Building, Reusing, and Generalizing Abstract Representations from Concrete Sequences

Shuchen Wu, Mirko Thalmann, Peter Dayan, Zeynep Akata, Eric Schulz

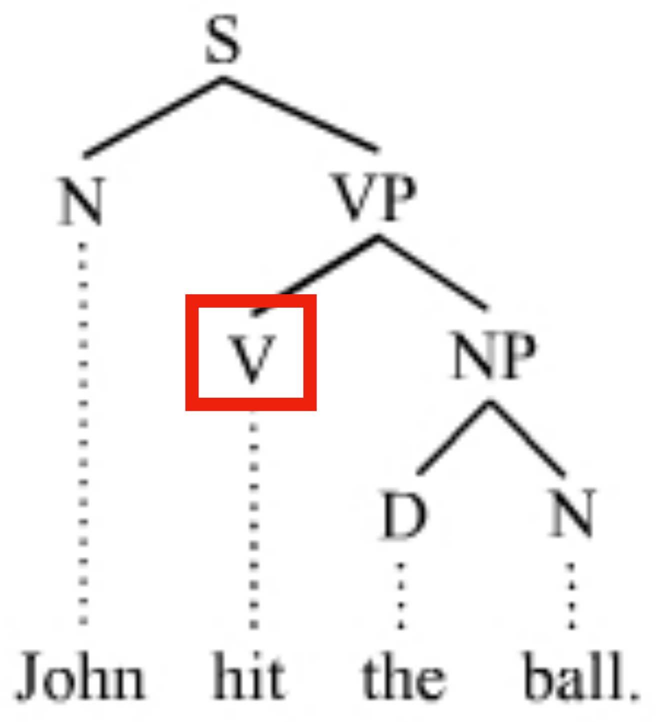
23/03/2025

Learning abstract concept is a hallmark of intelligence

How are abstract concepts learned?

$$\begin{array}{rc} & P \to Q \\ & P \\ \hline \therefore & Q \end{array}$$

Aristotle: concepts and categories are ingredients for deduction and induction



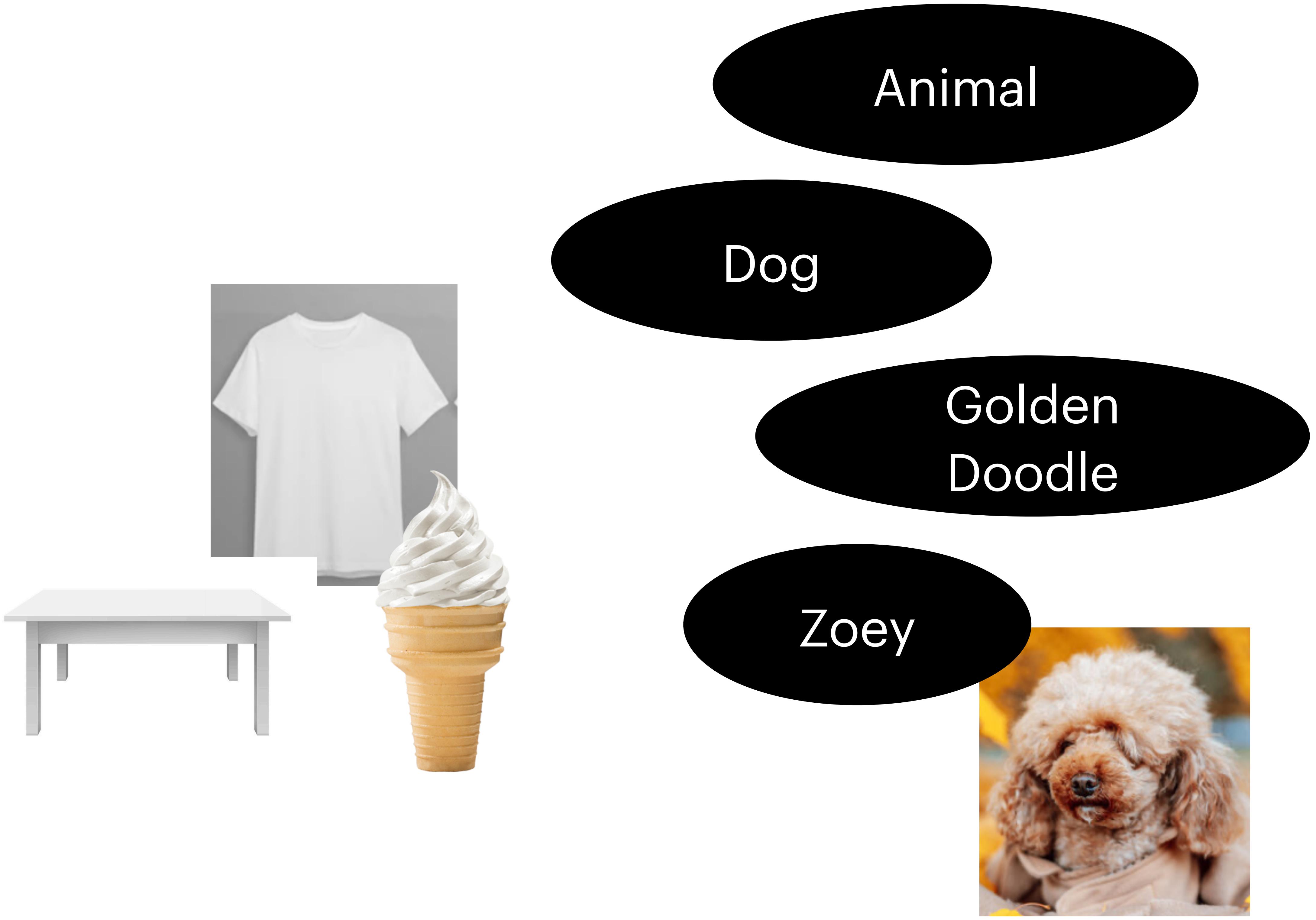
“Language is a process of free creation; its laws and principles are fixed, but the manner in which the principles of generation are used is free and infinitely varied.” —— Noam Chomsky

