Marr’s Levels of Analysis

Stephen Bennett

Overview

Marr's three levels

- Hardware - Neuroanatomy
- Algorithmic - What people do
- Computational* - Why they do it that way & what is learned

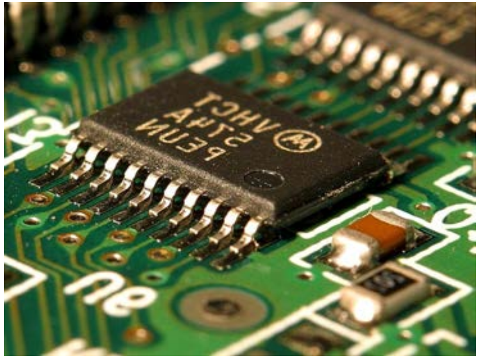
Computational Level & Language Acquisition

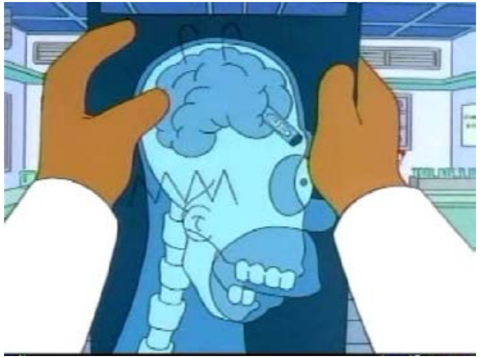
- Cognitive Modeling as a Computational Level Explanation
- Bayesian Babies
- Hypothesis Spaces
- Examples

Marr’s Levels - Hardware

How can the algorithm be performed, physically?

In a language acquisition context, which languages are learnable and what are the neuroanatomical underpinnings that drive the process?

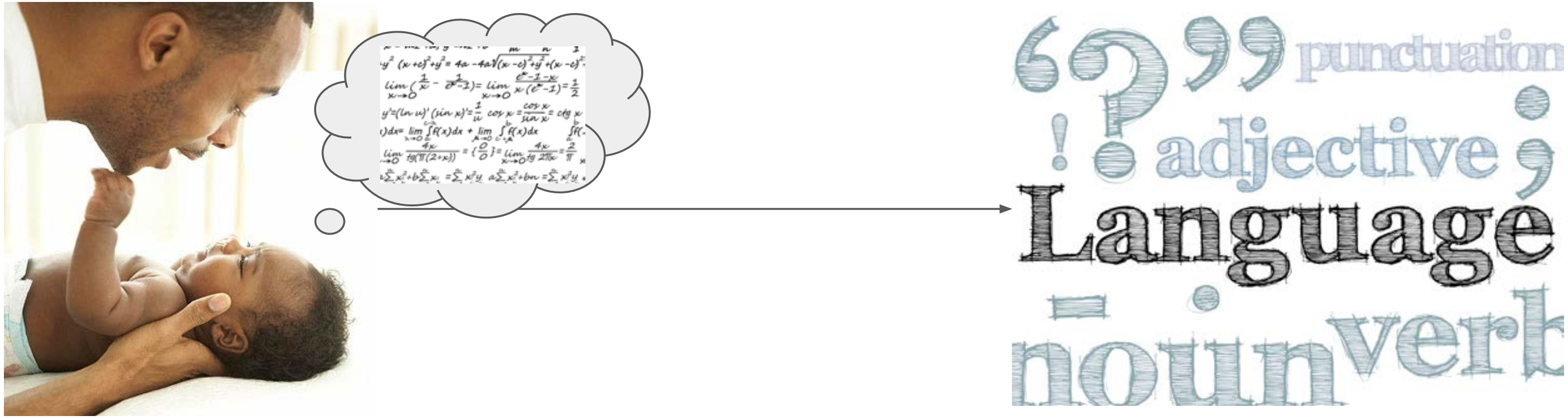




Marr’s Levels - Algorithmic

What input becomes certain outputs, and what computations do we do to get there?

