

# Sentence processing in spiking neurons: A biologically plausible left-corner parser

Terrence C. Stewart (tcstewar@uwaterloo.ca)

Xuan Choo (fchoo@uwaterloo.ca)

Chris Eliasmith (celiasmith@uwaterloo.ca)

Center for Theoretical Neuroscience, University of Waterloo  
Waterloo, ON, Canada N2L 3G1

## Abstract

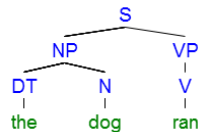
A long-standing challenge in cognitive science is how neurons could be capable of the flexible structured processing that is the hallmark of cognition. We present a spiking neural model that can be given an input sequence of words (a sentence) and produces a structured tree-like representation indicating the parts of speech it has identified and their relations to each other. While this system is based on a standard left-corner parser for constituency grammars, the neural nature of the model leads to new capabilities not seen in classical implementations. For example, the model gracefully decays in performance as the sentence structure gets larger. Unlike previous attempts at building neural parsing systems, this model is highly robust to neural damage, can be applied to any binary-constituency grammar, and requires relatively few neurons (~150,000).

**Keywords:** Neural engineering framework; vector symbolic architectures; left-corner parsing; syntax; binary trees; computational neuroscience

## Introduction

Human language processing requires not only the ability to represent complex structures, but also the ability to create these representations out of a serial sequence of words. This system is flexible enough to work for a huge variety of possible sentences, including (crucially), novel sentences.

Modern linguistics is founded on the principle that these structured representations are tree-like, and that this tree structure imposes useful order on the sentence. For example, “the dog ran” can be parsed as follows:



Here the sentence (S) is divided into a noun phrase (NP) and a verb phrase (VP). The noun phrase is divided into a determiner (DT) “the” and a noun (N) “dog”. The verb phrase consists of a single verb (V) “ran”.

To describe the space of possible trees we define *constituency rules* indicating how the parts can fit together. For example, the above sentence is consistent with the following constituency grammar rules:

|                         |                               |
|-------------------------|-------------------------------|
| $S \rightarrow [NP VP]$ | $DT \rightarrow \text{“the”}$ |
| $NP \rightarrow [DT N]$ | $N \rightarrow \text{“dog”}$  |
| $VP \rightarrow [V]$    | $V \rightarrow \text{“ran”}$  |
| (Structural rules)      | (Vocabulary)                  |

By adding more structural rules (on the left), we can build more complex sentences. By adding more words (on the right), we can increase our vocabulary.

## Left-Corner Parsing

As more rules are included in the grammar, parsing becomes more complicated. More rules mean more possibilities, and finding a tree that is consistent with the rules can become computationally expensive. For example, there may be the two rules  $VP \rightarrow [V]$  and  $VP \rightarrow [V NP]$ , leading to an explosion of possible structures to search through. Furthermore, a single word may have multiple interpretations, such as  $N \rightarrow \text{“dog”}$  and  $V \rightarrow \text{“dog”}$ .

The Left-Corner parsing algorithm addresses this problem by combining bottom-up and top-down information. The top-down information is what part of speech we are currently looking for (a sentence, a noun phrase, a determiner, etc.). The bottom-up information is the single word we are currently processing. So, if we are currently looking for a noun, we will first try interpreting the word “dog” as a noun, rather than a verb. If this does not lead to a successful parse, then the algorithm will backtrack to this point and try the other option. This approach drastically reduces the amount of backtracking needed.

For “the dog ran”, the algorithm proceeds as follows:

- Top-down: look for S
- Bottom-up: see “the”
- Apply rule  $DT \rightarrow \text{“the”}$
- Apply rule  $NP \rightarrow [DT N]$
- Store the partially completed tree
  - Top-down: look for N for previous rule
  - Bottom-up: see “dog”
  - Apply rule  $N \rightarrow \text{“dog”}$
  - Merge with previously stored tree
- Apply rule  $S \rightarrow [NP VP]$
- Store the partially completed tree
  - Top-down: look for VP for previous rule
  - Bottom-up: see “ran”
  - Apply rule  $V \rightarrow \text{“ran”}$
  - Apply rule  $VP \rightarrow [V]$
  - Merge with previously stored tree

As has been pointed out (e.g. Johnson-Laird, 1983), this algorithm matches well with observed human sentence comprehension. For example, it has difficulty with “garden-path” sentences such as the classic “the horse raced past the barn fell”, which is difficult for most people to interpret even though it has a similar form as the easier sentence “the deer shot by the hunter fell”. A left-corner parser has the same tendency as people do to connect the ambiguous word “raced” as part of a  $S \rightarrow [NP VP]$  rule.

