Editors' Note: Machine Learning has been a constant, theme throughout AI's two decades of existence. In this overview the authors analyze various aspects including the major methodological approaches advocated in Machine Learning research, and how they have related to major contemporary themes in “mainstream” AI Research. In a subsequent issue we plan to include a sequel to this article which will give the authors' views on current research directions in Machine Learning.

In the meanwhile, we are very anxious to get readers' reactions to this and all earlier contributions to this column — Derek Sleeman and Jaime Carbonell

Abstract

Machine learning has always been an integral part of artificial intelligence, and its methodology has evolved in concert with the major concerns of the field. In response to the difficulties of encoding ever-increasing volumes of knowledge in modern AI systems, many researchers have recently turned their attention to machine learning as a means to overcome the knowledge acquisition bottleneck. This article presents a taxonomic analysis of machine learning organized primarily by learning strategies and secondarily by knowledge representation and application areas. A historical survey outlining the development of various approaches to machine learning is presented from early neural networks to present knowledge-intensive techniques.

This paper is a modified and extended version of the first chapter of Machine Learning: An Artificial Intelligence Approach, with permission of the publisher, Tioga Press (Palo Alto, CA). The research described here was sponsored in part by the Office of Naval Research (ONR) under grant number N00014-79-C-0661, and in part by the National Science Foundation grant MCS82-05166

Learning is a many-faceted phenomenon. Learning processes include the acquisition of new declarative knowledge, the development of motor and cognitive skills through instruction or practice, the organization of new knowledge into general, effective representations, and the discovery of new facts and theories through observation and experimentation. The study and computer modelling of learning processes in their multiple manifestations constitute the subject matter of machine learning.

Although machine learning has been a central concern in artificial intelligence since the early days when the idea of “self-organizing systems” was popular, the limitations inherent in the early neural network approaches led to a temporary decline in research volume. More recently, new symbolic methods and knowledge-intensive techniques have yielded promising results and these in turn have led to the current revival in machine learning research. This article examines some basic methodological issues, proposes a classification of machine learning techniques, and provides a historical review of the major research directions.

The Objectives of Machine Learning

The field of machine learning can be organized around three primary research foci:

- Task-Oriented Studies - the development and analysis of learning systems oriented toward solving a
predictor set of tasks (also known as the “engineer-
ing approach”)

- Cognitive Simulation—the investigation and com-
puter simulation of human learning processes (also
known as the “cognitive modelling approach”)

- Theoretical Analysis—the theoretical exploration
of the space of possible learning methods and algo-
rithms independent of application domain.

Although many research efforts strive primarily towards
one of these objectives, progress in one objective often leads
to progress in another. For instance, in order to investigate
the space of possible learning methods, a reasonable start-
ing point may be to consider the only known example of
robust learning behavior, namely humans (and perhaps other
biological systems). Similarly, psychological investigations
of human learning may be helped by theoretical analysis that
may suggest various plausible learning models. The need to
acquire a particular form of knowledge in some task-oriented
study may itself spawn new theoretical analysis or pose the
question: “How do humans acquire this specific skill (or
knowledge)?” The existence of these mutually supportive ob-
jectives reflects the entire field of artificial intelligence, where
expert systems research, cognitive simulation, and theoretical
studies provide some cross-fertilization of problems and
ideas.

**Applied Learning Systems: A Practical Necessity.**

At present, instructing a computer or a computer-controlled
robot to perform a task requires one to define a complete and
correct algorithm for that task, and then laboriously pro-
gram the algorithm into a computer. These activities typi-
cally involve a tedious and time-consuming effort by specially
trained personnel.

Present-day computer systems cannot truly learn to per-
form a task through examples or by analogy to a similar,
previously-solved task. Nor can they improve significantly
on the basis of past mistakes, or acquire new abilities by
observing and imitating experts. Machine learning research
strives to open the possibility of instructing computers in
such new ways, and thereby promises to ease the burden of
hand-programming growing volumes of increasingly complex
information into the computers of tomorrow. The rapid ex-
ansion of applications and availability of computers today
makes this possibility even more attractive and desirable.

