Title
Bayesian Approaches to Cognitive Modeling

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

Authors
Tenenbaum, Joshua
Mozer, Michael C.

Publication Date
2000

Peer reviewed

Bayesian Approaches to Cognitive Modeling

Joshua Tenenbaum
Department of Psychology
Stanford University
Stanford, CA 94305

Michael C. Mozer
Department of Computer Science
University of Colorado
Boulder, CO 80309–0430

Many, if not all, aspects of human cognition depend fundamentally on inductive inference: evaluating degrees of belief in hypotheses given weak constraints imposed by observed data. In logic-based models of cognition, the currency of belief is a binary truth value. In connectionist models of cognition, the currency of belief is an activation level. In Bayesian models of cognition, the currency of belief is a probability. The term "Bayesian" comes from Thomas Bayes, an 18th century minister who introduced a key theorem which serves as the mathematical basis of probabilistic inference. Under the single assumption that degrees of belief be represented as probability distributions, Bayes' theorem describes how the degree of belief in a hypothesis, *h*, should be updated as a result of some new evidence, *e*:

$$P(h|e) = P(e|h)P(h) \Big/ \sum_{h' \in H} P(e|h')P(h')$$

where $P(h|e)$ denotes the conditional (*posterior*) probability that *h* is true given that *e* is true, $P(h)$ denotes the unconditional (*prior*) probability that *h* is true, and $P(e|h)$ denotes the *likelihood* of observing *e* given that *h* is true. *H* denotes a set of mutually exclusive and exhaustive alternative hypotheses that could be invoked to explain *e*. A Bayesian's belief in *h* given *e* is thus a measure of how well *h* explains *e* relative to how well alternative hypotheses $h' \in H$ explain *e*.

As a normative theory of inductive inference, the Bayesian paradigm provides a principled, general-purpose framework for constructing rational models of cognition across a wide range of domains (Anderson, 1990; Knill & Richards, 1996; Oaksford & Chater, 1998). This symposium will provide a forum for representatives of Bayesian approaches from various areas of cognitive science—perception, learning, reasoning, memory, and language acquisition—to discuss both the successful aspects and the open challenges of the Bayesian paradigm. Questions to be addressed include:

- How does a Bayesian analysis provide a rational explanation for phenomena that have previously been addressed by mechanistic models? When and why does Bayes predict new phenomena that mechanistic models fail to predict? When do Bayesian analyses result in emergent predictions that are not intuitively obvious from the model's design?
- How does Bayes support the integration of disparate sources of information into a single coherent inference?
- How does Bayes allow the unification of two or more apparently distinct modes of processing into a single computational framework?
- Where does a Bayesian agent's hypothesis space come from? What kind of extra-Bayesian assumptions are needed in deriving the probabilistic generative model (prior probabilities and likelihoods) that is the foundation of a Bayesian analysis?
- The Bayesian paradigm conceives of perception and cognition as being adapted to the structure and statistics of the environment, but the mechanisms of this adaptation may vary across domains. What are the roles of evolution, learning, and habituation in adapting a Bayesian agent to the structure of a particular domain?
- There are typically many different ways to give a Bayesian analysis of a particular task. Is there always one "correct" Bayesian model? What are the criteria for deciding that one is correct?
- How can Bayesian models be tested empirically? Is the Bayesian approach falsifiable? Should it be?
- How can we reconcile the success of Bayesian models of cognition with the well-known findings from the heuristics and biases literature that "people are not Bayesian"? Could these discrepancies reflect different ways of formulating Bayesian analyses of the same tasks?
- Bayesian models, when fully implemented, are often computationally intractable. What are the implications of this intractability for a model's psychological or neural plausibility? What are the possibilities for principled approximations that might preserve the rigor of the approach in a more tractable setting? How might familiar, cognitively plausible heuristics be viewed as approximations to the full Bayesian competence?
- "Probability is not really about numbers; it is about the structure of reasoning" (G. Shafer, as quoted in Pearl, 1988). How might the structural aspects of Bayesian inference, as captured in Bayes nets and other graphical models, be important for understanding human cognition?

Speakers at the symposium will include: Michael Brent (*Bayesian modeling of segmentation and word discovery*), Evan Heit (*A Bayesian account of category-based induction*), Michael Mozer (*Temporal dynamics of information transmission in a Bayesian cognitive architecture*), and Joshua Tenenbaum (*Rules and similarity in concept learning*).