Constraints on Bayesian Explanation

Johan Kwisthout (j.kwisthout@donders.ru.nl), Iris van Rooij (i.vanrooij@donders.ru.nl) [moderator]
Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour
Montessorilaan 3, 6525 HR Nijmegen, The Netherlands

Matteo Colombo (m.colombo@uvt.nl)
Tilburg Center for Logic and Philosophy of Science, Tilburg University
PO Box 90153, 5000 LE Tilburg, The Netherlands

Carlos Zednik (czednik@uos.de)
Institute of Cognitive Science, University of Osnabrück
49069 Osnabrück, Germany

David P. Reichert (david_reichert@brown.edu)
Brown University
190 Thayer St., Providence, RI 02912, United States of America

Keywords: Bayesian inference, levels of explanation, constraints, tractability, coherent infomax, computational complexity, unification, philosophy of science

Introduction

The hypothesis that human cognition may be well characterized as a set of Bayesian computations has been the topic of considerable debate over the last two decades. Recently, critics have argued that this hypothesis is either unlikely to be true or otherwise too unconstrained to be particularly useful for explaining cognition (e.g., Bowers & Davis, 2012), whereas proponents have defended their position by stating that the Bayesian perspective has been misunderstood, is not necessarily in conflict with other perspectives on cognition, and can still be explanatorily useful as a framework for cognitive science even if under-constrained in many ways (e.g., Griffiths, Chater, Norris, & Pouget, 2012). Our position in this debate is that both sides of this debate may be right as well as wrong: Proponents may be right that the Bayesian perspective has something uniquely useful to bring to cognitive science (and then the critics are wrong in their denial of this); yet, the critics may be right that cognitive theories are explanatorily useful only if properly constrained (and then proponents are wrong in their denial of this).

With this perspective in mind, we wish to move the debate forward in a constructive way by bringing in new perspectives and proposing novel constraints that can be exploited for purposes of improving the explanatory values and virtues of Bayesian explanations of cognition. Specifically, with this symposium we aim to focus on how constraints on Bayesian explanations can be exploited in ways that are yet underrepresented and underexplored.

The symposium brings together researchers from various disciplines, contributing a variety of perspectives on how Bayesian explanations can be fruitfully constrained, drawing on theories, analyses, and results from philosophy of science, cognitive neuroscience, information theory, machine learning, and theoretical computer science.

A complexity-theoretic perspective on the preconditions for Bayesian tractability

Johan Kwisthout (joint work with Iris van Rooij)

Many Bayesian computations have been proven to be computationally intractable (NP-hard) for unconstrained input domains, even if only an approximate solution is sought. Informally, this means that computations postulated by Bayesian models can take astronomical amounts of time for their completion even for realistic sized inputs. This property seems to be in strong contrast with the ease and speed with which humans can typically make the inferences that are modeled by Bayesian models. Some critics of the Bayesian approach have taken this property of Bayesian models as a reason to reject the entire approach (e.g., Gigerenzer, 2008). In contrast, I propose that it means that tractability forms a useful constraint on Bayesian explanations of cognition. In this talk, I will elucidate the use of complexity-theoretic concepts and techniques for making Bayesian models meet the tractability constraint, building on known results from theoretical computer science (e.g., Kwisthout, 2011). I will furthermore report on recent complexity results that have lead to novel hypotheses about the conditions under which Bayesian inferences can be tractably approximated (Kwisthout & van Rooij, 2013).

Bayesian cognitive science, unification, and explanation

Matteo Colombo (joint work with Stephan Hartmann)

A recurrent claim is that the greatest value of studying cognitive phenomena such as perception, action, categorization, and decision-making, within the Bayesian framework consists in its unifying power. Several Bayesian cognitive scientists, however, implicitly assume that unification is obviously linked to explanatory power. But this link is not obvious (e.g., Morrison, 2000).

A crucial feature of adequate explanations in the cognitive sciences is that they reveal aspects of the causal structure of the mechanism that produces the phenomenon to be explained. The kind of unification afforded by the Bayesian framework to cognitive science does not necessarily reveal the causal structure of a mechanism (cf. Colombo & Seriès, 2012). Bayesian unification is the product of the mathematics rather than of a causal hypothesis concerning how different cognitive phenomena are brought about by a single type of mechanism. Nonetheless, Bayesian unification can place fruitful constraints on causal mechanical explanation, which will be elucidated in this talk.

