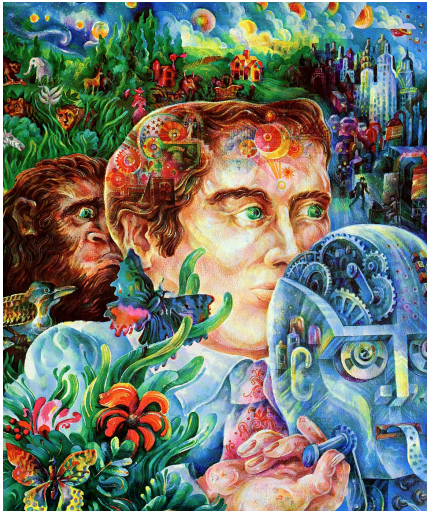
SYDE 556/750

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

Chris Eliasmith

November 6 & 11, 2024

- Slide design: Andreas Stöckel
- Content: Terry Stewart, Andreas Stöckel, Chris Eliasmith



UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Classical Representation of Knowledge

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

$$\mathbf{isSucc}(\texttt{NINE}, \texttt{EIGHT})\,.$$

- "All dogs chase cats":

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

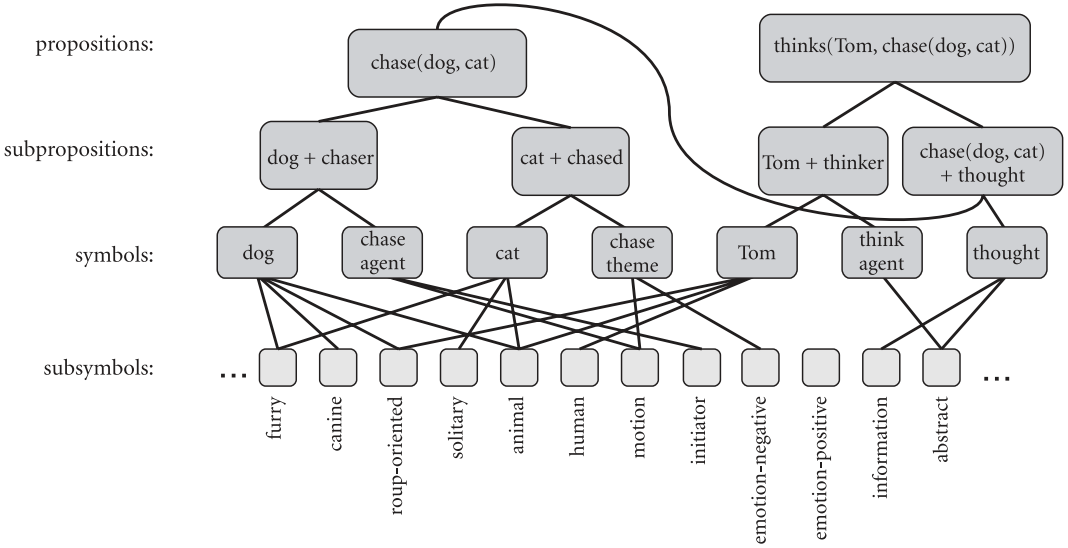
- "Anne knows that Bill thinks that Charlie likes Dave":

$$\mathbf{knows}\Big(\texttt{ANNE}, \text{``}\mathbf{thinks}\big(\texttt{BILL}, \text{`}\mathbf{likes}(\texttt{CHARLIE}, \texttt{DAVE})\text{'}\big)\text{''}\Big)\,.$$

Jackendoff's Challenges

- The Binding Problem
- The Problem of Two
- The Problem of Variables
- Working Memory versus Long-Term Memory

Solution Attempt 1: Neural Synchrony (I)

