In AI, psychology, and education, a growing body of research supports the view that learning is a goal-directed process. Psychological experiments show that people with varying goals process information differently, studies in education show that goals have a strong effect on what students learn, and functional arguments in machine learning support the necessity of goal-based focusing of learner effort. At the Fourteenth Annual Conference of the Cognitive Science Society, a symposium brought together researchers in AI, psychology, and education to discuss goal-driven learning. This article presents the fundamental points illuminated at the symposium, placing them in the context of open questions and current research directions in goal-driven learning.

Learning is a central area of study for researchers interested in human cognition as well as those interested in machine intelligence. Its study has benefited greatly from the multiple perspectives provided by disciplines such as psychology, AI, and education. In AI, machine-learning research has developed a rich repertoire of learning mechanisms. However, less attention has been given to understanding the issues involved in applying these methods—when learning should occur, what knowledge should be learned, and which learning strategies are appropriate in a given context. Standard machine-learning systems address the question of when to learn by attempting to learn in response to every input; they address the question of what to learn by a user-supplied target concept (either explicit in the input provided to the system or implicit in the user’s choice of training examples); and they address the question of how to learn by applying a single, fixed learning method.

Although such systems provide a useful test bed for examining individual learning mechanisms, they are inadequate for use as real-world learners. The problem is that real-world situations offer countless opportunities for learning, and each of these opportunities licenses the learning of infinitely many concepts, few of which are actually useful. Consequently, an indiscriminate learning system will expend enormous processing effort to learn things that might provide little or no benefit. To learn effectively in such situations, a learning system needs ways to focus the learning process.

The need for focusing concept formation is widely recognized in AI, and standard focusing methods have emerged. Inductive learning systems depend on built-in biases to constrain what they learn (Michalski 1983), and explanation-based learning systems depend on prespecified target concepts and appropriate operationalization criteria (Mitchell, Keller, and Kedar-Cabelli 1986). In most systems, these focusing criteria are fixed (contrast Utgoff 1986). However, as circumstances change, the need for learning changes as well. Because inappropriate learning might actually degrade system performance (Minton 1988), effective performance depends on assuring that what is learned actually furthers system goals.

Goal-driven learning takes system goals as a starting point in the learning process. The idea of goal-driven learning is that because the value of learning depends on how well it satisfies system goals, system goals should direct decisions of when and what to learn. In this way, goal-driven learning follows the spirit of research on failure-driven learning systems, in which learning is motivated by deficiencies in system performance (Sussman 1975; Riesbeck 1981; Schank 1982; Birnbaum et al. 1990; Hammond 1989; Ram and Cox 1993; Schank and Leake 1989). Likewise, goal-driven learning is in the spirit of explanation-based learning research into forming useful target concepts (Kedar-Cabelli 1987) and judging the utility of learning (Keller 1987; Minton 1988). Goal-driven learning, however, takes a broader view, examining the relationships between the many possible motivations for learning and the many strategies to achieve it.

The effectiveness of goal-driven learning depends on being able to make good decisions about when and what to learn and on selecting the best strategies for achieving the desired learning. Unlike the passive and static process used in many learning systems, goal-driven learning is itself a planful process in which selection of target concepts and learning strategies is guided by desires and needs for knowledge (Hunter 1990).

Recent research provides growing support for goal-driven approaches to learning, both on cognitive and on functional grounds. In psychology, learner goals have been shown to have strong effects on the human learning process (Barsalou 1991;
Symposium Report

What Is Goal-Driven Learning?

One of the purposes of the symposium was to clarify the nature of goal-driven learning. We cannot simply describe goal-driven learning as learning that reflects system goals; this definition includes all learning systems because any learning system can be viewed as having a built-in goal to perform a particular type of learning. Thus, built-in attitudes, general desires to learn, or goals for self-improvement are equivalent to the implicit focuses of traditional machine-learning systems. To clarify the difference between goal-driven learning and traditional approaches, Thagard proposed distinguishing between goal-relevant learning, in which learning is relevant to goals in a weak sense, and goal-directed learning, in which the content of what is learned is driven by the general goals of the learner. All learning in real systems is goal relevant in some sense but not necessarily goal directed.

To the symposium participants, a crucial aspect of goal-driven learning was that it be dynamically focused by current information needs. This point sharpens the definition of goal-driven learning, specifying that a goal-driven learner chooses its own target concepts in response to information needs from goals outside the learning process. In this view, non-goal-driven learning is done by a system as routine processing, regardless of changes in its overarching goals; goal-driven learning is learning tailored toward providing the specific information (or type of information) currently needed to further overarching goals.

