ACQUIRING SCHEMATA THROUGH UNDERSTANDING AND GENERALIZING PLANS

Gerald DeJong
Coordinated Sclenoe Laboratory
1101 West Springfield Avenue
University of Illinois
Urbana, IL 61801

ABSTRACT

This paper discusses explanatory schema acquisition, a learning technique with several interesting properties. It does not require a teacher or concept matching predicate to be provided. It does not rely on searching a concept space to produce generalizations. It can acquire a new concept based on only one input example, although later inputs might result in refinement of learned concepts. These features are made possible by taking a very knowledge-based approach.

I. INTRODUCTION

This paper gives an overview of a learning technique being developed at the University of Illinois. The technique, called explanatory schemes acquisition has some interesting properties. For example, it does not require a teacher or other oracle to seleot important examples; it is capable of one trial learning; and contrary to Mitohell's recent taxonomy of learning systems [13] it does little or no searching in the process of acquiring a new concept.

Before describing the technique we will pause briefly to consider what we might call the "standard theory of concept formation." This approach underlies much of the concept learning work in psychology and AI. In the standard theory, a system is given a number of inputs. Each input has some structure. Part of the structure is essential for the concept, part can be varied. An input with the proper structure is an instance of the concept; otherwise it is not an instance. A teacher, usually a human, supplies sample inputs to the system together with the information of which concept (if any) the input is a concept of. The system's task is to discover the structure that defines each concept. This approach has been fruitfully applied to many diverse domains (for example, [11], [12], [18]) and is a cornerstone of the field of inductive inference.

Now we can ask how we might construct an untutored concept learning system. At first this seems a bit of a non sequltur. Removing the teacher appears to result in no learning at all. Also the notion of forming a concept from a single input seems suspicious.

This work was supported in part by the Air Foroe Office of Scientific Researoh under grant F49620-82-K-0009 and in part by the National Sclenoe Foundation under grant NSF 1ST 81-20254.

The key is to adopt a much more knowledge-based approach. The learning algorithm to be described requires access to a large amount of domain knowledge. It is through reconciling a new input to the domain knowledge that learning and generalization occurs.

This is NOT to say that the proposed learning technique is domain specific. Techniques specific to a particular domain would be ad *hoc* and of very limited theoretical interest. Rather, explanatory schema acquisition is domain independent. Indeed, it has already been applied to three very different domains. The approach does, however, require access to a rich domain model. It is interaction with this rich domain information that determines whether or not concept acquisition is possible or desirable for a new input. The interaction also guides the generalization process.

II. EXPLANATORY SCHEMA ACQUISITION

The process involves three logically distinct (but possibly concurrent) processes:

1) The new input is understood.
2) The input is evaluated to see if schema formation is warranted.
3) The input is generalized to a new schema.

For illustrative purposes we will assume that the input is the following brief natural language story. The assumption of natural language input is not necessary and, indeed, one of the current applications involves robot arm planning which, of course, is non-linguistic.

> John, a bank teller, discovered that his boss, Fred, had embezzled $100,000. John sent Fred an inter-office memo saying that he would inform the police unless he was given $15,000. Fred paid John the money.

A. Understanding the Input

The requirements on the understanding prooess are not controversial. By "understanding" we mean nothing more than oonstruoting a causally complete representation of the input. This requires that any crucial information missing from the story must be Inferred and that the causal relations between components be discovered and made explicit. While this is not an easy task, it is one which has been the focus of a good deal of researoh, particularly for natural language texts ([2], [9]).

We require that our representation have one component that is not generally inoluded by understanding systems. We require that the understander maintain data dependency links ([3], [5]) justifying each element in the representation. The links connect each representation event with all of the inference rules from the domain model that were used to Justify the event during the understanding process. This includes all causal information, goal enablements, planning information, etc. This makes explicit in the final story representation the reasons the system had for connecting events in a particular way. For example, in the blackmail story John's demanding that Fred give him $15,000 is explicitly mentioned. The system must infer that John has the goal of possessing the $15,000. This is a necessary inference. A system cannot be said to have "understood" the input (in any sense of the word) if it does not make this inference.

By and large, current understanding systems do not include these backpointers to inference rules in the final representation. We will insist that they be explicitly stated in the understood representation. We call the amalgam of all of these data dependency links the Inference Justification Network.

