Simple Models of your Brain

Modeling Intelligence with Recurrent Neural Networks

Enrico Ventura

PhD Seminars – December 6th 2023
Università degli Studi "La Sapienza"

Sapienza Università di Roma

ENS | PSL

Outline of the talk

1. The simple Brain picture.
2. The emergence of Memory.
3. The emergence of “Creativity”.
4. Conclusions.

Modeling the Brain with Recurrent Neural Networks

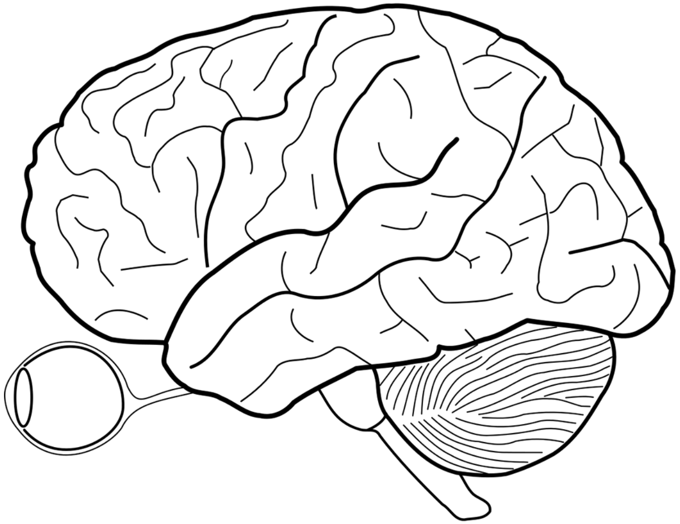
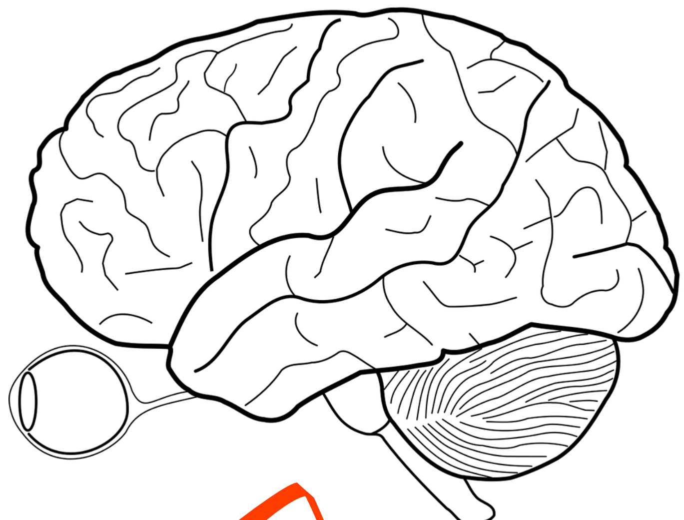


Image from Wikipedia.

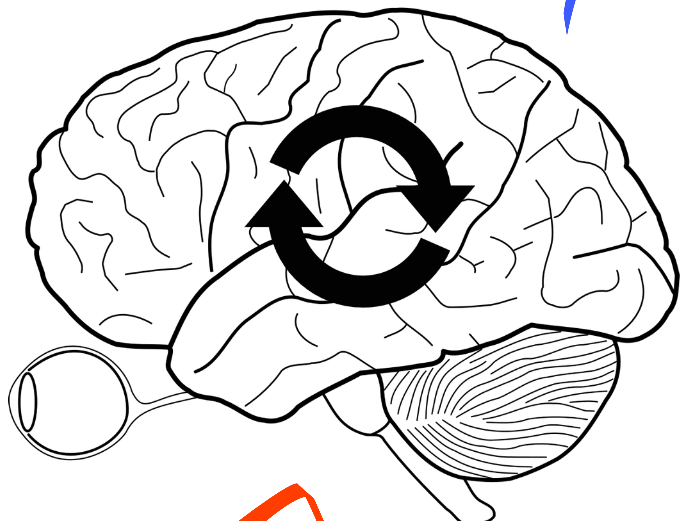
Modeling the Brain with Recurrent Neural Networks



External Stimuli

Images from Web.