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}$$
$$\nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}$$

“The propose of science is to find meaningful simplicity in the midst of disorderly complexity” —— Herbert Simon

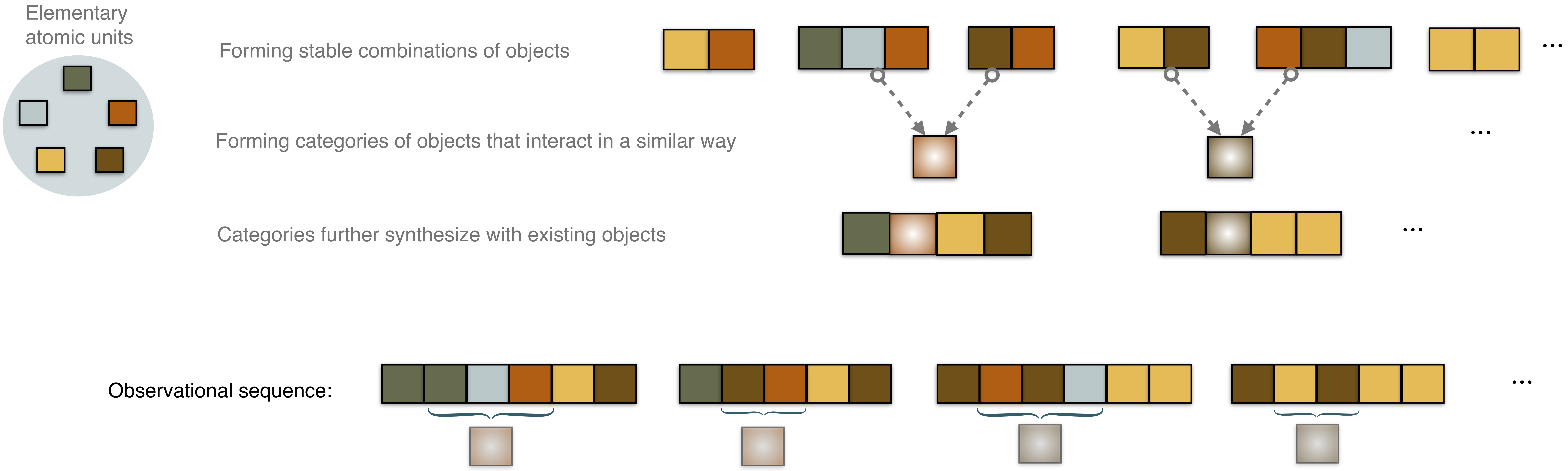
Charactieristics of learning abstract concepts

