Book Reviews

Similarity and Categorization
A Review

Bradley C. Love

What is the nature of human categories? How do we form categories? What is the role of similarity in categorization? Can we formalize the answers to these questions to derive further insights and develop useful software systems? These were the questions addressed at an interdisciplinary meeting attended by psychologists, computer scientists, anthropologists, statisticians, and philosophers held at the University of Edinburgh. The edited volume *Similarity and Categorization* arises from this meeting.

The publication of *Similarity and Categorization* is timely because the study of categorization is at a theoretical crossroads. In the 1970s, similarity was thought to be the basis of categorization—an object was assigned to the category to which it was most similar. In the 1980s, a new view emerged that held that similarity was too weak and vague a construct to ground human categorization. On this view, our categories cohere by virtue of being embedded in explanatory systems (that is, our knowledge of the world is organized around theories akin to scientific theories). According to the theory-based view, judgments of similarity are largely governed by theories that determine the relevant properties for evaluation. Recently, dissatisfaction has emerged with the theory-based account of categorization, and

researchers are adopting richer and more well-developed similarity-based approaches. These second-wave similarity-based approaches address some of the shortcomings of previous approaches. The 13 chapters that form this volume are broad in scope but can be characterized as couching categorization in terms of more sophisticated and precise notions of similarity. This work shows promise in elucidat-

---

Similarity and Categorization, Ulrike Hahn and Michael Ramscar, editors, Oxford University Press, New York, 279 pp., 2001, ISBN 0-19-850628-7.

---

ing the nature of human categorization.

The editors, Ulrike Hahn and Michael Ramscar, do an excellent job in the introductory and concluding chapters of explicating the historical shifts in the perceived relation between similarity and categorization. I briefly overview these shifts and how the contributed chapters fit into the latest movement.

The idea that similarity is the basis for categorization is intuitive given that similar objects tend to be in the same category. On this view, the cate-

gories we have are natural partitions of the world. This position is supported by data collected by Eleanor Rosch and her colleagues in the 1970s. In other words, we have the categories we do because they preserve existing similarities among objects and are therefore informative. For example, our categories tend to be inductively powerful (for example, If I know an object is a bird, I can assume it has feathers and probably flies).

Critics of this view contend that the notion of similarity is vague and ill defined and carries no explanatory force on its own. As the philosopher Nelson Goodman (1972) noted, objects can be similar in an unlimited number of ways (for example, weigh an odd number of grams, are less than 100 meters long, are less than 101 meters long). According to Goodman, one must specify in what respect two objects are similar, or the statement is empty. If two objects are similar only because they are in the same category, then similarity is a vacuous notion, and any account of categorization based on similarity is circular. A second line of attack on similarity-based accounts is that similarity is not a powerful enough construct to account for human categorization. For example, a man who jumps into the swimming pool at a party while he is wearing a tie and sports coat is classified as drunk not because he is similar (at least in a straightforward sense) to other examples of drunk people. Rather, he is classified as drunk because this behavior is in accord with our theories of the kinds of things that drunk people do. This theory-based view of categorization has been popularized by Murphy and Medin (1985) and is reflected in AI techniques such as explanation-based learning (DeJong and Mooney 1986).

Although the theory-based view of categories does touch on some shortcomings of the earlier similarity-based approaches, the theory-based view has itself proven vague and has not made large headway in understanding human categorization. In contrast, the contributors to this volume have made headway by considering richer accounts of similarity.

For example, Arthur Markman con-
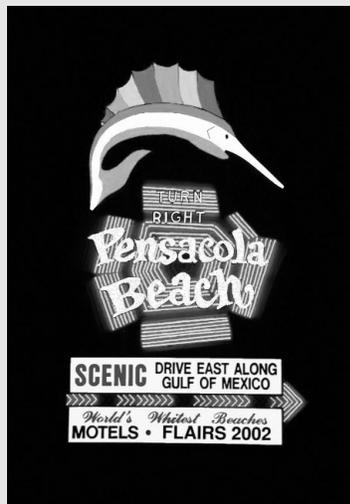
FLAIRS-2002—Pensacola, Florida

Proceedings of the Fifteenth International Florida Artificial Intelligence Research Society Conference

Edited by Susan Haller and Gene Simmons

ISBN 1-57735-141-X 500 pp., illus., index

The topics covered in this year's proceedings include interactive tutoring, machine learning, constraint satisfaction and theorem proving, robotics, genetic algorithms, agent architectures, automated reasoning, and natural language processing. The proceedings also features papers from fourteen special tracks on artificial intelligence in educational information technology, spatiotemporal reasoning, semantic web, neural networks, machine learning, evaluation of intelligent systems, categorization and concept representation models and implications, imprecise and indeterminate probabilities in artificial intelligence, uncertain reasoning, integrated intelligent systems, knowledge management, case-based reasoning, artificial intelligence in areospace, and behavioral characteristics and AI systems.

siders analogical forms of similarity computed over structured representations, as opposed to previous approaches in which representations are assumed to be spatial or featural. Markman distinguishes between differences that are related to commonalities (*alignable differences*) and differences that are not (*nonalignable differences*). For example, the number of wheels is an alignable difference for cars and motorcycles that arises out of the commonality of having wheels. An example of a nonalignable difference for this pair is a car's seat belt. Psychologically, these two types of differences are distinct. For example, high-similarity object pairs actually have more alignable differences than low-similarity pairs. Alignable differences are also remembered better than nonalignable differences and play a larger role in preference formation. Markman ponders what the metacognitive function of detecting similari-

ties is. He suggests (as do some theories of memory) that the "feeling" of similarity plays a role in coordinating retrieval strategies.