When approaching a task-oriented knowledge acquisi-
tion task, one must be aware that the resultant computer sys-
tems must interact with humans, and therefore should closely
parallel human abilities. The traditional argument that an
engineering approach need not reflect human or biological
performance is not truly applicable to machine learning.
Since airplanes, a successful result of an almost pure en-
geineering approach, bear little resemblance to their biological
counterparts, one may argue that applied knowledge acquisi-
tion systems could be equally divorced from any considera-
tion of human capabilities. This argument does not apply
here because airplanes need not interact with or understand
birds. Learning machines, on the other hand, will have to
interact with the people who make use of them, and conse-
quently the concepts and skills they acquire— if not neces-
sarily their internal mechanisms—must be understandable
to humans.

**Machine Learning as a Science.** The question of
what are the genetically-endowed abilities in a biological sys-
tem (versus environmentally-acquired skills or knowledge)
has fascinated biologists, psychologists, philosophers and
artificial intelligence researchers alike. A clear candidate for
a cognitive invariant in humans is the learning mechanism
the innate ability to acquire facts, skills and more abstract
concepts. Therefore, understanding human learning well
enough to reproduce aspects of that learning behavior in
a computer system is, in itself, a worthy scientific goal.
Moreover, the computer can render substantial assistance
to cognitive psychology, in that it may be used to test the
consistency and completeness of learning theories, and en-
force a commitment to fine-structure process-level detail that
precludes meaningless, tautological or untestable theories

The study of human learning processes is also of con-
siderable practical significance. Gaining insights into the
principles underlying human learning abilities is likely to
lead to more effective educational techniques. Thus, it is not
surprising that research into intelligent computer-assisted in-
struction, which attempts to develop computer-based tutor-
ning systems, shares many of the goals and perspectives with
machine learning research. One particularly interesting de-
velopment is that computer tutoring systems are starting to
incorporate abilities to infer models of student competence
from observed performance. Inferring the scope of a stu-
dent’s knowledge and skills in a particular area allows much
more effective and individualized tutoring of the student
(Sleeman 1983).

An equally basic scientific objective of machine learn-
ing is the exploration of possible learning mechanisms, in-
cluding the discovery of different induction algorithms, the
scope and theoretical limitations of certain methods, the in-
formation that must be available to the learner, the issue of
coping with imperfect training data, and the creation of
general techniques applicable in many task domains. There
is no reason to believe that human learning methods are the
only possible means of acquiring knowledge and skills. In
fact, common sense suggests that human learning represents
just one point in an uncharted space of possible learning
methods—a point that through the evolutionary process is
particularly well suited to cope with the general physical
environment in which we exist. Most theoretical work in
machine learning has centered on the creation, characteriza-
tion and analysis of general learning methods, with the major
emphasis on analyzing generality and performance rather
than psychological plausibility.

Whereas theoretical analysis provides a means of explor-
ing the space of possible learning methods, the task-oriented
approach provides a vehicle to test and improve the per-
formance of functional learning systems. By constructing
and testing applied learning systems, one can determine the cost-effectiveness trade-offs and limitations of particular approaches to learning. In this way, individual data points in the space of possible learning systems are explored, and the space itself becomes better understood.

Knowledge Acquisition versus Skill Refinement. There are two basic forms of learning: knowledge acquisition and skill refinement. When we say that someone learned physics, we mean that this person acquired concepts of physics, understood their meaning, and their relationship to each other as well as to the physical world. The essence of learning in this case is the acquisition of knowledge, including descriptions and models of physical systems and their behaviors, incorporating a variety of representations—from simple intuitive mental models, examples and images, to completely tested mathematical equations and physical laws. A person is said to have learned more if his knowledge explains a broader scope of situations, is more accurate, and is better able to predict the behavior of the physical world (Popper 1968). This form of learning is typical to a large variety of situations and is generally termed knowledge acquisition. Hence, knowledge acquisition is defined as learning new symbolic information coupled with the ability to apply that information in an effective manner.

