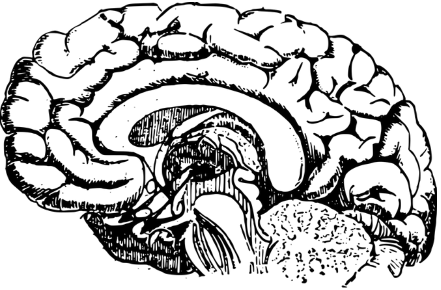
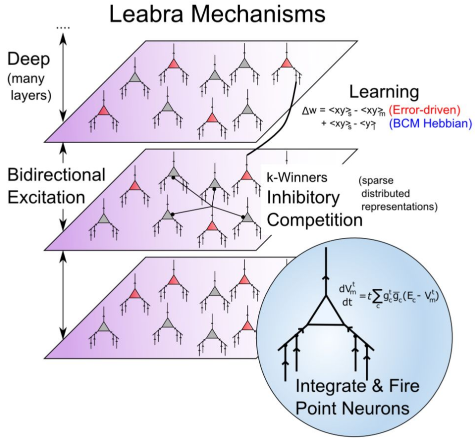
Biologically Based Models of Cortical Cognition



Modeling in Neuroscience
- Project Presentation -



Clément Bonnet,
Supervised by Jean-Pierre Nadal

Contents

- Introduction
- 6 principles for biologically based models of cortical cognition
- Interactions between principles
- LEABRA algorithm
- Experiments
- Conclusion

Introduction

- Number of principles developed for neural-network models of cognition
- Machine learning ideas taken from cognitive neuroscience
- Article provides a set of six principles integrated into a coherent overall framework

→ Presentation of the paper, the overall framework and some experiments

Principles for Biologically Based Models of Cortical Cognition

1) Biological Realism
    - Models should be constrained and informed by biological properties of the cortex
    - Computational mechanisms that violate biological properties should not be relied upon

Principles for Biologically Based Models of Cortical Cognition

2) Distributed Representations

   - Multiple active neuron-like processing units
   - Same unit used in multiple representations

→ Efficiency: combinations of small set of units

→ Robustness: redundancy of items represented by many units

→ Accuracy: representation of graded values

→ Similarity: shared units between similar items

Principles for Biologically Based Models of Cortical Cognition

3) Inhibitory Competition

- 20% of cortex neurons are inhibitory interneurons → controls explosion of activations (epilepsy)
- Allows for sparse distributed representations: general structure of the environment

Principles for Biologically Based Models of Cortical Cognition

4) Bidirectional Activation Propagation (Interactivity)

- Both bottom-up and top-down flows
- Iterative steps through a settling process to stabilize the flows of information

Principles for Biologically Based Models of Cortical Cognition

5) Error-driven Task Learning

- Supervised learning: minimize error between desired outcome and network output → back-propagation algorithm
- Not biologically plausible, except if using interactivity (GeneRec algorithm)
- Propagates 2 signals and locally take their difference at each unit
  - Expectation phase: expected consequences
  - Outcome phase: experience actual consequences

→ Teaching signal: difference between expectation and outcome

Principles for Biologically Based Models of Cortical Cognition

6) Hebbian Model Learning

- Self-organizing or unsupervised learning
- Internal representations of statistical structure of the environment
- Correlational structure → Hebbian
- Biologically plausible: NMDA-mediated LTP through changes of synaptic strength
- Co-occurrence of items ⇒ causal relationships (PCA)

Interactions Among the Principles

1) Interactivity and Noise

In an identification scenario

- Bidirectional activation propagation

  VS

- Exhibit independent contributions from context

→ Because of correlations between neurons

Interactions Among the Principles

2) Distributed Representations and Competition

- Distributed representations with multiple active neurons

VS

- Competition that inhibits cooperativity

→ Leads to sparse representation (⚠️ combinatorial explosion)

KWTA: enforces competition while allowing for sparse distributed representation

⚠️ Not treated within a probabilistic learning framework

Interactions Among the Principles

3) Learning Principles

- Computational objectives of learning
  - Task learning
  - Model learning
- Implementational mechanisms
  - Error-driven learning
  - Hebbian learning

