

Knowledge Acquisition for a Diagnosis-Based Task

Daniel E. O'Leary and Paul R. Watkins
Graduate School of Business
University of Southern California, Los Angeles, U.S.A.

Summary

The basic research task discussed in this paper is the investigation of the relative effectiveness of alternative knowledge acquisition methodologies at eliciting different types of knowledge for a diagnosis task. In order to investigate this task, a field study was used.

The scope of the study was to elicit the knowledge required to build an expert system to solve a diagnosis task. The study was designed to solve a specific problem within a specific organization. However, because of the generic diagnosis nature of the task, the study and design of the system offers insights into knowledge acquisition in similar settings.

The study used four basic forms of knowledge acquisition and identified at least five different types of knowledge used in the solution of the diagnosis problem. Further, it was found that different forms of knowledge acquisition were more effective in the elicitation of the different types of knowledge. Finally, an entire portfolio of approaches was necessary to elicit the knowledge to construct the system.

1 Introduction

Recently, there has been dissatisfaction with the knowledge elicitation and acquisition literature. In particular, there has been concern over which knowledge acquisition approaches should be used. Researchers have responded with theoretical discussions of which types of approaches should be used. Unfortunately, for each approach that is suggested as plausible, alternative arguments can be presented to make other approaches equally as plausible. Researchers also have begun to employ experimental techniques in an effort to see which technique can be used to 'best' establish the knowledge base for an expert system (e.g., Burton et al. (1987)). Clearly, it is important to use appropriate methodologies to examine the problems of knowledge acquisition: purely logical or theoretical approaches are incomplete. However, it is also true that a single approach to knowledge acquisition rarely is used to build an expert system. In addition, the use of a single approach is likely to result in a single dimensional system. As a result, it is important to study the process of building expert systems and the employment of a portfolio of knowledge acquisition approaches.

As with previous experimental research, such an inquiry is likely to be specific to the generic type of system (e.g., diagnostic) and the domain (e.g., business). However, it is also important to not divorce the investigation too much from the setting and the resulting experiences in which the experts are based (e.g., Kolodner and Riesbeck (1986)). Thus, in order to capture the context, such systems could be developed in the particular firm or organization.

1.1 Field Study Research

Even though there appears to be a generic nature to particular tasks (e.g., Hayes-Roth et al. (1983)), the domain and the organizational setting of the task can play a critical aspect in the expertise, and thus, the knowledge acquisition, associated with the task. One way to study the issue of the effectiveness of alternative knowledge acquisition methods for such generic-like systems, yet still allow for context is to use a field study.

That is the methodology used in this paper. As discussed by Kaplan (1964), "Much of . . . science is occupied with field studies, that is, with direct or indirect observation of behavior in the circumstances in which it occurs without any significant intervention on the part of the observer. . . . Of course in his role as a consulting expert the . . . scientist may be taking part in some part of an experiment. We should not overlook the opportunities for genuine and important scientific work afforded by this social role".

The field study approach to expert systems has been used in previous research. For example, Watkins and O'Leary (1987) present a paper that employs a field study approach to the study of the development of expert systems by consulting experts.

1.2 This Paper

This paper proceeds as follows. Section 1 has presented an introduction to the paper and the approach. Section 2 discusses diagnosis-based systems, in general, and for the particular application of this paper. Section 3 reviews the four basic knowledge acquisition approaches used to develop the system discussed in this paper. Section 4 discusses the findings. Section 5 provides a brief summary of the paper.

2 Diagnosis and Diagnosis Systems

Typically, in diagnosis, a situation is presented to an investigator. The investigator is then asked to determine what is wrong or how to remedy the situation. This requires eliciting data on the situation and then recognizing patterns of problems from that data. Based on the patterns recognized, possible solutions are suggested, as part of the diagnosis, to solve the problem in the situation. Thus, the quality of the diagnosis is based on the ability to solicit the appropriate situation information and to recognize the resulting patterns. One without the other is useless.

Expert systems designed to solve diagnosis problems have been developed for applications in medicine (e.g., Buchanan and Shortliffe (1985)), mining (e.g., Duda et al. (1979)), and business (Meservy (1985)). Hayes-Roth et al. (1983) describe diagnosis expert systems in the following manner: "Diagnosis systems infer system malfunctions from observables. . . . Diagnosis systems typically relate observed behavioral irregularities with underlying causes . . .".

