Induction as model selection

Keith J. Holyoak*
Department of Psychology, University of California, Los Angeles, CA 90095-1563

All intelligent systems, whether children, scientists, or futuristic robots, require the capacity for induction, broadly defined to encompass all inferential processes that expand knowledge in the face of uncertainty (1). Any finite set of data is consistent with an infinite number of inductive hypotheses. The apparent accuracy of many everyday inferences therefore suggests that humans have, as the philosopher Charles Peirce put it, "special aptitudes for guessing right" (2). How can people, often restricted to sparse and noisy data, achieve some significant degree of success in discerning the underlying regularities in the world? The answer seems to require specifying inductive constraints. The report by Kemp and Tenenbaum in this issue of PNAS (3) represents an important advance in understanding the constraints that guide successful induction across a broad set of domains.

The work reported in ref. 3 is consistent with the longstanding although controversial claim that laypeople resemble "intuitive scientists" (4) in the ways in which they discover orderly patterns in the world. Sometimes lay and scientific understanding of the world exhibit striking convergence, as exemplified by the basic biological concept of a species. Evolutionary biologist Ernst Mayr, beginning in the late 1920s, spent several years in New Guinea collecting and classifying specimens of birds (5). Looking back on his findings decades later, Mayr observed that he had identified 137 species of birds, for which the natives had 136 names, conflating just two species. "The coincidence of what Western scientists called species and what the natives called species was so total that I realized the species was a very real thing in nature" (6). Mayr's observations anticipated recent experimental evidence that Western experts and indigenous people exhibit broad agreement in their classifications of bird species (7).

Kemp and Tenenbaum (3) provide a detailed computational account of how a variety of basic structural forms (e.g., partitions, chains, trees, and grids) can be inferred from various types of data (e.g., feature sets, similarity matrices, and counts of relational frequencies). The key idea is to distinguish representations explicitly at different levels of abstraction. The overall approach is to use Bayesian inference to identify a hierarchical generating model that best accounts for the observed data (Fig. 1). The optimal model will have a certain abstract form (e.g., a tree), a certain specific structure (e.g., a set of nodes and edges that constitute a particular tree), specific attributes (e.g., features associated with the objects represented by the terminal leaves of the tree), and perhaps specific parameter values (e.g., feature weights). The focus is on the two most abstract levels of the model. The algorithm systematically generates