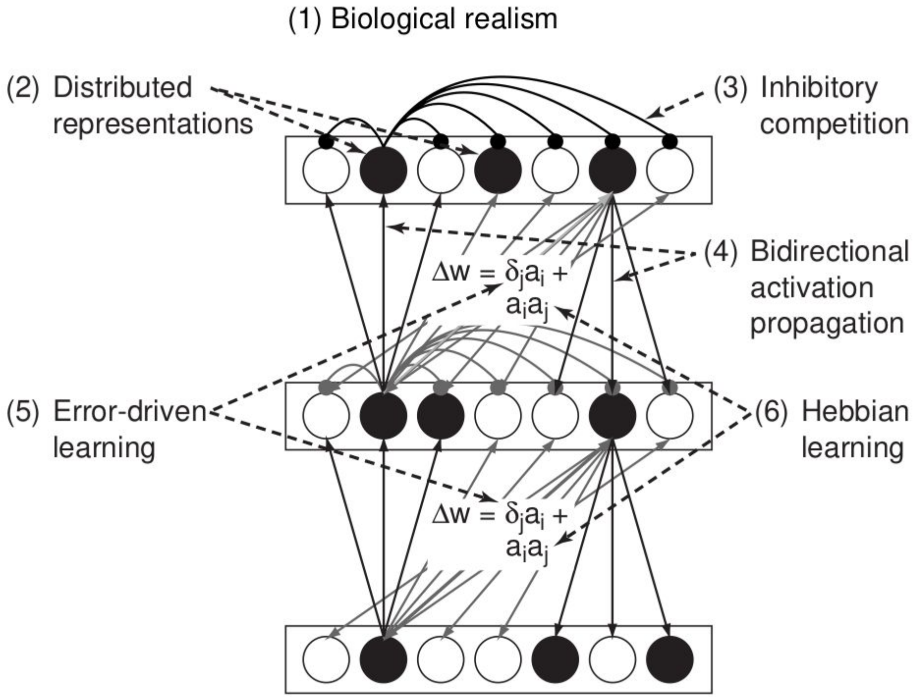
→ Quite similar

Only error-driven learning may map arbitrary inputs to outputs.

But may include a Hebbian bias towards representing co-occurrence statistics.

Leabra Algorithm

→ **L**earning in an **E**rror-driven and **A**ssociative, **B**iologically **R**ealistic **A**lgorithm



Leabra Algorithm

- Point-Neuron Activation Function

$$\frac{dV_m}{dt}(t) = \tau \sum_c g_c(t)\, \overline{g_c}(E_c - V_m(t))$$

Channels: excitatory input, leak current, inhibitory input

Communication to other cells: $y_j(t) = \dfrac{1}{1 + \dfrac{1}{\gamma[V_m(t) - \Theta]_+}}$

Leabra Algorithm

- K-Winners-Take-All Inhibition (KWTA)

Inhibitory current: $g_i = g^{\Theta}_{k+1} + q(g^{\Theta}_k - g^{\Theta}_{k+1})$ with $\begin{cases} q & \approx 0.25 \\ g^{\Theta} & = \dfrac{\sum_{c \neq i} g_c \overline{g_c}(E_c - \Theta)}{\Theta - E_i} \end{cases}$

Achieves sparse distributed representation.

- Using uniform level of inhibition ⇒ no more than k units active at a time

Leabra Algorithm

- Error-driven Learning

→ Symmetric version of the biologically plausible GeneRec algorithm

$$\Delta w_{ij}^{E} = x_i^{+} y_j^{+} - x_i^{-} y_j^{-}$$

2 phases:

- Expectation phase (-)
- Outcome phase (+)

Leabra Algorithm

- Hebbian Learning

Variant of Oja normalization: $\Delta w_{ij}^{H} = x_i y_j - y_j w_{ij} = y_j (x_i - w_{ij})$

→ Compute expected value of sender conditional on receiver: $w_{ij}^{H} \approx < x_i | y_j >$

