

Knowledge acquisition techniques for visual expertise: a study of oil-reservoir evaluation

Mara Abel

Universidade Federal do
Rio Grande do Sul
Instituto de Informática
91.501-970 Porto Alegre, Brazil
marabel@inf.ufrgs.br

John A. Campbell

University College London
Dept. of Computer Science
Gower Street WC1E 6BT
London, UK
j.campbell@cs.ucl.ac.uk

Luis Fernando De Ros

Universidade Federal do
Rio Grande do Sul
Instituto de Geociências
91.501-970 Porto Alegre, Brazil
deros@inf.ufrgs.br

Content Areas: knowledge acquisition, expertise, visual knowledge, petroleum exploration

Abstract

This work presents the results of a study about the nature of expertise in geological domains that demands primarily visual recognition and only secondarily uses search and analytical methods to describe and interpret petroleum reservoir rocks. An experiment was developed using rock images and analyzing how geologists with distinct levels of expertise describe them. The study demonstrated that experts develop a wide variety of representations and hierarchies, which differ from those found in the domain literature. They also retain a very large number of symbolic abstractions for images. These abstractions (which we call visual chunks) play an important role in guiding the inference process and integrating collections of tacit knowledge of the geological experts. We propose that the knowledge acquisition process in this domain should consider that inference objects and domain objects are parts of distinct ontologies. Our domain ontology has been developed through the analysis of cases, and knowledge graphs have been used to connect inference objects with the domain ontology. We establish a set of rules to be followed in conducting knowledge acquisition projects for the development of expert systems for visual (i.e. image-based) domains.

1 Introduction

Knowledge about the evaluation of oil-reservoir rocks is crucial in petroleum exploration, since it can substantially decrease the risks of exploration, and increase the

efficiency of hydrocarbon production. Geologists apply high levels of expertise to answer questions such as: What kind of reservoir rocks should be sought? Which are the criteria and parameters to be used to evaluate the reservoir? Which methods are needed for oil field development? Capturing this knowledge and representing it in a formal model is a permanent aim for knowledge management in petroleum companies.

During the exploration for new petroleum accumulations and the development and production of oilfields, two sets of information have fundamental importance. The first refers to the external geometry of the reservoir which describes the shape and extension of the sedimentary units that contain oil, and the second refers to the internal structure and characteristics of the reservoir rocks. Although not essentially, most of the data relevant for geological interpretation of oil reservoirs consist of visual information that cannot be described only through geometric components, such as size and format. Many of the aspects recognized by a geologist during the interpretation task have no formal denomination and are learnt through an implicit process during training and field experience.

The use of features without names in supporting problem solving is a current practice in many natural domains that demand pattern matching as a part of the problem-solving process. These objects constitute the implicit body of knowledge, also called tacit knowledge by Nonaka and Takeuchi [Nonaka, Takeuchi, 1995] when referring to the unarticulated knowledge that someone applies in daily tasks, but is not able to describe in words. The articulated or explicit knowledge refers to the consciously recognized objects and how these objects are organized. This portion of knowledge in the context of Artificial Intelligence is called ontology. Tacit and explicit knowledge should be seen as two separate aspects of knowledge and

not different sorts of it, although is evidently useful to treat them in separate ways.

The dynamic conversion of tacit knowledge to explicit knowledge in the human mind allows us to capture the essentials of a situation on a reduced scale, which lends itself more easily to imaginative manipulation and mental organization than the ungainly original. It thereby makes possible the development of engineering and the socialization and consequent multiplication of knowledge in a social dimension. The collective conversion of unconscious knowledge into explicit, organized knowledge is an essential part of the capture of a domain ontology and the proposal of knowledge representations for it. The word ontology has been taken from the technical vocabulary of philosophy, where it means a systematic explanation of the world which is shared by a community. Nowadays, the best-known definition of ontology is that of Gruber, as slightly modified by Borst: [Borst, 1997] "Ontologies are defined as a formal specification of a shared conceptualization".

Extracting the ontology of a domain is one important objective of knowledge acquisition techniques; it is sometimes misunderstood to be the only one. Along with ontology, which represents the explicit part of knowledge, it is necessary to identify the tacit knowledge applied by experts, and also the dynamic mechanisms of problem-solving. In Geology, as in many other natural domains, the role of tacit knowledge in the expert problem-solving process has a dominating influence of the results achieved. Therefore, it deserves special attention in knowledge engineering. This paper describes a study in the petroleum domain: that identifies what kind of tacit knowledge is applied by geologists in the evaluation of oil reservoir, as part of the PetroGrapher project for the development of an expert system to support reservoir-rock interpretation. The study acknowledges and identifies differences between the cognitive mechanisms of novice and expert geologists during the rock interpretation process. Further, we discuss how these differences can influence the choice and application of knowledge acquisition techniques.

2 Knowledge acquisition and knowledge elicitation techniques

Knowledge acquisition amounts to the process of collection, elicitation, interpretation and formalization of the data regarding to the functioning of expertise in a particular domain. Its objective is to reduce the communication gap between the expert or knowledge worker and the knowledge engineer, allowing the knowledge to become independent of its sources. The traditional motivation for knowledge acquisition, which underlies the development of any expert system, is complemented by other objectives from software engineering – accurate capturing of user requirements - and knowledge management – collecting and sharing the information of knowledge workers within organizations.

The main classes of knowledge acquisition techniques are described briefly below.

