Hopfield networks

Models of content-addressable memory

Hopfield networks - a model of content-addressable memory

What is content-addressable memory?

One small step for man...

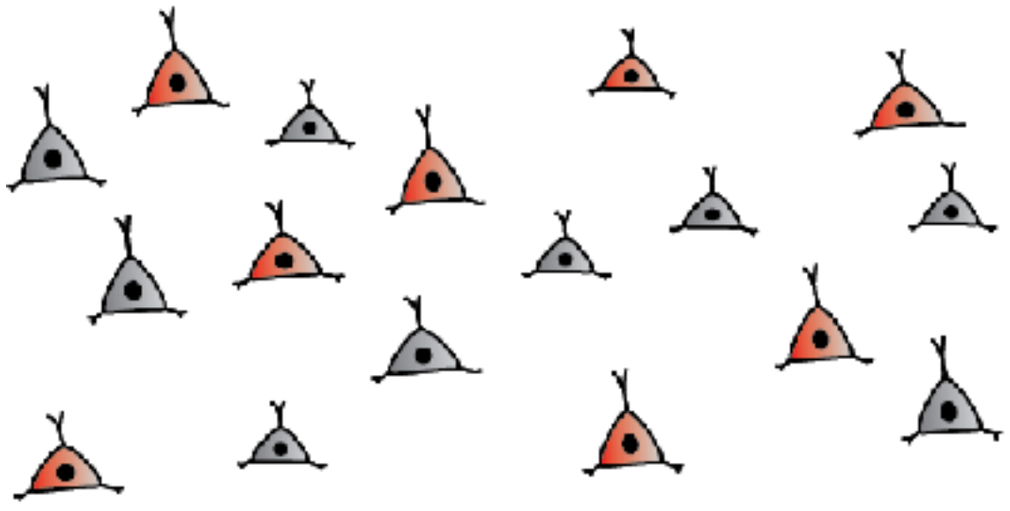
Can recover also when partial/distorted:

One tiny step for man…

(Computer: retrieve quote #12847)

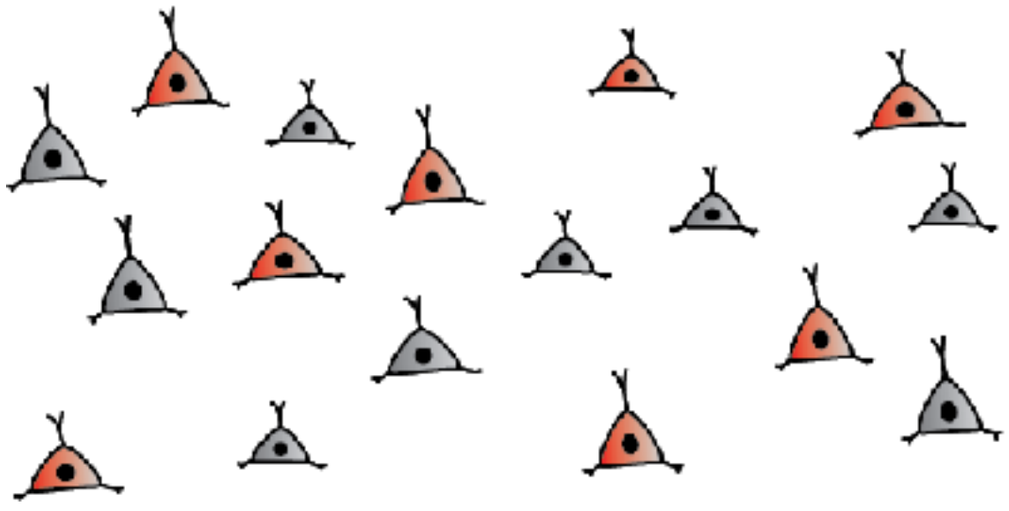
Ensemble dynamics view (cartoon)

One small step for man
one giant leap for mankind

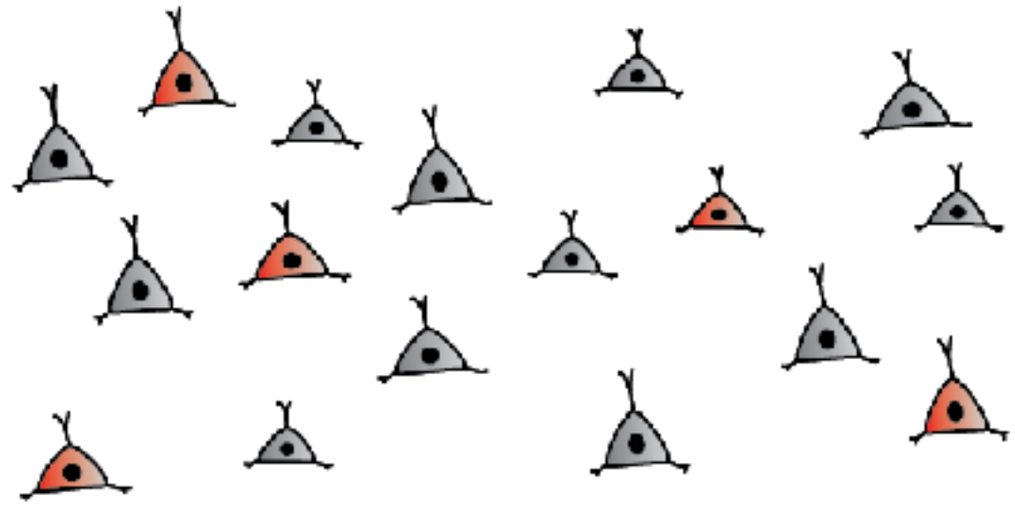


Ensemble dynamics view (cartoon)

