A predictive coding model of transitive inference

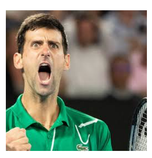
Moritz Moeller, Sanjay Manohar & Rafal Bogacz **MRC BNDU,** University of Oxford, Mansfield Road, Oxford OX1 3TH, UK

Introduction

GOAL To develop a model that tells us how transitive inference works. Which strategy is used? How is that strategy implemented?

What is transitive inference?

Try it yourself! Consider these facts:

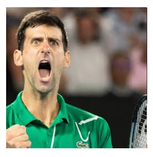
Premise 1  **vs.** 

Winner!

Premise 2  **vs.** 

Winner!

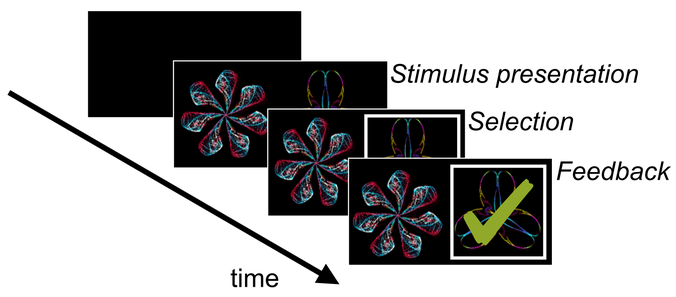
Now make a prediction:

Inference  **vs.** 

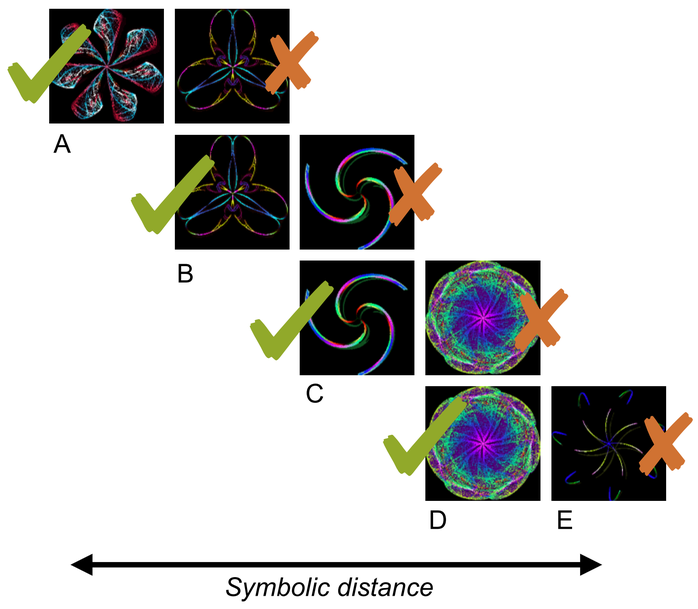
N-Series Task

A task to measure the ability to do TI in animals and infants

Trial structure



Reinforcement structure during training:



The test:

[1] Regina Paxton Gazes, Nicholas W. Chee, and Robert R. Hampton. Cognitive mechanisms for transitive inference performance in rhesus monkeys: measuring the influence of associa- tive strength and inferred order. Journal of Experimental Psychology. Animal Behavior Processes, 38(4):331–345, October 2012.

A Bayesian Model

A linear model of reasoning

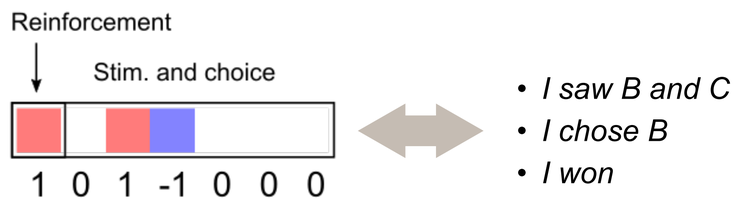
FACT = PREMISE1 + PREMISE2

Vector space of facts

$$\vec{y} = x_1\vec{w}_1 + .. + x_n\vec{w}_n = W\vec{x}$$

Encoding the fact B *is better than C*

Need a way to encode the experiences such that they can be added.



Learning premises via Bayesian inference

Likelihood function

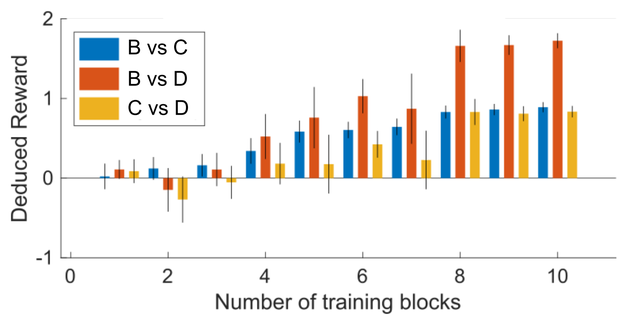
$$\vec{y} \sim \mathcal{N}(W\vec{x}, 1)$$

defines generative density and free energy:

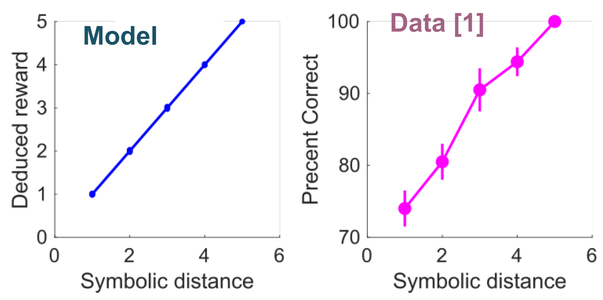
$$P(\vec{y}, W, \vec{x}) \qquad F \propto \frac{1}{2}\left(\vec{y} - W\vec{x}\right)^2$$

Network Performance

A linear bayesian model correctly learns to infer new facts while learning only the premises:

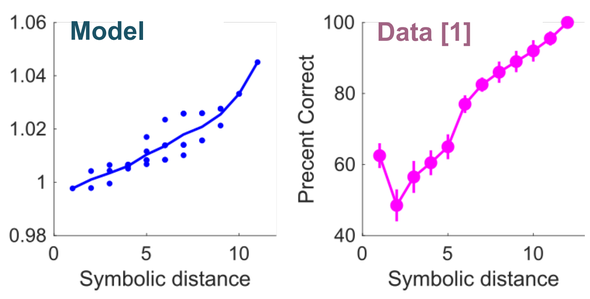


Symbolic distance effect

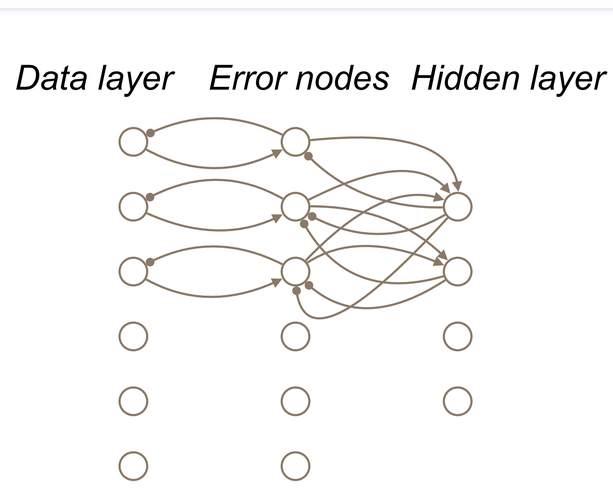


List linking *(other models cannot solve this!)*

1. A > B > C > D > E V > W > X > Y > Z
2. E > V



A Network Model

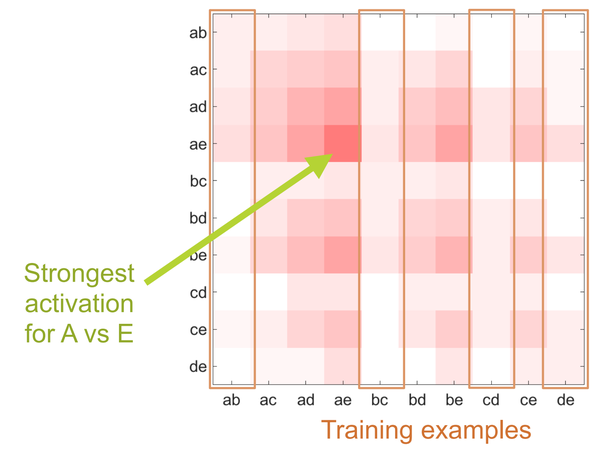


Neural Results

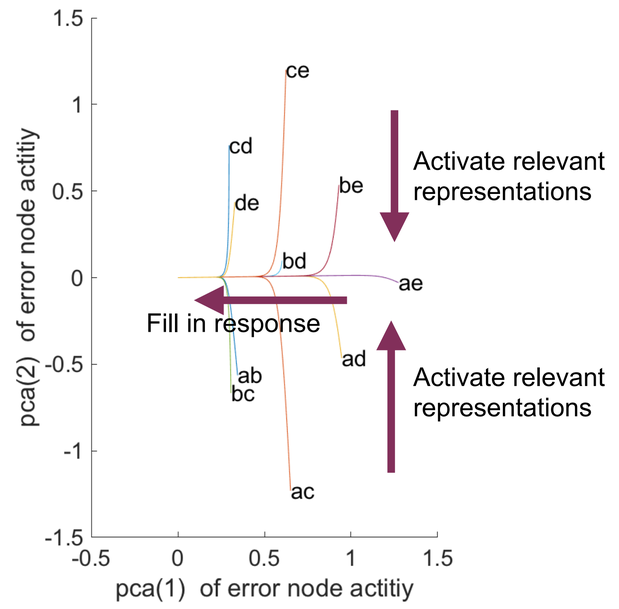
Representational similarity

Hidden unit representations mirror the underlying ordinal similarity structure:

$$R_{i,j} = \left\langle a^i_{\text{hidden}}, a^j_{\text{hidden}} \right\rangle$$



Dynamics in the network: 2 stages of inference



Discussion

- By choosing a suitable vector representation of propositions like "A>B", transitive inference can be performed.
- This kind of representation reproduces key behavioural effects.
- A neural implementation of this provides testable predictions and gives insight into the network computations.