Special Section

Learning: Association or Computation? Introduction to a Special Section

Alan M. Leslie[1]

Department of Psychology and Center for Cognitive Science, Rutgers University, Piscataway, New Jersey

Abstract

The fundamental nature of learning is a central problem in psychology. Traditionally, psychologists have assumed that learning must involve the formation of associations. Early last century, the pioneering work of Pavlov on conditioned learning in animals seemed to put this assumption beyond doubt. More recently, many psychologists came to believe that a different kind of process must underlie complex learning, such as language learning in humans, and that this process must be described as computational rather than associative. Whether complex human learning is associative or computational continues to be a subject of intense research. The articles in this Special Section turn this debate on its head by asking whether simple animal learning is associative or computational. Surprisingly, the question is still very much open, and excitingly, it appears quite tractable.

Keywords

learning; association; computation

Among the first things we hear about in Psychology 101 is how Pavlov's dog learned to associate the sound of a bell with food.[2] Pavlov's pioneering work on animal learning at the beginning of the last century became emblematic of the new experimental methods that underpinned psychology's emergence as a science. It also underpinned the most enduringly influential concept in psychology, the *association*. The idea of the association was introduced by the British empiricist philosophers in the 17th and 18th centuries. But it was in the 20th century, following Pavlov's lead, that this great idea was placed upon a solid experimental foundation by work on conditioned learning. From there, the idea came to influence contemporary neuroscience, for example, through Hebb's (1949) notion of functional cell assemblies formed by the tendency of neurons that fire together to become associated. Indeed, if there is anything both fundamental and indubitable that psychology has to say about the mind-brain, it is surely that association is its basic principle of operation.

Or is it? The postwar "cognitive revolution" posed a serious challenge to associationism. The new cognitivism, built upon ideas from electrical engineering, computer science, and the study of formal systems, emphasized specialized processing of information that was represented symbolically in a physical symbol system (Fodor, 1975; Newell, 1980). As this revolution consigned behaviorism to the wastebasket of history, many people thought that associationism would necessarily suffer the same fate. But that would have been to overlook both the resilience and the origins of the idea: In the hands of the philosopher David Hume, for example, the association was the cornerstone of a thoroughly representational (i.e., cognitive) theory of mind. And sure enough, though behaviorism all but disappeared, associationism has undergone a striking revival that started in the mid-1980s with "connectionist" models of associative learning in massively parallel networks. (A connectionist network models the activity of neurons as simple highly interconnected units that strengthen or weaken connections with their neighbors as a function of history of activity.)

Dedication

With great sadness, we note that on January 21, 2001, John Gibbon, pioneering scientist and author in this Special Section, passed away. The authors dedicate this Special Section to John's memory.

ASSOCIATION VERSUS COMPUTATION IN COGNITIVE SCIENCE

The struggle for hearts and minds continues today. In the study of language acquisition, for example, Rumelhart and McClelland (1986) put forward a pattern associator model in which past-tense forms of verbs are acquired and produced without the use of rules and variables, and claimed it could account for data from children's language learning. For example, according to this model, instead of learning a rule "to form a past tense, add -ed to the verb stem" and then having to learn that the verb *break* is an exception to this rule, children learn both regular and irregular past tenses in the same way, by associating present and past forms. This model was subjected to a detailed critique by Pinker and Prince (1988), who put forward a two-part model in which regular past-tense forms depend on a rule system, but irregular forms are learned and produced associatively (see also Marcus, 1996).

Young infants have been found to detect and learn strings of nonsense syllables rapidly, apparently by forming associations reflecting how likely one syllable is to follow another in the training set (Saffran, Newport, & Aslin, 1996). For example, if an infant listens to a long string of spoken syllables that have been arranged so that *ta* is statistically more likely to follow *pa* than to follow *ga*, whereas *ga* is statistically more likely to follow *ta*, the infant apparently learns the trisyllable string *pa-ta-ga* as a unit, as if it might be a word.