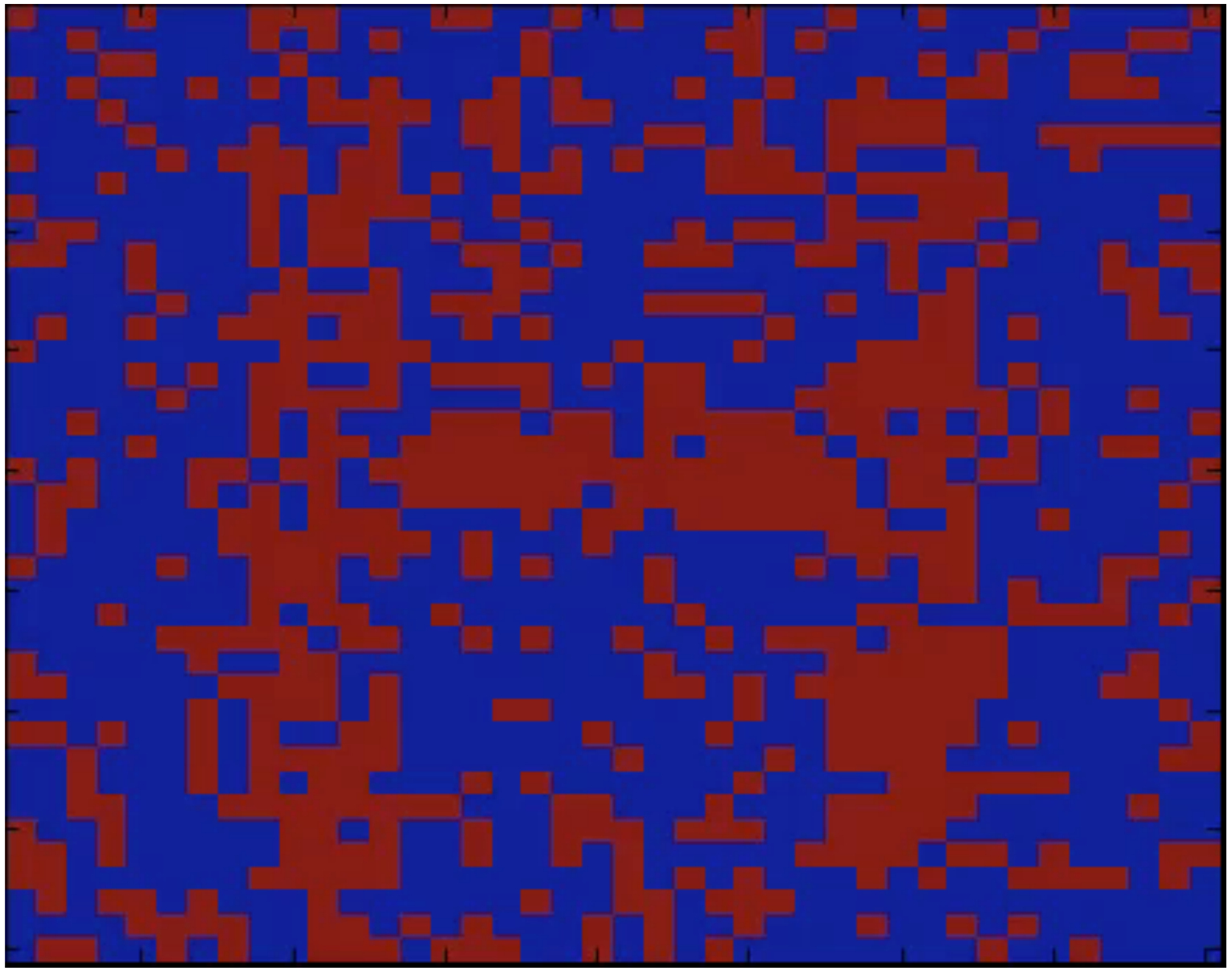
One small step for man
one giant leap for mankind



One small step for man...

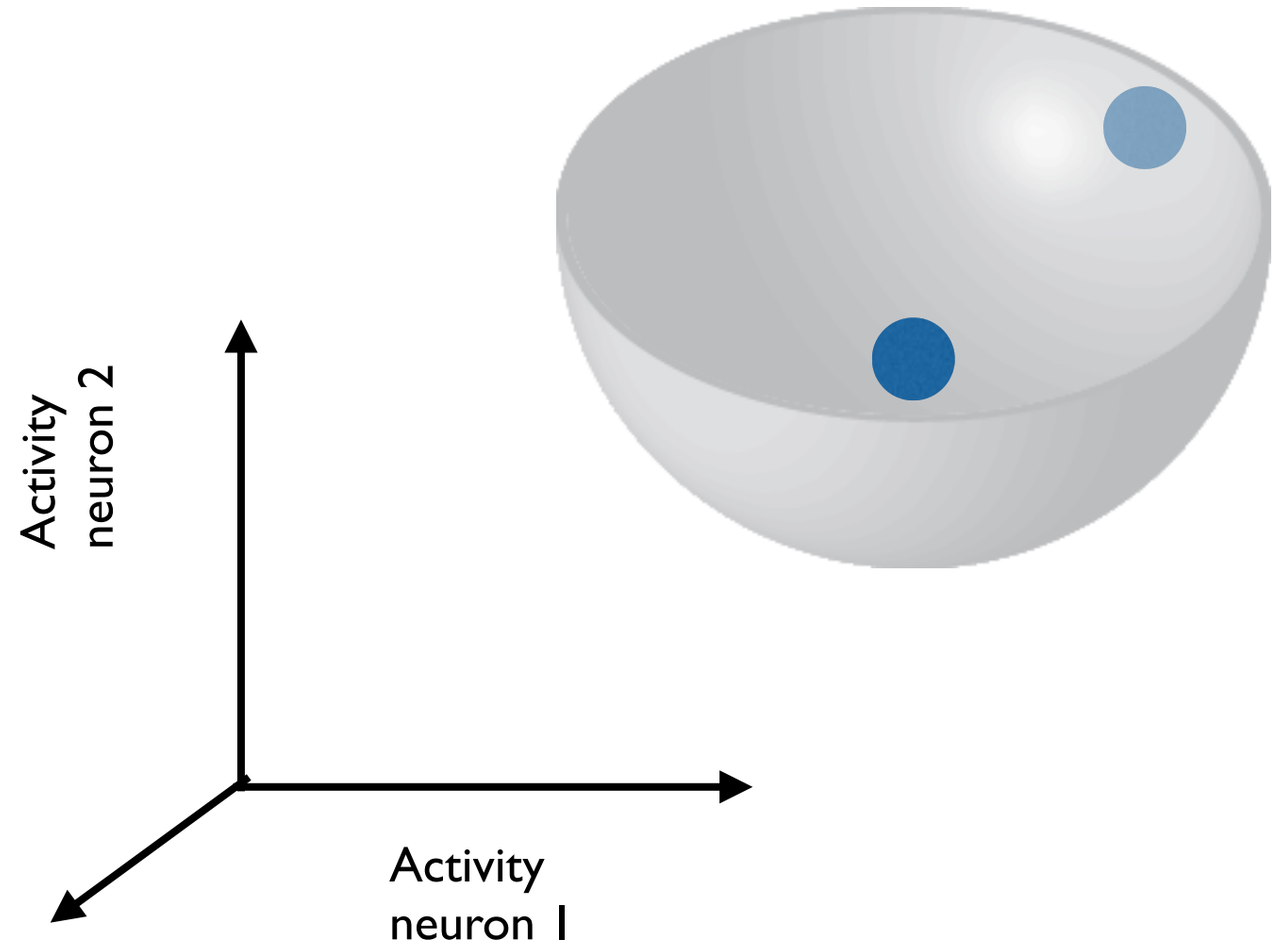


Hopfield network in action



Attractor dynamics

“Energy” landscape



What makes Hopfield networks interesting?