## Previous Models

Cognitive models of left-corner parsing already exist. For example, Lewis and Vasishth (2005) present an implementation using the ACT-R cognitive architecture. In that work, a set of IF-THEN production rules (for selecting grammar rules to apply) are combined with a declarative memory system (for storing the partially completed parts of the tree). The result is a computational model that shows how existing cognitive modules (for which there is strong prior evidence they exist in the human brain) can be repurposed to perform left-corner parsing. The model's reaction times and error patterns match well to human subjects. However, this work does not offer a neural explanation of how those particular modules and structures could be physically instantiated within the human brain. Indeed, the question of how language could be represented and manipulated by interacting neurons is a long-standing question in cognitive science (e.g. Jackendoff, 2002).

For a neural explanation of this process, van der Velde and de Kamps (2006) offer their Neural Blackboard Architecture. Here, a set of specific neural groups can temporarily come to represent different nouns, verbs, adjectives, and so on. This makes it possible to represent structured information such as sentences. While we have previously argued (Stewart & Eliasmith, 2012) that the neural structures proposed by this approach are inefficient and do not correspond to those seen in real brains, the core difference here is that they have only shown how specific cases of sentence patterns might be parsed, rather than presenting a general method for parsing arbitrary compositional grammar rules, as is attempted here.

Finally, we previously presented a system for parsing highly specific sentence patterns, but using a biologically realistic neural model (Stewart & Eliasmith, 2013). These parsed commands, such as “write two” or “if see eight write three”, could be successfully interpreted as commands and used to guide action (Choo & Eliasmith, 2013). However, as with the work by van der Velde and de Kamps, this model was restricted to parsing highly specific grammatical forms. The purpose of this paper is to generalize to significantly more complex grammars and to connect this neural work to the grammatical structures seen in established linguistic theory.

## The Semantic Pointer Architecture

The core contribution of this paper is an implementation of general-purpose left-corner parsing within a biologically realistic cognitive architecture. For that purpose, we use our Semantic Pointer Architecture (Eliasmith, 2013), which has been previously used to build Spaun, the first large-scale brain simulation capable of performing multiple tasks (Eliasmith et al., 2012). In the SPA, all parts of the model can be implemented using biologically realistic simulated spiking neurons. For this paper, we use the standard Leaky Integrate-and-Fire (LIF) neuron model. Each module in the SPA corresponds to a particular brain area, and the synaptic connections are optimized to compute some function.

For example, a common cortical module is a *buffer*. This is a group of neurons whose purpose is to store a value over time. We formalize this by expressing it as a differential equation as follows: the value to be represented is  $x$ , and there is some input to the group of neurons  $u$ . If we want a group of neurons that can store  $x$  over time, then we want  $x$  to not change at all when  $u$  is zero. In other words, if there is no input, do not change the value, and if there is an input, change the stored value by that much. Mathematically, this can be written as  $\frac{dx}{dt} = u$ .

The reason to express this as a differential equation is that any differential equation can be approximated by a group of spiking neurons using the Neural Engineering Framework (NEF; Eliasmith & Anderson, 2003). We do this by randomly generating a tuning curve (the neural activity for a given  $x$  value) for each neuron, consistent with observed firing patterns from that cortical area. For example, one neuron may fire at 1Hz for  $x = -0.5$ , but fire at 10Hz for  $x = 1.0$ . These tuning curves are randomly generated with a distribution of firing patterns consistent with empirical data. The NEF lets us do local optimization to find the optimal synaptic connections between two groups of neurons so as to achieve the empirically identified tuning curves.