To flexibly adjust learning strategies toward satisfying system goals, a learning system must be able to reason about the information it needs. Consequently, Ram proposed that in truly goal-driven learning systems, the goals must be represented explicitly, and the learning system must be able to reason about the goals to guide its search for information (Ram 1991; Ram and Cox 1993; Ram and Hunter 1992). In this view, goal-driven learning is a process in which the learner forms and executes plans for learning needed concepts, guided by its knowledge of what it needs to learn and how these learning strategies can be applied.

This description of goal-driven learning highlights key differences between goal-driven and non-goal-driven learning systems. Traditional learning systems take a fixed target concept as an input (Mitchell, Keller, and Kedar-Cabelli 1986) or rely on the user to select appropriate examples to define the appropriate concept (Winston 1975); they also apply fixed learning strategies regardless of circumstances. Goal-driven learning systems decide when to learn, formulate their own target concepts, and decide how best to carry out their learning.

Effects of Goal-Driven Learning on Human and Machine Learning

The intuitive appeal of goal-driven learning is clear—to focus learning according to what the learner needs to know. However, Barsalou pointed out that because almost any cognitive processing can be construed as serving some goal, it is vacuous to make a blanket statement that human learning is goal driven. In his view, the study of human goal-driven learning must analyze specific goals, the processing that achieves them, the learning produced during processing, and the positive effects of this learning on subsequent goal achievement. Thus, the study of human goal-driven learning must focus on the processing that involves learning as a side-effect.

Barsalou discussed two classes of goals—explicit problem-solving goals and implicit orientation goals—for maintaining a world model. He argued that learning that takes place during explicit problem solving tends to be goal specific in that the information stored centers on achievement of the particular goal being achieved. This type of learning differs from learning based on general goal orientations, which results in learning that can later serve a wide range of explicit goals.
Data collected by Ng substantiate the influence of particular classes of goals in shaping human learning (Ng and Bereiter 1991). Based on her studies of student learning, Ng distinguished goal-driven learning (the learning that results from setting and pursuing goals beyond the completion of an assignment or learning procedure) from learning that is not motivated by learning goals (for example, learning that occurs as an incidental outcome of other processing or learning that occurs when the student attempts to complete an assignment but not to learn the concepts that the assignment is intended to teach). Although some students carry out learning tasks to advance their knowledge, others perform learning tasks primarily to maintain favorable appearances, avoid criticism, or gain praise (Dweck 1985).

Ng collected protocols from students learning BASIC and divided the students into three classes according to the relative frequencies of the statements they made relating to the following types of goal: (1) goals simply to complete the current task successfully, (2) goals to learn what the assignment was intended to teach, and (3) goals to build knowledge relevant to an outside agenda for learning. Students with explicit knowledge-building goals were better able to deal constructively with problems, raised more questions, and tagged unsolved problems for future investigation.

Other panelists argued that goal-driven learning promises to have a profound effect on machine learning as well. Leake, Michalski, and Ram pointed out that the necessity to function effectively despite limited processing resources makes it important for programs to be able to decide what it would be useful to know. Ram discussed the issue of deciding whether to pursue the related knowledge goals immediately or to suspend it until a better opportunity to satisfy it arises in the future. Ram and Michalski also highlighted the need to reason about the ways in which learning goals can be achieved, that is, the different learning strategies that can be used. Thagard pointed out that this reasoning must consider not only how to achieve given goals but also how to reconcile conflicts between these goals; in this way, goal-driven learning both furthers goal achievement and clarifies what the system’s goals are.

**Issues in Goal-Driven Learning**

Building goal-driven learning systems depends on addressing concrete issues affecting the goal-driven learning process. The symposium panelists identified five fundamental questions that must be addressed by theories of goal-driven learning: (1) What are the types of learning goals? (2) How do learning goals arise? (3) How do learning goals affect the learning process? (4) How do different types of learning goals relate to each other? (5) How are learning goals represented?

**What Are the Types of Learning Goals?**

Developing goal-driven learning mechanisms depends on identifying the goals that can drive learning. Panelists proposed four ways of classifying learning goals.

- By overarching tasks: Most of the panelists classified learning goals by the tasks that give rise to them. Barsalou distinguished two types of goals: those serving explicit problem solving and those reflecting implicit orientations. Leake described a taxonomy of task-based information requirements relevant to the generation of explanations. Ng categorized the goals underlying student effort on class assignments into three types: task-completion goals, which are achieved by simply completing the assignment satisfactorily; instructional goals, which reflect what the program of instruction is intended to teach; and knowledge-building goals, which relate to the student’s own purposes or agenda for learning (for example, to use the knowledge that the exercise is designed to impart to further some other goal).

- By the learning that results: Ng distinguished five fundamental questions that must be addressed by theories of goal-driven learning: (1) What are the types of learning goals? (2) How do learning goals arise? (3) How do learning goals affect the learning process? (4) How do different types of learning goals relate to each other? (5) How are learning goals represented?