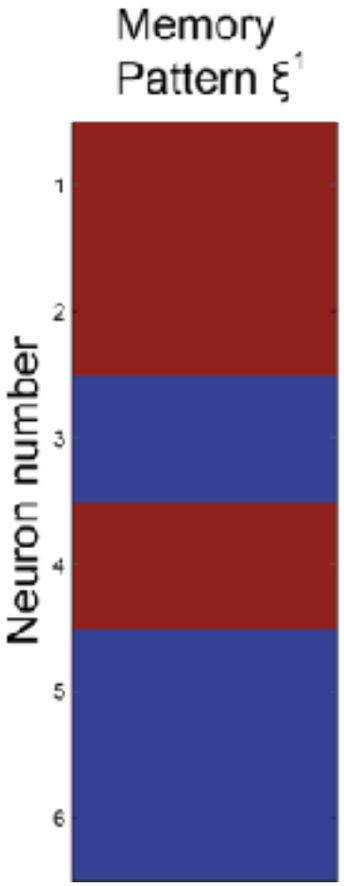
Implement content-addressable memory (synthesis of previous conceptual models of content-addressable memory)

One of the first (and almost only) concrete examples of computation with dynamical systems

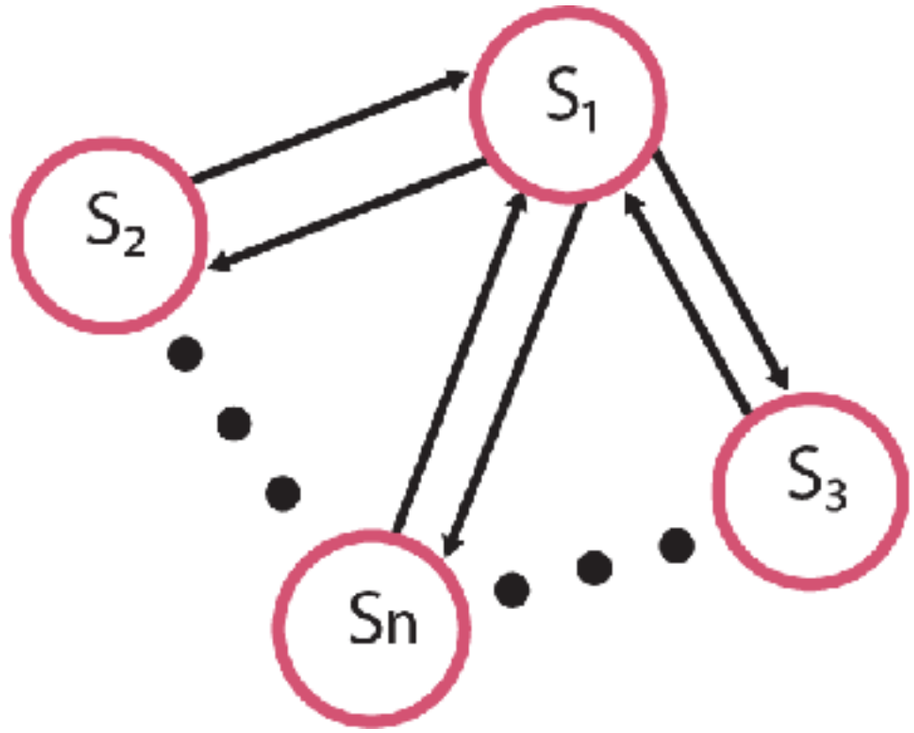
Provide a deep connection to spin-glass systems in physics

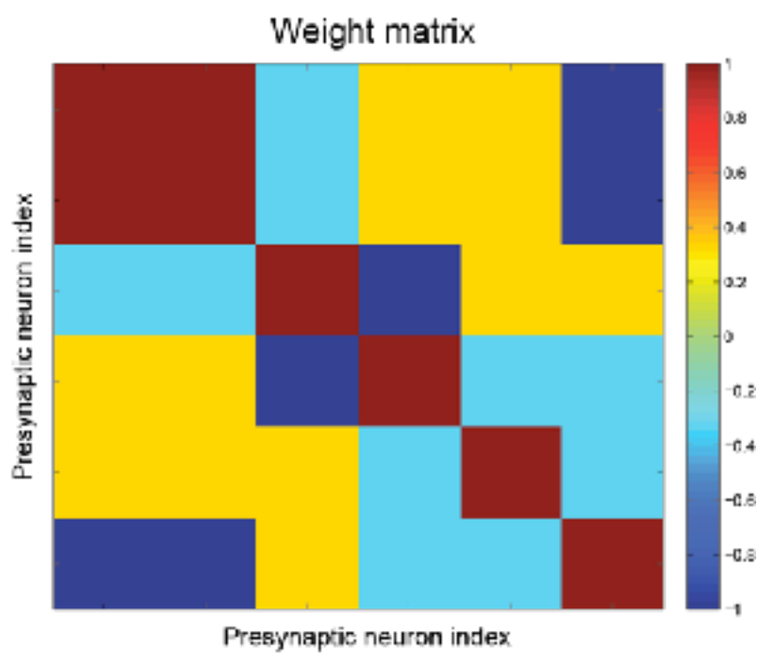
Hopfield network structure

Network is "Hebbian": memory ensembles fired together so they became wired together



Neuron	Activity
1	+1
2	+1
3	-1
4	+1
5	-1
6	-1





Hopfield network structure

Network is "Hebbian": memory ensembles fired together so they became wired together

Neuron	Activity of memory pattern #1
1	+1
2	+1
3	-1
4	+1
5	-1
6	-1

	Pre-synaptic neuron					
Post synaptic neuron						

Hopfield network structure

Network is "Hebbian": memory ensembles fired together so they became wired together

Neuron	Activity of memory pattern #1
1	+1
2	+1
3	-1
4	+1
5	-1
6	-1

		Pre-synaptic neuron				
Post synaptic neuron						
	+1					

Hopfield network structure

Network is "Hebbian": memory ensembles fired together so they became wired together

Neuron	Activity of memory pattern #1
1	+1
2	+1
3	-1
4	+1
5	-1
6	-1

		Pre-synaptic neuron				
Post synaptic neuron	+1					
	-1					

Hopfield network structure

Network is "Hebbian": memory ensembles fired together so they became wired together

Neuron	Activity of memory pattern #1
1	+1
2	+1
3	-1
4	+1
5	-1
6	-1

		Pre-synaptic neuron				
Post synaptic neuron						
	+1					
	-1					
	+1					
	-1					
	-1					

Hopfield network structure

Network is "Hebbian": memory ensembles fired together so they became wired together

Neuron	Activity of memory pattern #1
1	+1
2	+1
3	-1
4	+1
5	-1
6	-1

		Pre-synaptic neuron				
		+1	-1	+1	-1	-1
Post synaptic neuron	+1		-1	+1	-1	-1
	-1	-1		-1	+1	+1
	+1	+1	-1		-1	-1
	-1	-1	+1	-1		+1
	-1	-1	+1	-1	+1	

Hopfield network structure

Network is "Hebbian": memory ensembles fired together so they became wired together

$$J_{ij} = \xi_i^1 \xi_j^1$$

Neuron	Activity of memory pattern #1
1	+1
2	+1
3	-1
4	+1
5	-1
6	-1

		Pre-synaptic neuron				
		+1	-1	+1	-1	-1
Post synaptic neuron	+1		-1	+1	-1	-1
	-1	-1		-1	+1	+1
	+1	+1	-1		-1	-1
	-1	-1	+1	-1		+1
	-1	-1	+1	-1	+1	

