Special Issue: Probabilistic models of cognition

Probabilistic models of cognition: Conceptual foundations

Nick Chater[1], Joshua B. Tenenbaum[2] and Alan Yuille[3]

[1]Department of Psychology, University College London, Gower Street, London, WC1E 6BT, UK
[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Building 46-4015,
77 Massachusetts Avenue, Cambridge, MA 02139, USA
[3]Departments of Statistics and Psychology, UCLA, 8967 Math Sciences Building, Los Angeles, CA 90095-1554, USA

Remarkable progress in the mathematics and computer science of probability has led to a revolution in the scope of probabilistic models. In particular, 'sophisticated' probabilistic methods apply to structured relational systems such as graphs and grammars, of immediate relevance to the cognitive sciences. This Special Issue outlines progress in this rapidly developing field, which provides a potentially unifying perspective across a wide range of domains and levels of explanation. Here, we introduce the historical and conceptual foundations of the approach, explore how the approach relates to studies of explicit probabilistic reasoning, and give a brief overview of the field as it stands today.

Introduction

The history of probabilistic models of thought is, in a sense, as old as probability theory itself. Probability theory has always had a dual aspect, serving both as a *normative* theory for 'correct' reasoning about chance events, but also as a *descriptive* theory of how people reason about uncertainty – as providing an analysis, for example, of the mental processes of an 'intelligent' juror. The title of Bernouilli's great book, Ars Conjectandi [1], 'The Art of Conjecture', nicely embodies this ambiguity, suggesting both a 'how-to' guide for better reasoning, and a survey of how the 'art' is actually practiced. That is, from its origins, probability theory was viewed as both mathematics and psychology.

From a modern perspective, this conflation seems anomalous. Mathematics has shaken free of its psychological roots, and become an autonomous, and highly formal, discipline. The philosophical thesis of 'psychologism', that mathematics (including probability) is a description of thought, fell from favour by the end of the nineteenth century. Moreover, the mathematics and psychology of probability have become divorced. The normative mathematical theory has seen spectacular developments in rigor, generality, and sophistication, going far beyond unaided intuition (see Griffiths and Yuille, Technical Introduction: Supplementary material online). Yet the descriptive study of how people judge

Corresponding author: Chater, N. (n.chater@ucl.ac.uk).

probabilities has focussed on apparently systematic patterns of fallacious reasoning about chance [2].

This Special Issue is based on the premise that reconciliation is long overdue and that the mathematics of probability is a vital tool in building theories of cognition. The articles in this issue illustrate how probability provides a rich framework for vision and motor control, learning, language processing, reasoning, and beyond. Moreover, probabilistic models can be applied in various ways – ranging from analyzing a problem that the cognitive system faces, to explicating the function of the specific neural processes that solve it. Rather than advocating a monolithic and exclusively probabilistic view of the mind, we suggest instead that probabilistic methods have a range of valuable roles to play in understanding cognition. We hope that this Special Issue will help further inspire researchers in the cognitive and brain sciences to join the project of illuminating cognition from a probabilistic standpoint; and encourage mathematicians, statisticians and computer scientists to deploy the recent remarkable conceptual and computational armoury that they have developed to help understand cognition.

The ubiquity of probabilistic inference

The cognitive sciences view the brain as an information processor; and information processing typically involves inferring new information from information that has been derived from the senses, from linguistic input, or from memory. This process of inference from old to new is, outside pure mathematics, typically uncertain. Probability theory is, in essence, a calculus for uncertain inference, at least according to the subjective interpretation of probability (Box 1). Thus, prima facie, probabilistic methods have potentially broad application to uncertain inferences from sensory input to environmental layout; from speech signal to semantic interpretation; from goals to motor output; or from observations and experiments to regularities in nature.

Probability has, however, only recently become a major focus of attention in the cognitive sciences. One reason is that the field has often focussed on computational architecture (e.g. symbolic rule-based processing vs. connectionist networks), rather than the nature of the inferences, probabilistic or otherwise, implemented in that

Box 1. Subjective probability in a nutshell

The mathematical properties of probability are relatively uncontroversial. But the *interpretation* of probability is not [44]. Most scientists are familiar with the 'frequentist' interpretation: that probabilities are limiting relative frequencies of repeated identical 'experiments', such as coin flips or dice rolls.