A second kind of learning is the gradual improvement of motor and cognitive skills through practice, such as learning to ride a bicycle or to play the piano. Acquiring textbook knowledge on how to perform these activities represents only the initial, and not necessarily critical, phase in developing the requisite skills. The bulk of the learning process consists of refining the acquired skill, and improving the mental or motor coordination by repeated practice and a correction of deviations from desired behavior. This form of learning, often called skill refinement, differs in many ways from knowledge acquisition. Whereas the essence of knowledge acquisition may be a conscious process whose result is the creation of new symbolic knowledge structures and mental models, skill refinement occurs by virtue of repeated practice without concerted conscious effort. Most human learning appears to be a mixture of both activities, with intellectual endeavors favoring the former, and motor coordination tasks favoring the latter.

Present machine learning research focuses on the knowledge acquisition aspect, although some investigations, specifically those concerned with learning in problem-solving and transforming declarative instructions into effective actions, touch on aspects of both types of learning. Whereas knowledge acquisition clearly belongs in the realm of artificial intelligence research, a case could be made that skill refinement comes closer to non-symbolic processes, such as those studied in adaptive control systems. It may indeed be the case that skill acquisition is inherently non-symbolic in biological systems, but an interesting symbolic model capable of simulating gradual skill improvement through practice has been proposed by Newell and Rosenbloom (1981). Hence, perhaps both forms of learning can be captured in artificial intelligence models.

A Taxonomy of Machine Learning Research

This section presents a taxonomic road map to the field of machine learning with a view towards presenting useful criteria for classifying and comparing most artificial intelligence-based machine learning investigations. Later, the main directions actually taken by researchers in this area over the past twenty years are surveyed.

One may classify machine learning systems along many different dimensions. We have chosen three dimensions as particularly meaningful:

- Classification on the basis of the underlying learning strategy used. The strategies are ordered by the amount of inference the learning system performs on the information provided to the system.
- Classification on the basis of the type of representation of knowledge (or skill) acquired by the learner.
- Classification in terms of the application domain of the performance system for which knowledge is acquired.

Each point in the space defined by the above dimensions corresponds to a system employing a particular learning strategy, a particular knowledge representation, and applied to a particular domain. Since many existing learning systems employ multiple strategies and knowledge representations, and some have been applied to more than one domain, such learning systems are characterized by a collection of points in the space.

The subsections below describe explored values along each of these dimensions. Future research may well reveal new values on these dimensions as well as new dimensions. Indeed, the larger space of all possible learning systems is still only sparsely explored and partially understood. Existing learning systems correspond to only a small portion of the space because they represent only a small number of possible combinations of the values.

Classification Based on the Underlying Learning Strategy. Since we distinguish learning strategies by the amount of inference the learner performs on the information provided, we first consider the two extremes: performing no inference, and performing a substantial amount of inference. If a computer system is programmed directly, its knowledge increases, but it performs no inference whatsoever on the new information; all cognitive effort is on the part of the programmer. Conversely, if a system independently discovers new theories or invents new concepts, it must perform a very substantial amount of inference; it is deriving organized knowledge from experiments and observations. An intermediate point in the spectrum would be a student determining how to solve a mathematics problem by analogy to worked-out examples in the textbook—a process that requires inference, but much less than discovering a new branch of mathematics without guidance from teacher or textbook.

THE AI MAGAZINE Fall 1983 71
As the amount of inference that the learner is capable of performing increases, the burden placed on the teacher or external environment decreases. It is much more difficult to teach a person by explaining each step in a complex task than by showing that person the way that similar tasks are usually handled. It is more difficult yet to program a computer to perform a complex task than to instruct a person to perform the task; as programming requires explicit specification of all requisite detail, whereas a person receiving instruction can use prior knowledge and common sense to fill in most mundane details. The taxonomy below captures this notion of trade-offs in the amount of effort required of the learner and of the teacher.

**Rote Learning and Direct Implanting of New Knowledge.** In rote learning no inference or other transformation of the knowledge is performed by the learner. Variants of this strategy of knowledge acquisition method include:

- **Learning by being programmed, constructed, or modified by an external entity,** (for example, the usual style of computer programming)
- **Learning by memorization of given facts and data with no inferences drawn from the incoming information** (for example, as performed by existing database systems). The term “rote learning” is used primarily in this context.