Hopfield network structure

Network is "Hebbian": memory ensembles fired together so they became wired together

$$J_{ij} = \xi_i^1 \xi_j^1$$

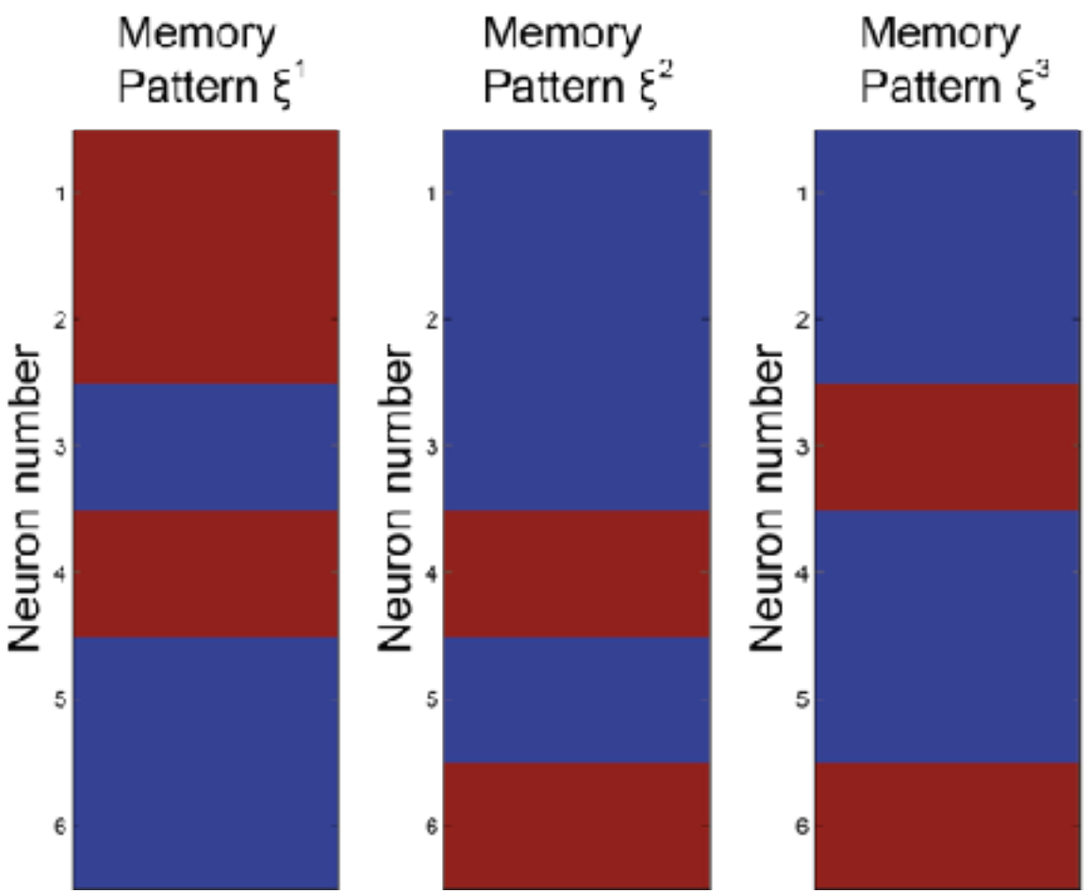
Neuron	Activity of memory pattern #1
1	+1
2	+1
3	-1
4	+1
5	-1
6	-1

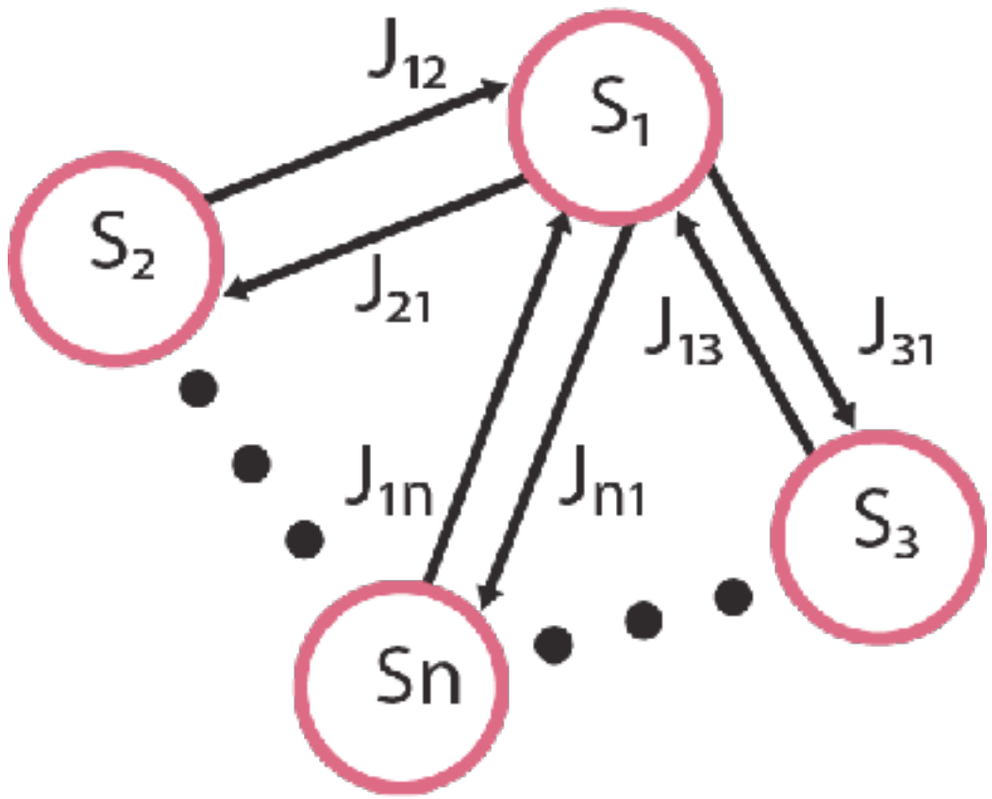
		Pre-synaptic neuron					
		+1	+1	-1	+1	-1	-1
Post synaptic neuron	+1	+1	-1	+1	-1	-1
	-1	-1	+1	-1	+1	+1
	+1	+1	-1	+1	-1	-1
	-1	-1	+1	-1	+1	+1
	-1	-1	+1	-1	+1	+1

Hopfield network structure

Network is "Hebbian": weights are a sum over correlation between neuron activation in the memory patterns

Synaptic weights defined as: $$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$$





Hopfield network structure

Network is "Hebbian": weights are a sum over correlation between neuron activation in the memory patterns

Synaptic weights defined as: $$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$$