- By the learning activity: Michalski suggested that learning goals can also be characterized by the particular learning activities that they involve, such as generalizing given knowledge, discovering regularities, or placing the knowledge in more operational form.

- By the learning goals that result: Leake (1992) discussed the role of failure in triggering learning, Leake concentrating on learning in response to anomalies during understanding and Ram developing a more general taxonomy of failures that result from multiple types of tasks. In their view, the goal-activation process can be characterized by the sequence in Figure 1.

- By the knowledge gap or failure: Michalski distinguished between learning goals that are hard wired and those that arise from other influences in human learners. Ng pointed out that in human learners, culture can have significant effects on back-
ground learning goals; consequently, developing effective instructional materials depends on understanding and influencing background learning goals.

How Do Learning Goals Affect the Learning Process?

Focusing on storage of information during processing, Barsalou observed that most psychological theories assume that the storage of information is done unintentionally; a problem solver attempting to solve a problem simply stores a trace of its processing without attention to its future relevance. However, Ng's previously mentioned studies show that for a different class of task, learning goals have a strong effect on the learning performance of human learners. A future question is to identify the limits of goal-driven processing in human learners.

From an AI perspective, other panelists focused on how connections are made between needs for information and relevant learning strategies. Ram pointed out that goals can affect many parts of the learning process, such as focusing attention, controlling what is learned, and determining how to select and combine learning strategies (Ram and Cox 1993). He stressed the importance of having an explicit representation of learning goals to support the goal-driven learning process. This representation would allow the reasoner to notice and avail itself of unexpected opportunities to learn something that was previously identified as important or interesting. He also advanced that the goal-driven learner's decision process must involve ways to reason about the relative priorities of pending goals, select and combine learning strategies, and suspend and opportunistically trigger learning goals when circumstances make them appropriate.

In the context of explanation, Leake identified broad classes of tasks, such as prediction, prevention, and repair tasks, and connected them to requirements for information; these connections allow an explainer to judge whether candidate explanations provide the information needed for useful learning. In his model, requirements for filling system knowledge gaps also direct explanation generation by guiding retrieval and revision of explanations during case-based explanation construction (Leake 1992). In the context of analogical mapping, Thagard pointed out that goals, semantic constraints, and syntactic constraints all affect analogical mapping (Holyoak and Thagard 1989) and the retrieval of potential analogs (Thagard et al. 1990).

Michalski described the inferential theory of learning, a theoretical framework in which learning depends on input information, prior knowledge, and learning goals. In the theory, current learner knowledge is transformed into desired knowledge according to a set of transmutations for searching knowledge space, such as generalization, discrimination, and reformulation of concepts (Michalski 1993). As the bridge between learning goals and learning strategies, Michalski presented a taxonomy associating inference types with knowledge transmutations that serve these types of inferences.

How Do Different Types of Learning Goals Relate to Each Other?

Both Michalski and Ram advanced models that treat learning as a planful process. In their models, explicit reasoning about information needs and information requirements determines subgoals for learning activity. Thagard focused on the need to reason about the relationship between learning goals to determine which goals are relevant in a given situation: Agents often simultaneously have a number of interrelated goals that must be balanced and reasoned about to determine which goals to keep, which goals to abandon, and what goal orderings apply. He described current research on the coherence theory of decision and on the DEC0 system, which uses connectionist algorithms for parallel constraint satisfaction to both make decisions about actions and adjust goal priorities.

How Are Learning Goals Represented?

For a system to reason about its information needs, it must be able to represent what these needs are. Ram proposed representations that include the desired knowledge (possibly partially specified) and the reason that the knowledge is sought. Leake focused on the representation of the knowledge required to resolve anomalies (which depends on a vocabulary of anomaly characterization structures to describe the information needed to resolve an anomaly) and on dimensions for representing the types of information that must be provided by explanations for different tasks. Michalski presented the goal-dependency network, a representation that includes the general goal being served, subordinate goals, attributes relevant to the goals, and the relationships connecting them. He also showed how a learning system using such a goal-dependency network creates different conceptual classifications of input data depending on the top-level goal and associated subordinate goals.

Properties of Goal-Driven Learners

The properties of goal-driven learning described earlier suggest some basic design principles for goal-driven learning systems. First, because the systems reason about current goals to
Innovative Applications of Artificial Intelligence 5: Proceedings of the IAAI–93 Conference
Phil Klahr and Elizabeth Byrnes, Editors

The applications described in this book not only demonstrate the commercial value of software systems employing AI techniques but also delineate the process by which these software systems are developed and deployed. The book points out the techniques or combination of techniques that are useful in particular circumstances or the difficulties encountered in applying or integrating different techniques. It illustrates the importance of understanding and addressing the needs of the user community. Each application presents a case study that contains valuable lessons for those who will develop future AI applications.