In language acquisition, can we describe the various effects we see in people?



Marr’s Levels - Computational

To what end is all of this done? What is the overarching strategy behind everything?

For language acquisition, this would translate to a theory.

Marr puts emphasis on this level, as an understanding of a process is best achieved by first understanding its goals. This allows for generalization.

Cognitive Modeling as a Computational Explanation

(1) Bayes' Theorem

$$P(h \mid d) = \frac{P(d \mid h)P(h)}{P(d)}$$

(2) $P(h \mid d) \propto P(d \mid h)P(h)$

Defining a Bayesian model usually involves three steps:

(1) Defining the hypothesis space: Which hypotheses does the learner consider?
(2) Defining the prior distribution over hypotheses: Which hypotheses is the learne
biased towards or against?
(3) Defining the likelihood function: How is the observed data generated under a
given hypothesis?

Bayesian Babies

There is a set of possible grammars that can be learned, each with an associated difficulty (probability)

Babies repeatedly update which grammars are likely being used by those around them based on the data they observe

Including indirect evidence (Poverty of the Stimulus)

This process naturally allows the baby to generalize by producing viable utterances from what the baby believes to be a viable grammar.

Hypothesis Space

In linguistics, it remains an open question what the possible grammars are.

As we learn more about how people learn language, we can develop more “hard” rules that govern what the Hypothesis Space a baby born into the world has.

Bootstrapping allows progress in one section of the Hypothesis Space to inform work in another

Overhypotheses allow the individual to infer general principles of their environment - e.g. Head-first structure & [penguins [on icebergs]] are cute.

Algorithms

The Bayesian approach does not specify the algorithm by which humans perform analysis of language, although it does specify the inputs and outputs.

Rational Process Models attempt to answer the *how?*

e.g. Exemplar models

The Utility of Computational Models

Computational modeling augments some of the *learnability* research of early proof-oriented computational studies.

Algorithmic work needs to be done to formalize *how* a computational model can be applied to syntax.

Despite all this, probabilistic models are relatively new for linguistics.

Example 1: Bootstrapping

Goal: Identify distinct phonemes

Solution: Add another goal: Identify the words those phonemes are a part of

How it works: by realizing that a certain distribution of sounds are split by the contexts in which they occur, we can infer that they are distinct phonemes.

A bit of a problem: There are several languages (Inuktitut, Arabic et al.) have context dependent phonemes (allophones) and under this model would be categorized as the same phoneme

Example 2: Purely Computational

How do the methods and assumptions shape word segmentation?

Assumptions: Word dependency vs no word dependency in a word-dependent language

Without word dependency, the model did not find all of the words.

Babies must, then, recognize the dependencies

Methods: Model humans as limited cognitive processors, only processing a single utterance at a time.

They still do pretty well

Sometimes outperform (When the model “overthinks”)

Example 3: Generative Models

Goal: In addition to describing effects of learning, our theory should generate novel utterances

Solution: It already does!

Bayesian models are inherently generative. Its understanding of language by definition allows it to “speak”.

Additionally, people have an intuition that language is developed in this way and coach children by selecting the maximally informative word for a given context.

Example 4: Suspicious Coincidences

“Look! A black cat. Oh, look - there’s another one!”

The lack of data in favor of a hypothesis despite encountering situations where an utterance consistent with that hypothesis would be said is evidence against said hypothesis.

“Look! A black cat.”

References

Clark, A., & Sakas, W. G. (2010). Computational Models of First Language Acquisition Special Issue of Research on Language and Computation. *Research on Language & Computation*, *8*(2), 101-106.

Marr, D., & Vision, A. (1982). A computational investigation into the human representation and processing of visual information. *WH San Francisco: Freeman and Company*.

Pearl, L., & Goldwater, S. Statistical learning, inductive bias, and bayesian inference in language acquisition.