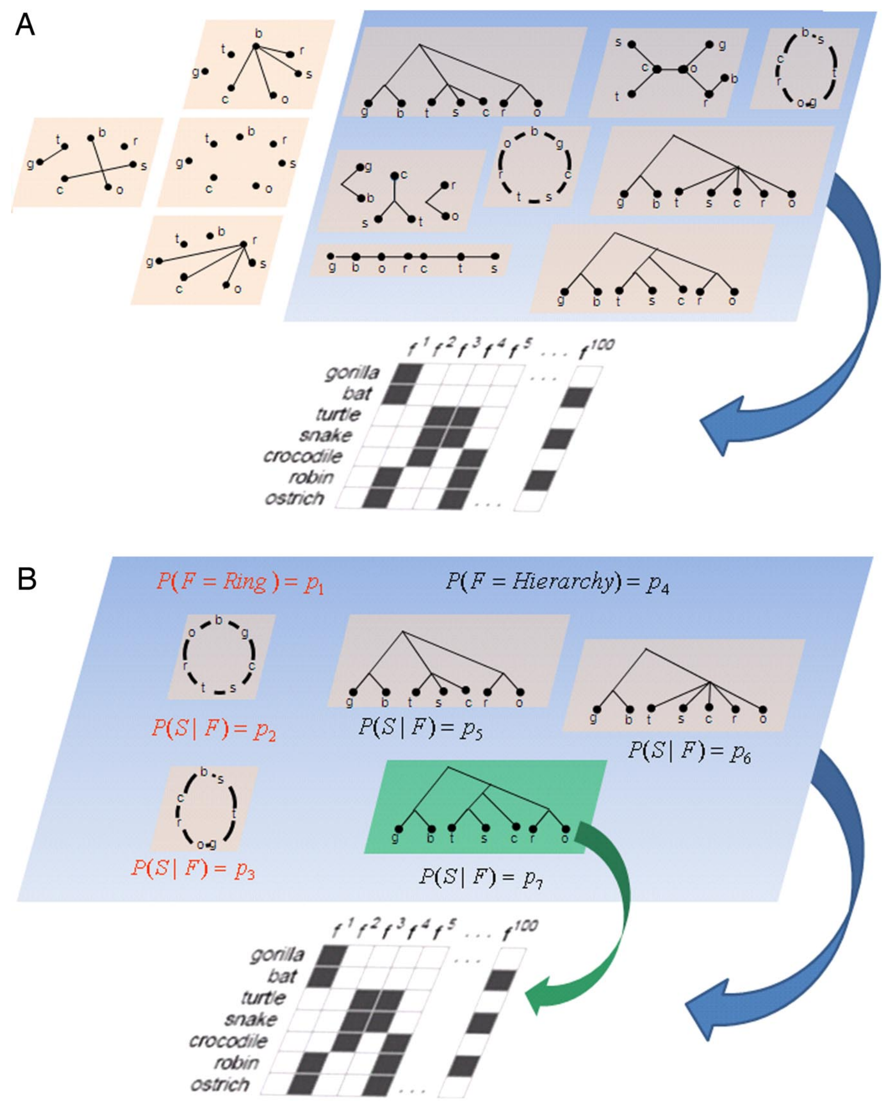

Fig. 1. Overview of hierarchical Bayesian approach to learning structural form proposed by Kemp and Tenenbaum (3), using examples of similarities among a set of animals. (*A*) The data at the bottom, in the form of a feature vector for each animal, can potentially be produced by alternative forms (ring, partition, tree, order, hierarchy) that can take on many different structures (defined by nodes and edges in graph). Likelihoods constrain the possible structural forms to those consistent with the data of feature vectors (blue background), but the set of possibilities may remain large. (*B*) The set of possible structural forms is further constrained by the prior probability of each form and by the prior conditional probability of each structure given a form. The priors for structures conditional on forms favor simpler structures (those with fewer nodes). Bayesian inference identifies the specific structure (hierarchy in green) that has maximal probability as determined by the product of the likelihood and prior knowledge: $P(S, F|D) \propto P(D|S)P(S|F)P(F)$.

The author declares no conflict of interest.

See companion article on page 10687.

*E-mail: holyoak@lifesci.ucla.edu.

candidate models from graph grammars, computes the probability of the data given each candidate model, and identifies the model with maximum posterior probability given the data. The approach captures the intuitive structural forms for a broad range of datasets, including a similarity matrix of animal species (yielding a tree), votes of Supreme Court justices (a chain), friendship relations among prisoners (a partition, reflecting a set of social cliques), and power interactions within a troop of monkeys (an order).

Kemp and Tenenbaum's proposal (3) makes explicit certain general inductive constraints that favor selection of the intuitive structural form. Perhaps the most basic is the constraint imposed by availability of models; as Jaynes put it, "If we hope to detect any phenomenon, we must use a model that at least allows for the *possibility* that it exists" (ref. 8; italics in original). The proposal assumes that prior knowledge provides a "library" of forms that are available to fit any dataset. Although these forms might be viewed as innate platonic ideals, the authors show that they could be generated systematically from a more abstract grammar of graphical structures. An important additional constraint is simplicity, inherent in Bayesian inference (where it is sometimes called the "Occam factor" after Occam's famous razor). For example, because any chain is a special case of a grid, there are necessarily more possible grids than chains for any fixed number of objects; it follows mathematically that, *ceteris paribus*, the simpler chain form will be preferred to a grid. Another aspect of simplicity is captured by penalizing larger structures (i.e., assigning lower prior probabilities to structural forms that contain more nodes). Various quantifiable measures of simplicity figure prominently in recent work on different types of human inductive inferences (9–11).

The project reported in ref. 3 extends a substantial body of research applying Bayesian methods to provide rational analyses of cognitive processes (e.g., 12–14). The most significant advance that Kemp and Tenenbaum's proposal makes over previous methods for inducing structural form stems from its hierarchical conception of models. This framework allows alternative forms to compete with one another to explain any given set of data rather than requiring an *a priori* assumption about the form appropriate for a specific dataset. Perhaps the most striking demonstration of the flexibility provided by the hierarchical approach is the finding that the model chosen for the dataset of animal similarities undergoes a qualitative shift as the number of available features increases, moving from a set of disconnected clusters to an integrated tree structure. This shift is consistent with evidence concerning the trajectory of children's acquisition of word meanings.

Any finite set of data is consistent with an infinite number of inductive hypotheses.