Modeling the Brain with Recurrent Neural Networks

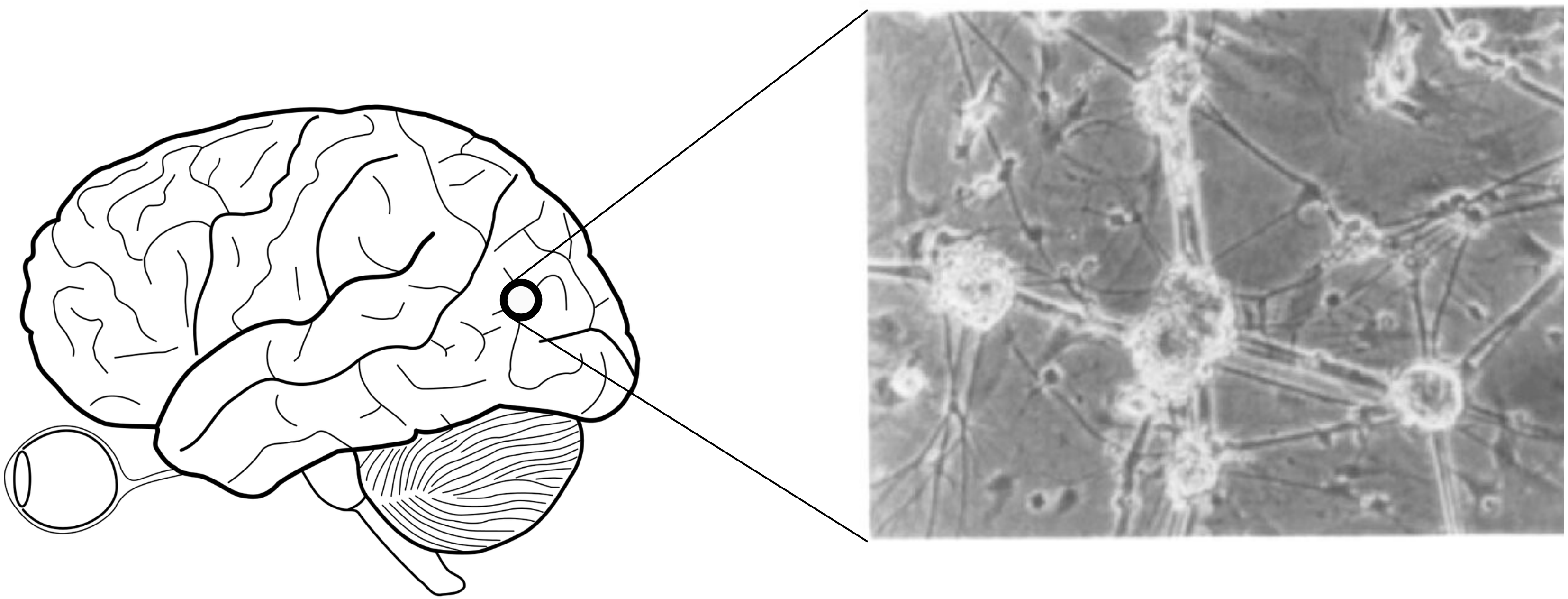


Decisions/
Tasks

External
Stimuli

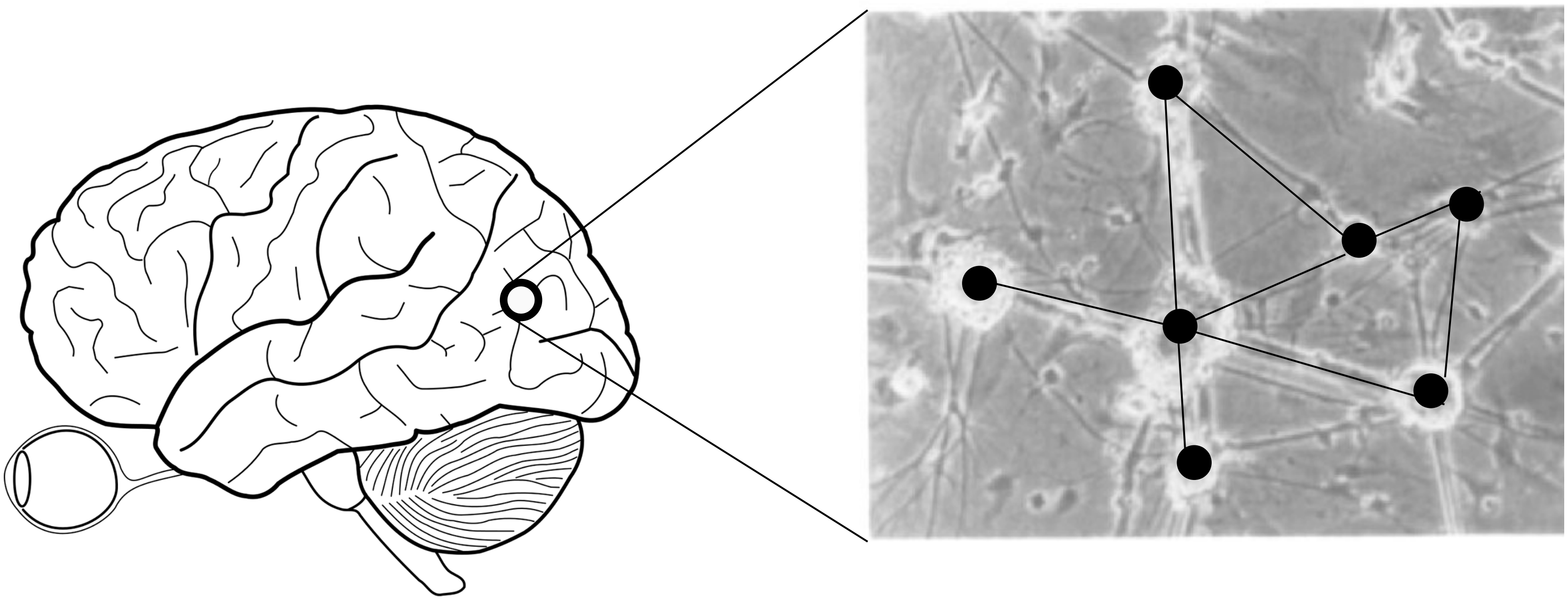
Images from Web.