B. To Generalize or Not to Generalize

There are five aspects to be considered when deciding whether or not to generalize an input into a new schema. By hypothesis we will assume that the input did not match an existing schema (if it had then the system already possesses the desired schema and indeed that schema would have been used to process the story). If any of these five conditions does not hold, constructing a new schema from this input is inappropriate.

The criteria are:

1) Is the main goal of a character achieved?
2) Is the goal a general one?
3) Are the resources required by the goal achiever generally available?
4) Is this new method of achieving the goal at least as effective as the other known volitional sohemata to achieve this goal?
5) Does the input match one of the known generalizable patterns?

These criteria are tested for all goals in the story. The first criterion "Was the goal achieved?" is self explanatory and easily judged. The second "Is it a general goal?" and the third "Are the resources generally available?" require some discussion.

Novelty alone in an approaoh to aohieving a goal is not sufficient to warrant constructing a new sohema. Consider, for example, a plot from the "Mission Impossible" television series. These plots are very novel but too specific to be useful again.

How can the utility of a particular goal be Judged? The answer to this is olosely tied to where goals come from. Aohieving a goal which arises from general conditions important to an Individual's well-being and using readily available resources is likely to result in an interesting new schema, one which will arise again and again. For the solution, we use an aspect of Schank and Abelson's theory of planning [15]. In their view themes give rise to the highest level goals (goals which are not simply subgoals in the achievement of other goals). Interpersonal and Life themes are what we are Interested in. An example of the former is a husband offering (and therefore, at some level, wanting) to type a term paper for his over-worked student wife. People often work to satisfy the goals of their loved ones. The system should realize that this requires no further explanation on the part of the husband. Examples of the latter are attempting to satisfy one's hunger, to gain money, or to relieve boredom. Life themes give rise to goals that require no further justification. Our example, which demonstrates a new way to gain money, relates directly to a life theme and therefore satisfies this criterion.

Criterion 4 is self-explanatory. The idea is that the system should not bother constructing schemata that are much less efficient than similar already-known schemata.

The fifth criterion has been discussed elsewhere [1]. As this is a short paper describing on-going research it is not appropriate to repeat it here. Suffice it to say that there is a taxonomy of explanatory acquisition techniques. The teohnique that is matched has implications for exactly how the generalization is performed.

c. The Generalisation Process

Assuming the input is completely understood (with data dependency links to inference rules justifying the understanding) and the five tests for learning have all been met, we must now perform the actual generalization. The generalization process oonsists of replacing the objects and actions in the understood representation with abstract counterparts. These counterparts are the most abstract possible while still preserving the validity of the inference justification network.

Consider again the example of John blackmailing Fred. One proposition that is a part of the understood representation is that Fred decided to pay John $15,000. This action is justified to the system by a number of other propositions. Among these supporting propositions are some supplied by the schema DECIDE (which we assume the system already possesses). These inferences from DECIDE are: 1) the decider must be at least a higher animate, 2) the decider must be oapable of a number of alternative possible actions, 3) the decider must know what the alternatives are, and 4) the chosen alternative will be among the most beneficial/least detrimental to the decider.

Thus, these (and other) justifications are tied to the representation of Fred's decision through data dependenoy links. Fred's deoision is believable to the system beoause Fred, in fact, is a higher animate, he knows at least two alternatives - paying John or losing the $100,000 and being arrested, and 3) he probably sees losing $15,000 as less detrimental than losing $100,000 and going to Jail. These justifications are supplied in the form of pointers to the above inferenoe rules during the understanding procedure.

The generalization prooess substitutes general entities for the specific objects and events that occurred in the story. The entities are the most general possible that still preserve the validity of the data dependency inference.

Through these and other generalizations the system can construct a first version of a BLACKNAIL schema. The schema might not be perfect. There may be later stories that do not quite fit and require further modification of the schema. However, it is a reasonably general schema that is likely to help a good deal in processing future similar stories.

III.CONCLUSION

There are several concluding points

1) Unlike many learning systems (e.g., [6], [11], [18]) explanatory schema acquisition does not depend on correlational evidence. It is capable of one trial learning, but the learning is not based on analogical reasoning like [17] and [19]. It is somewhat similar to Soloway's view of learning [16]. There is also some resemblance to the MACROPS notion in the STRIPS system [4].

