Binding of Sparse Distributed Representations in Hierarchical Temporal Memory

Luke Boudreau
Advisor: Dhireesha Kudithipudi
Neuromorphic AI Lab
Rochester Institute of Technology
Hierarchical Temporal Memory

HTM Structure

• Pyramidal neurons (cells) form Columns
• A collection of Columns form a Region, which models a Cortical Column or layers 2/3 of the Neocortex

HTM theory is continually evolving

• HTM has two major algorithm portions
• Spatial Pooler - processes Feedforward Input
• Temporal Memory - processes Feedforward & Contextual Input
• Numenta (Jeff Hawkins et. al) is driving the HTM research

Source: Numenta
**HTM Spatial Pooler**

**Spatial Pooler Structure**
- *Columns of cells* become active with sufficient feedforward input
- Each *cell* in a *column* receives the same feedforward input
- Input flows through *proximal synapses*
- Hebbian learning governs connectivity of *proximal synapses*
- Online learning - adapts to changes in input data
Sparse Distributed Representation (SDR)

Sparse Distributed Representations are core part of HTM
- Large Binary Vector (2048 bits; the number of columns in the Region)
- Sparsity ~ 2% (low # of ‘1’s)
- Distributed: non-localist and resilient to noise

Spatial Pooler learns to map similar inputs to similar outputs
- Two inputs that are similar should have some degree of overlap in their SDRs
- Similarity between SDRs can be computed with a dot product
Spatial Pooler Algorithm

1. Overlap
2. Inhibition
3. Learning

Notation:
- Input $\mathbf{X} \in \{0, 1\}^{m \times q}$
- Proximal Synapses $\Phi \in \mathbb{R}^{m \times q}$
- Connected Synapses $\mathbf{Y} \equiv I(\Phi_i \geq \rho_s) \ \forall i$

\[ \alpha \equiv \begin{cases} \alpha_i b_i, & \alpha_i \geq \rho_d, \\ 0, & \text{otherwise} \end{cases} \ \forall i \]

\[ \tilde{\alpha}_i \equiv \mathbf{X}_i \cdot \mathbf{Y}_i \]

\[ \tilde{\gamma} \equiv \max(k_{\max} (H \odot \tilde{\alpha}, \rho_c), 1) \ \forall i \]

\[ \tilde{c} \equiv I(\tilde{\alpha}_i \geq \tilde{\gamma}_i) \ \forall i \]

\[ \tilde{\alpha} \equiv \text{clip} (\Phi \odot \delta \Phi, 0, 1) \]

\[ \delta \Phi \equiv \mathbf{c}^T \odot (\phi_+ \mathbf{X} - (\phi_- \mathbf{X})) \]

Motivation

• Numenta’s research is focused on bioplausibility and emulation of the pathways in the Neocortex

• Binding operation is the basis for *Content Addressable Memory*, which could help facilitate long term storage and retrieval of SDRs

• Combine multimodal data without increasing dimensionality
Background on Binding

Binding problem for language and vision (Jackendoff 2003, Malsburg 1995)
• Vector Symbolic Architectures (Proposed by Gayler 2004)
• Vector operations (Binding, Superposition, and Permutation)
• Vectors are the same dimensionality; both atomic and complex representations

Different Implementations of VSA
• Holographic Reduced Representations (HRR)- Plate
• Binary Spatter Codes / Hyperdimensional Computing – Kanerva
• Multiply, Add, Permute (MAP) – Gayler
(Laiho & Kanerva 2015) proposed a binding operation for sparse binary distributed representations.

- Segmentation of sparse binary vector

- *Maximally Sparse* – requires controlled density

- Generalization of circular convolution for Plate’s HRRs in the frequency domain.

---

Binding

\[ C = A \otimes B \]

\[
A = 
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\]

\[
B = 
\begin{bmatrix}
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

\[
C = 
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\]

\[
a_s = \min(\arg\max_j (A_{(s,j)}))
\]

\[
C(s,j) \equiv B(s,j - \alpha a_s (L))
\]

\[
N = 20
\]

\[
L = 5
\]

\[
S = \frac{N}{L}
\]
Unbinding

B = A \bigcirc C

\begin{align*}
A & \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \\
C & \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \\
B & \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0
\end{align*}

N = 20
L = 5
S = N/L
Spatial Pooler – Maximally Sparse SDRs

Use Local Inhibition in the Spatial Pooler Algorithm

• Use neighborhood masks, $H$, to establish segments in the HTM Region
• Select the most active column in each neighborhood/segment instead of top $k$ columns
• Initialization of column to input connectivity is based on the dataset. Requires a mapping that preserves the distributed nature (topology) of SDRs
FREAK – Fast Retina Keypoint

FREAK Encoding

- 512 bit binary descriptor
- Difference of Gaussians (No Backprop!)
- Coarse to fine grained information
- Inspired by the retina, but no biological significance

Encoding small NORB

Spatial Pooler – Initialization

Significant Hyperparameters:
- 2048 columns (mini-columns)
- 64 active bits (~3% Sparsity)
- 1 epoch of learning on NORB training set

HTM Columns

![Segment 1 input mapping graph]

Coarse  FREAK Descriptor  Fine
Spatial Pooler – Learning Metrics

Spatial Pooler performance while training:
- Local Inhibition has less column participation because of the stricter inhibition rule
- Local Inhibition doesn’t fully stabilize after one epoch because of random selection between columns with equal overlap/activation levels.
Spatial Pooler – Cluster Analysis with t-SNE

Entangled Nature of SDRs:
- Global Inhibition produces more entangled representations than the raw FREAK Descriptors
- Local Inhibition helps preserve more of the class similarity (i.e. larger and identifiable clusters)
Binding Experiments

Binding Location & Feature SDRs

• Bind each features SDR with a location SDR (Supervised)
• Randomly generate 18 *maximally sparse* SDRs for each possible azimuth

Binding & Superposition

• Integrate the bound representation as a rotation around the object (repeat for each elevation)
• Superposition of SDRs (Logical OR)

\[
Rotation SDR = \sum_{a=0}^{17} C_a
\]
• Same Dimensionality and Sparseness as component SDRs
• These SDRs are Content Addressable:
  • Unbind with an azimuth SDR => “What feature(s) is located at this Azimuth?”
  • Unbind with a feature SDR => “What azimuth(s) is associated with this feature?”
Left: Superposition of Features without binding
- Larger density (~16% active bits)
- SDRs from different class/instances are similar (more overlapping bits)

Right: Superposition of Features bound with corresponding azimuth.
- Much larger density (~50% active bits)
- Binding gives more structure to superimposed representations
Conclusions

Basis for Content Addressable Memory

• Binding can give structure to entangled representations, but requires a supervised approach
• Density of superimposed representations will be a limiting factor for recall or retrieving vectors from superimposed SDRs
• Novel vector, non-similar to either component

Local Inhibition

• Minor modification to inhibition in Spatial Pooler Algorithm