**Learning from Instruction.** Acquiring knowledge from a teacher or other organized source, such as a textbook, requires that the learner transform the knowledge from the input language to an internally-usable representation, and that the new information be integrated with prior knowledge for effective use. Hence, the learner is required to perform some inference, but a large fraction of the burden remains with the teacher, who must present and organize knowledge in a way that incrementally augments the student's existing knowledge. Learning from instruction, also termed “learning by being told,” parallels most formal education methods. Therefore, the machine learning task is one of building a system that can accept instruction or advice and can store and apply this learned knowledge effectively.

**Learning by Analogy.** Learning by analogy is the process of transforming and augmenting existing knowledge (or skills) applicable in one domain to perform a similar task in a related domain. For instance, a person who has never driven a small truck, but drives automobiles, may well transform his existing skill (perhaps imperfectly) to the new task. Similarly, a learning-by-analogy system might be applied to convert an existing computer program into one that performs a closely-related function for which it was not originally designed. Learning by analogy requires more inference on the part of the learner than does rote learning or learning from instruction. A fact or skill analogous in relevant parameters must be retrieved from memory; then the retrieved knowledge must be appropriately transformed, applied to the new situation, and stored for future use.

**Learning from examples.** Learning from examples is a special case of inductive learning. Given a set of examples and counterexamples of a concept, the learner inducts a general concept description that describes all of the positive examples and none of the counterexamples. Learning from examples is a method that has been heavily investigated in artificial intelligence. The amount of inference performed by the learner is much greater than in learning from instruction, as no general concepts are provided by a teacher, and is somewhat greater than in learning by analogy, as no similar concepts are provided as “seed” from which the new concept may be grown. Learning from examples can be subcategorized according to the source of the examples:

- **The source is a teacher who knows the concept and generates examples of the concept that are meant to be as helpful as possible.** If the teacher also knows (or, more typically, infers) the knowledge state of the learner, the examples can be generated to optimize convergence on the desired concept (as in Winston’s (1975) near-miss analysis).
- **The source is the learner itself.** The learner typically knows its own knowledge state, but clearly does not know the concept to be acquired. Therefore, the learner can generate instances (and have an external entity such as the environment or a teacher classify them as positive or negative examples) on the basis of the information it believes necessary to discriminate among contending concept descriptions.

For instance, a learner trying to acquire the concept of “ferromagnetic substance,” may generate as a possible candidate “all metals.” Upon testing copper and other metals with a magnet, the learner will then discover that copper is a counterexample, and therefore the concept of ferromagnetic substance should not be generalized to include all metals. Mitchell’s LEX system (1983) and Carbonell’s plan generalization method (1983) illustrate the process of internal instance generation.

The source is the external environment. In this case the example generation process is operationally random, as the learner must rely on relatively uncontrolled observations. For example, an astronomer attempting to infer precursors to supernovas must rely mainly upon unstructured data presentation. Although the astronomer knows the concept of a supernova, he cannot know a priori where and when a supernova will occur, nor can he cause one to exist. Michalski’s STAR methodology (1983) exemplifies this type of learning.

One can also classify learning from examples by the type of examples available to the learner:

- **Only positive examples available.** Whereas positive examples provide instances of the concept to be acquired, they do not provide information for preventing overgeneralization of the inferred concept. In this kind of learning situation, overgeneralization might be avoided by considering only the minimal generalizations necessary, or by relying on a priori domain knowledge to constrain the concept to be inferred.
• **Positive and negative examples available** In this kind of situation, positive examples force generalization whereas negative examples prevent overgeneralization (the induced concept should never be so general as to include any of the negative examples). This is the most typical form of learning from examples.

