Cognitive Modeling

Lecture 9: Intro to Probabilistic Modeling: Rational Analysis

Sharon Goldwater

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

February 8, 2010

Reading: Anderson (2002).

Mechanistic Modeling

Traditional *mechanistic approach* to cognitive modeling (Chater and Oaksford 1999):

- analyze cognitive phenomena (memory, reasoning, language) regarding their causal structure;
- stipulate architectures and algorithms;
- develop either symbolic or connectionist computational models;
- experimental and neuroscientific data provide constraints on these models.

Mechanistic Modeling

Problems with the mechanistic approach:

- cognitive systems are seen as an assortment of arbitrary mechanisms;
- they are subject to arbitrary constraints;
- the purpose or *goal structure* of the cognitive systems is left unexplained;
- the fact that cognitive systems are well adapted to the task they are solving and the environment they operate in is left unexplained.

Rational Analysis

Alternative: *Rational Analysis* approach to cognitive modeling:

- provide *purposive* explanations: analyze cognitive system as to its goal and function;
- specify the *task* a cognitive system solves and the nature of its *environment;* assume the system is optimally adapted to task and environment;
- derive an *optimal (rational) solution* to the task, subject to constraints (resource limitations);
- historically, this approach is related to probability theory; Bayesian mathematics often used to formulate models.

Rational Analysis

Methodology (Anderson 1990, 2002):

1. **Goals:** specify precisely the goals of the cognitive system.
2. **Environment:** develop a formal model of the environment to which the systems is adapted.
3. **Computational Limitations:** make minimal assumptions about the computational limitations.
4. **Optimization:** derive the optimal behavior function, given (1)–(3).
5. **Data:** examine the empirical evidence to see whether the predictions of the behavior function are confirmed.
6. **Iteration:** repeat (1)–(5); iterative refinement.

Memory Retrieval

Items in memory decay gradually over time:

- traditional explanation (modal model) in terms of the architecture of the memory system (short term vs. long term store);
- alternative explanation: recent items are more likely to be needed again soon;
- the memory system is optimally adapted to this decline in *need probability* over time.

Example: if you read a fact about Iraq one sentence ago, then it's likely that you'll need this fact for understanding the next sentence.

Rational Analysis of Memory Retrieval

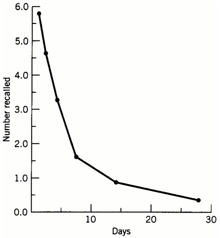
1. **Goals:** efficient retrieval of items in memory; specifically: availability of an item should match the probability that it will be needed.
2. **Environment:** need-probability p for an item is determined by the environment; items with high p should be most available.
3. **Computational Limitations:** items are searched sequentially, with a fixed cost C with searching each item.
4. **Optimization:** stop retrieving items when $pG < C$, where G is the gain associated with retrieving an item; p depends on current context and item's history of use.

Rational Analysis of Memory Retrieval

5. **Data:** need to account for two basic facts:
   - power law of forgetting: memory items decay exponentially over time: predicts need-probability decays as a power function;
   - power law of practice: reaction time decreases exponentially with no. of trials: predicts need-prob. increases as a power function of frequency of use.
6. **Iteration:** experiments that test the model:
   - investigate the role of context: recurrence of items in newspaper headlines;
   - manipulate need-probability experimentally; measure change in forgetting curves.

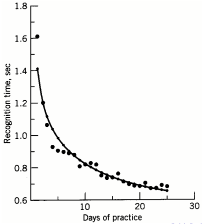
Background: Power Law of Forgetting

Number of items recalled decreases exponentially with time.



Background: Power Law of Practice

Reaction time (latency) for a given task decreases exponentially with number of practice trials.



Formalization

Anderson (1990) proposes that the need-probability p of an item A depends on its *history of use* H_A and the set of *contextual cues* Q that are present:

$$p = P(A|H_A, Q)$$

Assuming that the cues are independent of the history given A,

$$p \propto P(A|H_A)P(Q|A)$$

- $P(A|H_A)$: probability that A will be needed given its usage history;
- $P(Q|A)$: probability of observing the cues when A is needed (strength of association between A and Q).

History factor

Anderson's (1990) model of history is based on earlier model of library borrowings (Burrell 1980). Model predicts that $P(A|H_A)$

- decreases as a power function of time t since last use:

$$P(A|H_A) \propto t^{-k}$$

- increases as a power function of number of previous uses n.
- is maximized when t is equal to the interval between previous two uses.

all of which match subjects' memory behavior.

Schooler (1998) shows that these properties also hold for items in newspaper headlines.

Context factor

Holding history constant, need-probability is proportional to $P(Q|A)$.

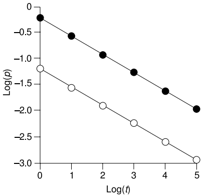
- $P(Q|A)$ is a product of separate cue strengths $P(q_i|A)$.
- Strength of cue i depends on direct association with A and association with items similar to A.

Model predicts various effects, including

- Memories are more accessible in the presence of related elements (priming).
- More subtle effects of prime frequency, number of related elements, etc.