Crucially, this interpretation is *not* in play here – in cognitive science applications, probabilities refer to 'degrees of belief'. Thus, a person's degree of belief that a coin that has rolled under the table has come up heads might be around 1/2; this degree of belief might well increase rapidly to 1 as she moves her head, bringing the coin into view. Her friend, observing the same event, might have different prior assumptions and obtain a different stream of sensory evidence. Thus the two people are viewing the same event, but their belief states and hence their *subjective* probabilities might differ. Moreover, the relevant information is defined by the specific details of the situation. This particular pattern of prior information and evidence will never be repeated, and hence cannot define a limiting frequency.

Probabilistic analyses of perceptual, linguistic, learning or motor tasks typically follow this pattern – the issue is to understand what is *believed*, and what can be *inferred*, about the objects in the environment [3], the future state of the motor system (see Körding and Wolpert, in this issue [45]), the message being conveyed [17], or the regularities linking cause and effect (see Courville et al. [46] and Tenebaum et al. [47], in this issue).

Why should degrees of belief follow the laws of probability? There are various convergent justifications, but two of the more notable are Cox's axioms and the 'Dutch book' argument. Cox proposes several qualitative axioms that any reasonable measure of degree of belief should satisfy, and it can then be proven that only probability measures satisfy those axioms. The 'Dutch book' argument suggests that any violation of the laws of probability leads to trouble: for example, combinations of gambles that each appear fair on their own but which, together, guarantee a loss.

The subjective interpretation of probability generally aims to evaluate conditional probabilities, $Pr(h_j|d)$, that is, probabilities of alternative hypotheses, h_j (about the state of reality), given certain data, d (e.g. available to the senses). By the definition of conditional probability, for any propositions, A and B, the probability that both are true, $Pr(A, B)$, is by definition the probability that A is true, $Pr(A)$, multiplied by the probability that B is true, given that A is true, $Pr(B|A)$. Applying this identity, simple algebra gives Bayes' theorem:

$$Pr(h_j|d) = \frac{Pr(d|h_j)Pr(h_j)}{Pr(d)}$$

The centrality of Bayes theorem to the subjective approach to probability has led to the approach commonly being known as the Bayesian approach. But the real content of the approach is the subjective interpretation of probability; Bayes' theorem itself is just an elementary, if spectacularly productive, identity in probability theory.

architecture. A second reason is that formal approaches to uncertain reasoning in psychology and artificial intelligence have often been studied using non-probabilistic methods, such as default logics, non-monotonic logics, or various heuristic techniques. A third reason is that probabilistic methods have typically been viewed as too restricted in scope to be relevant to cognitive processes defined over linguistic structural descriptions, logical representations, and networks of interconnected processing units. These restrictions have been substantially reduced by remarkable technical progress in the mathematics and computer science of probabilistic models (e.g. Yuille and Kersten, this issue [3], Griffiths and Yuille, Technical Introduction: Supplementary material online).

The focus in this Special Issue is modelling cognitive abilities using sophisticated forms of probabilistic inference. The term 'sophisticated' is intended in at least two ways. First, the knowledge and beliefs of cognitive agents are modeled using probability distributions defined over structured systems of representation, such as graphs, generative grammars, or predicate logic. This development is crucial for making probabilistic models relevant to cognitive science, where structured representations are frequently viewed as theoretically central. Second, the learning and reasoning processes of cognitive agents are modeled using advanced mathematical techniques from statistical estimation, statistical physics, stochastic differential equations, and information theory.

Early examples of sophisticated probabilistic models include Grenander's pattern theory [4] and Pearl's work on Bayesian networks [5]. This approach has led to broad advances in the design of intelligent machines, with implications for computer vision, machine learning, speech and language processing, and planning and decision making. Applying these ideas to modeling aspects

of human cognition was not straightforward, despite pioneering work by Shepard [6] and Anderson [7]. Indeed, classic work in cognitive psychology by Kahneman, Tversky and their colleagues suggested that human cognition might be non-rational, non-optimal, and non-probabilistic in fundamental ways (Box 2).

Box 2. How can probability theory be hard for a probabilistic mind?

