SYDE 556/750

**Simulating Neurobiological Systems
Lecture 10: Symbols and Symbol-like
Representations**
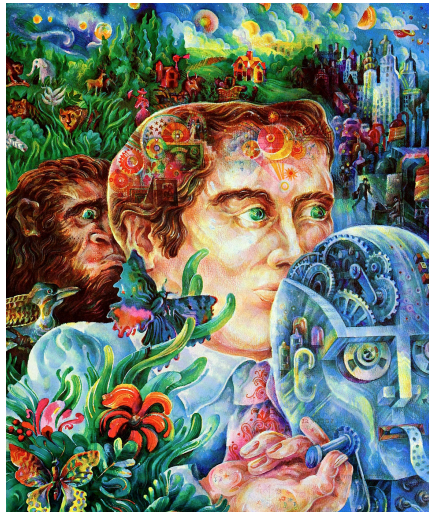
Andreas Stöckel

March 10 & 12, 2020

UNIVERSITY OF
WATERLOO | FACULTY OF
ENGINEERING

Classical Representation of Knowledge

▶ "The number eight comes after the number nine":
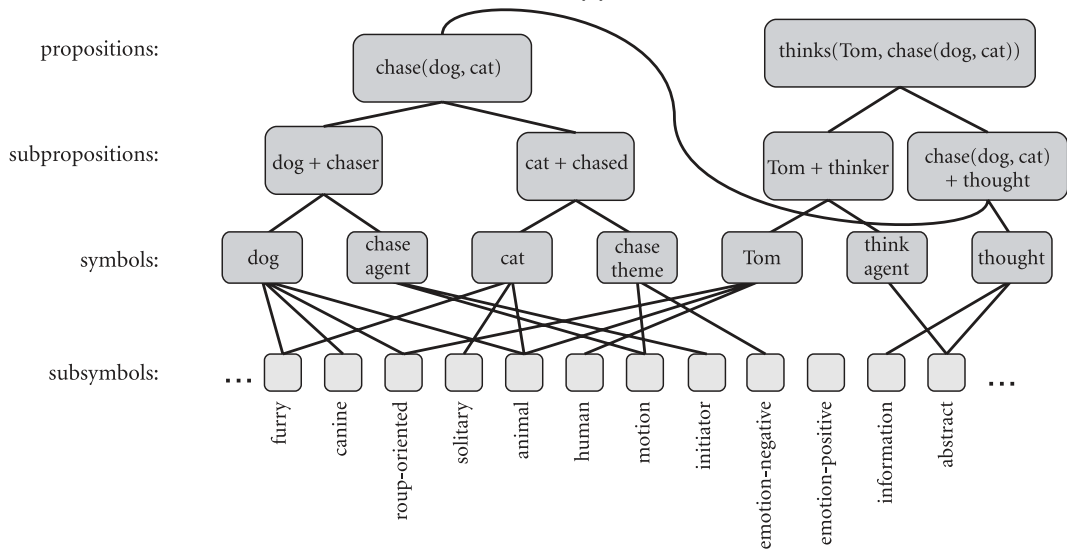
$$\textbf{isSucc}(\texttt{EIGHT}, \texttt{NINE}).$$

▶ "All dogs chase cats":

$$\forall x \forall y \left(\textbf{isDog}(x) \wedge \textbf{isCat}(y)\right) \rightarrow \textbf{doesChase}(x, y).$$

▶ "Anne knows that Bill thinks that Charlie likes Dave":