Predictions

Relationship betw. need probability p and retention interval t:



Filled dots: strong cue associations; open dots: weak cue associations.

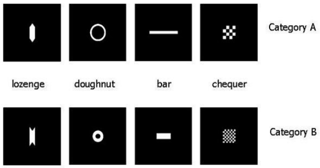
(Chater and Oaksford 1999)

Discussion

- Controversy about power laws: can arise as an artifact of averaging over subjects.
  - But, evidence that power laws of forgetting and practice also hold for individual subjects.
- Experimental evidence for both context and history factors;
- Some effects (e.g. primacy) are not predicted by the model.
  - Need to take into account underlying mechanism (capacity of short-term memory).
  - Attempts to integrate cognitive architectures with rational explanations (ACT-*R*).

Categorization

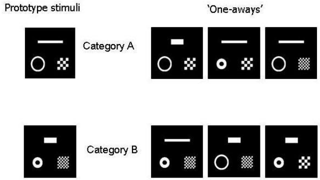
Features associated with categories:



(Lea and Wills 2008)

Categorization

Training stimuli:



(Lea and Wills 2008)

The purpose of categories

Anderson (1990) argues that psychologists often confuse *categorization* with *labeling*.

- In the real world, purpose of categories is *prediction*: objects in the same category behave similarly or have similar properties.
- The label assigned to an object is simply another feature of that object.
- Subjects' predictions may be based on a categorization that is different from the labeling used in an experiment.

Rational Analysis of Categorization

1. **Goals:** Predict features of a new object.
2. **Environment:** Disjoint partitioning of objects (species), independent variation of features within categories.
3. **Computational limitations:** Items are categorized sequentially.
4. **Optimization:** Probability that nth object has value j for feature i:

$$\sum_x P(ij|x)P(x|F_n)$$

- x: a partition, F_n: features of the n objects.

Rational Analysis of Categorization

5. **Data:** Many experimental phenomena, including effects of
   - similarity to "central tendency" of category (prototype effect);
   - similarity to specific instances in category (exemplar effect);
   - category size;
   - feature correlations within categories;
   - number of non-matching features (exponential function).

Model of Categorization

Under sequential categorization, we assume that categories of previous objects are fixed. Then

$$P(ij) = \sum_k P(ij|k)P(k|F_n)$$

- $P(ij|k)$: probability of nth object taking on jth value for feature i, given that it belongs to category k. *Depends on feature values for other objects in k*.
- $P(k|F_n)$: probability that nth object belongs to category k, given features observed for all objects. *Depends on relative sizes of categories and feature values observed for different categories*.

Discussion

- Model assumes categories are defined by items with similar features; category labels are simply features.
- Correctly predicts many experimental phenomena, including both "prototype" and "exemplar" effects, by learning multiple categories for a single label when appropriate.
- Assumes objects fall into disjoint categories; less true for non-species categories (artifacts, etc.).
- Ongoing work examining non-optimal categorization due to sequential constraints.

General discussion: Rational or irrational?

Many experiments conclude that people are 'irrational'.

- **Decision-making**: subjects don't integrate information about probability of events (*base rate neglect*).
- **Deductive reasoning**: subjects don't follow rules of logic (*Wason selection task*).

But: behavior is often far more optimal when probabilities are *experienced* or rules are framed in real-world scenarios.

Experiments often assume information is certain; real world is uncertain.

Adaptive rationality

Rational analysis assumes organisms are adapted to real world environments.

- Behavior is optimized over a range of situations, and given certain costs.
- Behavior may be non-optimal in specific situations (experiments).
- Example: Choice of local optimum over global optimum for reinforcement.

'Irrational' behavior may be the result of unnatural or unusual situations.

Summary

- Traditional modeling approaches treat the cognitive system as a collection of arbitrary mechanisms, with arbitrary performance limitations;
- they don't explain why these mechanisms cope with a complex and changing environment;
- rational analysis provides such explanations: analyze the task that a cognitive system solves, and its adaptation to the environment;
- optimal behavior functions explain why cognitive mechanisms are the way they are; provide constraints on possible theories and predict new data;
- successfully applied to memory, categorization, and other tasks.

References

Anderson, John R. 1990. *The Adaptive Character of Thought*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Anderson, John R. 2002. Is human cognition adaptive? In Thad A. Polk and Colleen M. Seifert, editors, *Cognitive Modeling*, MIT Press, Cambridge, MA, pages 1193–1228.

Burrell, Q. 1980. A simple stochastic model for library loans. *Journal of Documentation* 36:115–32.

Chater, Nicholas and Mike Oaksford. 1999. Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences* 3(2):57–65.

Lea, S. and A. Wills. 2008. Use of multiple dimensions in learned discriminations. *Comparative Cognition and Behavior Reviews* 3:115–133.

Schooler, L. 1998. Sorting out core memory processes. In Nicholas Chater and Mike Oaksford, editors, *Rational Models of Cognition*, Oxford University Press, Oxford, pages 128–155.