Terming probabilities as degrees of *belief* invites comparison with the folk psychological notion of belief, in which our everyday accounts of each other's behaviour are framed. This in turn suggests that people might reasonably be expected to introspect about the probabilities associated with their beliefs. In practice, people often appear poor at making such numerical judgments; and poor, too, at numerical probabilistic reasoning problems, where they appear to fall victim to a range of probabilistic fallacies [2]. The fact that people appear to be such poor probabilists might seem to conflict with the thesis that many aspects of cognition can fruitfully be modelled in probabilistic terms.

Yet this conflict is only apparent. People struggle not just with probability, but with all branches of mathematics. But the fact that, for example, Fourier analysis, is hard to understand does not imply that it, and its generalizations, are not fundamental to audition and vision. The ability to introspect about the operations of the cognitive system are the exception rather than the rule – hence, probabilistic models of cognition do not imply the cognitive naturalness of learning and applying probability theory.

Indeed, probabilistic models may be most applicable to cognitive process that are particularly well-optimized, and that solve the probabilistic problem of interest especially effectively. Thus, vision or motor control are especially tractable to a probabilistic approach; and our explicit attempts to reason about chance might often, ironically, be poorly modelled by probability theory [48]. Nonetheless, some conscious judgments have proven amenable to probabilistic analyses, such as assessments of covariation or causal efficacy [23,25], uncertain reasoning over causal models [49,50], or predicting the extent of everyday events [51]. But unlike textbook probability problems, these are exactly the sorts of crucial real-world judgments for which human cognition should be expected to be optimized.

Yet it seems increasingly plausible that human cognition may be explicable in rational probabilistic terms and that, in core domains, human cognition approaches an optimal level of performance. Nevertheless, these new ideas remain unfamiliar to most cognitive scientists and it is only in the last five years that they have started making a significant impact on the field.

Vision is the subfield of cognitive science where these models are most advanced (e.g. [8]). Recent work [9] has used these techniques to extend classical ideal observer models (which, by definition, perform optimally) to complex stimuli, using Bayesian decision theory, and has shown that these models can account for many aspects of human visual perception. There have also been successes in formulating the classic Gestalt laws of perceptual organization in terms of probabilistic models [10] that relate to earlier psychological models of grouping and scene perception [11,12].

Researchers have begun to explore 'grammatical' models of vision using compositional representations [13], and developed stochastic grammars for image parsing [14]. These provide links between vision and probabilistic approaches to language processing [15] which are becoming increasingly successful at modeling experiments in psycholinguistics [16]. Indeed, whereas hierarchical symbolic representations have been viewed as problematic for probabilistic approaches, recent work in both vision and language has focussed instead on taking these as the structures over which sophisticated probabilistic models are defined. Specifically, determining which structure is most likely to underlie image or speech data requires using Bayes' theorem to combine a prior probability distribution over structures, and the probability of the data, given each structure. Computing this latter quantity amounts to 'synthesizing' the data, from candidate structures. Thus, in both language and vision, 'analysis-by-synthesis' becomes natural from a probabilistic viewpoint (Yuille and Kersten [3] and Chater and Manning [17], in this issue).

An advantage of the probabilistic perspective is that it leads to techniques for coupling different sensory modes and for integrating perception with planning. Recent work (e.g. [18]) has built on theoretical studies [19] to model the integration of visual and haptic cues, yielding good fits with experimental data. Stankiewicz *et al.* [20] have made use of modeling by Kaelbling *et al.* [21] to design an ideal observer model for how humans navigate through mazes and demonstrated that this model fits data where people navigate mazes in virtual reality.

More recently, these ideas have started making an impact in causal learning and inference. This work has built on artificial intelligence approaches to probabilistic and causal reasoning using Bayesian networks [5,22]. Cheng's causal power model [23] explains people's judgements about the strength of causal relations, as a form of parameter estimation in a simple Bayesian network. Tenenbaum, Griffiths and colleagues [24–26] show that judgments about causal structure – which variables are causes of which other variables – could be explained using Bayesian model selection among a set of candidate Bayesian networks of the same class. Gopnik

and colleagues [27] argue that children's causal learning could be modeled in this way.