$$\textbf{knows}\Big(\texttt{ANNE}, \text{``}\textbf{thinks}(\texttt{BILL}, \text{`}\textbf{likes}(\texttt{CHARLIE}, \texttt{DAVE})\text{')''}\Big).$$
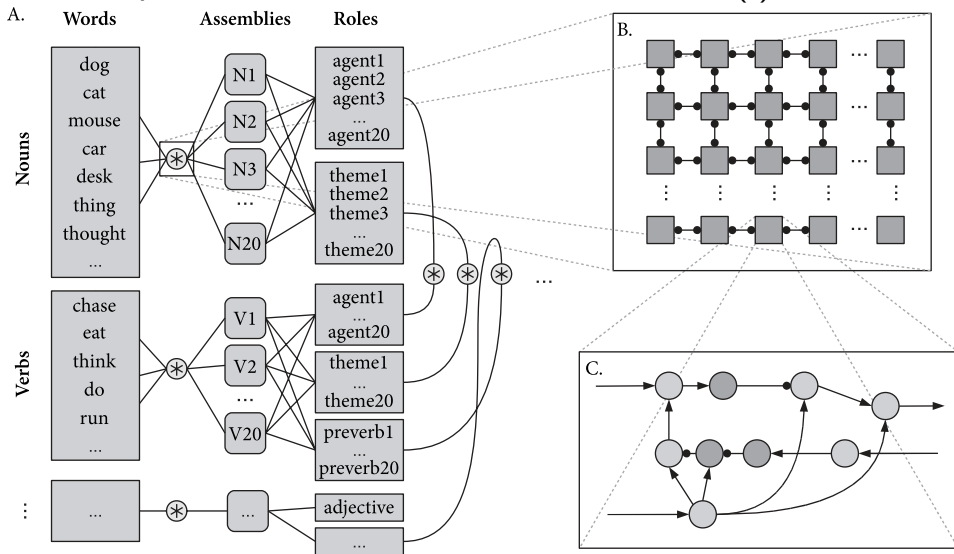
Solution Attempt 1: Neural Synchrony (I)

Solution Attempt 1: Neural Synchrony (II)

➕ Solves the binding problem

🟡 Localist representation

🟡 Unclear how to solve problems 2 to 4

➖ Unclear how these oscillations are generated and controlled

➖ Unclear how the representations are processed

➖ Exponential explosion of neurons required to represent concepts

Solution Attempt 2: Neural Blackboard Architecture (I)

A. Words, Assemblies, Roles

Nouns: dog, cat, mouse, car, desk, thing, thought, ...
N1, N2, N3, ..., N20
agent1, agent2, agent3, ..., agent20
theme1, theme2, theme3, ..., theme20

Verbs: chase, eat, think, do, run, ...
V1, V2, ..., V20
agent1, ..., agent20
theme1, ..., theme20
preverb1, ..., preverb20

adjective

B.

C.

Solution Attempt 2: Neural Blackboard Architecture (II)

🟢 Fewer resources than LISA

🟢 Solves all four of Jackendoffs challenges (according to the authors)

🟢 Explains limitations of human sentence representation

🟡 (At least partially) localist representation

🔴 Particular structure; does not match biology

🔴 Large number of neurons; about 500×10^6 to represent sentences

🔴 Only considers *representation*, no control structures

Solution Attempt 3: Vector Operators

Idea: High-dimensional vectors $\mathbf{x} \in \mathbb{R}^d$ represent symbols; bind using tensor product
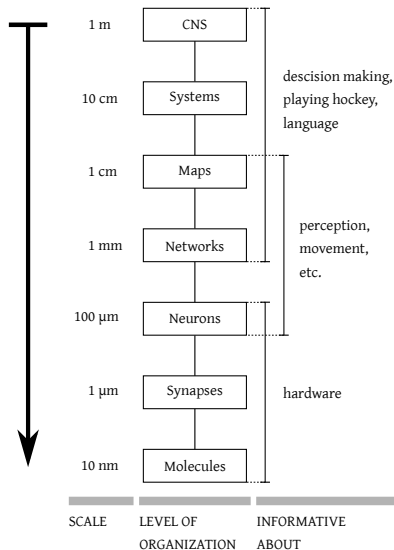
$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix} \qquad \text{(Outer product)}$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} & a_{12} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \\ a_{21} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} & a_{22} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \end{pmatrix} \qquad \text{(Tensor product)}$$

$$= \begin{pmatrix} a_{11} b_{11} & a_{11} b_{12} & a_{12} b_{11} & a_{12} b_{12} \\ a_{11} b_{21} & a_{11} b_{22} & a_{12} b_{21} & a_{12} b_{22} \\ a_{21} b_{11} & a_{21} b_{12} & a_{22} b_{11} & a_{22} b_{12} \\ a_{21} b_{21} & a_{21} b_{22} & a_{22} b_{21} & a_{22} b_{22} \end{pmatrix}$$

🔴 Scales extremely poorly d^n for n binding operations

A Deeper Problem: Cognitive Science vs. Neuroscience

▶ Trying very hard to map purely symbolic architectures onto neurons.

▶ Neural aspects are treated as *mere implementation details*.

▶ Instance of **top-down modelling**:
High-level cognitive architectures are mapped onto biology.

▶ Hope of many cognitive scientists:
If successful, **neurons do not matter**.

