Symbolic Reasoning in Spiking Neurons:

A Model of the Cortex/Basal Ganglia/Thalamus Loop

Terrence C. Stewart
Xuan Choo
Chris Eliasmith
Centre for Theoretical Neuroscience
University of Waterloo

University of Waterloo

Goal

- Create a neural cognitive architecture
  - Biologically realistic
    - spiking neurons, anatomical constraints, neural parameters, etc.
  - Supports high-level cognition
    - symbol manipulation, cognitive control, etc.
- Advantages
  - Connect cognitive theory to neural data
  - Neural implementation imposes constraints on theory

Required Components

Representation

- Distributed representation of high-dimensional vectors

Transformation

- Manipulate and combine representations

Memory

- Store representations over time

Control

- Apply the right operations at the right time

Representation

Assumption

Cognition uses high-dimensional vectors for representation

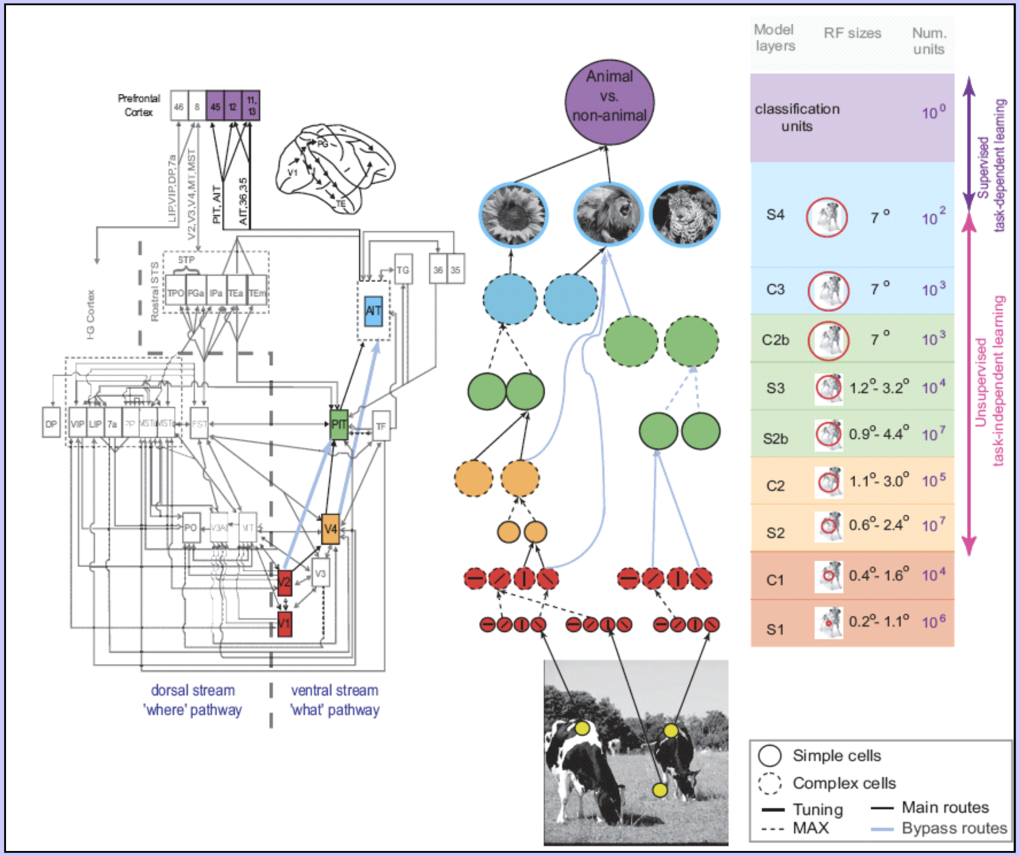
[2, 4, -3, 7, 0, 2, ...]

Top level of many hierarchical object recognition models

Compressed information

Different vectors:

DOG, CAT, SQUARE, TRIANGLE, RED, BLUE, SENTENCE, etc.