- Originate from sensorimotor correlates (Barsalou et al. 2018, Gentner & Asmuth 2017), some are nonperceptual (Smith 1992)
- Discrete concepts (Ohlsson et al. 1997), common features (Yee 2019)
- Developed in a graded Level (Pexman 2017), high degree of generality and variability (Yee 2019)
- Can be assembled, more complex abstract concepts can be built up from simpler ones (Ohlsson et al. 1997, Piaget 1977)



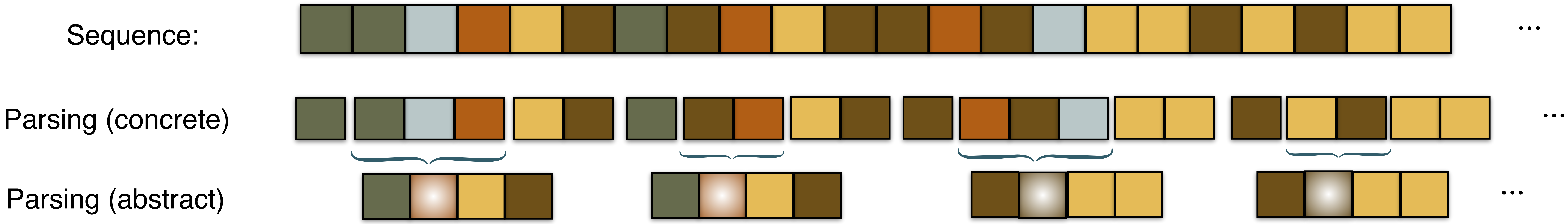
Generative Model:

Sequences are made up from hidden abstract categories



Learning abstractions from sequences

Via a conjunction of chunk proposal and variable discovery



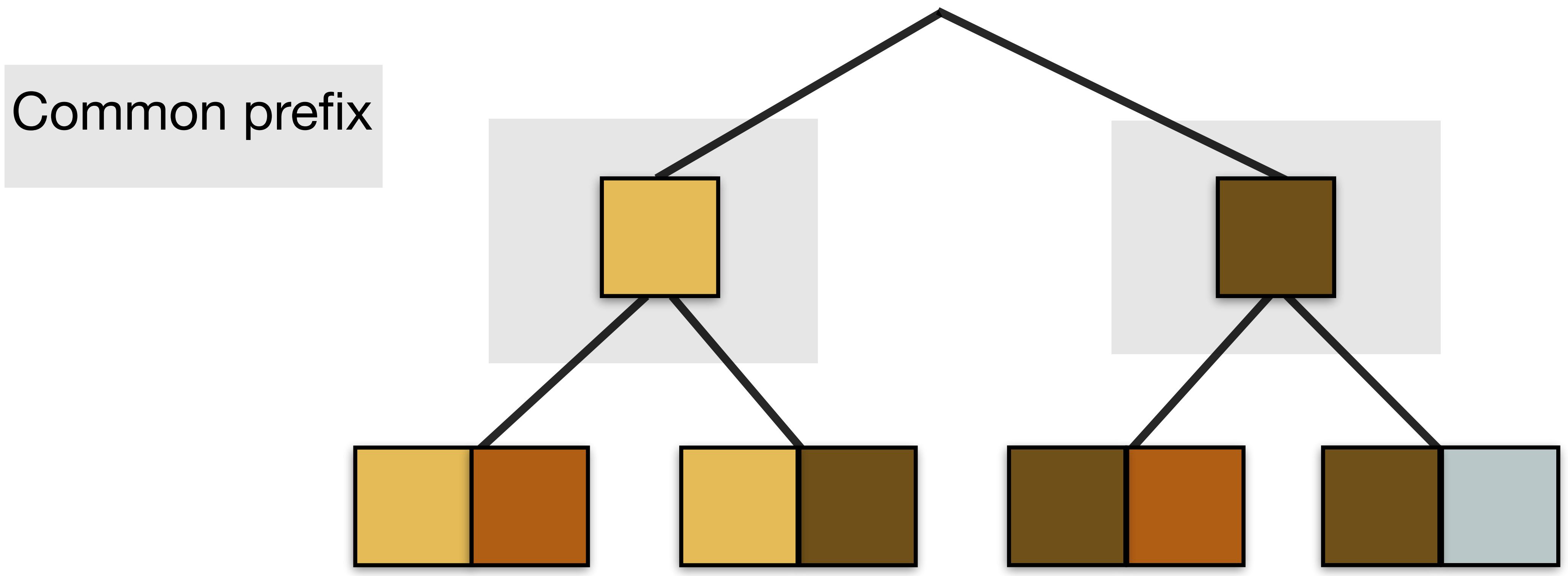
HVM learns a structured inventory of identifiable patterns and use these patterns as entities to parse the sequence