However, infants have also been found to learn abstract rules for syllable patterns in a way that cannot be accounted for by simple statistical learning and that instead requires learning the values of variables. In one study (Marcus, Vijayan, Rao, & Vishton, 1999), infants were familiarized with a training set of 16 trisyllable strings with an XYX structure (e.g., *ga ti ga*). After each string was repeated three times, the infants were tested on a novel set of 12 trisyllables. Half the test trisyllables conformed to the XYX structure of the training set (e.g., *de ko de*), and half presented a new XYY structure (e.g., *wo fe fe*). Although all the test strings presented novel syllables, the infants attended preferentially to the strings exhibiting the novel structure. What the infants learned, apparently, was a structure, or rule, defined over variables or placeholders (e.g., "item$_i$ is the same as item$_j$"). Such "algebraic" learning is a key aspect of computational models. Marcus et al. (1999) proposed that human infants possess both associative-learning and rule-learning capacities. This debate in the language heartland of classical, computational cognition is continuing. Happily, it is generating a significant amount of light along with some inevitable heat, not to mention a best-seller on an unlikely topic (Pinker, 1999).

There is no point in formulating the associationist and the computational frameworks so abstractly that no basis could ever be found for falsifying either framework. Fortunately, the key competing ideas can be rigorously conceptualized in terms of different classes of physically instantiable machines. Through its long and versatile history, the notion of the association has been applied to links between quite different entities, but the core idea has remained the same. Associations have always been thought of as conductive links of variable strength. Classically, associations were said to link ideas; in behaviorist theories, associations held between memory traces for different stimuli or between stimuli and responses; and currently in connectionist network theory, associative links are formed between processing units or nodes. The strength of an associative link refers to the level of activation conducted between one unit and another or to the probability of activation being so conducted. Following the work of the mathematician Alan Turing, computation refers to any algorithmic process that systematically maps between sets of symbols by storing symbols in and retrieving them from a memory. The electronic computer, of course, is designed to operate according to Turing's principles. Within this framework, to-be-computed functions (programs, processes, etc.) are defined over symbol systems, employing the notion of variables to which specific symbol values can be bound. Both associative and computational systems can be physically realized. However, their defining features have to do with their mode of operation, rather than with how they are physically realized.

The question whether learning involves association or computation is one of the most basic questions we are in a position to ask in psychology. Last century, this question became for the first time both intelligible and empirically tractable. Perhaps the present century will answer it. If so, the study of animal learning will provide an important test bed. The phenomena of conditioned learning in animals are richly documented, and experimental techniques are well developed. What is needed are programs that formulate the critical questions that experimentation will answer. The articles in this Special Section give a glimpse of the work that is in progress. Cognitive scientists working on human cognition should be alert to this work for at least three reasons. First, the answer to the question of whether neural tissue associates stimuli or computes over representations will determine both the overall direction that neuroscience and cognitive science will take and the relationship between them. And it will do this even if the answer is, ultimately, that neither view adequately captures the process of learning. Second, models of animal learning provide the bedrock for the assumption that (at least some) learning is associative in humans. If it turns out that Pavlovian learning in animals is actually accomplished by computational processes, that bedrock disappears. Third, Pavlovian learning occurs even in the lowliest of animal species. This suggests that, whether the character of the underlying mechanism of Pavlovian learning is associative or computational, that character is a fundamental property of neural tissue.

THE SPECIAL SECTION

This Special Section begins with Dickinson's article, in which he contrasts associative and computational models of human causal judgment in light of recent empirical findings from his lab and argues that causal judgment in humans can be understood using associative notions derived from the study of animals. Church points out that associative and computational models of learning can be pitted against one another by comparing the outputs they generate with the outputs of the animal. He then suggests a sophisticated way of doing this with respect to competing behavioral (associative) and scalar (computational) theories of timing. Killeen advances a philosophical analysis of the issue in which neither associations nor computations are located in the organism but are simply devices of the modeler. Miller and Escobar review a range of models—some associative, some computational, some hybrid—pointing out strengths and weaknesses in each case. The important thing for these authors is that the competition between learning models should stimulate new and surprising discoveries.

Finally, Gallistel and Gibbon list a number of major points on which associative and computational models of conditioned learning differ, highlighting one potentially critical empirical property of conditioned learning, namely, its *time-scale invariance*. Time-scale invariance refers to the fact that the rate of conditioned learning in the animal, the so-called learning curve, is not sensitive to the absolute values of the time intervals between the critical events in the learning situation, for example, between the cue stimulus and the reinforcer (i.e., the reinforcement delay). Instead, the learning process is sensitive only to the ratios of time intervals in the situation, for example, the ratio between the reinforcement delay and the intercue interval.