Diagnosis tasks have been among the most successfully modeled activities using expert systems technology (Hayes-Roth et al. (1983)). As with human diagnosis, computer diagnosis is dependent on the ability of the system to solicit the appropriate information, from which patterns can be elicited.

2.1 Particular Diagnosis Task Used in the Field Study

Because of the widespread nature of such tasks, we choose a diagnosis situation to determine what knowledge acquisition approaches "worked best". The particular task was diagnosing communications problems for a large timesharing software company.

In particular, the company has a number of clients that use programs on a time sharing basis, via communications to the company's central computer. The users of those systems are not sophisticated computer users. As a result, the company employs a large base of technicians to assist these users with problems they may have using the software. Included in this base of problems that technicians must address are communications problems (modems, etc.).

The technicians are located at a central location to which users refer all questions on their use of the software. All questions are answered over the phone. Either the questions are solved while the user is on the initial call or the technician will call back with the information. Sometimes this is an iterative process, with the technician offering a solution, getting feedback from the user and then offering another solution.

Since the diagnosis is off-site, the information must be gathered by someone other than the person doing the diagnosis. This person typically is not a sophisticated communications expert.

Unfortunately, the 'technicians' often come into their role without adequate understanding of the technology. In addition, the underlying technology continually changes. Further, these technician positions experience rapid turnover. When they completely understand the technology, they are attractive to other departments or firms. As a result, the company must continually educate the technicians about communications and other problems. This education process takes the form of direct education and indirect apprenticeship-like learning.

In each case the training is facilitated by documentation, that also is useful in the assistance of technicians on a day-to-day basis. The documentation consisted of segments of manufacturer's documentation, additional documentation on internally developed software and 'tips' on recent problems. Typically, these 'tips' are developed by the more expert technicians in response to recent events that have the potential to affect a number of different clients, and thus, technicians.

2.2 Scope of This Paper

The scope of this paper is to investigate multiple forms of knowledge acquisition for the development of a system. This is in contrast to much of the formal investigation of alternative forms of knowledge acquisition (e.g., Burton et al. (1987)) which has examined tasks that may or may not ultimately lead to the development of an actual system.

3 Knowledge Acquisition for a Diagnosis System

As part of the construction of the system, knowledge acquisition was initiated. In that field study, a portfolio of four different approaches knowledge acquisition process were employed: Interview, Observation (watching technicians and listening to phone calls),

Phone Taps (with scripts) and Examination of Documentation (used for education of the technicians and reference by the technicians).

3.1 Interviews

The initial approach used was interviews. This took two primary forms:

- interviews with individuals and
- a meeting with a group of seven technicians of varying experience.

The interviews were unstructured. However, the focus of the interviews was on how experienced technicians identify and solve communications problems.

The interview approach led to the understanding that this was a complex diagnosis problem. Again and again, the technicians indicated that,

1. there were many different communications problems,
2. in some cases, the problems were difficult to identify over the phone and
3. that there were alternative ways to solve these different problems.

The group interview process suggested that some solution approaches were better than others. Further, more experienced technicians (and higher in the managerial hierarchy) kept indicating that there was a substantial difference in their understanding of communications problems, as compared to lower level personnel.

The interview process was employed with the understanding that the context of the firm, clients and experiences are critical (e.g., Kolodner and Riesbeck (1986)). As a result, the interviews were done on site at the firm and in the department. In addition, the interviews stressed the use of examples to illustrate discussions and facilitate further discussion of given points. Thus, interviewees would relate their experiences with particular clients to illustrate certain concepts.

3.2 Observation

Observation at the firm's location of the technicians took two primary forms.

1. The technicians were observed for their 'external' problem solving process. That is, what kinds of behavior did they use when they solved problems.
2. We were permitted to listen in on calls from clients that dealt with communications issues, in order to understand what processes were used by the technicians to identify and solve client problems.

The first approach was particularly fruitful in seeing what the technicians *physically* do when they solve diagnosis problems, in general. It also was useful in determining what knowledge they used to solve problems and the form in which that knowledge was stored. In order to assist them in the decision process the technicians would, for example, go over to a modem on-site in order to visualize the problem. Buttons would be pressed, directories would be entered and files would be run. Since the technicians were not on site, it was found that they needed to visualize what was happening at the client's operations.