Learning from examples may be one-trial or incremental. In the former case, all examples are presented at once. In the latter case, the system must form one or more hypotheses of the concept (or range of concepts) consistent with the available data, and subsequently refine the hypotheses after considering additional examples. The incremental approach more closely parallels human learning, allows the learner to use partially learned concepts (for performance, or to guide the example generation process), and enables a teacher to focus on the basic aspects of a new concept before attempting to impart less central details. On the other hand, the one-step approach is less apt to lead one down garden paths by an injudicious choice of initial examples in formulating the kernel of the new concept.

**Learning from Observation and Discovery.** This "unsupervised learning" approach is a very general form of inductive learning that includes discovery systems, theory formation tasks, the creation of classification criteria to form taxonomic hierarchies, and similar tasks to be performed without benefit of an external teacher. Unsupervised learning requires the learner to perform more inference than any approach thus far discussed. The learner is not provided with a set of instances of a particular concept, nor is it given access to an oracle that can classify internally-generated instances as positive or negative examples of any given concept. Moreover, rather than focusing on a single concept at a time, the observations may span several concepts that need to be acquired, thus introducing a severe focus-of-attention problem. One may subclassify learning from observation according to the degree of interaction with an external environment. The extreme points in this dimension are:

- **Passive observation**, where the learner classifies and taxonomizes observations of multiple aspects of the environment (as in Michalski and Stepp's (1983) conceptual clustering.)
- **Active experimentation**, where the learner perturbs the environment to observe the results of its perturbations. Experimentation may be random, dynamically focused according to general criteria of interestingness, or strongly guided by theoretical constraints. As a system acquires knowledge, and hypothesizes theories it may be driven to confirm or disconfirm its theories, and hence explore its environment applying different observation and experimentation strategies as the need arises. Often this form of learning involves the generation of examples to test hypothesized or partially acquired concepts. This type of learning is exemplified in Lenat's AM and EURISKO systems (Lenat 1976, 1983; booklena82)

An Intermediate point in this dimension is the BACON system (Langley, Simon & Bradshaw, 1983), attention but does not design new experiments.

The above classification of learning strategies should help one to compare various learning systems in terms of their underlying mechanisms, in terms of the available external source of information, and in terms of the degree to which they rely on pre-organized knowledge.

**Classification According to Type of Knowledge Acquired.** A learning system may acquire rules of behavior, descriptions of physical objects, problem-solving heuristics, classification taxonomies over a sample space, and many other types of knowledge useful in the performance of a wide variety of tasks. The list below spans types of knowledge acquired, primarily as a function of the representation of that knowledge.

1. **Parameters in algebraic expressions**—Learning in this context consists of adjusting numerical parameters or coefficients in algebraic expressions of a fixed functional form so as to obtain desired performance. For instance, perceptrons adjust weighting coefficients for threshold logic elements when learning to recognize two-dimensional patterns (Rosenblatt 1958, Minsky & Papert 1969).

2. **Decision trees**—Some systems acquire decision trees to discriminate among classes of objects. The nodes in a decision tree correspond to selected object attributes, and the edges correspond to predetermined alternative values for these attributes. Leaves of the tree correspond to sets of objects with an identical classification. Feigenbaum's EPAM exemplifies this discrimination-based learning approach (Feigenbaum, 1963).

3. **Formal grammars**—In learning to recognize a particular (usually artificial) language, formal grammars are induced from sequences of expressions in the language. These grammars are typically represented as regular expressions, finite-state automata, context-free grammar rules, or transformation rules.

4. **Production rules**—A production rule is a condition-action pair C → A, where C is a set of conditions and A is a sequence of actions. If all the conditions in a production rule are satisfied, then the sequence of actions is executed. Due to their simplicity and ease of interpretation, production rules are a widely-used knowledge representation in learning systems. The four basic operations whereby production rules may be acquired and refined are:
   1. **Creation**: A new rule is constructed by the system or acquired from an external entity.
   2. **Generalization**: Conditions are dropped or made less restrictive, so that the rule applies in a larger number of situations.
   3. **Specialization**: Additional conditions are added to the condition set, or existing conditions made more restrictive, so that the rule applies to a smaller number of specific situations.
   4. **Composition**: Two or more rules that were applied in sequence are composed into a single larger rule.
Formal logic-based expressions and related formalisms—These general-purpose representations have been used to formulate descriptions of individual objects (that are input to a learning system) and to formulate resultant concept descriptions (that are output from a learning system). They take the form of formal logic expressions whose components are propositions, arbitrary predicates, finite-valued variables, statements restricting ranges of variables (such as “a number between 1 and 9”), or embedded logical expressions.