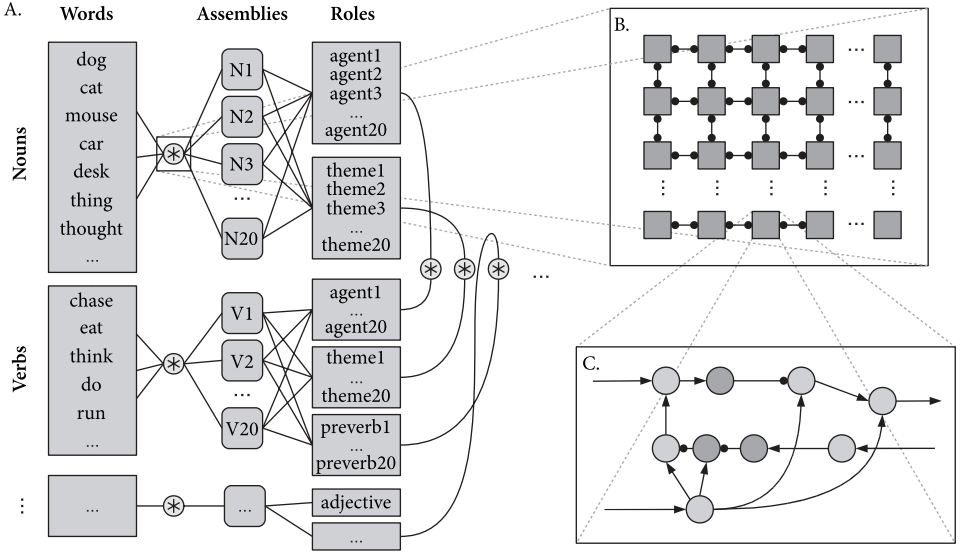


Solution Attempt 1: Neural Synchrony (II)

- ⊕ 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
- ⊖ Specific neural structure; does not match biology
- ⊖ Large number of neurons; about 500×10^6 to represent simple sentences
- ⊖ Only considers *representation*, no control structures

Solution Attempt 3: Vector Operators (I)

Idea: High-dimensional vectors $\mathbf{x} \in \mathbb{R}^d$ represent symbols; bind using tensor product

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

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

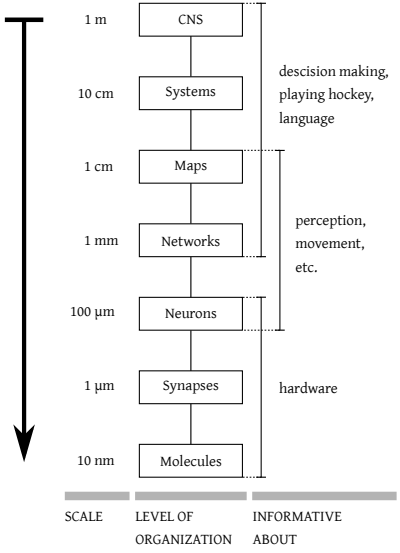
$$= \begin{pmatrix} a_{11} b_{11} & a_{11} b_{12} & a_{12} b_{11} & a_{12} b_{12} \\ a_{11} b_{21} & a_{11} b_{22} & a_{12} b_{21} & a_{12} b_{22} \\ a_{21} b_{11} & a_{21} b_{12} & a_{22} b_{11} & a_{22} b_{12} \\ a_{21} b_{21} & a_{21} b_{22} & a_{22} b_{21} & a_{22} b_{22} \end{pmatrix}$$

Solution Attempt 3: Vector Operators (II)

- ⊕ Solves the binding problem, the problem of two, and the problem of variables
- ● Unclear how to solve the working vs long-term memory problem
- ⊖ 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**.



Vector Symbolic Algebras

VSAs identify four key algebraic operations for capturing symbol-like representations:

1. Binding

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

2. Bundling

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

3. Permutation: We won't worry about this one.

4. Similarity

$$sim(x, y) : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}^1$$

Binding Operator Properties for Vector Symbolic Algebras

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 nine comes after the number eight":

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

- "The dog chases the cat":

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

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

$$\begin{aligned}&\texttt{SUBJ} \circledast \texttt{ANNE} + \texttt{ACT} \circledast \texttt{KNOWS} + \texttt{OBJ} \circledast \\ &\Big(\texttt{SUBJ} \circledast \texttt{BILL} + \texttt{ACT} \circledast \texttt{THINKS} + \texttt{OBJ} \circledast \\ &\Big(\texttt{SUBJ} \circledast \texttt{CHARLIE} + \texttt{ACT} \circledast \texttt{LIKES} + \texttt{OBJ} \circledast \texttt{DAVE}\Big)\Big)\,.\end{aligned}$$

Sentence Encoding Revisited

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

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

- "The dog chases the cat":

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

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

$$\begin{aligned}&\texttt{SUBJ} \circledast \texttt{ANNE} + \texttt{ACT} \circledast \texttt{KNOWS} + \texttt{OBJ} \circledast \\ &\Big(\texttt{SUBJ} \circledast \texttt{BILL} + \texttt{ACT} \circledast \texttt{THINKS} + \texttt{OBJ} \circledast \\ &\Big(\texttt{SUBJ} \circledast \texttt{CHARLIE} + \texttt{ACT} \circledast \texttt{LIKES} + \texttt{OBJ} \circledast \texttt{DAVE}\Big)\Big)\,.\end{aligned}$$

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

Using the Reversibility Property to Answer Questions

- "A blue square and a red circle:"

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

- "Which object is blue?"

