Computational Cognitive Science

Lecture 8: Models of Categorization

Chris Lucas
(Slides adapted from Frank Keller's)

School of Informatics
University of Edinburgh
clucas2@inf.ed.ac.uk

13 October, 2017

Categorization

Models of Categorization

- Generalized Context Model
- General Recognition Theory
- Deterministic Exemplar Model

Distinguishing between Models

- A Probabilistic Feedback Experiment
- Model Predictions
- Model Comparison

Reading: Chapter 7 of L&F.

Categorization

Categorization is the ability of humans (and some animals) to place objects into relevant groups (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;
- categories can be high-level (e.g., animal), mid-level (e.g., dog) and low level (e.g., pug).

Standard theories of categorization are based on *prototypes* (summary category descriptions)[1], *exemplars* (sets of similar instances), or *decision boundaries* (that delineate instances)

[1] Usually in the form of average or representative instances.

Categorization

Generalized Context Model

The Generalized Context Model (GCM; Nosofsky, 1986) is an exemplar model of categorization (see lecture 1):

- during training, the model stores every instance of a category;
- during testing, a new instance activates all stored exemplars depending on similarity;
- the response probability for a category depends on the sum of the similarity with the members of this category.

In this lecture, we will assume that instances only vary along one dimension (only one feature is used for categorization).

Generalized Context Model

The distance d_{ij} between two instances i and j with feature values x_i and x_j is:

$$d_{ij} = |x_i - x_j|$$

The similarity between i and j is (where c is a parameter that scales the drop-off in similarity with increasing distance):

$$s_{ij} = \exp(-c \cdot d_{ij})$$

Then the probability of classifying instance i into category A (rather than category B) is:

$$P(R_i = A|i) = \frac{\sum_{j \in A} s_{ij}}{\sum_{j \in A} s_{ij} + \sum_{j \in B} s_{ij}}$$

General Recognition Theory

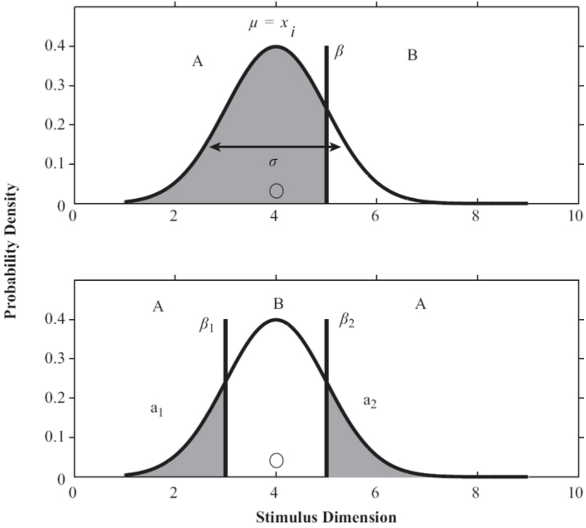
General Recognition Theory (GRT; Ashby and Townsend, 1986) is based on decision boundaries rather than exemplars:

- the feature space is divided into partitions (corresponding to categories) using decision boundaries;
- the perception of features is subject to random variation, so instances sometimes fall within a boundary, sometimes not.

We only have one feature, so category boundaries are points on a line representing the feature values. We assume that variability is normally distributed with standard deviation σ.

General Recognition Theory

Assume an instance with feature value x_i and two categories A and B with boundaries β or β_1 and β_2:



General Recognition Theory

To compute the probability of an instance being categorized as A, we need to work out the area under the curve (shaded gray):

$$P(R_i = A|i) = \int_{-\infty}^{\beta} N(x_i, \sigma) = \Phi\left(\frac{\beta - x_i}{\sigma}\right)$$

where N is the normal distribution with mean x_i and SD σ, and Φ is the cumulative normal distribution.

If there are two decision boundaries β_1 and β_2:

$$P(R_i = A|i) = \Phi\left(\frac{\beta_1 - x_i}{\sigma}\right) + \left(1 - \Phi\left(\frac{\beta_2 - x_i}{\sigma}\right)\right)$$

Note that the βs are free parameters in GRT.

Deterministic Exemplar Model

It's difficult to distinguish between GCM and GRT:

- they assume not only distinct representations (exemplars vs. decision boundaries);
- but also generate responses differently (using a probabilistic function vs. deterministically).

The Deterministic Exemplar Model (DEM; Ashby and Maddox, 1993) deconfounds these two aspects.

Deterministic Exemplar Model

DEM is identical to GCM, but with a deterministic response rule:

$$P(R_i = A|i) = \frac{\left(\sum_{j\in A} s_{ij}\right)^{\gamma}}{\left(\sum_{j\in A} s_{ij}\right)^{\gamma} + \left(\sum_{j\in B} s_{ij}\right)^{\gamma}}$$

the parameter γ controls how deterministic the model is:

- $\gamma = 1$: same response as GCM;
- $\gamma < 1$: responses are increasingly random;
- $\gamma > 1$: responses are increasingly deterministic.

