GENERALIZATIONS BASED ON EXPLANATIONS

Gerald DeJong

Coordinated Science Laboratory
University of Illinois
1101 West Springfield Avenue
Urbana, IL 61801

This paper describes a new project at the University of Illinois in computer learning. The phenomenon under study is a kind of "insight learning" of procedural schemata. The system described here is designed to grasp some principle underlying a natural language input. The underlying principle results in a new schema for the system. Once acquired, the schema serves the same purpose as the other schemata in the system: it aids in processing future natural language inputs.

The neutral term "schema" rather than "frame" (Minsky (1975), Charniak (1976)) or "script" (Schank A Abelson (1977)) is used to refer to knowledge chunks because a frame (which is used to describe static objects as well as progressions of world situations) is too general a notion, and the notion behind a script is- too specific. Scripts cannot be made to fit a novel situation nor are they designed to represent the more abstract concepts that this system will need.

The process that the system uses is called <u>explanatory schema acquisition</u> . The basic idea behind it is that the causal connections in an understood representation of a new input can be used to propose and propagate constraints on slot fillers. That is, from one particular instance of a situation the system can "reason out" the a general structure underlying that instance. The system is therefore capable of learning from just one example.

There has been much work in the past on learning systems which start from a state of very little or no initial knowledge (for example, Minsky and Papert (1969), Meisel (1972)). To a large extent, these systems failed to live up to the expectations of researchers. More recently, learning and knowledge acquisition systems have been constructed on a firm foundation of rich initial knowledge (Lenat (1976), Soloway (1977), Buchanan and Mitchell (1973), McDermott (1979), Selfridge (1980)).

The work reported in this paper was supported in part by the Office of Naval Research under conract N00014-7S-C-0612.

The system described here is a knowledge based one. It must contain much background information on how the world behaves and on the goals and motivations of people. Prom this, it is able to learn more about the behavior of the world.

In the example below assume the system does not yet have schemata for kidnap or extortion. It does, however, possess a considerable quantity of background information about stealing, bargaining, the use of normal physical objects, and goals of people, companies and institutions. Some of this knowledge is in the form of other schemata already built into the system.

> Paris police disclosed Tuesday that a man who identified himself Jean Maraneux abducted the 12 year old daughter of wealthy Parisian businessman Michel Boullard late last week. Boullard received a telegram demanding that 1 million francs be left in a lobby waste basket of the crowded Pompidou Center in exchange for the girl. Asking that the police not intervene, Boullard arranged for the delivery of the money. His daughter was found wandering blindfolded near his downtown office on Monday.

In processing the story the system will be able to learn a general schema for processing kidnap stories from this one instance. Using explanatory schemata acquisition, understanding is viewed as the process of explaining input events. The explanations can then be used to generalize a single event into a new schema.

In processing this example without a KIDNAP or EXTORTION schema the system cannot explain all of the events via existing schemata. In particular, existing schemata cannot explain why Maraneux might steal Boullard's daughter. While this is quite clearly an instance of taking something that belongs to someone else, there is no motivation for it. The daughter has no apparent value to Boullard, and a person, unlike money, cannot be used to acquire other valued goods. The system requires motivations for major volitional actions (su^h as a character invoking the STEAL schema). Therefore it is confused at this point.

The confusion is resolved by the next sentence. This input invokes the BARGAIN schema. The system understands the motivation for Maraneaux to initiate the bargaining event: he is trying to acquire money which it knows to be a possible goal of any human. Furthermore, this provides the motivation for the STEAL event. Maraneaux used the STEAL schema to satisfy the precondition of the BARGAIN schema of possessing the item to be traded.

Resolving the confusion causes the system to invoke its explanatory schemata acquisition procedure. This procedure does two things. First, it constructs a new schema composed of a STEAL event and a BARGAIN event where the STEAL is used to satisfy a precondition of the BARGAIN. Second, constraints on the slots for the new schema are derived from the knowledge in the systems STEAL and BARGAIN schemata and the story as follows:

1) the slot filled by 1 million francs is generalized to be any amount of money.
2) the slot filled by Maraneaux (the kidnapper) is generalized to be any adult human.
"O the slot filled by the daughter is generalized to be anyone with close personal ties with (4).

