Lecture 5: Human Concept Learning

Cognitive Systems - Machine Learning

Part I: Basic Approaches of Concept Learning

Rules, Prototypes, Examples

last change October 2, 2014

Outline

- Categories
- Similarity-based theories
  - Rule-based
  - Prototypes
  - Exemplars
- Summary

Function of Categories

- Grouping objects to categories/concepts is a basic cognitive competence
- sometimes: category in the world, concept in the mind
- Grouping of objects with “similar characteristics”
  (it is a bit vague what we mean by similar and by characteristic)
- Using knowledge about known exemplars of a category to govern behavior with new objects/in new situations
- Categories as prerequisite for understanding and inference (expectation about an object, a situation)
- Communication of knowledge about categories via language

See: *M. Waldmann (2002), Kategorisierung und Wissenserweb. In J. Müsseler and W. Prinz (Eds.), Allgemeine Psychologie. Spektrum Verlag.*

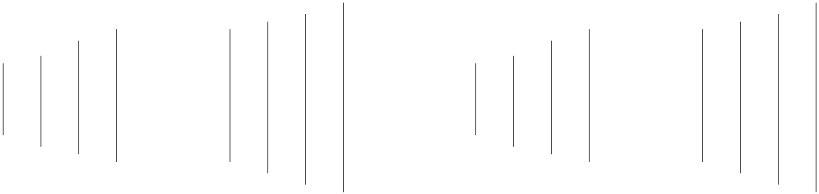
Mis-Conceptions

- People show tendencies to perceive differences between objects within a category as smaller and between objects of different categories as larger
- Tajfel and Wilkes (1963)
  - Imposition of a category directly affects judgement:
  - Lines shown without any category label: errors of judgement mainly random
  - Lines labelled with "A" and "B": perceivers tend to judge lines within category as more similar than they are and between categories as more different as they are
- Tajfel related these findings to social judgements/stereotypes and prejudice: category distinctions (e.g. between nations, gender) tend to overemphasize differences between members of the own group and others

Tajfel & Wilkes, 1963

Classification and quantitative judgement

experimental condition: control condition:



Category A Category B

Mental Representation of Categories

- Most theories concerned with representation and classification also some ideas on acquisition of categories!
- Two perspectives:
  - Similarity based: “bottom-up” grouping of exemplars
  - Theory guided: categories as intuitive theories, exemplars as data (looking at attributes and relations)
- Dominant perspective is similarity based:
  - Rule-based (definitory attributes)
  - Prototypes (probability of membership)
  - Exemplar-based (no generalization/aggregation)

Rule-Based Categories

- Concept specifies necessary and sufficient conditions for membership
- Mostly studied with artificial categories
- E.g., stimuli used by Bruner, Goodnow and Austin (1956)
  - Form: cross, circle, square
  - Color: unfilled, filled, gray
  - Number of objects: one, two, three
  - Frame: one, two, three

Stimuli of Bruner et al.

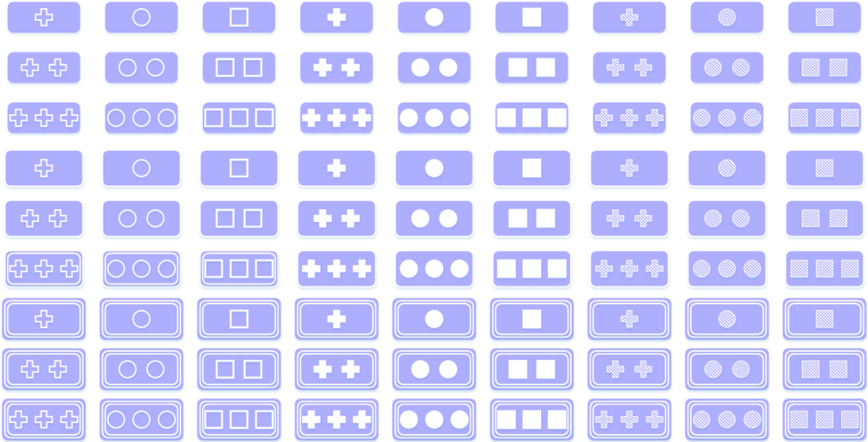


FIGURE 10.2 Material used by Bruner et al. (1956) to study concept identification.
Source: From J. S. Bruner, J. J. Goodnow, and G. A. Austin. *A study of thinking*.