The technicians also made extensive use of the documentation. In general, the most experienced analysts did not use the documentation, but they were able to refer other technicians to where in the documentation certain information could be found. Less experienced technicians typically made use of personal notes to find where they had previously found particular items in the documentation. The most inexperienced technicians found movement through the documentation difficult, since they did not understand many of the basic concepts and had no experience with much of its contents.

The second approach was less fruitful. Since communications calls occurred only seldom during the period of the study (which was purposely not during the 'rush' season), there were hours spent just waiting for a phone call to occur, only to find that it was not a communications problem after all.

Since it was unpredictable when communications calls would occur, we left tape recorders so that the technicians would be able to tape the appropriate phone calls. The tape recorders were left with multiple levels of technicians, ranging from relatively new technicians to the more experienced technicians. This was done so that alternative problem solving styles could be observed and summarized.

Strangely enough, whenever a communications call occurred either the technician 'forgot' to turn on the tape recorder or the technician could not get the tape recorder to 'work'. This led to an alternative approach to ensure that the calls would be recorded.

3.3 Phone Taps

Phone taps were made to further analyze the technician's solution of diagnosis problems. Skeleton scripts were prepared in conjunction with management, and given to another department that understood communications problems. The phone calls from the scripts were represented to the technicians as coming from legitimate clients. The scripts even went to the extent of providing legitimate client numbers so that it would appear that the calls were from clients. The technicians did not know that the calls were staged.

The same scripts were used to interrogate different technicians to determine the similarity of alternative solution generation processes. The scripts were developed to test the expected range of anticipated responses of the technicians. The scripts were developed to elicit both surface knowledge and deep knowledge about the application. Then the scripts were compared for differences used in problem solving.

The phone tap scripts had the advantage of having an external party record the calls. In addition, the use of the scripts also guaranteed the ability to have multiple responses to the same questions, from different technicians. This allowed the examination of the similarity of the technicians problem solving processes.

3.4 Examination of Documentation

The final form of knowledge acquisition had the best initial structure. It entailed a detailed examination of the existing documentation.

Since there is substantial turnover of personnel, the firm has been concerned with documenting the communications diagnosis process for educational and reference purposes. As a result, the documentation serves a purpose of providing reference material for the technicians, should they forget or not understand a particular process. Further,

changes to the communications environment are constantly occurring and thus, need to be documented. The communications process documentation included both the original manufacture's documentation on, e.g., modems and the time-sharing vendors additional documentation.

The documentation took the form of checklists, to-do lists and straight summaries of information. Unfortunately, the knowledge was highly fragmented. Information on communications appeared in a number of places in the written documentation.

Since the documentation was detailed and designed to educate the technician about communications problems, it is also useful for 'educating a computer program'—not just a human technician. Further since it was designed for educating technicians, it was sufficiently detailed so that an understanding of the documentation yielded important understanding for the particular communications problem investigated.

4 Findings

The field study scope included the development of an expert system and the corresponding necessary knowledge acquisition to solve the diagnosis problem. The knowledge acquisition effort led to the determination of at least five different types of knowledge and the finding that different knowledge acquisition methods were more effective at identifying those different kinds of knowledge. The use of a single approach would have been impossible or yielded an ineffective system. After using all four of these knowledge acquisition methodologies we identified five primary findings.

4.1 Extent of Knowledge

The diagnosis process is, in part, a knowledge intensive process. Expertise is a function of knowledge about certain key characteristics. For example, the technician must be familiar with the particular types of modems. Lack of knowledge of a particular type of modem differentiates the 'more expert' from the 'less expert'. Thus, a critical issue in building this diagnosis system was obtaining enough relevant and timely knowledge for the system. For example, in order to 'make a difference' knowledge about each of the modems used by their clients must be in the system.

The interview process was helpful in identifying the extent of knowledge necessary to solve the problem. However, because there was such extensive documentation, the documentation established clearly the extent of knowledge required.

4.2 Integrative Knowledge

Even though the documentation in its entirety provides virtually all the relevant technical facts, it is highly fragmented. That fragmentation does not inhibit the more expert users, who have developed a familiarity with the fragmentation. However, that fragmentation does have an impact on the less experienced technicians, who have difficulty finding information. The system that was developed mitigated that difficulty in the communications area by providing a knowledge base that integrated that knowledge.

