

SYDE 556/750

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

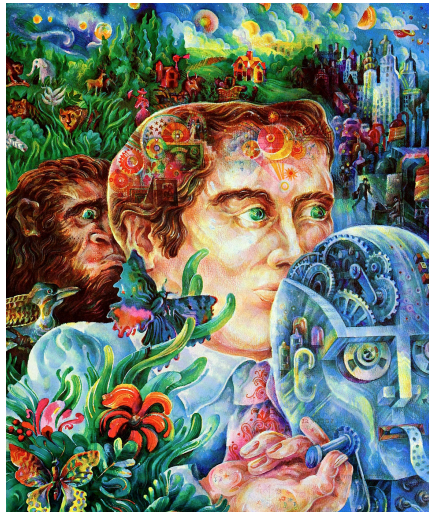
Andreas Stöckel

March 10 & 12, 2020



UNIVERSITY OF
WATERLOO

FACULTY OF
ENGINEERING



Classical Representation of Knowledge

- ▶ “The number eight comes after the number nine”:

isSucc(EIGHT, NINE) .

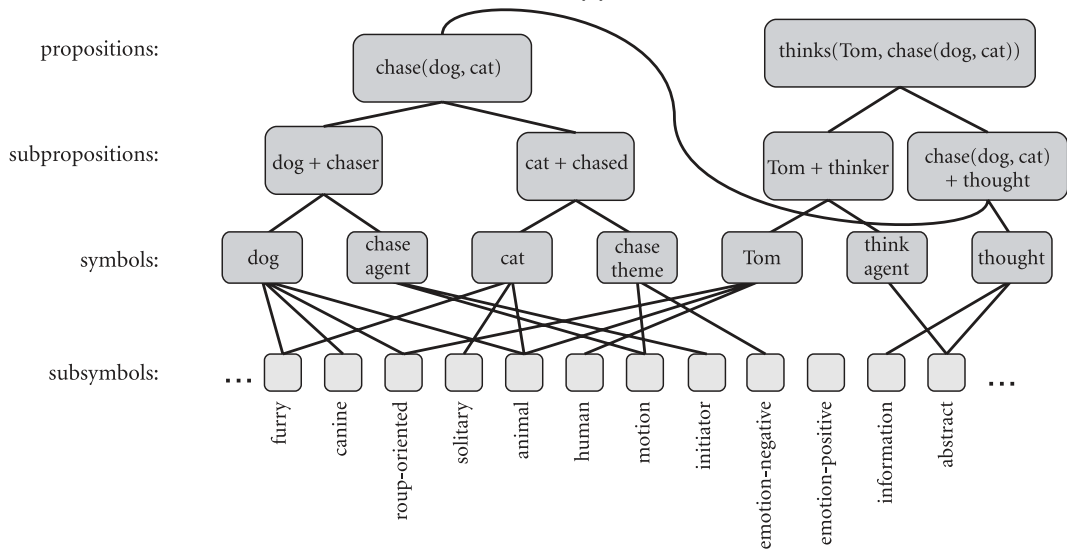
- ▶ “All dogs chase cats”:

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

- ▶ “Anne knows that Bill thinks that Charlie likes Dave”:

knows(ANNE, “**thinks**(BILL, ‘**likes**(CHARLIE, DAVE)’)”) .

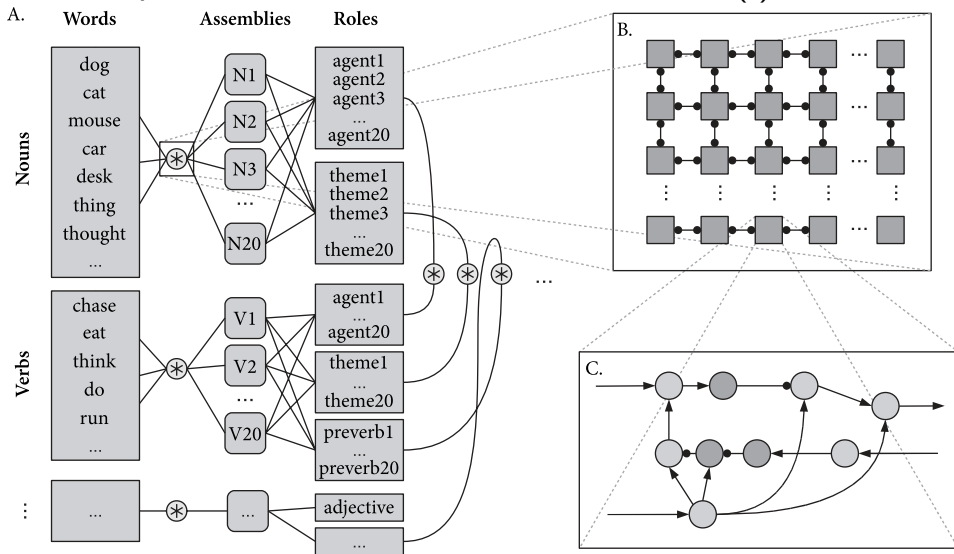
Solution Attempt 1: Neural Synchrony (I)



