Effect of Spatial Pooler Initialization on Column Activity in Hierarchical Temporal Memory

Mackenzie Leake[1], Liyu Xia[2], Kamil Rocki[3], Wayne Imaino[3]

1. Scripps College 2. University of Chicago 3. IBM Research

Abstract

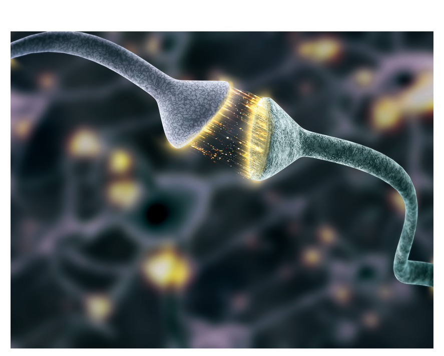
In the Hierarchical Temporal Memory (HTM) paradigm the effect of overlap between inputs on the activation of columns in the spatial pooler is studied. Numerical results suggest behavior such that overlapping inputs map to similar sets of columns and non-overlapping inputs map to distinct columns. We show that the spatial pooler exhibits this behavior under certain conditions for the synaptic and proximal thresholds.

HTM Background

Hierarchical Temporal Memory (HTM), a neurologically-inspired machine learning system originally proposed by Numenta (1), is a developing area of modern machine learning. Prior work has explored the use of HTM for

- land use recognition (2)
- speech recognition (3)
- content-based image retrieval (4)
- stock trading (5)

Our goal is to study the specific initial conditions for the parameters in order to achieve the desired behavior of grouping overlapping inputs into similar sets of columns and non-overlapping inputs into different columns.



Synapse (6)

0.1	0.4	0.4	0.1
0.2	0.1	0.5	0.6
0.1	0.7	0.9	0.8

Permanence Matrix

Fig. 1. Modeling synaptic activity through permanences.

Variables

Symbol	Name	Type	Description
x^k	input k	vector	encoded input
ov_{ij}	overlap between i & j	vector	overlap between inputs and/or columns
P	permanence	matrix	stores synaptic activity
C	connectivity	matrix	$P > \tau_C$
τ_C	synaptic threshold	scalar	converts P to C
τ_o	proximal threshold	scalar	controls column activation

Goal

How do we ensure different inputs have different representations?