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

Using the Reversibility Property to Answer Questions

- "A blue square and a red circle:"

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

- "Which object is blue?"

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

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

VSAs: Potential Binding Operators (I)

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

(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)

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

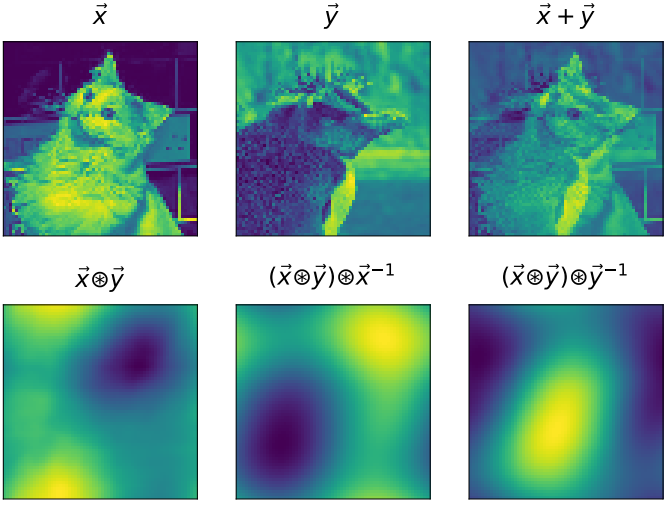
(Circular Convolution)

Circular Convolution is a "compressed" outer product:

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

(Outer Product)

Circular Convolution: Dissimilarity and Reversibility



Circular Convolution: Encoding Numbers

- Spaun uses an interesting encoding to capture number relations.
- Start with a random vector ONE, then

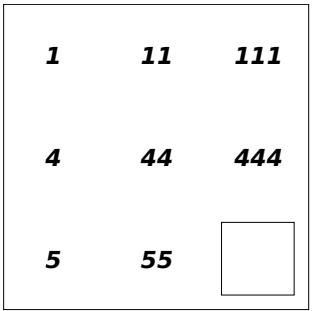
$$\begin{aligned} \texttt{TWO} &= \texttt{ONE} \circledast \texttt{ONE}\,, \\ \texttt{THREE} &= \texttt{ONE} \circledast \texttt{TWO} = \texttt{ONE} \circledast \texttt{ONE} \circledast \texttt{ONE}\,, \\ \texttt{NUMBER-}k &= \underbrace{\texttt{ONE} \circledast \texttt{ONE} \circledast \ldots \circledast \texttt{ONE}}_{k\text{-times}} \\ &= \texttt{ONE}^k\,. \end{aligned}$$

- Which is a slight abuse of notation. Also note,

$$\texttt{ONE}^k = \mathcal{DFT}^{-1}\big(\mathcal{DFT}(\texttt{ONE})^k\big)\,,$$

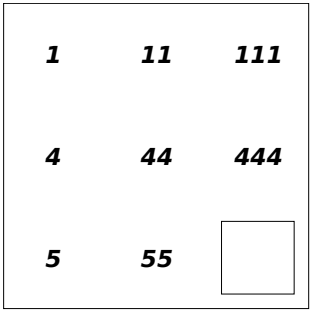
- Where inside the bracket is a Hadamard exponent.

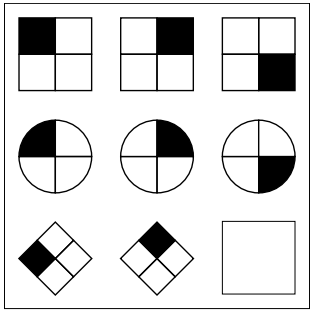
Raven's Progressive Matrices (I)



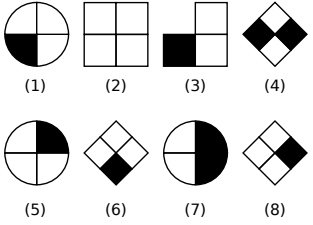
3	***55***	***111***	***44***
(1)	(2)	(3)	(4)
444	***555***	***999***	***33***
(5)	(6)	(7)	(8)

Raven's Progressive Matrices (I)