Many other cognitive abilities might be explicable within this framework. How people learn the forms and meanings of words from linguistic and perceptual experience has been the subject of recent work (e.g. [28]) that draws on, and advances, state-of-the-art techniques developed in information retrieval, computational linguistics, and machine learning [15]. In earlier, related work, Anderson [7] and Shiffrin and Steyvers [29] considered how people form long-term memories, and prioritize the retrieval of memories, as a function of the statistics of their experience with the relevant events (see also Steyvers *et al.*, this issue [30]). Work on concept learning by Tenenbaum and Griffiths [31,32] showed that many phenomena of inductive generalization and similarity could be explained in terms of Bayesian inference over a hypothesis space of candidate concepts, on the assumption that the observed examples of a concept are a random sample from the concept's extension. In reasoning, work by Chater and Oaksford [33] and Krauss, Martignon, and Hoffrage [34] helps explains why people use simple heuristics for certain judgment and decision tasks as approximations to Bayesian inference. This approach relates to theoretical analyses (e.g. [35]) showing that simple heuristics can be sometimes serve as surprisingly good approximations.

Finally, studies of the temporal characteristics of human causal learning [36] suggest relationships to stochastic differential equations. Causal relationship models can be learnt by variants of the Rescorla-Wagner associative learning model. The equilibria of this model have recently been classified [37] and convergence rates analyzed using stochastic approximation theory [38]. Related techniques have been applied by Dayan and colleagues [39] to analyzing the dynamics of animal learning behavior, and to understanding connections between basic human and animal learning processes. This approaches promises to provide a deeper understanding of learning phenomena that have typically been viewed in purely mechanistic, associative terms. More generally, the probabilistic viewpoint may help explain why the computational and neural mechanisms of the brain have the structure they do.

Levels of probabilistic explanation

Sophisticated probabilistic models can be related to cognitive processes in a variety of ways. This variety can usefully be understood in terms of Marr's [40] celebrated distinction between three levels of computational explanation: the *computational* level, which specifies the nature of the cognitive problem being solved, the information involved in solving it, and the logic by which it can be solved; the *algorithmic* level, which specifies the representations and processes by which solutions to the problem are computed; and the *implementational* level, which specifies how these representations and processes are realized in neural terms.

The probabilistic models and methods described in this Special Issue have potential relevance at each of these levels. As we have noted, the very fact that much cognitive processing is naturally interpreted as uncertain inference immediately highlights the relevance of probabilistic methods at the computational level. This level of analysis is focussed entirely on the nature of the problem being solved – there is no commitment concerning how the cognitive system actually attempts to solve (or approximately to solve) the problem. Thus, a probabilistic viewpoint on the problem of, say, perception or inference, is compatible with the belief that, at the algorithmic level, the relevant cognitive processes operate via a set of heuristic tricks, rather than explicit probabilistic computations.

One drawback of the heuristics approach, though, is that it is not easy to explain the remarkable generality and flexibility of human cognition. Such flexibility seems to suggest that cognitive problems involving uncertainty may, in some cases at least, be solved by the application of probabilistic methods. Thus, we may take models such as stochastic grammars for language or vision, or Bayesian networks, as candidate hypotheses about cognitive representation. Yet, when scaled-up to real-world problems, full Bayesian computations are intractable, an issue that is routinely faced in engineering applications. From this perspective, the fields of machine learning, artificial intelligence, statistics, informational theory and control theory can be viewed as rich sources of hypotheses concerning tractable, approximate algorithms that might underlie probabilistic cognition.

Finally, turning to the implementational level, one may ask whether the brain itself should be viewed in probabilistic terms. Intriguingly, many of the sophisticated probabilistic models that have been developed with cognitive processes in mind map naturally onto highly distributed, autonomous, and parallel computational architectures, which seem to capture the qualitative features of neural architecture. Indeed, computational neuroscience [41] has attempted to understand the nervous system as implementing probabilistic calculations; and neurophysiological findings, ranging from spike trains in the blow-fly visual system [42], to cells apparently involved in decision making in monkeys [43], have been interpreted as conveying probabilistic information. How far it is possible to tell an integrated probabilistic story across levels of explanation, or whether the picture is more complex, remains to be determined by future research.

Conclusion

Sophisticated probabilistic models are finding increasingly wide application across the cognitive and brain sciences. Much of cognition is concerned with dealing, highly effectively, with spectacularly complex problems of probabilistic inference. We suggest that probabilistic methods are likely to be increasingly important theoretical tools for understanding cognition. We hope that the articles in this Special Issue will inspire future researchers to contribute further to the project of building probabilistic models of mind.