Modeling the Brain with Recurrent Neural Networks

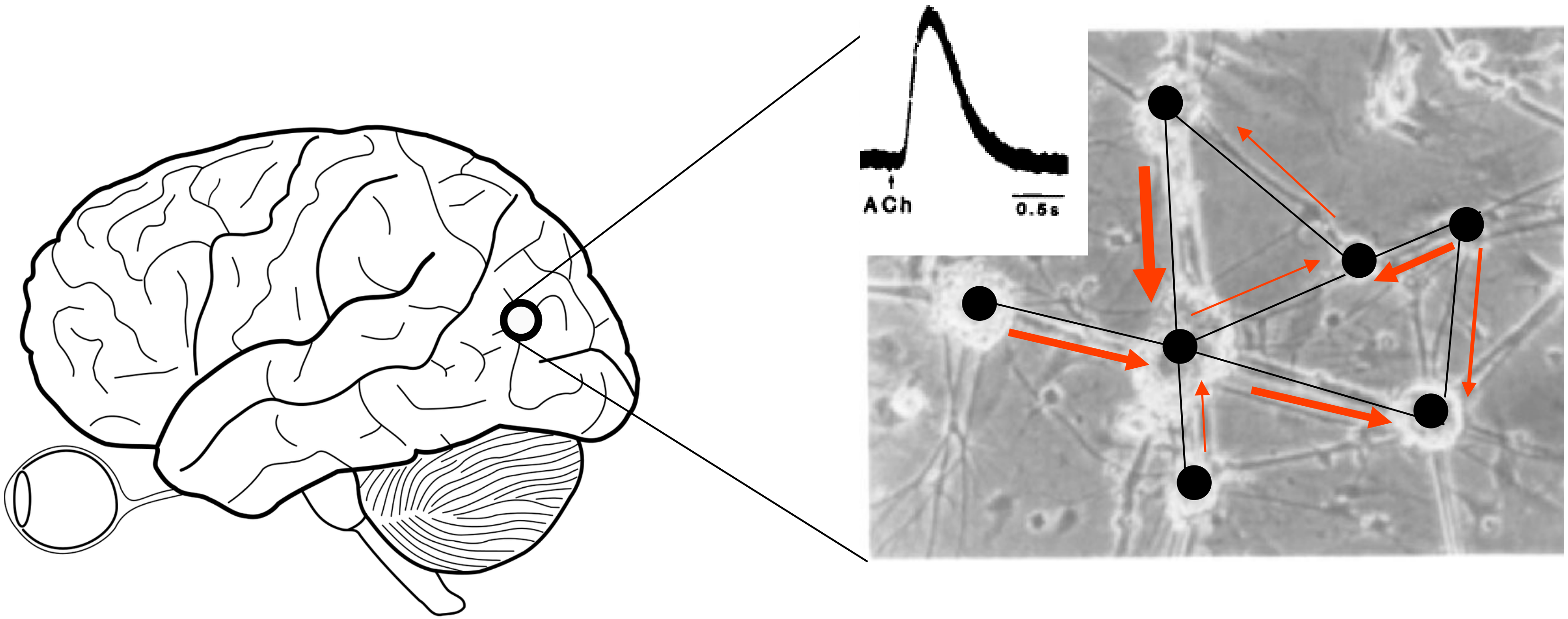


Rosen et al., 1992.