For large γ, the model is deterministic: if $\sum_{j\in A} s_{ij} \gg \sum_{j\in B} s_{ij}$, then $P(A) \approx 1$; if $\sum_{j\in A} s_{ij} \ll \sum_{j\in B} s_{ij}$, then $P(A) \approx 0$.

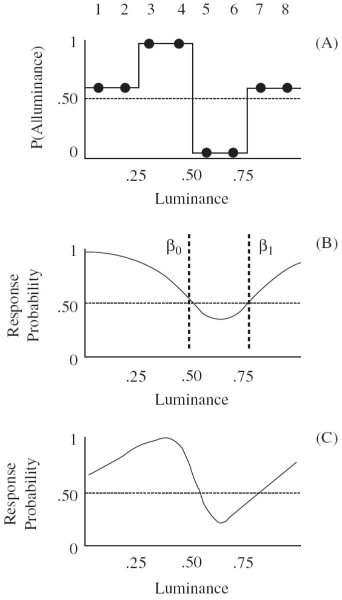
A Probabilistic Feedback Experiment

Experimental data to compare the three categorization models (Rouder and Ratcliff, 2004):

- items consist of 640×480 grids of pixels (each either white, black, or gray);
- items vary in luminance (proportion of black vs. white pixels);
- participants have to categorize each item as light or dark (A or B) and receive feedback during training;
- during feedback, items at extreme ends of the luminance space were A 60% of the time; items in the middle of the space were either A or B 100% of the time (Fig. A).

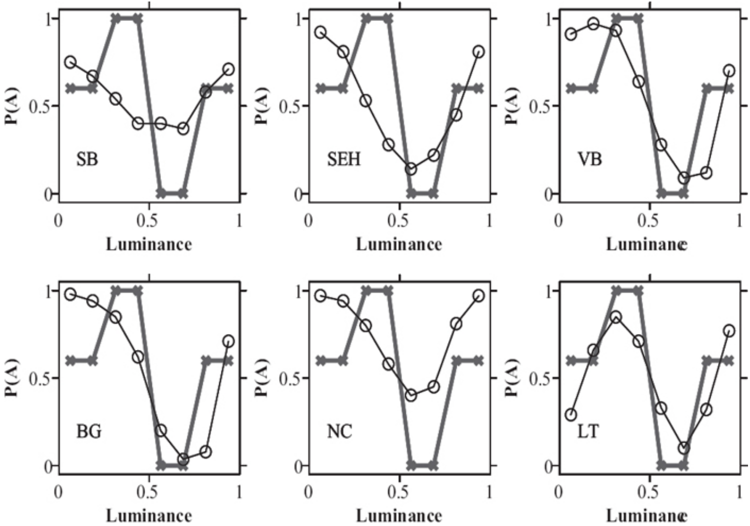
Boundary models such as GRT use simple response function (B), exemplar models such as GCM track feedback probabilities (C).

A Probabilistic Feedback Experiment



A Probabilistic Feedback Experiment

Proportion of trial categorized as *A* across six participants (circles) and feedback probabilities during training (crosses).