Supplementary data

Supplementary data associated with this article can be found at doi:10.1016/j.tics.2006.05.007

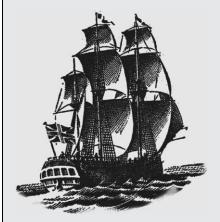## References

1 Bernoulli, J. (1713) *Ars conjectandi,* Thurnisiorum, Basel
2 Kahneman, D. and Tversky, A., eds (2000) *Choices, Values, and Frames*, Cambridge University Press
3 Yuille, A. and Kersten, D. (2006) Vision as Bayesian inference: analysis by synthesis? *Trends Cogn. Sci.* DOI:10.1016/j.tics.2006.05.002
4 Grenander, U. (1993) *General Pattern Theory: A Mathematical Study of Regular Structures*, Oxford University Press
5 Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan-Kaufman
6 Shepard, R.N. (1987) Towards a universal law of generalization for psychological science. *Science* 237, 1317–1323
7 Anderson, J.R. (1990) *The Adaptive Character of Thought*, Erlbaum
8 Kersten, D. and Yuille, A. (2003) Bayesian models of object perception. *Curr. Opin. Neurobiol.* 13, 150–158
9 Weiss, Y. *et al.* (2002) Motion illusions as optimal percepts. *Nat. Neurosci.* 5, 598–604
10 Zhu, S.C. (1999) Embedding Gestalt laws in Markov random fields. *IEEE Trans. Pattern Anal. Mach. Intell.* 21, 1170–1187
11 Shipley, T.F. and Kellman, P.J., eds (2001) *From Fragments to Objects: Segmentation and Grouping in Vision*, Elsevier
12 Chater, N. (1996) Reconciling simplicity and likelihood principles in perceptual organization. *Psychol. Rev.* 103, 566–581
13 Geman, S. (2002) Composition systems. *Q. Appl. Math.* LX, 707–736
14 Tu, Z. *et al.* (2005) Image parsing: unifying segmentation, detection, and object recognition. *Int. J. Comput. Vis.* 63, 113–140
15 Manning, C. and Schütze, H. (2000) *Foundations of Statistical Natural Language Processing*, MIT Press
16 Jurafsky, D. (2003) Probabilistic modelling in psycholinguistics: linguistic comprehension and production. In *Probabilistic Linguistics* (Bod, R. *et al.*, eds), MIT Press
17 Chater, N. and Manning, C.D. (2006) Probabilistic models of language processing and acquisition. *Trends Cogn. Sci.* DOI:10.1016/j.tics.2006.05.006
18 Ernst, M.O. and Banks, M.S. (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433
19 Clark, J.J. and Yuille, A.L. (1990) *Data Fusion for Sensory Information Processing Systems*, Kluwer Academic Publishers
20 Stankiewicz, B.J. *et al.* Lost in virtual space: human and ideal wayfinding behavior. *J. Exp. Psychol. Hum. Percept. Perform.* (in press)
21 Kaelbling, L. *et al.* (1998) Planning and acting in partially observable stochastic domains. *Artif. Intell.*, 101
22 Pearl, J. (2000) *Causality: Models, Reasoning and Inference*, Cambridge University Press
23 Cheng, P.W. (1997) From covariation to causation: a causal power theory. *Psychol. Rev.* 104, 367–405
24 Tenenbaum, J.B. and Griffiths, T.L. (2001) Structure learning in human causal induction. In *Advances in Neural Information Processing Systems Vol. 13*. (Leen, T. *et al.*, eds), pp. 59–65, MIT Press
25 Griffiths, T.L. and Tenenbaum, J.B. (2005) Structure and strength in causal induction. *Cogn. Psychol.* 51, 334–384
26 Steyvers, M. *et al.* (2003) Inferring causal networks through observations and interventions. *Cogn. Sci.* 27, 453–489

27 Gopnik, A. *et al*. (2004) A theory of causal learning in children: causal maps and Bayes nets. *Psychol. Rev.* 111, 3–32