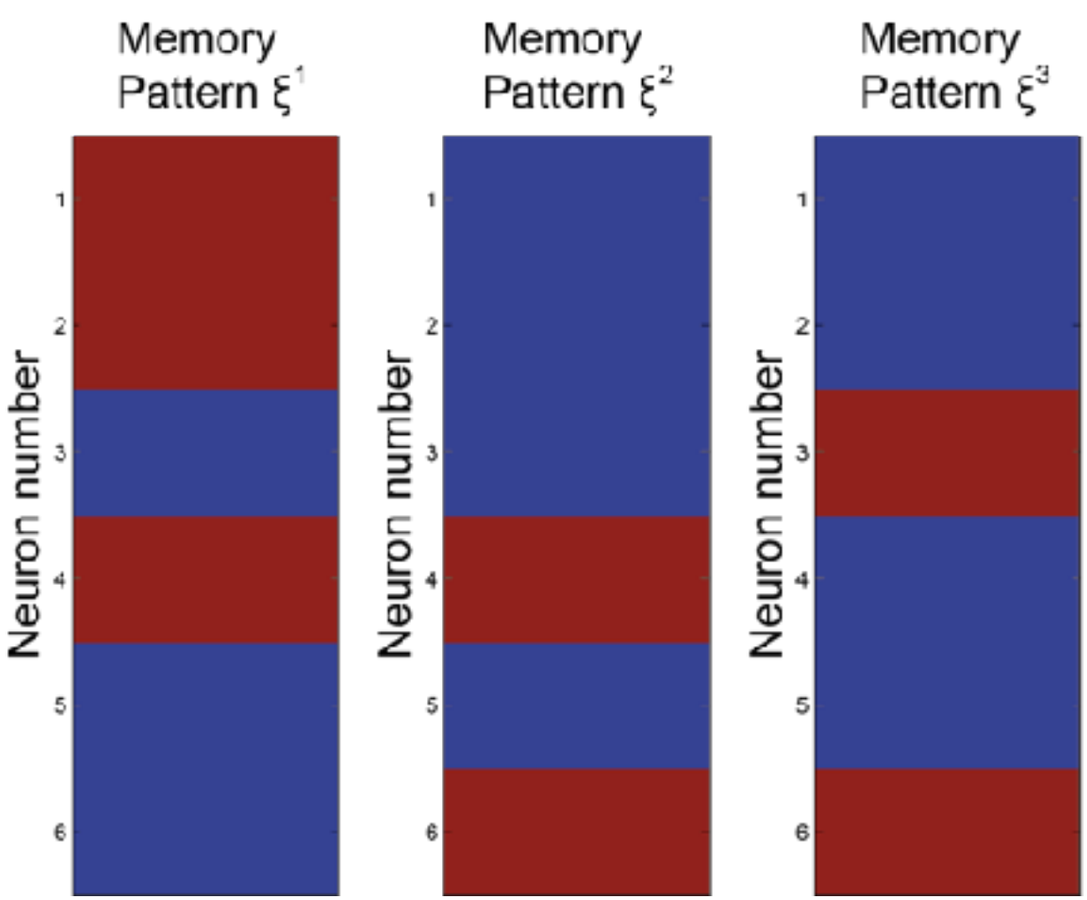
Neuron	Activity of memory pattern #2
1	-1
2	-1
3	-1
4	+1
5	-1
6	+1

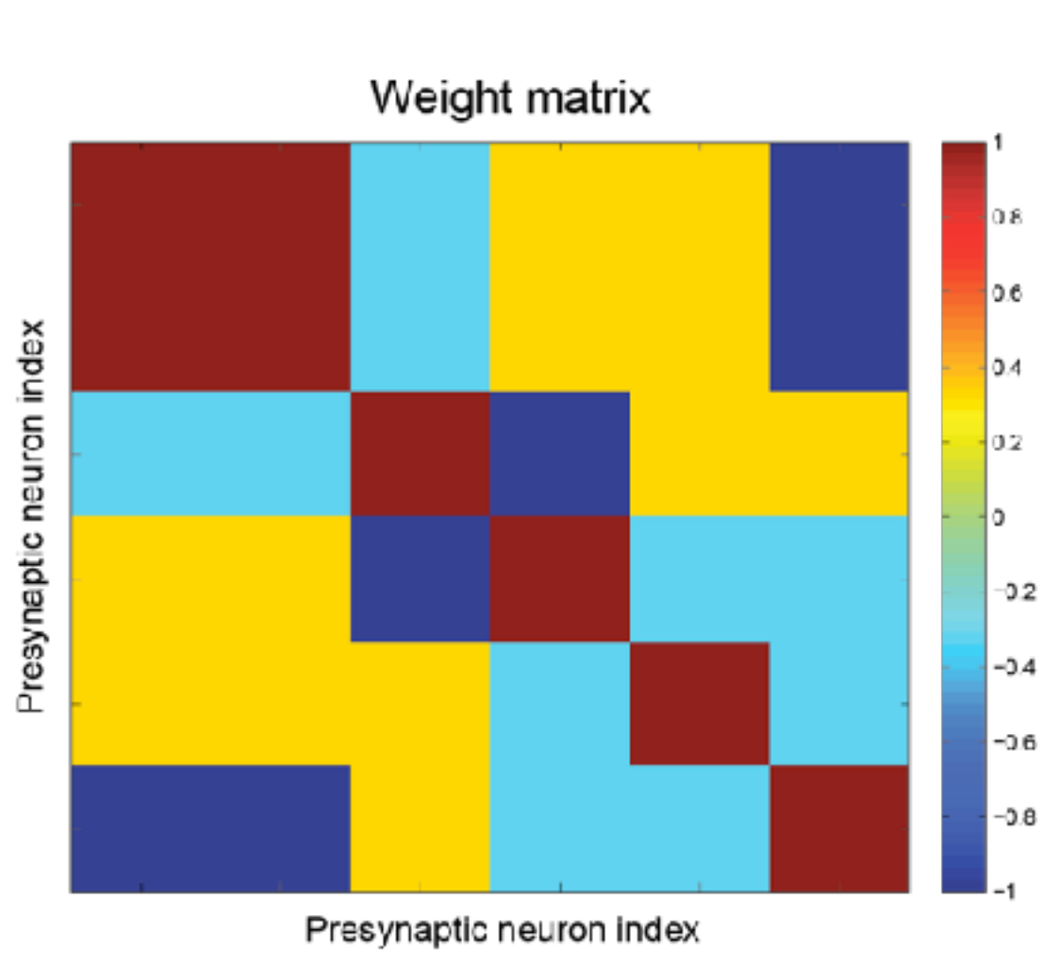
	Pre-synaptic neuron					
Post synaptic neuron	+1	+1	+1	-1	+1	-1
	+1	+1	+1	-1	+1	-1
	+1	+1	+1	-1	+1	-1
	-1	-1	-1	+1	-1	+1
	+1	+1	+1	-1	+1	-1
	-1	-1	-1	+1	-1	+1

Hopfield network structure

Network is “Hebbian”: weights are a sum over correlation between neuron activation in the memory patterns

Synaptic weights defined as: $$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$$