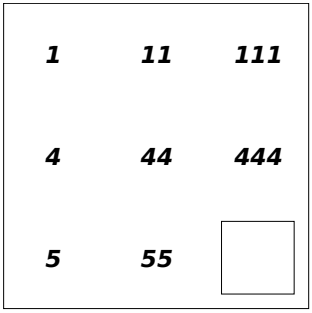


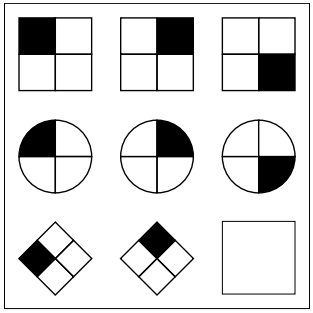


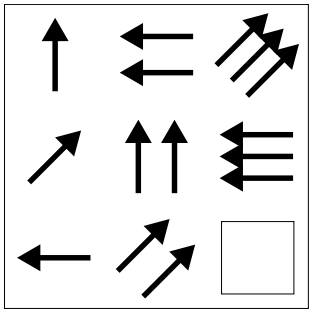
3	55	111	44
(1)	(2)	(3)	(4)
444	555	999	33
(5)	(6)	(7)	(8)



Raven's Progressive Matrices (I)

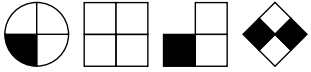






3 55 111 44

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

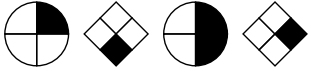


(1) (2) (3) (4)