Modeling the Brain with Recurrent Neural Networks

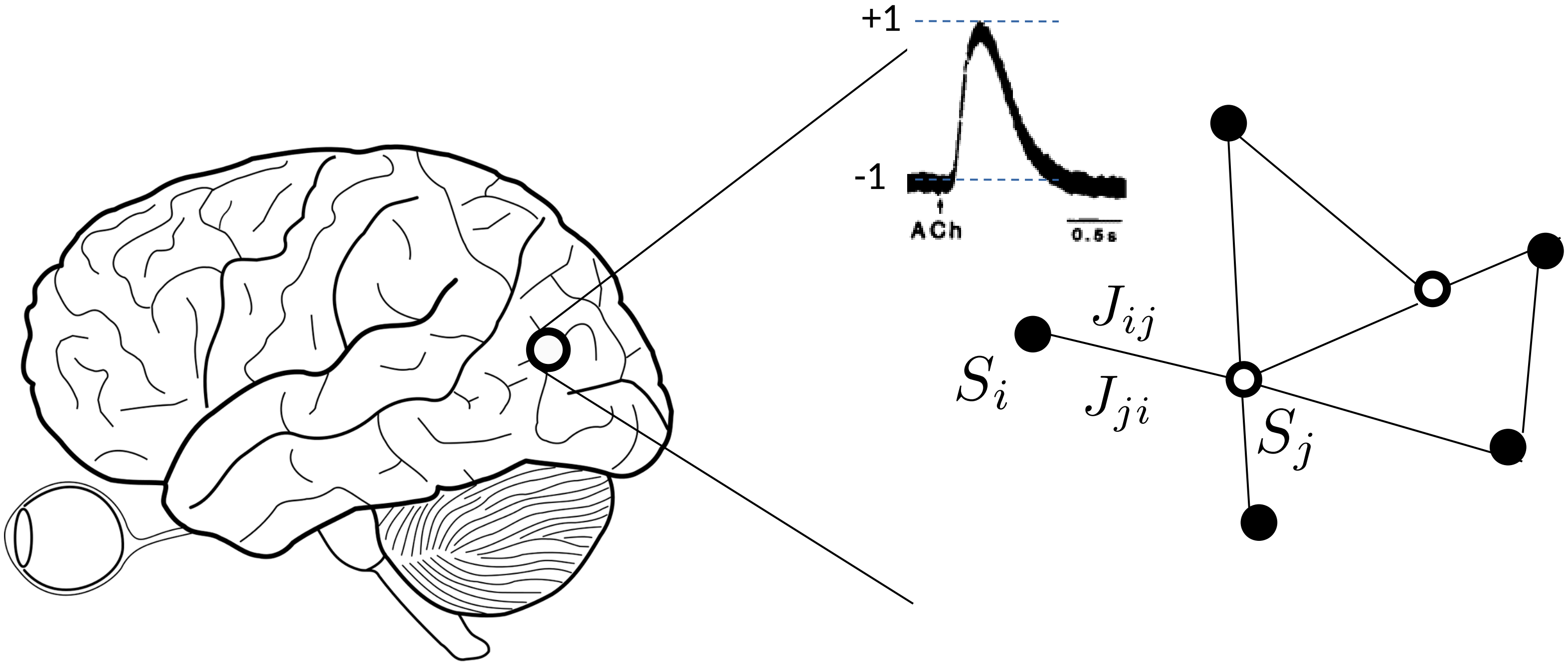


Modeling the Brain with Recurrent Neural Networks



Wood, 1993.

Modeling the Brain with Recurrent Neural Networks



Complex Behaviour

$$\begin{cases} \vec{S} = \{\pm 1\}^N & \text{neural activity} \\ \boldsymbol{J} \in \mathbb{R}^{N \times N} & \text{disorder/frustration} \end{cases}$$

Modeling the Brain with Recurrent Neural Networks

Learning = the process that shapes the parameters J depending on:

- a set of external stimuli (or *training data*).
- a specific task to be performed (which is linked to a specific *dynamics*).