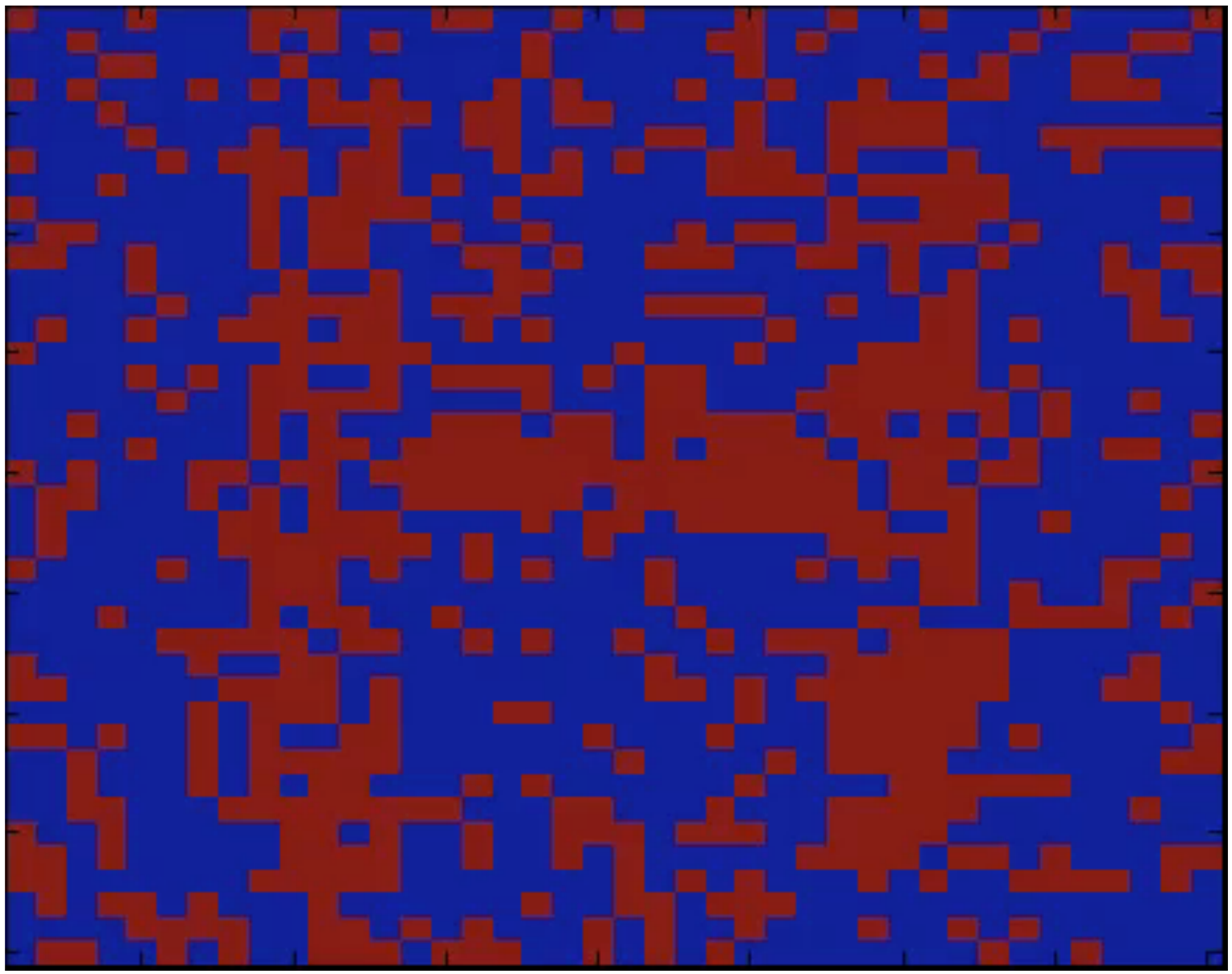




Hopfield network activity

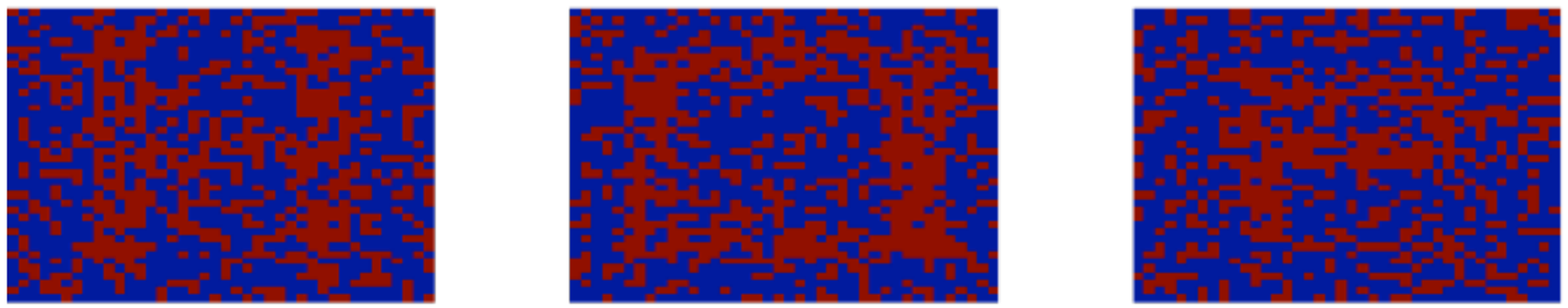
Binary neurons (active if input greater than threshold)