To make this more concrete, imagine you are training Pavlov's dog to salivate when a bell is sounded by puffing meat powder into his mouth 1 s after the bell sounds. You repeat this procedure so that there are 4 s between bell rings. You note how many such pairings of bell and powder are required before the dog has learned to salivate vigorously to the sound of the bell alone. Then you train a second dog in the same way, except you increase the interval between bell and powder to 2 s instead of 1. Now you will find that it takes the second dog longer to learn: It will require more pairings than the first dog to achieve the same slabbering to the bell alone. At first, you may assume that the slower learning is because you have reduced the contiguity between bell and powder by increasing the time between them. But you have changed something else as well—something that is less obvious. You have also changed the ratio between this interval and the interval between bell rings—the interval between bell rings remained at 4 s, so the 4:1 ratio with the first dog became a 2:1 ratio with the second dog. If you train a third dog with the interval between bell rings increased to 8 s and the bell-puff interval at 2 s, thus reinstating the 4:1 ratio, then the third dog will learn as quickly (in as many trials) as the first. Gallistel and Gibbon argue that the property of time-scale invariance in animal learning has wide ramifications. It implies the existence of brain mechanisms that time intervals, count reinforcers, store these values, and calculate ratios between them. Gallistel and Gibbon argue that time-scale invariance is unexpected in associative models, but follows naturally from their computational theory.

The articles in this section span a wide range of views on the issue of whether learning is associative or computational. Settling this issue will be neither quick nor easy. However, it is clear that the study of conditioned learning in animals deserves a more central place in the cognitive sciences than it has enjoyed of late. There are always indefinitely many functions, and therefore models, that can generate the same input-output pairing as does the animal. The only input-output function that is ultimately of scientific interest is the very function that the animal itself actually instantiates. We have learned that the search for this *strong equivalence* (Pylyshyn, 1984) between our models and reality is never quick or easy, but this search is the central business of modern cognitive neuroscience. The good news is that the study of conditioned learning is alive and well, both experimentally and theoretically. Surprising though it may be, Pavlov's old dog may yet have new tricks to teach us.

Recommended Reading

Marcus, G.F. (2001). *The algebraic mind: Integrating connectionism and cognitive science.* Cambridge, MA: MIT Press.
Pinker, S. (1999). (See References)
Skinner, B.F. (1999). Some responses to the stimulus "Pavlov." *Journal of the Experimental Analysis of Behavior*, *72*, 463–465. (Original work published 1972)

Notes

1. Address correspondence to Alan M. Leslie, Department of Psychology and Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Rd., Piscataway, NJ 08854; e-mail: aleslie@ruccs.rutgers.edu.

2. This Special Section grew out of a conference that Charles Flaherty, Kathleen Krauss-VanderGoot, and I organized at Rutgers University in November 1998 to address the issue of whether the best available models of animal learning are associationist or computational. (An account of the conference, with abstracts, is available on the World Wide Web at http://ruccs.rutgers.edu/~aleslie/, click on "Rutgers Symposium on Learning I.") As an outsider to animal learning, I was struck first by how vibrant the ideas in this field continue to be, and second by the willingness of its practitioners to examine fundamental assumptions in light of new evidence. These same properties are evident in the articles that follow.

References

Fodor, J.A. (1975). *The language of thought*. New York: Crowell.

Hebb, D.O. (1949). *The organization of behavior*. New York: John Wiley & Sons.

Marcus, G.F. (1996). Why do children say "breaked"? *Current Directions in Psychological Science*, *5*, 81–85.

Marcus, G.F., Vijayan, S., Rao, S.B., & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, *4*, 135–183.

Pinker, S. (1999). *Words and rules: The ingredients of language.* New York: Basic Books.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *29*, 73–193.

Pylyshyn, Z.W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press/Bradford Books.

Rumelhart, D.E., & McClelland, J.L. (1986). On learning the past tenses of English verbs. In D.E. Rumelhart & J.L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2. Psychological and biological models* (pp. 216–271). Cambridge, MA: MIT Press.

Saffran, J.R., Newport, E.L., & Aslin, R.N. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.