Types of Concepts

- conjunction of *n* attributes
  e.g., "circles with two frames" as two-dimensional conjunctive concept
- disjunction of *n* attributes
  e.g., one object and/or filled
- Task: Detect relevant dimensions
  (Subjects are informed which type of hypothesis with how many dimensions is searched for)

Experimental Settings

- Selection task:
  all cards are presented (in ordered form), after instructor shows a positive example, subjects select another example of which they assume that it is positive, instructor gives feedback, etc. until correct rule can be named
- Reception task:
  instructor presents one example together with the information whether it belongs to the searched for concept or not, subjects must name hypothesis after each step until correct rule is found

Identified Strategies

- *conservative focusing:*
  form hypothesis, select positive example, if not correct, vary one dimension (preferred)
- *focus gambling:*
  vary more than one dimension in one step
  (might lead to success faster, never slower)
- *successive scanning:*
  start with one hypothesis, keep until falsified, select a new hypothesis compatible with all examples seen so far
  (working memory problem)
- *simultaneous scanning:*
  start with all hypotheses, delete falsified ones

Variations

- *Selection from unordered set (higher difficulty):* results in successive scanning
- *Use meaningful material:* results in successive scanning (“plausible” hypotheses)
- with disjunctive concepts it is easier to focus on negative examples, humans have massive problems learning disjunctive categories (less than 20% if first example is positive; about 50% if first example is negative)

Hypotheses-Theories

- Most theories and studies are concerned with the question how the next hypothesis is constructed
- Assumption: search in space of all possible hypotheses (either with perfect memory about hypotheses tested so far or always a single hypothesis)
- Important findings (Trabasso & Bower, 1964; Levine, 1966):
  - No all-or-nothing learning: space of hypotheses is gradually reduced
  - Learning, i.e. generating a new hypothesis, not only after negative feedback (positive feedback helps to get rid of some hypotheses still left in hypothesis space)
- Often assumed: a switch to a new hypothesis instead of a refinement of the current hypothesis (as realized in decision tree algorithms)!

Critique

- Critique of the rule-based approach:
    - Difference between artificial and natural categories:
        - fuzzy category borders
        - objects are judged as stronger/weaker members of a category
    - Even for relatively straight-forward concepts it is often hard to give a definition using necessary and sufficient conditions (Quine, 1960):
        - Bachelor: male and unmarried
        - is the pope a bachelor?

Excursus: Necessary and Sufficient Conditions

- N is necessary for S, if it holds that $S \rightarrow N$
  - N is "weaker" than S, N cannot occur without S, $S \subset N$
  - Example "If something is a bird, it is an animal."
    It is necessary to be an animal for being a bird.
- S is sufficient for N, if it holds that $S \rightarrow N$
  - S is "stronger" than N
  - It is sufficient to be a bird for being an animal (Each bird is an animal).
- A necessary and sufficient condition requires that $S \implies N$ and $N \implies S$ hold.
  - If it rains, the street is wet. (Rain is sufficient but not necessary)
  - If it is Oct. 31, it is Halloween. (31. Oct. is necessary and sufficient for it being Halloween)

Rule-based Concepts and Decision Tree Learning

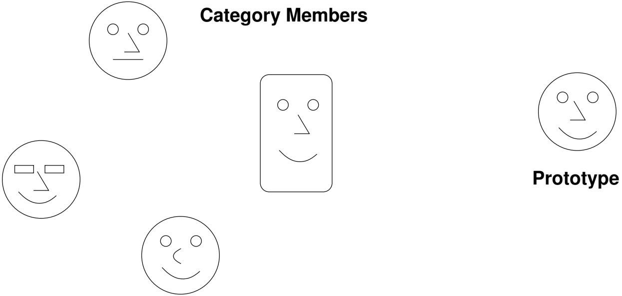
- Bruner et al.'s work inspired the first decision tree algorithm by Hunt, Marin & Stone, Experiments in Induction, 1966
- This Concept Learning System (CLS) explored at each level all possible DTs of this depth and selected the one with the least mis-classification costs.