4) the slot filled by Boullard is generalized to be any human who both has the amount of money to fill (1) and a person with close personal ties to fill

Thus the system now has a schema that can be used to process a new story about a person stealing another person in order to trade him back for money. This is, of course, a first approximation to a KIDNAP schema.

One might make the argument that learning the KIDNAP schema is unnecessary since the system could, after all, correctly process the example input without it. Why not simply rely on the more general planning/goal knowledge and forget about the schemata? There are three reasons.

First, there is an efficiency consideration. The more a system can rely on pre-packaged schemata, the less goal following and complex planning it will have to do.

Second, planning systems have proved to be more fragile and distinctly less successful than script/frame/schema systems. It might well not be possible to build a successful planning type system to deal with an interesting domain. The problem is that the fragility of a planner is multiplied by long reasoning chains, a domain rich in alternatives (as the real world tends to be) and the need to invoke the planner many times in the course of processing an input. In the proposed system, most of the processing is handled by the schema knowledge. Only when a new schema is being constructed or refined must the system resort to less reliable knowledge sources.

Finally, script/frame/schema systems have been much criticized as 'ad hoc*. This is due to the fact that programmers invariably buill in all of the necessary (and no unnecessary) knowledge structures. A system that could acquire these knowledge structures for itself would help dispel that criticism.

This is not the end of the learning process for the KIDNAP schema. The system is designed to refine and expand initial schemata in the course of processing more stories. We hope that the system will be able to expand the initial KIDNAP schema into a general EXTORT schema by recognizing that other things besides people ran be stolen and traiel back.

The system is only in the beginning stages. As yet it is not implemented.

In conclusion, it should be pointed out that this kind of learning does not depend on correlational evidence. That is, the new schema is not constructed by noticing the similarities and differences among a large number of inputs (like Winston (1970) or Fox & Reddy (1977)). Nor does it construct special purpose knowledge structures from mor« general ones already existing within the system (e.g., Lebowitz (1981)). Rather, the system is capable of "one trial" learning not unlike Soloway's (1977). From just one kidnap story, the system constructs a plausible KIDNAP schema. This schema can later be refilled and expanded, but from the beginning it is a viable new schema capable of aiding future processing.

REFERENCES

Charniak, E., (1976) A framed PAINTING: the representation of a common sense knowledge fragment. Working paper 29, Institute for semantic and cognitive studies.

Fox, M. and Reddy, R., (1977) Knowledge-guided learning of structural descriptions. Proceedings of the fifth IJCAT, Cambridge.

Lebowitz, M., (1980) Gneralization and memory in an integrated understanding system. Computer science research report 136, Ph. D. dissertation, Yale University, New Haven, CT.

Lenat, D., (1976) AM: an artificial intelligence approach to discovery in mathematics as heuristic search. AIM-2S6, AT Laboratory, Stanford University, Stanford, CA.

McDermott, ,]., (1979) Learning to use analogies. Proceedings of the sixth IJCAI, Tokyo.

Mei9el, W., (1972) Computer oriented approaches to pattern recognition. Academic Press, New York.

Minsky, M. & Papert, S., (1969) Perceptrons. MIT Press, Cambridge, MA.

Minsky, M., (1974) A framework for the representation of knowledge. MIT AI report TR-^06, MIT, Cambridge, MA.

Schank, R. & Abelson, R., (1977) Scripts Plans Goals and Understanding. Lawrence Earlbaum Associates, Hillsdale, NJ.

Selfridge, M., (1980) A process model of language acquisition. Computer science research report 17?, Ph. D. dissertation, Yale University, New Haven, CT.

Soloway, E., (1977) Knowledge directed learning using multiple levels of description. Ph. D. dissertation, Computer Science Department, University of Massachusetts, Amherst, MA.

Winston, P., (1970) Learning structural descriptions from examples. MIT Project MAC report 76, Cambridge, MA.