Learning New Features of Representation

Robert Goldstone
Psychology/Cognitive Science
Indiana University
Bloomington, Indiana. 47405
rgoldsto@indiana.edu

Philippe Schyns
Psychology
University of Montreal
schyns@ai.mit.edu

Introduction

One productive and influential approach to cognition maintains that categorization, object recognition, and higher-level cognitive processes operate on the output of lower-level perceptual processing. That is, our perceptual systems provide us with a set of fixed features. These features are the inputs to higher-level cognitive processes.

Recently, researchers in psychology, computer science, and philosophy have questioned this unidirectional approach, arguing that in many situations, the high-level cognitive process being executed has an influence on the lower-level features that are created. For example, in addition to categorization being based on featural descriptions of objects, it might also be the case that the categorization process partially creates the featural descriptions that are used. Rather than viewing the "vocabulary" of primitives to be fixed by low-level processes, this view maintains that the vocabulary is dependent on the higher-level process that uses the vocabulary. This symposium will investigate several issues related to bidirectional interactions between high-level and low-level cognitive processes. First, several foundational issues will be considered. What would it mean for a system's perceptual vocabulary to be changed? How are various conceptions of perceptual change related? What experimental results would count as evidence for particular types of perceptual change?

Symposium participants will also describe new research on questions related to these broad issues. Can high-level cognitive processes (including categorization, object recognition, reasoning, comparison-making) ever cause a new vocabulary of features to emerge? If so, is the influence of the high-level process genuinely perceptual? If new vocabularies of features can be created, what constraints are there on the creation of new features? What computational mechanisms underlie the interactions (or lack of interactions) between different levels of cognitive processing?

The Pervasiveness of Constructive Processes

Douglas L. Medin
Psychology Department
Northwestern University
medin@nwu.edu

Computational models of cognition can often be thought of as consisting of representation-process pairs. And it seems almost axiomatic that one needs to determine or "fix" representations if this enterprise is to succeed. Indeed cognitive scientists are often quite skilled at creating situations where conjectures about representations receive support and attention can focus on processing principles. Recent research, however, undermines this strategy both with respect to generalizability and with respect to modeling human cognition. The talk reviews evidence from the areas of similarity comparison, decision making, and categorization illustrating the pervasiveness of constructive processes in higher level cognition. As one example, Medin, Goldstone, and Gentner (1993) show that the representation of some object or entity is in part determined by what it is compared with. Related work by Wisniewski and Medin (in press) suggests that feature interpretation is guided by domain theories. Observations such as these challenge the sharp separation between structure and process. Constructive processes are systematic, principled and they provide important constraints for computational models.

Role of Variation and Knowledge on Stimuli Segmentation: Developmental Aspects

Jean-Pierre Thibaut
Psychology
University of Liege
5, Boulevard du Rectorat
4000 LIEGE. Belgium
u017301@vm1.ulg.ac.be

Models of concepts generally assume that concepts are represented in memory in terms of components (features) that are parts of a code used to describe them. An obvious question is why people select certain features to encode things or to describe them and not others? Do they form an independent set that exists prior to experience with the categories or are they built during our interactions with objects?

Authors generally describe concept acquisition in terms of features, but if they postulate that some features are absent in children's concepts they have to explain how do children know what are the relevant features for the categories they have to learn.

When a subject has to segment an entity into relevant attributes, he can use different sources of information: low level data, the structure of the entity, the structure of the other entities in the category, and theories and expectations about the category (Wisniewski & Medin, in press).

We compared the segmentations obtained for a stimulus presented in different contexts where different sources of information were available. The structure of the category was supposed to provide information that would help subjects find a coherent segmentation of the category that was different from the one obtained in a neutral condition (a target stimulus presented alone). The hypothesis we address is that stimulus variations (i.e., the fact that different members of a category are not exactly the same; in particular, each occurrence of one potential feature can vary from one stimulus to the other) provide information that subjects can use to identify attributes of a stimulus that can differ from the ones obtained when a stimulus is presented alone. The role of expectations about an item or a category was considered by providing (or not) a theme (a category name). Results indicate that the role of the theme interacted with the context of presentation of stimuli: the segmentation of an entity presented in a category was less affected by the theme than the one obtained for a stimulus presented alone.

Developmental aspects of segmentation were considered by comparing the preceding data with data obtained with children aged 4;0 to 6;0. Results seem to indicate that children select more global features and provide less coherent segmentations. In the category condition, segmentations obtained did not differ when a theme was provided. Children's results are interpreted in terms of analytic skills and knowledge about the world.