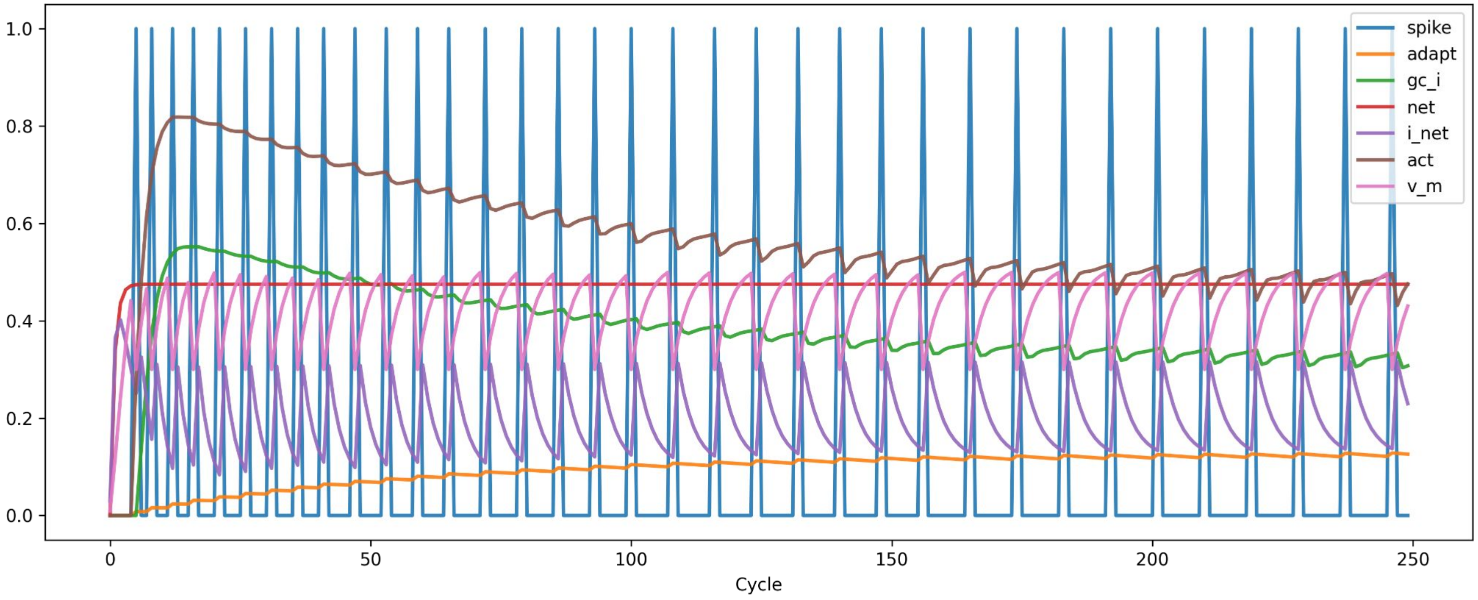
Total update: $\Delta w_{ij} \propto \Delta w_{ij}^{E} + \lambda \Delta w_{ij}^{H}$

Experiments

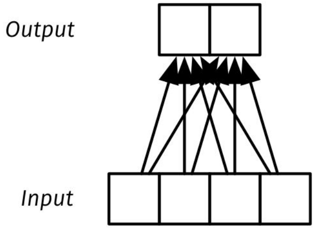
- Two neurons
- Pattern association
- Iris dataset classification

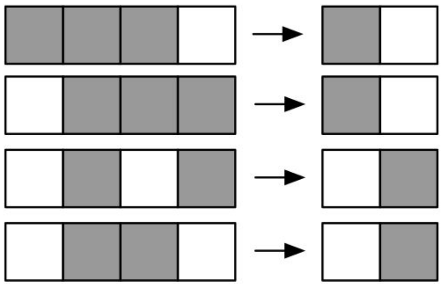
Two Neurons

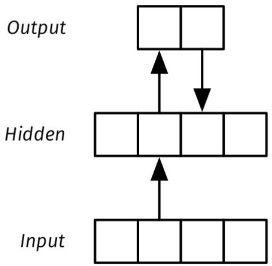
- Weight: 0.5
- Input activation: 1
- 1 cycle: 1ms
- V_m > thresh ⇒ spike

