TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 6 (2014) 198–203

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12089

Utility Maximization and Bounds on Human Information Processing

Andrew Howes,[a] Richard L. Lewis,[b] Satinder Singh[c]

[a]*School of Computer Science, University of Birmingham*
[b]*Department of Psychology, University of Michigan*
[c]*Computer Science and Engineering, University of Michigan*

Received 25 June 2013; received in revised form 9 September 2013; accepted 18 September 2013

Abstract

Utility maximization is a key element of a number of theoretical approaches to explaining human behavior. Among these approaches are rational analysis, ideal observer theory, and signal detection theory. While some examples of these approaches define the utility maximization problem with little reference to the bounds imposed by the organism, others start with, and emphasize approaches in which bounds imposed by the information processing architecture are considered as an explicit part of the utility maximization problem. These latter approaches are the topic of this issue of the journal.

Keywords: Utility maximization; Rationality; Optimality; Bounds

Utility maximization is a key element of a number of theoretical approaches to explaining human behavior. Among these approaches are rational analysis (Anderson, 1990), ideal observer theory (Geisler, 2011), and signal detection theory (Swets, Tanner, & Birdsall, 1961). However, these approaches vary in the extent to which they consider bounds imposed by the information processing architecture as an explicit part of the utility maximization problem. At one extreme, rational analyses of, for example, logical reasoning (Oaksford & Chater, 1994) focus on the explanatory force of statistical distributions in the environment and require rather little in the way of information processing bounds. At the another extreme, utility maximization analyses of eye movements explain behavior in terms of limits such as the distribution of rods and cones on the retina (Geisler, 2011). The current issue of *topiCS* is focused on examples of approaches that require not only perceptual information processing bounds but other

Correspondence should be sent to Andrew Howes, School of Computer Science, The University of Birmingham Edgbaston, Birmingham, B15 2TT United Kingdom. E-mail: HowesA@bham.ac.uk

sources of bounds, including cognitive and neural bounds. While there is still substantial variation, in all cases the articles in this issue consider the limits that these computational mechanisms impose on utility maximization.

One reaction to this *topiCS* issue might be surprise that there is anything new to be said about the relationship between utility maximization, one way to determine a rational behavior, and bounds. The importance of bounds in limiting human adaptation and therefore rationality has been recognized since at least the 1950s (Simon, 1955). Approaches to explaining this relationship include heuristics (Gigerenzer, 2004), cognitive architectures (Anderson, 1990; Newell, 1990), neurally inspired architectures (McClelland et al., 2010), and the correct inference approximation algorithms studied by (Sanborn, Griffiths, and Navarro (2010). However, the details and theoretical implications of the relationship have varied greatly in these different approaches. We contend that the working definitions that underpin the articles in the current issue offer something new.

There has been much controversy over the past 20 years of the value of utility maximization and other optimality-based approaches (Anderson, 1990; Bowers & Davis, 2012; Chase, Hertwig, & Gigerenzer, 1998; Griffiths & Tenenbaum, 2006; Griffiths, Chater, Norris, & Pouget, 2012; Oaksford & Chater, 2007; Simon, 1992). The debates have sometimes been polarized, focusing repeatedly on whether or not people are optimal, or on the relationship between mechanistic theories and rationalistic theories. For Bowers and Davis (2012), for example, concerns with Bayesian approaches include the absence of empirical evidence, the arbitrariness of assumptions concerning priors and utility functions, and the neglect of evolutionary and biological constraints. See Griffiths et al. (2012) for a reply. For Gigerenzer and Todd (1999), the concern was with the computational intractability of optimization and the failure of optimization under constraints to address this problem. Simon (1992) offered a critique of Anderson's rational analysis in which he emphasized that the study of an adaptive system is not a "logical study of optimization," but an empirical study of the constraints that limit the approach to the optimum. In contrast, the articles in the current issue have put this debate to one side, assumed that utility maximization is a valuable theoretical tool, and focused on the way in which it is used in conjunction with theories of information processing bounds (the constraints) to explain behavior across a broad range of domains. The issue thereby puts greater emphasis on cognitive information processing mechanisms than has been evident in many previous optimization accounts.