A Taxonomy of Abstractions

Memory Abstraction

Shared content across concrete chunks

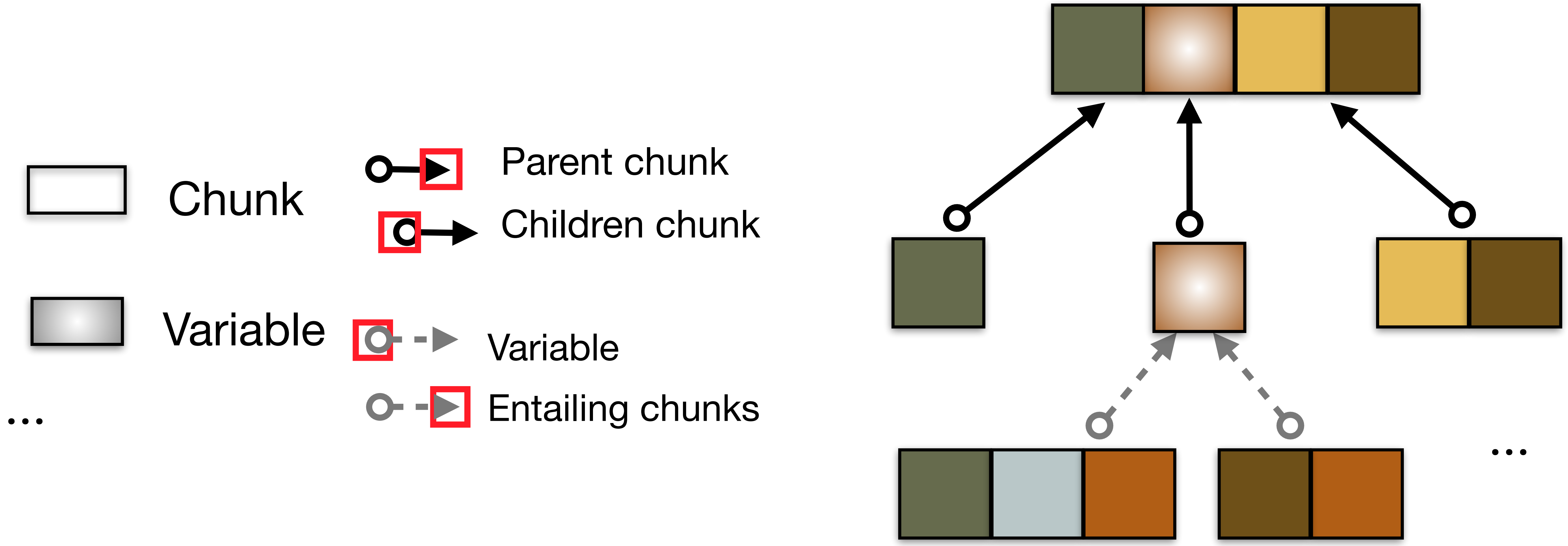
Speed up chunk identification



Variable Abstraction