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

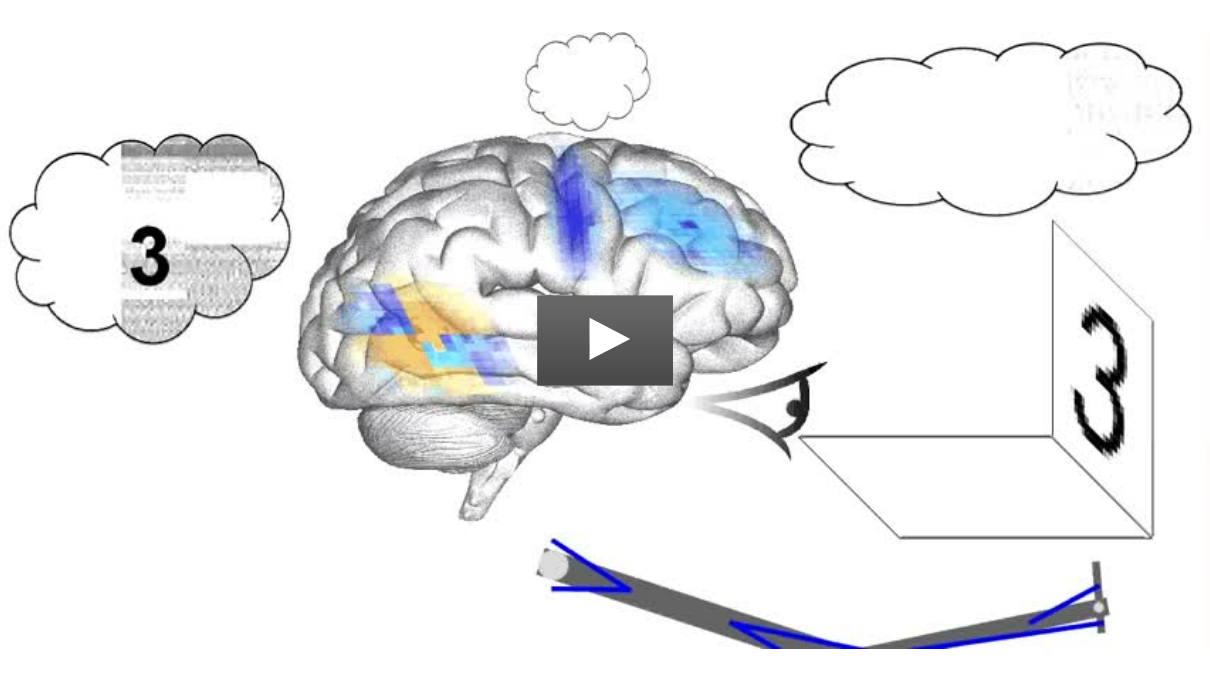
444 555 999 33

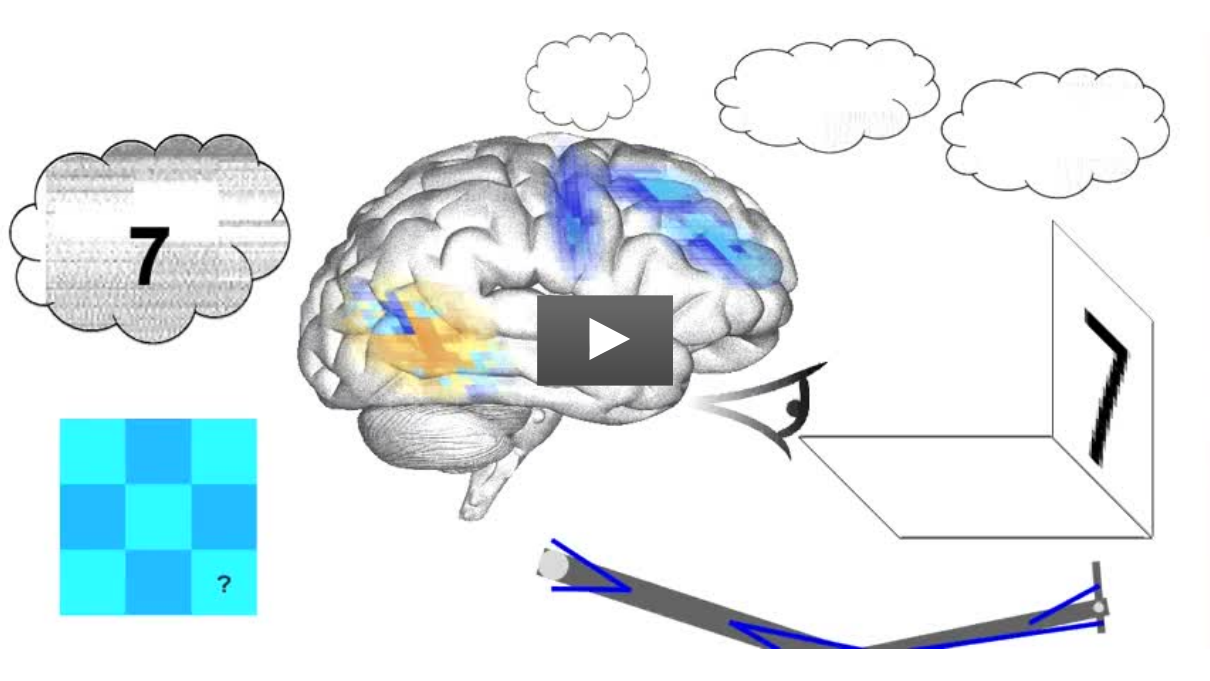
(5) (6) (7) (8)



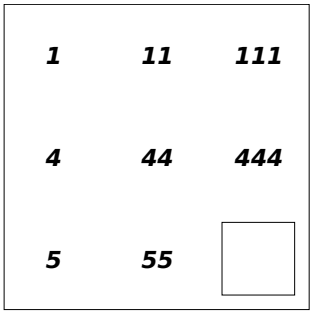
(5) (6) (7) (8)

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


3
3


7
?
7

Raven's Progressive Matrices (II)

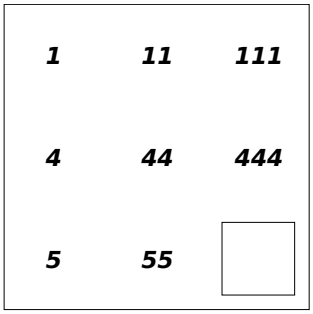


3	***55***	***111***	***44***
(1)	(2)	(3)	(4)
444	***555***	***999***	***33***
(5)	(6)	(7)	(8)

Representing cells:

$$
\begin{aligned}
\texttt{C1} &= \texttt{ONE} \circledast \texttt{P1}\,,\\
\texttt{C2} &= \texttt{ONE} \circledast \texttt{P1} + \texttt{ONE} \circledast \texttt{P2}\,,\\
\texttt{C3} &= \texttt{ONE} \circledast \texttt{P1} + \texttt{ONE} \circledast \texttt{P2} + \texttt{ONE} \circledast \texttt{P3}\,,\\
\texttt{C4} &= \texttt{FOUR}\,,\\
\texttt{C5} &= \texttt{FOUR} \circledast \texttt{P1} + \texttt{FOUR} \circledast \texttt{P2}\,,\\
\texttt{C6} &= \texttt{FOUR} \circledast \texttt{P1} + \texttt{FOUR} \circledast \texttt{P2} + \texttt{FOUR} \circledast \texttt{P3}\,,\\
\texttt{C7} &= \texttt{FIVE} \circledast \texttt{P1}\,,\\
\texttt{C8} &= \texttt{FIVE} \circledast \texttt{P1} + \texttt{FIVE} \circledast \texttt{P2}\,.
\end{aligned}
$$

Raven's Progressive Matrices (III)



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

Extracting the horizontal rule:

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

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

Making a prediction:

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

Jackendoff's Challenges with a VSA

- The Binding Problem
- The Problem of Two
- The Problem of Variables
- Working Memory versus Long-Term Memory

Image sources

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