HTM Stability
Written by David McDougall, 2017

Introduction

This is a response to the paper: "Why Does the Neocortex Have Columns, A Theory of Learning the Structure of the World" Hawkins, Ahmad, Cui, 2017, http://dx.doi.org/10.1101/162263. The paper describes how two layers of HTM neurons can be used to learn viewpoint invariant representations. The model however does not form the viewpoint invariant representations; it assigns them and forces the output layer to hold them active during supervised learning. I have found an unsupervised method of forming viewpoint invariant representations. The method is implemented and tested on random data. It appears to work as intended. The purpose of these experiments is to show that the method could work; these experiments do not use real world data. In the future I hope to implement the full two layer circuit, incorperate this method into it, and test it on visual data.

The source code for these experiments is posted at github.com/ctrl-z-9000-times/HTM_experiments, see file stability_experiment.py.

Model

All of the changes to the model are in the output layer spatial pooler (S.P.) and so only that SP is implemented and tested. This model builds on the existing spatial pooler model described by (Cui, Ahmad, Hawkins, 2017, "The HTM Spatial Pooler - a neocortical..."). This model adds two critical features to their SP, 1) a new mechanism for stability and 2) proximal dendrites now have multiple segments.

1) A mechanism for stability. The stability of an SDR is defined as the overlap in its activations between consecutive moments in time. The stability of the output layer mini-columns is then the amount of overlap in the output layer's mini-column activations between consecutive moments in time. There is a target output layer mini-column stability which the mechanism maintains by forcing mini-columns to remain active if too few active mini-columns from the previous cycle remained active.
Mini-columns remain active if they fulfill the following criteria:
A) Was active in the previous time step,
B) Is no longer activated by feed forward input,
C) Is one of the X most proximally excited mini-columns, among the mini-columns which meet the criteria A & B, where X proportional to the target stability minus the actual stability.
The purpose of this mechanism is to make the output layer mini-columns differentiate between objects on the basis of how fast and with long term reproducibility the input sensor moves between the objects.

2) Proximal dendrites have multiple independent segments. The spatial pooler models the proximal dendrites of a pyramidal neuron. Numenta models the surface of these dendrites as a single large surface area. In this new model the proximal dendrites have been split into multiple distinct areas which operate independently from each other. Splitting the dendrites into independent segments offers the SP some of the same benefits as it offers the TM, such as protecting synapses from learning when they are not in use because inactive segments do not learn. I hypothesis that multiple proximal segments are critical for L2/3 SP stability because they allow columns to recognize a larger number of distinct inputs. These inputs may have little to no semantic similarity which means that their ability to share proximal dendrite segments is limited.
Implementation: A mini-columns overall excitement is the excitement of its most excited segment. When a mini-column activates only the most excited segment is allowed to learn. Boosting is applied at the segment level.

Dataset
The model is tested using randomly generated data. The dataset contains randomly generated sets of neuron activations. Each datum represents a set of activations in the input layer. The dataset is organized into 100 objects with 10 elements each. This is a time series dataset; the dataset picks a random object, plays 20 elements from it in random order, and repeats indefinitely. The model trains on the dataset for 100,000 cycles. A new dataset is generated every time the program is run.

Parameter Optimization
The parameters in the model were optimized using an evolutionary algorithm. The model was evaluated based on a combination of the following metrics:
1) Object classification accuracy. A statistical classifier was trained to recognized to current object. This metric is maximized.
2) Intra-object overlap. This is the average amount of output layer mini-column overlap between elements of the same object. This is exhaustively measured before and after training on the dataset. This metric is maximized.
3) Inter-object overlap. This is the average amount of output layer mini-column overlap between different objects. This is not exhaustively measured, 1/25 of the possible overlaps are sampled for efficiency. This metric is minimized.
4) Memory consumption. This metric is minimized if the program exceeds a certain memory usage.

Results
The evolutionary algorithm suggested the following parameters.

Parameter	Value	Description & Notes
enc_size	1950	Number of neurons in the input layer.
enc_sparsity	1.70122738181%	Sparsity of input layer activations.
init_dist[0]	0.049742999835167091	Average permanence of new synapses.
init_dist[1]	0.032157580415712081	Standard deviation of permanence of new synapses.
alpha	0.00694063418348	Controls the strength of the Boosting algorithm.
cols	2000	Number of output layer mini-columns. Note: this parameter was this was not evolved.
col_sparsity	1%	Output layer mini-column sparsity. Note: this parameter was this was not evolved.
pp	1394	Size of potential pool.
dec	0.00357969954117	Permanence decrement.
inc	0.0620456340463	Permanence increment.
thresh	0.138191824662	Connected synapse permanence threshold.
prox_segs	7	Number of proximal segments per mini-column.
min_stab	0.642926775309	Target output layer mini-column stability.

Table 1. Model Parameters.

Running the model yielded following results.

Metric	Before Training	After Training
Inter object overlap	0.011848	0.0093156
Intra object overlap	0.012044	0.12281
Object classification accuracy	1%	97.7%
Average synapses on proximal segments	4	60

Table 2. Model performance. The intra object overlap is significantly greater than the inter object overlap after training.

Experiment: I turned off the mechanism for stability and the model yielded the following results.

Metric	Before Training	After Training
Inter object overlap	0.011494	0.0092702
Intra object overlap	0.011899	0.0089444
Object classification accuracy	1.2%	96.7%
Average synapses on proximal segments	4	37

Table 3: Without the mechanism for stability, training does not increase the intra object overlap. Also the SP is able to classify the objects using fewer synapses.

Experiment: I reduce the number of proximal segments to one and the model yielded the following results.

Metric	Before Training	After Training
Inter object overlap	0.013396	0.0094570
Intra object overlap	0.013288	0.017488
Object classification accuracy	0.8%	97.3%
Average synapses on proximal segments	4	182

Table 4. With fewer proximal segments, each proximal segment takes on many more synapses. The intra object overlap is significantly reduced but is still greater than it is without the mechanism for stability.