That is, if one group of neurons have a set of tuning curves dependent on  $x$ , while another group of neurons need tuning curves based on  $y$ , then we can find a set of connection weights from  $x$  to  $y$  such that the neurons approximate the computation  $y = f(x)$ . When recurrent connections are introduced, the NEF allows for the approximation of any function  $\frac{dx}{dt} = f(x, u)$ . Importantly, the accuracy of this approximation is dependent on the number of neurons and the complexity of the function (roughly, how non-linear it is).

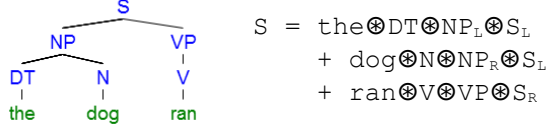
Since  $x$  can be a vector, the NEF allows for the implementation of neural models that manipulate vectors in desired ways. However, for a parsing system, we need to manipulate symbolic structured information using vectors.

For this, we turn to Vector Symbolic Architectures (VSAs; Gayler, 2003). These are a set of algorithms where both symbols and symbol structures can be expressed as high-dimensional vectors. For example, we might have one 1000-dimensional vector than means “dog”, and another than means N (noun), and even another that means “the”. These vectors can be randomly chosen (as they are here), or they can be chosen such that similar terms (“dog” and “wolf”, for example) could have similar vectors.

We now need a way to combine these vectors. We could, for example, simply add them together to produce a new vector. However, this would result in lost information; `dogs+chase+cats` would be the same as `cats+chase+dogs`. Instead, VSAs introduce a *binding* operation. Different VSAs choose different binding operators, and for our work we chose circular convolution (Plate, 2003), written as  $\otimes$ , as it can be efficiently implemented with the NEF. We can now encode the sentence “dogs chase cats” as `S=dogs $\otimes$ subject+chase $\otimes$ verb+cats $\otimes$ object`. Importantly,

VSAs also define an inverse operation: given  $S$  we can determine the subject of the sentence by computing  $S \otimes \text{subject}^{-1}$ , which is approximately “dog”.

Here, we extend this approach to full parse trees. For example, “the dog ran” is represented as follows:



This computes a single vector  $S$  which stores that particular tree. The terms “the”, “dog”, and “ran” are randomly chosen vectors for each of those words. The terms  $DT$ ,  $N$ ,  $NP_L$ ,  $S_R$ , and so on are also randomly chosen vectors. These are used to indicate the structure of the tree. Note that there are different vectors for taking the left or right branch of the tree (the  $R$  or  $L$  subscript). This is so that the vector for the nonsensical “ran the dog” will be different than the vector for this tree ( $S_L$  and  $S_R$  would be swapped). A similar approach could be used for more than just binary (left or right) branching, but only binary rules are considered here.

Given the vector  $S$ , we can determine the verb, for example, by computing  $S \otimes (V \otimes VP \otimes S_R)^{-1}$ . A parse is only considered accurate if this vector is closer to “ran” than to any other word in its vocabulary. For this paper, vocabulary sizes are set at 10,000 other randomly chosen vectors.

The Semantic Pointer Architecture uses vectors as a generic method for passing information between cortical modules. For example, buffers can store information, sensory areas can turn stimuli into the appropriate vector, and motor areas take vectors describing the desired action and convert them into muscle movements (see Eliasmith, 2013 for more details). However, it also needs a method for controlling the flow of information between these neural modules. We achieve this via a model of the cortex-basal ganglia-thalamus loop (Stewart, Choo, & Eliasmith, 2010). This acts as an *action selection* and *action execution* system. Neural connections from the rest of the brain into the basal ganglia compute the *utility* of each of the possible actions that could be taken. The basal ganglia determine which of those utility values is largest, and pass that information to the thalamus. In the thalamus, neurons for every action except for the one that is chosen are suppressed.

