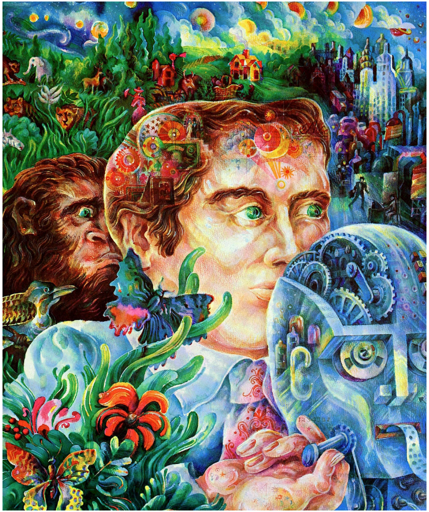
SYDE 556/750

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":

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

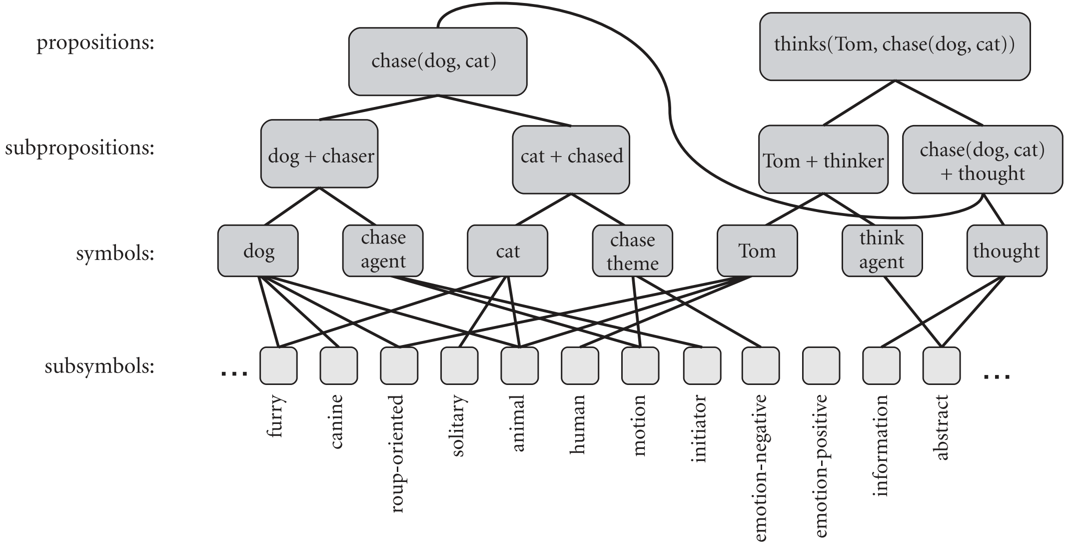
- "All dogs chase cats":

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

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

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

Solution Attempt 1: Neural Synchrony (I)

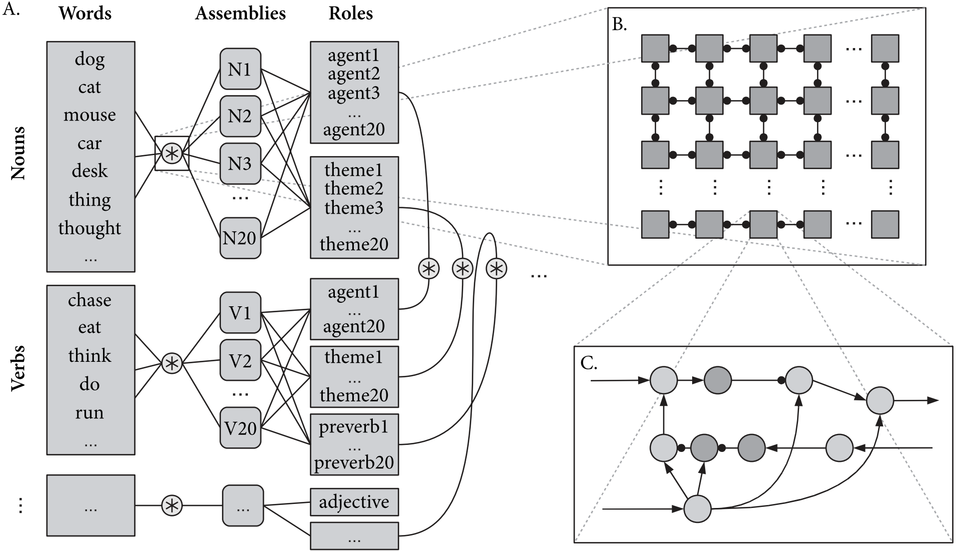


Solution Attempt 1: Neural Synchrony (II)

- 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)



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

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

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

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

Solution Attempt 3: Vector Operators

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

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

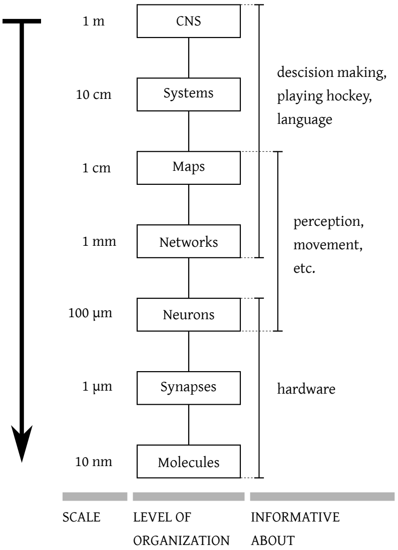
$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} & a_{12} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}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

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

⊖ 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**.