Other contributors consider the importance of causal relations in the perception of similarity. Keane, Smyth, and O'Sullivan go further and consider how the perceived similarity of two objects is affected by their processing histories. Items sharing similar processing contexts tend to be viewed as more similar. These findings parallel work demonstrating that thematic relations influence similarity. For example, a hammer and a nail are perceived as more similar because of their interaction. Work in latent semantic analysis that derives semantic representations of words through examination of usage patterns in large text corporal is a related approach. The aforementioned contributors all consider a wider range of input in the calculation of

similarity than previous similarity-based accounts.

Many of the contributors stress the importance of developing computational models. Rodriguez offers a review of case-based reasoning approaches, and Pothos and Chater develop a clustering approach to category formation. Their approach discovers the best partitioning of the data (in terms of preserving the pairwise similarity relations in the raw data). Pothos and Chater's method is based on finding the minimum–description-length solution for a set of pairwise similarity ratings that yields the greatest compression (in terms of bits). Gosselin, Archambault, and Schyns describe a model of categorization that explains the relative speed of classifying objects at different levels in a category hierarchy. Palmeri's chapter explores the time course of category learning. He evaluates how well two exemplar models (that is, models in

which each learning episode leaves an independent trace in memory, and subsequent items are classified by calculating pairwise similarities with every item stored in memory) predict the time course of classification. One model, EGCM, captures how stimulus properties become sequentially available during perception, affecting the activation of stored exemplars in memory (for example, form might be perceived before texture). The second model considered by Palmeri is the exemplar-based random walk model (EBRW). EBRW works by sequentially retrieving exemplars from memory that are fed into a diffusion decision procedure that settles on a classification when evidence for category membership crosses a decision boundary. The mathematics behind the model are well motivated, and the predictions are verified through behavioral experiments.

Other researchers strive to clarify what is meant by the terms *similarity* and *category*. Hampton's chapter distinguishes between concepts that are social constructs and those that are not. He argues that similarity-based accounts of categorization fair better when applied to concepts that are not culturally defined. For example, a bank note is not categorized as legal tender because it is similar to other acceptable notes. Rather, the note's status is tied to its origin. The properties that need to be evaluated to determine origin are not accessible to nonexperts. Hampton argues that similarity-based accounts, although viable for many routine categorization tasks, are not appropriate for such situations. Along these lines, Sloman, Malt, and Friedman show that certain measures of similarity do not predict object naming (a categorization task) and that differences occur across cultures. They argue that social convention plays a role in naming, which leads to divergences between similarity and naming. Certainly, there are some categories in which similarity does not provide the definition. Some (very few in reality) categories have tractable definitions, such as the concept of triangle. Interestingly, even when a definition is available, all category members are not treated equally (consistent

Is similarity so unconstrained that it will never provide a full account of categorization?

…

Similarity itself will not explain categorization, but understanding how similarities are processed and represented will go a long way toward understanding categorization.

with similarity-based accounts). For example, people judge canonical members of the category of odd numbers, such as the number three, as more typical of odd numbers than other odd numbers, such as the number 687, despite the fact that a clear criterion for category membership is available (Gleitman et al. 1996).

The concluding chapter offers an informative discussion of the relation between similarity and theory-based views. The editors note that the notion of theory is at least as unwieldy and vague as the notion of similarity. They also note that many of our successful theories of the world arise from, or conform to, similarity-based accounts. For example, biological taxonomies of animals ordered by DNA roughly conform to previous taxonomies that are ordered by the salient properties of the animals. The editors also note that the application and interpretation of a theory relies on similarity-based operations. For example, there is no definitional sequence of DNA that specifies what a tiger is. Evan Heit's chapter further blurs the theory-similarity dichotomy. Heit's model is exemplar based (a similarity-driven model) but captures the effects of prior knowledge and theories by seeding the model with appropriate exemplars. The model is successful in addressing a range of human data that until now were out of the reach of similarity-based models (although there are some unresolved theoretical issues surrounding how the set of initial exemplars is constructed). Heit's model

goes far in integrating theory and similarity-based accounts and demonstrates that similarity-based accounts are not as limited as some might think.

Still, the question remains, is similarity an imposter as Goodman argues? Is similarity so unconstrained that it will never provide a full account of categorization? The answer is both yes and no. Similarity itself will not explain categorization, but understanding how similarities are processed and represented will go a long way toward understanding categorization. The current volume makes substantial progress on this front.

References

DeJong, G., and Mooney, R. 1986. Explanation-Based Learning: An Alternative View. *Machine Learning* 1(2): 145–176.

Gleitman, L. R.; Gleitman, H.; Miller, C.; and Ostrin, R. 1996. Similar, and Similar Concepts. *Cognition* 58(3): 321–376.

Goodman, N. 1972. Seven Strictures on Similarity. In *Problems and Projects,* 22–32. Indianapolis, Ind.: Bobs-Merrill.

Murphy, G. L., and Medin, D. L. 1985. The Role of Theories in Conceptual Coherence. *Psychological Review* 92(3): 289–316.

Brad Love is an assistant professor in cognitive psychology at the University of Texas at Austin. He received his Ph.D. in cognitive psychology from Northwestern University and his B.S. in cognitive and linguistic sciences from Brown University. His main research interest is empirical and computational explorations of human category learning. His e-mail address is love@psy.utexas.edu.