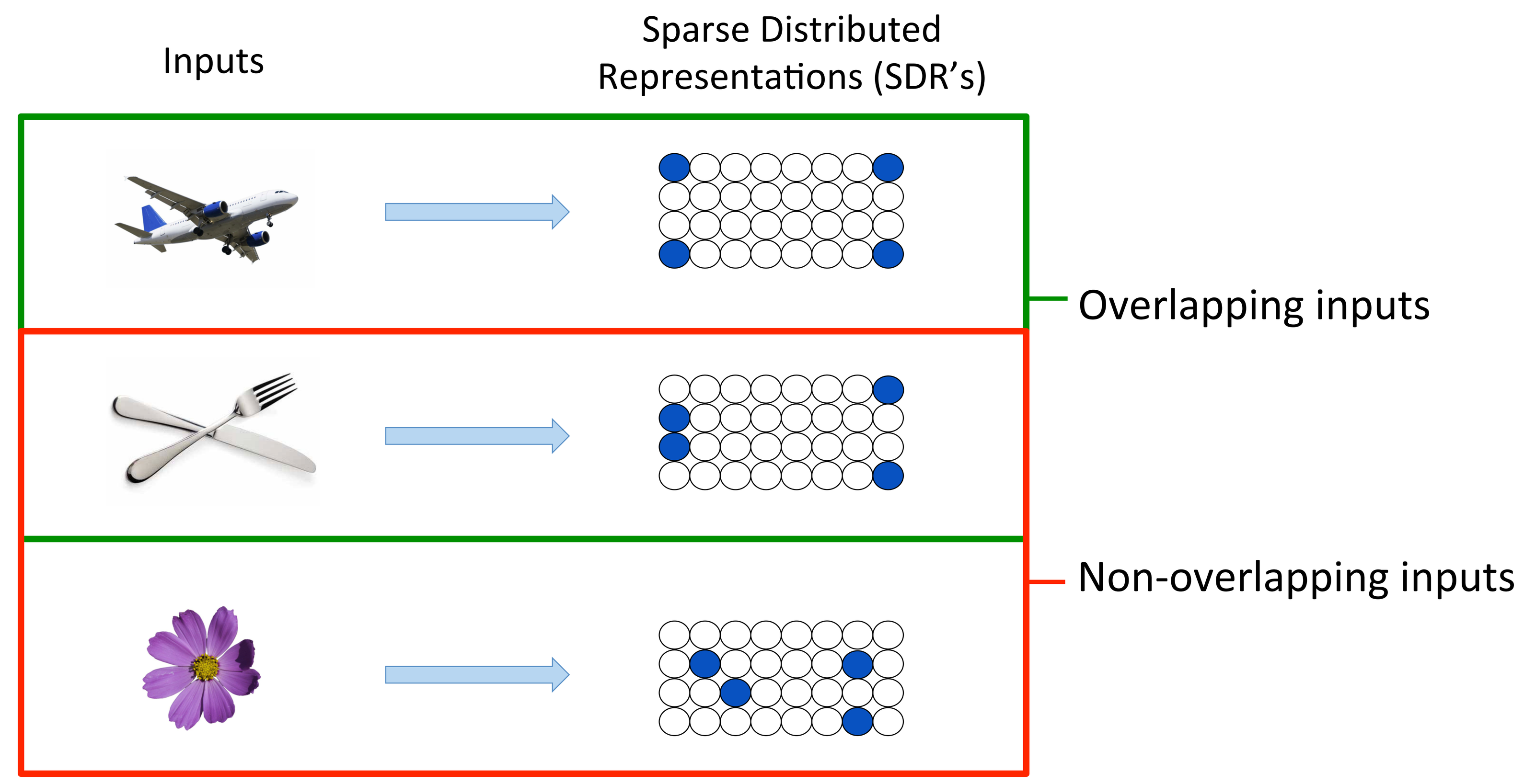


Fig. 2. SDR's for overlapping and non-overlapping inputs.

Non-Overlapping Inputs

Here we compute the expected value of overlap between x^b and column C_j

Let us denote the following assumptions by event Z:

- Each input is equally likely
- x^a activates C_j
- x^a and x^b are non-overlapping
- x^a=[1,…,1,0,…,0] with *d* 1's at the beginning

Using Bayes' theorem:

$$\mathbb{E}[\sum_{k=1}^{L} x_k^b C_{jk}|Z] = \sum_{k=d+1}^{L} \mathbb{P}[x_k^b = 1|C_{jk} = 1, Z] \times \mathbb{P}[C_{jk} = 1|Z]$$

Considering all possible x^b, the probability of each entry of x^b to be 1 is $\frac{d}{L-d}$, so:

$$= \sum_{k=d+1}^{L} \frac{d}{L-d} \times \mathbb{P}[C_{jk} = 1|Z] = d \times \mathbb{P}[C_{jk} = 1|Z]$$

The expected number of 1's within a column is $L(1-\tau_C)$, and the number of 1's left for the column to overlap with x^b is equal to $L(1-\tau_C)-ov_{ac}$, so:

$$\mathbb{E}[\sum_{k=1}^{L} x_k^b C_{jk}|Z] = d \times \frac{L(1-\tau_C) - ov_{ac}}{L-d}$$

The proximal threshold, τ_o, gives a loose upper bound for the expected overlap score between x^b and C_j. This results in:

$$d \times \frac{L(1-\tau_C) - \tau_o}{L-d} \leq \tau_o$$

$$\boxed{1 - \tau_C \leq \frac{1}{d}\tau_o.}$$

The preceding relationship provides general guidelines for the initial selection of τ_C and τ_o values that on average will prohibit input x^b from activating column C_j, given that x^a already activates C_j.

Note: The overlapping case is a generalization of the non-overlapping case.