Prototypes

- Characteristic attributes (instead of defining attributes)
- e.g.: characteristic for a bird is building a nest, that it can fly, even if not all birds have these characteristics (penguin, ostric)
- Prototype theory (Medin, Rosch):
  - There is no attribute which must be shared by all members of a category, but there are characteristic attributes shared by large subsets of objects
  - "Family resemblance": cf. Wittgenstein "Spiel"
  - Prototype is an "average" object, having all characteristic features
  - Prototype itself typically has NO correspondence to a real object

Family Resemblance

- Example (Medin et al. 2001)



Posner and Keele, 1968

- *Study Items:*
  High distortions of a prototype pattern
- *Test Items:*
  Prototype, low distortion, high distortion, random
- *Result:*
  Prototype is classified to belonging to the learned category!
- *Interpretation:*
  Similarity-based creation of prototypes as mean of the features of the exemplars

Dot Patterns

Study items

Test items

Critique

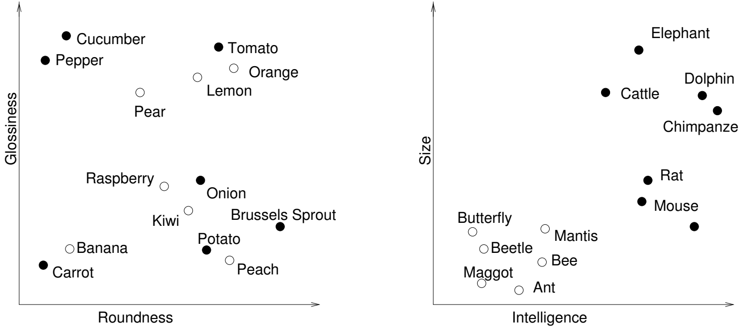
- If only a prototype is represented (exemplars forgotten), then
  - no information about variability (e.g. standard deviation of characteristic attributes)
  - no information about relative size of category
  - no consideration of attribute correlations (smaller birds typically can sing, larger birds not)
- Experiments show that humans use these kinds of information
- Context effects: e.g. typical beverage (office: coffee; construction site: beer)

Context Effects

- Labov (1972): The concept "cup"
- Naming experiment: when is an object called "cup"?
- Width-to-depth ratio and function strongly influence categorization
- Flower context necessary for classification as "vase"; but objects with high width-to-depth ratio are still classified as cup
- Holding mashed potatoes as context for "bowl" . . .

Linear Separability

- *Linear separability of categories:* only for linear separable categories, prototypes can be defined which guarantee that the prototype is more similar to all exemplars belonging to the category than to exemplars belonging to another category
- *Assumption:* linear separable categories are easier to learn (no conclusive empirical evidence)



Exemplar Theories

- e.g., Nosofsky & Johansen, 2000
- Learning as remembering all exemplars together with a category name
- Categorization
  - *Rule-based:*
    does a new object have the defining attributes?
  - *Prototype:*
    greatest similarity of feature vector with a certain prototype
  - *Exemplar-based:*
    comparison of similarity between new object and exemplars in memory
- Diagnosis of skin diseases are strongly influenced by memory of patients with similar symptoms (Brooks, Norman, Allen, 1991)

Critique

- Positive:
  Some experimental results which cannot be explained with prototypes can be explained with exemplar theory (variability of a category, context effects, correlative features)
- BUT:
  - Unrealistic assumptions about memory capacity (?)
  - What are the criteria to make a category coherent?

RULEX Theory

- Combination of rule- and exemplar-view
- Nosofsky, Palmeri & McKinley, 1994
- Learning as sequential process:
  - Search for a simple rule (one dimensional) which separates the categories
  - If no such rule exists: search for a partially successful rule and remember the exceptions explicitly

Problems with Similarity Theories

- Similarity obeys an underlying distance metric: identity, symmetry, triangle equality
- Humans perform asymmetric similarity judgements! (Tversky)
  - Which of these two pairs is more similar?
    (a) China and Russia
    (b) Vietnam and Poland
  - Which of these two pairs is less similar?
    (a) China and Russia
    (b) Vietnam and Poland
  - Answer (a) in both cases

Problems with Similarity Theories cont.

- Similarity is variable and context-dependent
  - Meeting in USA: Maine and Florida are judged as not very similar
  - Meeting in Tokyo: they are judged as similar
- If there are no restrictions for appropriate features, every two objects in the universe can be made arbitrarily similar or dissimilar

Summary

- Similarity-based theories of human concept learning can be distinguished in three groups
  - Rule-based approaches: identify defining characteristics (often too brittle for natural categories) – *compare to decision trees*
  - Prototypes: are based on "average objects" as representant for a concept
  - Exemplars: concept as collection of examples – *compare to k-nearest neighbor approaches and cluster analysis*