28 Tenenbaum, J.B. and Xu, F. (2000) Word learning as Bayesian inference. In *Proc. 22nd Annu. Conf. Cogn. Sci. Soc.* (Gleitman, L.R. and Joshi, A.K., eds), pp. 517–522, Erlbaum

29 Shiffrin, R.M. and Steyvers, M. (1997) A model for recognition memory: REM: Retrieving effectively from memory. *Psychon. Bull. Rev.* 4, 145–166

30 Steyvers, M. *et al.* (2006) Probabilistic inference in human semantic memory. *Trends Cogn. Sci.* DOI:10.1016/j.tics.2006.05.005

31 Tenenbaum, J.B. (1999) Bayesian modeling of human concept learning. In *Advances in Neural Information Processing Systems Vol. 11* (Kearns, M. *et al*., eds), pp. 59–68, MIT Press

32 Tenenbaum, J.B. and Griffiths, T.L. (2001) Generalization, similarity, and Bayesian inference. *Behav. Brain Sci.* 24, 629–641

33 Chater, N. and Oaksford, M. (1999) The probability heuristics model of syllogistic reasoning. *Cogn. Psychol.* 38, 191–258

34 Krauss, S. *et al*. (1999) Simplifying Bayesian inference: the general case. In *Model-based Reasoning in Scientific Discovery* (Magnani, L. *et al*., eds), pp. 165–179, Kluwer Academic/Plenum Press

35 Coughlan, J.M. and Yuille, A.L. (2002) Bayesian A* tree search with expected O(N) node expansions: applications to road tracking. *Neural Comput.* 14, 1929–1958

36 Danks, D. *et al*. (2003) Dynamical causal learning, *Advances in Neural Information Processing Systems Vol. 15*. (Becker, S. *et al.*, eds), pp. 67–74, MIT Press

37 Danks, D. (2003) Equilibria of the Rescorla–Wagner model. *J. Math. Psychol.* 47, 109–121

38 Yuille, A.L. (2004) The Rescorla–Wagner algorithm and maximum likelihood estimation of causal parameters. *Adv. Neural Inf. Process. Syst.* 17, 1585–1592

39 Dayan, P. *et al*. (2000) Learning and selective attention. *Nat. Neurosci.* 3, 1218–1223

40 Marr, D. (1982) *Vision*, W.H. Freeman

41 Dayan, P. and Abbott, L.F. (2001) *Theoretical Neuroscience: Computational and Mathematical Modelling of Neural Systems*, MIT Press

42 Rieke, F. *et al*. (1999) *Spikes*, MIT Press

43 Shadlen, M.N. and Gold, J.I. (2004) The neurophysiology of decision-making as a window on cognition. In *The Cognitive Neurosciences*, 3rd edn (Gazzaniga, M.S., ed.), pp. 1229–1241, MIT Press

44 Hájek, A. (2003) Interpretations of probability. In *The Stanford Encyclopedia of Philosophy* (Zalta, E.N., ed.), http://plato.stanford.edu/archives/sum2003/entries/probability-interpret/

45 Körding, K.P. and Wolpert, D.M. (2006) Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.* DOI:10.1016/j.tics.2006.05.003

46 Courville, A.C. *et al.* (2006) Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* DOI:10.1016/j.tics.2006.05.004

47 Tenenbaum, J.B. *et al.* (2006) Theory-based Bayesian models of inductive learning and reasoning. *Trends Cogn. Sci.* DOI:10.1016/j.tics.2006.05.009

48 Stewart, N. *et al.* Decision by sampling. *Cogn. Psychol.* (in press)

49 Krynski, T.R. and Tenenbaum, J.B. (2003) The role of causal models in statistical reasoning. In *Proc. 25th Annu. Conf. Cogn. Sci. Soc.* (Alterman, R. and Kirsh, D., eds), pp. 693–698, Erlbaum

50 Sloman, S. and Lagnado, D. (2004) Do we 'do'? *Cogn. Sci.* 29, 5–39

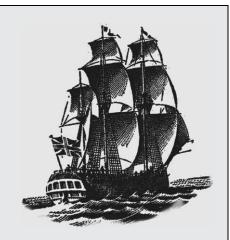51 Griffiths, T.L. and Tenenbaum, J.B. Optimal predictions in everyday cognition. *Psychol. Sci.* (in press)