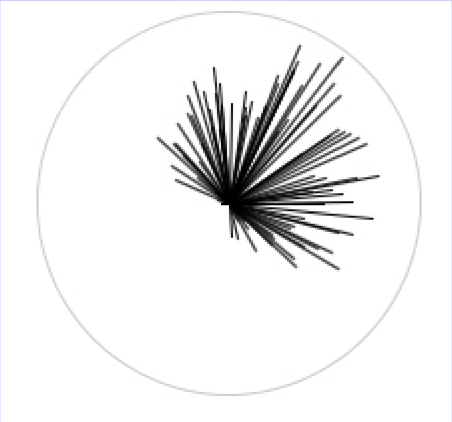
Serre et al., 2007 PNAS

Representation

How can a group of neurons represent vectors?

Visual and motor cortex (e.g. Georgopoulos et al., 1986)

- Representing spatial location (2D)
- Distributed representation
- Each neuron has a “preferred” direction
  - One direction it fires most strongly for
  - Uniformly distributed

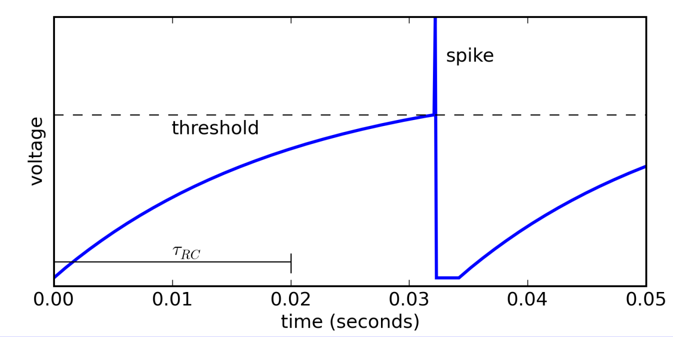


[run]

Representation

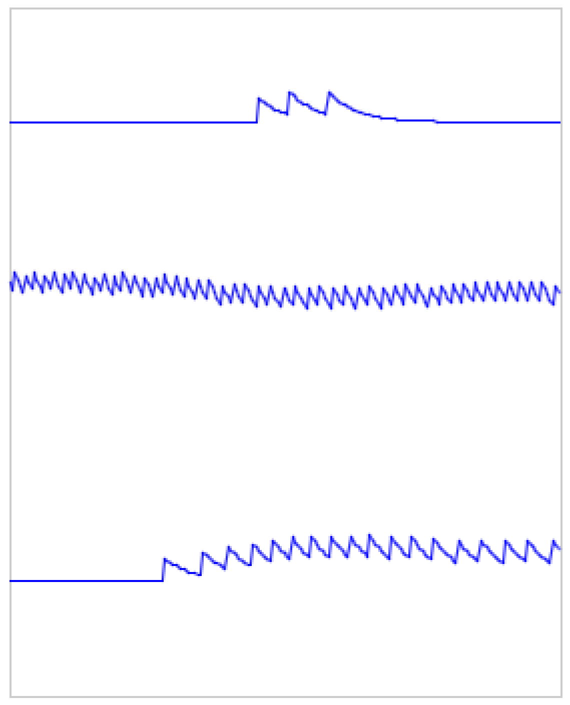
Neural representation

- Leaky Integrate-and-Fire neurons
- Input current: $J = \alpha \boldsymbol{e} \cdot \boldsymbol{x} + J_{bias}$

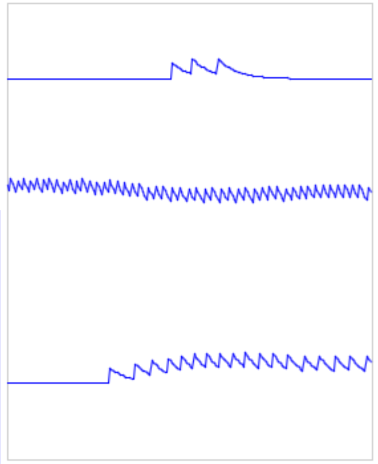


How good is the representation?