Computational Approaches to Functional Feature Learning

Mike Mozer
Computer Science
University of Colorado, Boulder
Boulder, CO. 80309-0430
mozer@cs.colorado.edu

In this talk, I discuss computational issues in the learning of new features of representation, as well as connectionist mechanisms that integrate feature discovery and other aspects of cognition.

An active area of connectionist research involves the investigation of unsupervised learning algorithms that discover regularities in their environment. Stimuli in the environment are represented by activity patterns over a set of primitive input features; these input features are transformed into an alternative feature encoding in which regularities are made explicit. Most unsupervised algorithms attempt to achieve a compact encoding while preserving as much information as possible about the stimuli. Examples of unsupervised learning algorithms include principle components analysis and its variations (Cottrell, Munro, & Zipser, 1987; Sanger, 1989), competitive learning (Grossberg, 1976; Rumelhart & Zipser, 1985), and minimum description length and information theoretical approaches (Becker & Hinton, 1992; Hinton & Zemel, 1994).

Underlying the unsupervised learning approach is an assumption that information processing can be divided into two stages: an early unsupervised stage that constructs a representation of the environment, and a later stage that uses this representation for categorization, object recognition, and higher-cognitive processes. The conjecture motivating this approach is that unsupervised learning can produce featural encodings that are in some sense more useful for perceptual and cognitive processes than the raw, input feature representation. One can always hope for this, but because the feature extraction stage operates without regard to the later processes, the utility of the representations it produces is far from assured.

An alternative to unsupervised feature discovery is functional feature learning, in which the goal is to produce features that have functional utility in subsequent perceptual and cognitive processing. Such a goal requires that the feature learning be influenced by the nature of the subsequent processing. In a connectionist framework, there is a straightforward way to achieve this. Consider the feature extraction stage as a feedforward neural network that maps the input feature representation to a set of functional features, and the subsequent processing as another neural net that maps the functional features to, e.g., classes in a categorization task. The two stages can be trained simultaneously using a supervised learning procedure in which a teacher provides a target output, based on this target an error signal is generated, and back propagation is applied to adjust the weights in both stages.

This framework raises a concern: Because each stage is a feedforward multilayered neural net, there is no qualitative distinction between the two stages. Thus, the boundary between feature extraction and subsequent processing is ill-defined. Unless one makes additional claims about the nature of feature extraction or subsequent processing, the approach remains ambiguous.

In our work, we have made a fairly strong claim: Categorization and other cognitive processes can be adequately and appropriately characterized by simple, direct, rule- or example-based procedures of specific forms. By imposing strong constraints on the form and complexity of the cognitive processing, we delineate the job of the feature extraction stage. We illustrate with two examples:

Fedrick and Mozer (1994) have proposed a neural net categorization model whose feature extraction stage remaps a high dimensional input feature space into a lower dimensional "psychological" space, and then classifies the input based on the psychological representation. The feature extractor is a standard feedforward neural net, which feeds into Kruschke's (1992) ALCOVE, a categorization model. We show that computationally, this sort of architecture can produce better learning and generalization performance than ALCOVE alone or a multilayered feedforward neural network alone.

McMillan, Mozer, & Smolensky (1991) have studied a neural net model, called RuleNet, that learns simple string-to-string mapping rules. RuleNet consists of two components: a feature extractor and a set of simple condition-action rules -- implemented in a neural net -- that operate on the extracted features. Based on a training set of input-output examples, RuleNet performs better than a standard neural net architecture in which the processing is completely unconstrained.

By using a neural net paradigm for modeling the higher cognitive processing as well as the feature extraction stage, we have achieved a principled, unified learning procedure for the entire system, and a concrete notion of how new functional features are learned in the context of cognition.

Representation-building in Analogical Reasoning

Bob French
Computer Science
Willamette University
french@willamette.edu

Most computer models of analogy-making put into correspondence the (representations of) objects and their associated relations in one situation with the objects and their associated relations in another situation. In this talk I will argue for extending this methodology to integrate representation-building in the process of correspondence-disovery. I hope to show that representation-building cannot be isolated, even in principle, from other cognitive processes, and in particular, from finding correspondences between two different situations [Chalmers, French, & Hofstadter, 1992]. In other words, traditional artificial-intelligence models of cognitive processes that suppose the existence of a "representation module" that provide ready-made a priori representations cannot be defended. Representations are highly task-dependent and no one representation of any object or situation is appropriate for all tasks. All too often, once the appropriate representations for a particular problem have been supplied ahead of time, the solution to the problem falls out without great difficulty.

The computer models of analogy-making [French & Hofstadter, 1991; Mitchell, 1993] that I will discuss interweave the two fundamental processes of representation-building and correspondence-discovery. This is achieved in a unique manner that relies on a continual stochastic interaction between the top-down pressures of an associative concept network and the bottom-up effects of a host of low-level "perceptual agents."

