EXPERT SYSTEMS
A USER'S PERSPECTIVE OF SOME CURRENT TOOLS

Susan P. Ennis
Amoco Production Company, Research Center
4502 East 41st Street, P.O. Box 591
Tulsa, OK 74102

ABSTRACT

The purpose of this paper is to report on one user's experience with several of the software tools for building an expert system. To the best of the author's knowledge, this is the first time a number of different expert system building tools have been applied to a single problem by a single analyst. A similar single problem/different tool experiment was performed at the Expert Systems Workshop in 1980, but each tool was used by its own proponents on the given problem.

I THE APPLICATION

The application used for this study involved the interpretation of x-ray powder diffraction spectra of rocks to determine the constituent minerals. The system is required to recognize only about 25 minerals -- these will resolve over 90% of the sedimentary rocks. The ability to recognize more minerals and thus handle a higher percentage of sedimentary rocks and some igneous rocks will be added to the system later.

Textbooks on x-ray diffraction present an algorithm for solving this problem [3]. This algorithm utilizes the characteristic spectrum for each mineral, which describes the location and relative amplitude of each peak. A published file of characteristic spectra data is readily available [8], as are programs that apply the algorithm. The peak locations can be readily reproduced in the laboratory, but the relative intensities vary considerably from the published norms as a function of the laboratory procedures. Control of these procedures for running several hundred rock samples per week is prohibitively expensive. Expert mineralogists understand the vagaries of the data and apply the algorithm with heuristic modifications gleaned from experience. The basic expert process to be modeled is an iterative process of selecting the highest peak in the spectrum, matching that peak to a mineral (called the determination phase), and subtracting the contribution of that mineral from the spectrum (called the reduction phase). The locations of the major and minor peaks of the minerals of concern are often very close or even coincident and, in addition, minor peaks of several minerals may combine to form the current highest peak.

This task is a good candidate for an expert system:

- Mineralogists handle the analysis well,
- Spectral problems have been solved with AI techniques [1],
- The system can be verified by running extensive tests on selected samples,
- The expert was available and willing to devote his time, and
- The correct analysis of rock samples is important.

II THE CRITERIA FOR COMPARING TOOLS

The various tools are compared and contrasted on their ability to provide the purported advantages of the expert system technology. These advantages include, but are not limited to, the following:

- The knowledge statements, notably the rules, can be read and understood by non-computer-oriented domain experts,
- It is easy to modify and expand the knowledge base,
- The systems can explain their reasoning -- both during the running of the system (via a trace) and after the final conclusions have been made (explain), and
- Only knowledge needs to be entered into the system. If rules are necessary to control the questioning or rule execution, then these "control rules" can be separated from the "domain rules."

A major objective was to create a distinct sub-set of rules for the mineralogist to examine and expand. In particular, the mineralogist did not want to be concerned either with the control of the iterative process or with some of the data manipulation involved. The systems are also rated by the ease with which application-specific extensions can be made since all of them are undergoing active development and it is unlikely that a single system will ever include all of the possible functions for each application.

Because of time limitations, knowledge-base construction ceased as soon as it was shown that the system could handle the problem and give correct results; the extensions to the various tools were not hardened for production use.

III THE TOOLS

A. EXPERT

For various reasons, primarily related to hardware availability, the problem was first approached using EXPERT [7]. This effort, lasting about one month, demonstrated several things. It showed that mineral determination rules could be written and it showed that the determination-reduction cycle must be run iteratively to prevent minor peaks of several minerals from combining to give false indications of another mineral. The attempt to use EXPERT was abandoned because EXPERT does not allow parameter values to be changed once they are set and each possible spectral peak had to be named in advance. It should be noted that no attempt was made to work with the developers to extend EXPERT to handle these problems.

B. LISP

At this point a system was written in LISP to perform the task. The system is domain-specific with much of the control and very generalized knowledge embedded. However the rules to perform the determination and reduction have an English-like syntax, the run can be traced, and the system can explain its reasoning to the user. The inference method and the control and generalized knowledge functions was finished in 1.5 months. A sufficiently 'user friendly' interface was completed in another two months and the entire system was given to the mineralogist for refinement of the rule base. The mineralogist was able to process data interactively, examine the run, change the rules using a full-screen text editor, and reprocess data using the revised rules.

Two months after receiving the system, the mineralogist had made changes to the rules for most of the minerals known to the system and added rules for two additional minerals. Although he was computer-naive, he asked for help only once. He was very pleased with the capability the system gave him for changing the rules without help from a programmer.

C. UNITS

The next system tried was Units [4], a very sophisticated data base system with inheritance along a tree structure and the capability to store rules written in a procedural language. A system that could handle four minerals was constructed in one month. The rules were conveniently grouped for a mineralogist to examine -- both in interactive and hard-copy form. The control rules were well separated from the domain rules. The primitive nature of the rule language required many statements where a more complete language would have allowed the flow of the mineralogy rules to be improved. Units does provide a table data type that was used to store and manipulate the spectral data without the need to pre-define each possible peak. There is a dearth of generalized relations and functions. New data types can be defined by the Knowledge Engineer (KE) in a very structured manner. However, the necessary relations and functions for using these data types require that existing system code be modified. One of the existing functions was more than 600 lines in PrettyPrint format. Units has no explanation facilities and poor run-time tracing facilities.