- Vector → spikes
- Can the input be recovered from the output?
- Post-synaptic current → vector



Representation



Linear decoding

- Weighted sum of neural outputs
- Need weights ***d*** (decoding) for optimal estimate of input

$$\mathbf{d} = \Gamma^{-1}\Upsilon \quad \Gamma_{ij} = \int a_i a_j \, dx \quad \Upsilon_j = \int a_j \mathbf{x} \, dx$$

- Extends to higher dimensions
  - Basis of Neural Engineering Framework
  - Decrease error by increasing number of neurons
  - Distributed representation
  - Robust to noise, neuron loss
  - (Eliasmith & Anderson, 2003)

[run]

Transformation

Using the representations

- Transfer information from one neural group to another

Simplest case:

- Form synaptic connections between group A and group B such that B will represent whatever A represents
- Communication channel
- Neural Engineering Framework
  - Optimal synaptic connection weights: $w_{ij}=\alpha_j \boldsymbol{e}_j \boldsymbol{d}_i$

[run]

Transformation

Linear transformation

Group A represents $\boldsymbol{x}$, want group B to represent $\boldsymbol{Mx}$

[x y z] → [2x 3x-2y+z]

$$w_{ij} = \alpha_j \boldsymbol{e}_j \boldsymbol{M} \boldsymbol{d}_i$$

Non-linear transform $f(\boldsymbol{x})$

$$\boldsymbol{d}^{f(\boldsymbol{x})} = \Gamma^{-1} \Upsilon \quad \Gamma_{ij} = \int a_i a_j dx \quad \Upsilon_j = \int a_j f(\boldsymbol{x}) dx$$

Some functions more accurately represented than others

What functions do we need?

Transformation

Binding operation

- We have vectors for RED, BLUE, TRIANGLE, CIRCLE
- How do we represent “a red triangle and a blue circle”?
- Solution via Vector Symbolic Architectures (Gayler, 2003)
  - Create a new vector for RED$\otimes$TRIANGLE
  - Should be highly dissimilar from other vectors
- Many suitable functions
  - Circular convolution (Plate, 2003)
  - Inverse operation: (RED$\otimes$TRIANGLE)$\otimes$RED'$\approx$TRIANGLE
    - where RED' is a rearrangement of RED (linear operation)

[run]

Memory

Need to store representations over time

- Short-term working memory
- Neurons need to maintain their firing over time
- Need recurrent connections
- Communications channel *back to itself*

$$w_{ij} = \alpha_j \boldsymbol{e}_j \boldsymbol{d}_i$$

  - Given no input, will maintain current value (some random drift)
  - Given input, will add input to current value (integrator)

[run]

Memory

Decay time

- Depends on number of neurons, time constant of neurotransmitter
- Matches spike data from PFC during memory task
  - (Romo et al., 1999; Singh & Eliasmith, 2006)

Storage capacity

- Scales exponentially in # dimensions
- ~50,000 neurons can represent 8 pairs of terms out of a vocabulary of 60,000 items
  - (see Stewart, Tang, Eliasmith, 2010)