→ Modeled by **Learning Algorithms.**

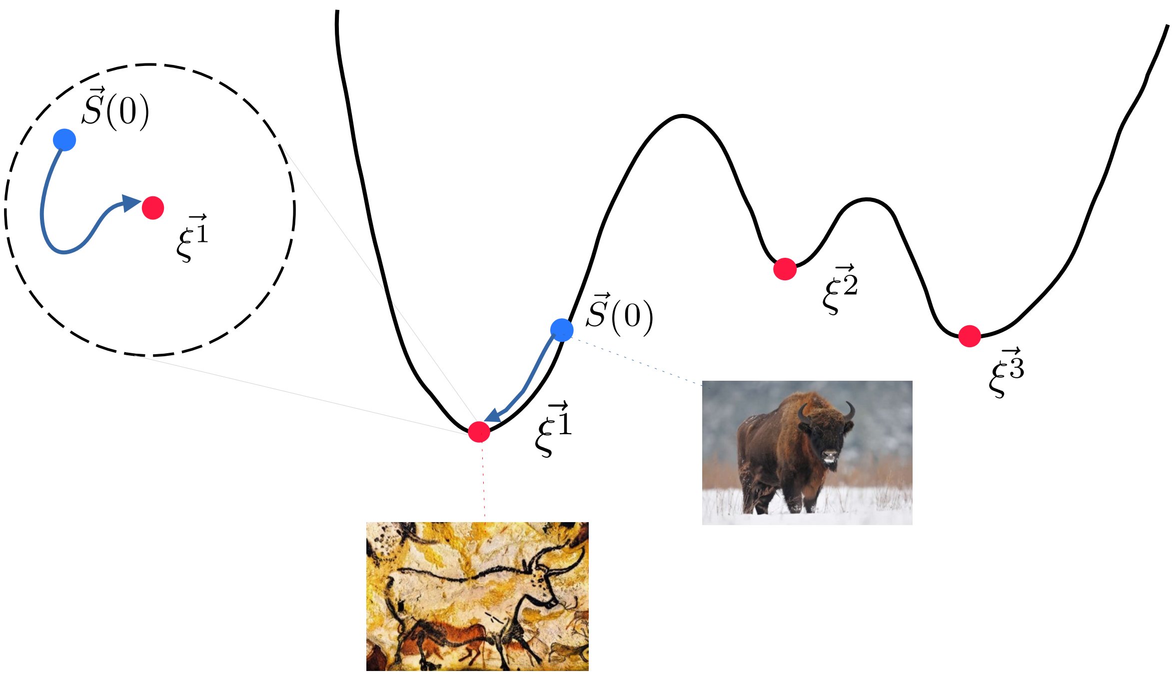
Modeling the Brain with Recurrent Neural Networks

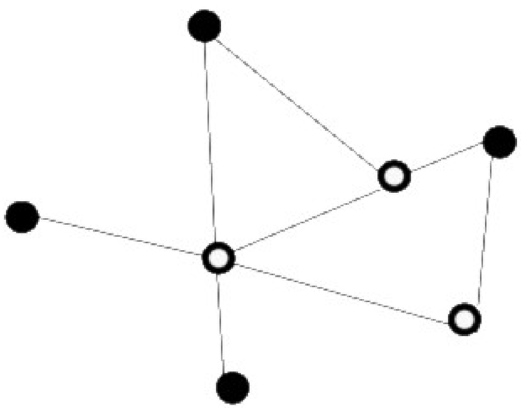
$$E[\vec{S}|\boldsymbol{J}] = -\sum_{i,j>i} J_{ij} S_i S_j$$