Chunks and variables

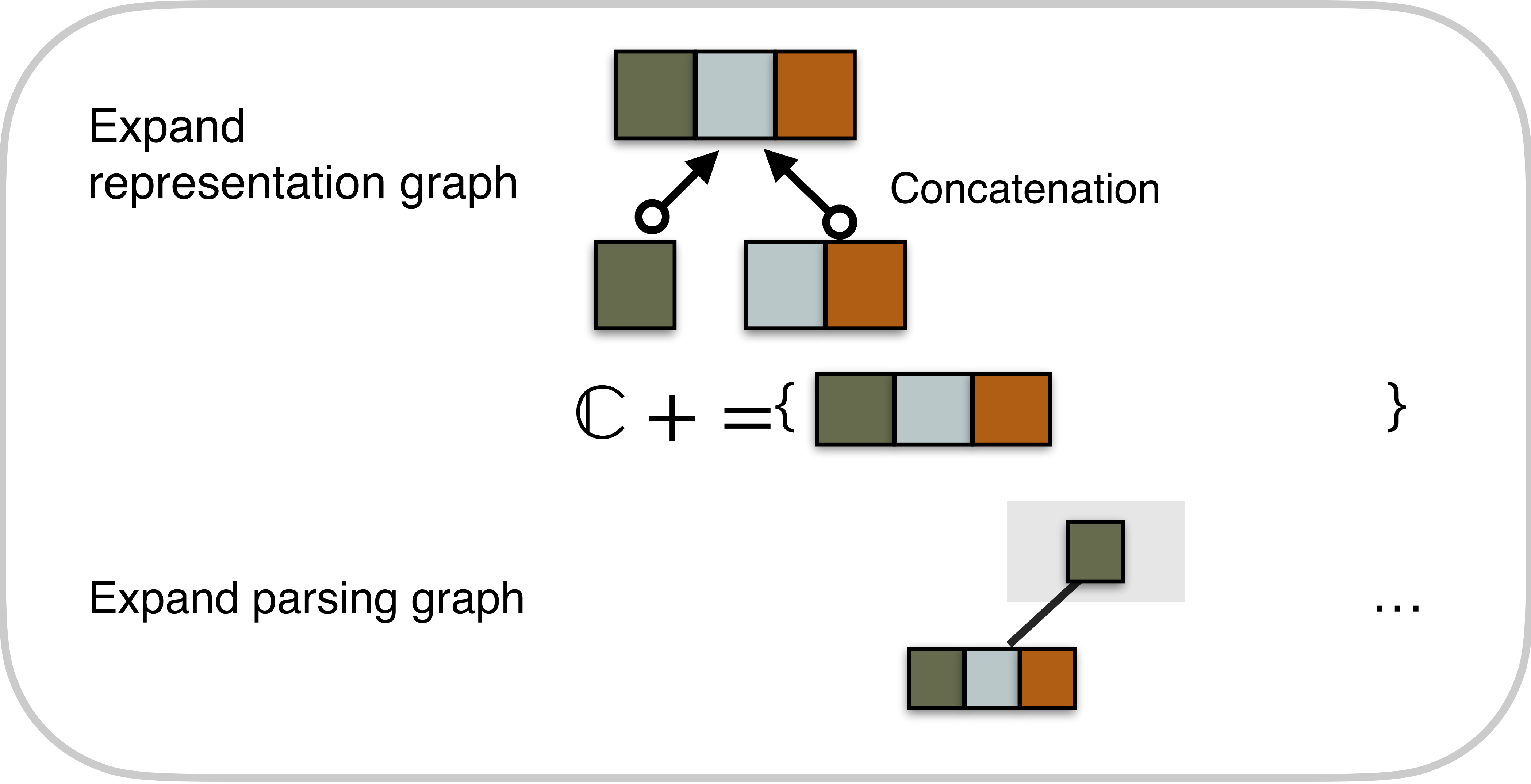
Multiple abstraction levels



Growing the dictionary

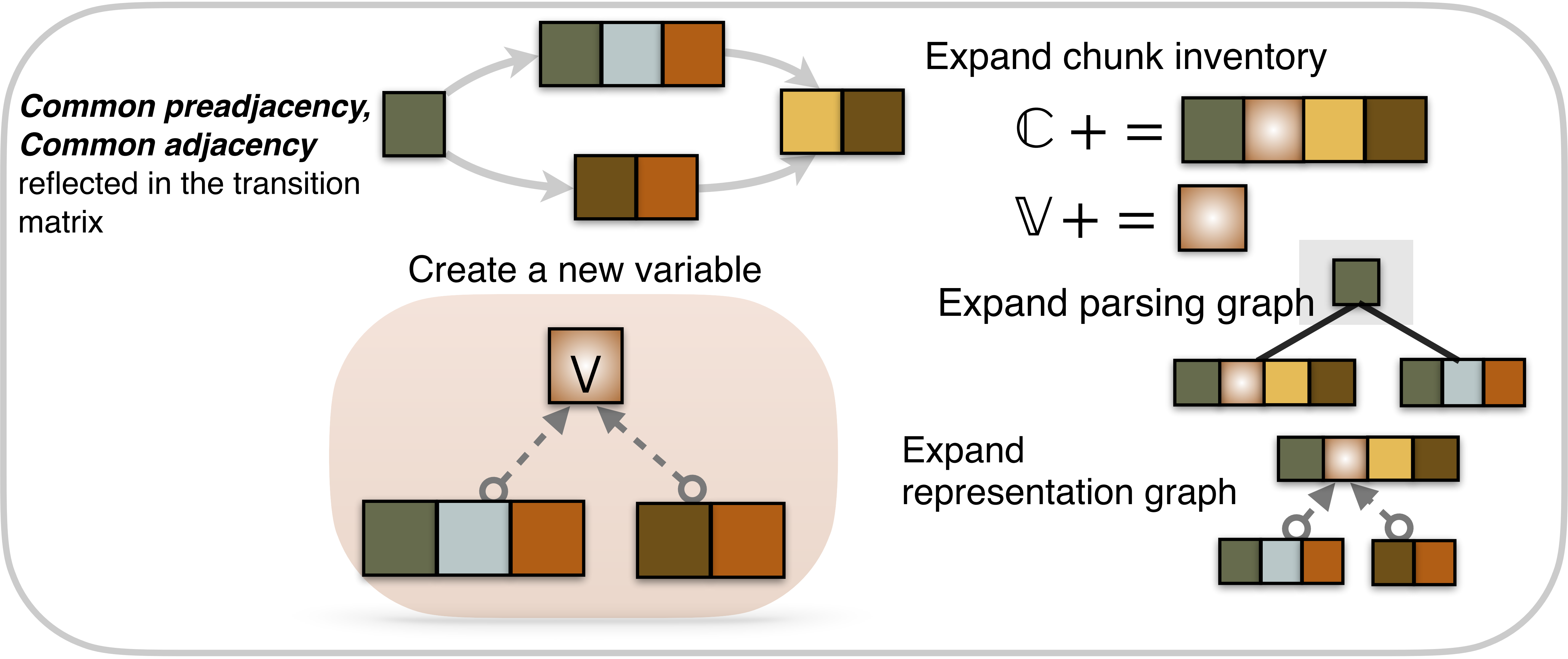
Chunk Proposal