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":

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

- "The dog chases the cat":

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

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

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

Sentence Encoding Revisited

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

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

- "The dog chases the cat":

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

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

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

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

Using the Reversibility Property to Answer Questions

- "A blue square and a red circle:"

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

- "Which object is blue?"

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

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?"

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

⚠ **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)

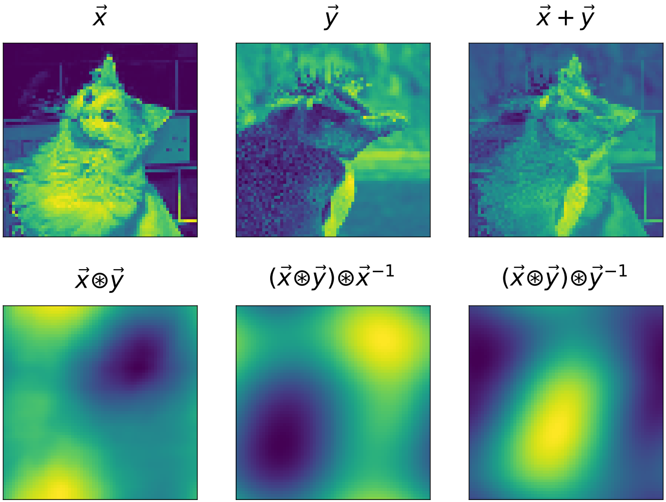
VSAs: Potential Binding Operators (II)

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

Circular Convolution is a “compressed” outer product:

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

Circular Convolution: Dissimilarity and Reversibility



Raven's Progressive Matrices (I)

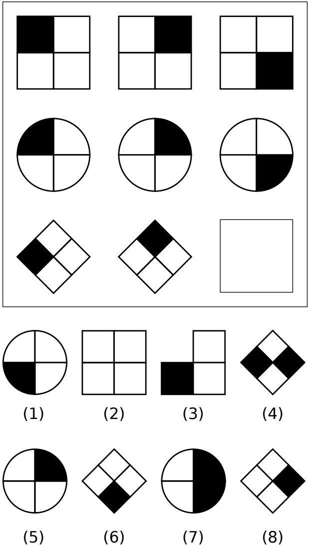
1	***11***	***111***
4	***44***	***444***
5	***55***	

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

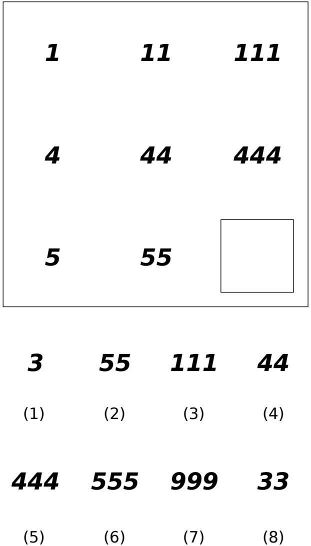
Raven's Progressive Matrices (I)

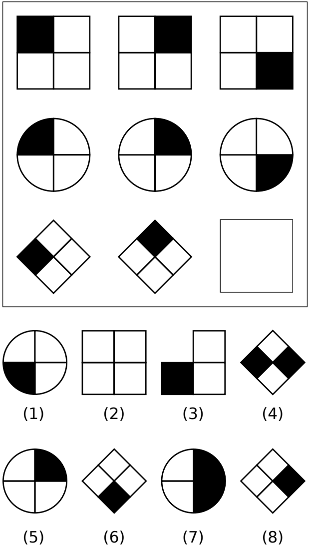
1	***11***	***111***
4	***44***	***444***
5	***55***	

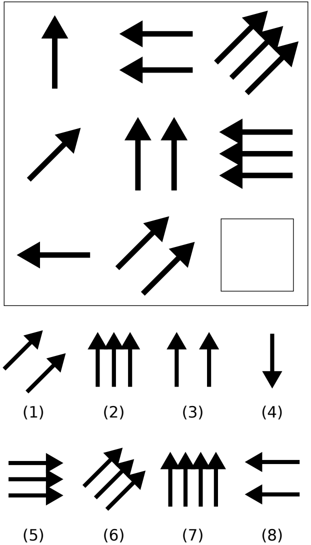
3	***55***	***111***	***44***
(1)	(2)	(3)	(4)
444	***555***	***999***	***33***
(5)	(6)	(7)	(8)