To build a model with the SPA, we thus define a set of actions that can be performed. For each action  $i$ , we define its effects  $E_i$  (what vectors should be sent from one cortical area to another) and its utility  $U_i$  (a function that should give a value near 1 when this action should be performed and smaller values otherwise). For example, the following rule would send the output from a *memory* module to the input to a *speech* module whenever the *vision* module sees a dog:

$$U_i: \text{vision} \bullet \text{dog} \quad (\text{this will be large when } \text{vision} = \text{dog})$$

$$E_i: \text{speech} \leftarrow \text{memory}$$

These rules are efficiently implemented via the NEF, requiring ~300 basal ganglia neurons per rule (Stewart, Choo, & Eliasmith, 2010).

## The Model: Left-Corner Parsing in SPA

To develop a neurally plausible implementation of left-corner parsing within the Semantic Pointer Architecture, we need to define cortical modules, their functions, and the basal ganglia/thalamus rules that coordinate the flow of information between these modules.

For cortical modules, we only need one basic component: a *buffer* capable of storing a semantic pointer. This is a module that stores a single high-dimensional vector (for this model, 1000 dimensions is sufficient) over time. We need three of these: one to store the tree being built (*tree*), one to store the current top-down goal (*goal*; what part of speech we are looking for), and one to store the partially completed trees (*partial*). Note that the neural modules needed to visually recognize words, or to move visual attention from one word to the next, are not considered here (see Tang & Eliasmith, 2010 and Bobier, Stewart & Eliasmith, 2011 for potential modules).

Since these buffers store vectors, we can use the approach of Vector Symbolic Architectures to store a tree within them. We assume words are presented sequentially, and that their vectors are stored in the buffer called *tree*.

We now need rules for the basal ganglia and thalamus system that cause the model to implement the parsing algorithm. We start with simple rules of the form  $X \rightarrow [Y]$ . For each of these, we need a system that says if we're currently parsing a  $Y$ , we should build a tree that consists of an  $X$  connected to a  $Y$ . In terms of vectors, this means building a new vector that is  $X + \text{tree} \otimes X$ . For example, if the *tree* buffer contains *ran* and we have a rule  $V \rightarrow \text{ran}$ , we want to compute  $V + \text{ran} \otimes V$  and store that in the *tree* buffer. This can be written in SPA form as follows:

$$U_i: \text{tree} \bullet Y$$

$$E_i: \text{tree} \leftarrow X + \text{tree} \otimes X$$

Note that the utility of this rule is the dot product of whatever is in the *tree* buffer and the  $Y$  part of the rule, so it will only be active when the *tree* looks like  $Y$ .

To see how this helps to build up the desired tree, consider what happens if there is another similar rule implementing  $VP \rightarrow [V]$ . Once the first rule is active, the *tree* buffer will contain  $V + \text{ran} \otimes V$ . This will cause the rule for  $VP \rightarrow [V]$  to have a high utility (since its utility is  $\text{tree} \bullet V$ ). The effect of the rule is  $\text{tree} \leftarrow VP + \text{tree} \otimes VP$ , so the result will be  $VP + (V + \text{ran} \otimes V) \otimes VP$ . This vector is highly similar to  $\text{ran} \otimes V \otimes VP$ , which is part of what we need for parsing the complete tree for “the dog ran”.

We also need to handle rules of the form  $X \rightarrow [Y Z]$ , which gives the branching capability to the tree. Here we need to not only build up a tree, but we also need to set a new top-down goal to find a  $Z$ . The utility is the same as in the previous rule (we want the rule to be active when *tree* contains a  $Y$ ), but we also want to store the partially completed tree and go on to processing the next word. An initial version of this rule's effect would be:

$E_i: \text{partial} \leftarrow X + \text{tree} \otimes X_L$   
 $\text{tree} \leftarrow \text{the next word}$   
 $\text{goal} \leftarrow Z$

This rule would successfully store the partially completed tree in the *partial* buffer, and then would go on to start trying to parse the next word, trying to find a *Z*. However, setting this new *goal* in this way would completely replace the old value in *goal*. Indeed, setting the new *partial* tree would completely erase any previous *partial* tree.