Energy

$$T = 0$$

Associative Memory





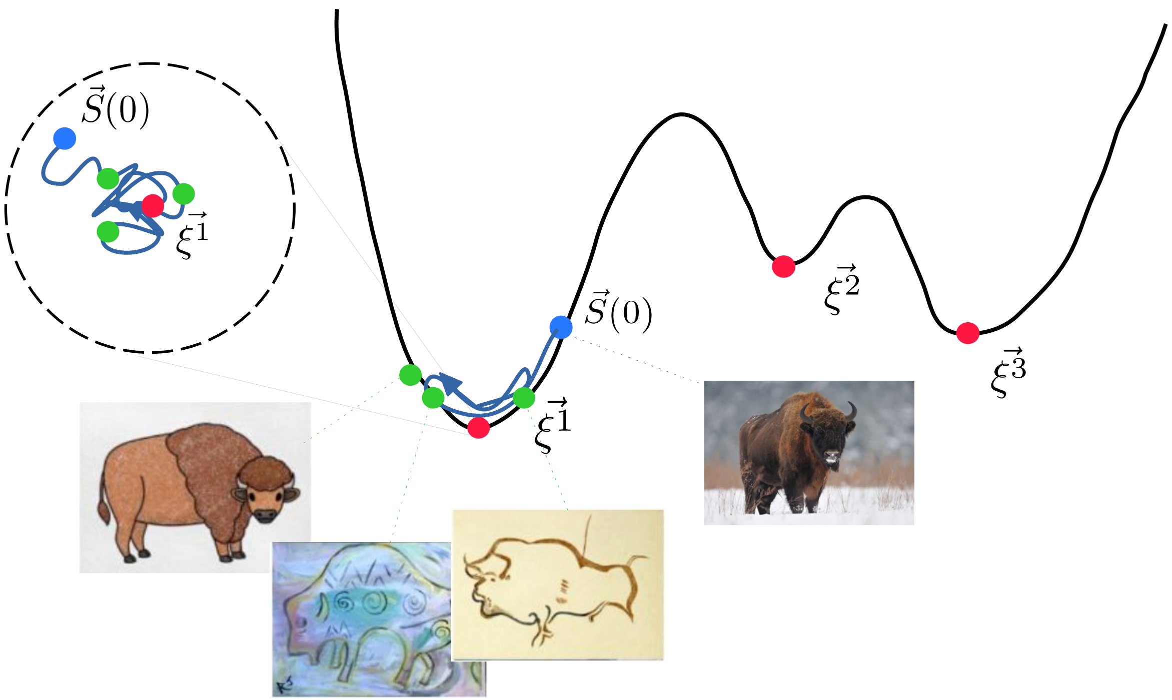
Modeling the Brain with Recurrent Neural Networks

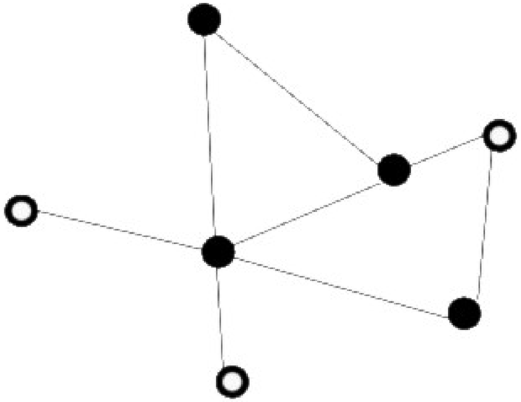
$$E[\vec{S}|\boldsymbol{J}] = -\sum_{i,j>i} J_{ij} S_i S_j$$