Graphs and Networks—In many domains graphs and networks provide a more convenient and efficient representation than logical expressions, although the expressive power of network representations is comparable to that of formal logic expressions. Some learning techniques exploit graph-matching and graph-transformation schemes to compare and index knowledge efficiently.

Frames and schemas—These provide larger organizational units than single logical expressions or production rules. Frames and schemas can be viewed as collections of labeled entities (“slots”), each slot playing a certain prescribed role in the representation. They have proven quite useful in many artificial intelligence applications. For instance, a system that acquires generalized plans must be able to represent and manipulate such plans as units, although their internal structure may be arbitrarily complex. Moreover, in experiential learning, past successes, untested alternatives, causes of failure, and other information must be recorded and compared in inducing and refining various rules of behavior (or entire plans). Schema representations provide an appropriate formalism.

Computer programs and other procedural encodings—The objective of several learning systems is to acquire an ability to carry out a specific process efficiently, rather than to reason about the internal structure of the process. Most automatic programming systems fall in this general category. In addition to computer programs, procedural encodings include human motor skills (such as knowing how to ride a bicycle), instruction sequences to robot manipulators, and other “compiled” human or machine skills. Unlike logical descriptions, networks or frames, the detailed internal structure of the resultant procedural encodings need not be comprehensible to humans, or to automated reasoning systems. Only the external behavior of acquired procedural skills become directly available to the reasoning system.

Taxonomies—Learning from observation may result in global structuring of domain objects into a hierarchy or taxonomy. Clustering object descriptions into newly-proposed categories and forming hierarchical classifications require that the system formulate relevant criteria for classification.

Multiple representations—Some knowledge acquisition systems use several representation schemes for the newly acquired knowledge. Most notably, some discovery and theory-formation systems acquire concepts, operations on those concepts, and heuristic rules for new domains. These learning systems must select appropriate combinations of representation schemes applicable to the different forms of knowledge acquired.

Classification by Domain of Application Another useful dimension for classifying learning systems is their area of application. The list below specifies application areas to which various existing learning systems have been applied.

1. Agriculture
2. Chemistry
3. Cognitive Modeling (simulating human learning processes)
4. Computer Programming
5. Education
7. Game Playing (chess, checkers, poker, and so on)
8. General Methods (no specific domain)
9. Image Recognition
10. Mathematics
11. Medical Diagnosis
12. Music
13. Natural Language Processing
14. Physical Object Characterizations
15. Physics
16. Planning and Problem-solving
17. Robotics
18. Sequence Extrapolation
19. Speech Recognition

Now that we have a basis for classifying and comparing learning systems, we turn to a brief historical outline of machine learning.

A Historical Sketch of Machine Learning

Over the years, research in machine learning has been pursued with varying degrees of intensity, using different approaches and placing emphasis on different aspects and goals. Within the relatively short history of this discipline, one may distinguish three major periods, each centered around a different paradigm:

- neural modeling and decision-theoretic techniques
- symbolic concept-oriented learning
- knowledge-intensive approaches combining various learning strategies
The Neural Modelling Paradigm. The distinguishing feature of the first paradigm was the interest in building general purpose learning systems that start with little or no initial structure or task-oriented knowledge. The major thrust of research based on this tabula rasa approach involved constructing a variety of neural model-based machines, with random or partially random initial structure. These systems were generally referred to as neural nets or self-organizing systems. Learning in such systems consisted of incremental changes in the probabilities that neuron-like elements (typically threshold logic units) would transmit a signal.

