

Rational Process Models

Edward Vul (evul@mit.edu), Joshua B. Tenenbaum (jbt@mit.edu)

Brain and Cognitive Sciences, 43 Vassar St.
Cambridge, MA 02139 USA

Thomas L. Griffiths (tom_griffiths@berkeley.edu)

Dept. of Psychology, Tolman Hall
Berkeley, CA 94720 USA

Roger Levy (rlevy@ucsd.edu)

Dept. of Linguistics, 9500 Gilman Drive
La Jolla, CA 92093 USA

Mark Steyvers (msteyver@uci.edu)

Dept. of Cognitive Science, 3151 Social Science Pl.
Irvine, CA 92697 USA

Craig R. M. McKenzie (cmckenzie@ucsd.edu)

Rady School of Management &
Dept. of Psychology, 9500 Gilman Drive
La Jolla, CA 92093-0553 USA

Keywords: Bayesian modeling; Computational modeling;
Process modeling; Algorithms

Summary

Rational, Bayesian accounts of cognition at the computational level have enjoyed much success in recent years: human behavior is consistent with optimal Bayesian agents in low-level perceptual and motor tasks as well as high level cognitive tasks like category and concept learning, language, and theory of mind. However, two challenges have thus far been ignored by these computational-level models.

First, the “process” challenge – Bayesian models often assume unbounded cognitive resources available for computation, yet cognitive psychology has emphasized the severe limitations imposed on human cognition: How do models at the computational level relate to traditional models from cognitive psychology concerned with psychological mechanisms such as memory and attention?

The second challenge is the “scaling” problem – research in machine learning and statistics has shown that exact computation is intractable for inference problems on the scale relevant to human cognition, indicating that people must be solving these problems approximately: How can Bayesian models of cognition scale up to problems of the size that the mind faces in the real world, beyond the small scales of typical laboratory tasks where these models are usually tested?

This symposium brings together researchers from Machine Learning, Cognitive Science, Linguistics, and Psychology, who are working at the interface between the computational and algorithmic levels of description. The overarching theme is a new approach to answering both the “process” and “scaling” challenges by rational reverse-engineering of Bayesian algorithmic-level models.

Rational or reverse-engineering analyses are by now familiar for computational-level questions, where they ask: “what is the ideal inference (or at least, what are rational inferences with good statistical properties) given the available information and task?” The answer to this computational-level question can often be described as some form of Bayesian inference, and models derived from these considerations have en-

joyed success in explaining some aspects of human behavior. This symposium proposes an approach that asks the same question at the algorithmic level: “What is the ideal way to implement this inferential computation given constraints on space, time, energy, the scale of problem, etc?”

The answer to these problems in Bayesian statistics and machine learning is usually some form of Monte Carlo. Monte Carlo sampling is a method for approximating probability distributions by simulating a stochastic process, with long-run properties reflecting the probability distribution being simulated. Sampling is a general strategy for approximating otherwise intractable statistical inferences with limited resources: This strategy may be applied to any inference problem and is more robust to the size of the problem than other numerical methods.

Based on such reverse-engineering considerations, the panelists suggest that in a variety of domains (categorization – Griffiths; learning temporal structure – Steyvers; parsing language – Levy; and multiple object tracking – Vul) people adopt sampling algorithms to approximate optimal inference. One specific suggestion that cuts across the fields and topics of the speakers is that instead of representing a full posterior distribution, people keep track of a few sampled hypotheses. In the sequential tasks considered here, a sample-based representation of the posterior may be updated online with a particle-filtering (sequential Monte Carlo) strategy. Across the different domains and models considered in this symposium, this domain-general algorithm provides a cognitively plausible mechanism for approximating Bayes-optimal computations online. What’s most exciting is that these models make contact with (and even extend) the rich empirical paradigms of traditional cognitive psychology and can account for interesting new aspects of human behavior.

The panelists in this symposium suggest that instead of producing ad hoc cognitive process models one at a time, one for each task, the development of process models can be guided by reverse-engineering considerations. Through rational analysis of algorithms for approximate Bayesian inference, we can link up Bayesian models with traditional process accounts in cognitive psychology and suggest how Bayesian

models might be scaled up to real-world cognitive problems.