A Functional Approach to Feature Learning

Philippe Schyns
Psychology
University of Montreal
schyns@ai.mit.edu

A complete theory of object representation not only concerns the way people combine features into object concepts, but it also concerns the problem of how these features could be arrived at, starting from perceptual data. Most theories of categorization assume a featural analysis of the input before categorization can start. Until recently, the features participating in the analysis were assumed to form a fixed set--a set existing before experience with object categories.

Recent studies on the interactions between perception and categorization have shown that higher cognitive processes influence feature extraction and other perceptual processes (e.g., Goldstone, in press; Schyns & Murphy, in press; Wisniewski & Medin, 1991). For example, when distinguishing between two categories, people might notice new differences; differences requiring new features for their categorization. This type of learning probably occurs when people interpret X-rays, determine the sex of chickens or categorize pre-Cambrian organisms, but it could also illuminate the way children or simple mechanisms learn the features that categorize everyday objects.

I will discuss a functional theory of features suggesting that the interplay of categorical and perceptual structures may account for the learning of a flexible vocabulary of object concepts. The functionality principle summarizes this view (Schyns & Murphy, in press): If a fragment of a stimulus categorizes objects (distinguishes members from nonmembers) the fragment is instantiated as a feature in the vocabulary of object concepts.

The functionality principle suggests that categorical structures influences the feature extraction stage. The categories people know set specific contrasts and similarities for the encoding of a new category with new features. But categorical structures are not the only source of constraints for feature extraction. Which features are extracted to represent a new categorical difference also depends on perceptual structures (not all categorical similarities/differences are equivalently salient.)

I will present experiments testing implications of the functionality principle for feature extraction and for the construction of simple conceptual hierarchies.

References

Becker, S., & Hinton, G. E. (1992). A self-organizing neural network that discovers surfaces in random-dot stereograms. Nature, 355, 161-163.

Chalmers, D., French, R., & Hofstadter, D., (1992). "High-level perception, representation, and analogy: A critque of artificial intelligenc methodology," Journal of Experimental and Theoretical Artificial Intelligence, Vol. 4(3), pp. 185-211.

Cottrell, G. W., Munro, P., & Zipser, D. (1987). Image compression by back propagation: An example of extensional programming (ICS Technical Report 8702). La Jolla, CA: University of California at San Diego, Institute for Cognitive Science.

Fedrick, K., & Mozer, M. C. (1994). Incorporating a psychologically- motivated model of categorization into a neural network improves generalization performance. Manuscript in preparation.

French, R. & Hofstadter, D. (1991). "Tabletop: An emergent, stochastic model of analogy-making," Proceedings of the Thirtennth Annual Cognitive Science Society Conference," pp. 708-713.

Goldstone, R. L. (in press). Influences of Categorization on Perceptual Discrimination. Journal of Experimental Psychology: General, 123.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding. I: Parallel development and coding of neural feature detectors. Biological Cybernetics, 23, 121-134.

Hinton, G. E., & Zemel, R. S. (1994). Using the minimum description length principle to discover factorial codes. To appear in J. Cowan, G. Tesauro, & J. Alspector (Eds.), Neural Information Processing Systems 6. San Mateo, CA: Morgan Kaufmann.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. Psychological Review, 99, 22-44.

McMillan, C., Mozer, M. C., & Smolensky, P. (1991). The connectionist scientist game: Rule extraction and refinement in a neural network. Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society (pp. 424-430). Hillsdale, NJ: Erlbaum.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. Psychological Review, 100, 254-278.

Mitchell, M. (1993). Analogy-making as Perception, MIT Press.

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. Cognitive Science, 9, 75-112.

Sanger, T. D. (1989). An optimality principle for unsupervised learning. In D. S. Touretzky (Ed.), Advances in Neural Information Processing Systems 1 (pp. 11-19). San Mateo, CA: Morgan Kauffman.

Schyns, P. G., & Murphy, G. L. (in press). The ontogeny of part representation in object concepts. In D. L. Medin (Ed.) The Psychology of Learning and Motivation, 31. Academic Press.

Wisniewski, E. J. & Medin, D. L. (in press). On the interaction of theory and data in concept learning. Cognitive Science.

Wisnieswki, E. J. & Medin, D. L. (1991). Harpoons and long sticks: The interaction of theory and similarity in rule induction In D. H. Fisher, M. J Pazzani, & P. Langley (Eds.) Concept formation: Knowledge and experience in unsupervised learning. San Mateo, CA: Morgan Kaufman.