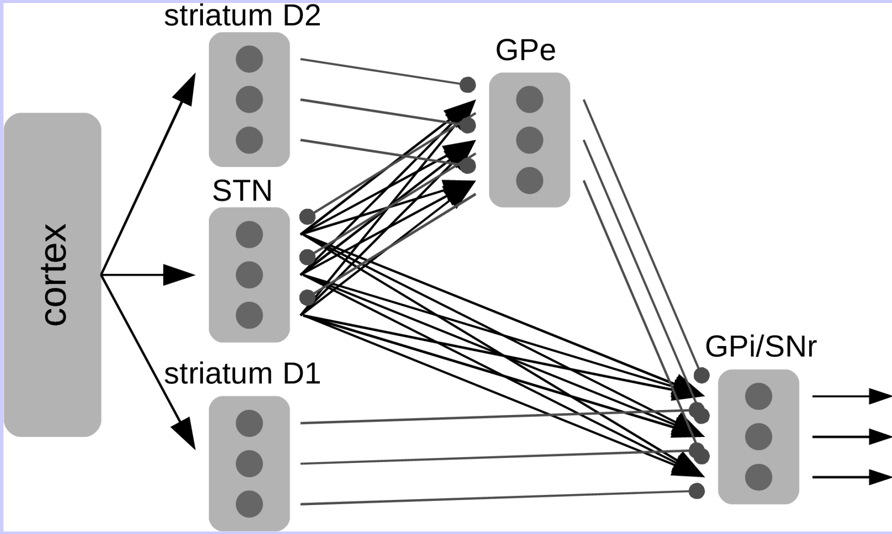
Control

Need to do different things at different times

- Set the inputs to the memory
- Set what to extract from memory

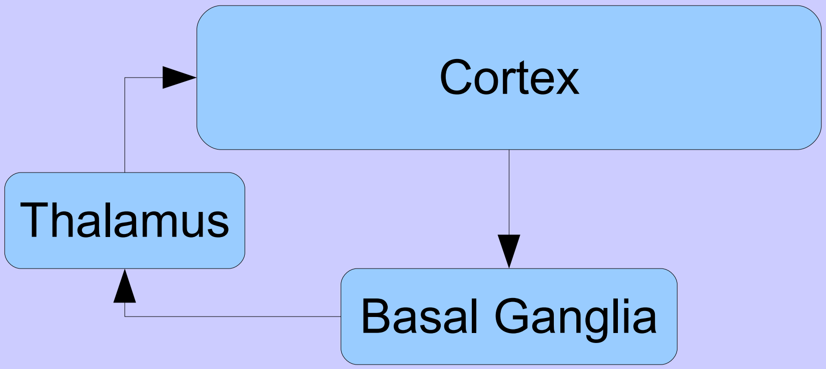
Action selection

- Choose one action out of many
- Basal ganglia
- Select action with greatest utility
- Non-spiking model (Gurney, Prescott, & Redgrave, 2001)
- Convert to spiking model via Neural Engineering Framework (Stewart, Choo, & Eliasmith, 2010b)



[run]

Control



Cortex

- Stores and transforms vectors

Basal ganglia

- Compare cortex state to optimum states for all actions
- Output inhibits all actions in thalamus except best action
  - (closest match)

Thalamus

- Executes chosen action
- Sends vectors to cortex; controls cortex transformations

Sequential Action

Simple example

- If working memory contains A, set working memory to B
- If working memory contains B, set working memory to C
- If working memory contains C, set working memory to D
- ...

Implementation

- Connect cortex to basal ganglia using $w_{ij} = \alpha_j \boldsymbol{e}_j \boldsymbol{M} \boldsymbol{d}_i$ where $\boldsymbol{M}$=[A B C ...]
- Connect thalamus to cortex using $\boldsymbol{M}$=[B C D ...]

[run]

Information Routing

- Add a visual area to cortex
- Add one action to basal ganglia
  - If LETTER in visual area, transfer contents of visual area to cortex
  - Add communication channel: visual to working memory
  - Connect action in thalamus: inhibit communication channel
- Start with nothing (no action chosen)
  - Present LETTER+C to visual area
  - Transfers C to working memory, continues sequence

[run]

Question Answering

Two actions

- If STATEMENT in visual, send to working memory
- If QUESTION in visual, send to working memory extraction area

Stimulus

- Present statement to visual area
  - STATEMENT+RED⊗TRIANGLE+BLUE⊗CIRCLE
- Remove statement
- Present question to visual area
  - QUESTION+RED

[run]

Results

- Successful implementation of controlled symbolic reasoning in spiking neurons
  - Symbols are vectors, vectors can be bound and unbound
  - Neural implementation of VSAs
- Scales to human-sized vocabulary
- Timing different for two types of actions
  - Simple: 34-44ms; Communication channels: 59-73ms
    - (see Stewart, Choo, Eliasmith, 2010b)
  - No free parameters (all fixed from neuroscience data)

Conclusions

Can perform symbol manipulation via spiking neurons

- Distinguishes “red triangle and blue circle” from “red circle and blue triangle”
- Implements IF-THEN action rules (productions)
  - IF cortex state is similar to X
  - THEN send particular values to particular areas in cortex and/or activate particular communication channels
- Conforms to neural anatomy, neuron properties
  - (firing rates, neurotransmitter time constants)
- Produces neural activity predictions
- Suggests change to the standard 50ms cognitive cycle time