

Editorial: On Machine Learning

The central role of learning

Although researchers in artificial intelligence and psychology have long recognized the importance of learning, this topic has not always been the central focus of these fields. In the first years of AI, considerable attention was given to learning issues, but as pattern recognition and AI developed separate identities, learning research became associated with the former while the latter concentrated on problems of representation and performance. A similar phenomenon occurred in psychology. The behaviorist paradigm was almost exclusively concerned with learning phenomena, but as information processing psychology gained in popularity, psychologists turned their sights towards memory and performance phenomena and all but abandoned efforts to explain the learning process.

However, the past five years have seen a resurgence of interest in learning within both artificial intelligence and cognitive psychology. This has resulted partly from dissatisfaction with pure performance models of intelligence. One of the major insights of both fields has been that, except in the simplest domains, intelligent behavior requires significant knowledge of those domains. Although this insight has led to successful applied AI systems and to accurate psychological models of domain-specific performance, it has not led to systems or theories having any great degree of generality. By refocusing their efforts on learning, many researchers hope to discover more general principles of intelligence. In the case of psychology, such principles would lead to more encompassing theories of human behavior that move beyond particular domains. In the case of applied AI, general learning methods might let one automate the construction of knowledge-intensive systems, saving man-years of effort for each application area.

Yet dissatisfaction with performance models is not sufficient to account for the explosion of research on computational approaches to learning. One must also credit the advances made on representational and performance issues within the two fields over the past two decades. Since any learning system must incorporate representational and performance components, learning researchers have much to glean from results in these areas. There is a growing perception within AI and cognitive

psychology that we now understand enough about representation and performance to look anew at learning issues, and some of the most interesting developments have resulted from attempts to integrate these approaches to understanding intelligence.

The need for a journal

One could attempt to identify other reasons for the current interest in computational approaches to learning, but the fact remains that there is a growing community of researchers in this area. The members of this community read each other's papers, attend each other's talks at conferences, and since 1980 have interacted at biennial summer workshops. The new field has been given the name 'machine learning'. This term has the same benefits and problems as the term 'artificial intelligence', representing much of the field's flavor but misrepresenting other aspects. Nevertheless, both terms are here to stay and we must make the best of them.

The growing sense of community among learning researchers, together with the increasing amount of work in this area, has prompted us to establish the journal *Machine Learning*. We did not take this step lightly, for one does not introduce a new publication without significant justification. Two major forums already existed for publishing formal research results in AI and computational models of human behavior – the journals *Artificial Intelligence* and *Cognitive Science* – and both have included machine learning papers in their time. However, the frequency of published learning papers has been lower than one would expect given the activity of the field, and for researchers interested mainly in learning issues these journals have not been ideal sources. We perceived the need for a forum devoted to workers in machine learning, and we felt that a journal would foster quality research in the area, as well as improve the sense of community in this rapidly developing field.

Some readers will undoubtedly reason that the emergence of specialized journals like *Machine Learning* will serve to fragment the fields of AI and cognitive science, leading to narrowed focus rather than integration. However, the emerging field of machine learning is not narrow in any traditional sense. One can study learning phenomena in such diverse domains as problem solving, natural language understanding, and perception. In addition, one cannot study learning in isolation from work on models of representation and performance, so that integration with traditional areas of AI and cognitive science seems an inherent feature of the field. Machine learning is narrow only in the sense that it focuses attention on systems that improve over time, and we believe that this focus is an important one from which all research into the nature of intelligence would benefit.

Characterizing the field

The field of machine learning is as difficult to define as its parent fields. One might describe it as that field of inquiry concerned with the processes by which intelligent systems improve their performance over time. However, such hard and fast definitions are no more successful at describing scientific disciplines than they are useful in characterizing everyday concepts such as ‘chair’ and ‘food’. The best we can hope to accomplish is to describe the central tendency of the field, a tendency that may itself change as the field develops. For instance, machine learning shares with AI a bias towards symbolic representations rather than numeric ones, though it certainly does not exclude the latter. Similarly, much machine learning research employs heuristic approaches to learning rather than algorithmic ones, but the latter are not forbidden. These dimensions separate artificial intelligence and cognitive science from mainstream computer science and pattern recognition, and machine learning is much more closely associated with the former two areas than with the latter.

