

Models of Human Scientific Discovery

Robert L. Goldstone (rgoldsto@indiana.edu)

Department of Psychological and Brain Sciences, Indiana University
Bloomington, IN. 47401 USA

Alison Gopnik (gopnik@berkeley.edu)

Department of Psychology, University of California at Berkeley
Berkeley, CA. 94720 USA

Paul Thagard (pthagard@uwaterloo.ca)

Department of Philosophy, University of Waterloo
Waterloo, ON. N2L 3G1 Canada

Tomer D. Ullman (tomerrullman@gmail.com)

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology
Cambridge, MA. 02139 USA

Keywords: scientific discovery; computational modeling; philosophy of science; reasoning; simulations; cognitive development

Introduction

The scientific understanding of scientific understanding has been a long-standing goal of cognitive science. A satisfying formal model of human scientific discovery would be a major intellectual achievement, requiring solutions to core problems in cognitive science: the creation and use of apt mental models, the prediction of the behavior of complex systems involving interactions between multiple classes of elements, high-level perception of noisy and multiply interpretable environments, and the active interrogation of a system through strategic interventions on it – namely, via experiments. Over the past decades there have been numerous attempts to build formal models that capture what Perkins (1981) calls some of the “mind’s best work” – scientific explanations for how the natural world works by systematic observation, prediction, and testing. Early work by Hebert Simon and his colleagues (Langley, Simon, Bradshaw, & Zytkow, 1987) developed production rule systems employing heuristics to tame extremely large conjoint search spaces of experiments to run and hypotheses to test. Qualitative physics approaches seek to understand physical phenomena by building non-numeric, relational models of the phenomena (Forbus, 1984). Some early connectionist models interpreted scientific explanation in terms of emerging patterns of strongly activated hypotheses that mutually support one another (Thagard, 1992).

The last few years have seen rapid and exciting progress on modeling scientific understanding. The purpose of this symposium is to present some promising recent examples of models of scientific discovery, and describe their

applications to advancing both scientific understanding and our understanding of science. Common themes addressed by the talks include: bottom-up and top-down processes for detecting patterns, exploring new hypotheses versus testing existing hypotheses, scientific practice as a multi-level search process, perception and the postulation of hidden variables, and the relation between laboratory experiments and scientific reasoning “in the wild.”

Computational Models of Mental Models of Computational Models of the World

Robert L. Goldstone, Francisco Lara-Dammer, Douglas R. Hofstadter

In classroom and laboratory observations of students interacting with computer simulations to learn systems principles, we have observed systematic misinterpretations of these simulations. Students (and scientists) often discover erroneous patterns in the simulations, and construct underlying rules for the interactions among simulation parts that diverge substantially from the actual rules underlying the simulations. At the same time, students can also sometimes learn a considerable amount about the causal mechanisms underlying a simulation by interacting with it. To understand both the successes and failures of students’ interpretative efforts, we have developed a computational model of the process by which human learners discover patterns in natural phenomena. Our approach to modeling how people learn about a system by interacting with it follows three core design principles: 1) perceptual grounding, 2) experimental intervention, and 3) cognitively plausible heuristics for determining relations between simulation elements. In contrast to the vast majority of existing models of scientific discovery in which inputs are

presented as symbolic, often numerically quantified, structured representations, our model takes as input perceptually grounded, spatio-temporal movies of simulated natural phenomena. Given this relatively raw visual representation, instilling plausible (per/con)ceptual constraints is key to building apt and efficient relation detectors. We will consider the recognition of relations such as: collide, attract, repel, change state, transfer state, excite, and inhibit. An application of the model to the discovery of ideal gas laws will be described.

Paradigm Shifts, Hierarchical Bayesian Models and High-Temperature Search

Alison Gopnik, Thomas L. Griffiths, Christopher G. Lucas

One of the classic problems in philosophy of science involves the relationship between belief revision in everyday life and the broader changes in scientific theory formation, particularly the changes in “framework theories” leading to paradigm shifts. We articulate this relationship in terms of hierarchical Bayesian models. Higher-order beliefs constrain beliefs at the lower level, but higher-order beliefs can also be revised when evidence accumulates at a lower level. The basic mechanisms for both kinds of belief change are similar, but changes at the higher-level will appear to be more radical and “framework theory”-like than those at the lower level. We show that very young children can make inferences at the higher-level as well as the lower-level and that they are actually more willing to consider high-level changes than typical adults are. This difference may be explained in terms of different search strategies – a low-temperature narrow, “exploit” search that respects the constraints of higher-order beliefs and seeks accommodation to evidence through changes in lower-level beliefs versus a broader “high-temperature” “explore” search that moves further from current beliefs and considers both higher and lower level changes. Science may proceed by encouraging broad high-temperature search even in adulthood, and by providing conditions in which incentives for exploration outweigh those for exploitation.

Procedural Creativity in Scientific Discovery

Paul Thagard

Scientific discoveries concern not just new concepts and hypotheses, but also new methods, including: naturalistic explanation, experimentation, mathematical techniques such as calculus and statistical inference, instruments such as the telescope and spectroscopy, and taxonomy. All of these methods can be represented by rules, some of them multimodal. These rules can be generated by a cognitive process of procedural generalization. The important new biological method for gene editing, CRISPR/Cas9, illustrates the mental representations and processes that produce procedural scientific discoveries. In addition to

rules, the CRISPR case displays the importance of concepts, images, analogies, emotions, and social interactions.

Discovery Generative Programs and Approximations in Learning from Dynamical Scenes

Tomer Ullman, Andreas Stuhlmüller, Noah D. Goodman, and Joshua B. Tenenbaum

Scientific theory formation and intuitive concept discovery is often phrased over explicit logical rules, such as “if A is a magnet and B is a magnet, then A and B will interact”. However, this formulation does not track the fine-grained nature of mental simulation necessary for physical reasoning. We describe a hierarchical generative model for reasoning about intuitive physics, using a probabilistic program for going from abstract concepts of force and properties down to perceptual simulation. Within such a program, simple learning corresponds to inferences about lower-level parameters such as the particular mass or friction of an object, while more radical learning corresponds to inferences about higher-level definitions. We consider the application of this model for a relatively simple 2D domain involving attraction, repulsion, friction, and global forces. Even within such a limited domain, there is a vast number of possible theories, in the sense of a vast number of settings of the program that can potentially generate the data. We propose that efficient search of the space of possible theories, whether as a scientist or a naive adult, relies on an approximation to ideal reasoning, in the form of lower-level perceptual features trained by simulation. These features do not remove the need for a generative simulation, but they allow a reasonable learner to start with a reasonable guess of the correct theory, and to conduct a short(ish) search of the nearby mental space.

References

- Forbus, K. (1984). Qualitative Process Theory. *Artificial Intelligence*, 24, 85-168.
- Langley, P., Simon, H., Bradshaw, G., & Zytkow, M. (1987). *Scientific Discovery. Computational Explorations of the Creative Processes*. Cambridge: The MIT Press.
- Perkins, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.
- Thagard, P. (1992). *Conceptual revolutions*. Princeton, NJ: Princeton University Press.