Asynchronous dynamics - one neuron updated at a time



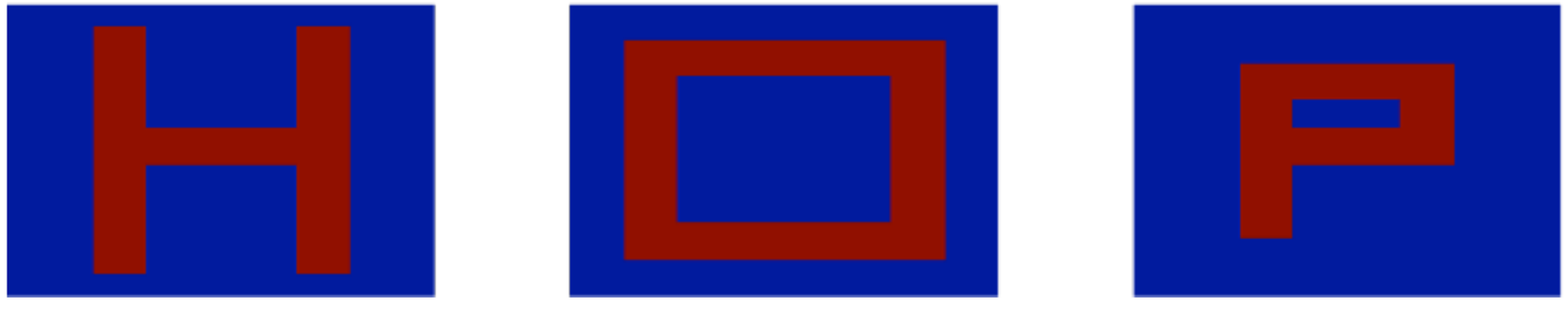
Hopfield network activity

Initial noisy pattern



Midway in dynamics

Final noise-free state

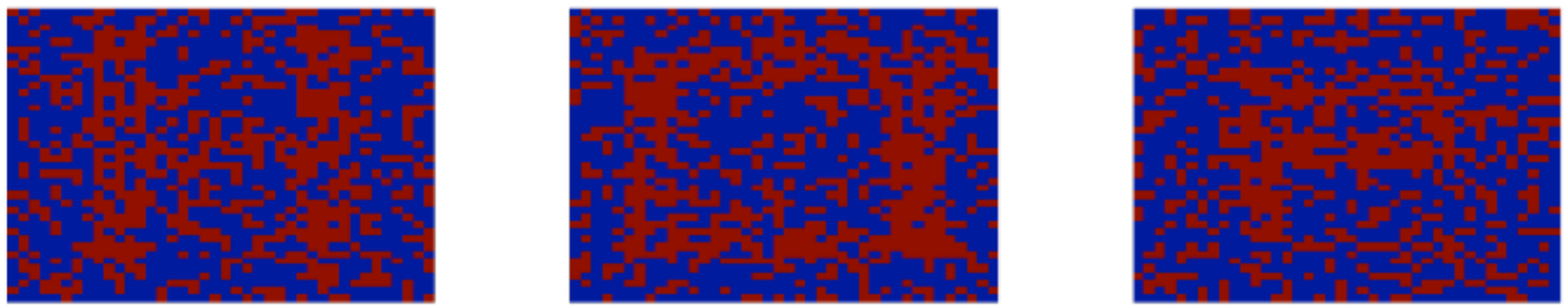


Why does the Hopfield network work?

Go to whiteboard

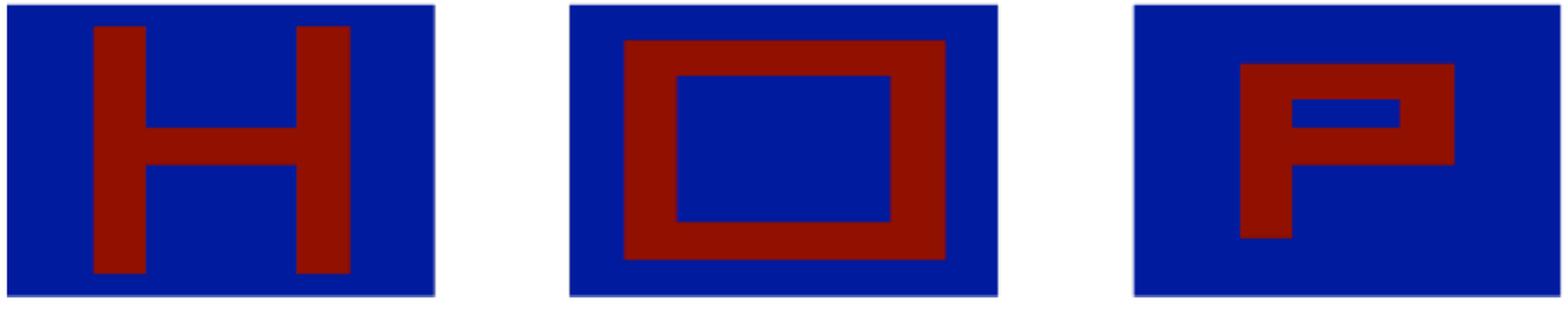
Hopfield network activity

Initial noisy pattern



Midway in dynamics

Final noise-free state



Assumptions?

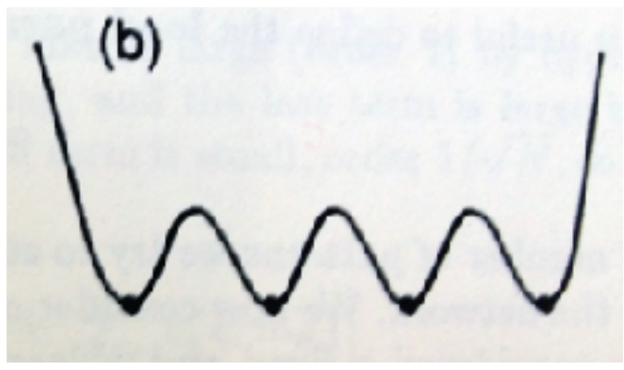
Are binary neurons a valid approximation?

Yes. Results similar with continuous units (Hopfield 1984, PNAS)

Additional questions with the Hopfield model

Except for stability of memory patterns what can we say about dynamics?

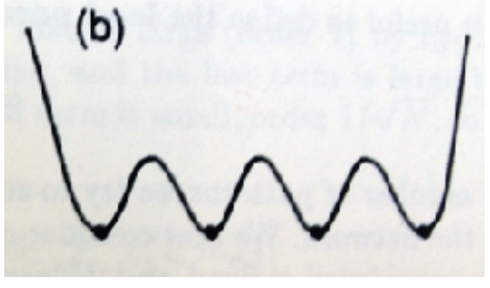
- Dynamics result in decrease in energy function, of which memory patterns are minima