Some examples will help clarify the central tendencies of the field. Traditional machine learning problems have included the tasks of learning concepts from examples (Winston, 1975; Hayes-Roth & McDermott, 1978; Michalski, 1980), learning heuristics for directing search (Anzai, 1978; Mitchell, Utgoff, & Banerji, 1983; Langley, 1985), and learning grammars from sample sentence-meaning pairs (Siklóssy, 1972; Anderson, 1977; Langley, 1982). Traditional representations used in this work have included propositions, discrimination networks, production rules, and augmented transition networks. The notion of heuristic search through a space of rules or hypotheses (Mitchell, 1982) has played a central role in much of the machine learning research.

Although tasks like scientific discovery involve no learning in the narrow sense of the term, research in this area (Lenat, 1977; Langley, 1981) is usually included within the fold of machine learning, since discovery problems involve many of the same issues as more obvious learning tasks. A similar inclusion holds for methods like reasoning by analogy (Winston, 1979; Carbonell, 1983), which often involve no learning but which employ mechanisms very similar to those used in learning systems.

The papers in our first issue represent a cross-section of research in the field. Early work in machine learning focused on tasks like learning from examples and language acquisition, employing methods that relied on significant amounts of data and that often involved considerable search. This data-intensive paradigm is alive and well, as shown by Quinlan’s survey of research on the induction of decision trees. However, this approach is now complemented by methods that derive maximum information from single instances, using analytic rather than empirical techniques. Mitchell, Keller, and Kedar-Cabelli’s paper on explanation-based learning summarizes some of the latest work in this area.

Despite the name ‘machine learning’, a significant fraction of the field has always been concerned with modeling human behavior, starting with Feigenbaum’s (1963)

EPAM model of verbal learning. In the current issue, this tradition is represented by Zytlow and Simon's paper on chemical discovery, in which the authors compare the behavior of their STAHL system with actual discoveries made by early chemists. In contrast, the paper by Laird, Rosenbloom, and Newell represents a distinct break with tradition in its attempt to specify a complete *cognitive architecture*, all the components of which are open to learning. In another sense, this work combines two venerable traditions – describing performance in terms of search through a problem space, and modeling learning in terms of adaptive production systems.

In addition to the themes represented in this issue, the field is moving beyond the traditional tasks and methods on a number of fronts. One innovation involves the notion of *integrated* models of learning, focusing on the interaction between relatively well-understood components such as concept learning, strategy acquisition, and language learning. Another emerging theme is the study of learning in complex, reactive environments that cannot be completely predicted, and which may require entirely new learning methods. There is also considerable interest in applied machine learning research, focusing on the automatic construction of knowledge-based systems. For instance, Michalski and Chilausky (1980) have worked on the induction of a knowledge base from examples, and Mitchell, Mahadevan, and Steinberg (1985) have outlined another approach based on learning apprentices. Only time will tell which of these paradigms will prove viable, but we mention them to provide further insight into the state of the field and the direction in which it seems to be moving.

Suggestions for authors

The four papers in this issue should give authors some idea of the type of manuscripts we would encourage them to submit, but a few additional words seem in order. Since machine learning is concerned with *computational* approaches to learning, we expect that most papers will describe a running system that shows evidence of improvement over time. Although papers need not describe actual systems, the computational metaphor is a central assumption of the field and ideas should be cast in these terms.

However, describing 'yet another program' is not by itself sufficient; the system should reveal *principles* of learning that go beyond the particular implementation. We feel that such principles should be a major goal of the field, and we welcome attempts to synthesize the results of previous research in these terms. Associated with principles is the notion of *generality*, and the ideal paper will discuss how an approach can be applied to different domains, different representations, and similar variations.

Although machine learning (like AI) is largely an empirical science, we encourage theoretical treatments, especially those which address tasks and methods that have been studied empirically. We welcome papers on computational models of human learning, though we do not expect the majority to fall into this category. We also en-

courage thoughtful replications of previous systems, provided they reveal new results not apparent in the original work. Finally, while the journal will emphasize basic research in machine learning, we will also consider applications of learning methods to real-world domains.

Aside from content, submitted papers should be clearly written. For instance, concrete examples often help to clarify abstract concepts, and systems can be usefully described in terms of their inputs and outputs, as well as the representations and processes they employ. Since the notion of heuristic search will be familiar to most readers of the journal, describing a learning system's behavior in these terms will usually lead to increased understanding. Science requires the ability to replicate results, and authors should describe their systems in enough detail to enable readers to reconstruct those systems and test them. However, one can achieve this level of detail without program listings. Judicious English paraphrases of rules and heuristics should be sufficient, combined with occasional commented traces of the systems' behaviors.