SCALE	LEVEL OF ORGANIZATION	INFORMATIVE ABOUT
1 m	CNS	
10 cm	Systems	descision making, playing hockey, language
1 cm	Maps	
1 mm	Networks	perception, movement, etc.
100 μm	Neurons	
1 μm	Synapses	hardware
10 nm	Molecules	

Binding Operator Properties

i. Preservation of Dimensionality

$$\circledast : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$$

ii. Approximately Reversible

$$\mathbf{x} \approx (\mathbf{x} \circledast \mathbf{y}) \circledast \mathbf{y}^{-1}$$

iii. Dissimilar to Inputs

$$0 \approx \langle \mathbf{x} \circledast \mathbf{y}, \mathbf{x} \rangle, 0 \approx \langle \mathbf{x} \circledast \mathbf{y}, \mathbf{y} \rangle$$

Sentence Encoding Revisited

▶ "The number eight comes after the number nine":

$$\text{NUMBER} \circledast \text{EIGHT} + \text{SUCC} \circledast \text{NINE} \,.$$

▶ "The dog chases the cat":

$$\text{DOG} \circledast \text{SUBJ} + \text{CAT} \circledast \text{OBJ} + \text{CHASE} \circledast \text{VERB} \,.$$

▶ "Anne knows that Bill thinks that Charlie likes Dave":

$$\text{SUBJ} \circledast \text{ANNE} + \text{ACT} \circledast \text{KNOWS} + \text{OBJ} \circledast$$

$$\Big(\text{SUBJ} \circledast \text{BILL} + \text{ACT} \circledast \text{THINKS} + \text{OBJ} \circledast$$

$$\big(\text{SUBJ} \circledast \text{CHARLIE} + \text{ACT} \circledast \text{LIKES} + \text{OBJ} \circledast \text{DAVE} \big) \Big) \,.$$

⚠ **Compression of information; graceful degradation; depends on d**

Using the Reversibility Property to Answer Questions

▶ "A blue square and a red circle:"

$$\mathbf{x} = \text{BLUE} \circledast \text{SQUARE} + \text{RED} \circledast \text{CIRCLE}.$$

▶ "Which object is blue?"

$$\mathbf{y} = \left(\text{BLUE} \circledast \text{SQUARE} + \text{RED} \circledast \text{CIRCLE}\right) \circledast \text{BLUE}^{-1}$$

$$= \left(\text{BLUE} \circledast \text{SQUARE}\right) \circledast \text{BLUE}^{-1} + \left(\text{RED} \circledast \text{CIRCLE}\right) \circledast \text{BLUE}^{-1}$$

$$\approx \text{SQUARE} + \underbrace{\text{RED} \circledast \text{CIRCLE} \circledast \text{BLUE}^{-1}}_{\text{"noise"}}$$

$$\approx \text{SQUARE}.$$

⚠ **Supposes that there is a set of valid symbols ⇒ "Cleanup Memory"**

VSAs: Potential Binding Operators (I)

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \qquad \text{(XOR)}$$

$$\begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} \odot \begin{pmatrix} E \\ F \\ G \\ H \end{pmatrix} = \begin{pmatrix} AE \\ BF \\ CG \\ DH \end{pmatrix} \qquad \text{(Hadamard Product)}$$

VSAs: Potential Binding Operators (II)

$$\begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} \circledast \begin{pmatrix} E \\ F \\ G \\ H \end{pmatrix} = \begin{pmatrix} AE & + & BH & + & CG & + & DF \\ AF & + & BE & + & CH & + & DG \\ AG & + & BF & + & CE & + & DH \\ AH & + & BG & + & CF & + & DE \end{pmatrix} \quad \text{(Circular Convolution)}$$

Circular Convolution is a "compressed" outer product:

$$\begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} \otimes \begin{pmatrix} E \\ F \\ G \\ H \end{pmatrix} = \begin{pmatrix} AE & AF & AG & AH \\ BE & BF & BG & BH \\ CE & CF & CG & CH \\ DE & DF & DG & DH \end{pmatrix} \quad \text{(Outer Product)}$$