Incremental Learning

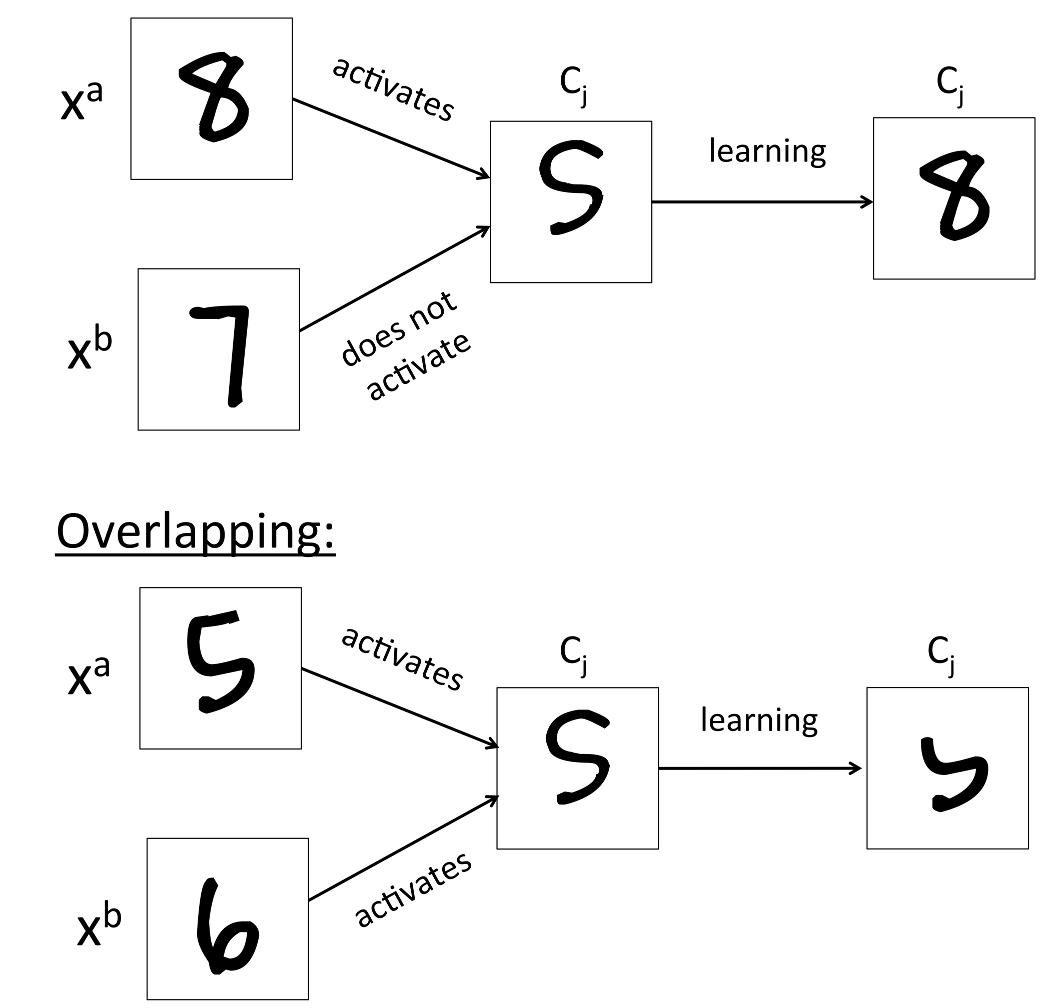


Fig. 3. Effects of learning on column activity.

Summary

We show that with careful initialization of parameters the spatial pooler maps distinct inputs into distinct SDR's as indicated by the column activity. Assuming appropriate parameter selection, we make observations about the learning dynamics for both non-overlapping and overlapping inputs.

Future Work

- Transition from probabilistic view to analysis of sparse distributed representations (SDR's)
- Examination of the properties of the spatial pooler as a map from SDR's to SDR's in a lower dimensional space
- Numerical experiments for the effect of learning on the spatial pooler

References

[1] J. Hawkins, et al.,"Hierarchical temporal memory Including HTM cortical learning algorithms," Ver. 0.2.1, 2011.
[2] A.J. Perea, J.E. Merono, and M.J. Aguilera, "Application of Numenta hierarchical temporal memory for land-use classification," S. Afr. J. Sci., vol. 105, pp. 370-375, Sept./Oct. 2009.
[3] J. van Doremalen and L. Boves, "Spoken digit recognition using a hierarchical temporal memory," presented at Interspeech, pp. 2566-2569, Brisbane, Australia, Sept. 22-26, 2008.
[4] B. Bobier and M. Wirth, "Content-based image retrieval using hierarchical temporal memory," in Proc. 16th ACM Int. Conf. on Multimedia, 2008, pp. 925-928.
[5] P. Gabrielsson, R. Konig, and U. Johansson, "Evolving Hierarchical Temporal Memory-Based Trading Models," EvoApplications 2013, Applications of Evolutionary Computing, Vienna April 3-5, 2013.
[6] R. Breslin, Synapse. Digital image. Bluesci. Cambridge University, n.d. Web. 30 July 2014.