To deal with this, we use the vector to store a list of trees (akin to a “stack”). This list can all be stored as a single vector by combining with circular convolution. For example, if the goal currently contains the vector for *S* (a sentence), but due to a rule like  $\text{NP} \rightarrow [\text{DT N}]$  we now need to look for an *N* (a noun), we can set the new *goal* to be  $\text{N} + \text{goal} \otimes \text{STACK}$ , where *STACK* is another random vector. In this case, the *goal* will now be  $\text{N} + \text{S} \otimes \text{STACK}$ . If we now need to look for a *V* (a verb), we would compute  $\text{V} + \text{goal} \otimes \text{STACK}$ , giving  $\text{V} + (\text{N} + \text{S} \otimes \text{STACK}) \otimes \text{STACK}$ , or  $\text{V} + \text{N} \otimes \text{STACK} + \text{S} \otimes \text{STACK} \otimes \text{STACK}$ . Note that to remove an item from the stack, we can compute  $\text{goal} \otimes \text{STACK}^{-1}$ , which would give us back an approximation of  $\text{N} + \text{S} \otimes \text{STACK}$ . This approximation will gradually become worse as the number of items in the stack increases.

The resulting rule for  $X \rightarrow [Y Z]$  is as follows:

$U_i: \text{tree} \bullet Y$   
 $E_i: \text{partial} \leftarrow X + \text{tree} \otimes X_L + \text{partial} \otimes \text{STACK}$   
 $\text{tree} \leftarrow \text{the next word}$   
 $\text{goal} \leftarrow Z + \text{goal} \otimes \text{STACK}$

Finally, we need a top-down rule to recognize when we have found a part of speech we are looking for. When it does so, we combine it with the partial tree on the stack, remove it from the stack, and continue. For a rule of the form  $X \rightarrow [Y Z]$ , we get:

$U_i: (\text{partial} \bullet X)(\text{goal} \bullet \text{tree})$   
 $E_i: \text{tree} \leftarrow \text{partial} + X_R \otimes \text{tree}$   
 $\text{goal} \leftarrow \text{goal} \otimes \text{STACK}^{-1}$   
 $\text{partial} \leftarrow \text{partial} \otimes \text{STACK}^{-1}$

So, if we are on the last word of “the dog ran”, the stored value in *partial* will be close to  $\text{the} \otimes \text{DT} \otimes \text{NP}_L \otimes \text{S}_L + \text{dog} \otimes \text{N} \otimes \text{NP}_R \otimes \text{S}_L$  and the value in *tree* will get built up to approximately  $\text{ran} \otimes \text{V} \otimes \text{VP}$ , as indicated previously. The goal will be *VP*. This means that the utility for this top-down version of the  $\text{S} \rightarrow [\text{NP VP}]$  rule will be high (since  $\text{partial} \bullet \text{S}$  will be large and  $\text{goal} \bullet \text{tree}$  will be large). The resulting tree will be  $\text{partial} + \text{S}_R \otimes \text{tree}$ , or approximately  $\text{the} \otimes \text{DT} \otimes \text{NP}_L \otimes \text{S}_L + \text{dog} \otimes \text{N} \otimes \text{NP}_R \otimes \text{S}_L + \text{ran} \otimes \text{V} \otimes \text{VP} \otimes \text{S}_R$ . This is the desired vector for the correct parse of the tree.

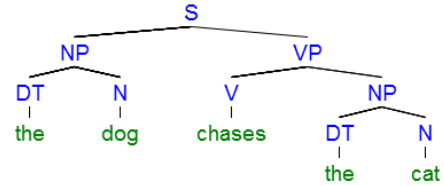
This algorithm will work for arbitrary binary construction grammar rules. However, it needs to be extended to deal with multiple possible rules that could be applied at once (ambiguous sentences), as is discussed below.

## Parsing Results

This biologically plausible left-corner parsing algorithm is capable of computing the vector-based tree representation of a sentence, given a set of constituency grammar rules. These rules are converted into connections between the cortex, basal ganglia, and thalamus, as per the Semantic Pointer Architecture. That neural structure is then capable of parsing sentences consistent with those rules.

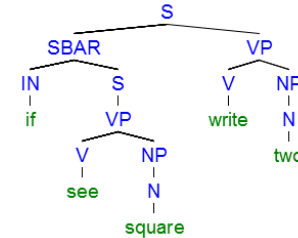
In addition to the “the dog ran” example given previously, the following set of rules allows the model to parse “the dog chases the cat”.