Circular Convolution: Dissimilarity and Reversibility
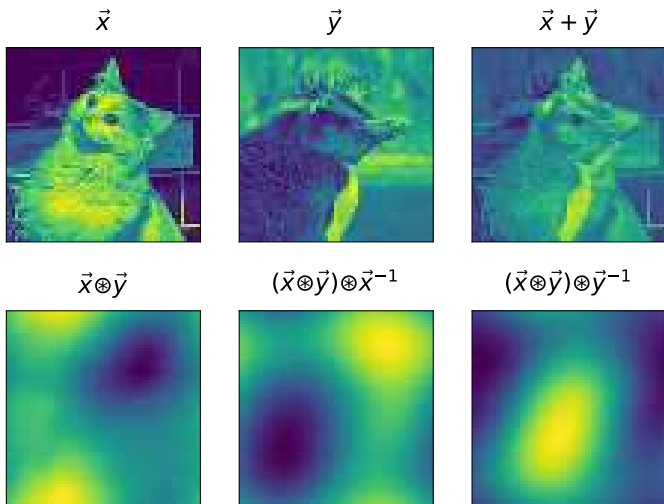
$\vec{x}$ $\vec{y}$ $\vec{x} + \vec{y}$
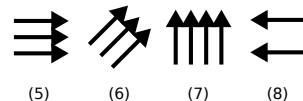
$\vec{x} \circledast \vec{y}$ $(\vec{x} \circledast \vec{y}) \circledast \vec{x}^{-1}$ $(\vec{x} \circledast \vec{y}) \circledast \vec{y}^{-1}$

Raven's Progressive Matrices (I)

Raven's Progressive Matrices (II)

1	*11*	*111*
4	*44*	*444*
5	*55*	
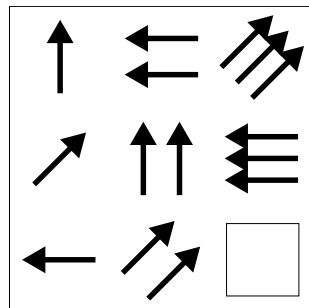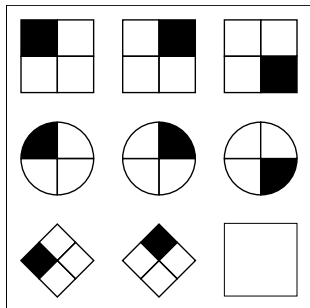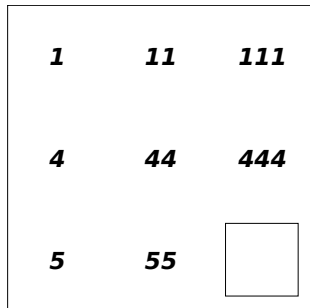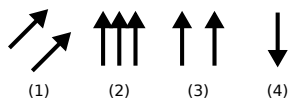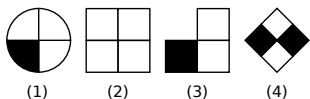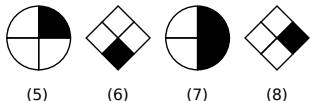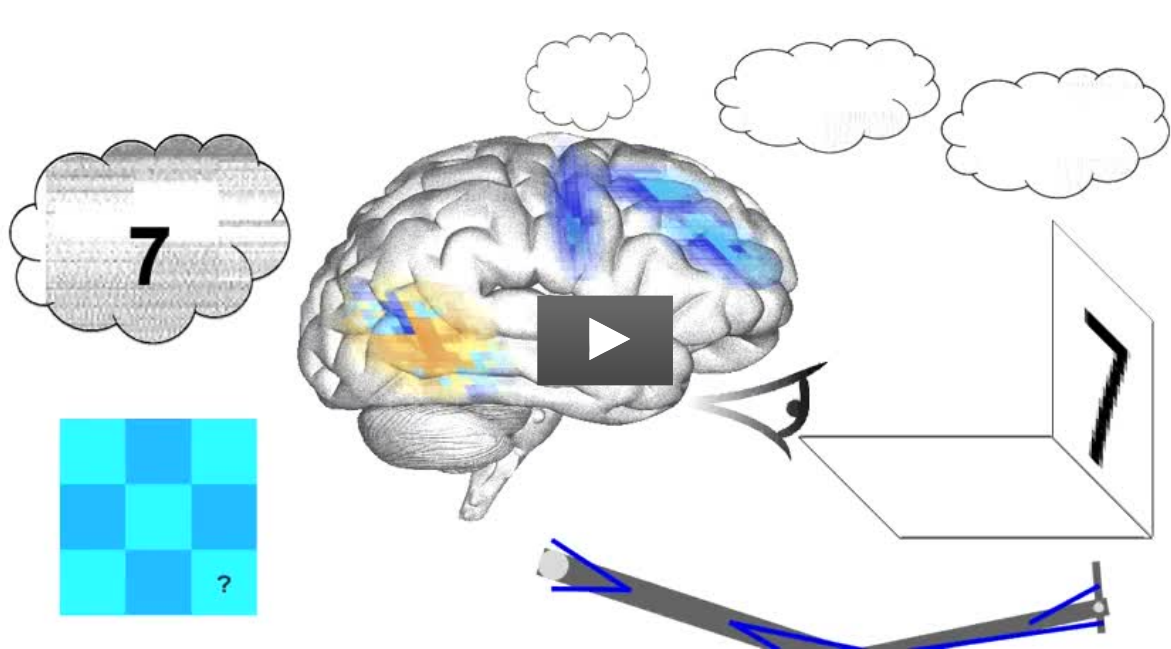
3	*55*	*111*	*44*
(1)	(2)	(3)	(4)
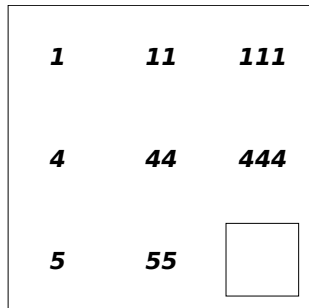
444	*555*	*999*	*33*
(5)	(6)	(7)	(8)