Bayesian modeling and heuristic strategies for model-development

Carlos Zednik (joint work with Frank Jäkel)

It is generally agreed that Bayesian models in cognitive science operate at Marr's computational level of analysis (Marr, 1982). Unfortunately, it remains unclear exactly how the computational, algorithmic, and implementation levels are related.

This talk explicates inter-level relationships in terms of heuristic strategies for model-development (Zednik, in press). Specifically, Bayesian computational-level models play the heuristic role of suggesting possible algorithms to compute a particular function, and of suggesting particular ways of delineating and interpreting the components of a physical mechanism. In turn, algorithmic and mechanistic models specify memory, time, and resource limitations that constrain the cognitive tasks described by Bayesian models. In contrast to the view that Bayesian computational-level modeling is independent of low-level considerations, on this view the development of Bayesian models is constrained by, and at the same time itself constrains, the development of models at lower levels of analysis.

From Bayesian ideal observers to approximate probabilistic inference in the cortex: the case of bistable perception

David P. Reichert

The recent debate concerning the merit of Bayesian models of cognition seems due in part to a disagreement, or even confusion, with regards to what Bayesian models of cognition `are about' (Bowers & Davis, 2012; Jones & Love, 2011). There appears to be some consensus however that there is a need for models that seek to explain how (approximate) probabilistic inference could be realized in the brain. An example of this type of approach is found in my own work on modeling bistable perception as emerging from sampling-based approximate probabilistic inference, implemented in a deep neural network (Reichert et al., 2011; Reichert, 2012). Using this work as a starting point and contrasting it to related approaches on the same topic, I will argue for a more fine-grained conceptual distinction between different probabilistic or Bayesian models. I will thus characterize several conceptual dimensions that distinguish between ideal observer models and the various types of models of psychological constructs or neuronal processing. On the basis of this clarification of what Bayesian models are about, I identify challenges faced by computational neuroscience models that seek to directly map Bayesian computations onto neuronal implementations.

References

Bowers, J.S., & Davis, C.J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin, 138*(3), 389–414.

Bowers, J.S., & Davis, C.J. (2012). Is that what Bayesians believe? Reply to Griffiths, Chater, Norris, and Pouget (2012). *Psychological Bulletin, 138*(3), 423–426.

Colombo, M., & Seriès, P. (2012). Bayes in the brain. On Bayesian modelling in neuroscience. *The British Journal for Philosophy of Science*, 63, 697–723.

Griffiths T., Chater N., Norris D., & Pouget A. (2012). How the Bayesians got their beliefs (and what those beliefs actually are). *Psychological Bulletin, 138*(3), 415–422.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives in Psychological Science 3*(1), 20–29.

Jones, M., & Love, B.C. (2011). Pinning down the theoretical commitments of Bayesian cognitive models. *Behavioral and Brain Sciences*, *34*(4), 215–231.

Kwisthout, J. (2011). Most probable explanations in Bayesian networks: Complexity and tractability. *International Journal of Approximate Reasoning, 52*(9), 1452–1469.

Kwisthout, J., & van Rooij, I. (2013). Bridging the gap between theory and practice of approximate Bayesian inference. *Cognitive Systems Research*, *24*, 2–8.

Marr, D. (1982). *Vision*. New York: Henry Holt & Co.

Morrison, M. (2000). *Unifying scientific theories,* Cambridge: Cambridge University Press.

Reichert, D.P. (2012). *Deep Boltzmann Machines as Hierarchical Generative Models of Perceptual Inference in the Cortex*. PhD thesis, University of Edinburgh, Edinburgh, UK.

Reichert, D.P., Seriès, P., & Storkey, A.J. (2011). *Neuronal Adaptation for Sampling-Based Probabilistic Inference in Perceptual Bistability*, Advances in Neural Information Processing Systems 24, 2357–2365.

Zednik, C. (in press). Heuristics, descriptions, and the scope of mechanistic explanation. In C. Malaterre & P-A. Braillard (Eds.), *How does biology explain? An enquiry into the diversity of explanatory patterns in the life sciences.* Springer.