$\text{VP} \rightarrow [\text{V NP}]$   
 $\text{S} \rightarrow [\text{VP}]$   
 $\text{S} \rightarrow [\text{NP VP}]$   
 $\text{NP} \rightarrow [\text{DET N}]$   
 $\text{NP} \rightarrow [\text{N}]$



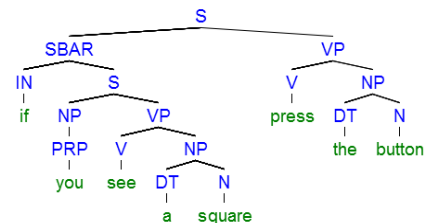
To demonstrate more complex parsing, we add rules for subordinate clauses. In previous work (Choo & Eliasmith, 2013) we had built neural models that could follow orders of the form “if see square, press button”. Rather than using the special-case neural parser we had built previously (Stewart & Eliasmith, 2013), the parser presented here can be extended by adding the following rules, where *IN* is a subordinating conjunction (such as “if”), and *SBAR* is a subordinating conjunctive phrase (“if see square”).

$\text{S} \rightarrow [\text{SBAR VP}]$   
 $\text{SBAR} \rightarrow [\text{IN S}]$



Importantly, this system can also parse more grammatically complete sentences. Indeed, if we also add the following rule for *PRP* (personal pronouns), then the system is able to successfully parse “if you see a square, press the button”.

$\text{NP} \rightarrow [\text{PRP}]$

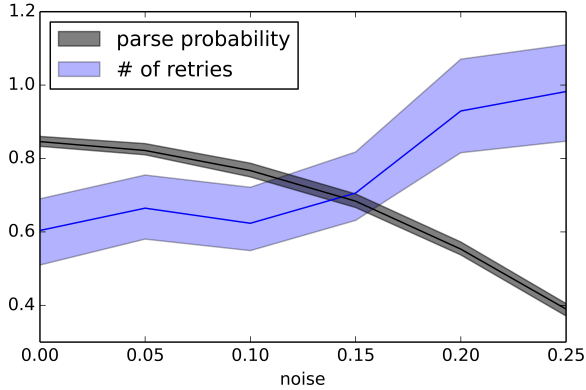




## Parsing Accuracy

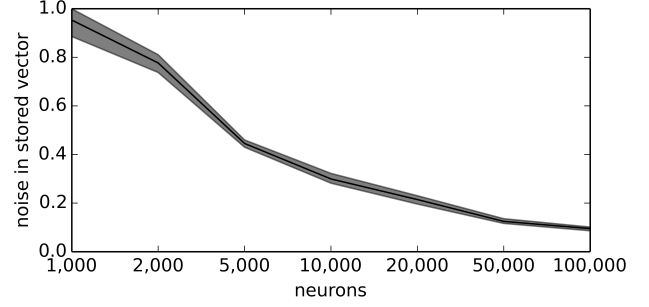
While the model is capable of parsing the previous sentences, its capabilities are not perfect. Indeed, the accuracy of the parsing is dependent on the accuracy of the neural representation. This is a concrete example of the competence vs. performance distinction: while the underlying algorithm may have the competence to parse those sentences, the neural implementation may not match that in terms of actual performance.

We can analyze this by determining the probability of correctly parsing the sentence as we increase the amount of noise. A word is considered to be correctly parsed if we can extract it from the vector for the sentence, using the standard approach to extracting in Vector Symbolic Architectures. For example, for the “the dog ran” the ideal resulting vector is  $S = \text{the} \otimes \text{DT} \otimes \text{NP}_L \otimes S_L + \text{dog} \otimes \text{N} \otimes \text{NP}_R \otimes S_L + \text{ran} \otimes \text{V} \otimes \text{VP} \otimes S_R$ . To determine if “the” was parsed correctly, we compute  $S \otimes (\text{DT} \otimes \text{NP}_L \otimes S_L)^{-1}$ . If this vector is closer to the vector for *the* than to any other vector in the vocabulary of 10,000 words, then the parse is considered to be correct. Noise is adjusted by adding random values to the stored *tree*, *partial*, and *goal* vectors after every rule. Finally, if the algorithm fails (i.e. no words can be extracted), then we retry it until it succeeds. Figure 1 shows that performance gets much worse with noise values of 0.15 or more, and that even small values of noise require an average of 0.6 retries.