Since different machine learning researchers have different goals – such as understanding human behavior or automatically constructing expert systems – we encourage authors to clearly state their goals at the outset of their papers. If their approach relies on particular assumptions (e.g., the absence of noise), they should make this clear as well. We also encourage the use of established terms, and clarification when multiple meanings exist for a term. For instance, the term 'generalization' has two distinct meanings within machine learning – the process of moving from a specific hypothesis to a more general one, and any process for constructing a general rule from data. The first of these senses is the opposite of 'discrimination' learning, while the second encompasses both the first sense and discrimination. The paper by Mitchell, Keller, and Kedar-Cabelli in this issue uses the second meaning of 'generalization'. While we will not attempt to legislate which terms the authors employ, we will insist on clear use of those terms.

Although the success of a new journal depends largely on its contributors, the editors and editorial board also have a significant role to play. We will do our best to give rapid reviews, and we will endeavor to provide useful feedback whether or not we decide to accept a paper for publication. The editors will revolve after three years to ensure the infusion of new blood and new ideas (and to give us more time for research), and we expect the editorial board to evolve over the years as new researchers and research themes emerge.

This is an exciting period for the field of machine learning, with new problems constantly being formulated and new answers being generated in response. We hope this journal can convey some of the excitement to its audience, bringing new researchers into the discipline and spurring veterans on to greater heights. We invite the reader to join us in our exploration of one of the central wonders of intelligence, as we attempt to learn about the nature of learning.

Pat Langley
University of California, Irvine

References

- Anderson, J.R. (1977). Induction of augmented transition networks. *Cognitive Science*, 2, 125–157.
- Anzai Y. (1978). Learning strategies by computer. *Proceedings of the Second National Conference of the Canadian Society for Computational Studies of Intelligence* (pp. 181–190). Toronto, Ontario, Canada.
- Carbonell, J.G. (1983). Learning by analogy: Formulating and generalizing plans from past experience. In R.S. Michalski, J.G. Carbonell, & T.M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach*. Los Altos, CA: Morgan Kaufmann.
- Feigenbaum, E.A. (1963). The simulation of verbal learning behavior. In E.A. Feigenbaum & J. Feldman (Eds.), *Computers and thought*. New York: McGraw-Hill, Inc.
- Hayes-Roth, F., & McDermott, J. (1978). An interference matching technique for inducing abstractions. *Communications of the ACM*, 21, 401–410.
- Langley, P. (1981). Data-driven discovery of physical laws. *Cognitive Science*, 5, 31–54.
- Langley, P. (1982). Language acquisition through error recovery. *Cognition and Brain Theory*, 5, 211–255.
- Langley, P. (1985). Learning to search: From weak methods to domain-specific heuristics. *Cognitive Science*, 9, 217–260.
- Lenat, D.B. (1977). Automated theory formation in mathematics. *Proceedings of the Fifth International Joint Conference on Artificial Intelligence* (pp. 833–842). Cambridge, Mass.: Morgan Kaufmann.
- Michalski, R.S., & Chilausky, R.L. (1980). Learning by being told and learning from examples: An experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis. *International Journal of Policy Analysis and Information Systems*, 4.
- Michalski, R.S. (1980). Pattern recognition as rule-guided inductive inference. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2, 349–361.
- Mitchell, T.M. (1982). Generalization as search. *Artificial intelligence*, 18, 203–226.
- Mitchell, T.M., Utgoff, P., & Banerji, R.B. (1983). Learning problem solving heuristics by experimentation. In R.S. Michalski, J.G. Carbonell, & T.M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach*. Los Altos, CA: Morgan Kaufmann.
- Mitchell, T.M., Mahadevan, S., & Steinberg, L. (1985). LEAP: A learning apprentice for VLSI design. *Proceedings of the Ninth International Joint Conference on Artificial Intelligence* (pp. 573–580). Los Angeles, CA: Morgan Kaufmann.
- Siklóssy, L. (1972). Natural language learning by computer. In H.A. Simon & L. Siklóssy (Eds.), *Representation and meaning: Experiments with information processing systems*. Englewood Cliffs, NJ: Prentice-Hall.
- Winston, P.H. (1975). Learning structural descriptions from examples. In P.H. Winston (Ed.), *The psychology of computer vision*. New York: McGraw-Hill.
- Winston, P.H. (1979). Learning and reasoning by analogy. *Communications of the ACM*, 12, 689–703.