Building on top of the hierarchical chunking model (HCM) (Wu et al., 2022)



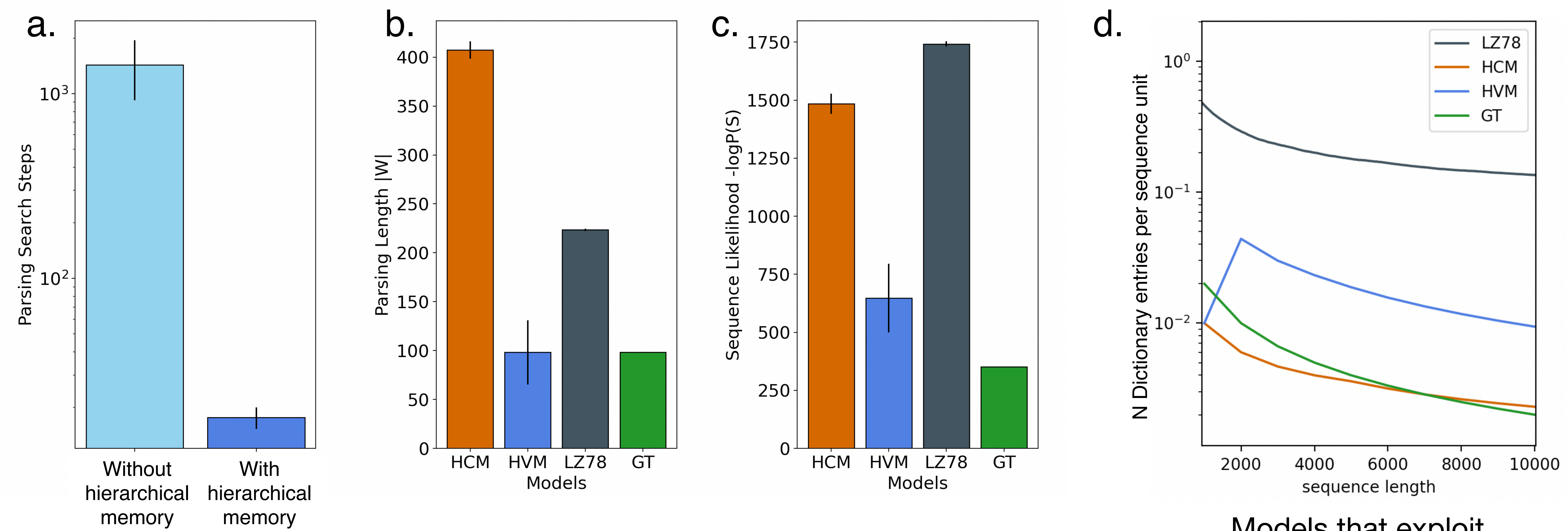
Propose chunks based on correlated transition entries