Energy

$$T > 0$$

Generative Task

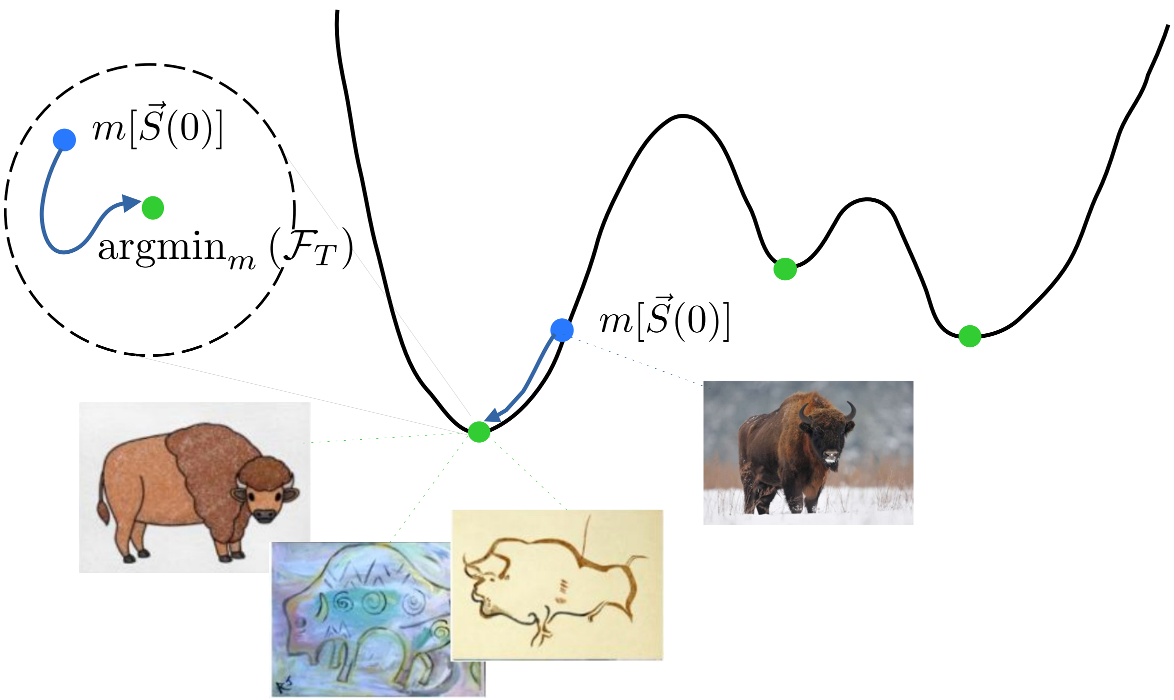




Modeling the Brain with Recurrent Neural Networks

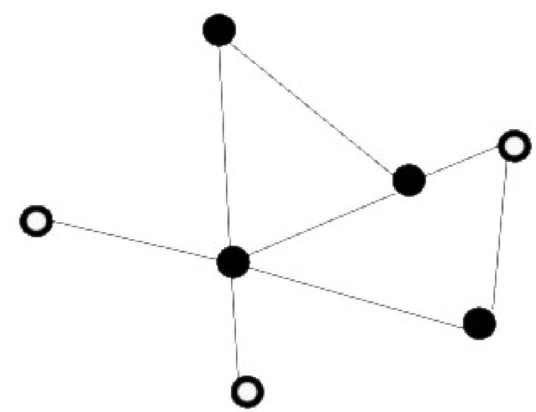
$$\mathcal{F}_T(m, \boldsymbol{J}) = E(m, \boldsymbol{J}) - T\mathcal{S}(m, \boldsymbol{J})$$ **Free Energy**

where $m = m[\vec{S}]$

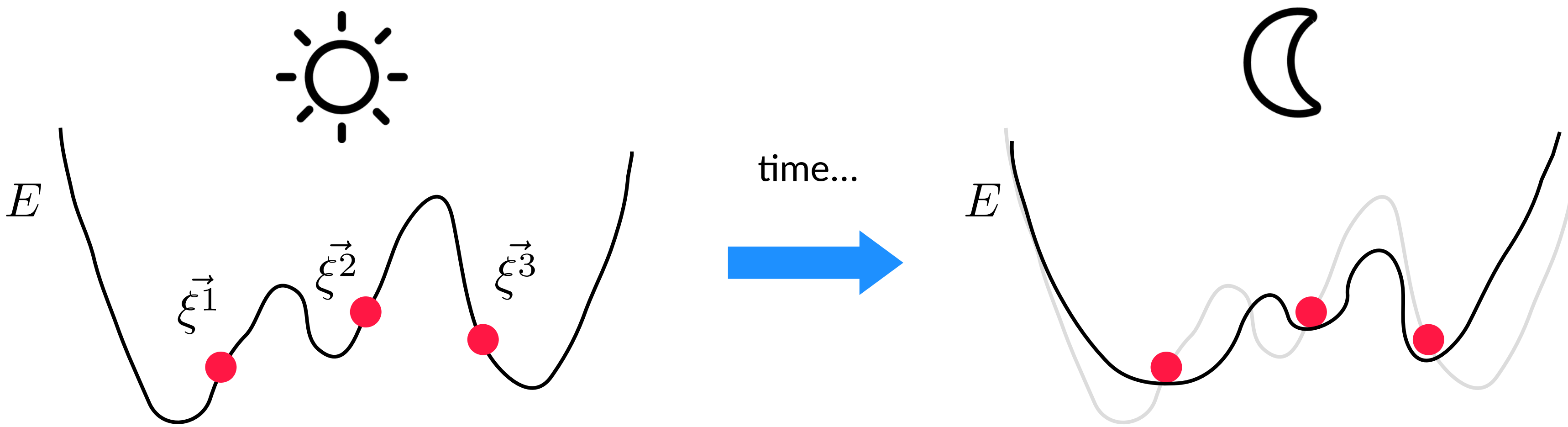


$T > 0$

Generative Task



Associative Memory: learning algorithms



A set of $p = \mathcal{O}(N)$ data $\{\vec{\xi^\mu}\}_{\mu=1}^{p}$ is learned as:

Hebb's Rule

$$J_{ij}^{(t=0)} = \langle S_i S_j \rangle_{data} = \frac{1}{p} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$$