In addition to making a commitment to bounds, the articles presented in this issue all use utility maximization, in some form, to determine a prediction. Lewis et al. provide a framework for understanding this variation. They argue that behavior can be explained by solving *optimal program problems* in which the selection of the programs that guide behavior is delimited by the bounds set by (a) experience of the environment; (b) information processing mechanisms; and (c) subjective utility. These three sources of bounds define the optimization problem and constitute the theory to be tested. Predictions are derived by assuming that utility maximizing programs, and therefore optimal strategies, are selected given these bounds. For these theories, the failure of a prediction to correspond to the data, a sub-optimality, is an opportunity to revise the theory of the bounds and generate further empirical tests. In this way, the optimality assumption is maintained (but not tested). It defines a rigorous class of explanation that can be used in conjunction with any well-formed theory of the bounds imposed by experience, mechanism, and utility to explain a corresponding behavior.

For Holmes and Cohen, the bounds are initially imposed by what is known about the computational properties of pools of neurons that are required to perform two alternative forced-choice tasks (2AFC). They derive an optimal performance curve from these constraints using drift diffusion assumptions and test against data. However, they find that human performance is sub-optimal *relative to this model* and they therefore propose two revisions. The first revision reflects human experience and constraints due to a noisy ability to time reward rates. Timing uncertainty corrupts estimates of reward rate, systematically biasing performance toward longer decision times. The second reflects the costs of control and the properties of spiking neurons in the cognitive and neural architectures. The idea is that participants include the costs associated with the fine adjustment of strategic parameters, the advantages of which may be small relative to the costs, in their estimates of reward rate. Holmes and Cohen suggest that adoption of these additional bounds has the potential to provide successful optimality-based explanation of observed two alternative forced-choice behaviors. This can be viewed as an illustration of the successful pursuit of an optimality-based program of research in which the focus is on incrementally uncovering the mechanism bounds, revealed by gaps between human behavior and various optimality analyses.

For Dayan, bounds are imposed by noise that is extant in the neural system but also by higher level neural mechanisms, such as working memory. These bounds have consequences for performance in tasks in which the choice of action is supported by either model-based or model-free learning. Given unlimited memory, these tasks could be solved with internal, model-based search, but unbounded optimization requires consideration of all future moves. As the number of future moves considered by humans must be bounded by the capacity and operation of the available memory structures, Dayan proposes that heuristic pruning may be an optimal strategy given these bounds. Dayan offers a rigorous and accessible overview of recent computational theories of how these bounds might explain the apparent irrationalities of choice.

For Hahn, the bounds are imposed primarily by experience. Hahn shows how experience of small samples limits perceptions of randomness. It seems that it is an empirical fact that human perceptions of randomness do not correspond to what would be expected from an unbounded mathematical analysis of probability. Hahn explains this fact by pointing out that people have a limited experience, so that it is an individual's exposure to a subset of the possible histories, through a distribution of task environments, that defines the bounds (Hahn & Warren, 2009; Hills & Hertwig, 2010). Experience, thereby, imposes a much more severe bound on the utility maximization problem faced by an individual than the assumption that people optimize to some abstract specification of the task environment.

The focus of Trimmer and Houston's article is on bounds imposed by the ecology and by natural selection. It is assumed that the constraints imposed by an animal's natural ecological setting are the adaptation environment. The assumption that an animal is optimized to its natural setting can explain departures from rationality in other settings, including laboratories. Further, the assumption that the process of natural selection itself constrains the acquisition of strategies can explain why organisms may sometimes appear sub-optimal even with respect to their natural ecological settings. Some strategies, for example, Rescorla–Wagner, may be easier to evolve than others. Again, behavior that is sub-optimal with respect to one set of constraints can be explained as optimal with respect to constraints that derive from a thorough analysis of the adaptive problem.

