Informavores: Active information foraging and human cognition

Doug Markant (*Moderator*) and **Todd Gureckis**
Dept. of Psychology, New York University

Björn Meder and **Jonathan D. Nelson**
Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development

Peter Pirolli
Palo Alto Research Center

Chen Yu
Dept. of Psychological and Brain Sciences, Indiana University

Keywords: active learning, self-directed learning, information search, sensemaking

Just as the body survives by ingesting negative entropy, so the mind survives by ingesting information. In a very general sense, all higher organisms are informavores. - Miller (1983)

Unlike a passive sponge floating in a sea of information, humans are active information foragers – informavores – who gather and consume new knowledge. From controlling the movement of our eyes to determining which sources of news to consult, judging the quality of alternative sources of information is a critical part of our behavior. The goal of this symposium is to bring together researchers who are working to understand the cognitive processes underlying active information foraging and how they interact with more general aspects of cognition.

The study of active information search is in the midst of a renaissance. Psychological research from diverse areas ranging from developmental psychology (Schulz & Bonawitz, 2007), to higher level cognition (Nelson, 2005) to visual perception (Najemnik & Geisler, 2005) have begun to understand information gathering strategies in terms of a common set of computational principles. Simultaneous developments in machine learning on "active vision" and "active learning" (Settles, 2009) have resulted in new algorithms that optimize their own learning by focusing on useful training data. Similarly, models from optimal foraging theory from biology are being brought to bear on cognitive search processes both within and outside the mind (Pirolli, 2007; Todd, Hills, & Robbins, 2012).

This symposium aims to bring together leading experts in this area to discuss how active information foraging can be understood from a diverse set of perspectives within cognitive science. Key themes include how prior knowledge influences search (Markant & Gureckis), how information and reward interact to determine choice (Meder & Nelson), developmental patterns in information seeking behavior (Nelson et al.), information foraging in complex sensemaking tasks (Pirolli), and the allocation of attention during statistical word learning (Yu). While each represents a distinct area of research, all discussants in the symposium share a core approach of applying computational models to understand information search in humans. The symposium should appeal to a broad set of attendees including educators, developmental psychologists, cognitive modelers, and computer scientists.

The influence of priors on sequential search decisions - Doug Markant and Todd Gureckis

Normative models of information acquisition predict that people's search decisions should be strongly influenced by their prior beliefs, which capture the set of alternative hypotheses they are considering. In the present experiments we tested whether people adjusted their information search behavior in response to sequential changes in the prior. Participants played a search game in which they had to identify the shape and location of multiple hidden targets in a display (similar to the board game *Battleship*). During the task they were told that the set of possible shapes had changed, and the key question was whether they would adjust their search decisions according to the predictions of a normative model. Manipulations of the prior included changes in the frequency of certain classes of targets as well as the introduction of higher-order constraints (e.g., that all targets would have the same shape). The results showed that an individual's prior could be recovered from their sequences of search decisions, but that there were notable differences in their ability to adjust to certain changes in the hypothesis space, an effect that is not predicted by the normative model. We discuss the implications of these findings for how people generate and represent hypotheses during the course of information foraging.

Is people's information search behavior sensitive to different reward structures? - Björn Meder and Jonathan Nelson

In situations where humans actively acquire information for classification, information search preferentially maximizes accuracy (Nelson et al., 2010). However, the goal of obtaining information to improve classification accuracy can strongly conflict with the goal of obtaining information for improving utility when there are asymmetries in costs and benefits for classification decisions (e.g., in many medical diagnosis situations). Is people's information search behavior sensitive to such asymmetries? We addressed this experimentally via multiple-cue probabilistic category-learning and information-search experiments, where the payoffs corresponded either to accuracy, with equal rewards associated with the two categories, or to an asymmetric payoff function with different rewards associated with each category. We found that people have difficulties identifying the reward-maximizing (rather than accuracy-maximizing) feature in search, following a neutral category learning task. Conversely, when trained to classify under asymmetric payoffs, they had difficulties conducting accuracy-maximizing queries when searching under symmetric rewards, where the accuracy-maximizing feature maximizes reward.

Finally, if words and numbers are used to convey environmental probabilities, neither reward nor accuracy consistently predicts search. These findings emphasize the necessity of taking into account peoples goals and search-and-decision processes during learning, thereby challenging current models of information search.

