A Bayesian Model of Joint Category and Feature Learning

Lea Frermann
l.frermann@ed.ac.uk
Institute for Language, Cognition, and Computation, School of Informatics, University of Edinburgh

The ability to generalize from experience
underlying a variety of common mental tasks,
such as learning, perception or language use

Learning categories of concepts
Is a scarf a piece of clothing?

Learning structured types of features
Do all pieces of clothing have color? or material?

Learning incrementally
Immediately utilizing novel insights and information

Large(r)-scale training and testing
Approximating the learning environment with text

We present the first cognitive Bayesian model which learns categories and their structured feature representations in a joint process.

The BCF model

Input
concept mentions in linguistic context

Output
1. categories of concepts
2. feature types
3. category-feature type associations

Assumptions
- each concept \( c \) belongs to a single category \( k \)
- each input refers to a single feature type \( g \)
- feature types capture one aspect of meaning
- categories differ in feature type associations

Experiments

Data – The CHILDES corpus
- speech from child-parent interaction
- we take child-directed speech only
- 21 English-speaking children
- age between 0y11m and 4y11m
- extract mentions of concepts in context

Procedure
- bucket data into 3-month intervals
- present them in chronological order to the model

Evaluate
1. learning behaviour \( \rightarrow \) do representations improve over time?
2. memory constraints \( \rightarrow \) how does limited memory (\# of particles) effect performance?
3. quality of learnt categories and features

Discussion
✓ first cognitive model of joint category and feature learning
✓ cognitively motivated learning algorithm
✓ model training and testing on a (more) realistic scale

X realistic model input – visual or pragmatic signals

A particle filter for the BCF model

Incremental inference: particle filtering (Doucet, 2008)

Sequential Monte Carlo
- incrementally approximate a target distribution through a sequence of intermediate distributions
- represent each distribution through a set of weighted samples (particles)
- recursively update each particle with information from novel observations
- approximate memory limitations: \# of particles, or allowed capacity for re-consideration of past decisions
- known issues: sample degeneracy and sample impoverishment

A particle filter shows how to approximate a target distribution through a sequence of intermediate distributions, recursively updating each particle with information from novel observations. Known issues include sample degeneracy and sample impoverishment.

Table of examples of honed categories, feature types, and their associations

Categories
- tray bread carrot chair bowl bag cheese biscuit potato crayon table

Feature types
- candle cake rose eye ear
- bed duck water eat yay
- clock

Qualitative examples of learnt categories, feature types, and their associations

Improvement of learnt categories over time and under memory constraints

Incremental inference: particle filtering

Diagram shows the improvement of learnt categories over time and under memory constraints, with different numbers of particles (1, 10, 50, 100). The graph illustrates how the performance metric increases as the number of particles increases, indicating better performance under more particles due to improved sampling.