Due to the primitive nature of computer technology at that time, most of the research under this paradigm was either theoretical or involved the construction of special purpose experimental hardware systems, such as perceptrons (Rosenblatt 1958), pandemonium (Selfridge 1959) and adaeine (widrow 1962). The groundwork for this paradigm was laid in the forties by Rashevsky and his followers working in the area of mathematical biophysics (Rashevsky 1948), and by McCulloch and Pitts (1943), who discovered the applicability of symbolic logic to modeling nervous system activities. Among the large number of research efforts in this area, one may mention many works such as Ashby (1960), Rosenblatt (1958, 1962), Minsky & Papert (1969), Block (1961), Yovits (1962), Widrow (1962), Culberson (1963), Kazmiereczak (1963). Related research involved the simulation of evolutionary processes, that through random mutation and "natural" selection might create a system capable of some intelligent behavior (for example, Fiedberg 1958, 1959; Holland 1980).

Experience in the above areas spawned the new discipline of pattern recognition and led to the development of a decision-theoretic approach to machine learning. In this approach, learning is equated with the acquisition of linear, polynomial, or related discriminant functions from a given set of training examples. Example include Nilsson (1965), Koford (1966), Uhr (1966), and Highleyman (1967). One of the best known successful learning systems utilizing such techniques (as well as some original new ideas involving non-linear transformations) was Samuel’s checkers program (Samuel, 1959, 1963). Through repeated training, this program acquired master-level performance. Somewhat different, but closely related, techniques utilized methods of statistical decision theory for learning pattern recognition rules (for example, Sebestyen 1962, Fu 1968, Watanabe 1960, Arkadie 1971, Fu 1972, Duda & Hart 1973, Kanal 1974).

In parallel to research on neural modeling and decision-theoretic techniques, researchers in control theory developed adaptive control systems able to adjust automatically their parameters in order to maintain stable performance in the presence of various disturbances, for example, Truxal (1955); Davies (1970); Mendel (1970); Tseykin (1968, 1971, 1973); and Fu (1971, 1974).

Practical results sought by the neural modeling and decision-theoretic approaches met with limited success. High expectations articulated in various early works were not realized, and research under this paradigm began to decline. Theoretical studies have revealed strong limitations of the “knowledge-free” perceptron-type learning systems.

The Symbolic Concept-Acquisition Paradigm. A second major paradigm started to emerge in the early sixties stemming from the work of psychologists and early AI researchers on models of human learning (Hunt et al., 1963, 1966). The paradigm utilized logic or graph structure representations rather than numerical or statistical methods. Systems learned symbolic descriptors representing higher level knowledge and made strong structural assumptions about the concepts to be acquired. Examples of work in this paradigm include research on human concept acquisition (Hunt & Hovland 1963, Feigenbaum 1963, Hunt et al. 1966, Hilgard 1966, Simon & Lea 1974) and various applied pattern recognition systems (Bongard 1970, Uhr 1966, Karpinski & Michalski 1966).

Some researchers constructed task-oriented specialized systems that would acquire knowledge in the context of a practical problem. For instance, the META-DENDRAL program (Buchanan & Feigenbaum 1978) generates rules explaining mass spectrometry data for use in the DENDRAL system (Buchanan et al. 1971).

Winston’s (1975) structurally learning system was an influential development in this paradigm. In parallel with Winston’s work, different approaches to learning structural concepts from examples emerged, including a family of logic-based inductive learning programs, AQVAL (Michalski 1972, 1973, 1978), and related work (Hayes-Roth 1974, Hayes-Roth & McDermott 1978, Vere 1973, Mitchell 1978). See Dietterich and Michalski (1983) and Michie (1982) for additional discussion of this paradigm.

The Modern Knowledge-Intensive Paradigm. The third paradigm represents the most recent period of research starting in the mid-seventies. Researchers have broadened their interest beyond learning isolated concepts from examples, and have begun investigating a wide spectrum of learning methods, most based upon knowledge rich systems. Specifically, this paradigm can be characterized by several new trends, including:

1. Knowledge-Intensive Approaches: Researchers are strongly emphasizing the use of task-oriented knowledge and the constraints it provides in guiding the learning process. One lesson from the failures of earlier tabula rasa and knowledge-poor learning systems is that to acquire new knowledge a system must already possess a great deal of initial knowledge.