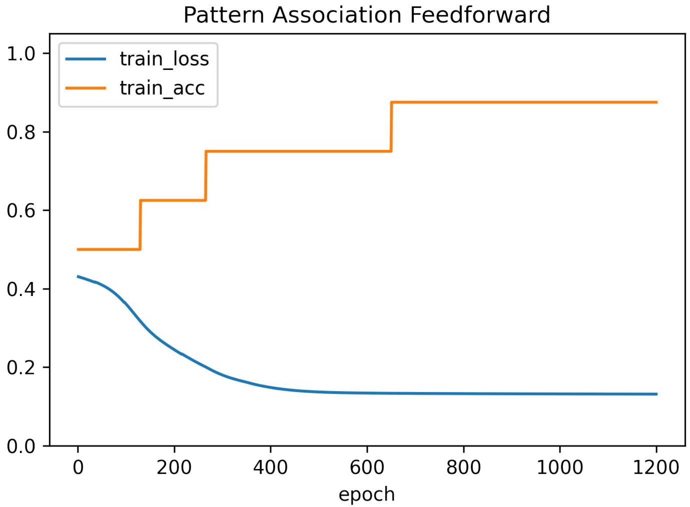


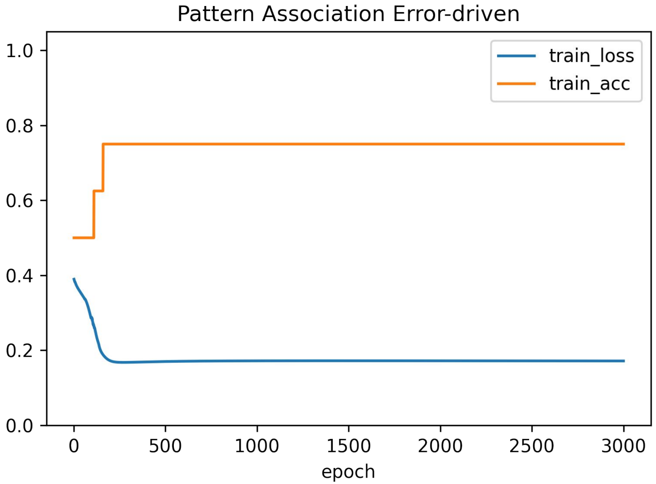
Pattern Association





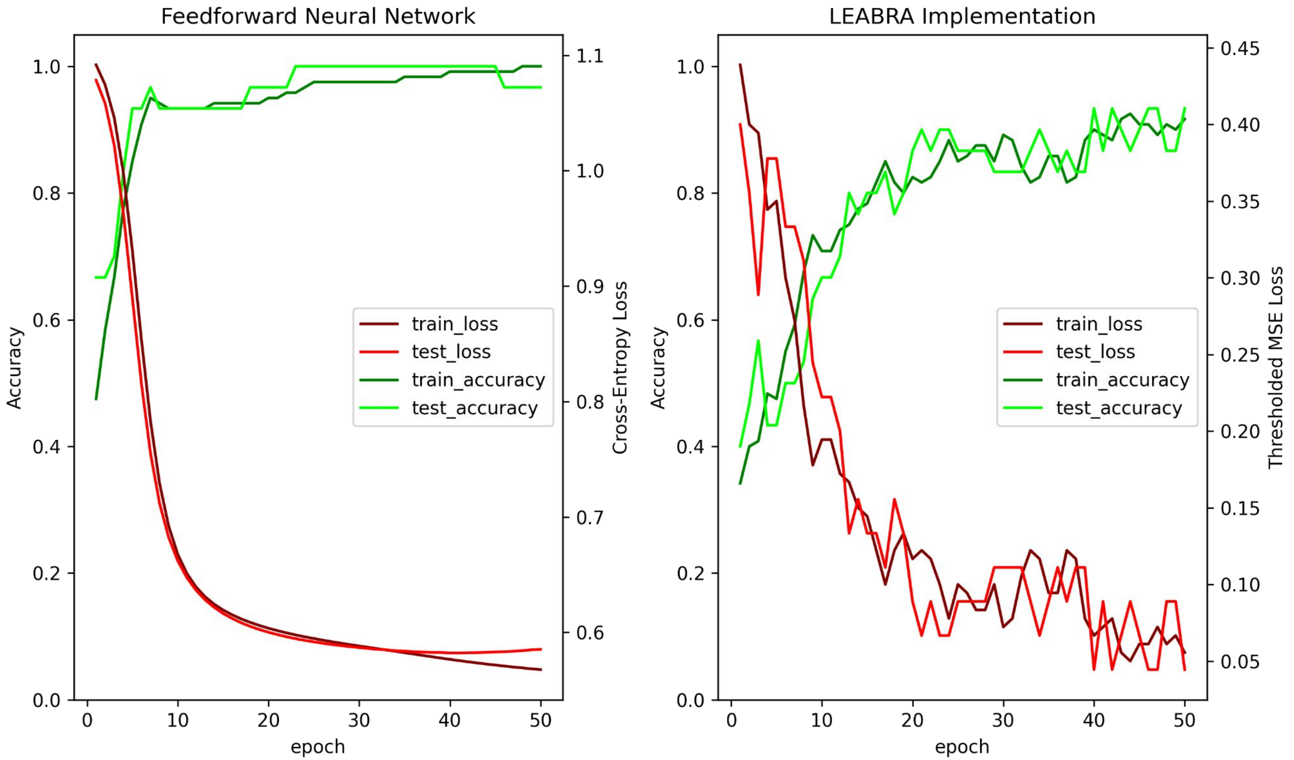






Iris Classification

- 40 input units
- 23 hidden units
- 3 output units



References

[1] Randall C. O'Reilly. "Six principles for biologically based computational models of cortical cognition". In: Trends in Cognitive Sciences 2.11 (1998).

[2] R. C. O'Reilly. "Biologically Plausible Error-Driven Learning Using Local Activation Differences: The Generalized Recirculation Algorithm". In:Neural Computation 8.5 (1996).

[3] Erkki Oja. "Simplified neuron model as a principal component analyzer". In: Journal of Mathematical Biology 15.3 (1982).

[4] C. Daniel Greenidge, Noam Miller, and Kenneth A. Norman. "Leabra7: a Python package for modeling recurrent, biologically-realistic neural networks". In: CoRRabs/1809.04166 (2018).

Conclusion

- Presented 6 principles
- Can be integrated into a coherent framework
- LEABRA does learn
- Training is unstable
- Analysis is hard (no probabilistic framework / sparsity / complex equations)
- Great potential for applications in ML and further research in cognitive neuroscience