

Concepts and Categories

Informatics 1 CG: Lecture 11

Mirella Lapata

School of Informatics
University of Edinburgh
mlap@inf.ed.ac.uk

February 4, 2016

Reading:

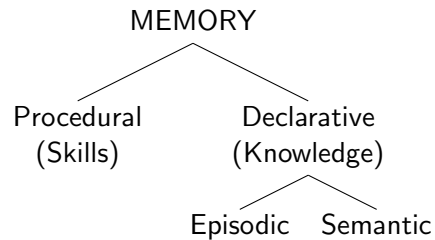
Trevor Harley (2001). The Psychology of Language, Chapter 10

Recap: The Problem of Learning Syntactic Categories

- What information sources are useful?
- Model of how distributional information aids the acquisition 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?



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

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, we 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.

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

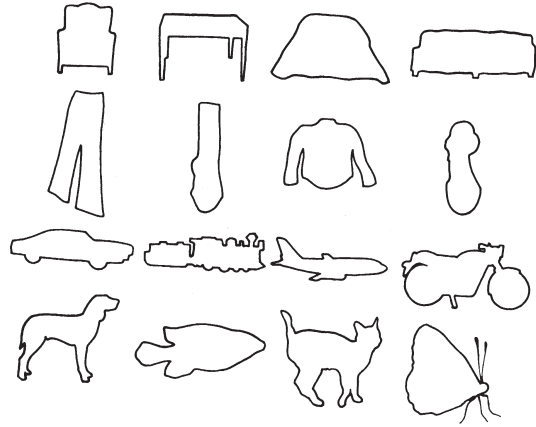
- All words have an underlying concept.
- Not all concepts are labeled by a word.
- We do not have a special word for BROWN DOG.
- How is the meaning of DOG represented and how do we pick instances of dogs in the environment?
- How do we categorize the world?

[We use SMALL CAPS to denote concepts or categories]

Is there a Preferred Level of Conceptualization?



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?

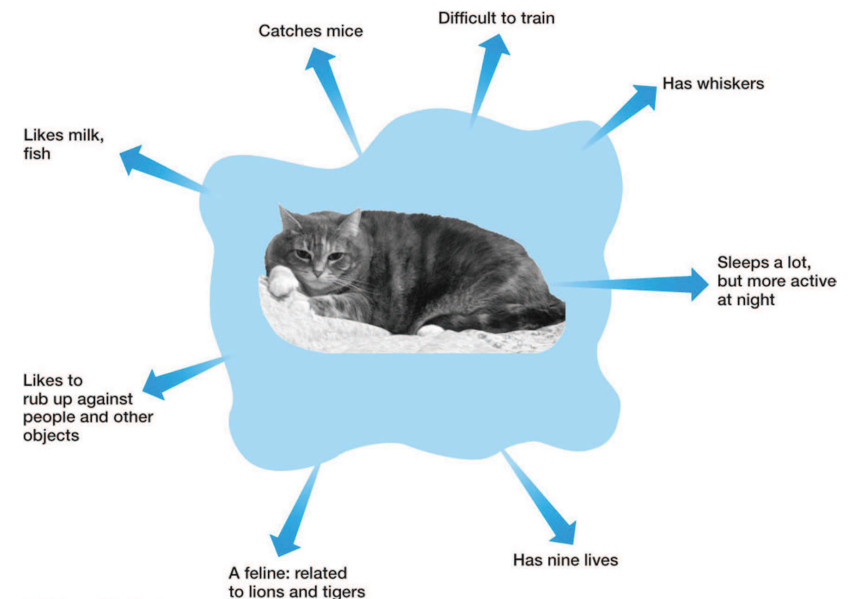
Categorization is the process by which things are placed into groups called **categories**.

- 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



Knowing the Category Provides a Lot of Information



Representation and Classification

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.

TEA CUP

1. *concrete object*
2. *concave*
3. *can hold liquids*
4. *has a handle*
5. *can be used to drink hot liquids*

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.

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.

Typicality Effects

Typical

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

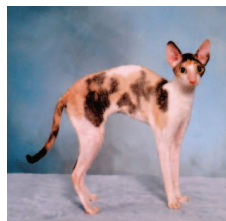
Atypical

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

Slower verification times for atypical items.



Is this a *chair*?



Is this a *cat*?



Is this a *dog*?

Representation and Classification

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).



© 2007 Thomson Higher Education

- Categories under prototype view are **"fuzzy"**
- Organized around typical properties or correlated attributes
- Category membership is **similarity-based**

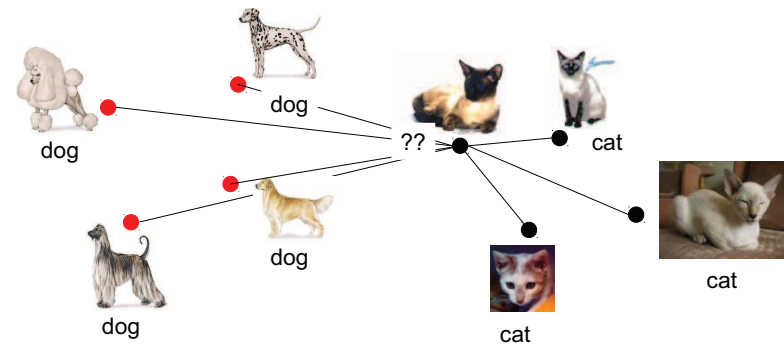


Is this a CUP or a BOWL?

- 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).

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.



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.

- 1 Retrieve memories of specific cats we have encountered.
- 2 Retrieve memories of relevant non-cats e.g., memory of a dog; memory of a stuffed animal; memory of a raccoon.
- 3 Compute total similarity of current instance to memories of positive and negative exemplars (exemplars of cats/non-cats).
- 4 Decide that exemplar is a cat if it is more similar to the memories of cats than to memories of relevant non-cats.

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).

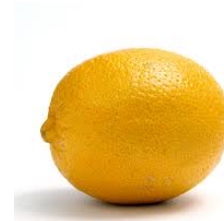


- Both plums and lawnmowers weigh less than a ton
- They are both found on earth
- They are both bigger than a grain of sand.

Theory-based View

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?

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 we paint it?

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 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**.

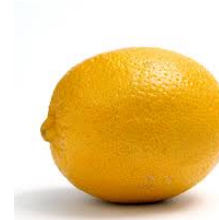


Or squash it?

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**.



Is this a lemon?

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

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**.



Is this a teapot?

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 a bird lives inside it, in a tree?

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 a bird lives inside it, in a tree?

Function also plays an important role in category membership.

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**.



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 we buy stuff with it?

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?

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!

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**.

- Classification is based on **explanatory relationships**
- “Deep” features more important than surface ones (if something functions as a home for birds, must be a birdhouse)
- Theories give us coherent concepts (“wings” and “can fly” go together for a reason)
- Medin (1989): we learn categories faster when we understand why the features go together.

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.