Published by AAAI Press
445 Burgess Drive • Menlo Park, California 94025-3496
ISBN 0-929280-46-6 $30.00, softcover
To order call: (415) 328-3123 or fax to (415) 321-4457. MasterCard and VISA accepted.

make decisions about what information to seek and how to seek it, learning in such systems is an active process. This learning process is unlike the passive process of systems that accept the information they are given and apply to it a fixed learning procedure. Likewise, this view requires that goal-driven learners reason about how to acquire needed information, making learning a planful process (Hunter 1990; Ram and Hunter 1992).

Finally, to generate learning goals, goal-driven learning systems must be introspective: They must be able to notice gaps in their knowledge and to reason about the information needed to fill these gaps. Developing introspective systems requires that the system have a representation of its own processes to detect deviations that show learning is needed (Ram and Cox 1993; Kruilwich, Birnbaum, and Collins 1992). Experimental results in the metacognition literature also suggest that introspective or metacognitive reasoning can facilitate human learning (Schneider 1985; Weinert 1987).

The Relationship between Goal-Driven and Non-Goal-Driven Learning

Goal-driven learning offers significant advantages over non–goal-driven methods, but it also plays a complementary role to these methods; a number of panelists argued that both are needed for successful performance. Goal-driven learning is important because in complex domains, learners are faced with an overwhelmingly large set of alternatives that could be learned. This statement led to functional arguments in favor of goal-driven learning, advanced by Barsalou, Leake, Michalski, and Ram, that focused on the importance of constraining the field of possible generalizations, thus restricting the effort of a learning system and allowing it to be applied more effectively (Leake 1992; Michalski 1993; Ram and Hunter 1992).

However, non–goal-driven learning is needed because it is impossible to anticipate all future needs for information; learning exclusively in service of current goals might not take advantage of opportunities for low-cost learning. Thus, controlled non–goal-driven learning can also be beneficial, provided that it can be done at sufficiently low cost. Along these lines, Barsalou observed that one form of human learning—storage of information—seems to be influenced by goals only indirectly, in that it maintains a trace of processing that might be goal directed. Both Barsalou and Leake stressed the importance of maintaining an accurate world model to support future goal-based activity, even though, in general, the task of maintaining the model is only indirectly related to goals that are active at the time of learning.

The trade-off between goal-driven and non–goal-driven learning was addressed by Thagard, who observed that if goals are the sole influence in deciding what to learn, to the exclusion of other semantic or pragmatic constraints, the objective accuracy of learning can be compromised. (An example of this sort of distortion is a student who is eager to avoid blame for a bad test score and attributes the score to the teacher's unfairness, even though the true explanation is that the student had not studied hard enough. The explanation might serve the student's goals, but what is learned is nevertheless false.) The issue here is the balance between learning what is most useful to the learner's current tasks and maintaining an accurate view of the world. Thagard discussed a model of analogical learning in which such influences are balanced.

Conclusions

Goal-driven learning systems learn in response to explicit goals for knowledge. Goal-driven learning allows flexibility of processing that is otherwise impossible in learning systems: A goal-driven learning system's choices of what to learn, when to learn, and which learning strategies to use can be tailored toward achieving effective learning. In addition to these functional supports for the goal-driven learning process, psychological experiments support its validity as a cognitive model.

The symposium on goal-driven learning revealed the common ground and differences between disparate research efforts on goal-driven learning. It identified fundamental questions about the goal-driven learning process and pinpointed a number of important avenues for future research. Major issues include developing appropriate representations for learning goals, developing principles for resolving contradictions among competing goals, and developing theoretical principles and practical mechanisms for reflecting goals in the learning process.

The symposium is only a first step toward developing a paradigm whose ramifications on learning systems are likely to be considerable. Viewing learning as a goal-driven process suggests a new generation of active,
planful, and introspective learning systems that can learn effectively in complex situations by reasoning about what, when, and how to learn.

Acknowledgments

The authors would like to thank Ryszard Michalski and Michael Cox for their helpful comments on a draft of this article.

References


David Leake is an assistant professor of computer science and a member of the cognitive science faculty at Indiana University. He received his Ph.D. from Yale University in 1990. His primary research interests are case-based reasoning and explanation, which are the focus of his book Evaluating Explanations: A Content Theory (Lawrence Erlbaum Associates, 1992). He is coediting (with Ashwin Ram) the forthcoming book Goal-Driven Learning (Bradford Books).

Ashwin Ram is an assistant professor in the College of Computing at the Georgia Institute of Technology and an adjunct professor in the School of Psychology. He received his Ph.D. from Yale University in 1989. His research interests are in the areas of machine learning, natural language understanding, explanation, and cognitive science. He is coediting (with David Leake) the forthcoming book Goal-Driven Learning (Bradford Books).