Variable Discovery



Propose variables between correlated chunk pairs

The benefit of learning abstraction



A parsing graph that exploit common prefixes significantly reduces average parsing search steps

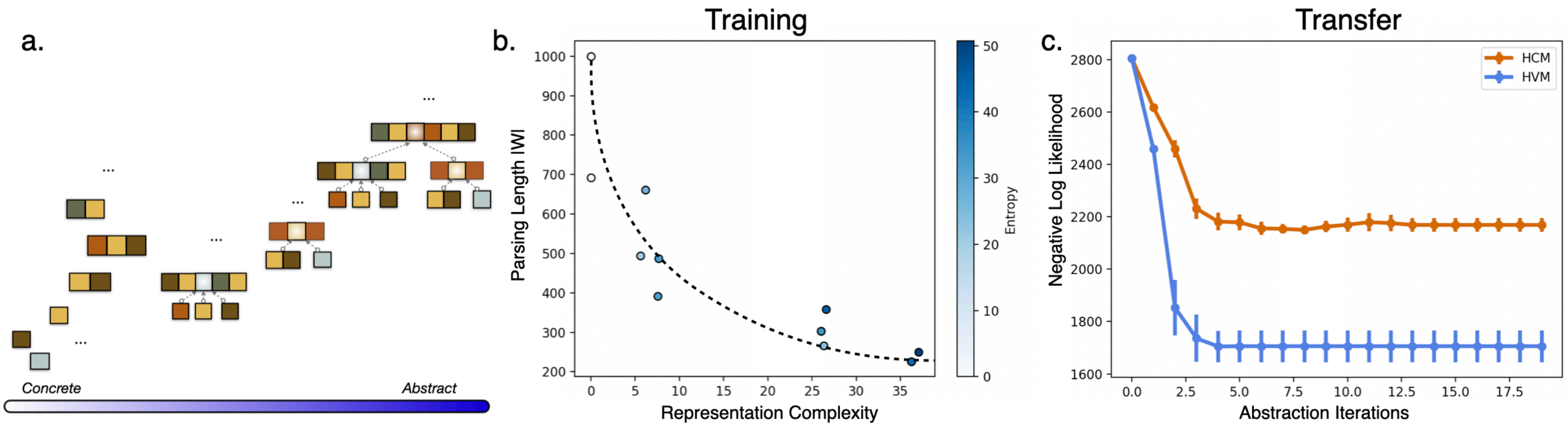
Learn longer patterns in sequences, albeit abstract

Parse sequences with higher likelihood

Models that exploit sequence hierarchies compress sequences more effectively than traditional compression methods.