Representing cells:

$$C1 = \text{ONE} \circledast P1,$$
$$C2 = \text{ONE} \circledast P1 + \text{ONE} \circledast P2,$$
$$C3 = \text{ONE} \circledast P1 + \text{ONE} \circledast P2 + \text{ONE} \circledast P3,$$
$$C4 = \text{ONE} \circledast P1 + \text{ONE} \circledast P2 + \text{ONE} \circledast P3,$$
$$C5 = \text{FOUR} \circledast P1 + \text{FOUR} \circledast P2,$$
$$C6 = \text{FOUR} \circledast P1 + \text{FOUR} \circledast P2 + \text{FOUR} \circledast P3,$$
$$C7 = \text{FIVE} \circledast P1,$$
$$C8 = \text{FIVE} \circledast P1 + \text{FIVE} \circledast P2.$$

Raven's Progressive Matrices (III)

1	*11*	*111*
4	*44*	*444*
5	*55*	

3	*55*	*111*	*44*
(1)	(2)	(3)	(4)

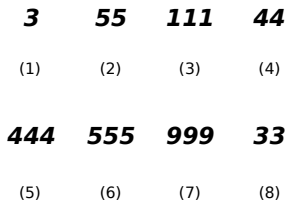| *444*	*555*	*999*	*33*
(5)	(6)	(7)	(8)
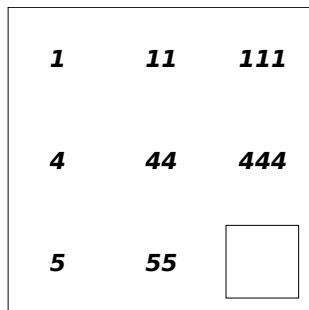
Extracting the horizontal rule:

$$T1 = C2 \circledast C1^{-1}, \qquad T4 = C6 \circledast C5^{-1},$$
$$T2 = C3 \circledast C2^{-1}, \qquad T5 = C8 \circledast C7^{-1},$$
$$T3 = C5 \circledast C4^{-1}.$$

$$T = \frac{T1 + T2 + T3 + T4 + T5}{5}.$$

Making a prediction:

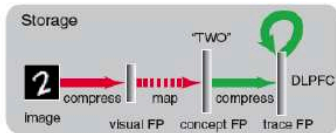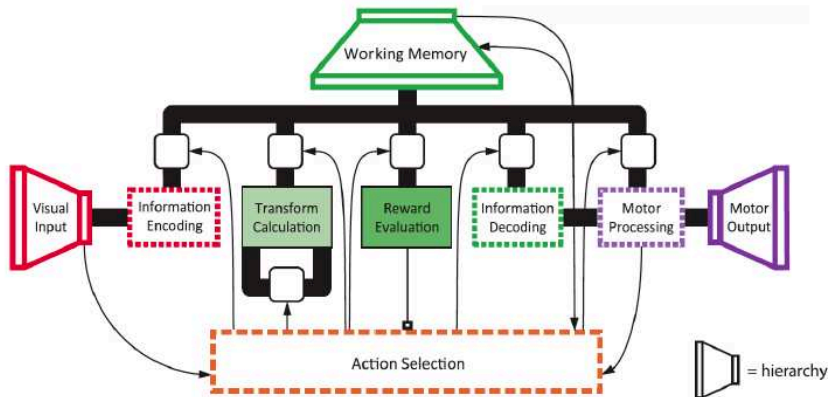$$C9 = C8 \circledast T$$
$$\approx FIVE \circledast P1 + FIVE \circledast P2 + FIVE \circledast P3.$$
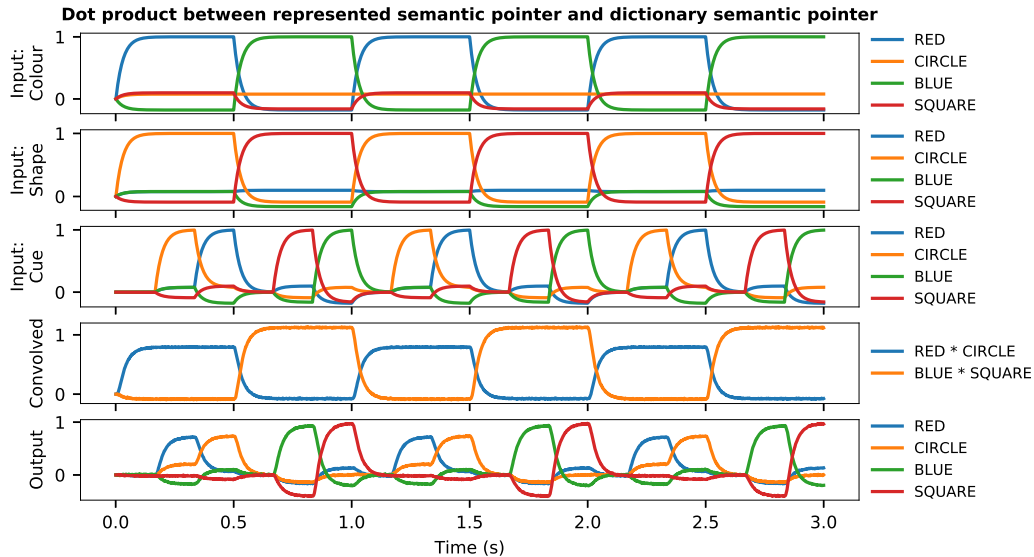
Nengo SPA Example

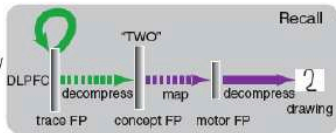Dot product between represented semantic pointer and dictionary semantic pointer

Spaun – Semantic Pointer Architecture Unified Network (I)

Spaun – Semantic Pointer Architecture Unified Network (II)
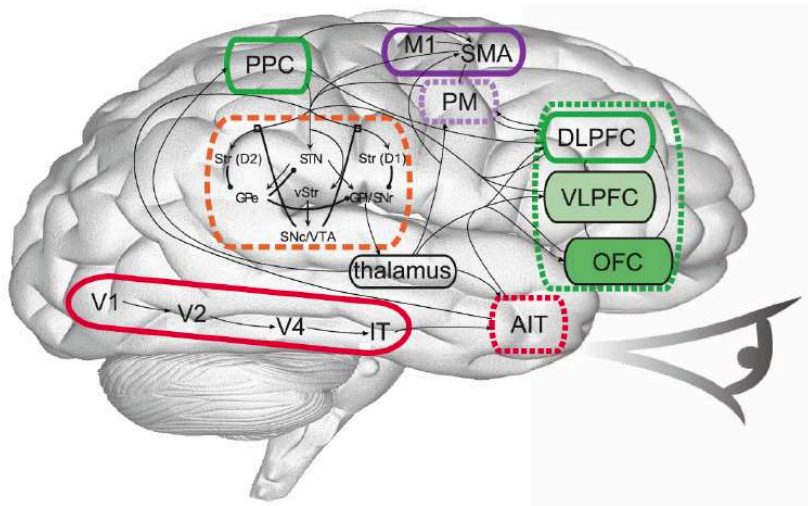
Image sources

Title slide
Bell telephone magazine, 1922, American Telephone and Telegraph Company
Wikimedia.