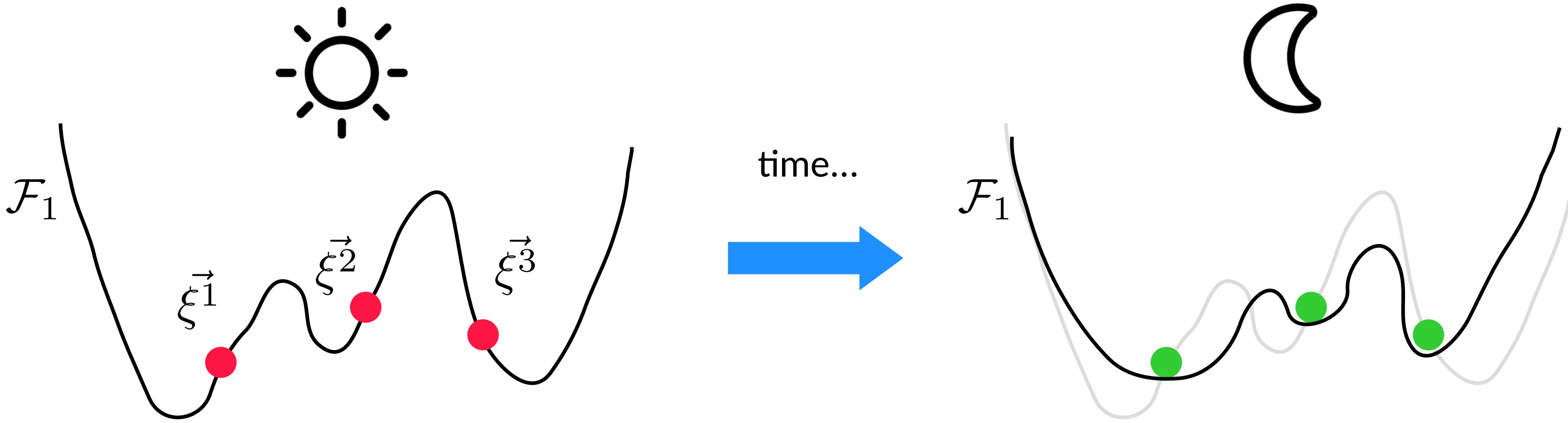
Data are consolidated as:

Unlearning Rule (Hopfield)

$$J_{ij}^{(t)} = J_{ij}^{(t-1)} - \lambda \langle S_i S_j \rangle_{T=0}^{(t)}$$

Hopfield, 1982-1983; Amit et al. 1985; Benedetti et al., 2022.

Generative Modeling: learning algorithms



A set of $p \gg N$ data $\{\vec{\xi^\mu}\}_{\mu=1}^{p}$ is learned as:

Hebb's Rule

$$J_{ij}^{(t=0)} = \langle S_i S_j \rangle_{data} = \frac{1}{p}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$$

Data are consolidated as:

Boltzmann Machine (Hinton & Sejnowski)

$$J_{ij}^{(t)} = J_{ij}^{(t-1)} - \lambda \langle S_i S_j \rangle_{T=1}^{(t)}$$

Hinton et al. 1985; Ventura et al., 2023.

Conclusions

- Both **associative memory** and the **generative task** emerge from the same type of algorithm, performed at *different temperatures*.

$$\langle S_i S_j \rangle_T \longrightarrow \langle S_i S_j \rangle_{data}$$

$$P_T[\vec{S}] = \frac{1}{Z_T} \exp\left(-E[\vec{S}]/T\right) \longrightarrow P_{data}[\vec{S}]$$

Moment Matching condition

- The structure of the algorithm suggests **daily-experience** followed by **night-sleep** might play the roles of *data-acquisition* and *consolidation* (Crick and Mitchison, 1983; Girardeau et al., 2020; Hoel, 2021; Hinton, 2023).

- Our brain is **not simple**: not symmetric, complex connectivity, different types of neurons etc.

Conclusions

- Both **associative memory** and the **generative task** emerge from the same type of algorithm, performed at *different temperatures*.

$$\langle S_i S_j \rangle_T \longrightarrow \langle S_i S_j \rangle_{data}$$

$$P_T[\vec{S}] = \frac{1}{Z_T} \exp\left(-E[\vec{S}]/T\right) \longrightarrow P_{data}[\vec{S}]$$

- The structure of the algorithm suggests **daily-experience** followed by **night-sleep** might play the roles of *data-acquisition* and *consolidation* (Crick and Mitchison, 1983; Girardeau et al., 2020; Hoel, 2021; Hinton, 2023).

- Our brain is **not simple**: not symmetric, complex connectivity, different types of neurons etc.

Thank you!