As a proposal focusing on the abstract level of computation (15), the work in ref. 3 opens up new questions at more detailed levels of analysis. Perhaps the most basic question is: In whose mind (if any) does the generating model explicitly exist? The case of the New Guinea natives who apparently anticipated the Western conception of species may not be typical. The elliptical form of planetary orbits and the double-helical form of DNA were inductive products of scientific reasoning, not of lay observation. The analogous question arises even for very simple forms. For example, the dominance order that describes the interactions of a troop of monkeys is clearly available to primatologists as an explicit representation, but whether it is explicit in the minds of the monkeys themselves is controversial and indeed dubious (16).

Future work will need to address the induction of structural forms based on heterogeneous types of relations, including functional and causal relations. Examples include the bauplans ("body plans") of actual or possible biological organisms, such as the form of a tetrapod or of a horse (17). In science, forms are sometimes proposed by analogy, as in the case of the wave theory of sound, developed by analogy to the behavior of waves emanating from a stone dropped into a pond (18). Even in cases where a form appears to be selected from a small set of alternatives, the induction process can be far more complex than calculating the "best fit" from among a set of forms provided by a grammar or library. Kepler's discovery of the form of planetary motion is a case in point. Having faith that "Geometry . . . supplied God with patterns for the creation of the world" (19), Kepler labored for years poring over copious and precise astronomical data provided by Tycho Brahe, progressively reconceptualizing the orbit of Mars as a circle, an oval, and finally an ellipse. As his investigation progressed, Kepler developed a crude physical model of the planetary system, according to which the sun somehow causes the motion of the planets, exerting a force that decreases with distance so that Mars moves slowest when furthest from the sun (ultimately quantified as Kepler's second law). Achieving a full understanding of the induction of structural form remains a great challenge for cognitive science.

ACKNOWLEDGMENTS. Hongjing Lu and Alan Yuille provided valuable comments on a draft, and Hongjing Lu also designed Fig. 1. This work was supported by Office of Naval Research Grant N000140810186.

1. Holland JH, Holyoak KJ, Nisbett RE, Thagard PR (1986) *Induction: Processes of Inference, Learning, and Discovery* (MIT Press, Cambridge, MA).
2. Peirce CS (1931–1958) *Collected Papers*, eds Hartshorne C, Weiss P, Burks A (Harvard Univ Press, Cambridge, MA), Vol 8.
3. Kemp C, Tenenbaum JB (2008) The discovery of structural form. *Proc Natl Acad Science USA* 105:10687–10692.
4. Kelley HH (1967) Attribution theory in social psychology. *Nebraska Symp Motiv* 15:192–238.
5. Mayr E (1941) *List of New Guinea Birds: A Systematic and Faunal List of the Birds of New Guinea and Adjacent Islands* (American Museum of Natural History, New York).
6. Yoon CK (Feb 5, 2005) Ernst Mayr, pioneer in tracing geography's role in the origin of species, dies at 100. *New York Times*.
7. Bailenson JN, Shum MS, Atran S, Medin DL, Coley JD (2002) A bird's eye view: Biological categorization and reasoning within and across cultures. *Cognition* 84:1–53.
8. Jaynes ET (2003) *Probability Theory: The Logic of Science* (Cambridge Univ Press, Cambridge, UK), p xxvi.
9. Chater N, Vitányi P (2003) Simplicity: A unifying principle in cognitive science? *Trends Cognit Sci* 7:19–22.
10. Lombrozo T (2007) Simplicity and probability in causal explanation. *Cognit Psychol* 55:232–257.
11. Lu H, Yuille AL, Liljeholm M, Cheng PW, Holyoak KJ (2008) Bayesian generic priors for causal learning. *Psychol Rev*, in press.
12. Anderson JR (1990) *The Adaptive Character of Thought* (Erlbaum, Hillsdale, NJ).
13. Chater N, Tenenbaum J, Yuille AL (2006) Probabilistic models of cognition: Where next? *Trends Cognit Neurosci* 10:292–293.
14. Griffiths TL, Tenenbaum JB (2006) Optimal predictions in everyday cognition. *Psychol Sci* 17:767–773.
15. Marr D (1982) *Vision* (Freeman, New York).
16. Penn DC, Holyoak KJ, Povinelli DJ (2008) Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behav Brain Sci* 31:109–178.
17. Arthur W (1997) *The Origin of Animal Body Plans: A Study in Evolutionary Developmental Biology* (Cambridge Univ Press, Cambridge, UK).
18. Holyoak KJ, Thagard P (1995) *Mental Leaps: Analogy in Creative Thought* (MIT Press, Cambridge, MA).
19. Kepler J (1997) *The Harmony of the World*, trans Aiton EJ, Duncan AM, Field JV (Am Philosoph Soc, Philadelphia), Book IV, p 304.