Sequential information search: Theoretical, developmental and psychological issues - Jonathan Nelson, Björn Meder, Bojana Divjak, Gudny Gudmundsdottir, Matt Jones, and Laura Martignon

We theoretically and empirically examine sequential search games in which the task is to identify an unknown target object by asking yes-no questions about its features. Globally optimal decision trees were identified using exhaustive search, in two task environments. This provided a benchmark for evaluating the efficiency of heuristic and stepwise optimal experimental design (OED) approaches for selecting questions. Some, but not all, OED approaches are useful for selecting queries. A heuristic strategy, the split-half heuristic, is mathematically equivalent to information gain, a stepwise-optimal OED method. We investigated 4th-grade childrens search strategies on this task. Results show that children have good intuitions regarding questions' usefulness and search adaptively, relative to the statistical structure of the task environment. Search was especially efficient in a task environment that was representative of real-world experiences. This suggests that children can use their knowledge of real-world environmental statistics to guide their search behavior.

One issue for future work is to characterize the circumstances under which people identify efficient search strategies, especially in environments in which no stepwise strategy is optimal. A related issue is whether directed play can foment generalizable insights or intuitions.

Some models of human information foraging and sensemaking - Peter Pirolli

Information Foraging Theory aims to explain and predict how people shape their information seeking behaviors to their information environments (e.g., the Web, Twitter, social tagging systems, etc.). Typically, the key steps in developing a model of information foraging involve: (a) a rational analysis of the task and information environment (often drawing on optimal foraging theory from biology) and (b) an ACT-R computational cognitive model. I will present work on individual information seeking (e.g., on the Web), and then discuss how this work has been expanded to an ACT-R simulation of a complex sensemaking task involving geospatial intelligence analysis. This map-based task requires seeking (choosing) various types of available intelligence information, and using that information to revise probability estimates about which insurgent groups might commit a future bombing attack. The model exhibits information-seeking patterns that are comparable to humans studied on this task and both model and people deviate from a rational model based on greedy maximization of expected information gain. The model also exhibits observed human biases in seeking and using information.

Active learning and selective attention in statistical word learning - Chen Yu

There are various kinds of statistical regularities in a real-world learning environment. Therefore, statistical learners have to be selective and actively gather just-in-time information required by internal learning processes and then update their internal learning states which will consequently influence their attention and selection in the next learning moment. The present study provides evidence for the operation of selective attention in the course of cross-situational learning with two main goals. The first was to show that selective attention is critical for the underlying mechanisms that support successful statistical learning. The second one was to test whether an associative mechanism with selective attention is sufficient to explain momentary gaze data in human learning.

Toward these goals, we collected eye movement data from participants engaged in a cross-situational statistical word-learning task. Various gaze patterns were extracted, analyzed and compared between strong learners who acquired more word-referent pairs through training, and average and weak learners who learned fewer pairs. Fine-grained behavioural patterns from gaze data reveal how learners actively control their attention to gather statistical information after hearing a word, how they attend to individual objects which compete for attention within a learning trial, and how statistical evidence is selected and accumulated moment by moment, and integrated across words, across objects, and across word-object mappings. Taken together, these findings provide new evidence on the real-time active learning mechanisms operating in the human cognitive system.

References

Miller, G. (1983). Informavores. In F. Machlup & U. Mansfield (Eds.), *The study of information: Interdisciplinary messages* (p. 111-113). Wiley-Interscience.

Najemnik, J., & Geisler, W. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391.

Nelson, J. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *114*(3), 677.

Nelson, J., McKenzie, C., Cottrell, G., & Sejnowski, T. (2010). Experience Matters: Information acquisition optimizes probability gain. *Psychological Science*, *21*(7), 960.

Pirolli, P. (2007). *Information Foraging Theory: Adaptive Interaction with Information*.

Schulz, L., & Bonawitz, E. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental psychology*, *43*(4), 1045.

Settles, B. (2009). *Active Learning Literature Survey* (Tech. Rep. No. 1648). University of Wisconsin-Madison.

Todd, P., Hills, T., & Robbins, T. (2012). *Cognitive search: Evolution, algorithms, and the brain* (Vol. 9). Cambridge, MA: MIT Press.