**Figure 1:** The probability of successful parsing and the number of retries needed as the random noise varies. 95% confidence intervals are shown.

These results give us a lower bound on the accuracy of the neural representation that is needed for the buffers. Since the Neural Engineering Framework indicates that accuracy is increased as more neurons are added, we can find how many neurons are needed for each buffer to achieve this level of accuracy. We do this by creating groups of neurons with connection weights between them optimized via the NEF to compute the function  $\frac{dx}{dt} = 0$  (i.e. the value  $x$  that is being stored should not change). The network is initialized to contain a randomly chosen vector  $x$ , and the model is run for 50 milliseconds. We have previously shown that 50 milliseconds is the average amount of time taken between rule activation (Stewart, Choo, & Eliasmith, 2010).



**Figure 2:** Amount of representational noise for different numbers of neurons. 95% confidence intervals are shown.

Figure 2 shows that 50 to 100 neurons are needed per dimension to achieve noise levels around 0.15. Since Figure 1 indicates that performance is much improved with noise below 0.15, and since we are using vectors with 1000 dimension, we build our final model with 50,000 neurons per buffer. The complete model has 3 buffers and 18 basal ganglia/thalamus/cortex rules (300 neurons each), resulting in a parsing system with 155,400 neurons. Importantly, Figures 1 and 2 also show that the model exhibits graceful degradation. If neurons are removed (or die), accuracy will gradually decrease.

## Using Parsed Commands

While the model presented thus far is capable of parsing sentences, the real test is to be able to parse a sentence and then make use of that information. In previous work (Choo & Eliasmith, 2013), we presented a spiking neural model that is capable of following instructions of the form “If <condition>, <do action>”. However, that model did not perform the parsing itself. It relied on a special-case neural parser that only worked on particular word patterns (Stewart & Eliasmith, 2013). Since they both make use of the Semantic Pointer Architecture, the general-purpose left-corner parser presented here can be combined with the instruction-following model.

In the instruction-following model, the instruction “if see square, write two” was encoded as:

```
CONDITION ⊗ (SENSE ⊗ VISION + SENSE_DATA ⊗ SQUARE)
+ ACTION ⊗ (MOTOR ⊗ WRITE + MOTOR_DATA ⊗ TWO)
```

With the parser presented here, we can directly use the parsed sentence as a command to the instruction-following model by making the following definitions:

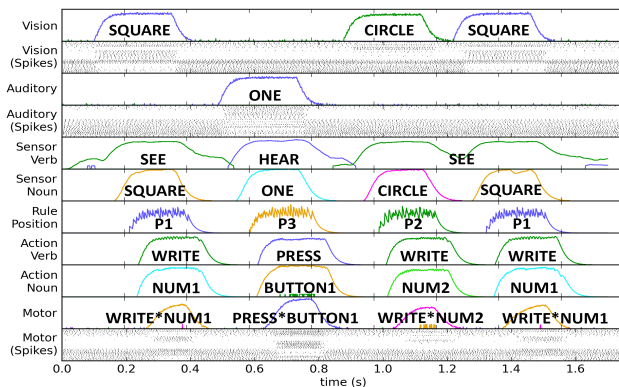
```
CONDITION = S_L
SENSE = V ⊗ VP_L ⊗ S ⊗ SBAR_R
SENSE_DATA = N ⊗ NP ⊗ VP_R ⊗ S ⊗ SBAR_R
ACTION = S_R
MOTOR = V ⊗ VP_L
MOTOR_DATA = N ⊗ NP ⊗ VP_R
```

The result is a biologically plausible neural model where we can feed in a set of commands as sentences, and the model can parse those sentences, remember them, and apply

them to incoming stimuli. For example, Figure 3 shows the effect of a system configured to follow these instructions:

- P1. If see square write 1
- P2. If see circle write 2
- P3. If hear one press button
- P4. If hear two press button2

The the visual and sensory inputs to the model change (top four rows of Figure 3), the model successfully responds as appropriate (bottom two rows). Of course, the model presented here does not include all the details necessary to actually perform the motor actions or the visual processing necessary to complete these tasks. The purpose here is to show that the model is capable of correctly selecting the task to perform.



