Recap: The Problem of Learning Syntactic Categories

- What information sources are useful?
- Model of how distributional information aids the acquisition of syntactic categories.
- Using agglomerative clustering on CHILDES corpus.
- General approach uses computationally explicit model of specific aspects of language acquisition.

Word Meaning

Q1: How do we organize our knowledge of the world?
Q2: What are concepts?
Q3: How are objects placed into categories?
Q4: How do we represent the meaning of words?

Reading:
Trevor Harley (2001). The Psychology of Language, Chapter 10
Memory (Tulving, 1972)

- **Episodic memory**: memory for events and particular episodes.
- **Semantic memory**: general knowledge (is abstracted from episodes that may happen many times).

Concepts and Categories

A **concept** is a mental representation of a class of objects or events. It determines how things are related or categorized.

- All words have an underlying concept.
- Not all concepts are labeled by a word.
- We do not have a special word for **brown dog**.
- How is the meaning of **dog** represented and how do we pick instances of dogs in the environment?
- How do we categorize the world?

[We use **small caps** to denote concepts or categories]

Functions of Concepts (Rosch, 1978)

- The way in which we categorize the world is **not arbitrary**.
- The categories we form are determined in part by the way in which we perceive the **structure** of the world.
- By dividing the world into classes of things to decrease the amount of information we need to learn, perceive, remember, and recognize: **cognitive economy**.
- Concepts permit us to make accurate **predictions**.
- Categorization serves a **communication** purpose.

Is there a Preferred Level of Conceptualization?

- **Superordinate level**
  - **Furniture**
  - **Chair**
  - **Windsor**
- **Basic level**
- **Subordinate level**
What’s Special about the Basic Level?
1. Most abstract level at which objects have similar shapes.
2. First words are learned at the basic level (%doggy%, %car%, %ball%).
3. We name objects at the basic level faster.
4. Participants list attributes of the basic level easily.

What is Categorization?
Categorization is the process by which things are placed into groups called categories.
- results in a compact and efficient representation of the world;
- presumably involves a process of abstraction or generalization;
- makes it possible to recognize a new object quickly and determine (for example) its use;
- we learn new categories (children learn new categories frequently; adults learn new categories from time to time).

Knowing the Category Provides a Lot of Information

- Likes milk, fish
- Has whiskers
- Sleeps a lot, but more active at night
- A feline, related to lions and tigers
- Has nine lives
- Difficult to train
- Catches mice
- Likes to rub up against people and other objects

© 2007 Thomson Higher Education
Definitional Theory

Originated with Aristotle. Categories are represented as list of features which are both necessary and sufficient. Category membership is determined by checking if an item possesses all requisite features.

<table>
<thead>
<tr>
<th>TEA CUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>concrete object</em></td>
</tr>
<tr>
<td>2. <em>concave</em></td>
</tr>
<tr>
<td>3. <em>can hold liquids</em></td>
</tr>
<tr>
<td>4. <em>has a handle</em></td>
</tr>
<tr>
<td>5. <em>can be used to drink hot liquids</em></td>
</tr>
</tbody>
</table>

Properties 4 and 5 are debatable (Chinese tea cups). If you drop 4 and 5, then there are many objects (bowls) that satisfy 1–3.

Typicality Effects

Typical
- is robin a bird?
- is dog a mammal?
- is diamond a precious stone?

Atypical
- is ostrich a bird?
- is a whale a mammal?
- is turquoise a precious stone?

Slower verification times for atypical items.

Family Resemblance

Ludwig Wittgenstein: the structure of a concept is like a family resemblance photo.
- Take multiple exposure photo of all family members.
- Only the common features are retained.
- Photo looks like everyone in family, but isn’t any one person.

Prototype Theory

Categories are organized around a category prototype. A prototype is an average family member. Potential members of the category are identified by how closely they resemble the prototype (best example of a concept).

- Categories under prototype view are “fuzzy”
- Organized around typical properties or correlated attributes
- Category membership is similarity-based
Fuzzy Boundaries

It can be both!
It is perhaps more prototypical of a bowl.
Fuzzy boundary means membership can be graded (0.9 bowl vs. 0.3 cup).

Is this a cup or a bowl?

Representation and Classification

Exemplar Theory
A category is represented by list previously encountered exemplars. New exemplars are compared to known exemplars — most similar item will influence classification the most.

Prototype vs Exemplar Theory

In both theories category membership is based on resemblance.

Prototype theory:
- category structure is based on prototypes;
- categorization based on similarity to prototype of category.

Exemplar theory:
- categorization based on total similarity of object to exemplars of the category versus total similarity of object to non-exemplars of the category;
- assumes only that we can retrieve memories of specific instances of a category;
- no abstraction of prototypes.

Maybe both theories are true but for different situations.
Problems for Resemblance-based Approaches

How do we know which properties to compare? (Murphy and Medin, 1985).

- Both plums and lawnmowers weigh less than a ton
- They are both found on earth
- They are both bigger than a grain of sand.

Theory-based View

Organization of categories is based on theories about the world; understanding a category involves having a theory about that category and why its members cohere. Classification is based on explanatory relationships.

- Is this a lemon?
- What if we paint it?
- Or make it sweet?
Theory-based View

**Theory Theories**
Organization of categories is based on theories about the world; understanding a category involves having a *theory* about that category and why its members cohere. Classification is based on *explanatory relationships*.

Or squash it?

Informatics 1 CG: Lecture 11  Concepts and Categories  21

---

**Theory Theories**
Organization of categories is based on theories about the world; understanding a category involves having a *theory* about that category and why its members cohere. Classification is based on *explanatory relationships*.

Is this a lemon?

People say its still a lemon – the substance is still “the same”.

Informatics 1 CG: Lecture 11  Concepts and Categories  21

---

**Theory Theories**
Organization of categories is based on theories about the world; understanding a category involves having a *theory* about that category and why its members cohere. Classification is based on *explanatory relationships*.

Is this a teapot?

Informatics 1 CG: Lecture 11  Concepts and Categories  21

---

**Theory Theories**
Organization of categories is based on theories about the world; understanding a category involves having a *theory* about that category and why its members cohere. Classification is based on *explanatory relationships*.

What if a bird lives inside it, in a tree?

Informatics 1 CG: Lecture 11  Concepts and Categories  21
Theory-based View

Organization of categories is based on theories about the world; understanding a category involves having a **theory** about that category and why its members cohere. Classification is based on **explanatory relationships**.

What if a bird lives inside it, in a tree?

Function also plays an important role in category membership.

What if we buy stuff with it?

What if it looked exactly like a real dollar?

Is a counterfeit dollar bill a dollar bill?
Theory-based View

**Theory Theories**

Organization of categories is based on theories about the world; understanding a category involves having a theory about that category and why its members cohere. Classification is based on explanatory relationships.

What if it looked exactly like a real dollar?

No amount of resemblance can make counterfeit dollar into real one!

Summary

**Categorization** is one of the classical problems in the field of cognitive science, one with a history dating back to Aristotle.

- Ability to generalize from experience underlies a variety of common mental tasks
- Perception, learning, and the use of language.
- Definitional, prototype, exemplar, and theories theory.
- Basic-level categories, prototype, family resemblance.

**Next lecture:** theories of word meaning.