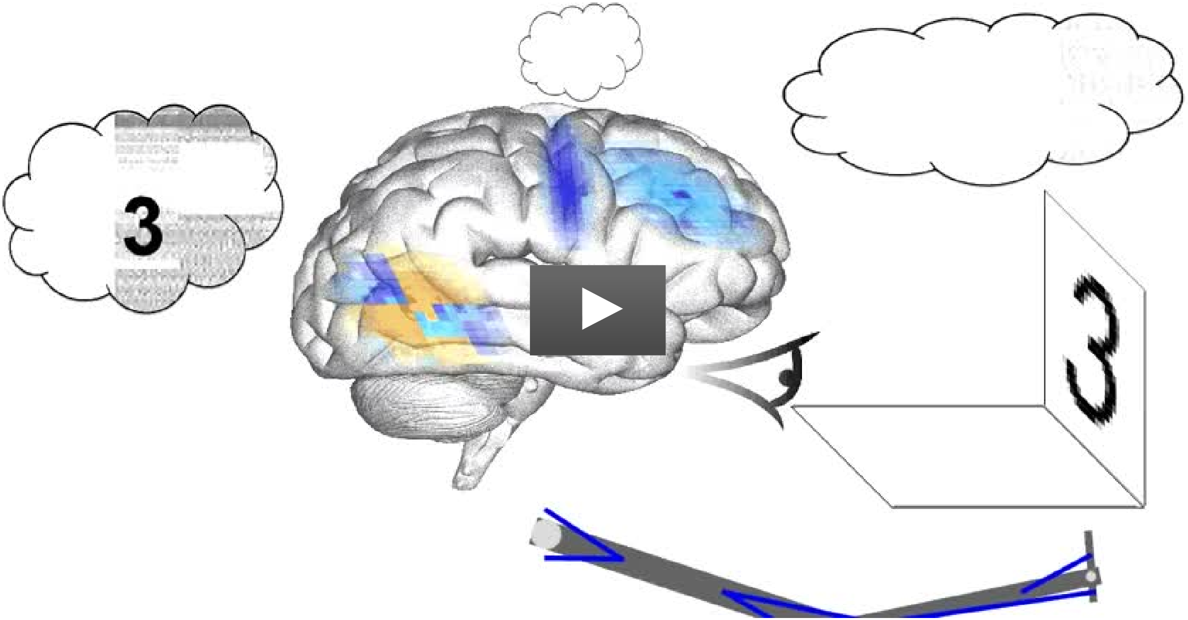


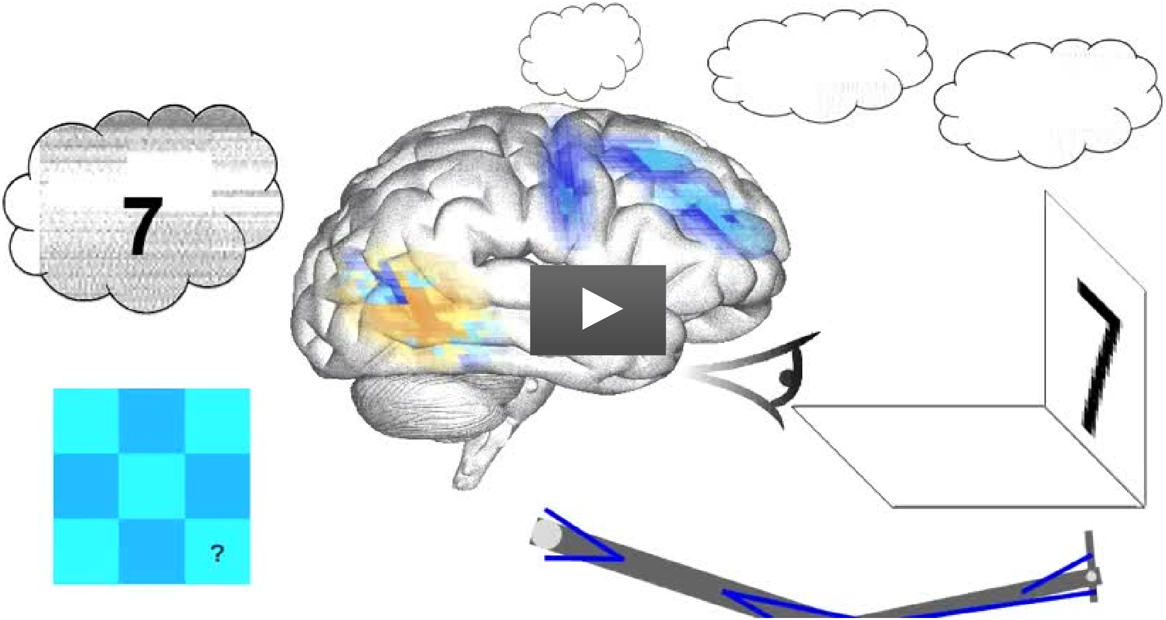
Raven's Progressive Matrices (I)










7
?

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:

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

Raven's Progressive Matrices (III)

1	***11***	***111***
4	***44***	***444***
5	***55***	$\square$

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

Extracting the horizontal rule:

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

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

Making a prediction:

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

Image sources

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