**Figure 3:** The instruction-following model. Each row is a different group of neurons. Labelled lines indicate the value represented by this group of neurons is close to the indicated vector. First the model is visually shown a SQUARE (top row), and it responds correctly with the motor action of WRITE\*NUM1 (“write the number 1”; bottom row). It then hears a ONE, is shown a CIRCLE, and a SQUARE again. The correct motor response is given each time.

## Conclusions and Future Directions

We have presented a basic version of left-corner syntactic parsing that can be implemented in biologically realistic spiking neurons. This makes use of the cortex-basal ganglia-thalamus loop in the brain, and is compatible with our ongoing development of large-scale neural models capable of performing cognitive tasks. This shows how our Semantic Pointer Architecture can be adapted to implement cognitive algorithms, and that the resulting models can accurately manipulate complex structured representations like parse trees.

However, in order to truly model syntactic parsing, the model needs to be able to deal with ambiguity. Right now, if the parsing fails (due to it choosing the wrong rule to apply when multiple rules could be applied at a given time), we simply reset the model and re-run it, hoping it will choose the correct action next time. We are instead exploring methods to suppress recently applied rules, so as to encourage the use of alternatives and greatly improve this recovery process. We are also examining applying

reinforcement learning to situation, allowing it to learn to adjust the utility values  $U_i$ , opening up the possibility of context-sensitive parsing.

Other ongoing work is in improving the accuracy of the system. In particular, it appears that the parse accuracy (Figure 1) can be improved by increasing the dimensionality of the vectors and at the same time decreasing the number of neurons per dimension.

## References

- Ball, J. (2011). A Pseudo-Deterministic Model of Human Language Processing. *33<sup>rd</sup> Cog. Sci. Society Conference*.
- Bobier, B., Stewart, T.C., and Eliasmith, C. (2011). The attentional routing circuit: receptive field modulation through nonlinear dendritic interactions. *Proceedings of Cognitive and Systems Neuroscience*.
- Choo, X. and Eliasmith, C. (2013). General Instruction Following in a Large-Scale Biologically Plausible Brain Model. *35<sup>th</sup> Cog. Sci. Society Conference*.
- Eliasmith, C. (2013). *How to build a brain*. Oxford University Press, New York, NY.
- Eliasmith, C. and Anderson, C. (2003). *Neural Engineering*. Cambridge: MIT Press.
- Eliasmith, C., Stewart, T.C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., and Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science*, 338:1202-1205.
- Gayler, R. (2003). Vector Symbolic Architectures Answer Jackendoff’s Challenges for Cognitive Neuroscience, in Slezak, P. (ed). *Int. Conference on Cognitive Science*, Sydney: University of New South Wales, 133–138.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford, UK.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.
- Lewis, R. L. and Vasisht, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29:375-419.
- Plate, T. (2003). *Holographic Reduced Representations*, CSLI Publications, Stanford, CA.
- Stewart, T.C., Choo, X., and Eliasmith, C. (2010). Dynamic Behaviour of a Spiking Model of Action Selection in the Basal Ganglia. *10<sup>th</sup> Int. Conf. on Cognitive Modeling*.
- Stewart, T.C., and Eliasmith, C. (2012). Compositionality and biologically plausible models. In *Oxford Handbook of Compositionality*. Oxford University Press, 2012.
- Tang, Y. and Eliasmith, C. (2010). Deep networks for robust visual recognition. *Proceedings of the International Conference on Machine Learning*.
- Stewart, T.C. and Eliasmith, C.. (2013). Parsing sequentially presented commands in a large-scale biologically realistic brain model. In *35th Annual Conference of the Cognitive Science Society*, 3460-3467.
- van der Velde, F. and de Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences*, 29, 37-70