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)



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

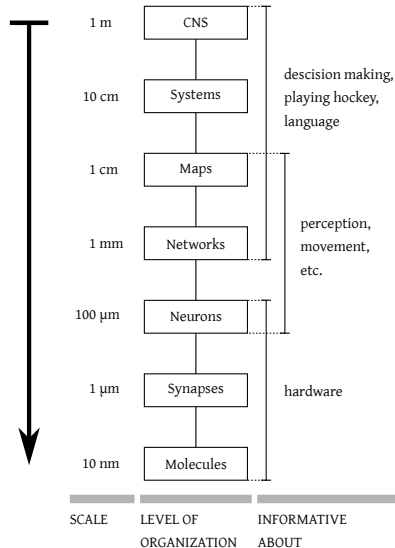
$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix} \quad (\text{Outer product})$$

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

⊖ 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**.



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”:

NUMBER * EIGHT + SUCC * NINE .

- ▶ “The dog chases the cat”:

DOG * SUBJ + CAT * OBJ + CHASE * VERB .

- ▶ “Anne knows that Bill thinks that Charlie likes Dave”:

SUBJ * ANNE + ACT * KNOWS + OBJ *
 (SUBJ * BILL + ACT * THINKS + OBJ *
 (SUBJ * CHARLIE + ACT * LIKES + OBJ * DAVE)) .



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?”

$$\begin{aligned} \mathbf{y} &= (\text{BLUE} \circledast \text{SQUARE} + \text{RED} \circledast \text{CIRCLE}) \circledast \text{BLUE}^{-1} \\ &= (\text{BLUE} \circledast \text{SQUARE}) \circledast \text{BLUE}^{-1} + (\text{RED} \circledast \text{CIRCLE}) \circledast \text{BLUE}^{-1} \\ &\approx \text{SQUARE} + \underbrace{\text{RED} \circledast \text{CIRCLE} \circledast \text{BLUE}^{-1}}_{\text{“noise”}} \\ &\approx \text{SQUARE}. \end{aligned}$$



Supposes that there is a set of valid symbols \Rightarrow “Cleanup Memory”

VSAs: Potential Binding Operators (I)

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

(XOR)

$$\begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} \odot \begin{pmatrix} E \\ F \\ G \\ H \end{pmatrix} = \begin{pmatrix} AE \\ BF \\ CG \\ DH \end{pmatrix}$$

(Hadamard Product)

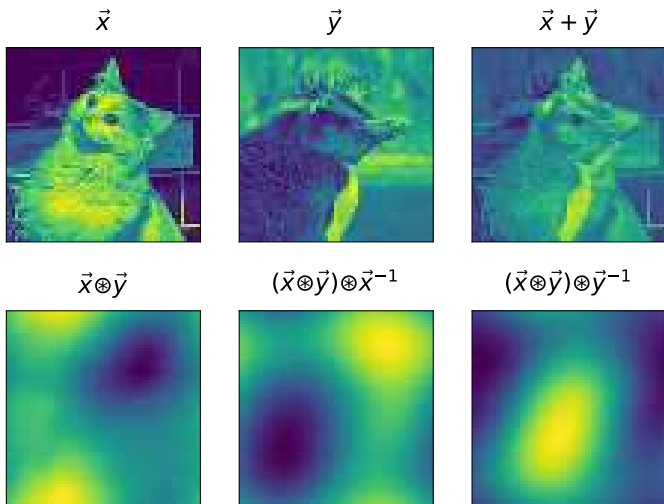
VSAs: Potential Binding Operators (II)

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

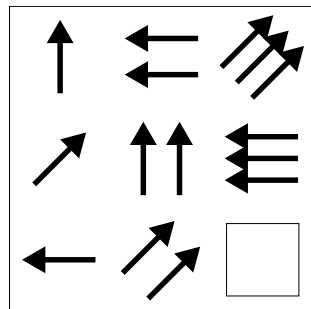
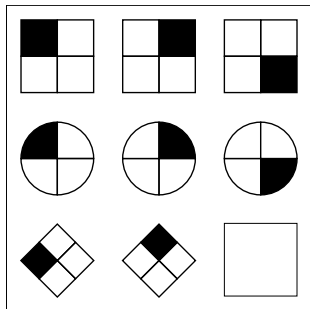
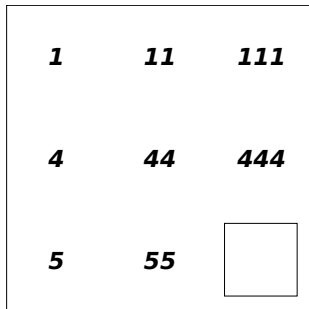
Circular Convolution is a “compressed” outer product:

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

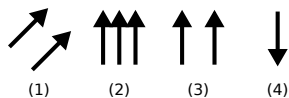
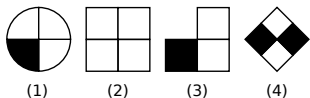
Circular Convolution: Dissimilarity and Reversibility



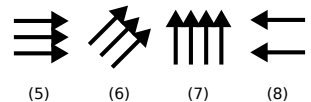
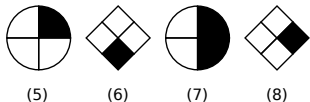
Raven's Progressive Matrices (I)

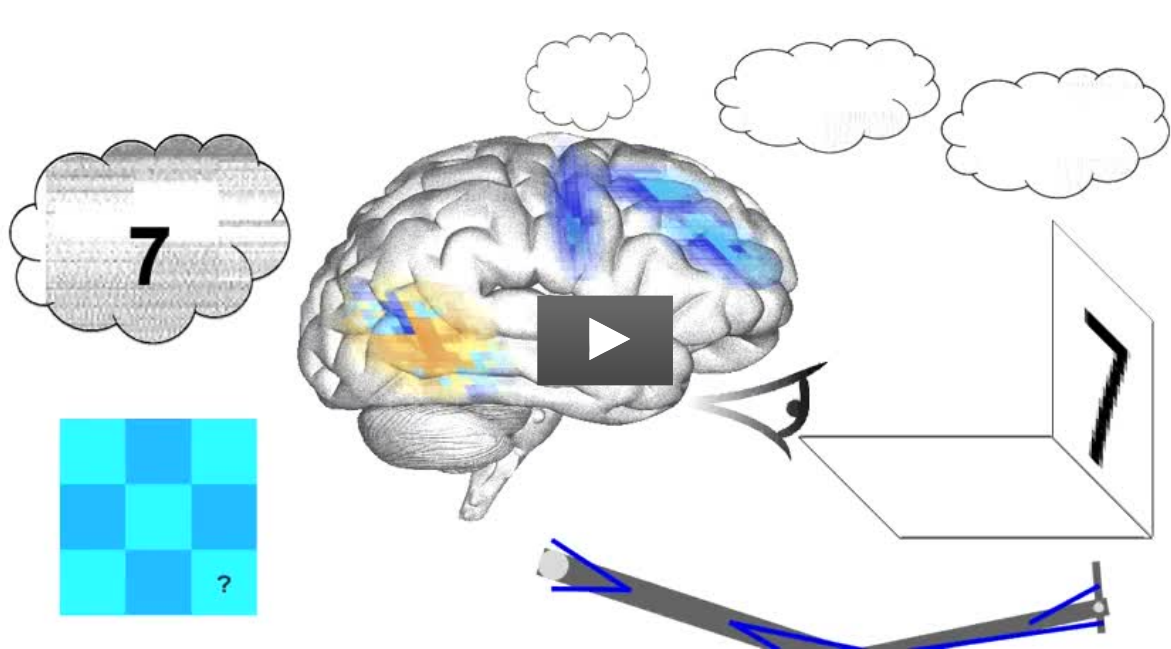


3	55	111	44
(1)	(2)	(3)	(4)



444	555	999	33
(5)	(6)	(7)	(8)





Raven's Progressive Matrices (II)

1	11	111
4	44	444
5	55	<div></div>

Representing cells:

$$C1 = \text{ONE} * P1,$$

$$C2 = \text{ONE} * P1 + \text{ONE} * P2,$$

$$C3 = \text{ONE} * P1 + \text{ONE} * P2 + \text{ONE} * P3,$$

$$C4 = \text{ONE} * P1 + \text{ONE} * P2 + \text{ONE} * P3,$$

$$C5 = \text{FOUR} * P1 + \text{FOUR} * P2,$$

$$C6 = \text{FOUR} * P1 + \text{FOUR} * P2 + \text{FOUR} * P3,$$

$$C7 = \text{FIVE} * P1,$$

$$C8 = \text{FIVE} * P1 + \text{FIVE} * P2.$$

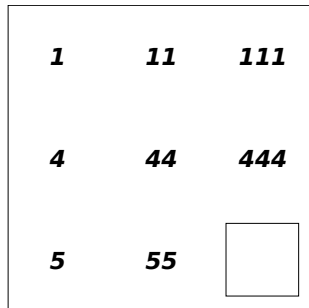
3 55 111 44

(1) (2) (3) (4)

444 555 999 33

(5) (6) (7) (8)

Raven's Progressive Matrices (III)



Extracting the horizontal rule:

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

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

Making a prediction:

$$\begin{aligned} C9 &= C8 \circledast T \\ &\approx \text{FIVE} \circledast P1 + \text{FIVE} \circledast P2 + \text{FIVE} \circledast P3. \end{aligned}$$

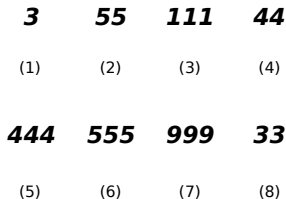


Image sources

Title slide

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