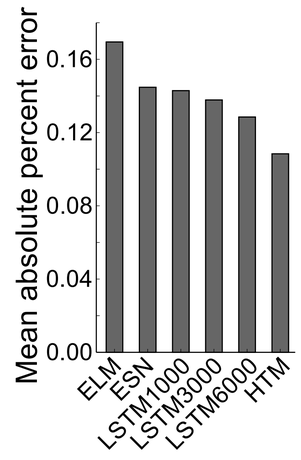
Properties of HTM sequence memory

Works well on real-world data:



HTM adapts quickly to changes due to ability to learn continuously

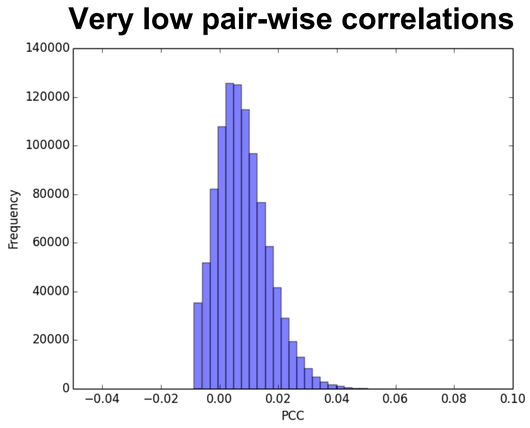




(Cui, Ahmad, & Hawkins 2016)

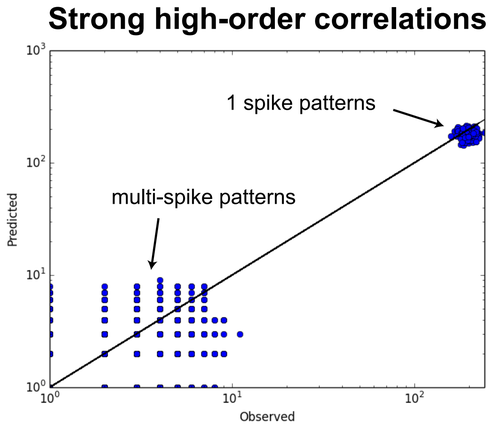
Correlation structure:

Very low pair-wise correlations



Consistent with experimental observations
(e.g., *Ecker et al., 2010*)

Strong high-order correlations



Observed frequencies of 10-bit patterns significantly different from frequencies predicted by poisson model.

Analogous to (*Schneidman et al, 2006*)