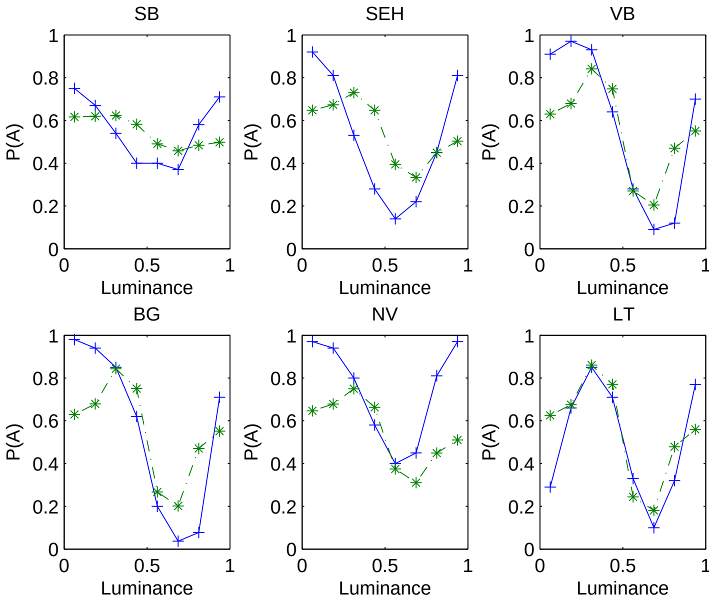


Model Predictions

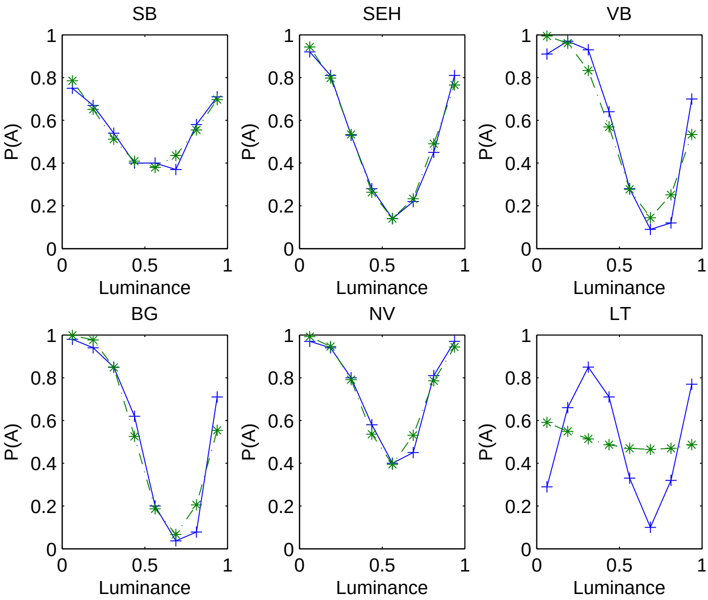
We use maximum likelihood estimation to determine the parameters for each model for each participant separately.

	GCM	GRT			DEM	
Participant	c	β_1	β_2	σ	c	γ
AB	2.29	0.29	0.80	0.29	2.44	0.95
SEH	5.29	0.33	0.82	0.17	4.89	1.13
VB	9.32	0.92	0.46	0.16	4.39	3.38
BG	9.42	0.45	0.92	0.13	4.65	3.24
NV	5.91	0.43	0.69	0.15	0.86	5.81
LT	10.28	0.12	1.26	0.78	13.51	0.69

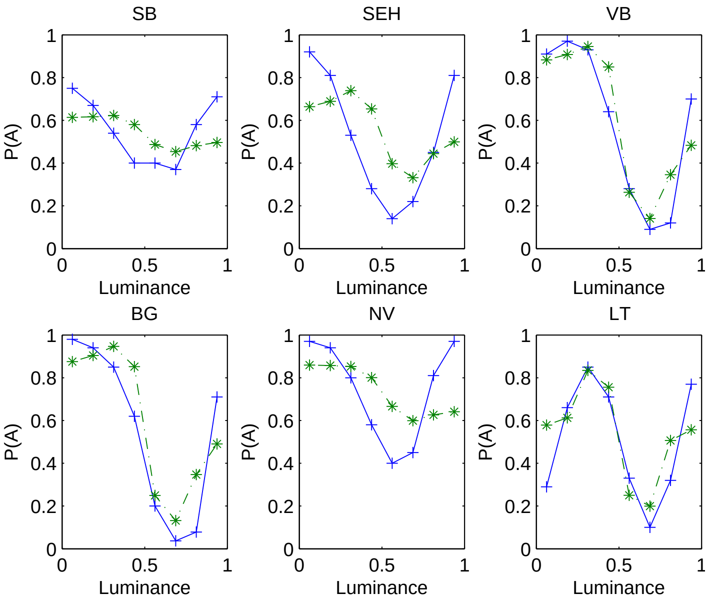
Model Predictions: GCM



Model Predictions: GRT



Model Predictions: DEM



Model Comparison

Note that the models differ in the number of parameters (GGM: 1, GRT: 3, DEM: 2), so we can't compare likelihoods directly.

We will instead compare ΔAIC, the difference in AIC with the best model (BIC gives similar results).

Participant	GCM	GRT	DEM
AB	178.42	0.00	180.38
SEH	715.39	0.00	715.83
VB	413.26	0.00	43.53
BG	703.67	0.00	337.16
NV	990.71	0.00	581.08
LT	18.78	527.95	0.00
Sum	3020.23	527.95	1857.99

What have we learned?

What have we learned about categorization?

- Overall, the data are consistent with boundary models (GRT), but not with exemplar models (GCM or DEM);
- this result can be attributed to the representations (boundary vs. exemplars), not to the decision function (probabilistic vs. deterministic): DEM is deterministic like GRT;
- however, there is individual variation: one participant (LT) is well modeled by exemplar models;
- alternative, LT's data could be captured by GRT if we assume a third decision boundary;
- the recent literature proposes *hybrid models:* new experiments suggests people sometimes use a boundary approach and sometimes an exemplar approach to categorization.

Summary

- Categorization is the process of grouping similar objects;
- it can be modeled by exemplar-based models (Generalized Context Model) or by models using decision boundaries (General Recognition Theory);
- GRT fits data on categorization with a single feature better than GCM (even though it has more parameters);
- however, the responses of some participants are better captured by GCM;
- this points to a hybrid model that combines exemplars and decision boundaries.

References

Ashby, F. G. & Maddox, W. T. (1993). Relations beween prototype, exemplar, and decision boundary models of categorization. *Journal of Mathematical Psychology*, *37*, 372–400.

Ashby, F. G. & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–179.

Lewandowsky, S. & Farrell, S. (2011). *Computational modeling in cognition: Principles and practice*. Thousand Oaks, CA: Sage.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–57.

Rouder, J. N. & Ratcliff, R. (2004). Comparing categorization models. *Journal of Experimental Psychology: General*, *133*, 63–82.