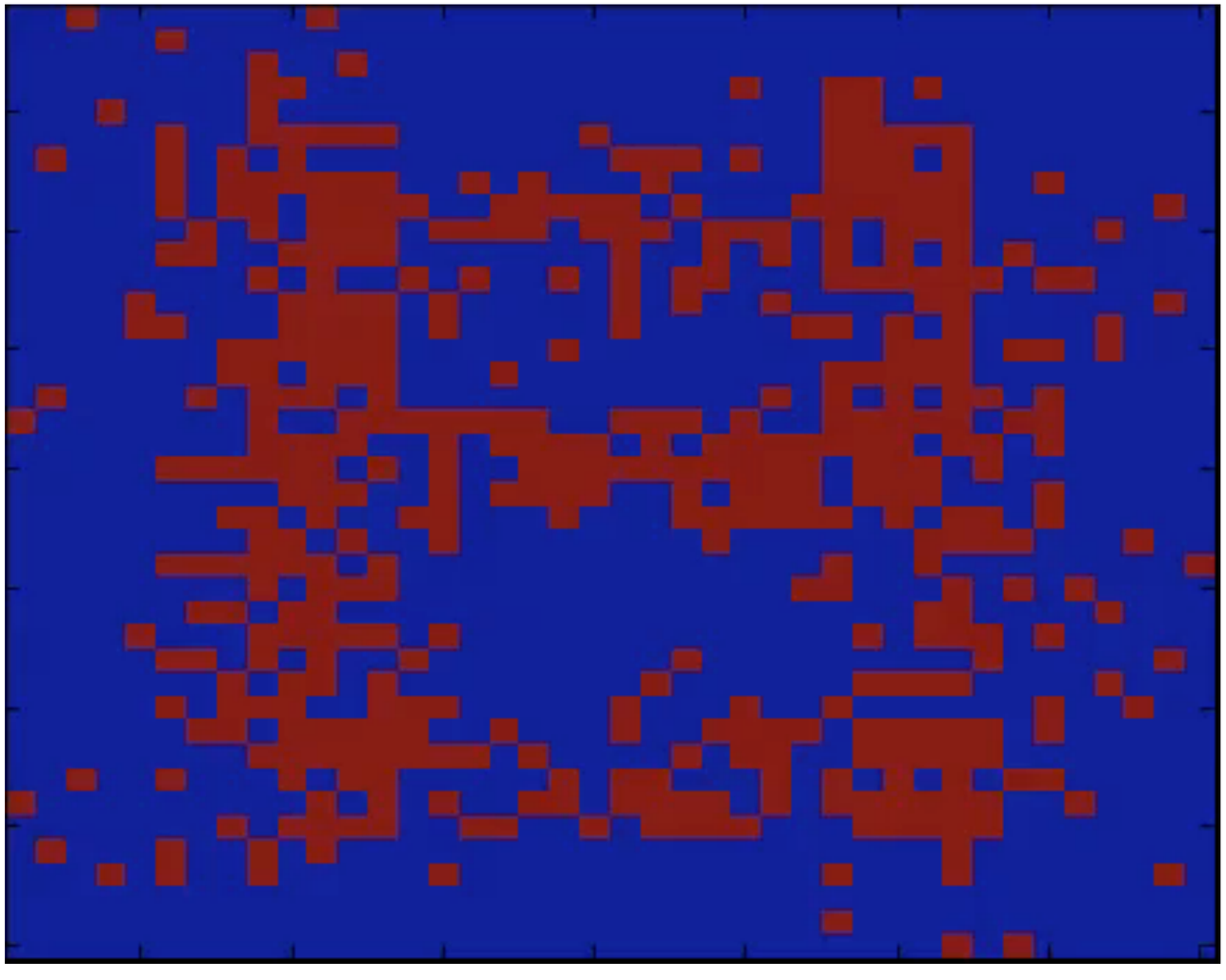


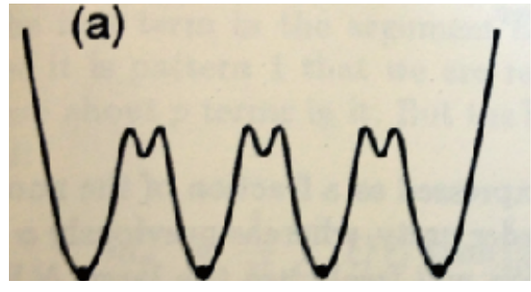
Problems with the Hopfield model

Memory states are stable, but are they the only stable states?

No! Problem with spurious states!



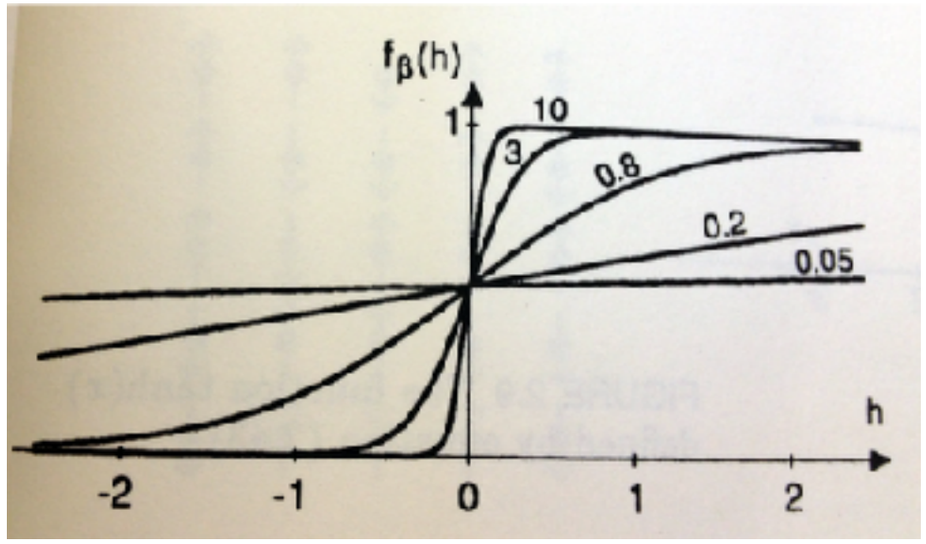




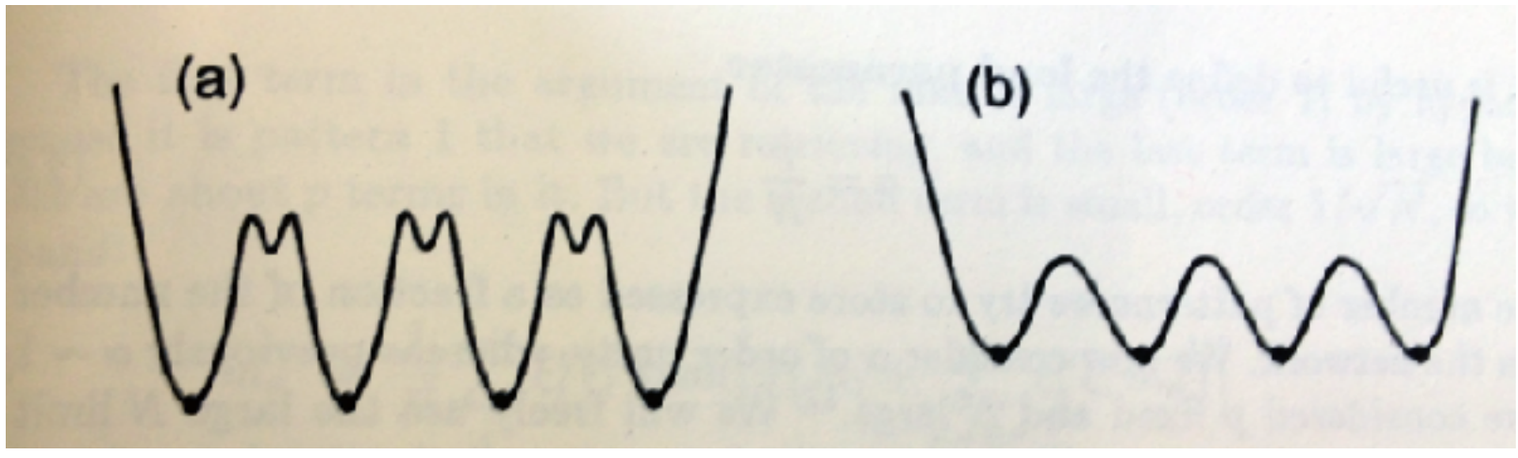
Noise wipes out spurious states

Amit, Gutfreund, Sompolinsky (1985 a,b)

Neurons spike probabilistically, proportional to input:



Noise "smoothes out" spurious states:



Beyond the Hopfield model

Can the dynamics be solved in detail?

Yes! See Amit, Gutfreund, Sompolinsky

Are the Hopfield weights the only weights that can support memory?

No! One can write conditions in terms of perceptron learning

For more details see Intro to Theoretical Neuroscience book (for physics see Amit book)

Summary

Hopfield model implements content-addressable memory

One of the first examples of computation with dynamical systems

One of the first examples of distributed computation

Opened up the world of "attractor neural networks"