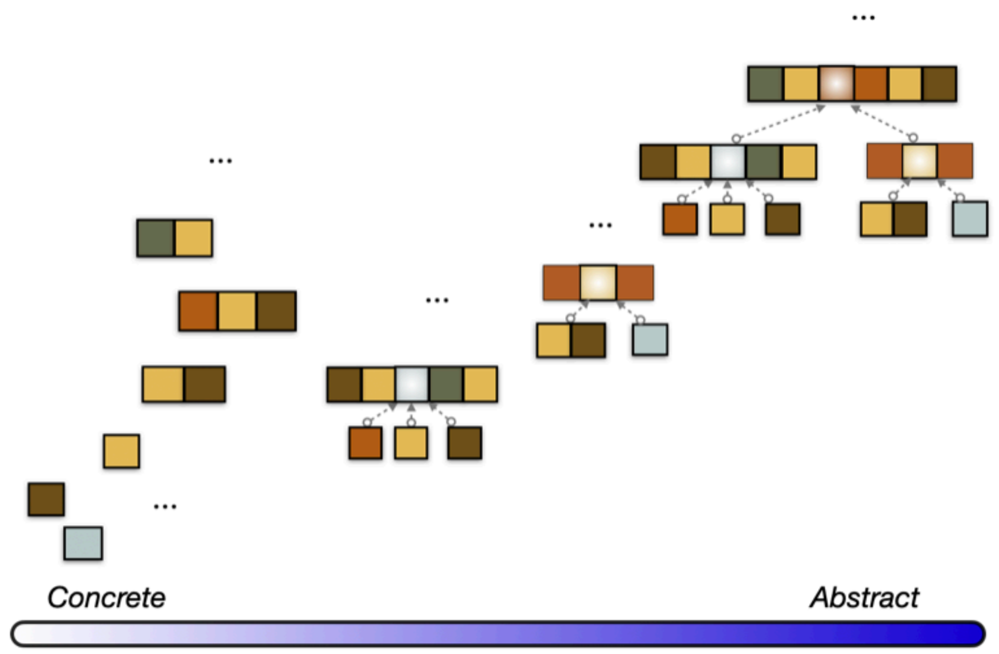
Model complexity relates to human recall time

Abstraction level relates to distortion and generalization



Summary

- We build a minimal cognitive model (HVM) that combines sequence chunking with variable extrapolation
- HVM that learns chunks from sequences and abstracts contextually similar chunks as variables, it efficiently organizes memory while uncovering abstractions
- HVM learns compact sequence representations compared to alternative models
- HVM's sequence likelihood correlates with human recall time in a sequence recall task that demands the transfer of variables
- From HVM's adjustable layer of abstraction, we demonstrate a precise trade-off between compression and generalization
- This cognitive model captures the learning and transfer of abstract representations in human cognition and differentiates itself from LLMs



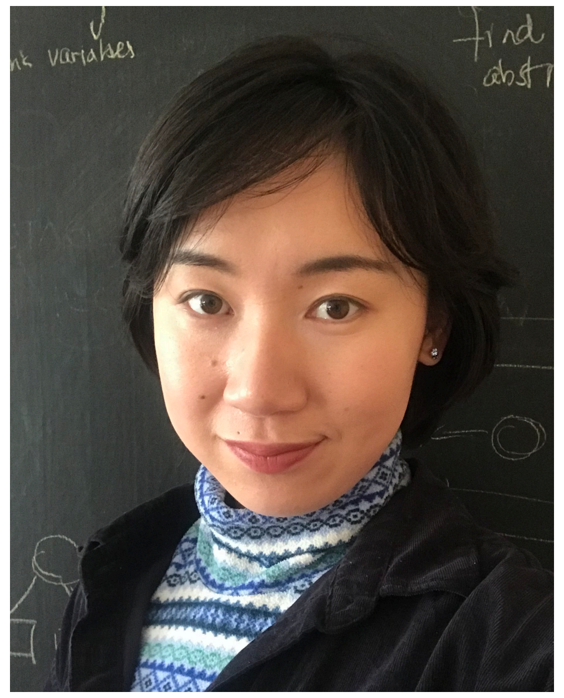
Paper:

HELMHOLTZ MUNICH

EMLMunich

HCAI

Thank you!



Shuchen Wu

Mirko Thalmann

Peter Dayan

Zeynep Akata

Eric Schulz

MAX-PLANCK-GESELLSCHAFT

Max-Planck-Institut für biologische Kybernetik