2. Exploration of alternative methods of learning: In addition to the earlier research emphasis on learning from examples, researchers are now investigating a wider variety of learning methods such as learning from instruction (e.g., Mostow 1983, Haas & Hendrix 1983, Rychener 1983), learning by analogy (e.g., Winston 1979, Carbonell 1983, Anderson 1983), and discovery of concepts and classifications (e.g.,
3 Incorporating abilities to generate and select learning tasks: In contrast to previous efforts, a number of current systems incorporate heuristics to control their focus of attention by generating learning tasks, proposing experiments to gather training data, and choosing concepts to acquire (e.g., Lenat 1976, Michalski & Stepp 1983, Hayes-Roth 1983, Quinlan 1983).

In contrast with the knowledge-free parametric learning methods used in the neural networks, and in contrast with the early symbolic methods that learned isolated, “disembodied” concepts, the current approaches use a wealth of general and domain-specific knowledge. However, the availability of large volumes of knowledge does not mean that the inductive inference processes are themselves domain dependent and non-generalizable. The generality lies in the inductive inference methods and the power is derived from their ability to use domain knowledge to focus attention and structure new concepts. The current methodological assumption is that machine learning systems, much like humans, must learn incrementally, slowly expanding a highly-organized knowledge base, rather than by some gestalt self-organization process. A recently published book on machine learning (Michalski, Carbonell & Mitchell, 1983) presents some of the major research directions in this general approach.

References


---

**Not just the findings but the complete memos**

The development of Artificial Intelligence took place at two great AI Laboratories. At MIT under the direction of Professor Marvin Minsky and at Stanford under the guidance of Professor John McCarthy. The complete, unabridged sets of reports, memos, idea exchanges and research findings at these two great laboratories is now available for the first time ever as a collection in two new sets of microfiche. The MIT set covers 1958-1979. The Stanford set covers 1963-1982. There's a special introduction to the MIT set by Professor Minsky. And over 10,000 pages of memoranda are represented.

**Over 800 reports organized, indexed and catalogued.**

The two breakthrough collections are carefully designed for research use. A fully annotated, indexed catalogue with abstracts covers such topics as seminal work in robots, LISP, expert problem solving, understanding language, vision, theories, systems and games. Retrieving reports or specific information has been greatly simplified by a team of top editors and research professionals. This important collection is a simple and easy to use reference.

---

**Essential for AI and Robotics libraries and laboratories.**

As a research tool, as a reference tool or simply as fascinating reading, these two first time collections will undoubtedly be in great demand. They are a must for the shelves of any institution, university or corporation that considers itself an entity in the world of Artificial Intelligence or Robotics. And they are clearly destined to be scientific classics to be used by engineers, experimental psychologists, linguists, computer scientists as well as AI professionals.

**MIT and Stanford: The basis for all future developments in AI.**

MIT, Stanford, Professor Minsky and Professor John McCarthy took a long but rewarding voyage into the exploration of thinking. On the way, they discovered more than they ever dreamed of. Now, you can discover the institutions and personalities that had the vision to see what the study of Artificial Intelligence would bring to the 80's. Share those great, scientific moments in two exciting microfiche collections. The MIT collection (1958-1979) sells for $2,450. The Stanford collection (1963-1982) sells for $2,750. Or better yet, order both sets now for $3,775 and save more than 25%.

---

**How to gain access to the complete scientific memoranda of MIT and Stanford's Artificial Intelligence Laboratories.**

Comtex Scientific Corporation
Department 101A
850 Third Avenue
New York, N.Y. 10022

- Please send _______ set(s) of the AI/MIT Memoranda (1958-1979) at $2,450 per set
- Please send _______ set(s) of the AI/Stanford Memoranda (1963-1982) at $2,750 per set
- Please send me more information about the Memoranda
- Please send both the AI/MIT and AI/Stanford Memoranda at the special discount price of $3,775.

Name ____________________________________________
Organization _______________________________________
Address __________________________________________
City __________________________ State ____________ Zip __________

Please feel free to call us Collect at 212-388-7200. Ask for Department 101A.