For Halpern, the concern is not to do with bounds imposed by neural information processing mechanisms, nor those imposed by natural selection, ecology, or experience; rather it is to do with how the fundamental theory of computation can be used to reason about bounded information processing systems in general, whether these systems are implemented in organisms or agents. To be sure, Halpern et al. use illustrations from human psychology, but the force of their proposal comes with its application of computational theory to investigate bounded information processing mechanisms. Computer Science has established tools for reasoning about computation, which is what brains do, irrespective of the physical substrate and it is these tools that Halpern et al. describe how to deploy. Halpern's contribution is to consider ways of using foundational ideas about the computational complexity of algorithms and the universality of finite state machines as tools for explaining the behavior of bounded organisms.

In addition, the issue contains two commentaries. Chater points out that optimality-based explanations are challenged by how little is known about how or whether brains represent subjective utility. Chater's article focuses on differences between theories of subjective utility as they have been construed and tested in cognitive science and biology and theories of preferences in economics. Barto discusses a number of issues. One is the fact that rather than directly delimiting behavior, processing constraints delimit programs, which in turn delimit behavioral histories. It follows that testing theories of processing constraints must involve testing the programs that mediate between processing constraints and behavior. Another issue is the extent to which each of the approaches is sufficient to account for the costs of computation.

The contributions to this issue point to many unresolved questions. Key among them is how to build a theory of the utility functions that biological organisms seek to maximize. While much has been done to answer this question (for a review see Johnson & Busemeyer, 2010), it is a question that has been relatively neglected by cognitive scientists, although see Singh, Lewis, Barto, and Sorg, (2010) in which they define an *optimal reward problem* that can be used to derive subjective utility given a bounded agent, environment, and objective/external utility function.

Another question concerns the implications of utility maximization for experimental design. For example, are instructions to be fast and accurate sufficient to control the strategic response to utility and reveal the signatures of invariant mechanisms in behavior? A broader question concerns how the behavioral sciences might systematize scientific knowledge concerning the disparate sources of bounds on behavior. Bounds imposed by physics, biology, evolution, neural substrates, task environments, cognitive architecture, and individual experience all have a role. At present, theories tend to focus on one or other source of bounds. Many more questions are articulated throughout this issue of the journal.

To end, we reflect on the fact that the usefulness of optimality analyses does not rest on the "truth" about a grand claim concerning the optimality or sub-optimality of humans. Rather, optimality analyses are used as a scientific tool for deriving the implications of theories of adaptive behavior. When behavior is predicted by such a theory, then that theory moves beyond description to explanation. The theory not only describes what might be happening but also explains why it happens, a point that has been made by many, including some of the authors in the current issue.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum.

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

Chase, V., Hertwig, R., & Gigerenzer, G. (1998). Visions of rationality? *Trends in Cognitive Sciences*, *2*(6), 206–214.

Geisler, W. S. (2011). Contributions of ideal observer theory to vision research. *Vision Research*, *51*(7), 771–781.

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. *Blackwell Handbook of Judgment and Decision Making*, 62–88.

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P.M. Todd, and the ABC Research Group, *Simple heuristics that make us smart* (pp. 3–34). New York: Oxford University Press.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science: A journal of the American Psychological Society /APS*, *17*(9), 767–73.

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

Hahn, U., & Warren, P. (2009). Perceptions of randomness: Why three heads are better than four. *Psychological Review*, *116*(2), 454–461.

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience. *Psychological Science*, *21*(12), 1787.

Johnson, J. G., & Busemeyer, J. R. (2010). Decision making under risk and uncertainty. *Wiley Interdisciplinary Reviews : Cognitive Science*, *1*(5), 736–749.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., et al. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences*, *14*, 348–356.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608–631.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, England: Oxford University Press.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2010). Rational approximations to rational models: Alternative algorithms for category learning. *Psychological Review*, *117*(4), 1144–1167.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118.

Simon, H. A. (1992). What is an explanation of behavior? *Psychological Science*, *3*(3), 150.

Singh, S., Lewis, R., Barto, A., & Sorg, J. (2010). Intrinsically motivated reinforcement learning: An evolutionary perspective. *IEEE Transactions on Autonomous Mental Development*, *2*(2), 70–82.

Swets, J., Tanner Jr, W., & Birdsall, T. (1961). Decision processes in perception. *Psychological Review*, *68*(5), 301–340.