Similarity and Analogy

Praveen Paritosh
CogSci 207: Fall 2004: Week 5
11/4/04

Similarity

- Fundamental to learning, knowledge and thought (Quine).
  - Making predictions, generating expectations
- Appraisals of representation provide a lens into internal representations/processes.
- But elusive. Goodman (1972):
  - Things can be similar in an unlimited number of ways.
  - One must specify in what respects two objects are similar, or the statement is empty.

Models of Similarity

- Spatial/Geometric (Torgerson, 1958; Shepard, 1962)
- Featural (Tversky, 1977)
- Structural (Gentner, 1983)

As we go through these, look for:

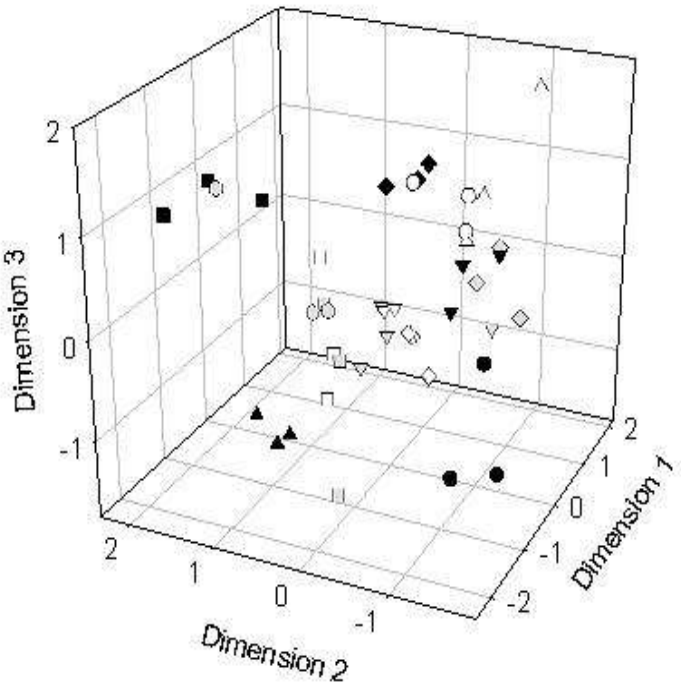
- Representational Assumptions
- Algorithms

Spatial/Geometric Models

- Objects represented by points in a n-dimensional space.
- Dissimilarity = distance between the two objects in this space.
- Distance metric = Euclidean (integral stimulus dimensions), city-block (separable).

Geometric Models – II

- Where do the dimensions come from?
  - MDS (Multidimensional Scaling): Constructs the "underlying psychological scales" from subject's pairwise similarity ratings.



Problems with Geometric Models

- Minimality: $D(A,B) \geq D(A,A)=0$
  - The letter "S" is more similar to itself than "W" is to itself (Podgorny and Garner, 1979)
  - M is more often recognised as H than M (Gilmore et al, 1979)
- Symmetry: $D(A,B)=D(B,A)$
  - North Korea is more similar to China than the other way around.
- The Triangle Inequality: $D(A,B)+D(B,C) \geq D(A,C)$

  where D(A,B) is interpreted as the dissimilarity between items A and B.

Featural Models

- Contrast Model (Tversky, 1977)
  - Entities represented by a set of features
  - Similarity a weighted combination of common and distinctive features
  - No restriction on whats a feature, e.g., Red, Symmetric, Useful, Pretty.

$$S(A,B) = q * f(A \cap B) + a * f(A - B) - b * f(B - A)$$

Featural Models - II

- Solves the problems with geometric models
- Explains how two things can be very similar and different at the same time.
  - East Germany and W Germany more similar/different than Ceylon and Nepal.
- Influencial idea, spread to fields like speech perception, pattern recognition, categorization, artificial intelligence (case-based reasoning).

Problems with Featural Models

- Entities, stories, etc. **not** an unstructured grab bag of features.
  - Above(circle, triangle) not the same as Above (triangle, circle).
  - X is a part of Y
  - X causes Y