2) The approach is heavily knowledge-based. A great deal of background knowledge must be present for learning to take place. In this respect explanatory schema acquisition follows the current trend in AI learning and discovery systems perhaps traoeable to Lenat [10].

3) The learning mechanism is not "failure-driven" as is the MOPs approaoh ([14], [7], [9]). In that view learning takes place in response to incorrect predictions by the system. In explanatory acquisition learning is usually stimulated by positive inputs which encounter no particular problems or prediction failures.

4) The absolute representation power of the system is not enhanced by learning new schemata. This statement is only superficially surprising. Indeed, Fodor [6] shows that this must be true of all self-consistent learning systems. Explanatory schema acquisition does, however, increase processing efficiency. Since all real-world systems are resource limited, this learning technique does, in fact, increase the system's processing power.

ACKNOWLEDGMENTS

I am indebted to the members of the CSL learning group: Chris Debrunner, Paul Harrington, Paul O'Rorke, and Alberto Segre.

REFERENCES

[1] DeJong, G., "Automatic Schema Acquisition in a Natural Language Environment," Proc. Nat'1. conf. on Artificial Intelligence. Pittsburgh, PA, August 1982, pp. 410-413.

[2] DeJong, G., "Prediction and Substantiation: A New Approaoh to Natural Language Processing," Cognitive Science 3 (1979) pp. 251-273.

[3] Doyle, J., "A Truth Maintenance System," Artificial intelligence 12:2 (1979) PP. 231-272. Teohnical report TR-419, MIT, Cambridge, MA.

[4] Fikes, R., P. Hart and N. Nilsson, "Learning and Executing Generalized Robot Plans," Artificial Intelligence 3:2 (1972) pp. 251-288.

[5] Fikes, R., "Deductive Retrieval Mechanisms for State Description Models," Proc. Fourth Int. Joint conf. on Artificial IntelllgfiAfie., 1975, pp. 99-106.

[6] Fodor, J., The Language of Thought. Thomas Y. Crowell Company, New York, 1975.

[7] Kolodner, J., "Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model," Computer Science Research Report 187, Ph.D. dissertation, Yale University, New Haven, CT, 1980.

[8] Langley, P., "Data-Driven Discovery of Physical Lava," Cognitlve Science 5 (1981) pp. 31-54.

[9] Lebowitz, M., "Generalization and Memory in an Integrated Understanding System," Computer Science Research Report 186, Ph.D. dissertation, Yale University, New Haven, CT, 1960.

[10] Lenat, D., "AM: An Artificial Intelligence Approach to Discovery In Mathematics as Heuristic Search," AIM-286, AI Laboratory, Stanford University, Stanford, CA, 1976.

[11] Michalskl, R., "A System of Programs for Computer Aided Induction: A Summary," Proc. Fifth Int'l. Joint conf. an Artificial Intelligence. Cambridge, MA, 1977, pp. 319-320.

[12] Mitchell, T., "Version Spaoes: A Candidate Elimination Approaoh to Rule Learning," Proc. Fifth Int'l. Joint Conf. .on Artificial Intelligence. Cambridge, MA, 1977, PP. 305-310.

[13] Mitchell, T., "Generalization as Searoh," Artificial Intelligence 18:2 (1982) pp. 203-226.

[14] Sohank, R., "Language and Memory," Cognitive Science. 4 (1980) pp. 243-283.

[15] Sohank, R. and R. Abelson, Scripts Plans Goals. and. Understanding. Lawrence Erlbaum Associates, Hillsdale, NJ, 1977.

[16] Soloway, E., "Knowledge Direoted Learning Using Multiple Levels of Description," Ph.D. dissertation, Computer Science Department, University of Massachusetts, Amherst, MA, 1977.

[17] Winston, P., "Learning and Reasoning by Analogy," Commun. ACM 23:12 (1980).

[18] Winston, P., "Learning New Principles from Precedents and Exerolses," Artificial Intelligence 19:3 (1982) pp. 321-350.

[19] Carbonell, J., "Experiential Learning in Analogical Problem Solving," Proc Nat'l. Conf. *on* Artificial Intelligence, Pittsburgh, PA, August 1982.