McKenzie: Challenges to rational process models

McKenzie will open the symposium by posing five important challenges to Bayesian accounts of cognition as they are extended to process level descriptions. First, the proposed process models will need to be compatible with other process-level constraints we know of in the cognitive system (e.g., limited memory and attention). Second, these models should be tractable for analogous “everyday” behaviors, not just the laboratory tasks being studied. Third, the authors should provide evidence that subjects process information in the manner described by the models, not just that subjects arrive at answers consistent with model predictions. Fourth, a successful argument for a specific rational process model should relate to alternative (perhaps heuristic) process-level descriptions. Finally, the authors will need to be clear on the ways that the approximate inferences carried out by a particular process model deviate from exact Bayesian inference, and the aspects of Bayesian computations that are preserved. McKenzie will also close the symposium by leading a discussion about the successes and shortcomings of the proposed process-level accounts of Bayesian cognition.

Griffiths: Monte Carlo as a Mechanism

Monte Carlo simulation provides a way to efficiently evaluate the probabilities of events in complex, high-dimensional probability distributions. It also provides a way to think about connecting rational models of cognition to psychological processes. Monte Carlo algorithms for many probabilistic models can be reduced to steps that are commonly used in psychological process models, such as storing items in memory and activating those items based on similarity to a target. These algorithms also provide a way to explore the effects of cognitive constraints, such as limitations on the capacity of working memory. Griffiths will summarize the basic idea behind modern Monte Carlo methods such as importance sampling and particle filtering, and illustrate how these approaches can be related to existing psychological process models.

Steyvers: Temporal change detection with particle filters

Many real-world environments involve complex changes over time where behavior that was previously adaptive becomes maladaptive. We investigate the computational problem of tracking changes in dynamic environments. We develop a particle filter model for online change detection that makes minimal demands on computational resources such as memory. The model can explain the large individual differences that we find in behavioral change detection tasks. In several experiments, we found that observers range in their overall performance from random to near optimal behavior. We model these individual differences by varying the number of particles, with more particles available for the good performers. The number of particles can be interpreted as the amount

of cognitive resources an observer is utilizing to solve a problem with a large number of particles leading to good approximations to the optimal solution. We also found that observers vary in their propensity to detect changes – some individuals detect too many and some too few changes. We model this by varying the prior belief about the probability of a change occurring in the data generating process. Overall, the particle filter model is able to model most of the observed individual differences. We also applied heuristic models to our data but found that the heuristic model cannot easily account for the range of individual differences.

Levy: Online language parsing

Language comprehension in humans is significantly constrained by memory, yet rapid, highly incremental, and capable of utilizing a wide range of contextual information to resolve ambiguity and form expectations about future input. In contrast, most of the leading psycholinguistic models and fielded algorithms for natural language parsing are non-incremental, have run time superlinear in input length, and/or enforce structural locality constraints on probabilistic dependencies between events. We present a new limited-memory model of sentence comprehension which involves an adaptation of the particle filter, a sequential Monte Carlo method, to the problem of incremental parsing. We show that this model can reproduce classic results in online sentence comprehension, and that it naturally provides the first rational account of an outstanding problem in psycholinguistics, in which the preferred alternative in a syntactic ambiguity seems to grow more attractive over time even in the absence of strong disambiguating information.

Vul: Visual attention and multiple object tracking

Multiple object tracking (MOT) is often used to measure the limits of human visual attention; however, the relationship between limited resources and performance is typically based on heuristic assumptions. We consider the computational-level solution to this problem (MOT is formally identical to an “aircraft tracking” problem), and implement an online Bayesian solution using a particle filter. The computational-level description of this problem is sufficient to account for many commonly observed phenomena in human MOT: tracking performance suffers with increases of object speed, number, and unpredictability of trajectories. However, the simple computational model does not account for the characteristic tradeoff between the speed and the number of targets that humans can track. To account for this behavior, we considered the resources limiting the process-level implementation of the computations, and asked, “how would an agent implementing this solution online allocate resources (memory, attention) to facilitate the solution?” This approach allows us to estimate what resources limit human performance without making heuristic assumptions about the relationship between resources and performance. We find that a memory limit on the amount of information that can be propagated through time accounts for human behavior in this online tracking task.