The need for integrative knowledge rapidly became apparent in the investigation of the documentation. Observation showed that the less experienced technicians had notes

to facilitate the use of the documentation, while the most inexperienced had to continually consult the expert users. A reader of the documentation needed knowledge that integrated, e.g., modems and occurrences to modems, different types of modems, etc.

4.3 Data Solicitation Knowledge

One type of knowledge only used to a limited extent in the documentation was what we called 'data solicitation knowledge'. As noted above, a critical variable in the recognition of patterns of causation, is the extent to which the appropriate data has been solicited. This knowledge was referred to data solicitation knowledge. In the system discussed here that is particularly critical since the clients that call the technicians may know absolutely nothing about the communication process or equipment.

Much of the documentation was aimed at what solutions were to particular problems, rather than what questions should be asked to determine what knowledge was required. These questions were as simple as "what has happened since the last time you used the system?". This type of knowledge differentiated the 'more' expert from the 'less' expert. It is critical to assuring that the technician moves down the correct path. This knowledge primarily was identified from the scripts and from the interview process. This knowledge was manifested in the system, by the rules designed to ferret out this information.

4.4 Ordering of the Groups of Knowledge

The results from the different scripts and the interviews indicated that 'more expert' technicians and 'less expert' technicians sometimes differed in their solution approaches to different problems. One cause of such differences was their ordering of the information search process. Different orderings can result in substantially different decision trees being established. This was manifested in the system by the order of the rules.

4.5 Word and Picture-Based Knowledge

Finally, the technicians used more than just 'word-based' knowledge to solve the diagnosis problems. Observation and interview found that technicians often had to 'see' what was happening.

In order for the system to meet the behavioral needs of the technicians, this means that pictorial-based knowledge facilitates understanding of the problem. That is, pictures and the ability to do things to the pictures were found to be a critical issue in the successful representation of the knowledge for the solution of the diagnosis problems the technicians solved.

5 Summary

This paper discussed the knowledge acquisition required for the development of a system to diagnose communication problems for users of a large mainframe time sharing system. The investigation employed a field study approach in order to exploit the experiences and existing knowledge structures of the technicians. The paper finds that multiple forms of

knowledge acquisition and representation were required for the development of the system. A portfolio of approaches was required.

No one form of knowledge acquisition was found to provide insight into each of the different types of knowledge that ultimately were used in the development of the system. Thus, this study suggests that individual task comparisons of different forms of knowledge acquisition may understate or misstate the problems of interest to developers of expert systems. Individual approaches each have their own strengths and weaknesses so that in a portfolio of approaches, weaknesses of one approach can be mitigated by strengths of other approaches.

References

- BUCHANAN, B., and SHORTLIFFE, E. (1985), *Rule-based Expert Systems*, Addison-Wesley, Reading, MA.
- BURTON, A., SHADBOLT, N., HEDGECOCK, A., and RUSS, R. (1987), *A Formal Evaluation of Knowledge Elicitation Techniques for Expert Systems*, Proceedings of the First European Workshop on Knowledge Acquisition for Knowledge-Based Systems, Reading University.
- DUDA, R., GASCHNIG, J., and HART, P. (1979), Model Design in the Prospector Consultant System for Mineral Exploration, in: *Expert Systems for the Micro Electronic Age*, ed. D. Mitchie, Edinburgh University Press.
- HAYES-ROTH, F., WATERMAN, D., and LENAT, D. (1983), *Building Expert Systems*, Addison-Wesley, Reading, MA.
- KAPLAN, A. (1964), *The Conduct of Inquiry*, Chandler Publishing, San Francisco.
- KOLODNER, J., and RIESBECK, C. (1986), *Experience, Memory and Reasoning*, LEA Hillsdale, NJ.
- MESERVY, R. (1985), *Auditing Internal Controls: A Computational Model of the Review Process*, Unpublished Ph.D. Dissertation, University of Minnesota.
- WATKINS, P., and O'LEARY, D. (1987), *Knowledge Acquisition for Small Scale Expert Systems from Consulting Experts: A Field Study Approach*, Proceedings of the First European Workshop on Knowledge Acquisition for Knowledge-Based Systems, Reading University.