Limited procedural attachment is allowed, though the system documentation is misleading on how to activate it. Units is written in INTERLISP which is upper/lower case sensitive. The KE interface does not always protect the KE from entering objects in the incorrect case. On the other hand, all of the messages generated for the user running the system are in upper case. It does offer a reasonable mechanism to separate parts of a problem, if the problem is separable.

D. EMYCIN

EMYCIN [5,6] could not handle the iteration required by the problem nor could it represent the spectral data without naming each possible peak. The developers added a single function to handle the iteration problem and a new data type (array) with its concomitant functions to handle the spectral data. A mutually cooperative debugging effort allowed the construction of a system able to handle four minerals in three weeks. Several application-specific functions were added with ease. The control rules separate nicely from the domain rules, though there is not an easy way to get a separated listing of the rules. The trace facility was adequate. Because of the heavy dependence on the array data type and the rushed implementation of that type, the Question/Answer module (explanation facility) of EMYCIN was not used in debugging the system.

The KE interface is extremely good. Although written in INTERLISP, the system handles the upper-lower case problem well. The EMYCIN system maintains a very large network of pointers. This ability of the system to cross-reference almost anything to anything else was invaluable for debugging the knowledge base.

E. OPS-5

The system was constructed under OPS-5 [2] in 1.5 weeks -- for only three minerals, but enough to show that it could work. Functions can be added in a straight-forward manner, although expansion of the documentation by a few examples would assist a new KE. Relations cannot be added and the rules can become rather complex. Cross-referencing of rules (productions in OPS terminology) and parameters is not supported. OPS offers splendid opportunities to the KE who wants to handle his own control, because the OPS control is extremely primitive. This leads to control and knowledge being interspersed in almost every production. The developers contend that the control is part of the knowledge and in a sense they are correct. However valid this assertion may be from an Artificial Intelligence viewpoint, it frustrated the attempt to separate control and domain knowledge.

It should be rather simple to write a domain-specific interface that would give the appearance of a separation of domain and control rules.

IV CONCLUSION

The x-ray problem was implemented on the tools in the order presented. The generally decreasing time to implement the system on each new tool would probably have occurred independent of order. This study did show that a knowledge base can be re-structured to run with various tools in a short time period. The tools do not, in general, provide what industry has been lead to expect. This is a shortcoming of the current implementations, rather than the basic technology. A large amount of work remains to be done before these research tools can be used effectively by industry.

The benefits to those who persevere are potentially large. A recommended approach is to select an appropriate tool and generate a prototype system. With luck the tool will be capable of supporting the full system, and with cooperation from the developers most of the tools can be expanded to support the full system. At worst the KE will have generated the specifications for a domain specific LISP-based system. Each and every industrial application -- if its conception problems are reported -- will further the development of tools suitable for industrial usage.

ACKNOWLEDGMENTS

To Amoco Production Company for providing the opportunity for me to do this study. To James C. Hoffman (Amoco) for developing the mineralogy rules. To Ellen Bailey and Jay Tompkins (Amoco) for assisting in the construction of the LISP system. To those who helped with the various tools: S. M. Weiss (Rutgers) for EXPERT, Peter Friedland (Stanford and Teknowledge) for Units, Jim Bennett and Carli Scott (Stanford and Teknowledge) for EMYCIN, and C. L. Forgy (Carnegie-Mellon) for OPS-5. To Jim Bennett and Jay Tompkins for reviewing drafts.

REFERENCES

[1] Buchanan, B. G. and E. A. Feigenbaum, "DENDRAL and Meta-DENDRAL." Artificial Intelligence 11:1 (1978) 5-24.

[2] Forgy, C. L., "OPS-5 User's Manual." CMU-CS-81-135, Dept. of Computer Science, Carnegie-Mellon University, July, 1981.

[3] Klug, H. P. and L. E. Alexander, X-ray Diffraction Procedures. New York, John Wiley & Sons, Rev. Ed., 1974.

[4] Smith, R. G. and P. Friedland, "A Users Guide to the Unit System." Heuristic Programming Project Memo HPP-80-28, Stanford University, December, 1980.

[5] van Melle, W., A domain-independent system that aids in constructing knowledge-based consultation programs. PhD dissertation, Computer Science Dept., Stanford University, June, 1980.

[6] van Melle, W., A. C. Scott, J. S. Bennett, and M. A. Peairs, "The EMYCIN Manual." Heuristic Programming Project Report HPP-81-16, Stanford University, November, 1981.

[7] Weiss, S. M. and C. A. Kulikowski, "EXPERT: A System for Developing Consultation Models." In Proc. IJCAI-79. Tokyo, Japan, August, 1979, pp. 942-947.

[8] Joint Committee on Powder Diffraction Standards, Mineral Powder Diffraction File. International Center for Diffraction Data, Swarthmore, PA.