- Interviews, observation and protocol analysis: A grouping of many different techniques that demand direct interaction with experts through retrospective or concurrent enquiry. In the retrospective interview, the expert narrates a memory of how a problem was solved. This description is commonly rationalized and omits many crucial details. In concurrent interviewing, e.g. via observation and protocol analysis, the expert verbalizes his/her reasoning *during* the problem-solving process while it is being recorded and observed. The result is more trustworthy than a direct interview, but the expert is usually unable to verbalize what he/she is doing when the inference requires both a high level of abstract reasoning and a low (i.e. concrete) level of sensorial activity. The collected information is commonly imperfect and needs to be complemented through further techniques.
- Classification techniques aim to identify the terms and concepts of the domain and how these concepts are organized in classes, groups or components, according to the expert. These include card-sorting and multidimensional scaling techniques [Shadbolt, O'hara, 1999].
- Collecting cases: A general label for all techniques that exploit recorded cases in knowledge acquisition, such as scenario analysis, event recovering, and the analysis of legacy cases for use in case-base reasoning systems.
- Extracting cause-effect relations. This includes techniques used to extract causal relations among concepts of the domain (such as evidence for conclusions, or problems and their applicable solutions). Variations on repertory grids [Gaines and Shaw, 2003], rule-extraction, knowledge graphs [Abel, Castilho, 1998, Leão and Rocha, 1990] and conceptual graphs [Sowa, 1984] belong in this class.
- Identifying the reasoning path. Problem-solving methods (PSM) [Benjamins and Fensel, 1998] and inference structures [Schreiber, Akkermans, 2000] are graphical representations of the inference process involved in problem solving, described at an abstract (though not generic) level.

The accepted knowledge acquisition techniques are effective in revealing the ontology underlined in expert reasoning, or the explicit component of knowledge. However, little progress on elucidating the unarticulated parts of expert knowledge has occurred, e.g. visual recognition or the integration of sensorial objects in the domain ontology. These issues are the main focus of the present work.

3 Investigation of the cognitive abilities of experts in petrography

Geological interpretation is strongly based on visual interpretation of features imprinted in rocks or landscape by physical phenomena. Since this process is improved mainly by practical experience, and not by supervised learning, geology is one of the areas in which experts develop themselves as strategic sources of knowledge.

The most familiar achievements in knowledge acquisition for geological domains are the development of the PROSPECTOR system [Duda, Hart, 1979] knowledge engineering; the XEOD expert system to identify detrital depositional environments [Schultz, Fang, 1988]; and the knowledge acquisition project SISYPHUS III, concerning igneous petrography [Gappa and Puppe, 1998]. The influence of the domain in the effectivity of the knowledge acquisition technique is remarkable, as was discussed by [Wagner, Chung, 2003]. The common aspect of all these projects was that they have drawn attention to the development of dealing with visual diagnostic features and also have shown the necessity of representing the objects that support inference as a separate ontology.

In the present study, our intention has been to elucidate the cognitive process and objects that support the expertise in sedimentary petrography, by using tests similar to those found occasionally in cognitive psychology, e.g. in [Ericsson and Smith, 1991].

The investigation was conducted over a group of 19 geologist with distinct levels of expertise in sedimentary petrography. The group was selected among lecturers undergraduate and graduate students of the Geoscience Institute of UFRGS and geologists from a petroleum company. Practical experience, instead of only theoretical knowledge, was a fundamental prerequisite for the selected group. The members of this group were first classified as novices, intermediates or experts. Novices were students or geologists who had received at least 100 hours of training in sedimentary petrography. Intermediates were geologists who used petrography as a daily work tool. Experts possessed at least 10 years of experience in the subject and directly utilized sedimentary petrography for more than 10 hours per week.

The group was requested to carry out 5 different tests, based on the presentation of images from rocks, using a high-definition video system attached to an optical microscope. The first experiment was designed to evaluate long-term memory. They were requested to describe fully a thin section that had been examined one hour before. The second test requested a full description of a thin section without time restriction, in order to evaluate the richness of technical vocabulary. The third test investigated short-term memory, by requiring description of a rock image shown just previously for a very short time. In the fourth experiment, recall of a first set of pictures of common objects (animals, land-

scapes and people) and a second set of sedimentary rocks was requested just after these had been shown for a short time. In the fifth experiment, the geologists were asked to divide another set of pictures into subsets (i.e., classify them) and explain their criteria for this subdivision.

The set of experiments was conceived in order to measure the association of 3 indicators with the predefined class of expertise. The indicators were:

- (1) amount of significant information in the description obtained after or during image exposition: it was expected that experts knew more about the domain and that their knowledge would be expressed through the use of technical and precise vocabulary;
- (2) intensity of use of interpreted features, instead of features having objective geometrical properties: it was expected that experts would develop image recognition in a more abstract level than novices and would be able to demonstrate this in recognizing features that needed only a short process of inference to be identified;
- (3) efficiency of organization and indexing of the domain: it was expected that experts were more effective in grouping and classifying new information related to the domain and could demonstrate this ability in experiments involving their memory.

We found no relation between the quantities of words or even of significant words (words really related to the domain) in descriptions produced by experts and by novices. On the average, experts record more efficiently the details of the rock inspected, but the overall relation is not so simple: some novices exhibit an expert level of performance and some experts perform like novices.

When considering the use of interpreted features in the description, experts clearly demonstrate a higher average, as shown in Figure 1. The vertical in this figure indicates how many clearly interpreted features were used in describing the rock. The general pattern was the same in the short-term memory experiments involving sedimentary rock samples. The faster the expert is requested to classify or interpret rocks, the more he/she will try to recognize diagnostic features. Our experimental pictures changes completely when another kind of rock was used, such as a metamorphic rock sample. The result can be seen in Figure 2, where no relation between level of expertise and interpreted features can be identified. The sample here includes a professor of metamorphic petrology among the sample, classified as a novice. As expected, the geologist has shown the expert pattern of behavior in his own domain of knowledge.

Another significant result was obtained from the experiments with the set of images. Experts have shown worst memory for common photographs than the novices, but obtain excellent results when the pictures concern their domain of expertise.

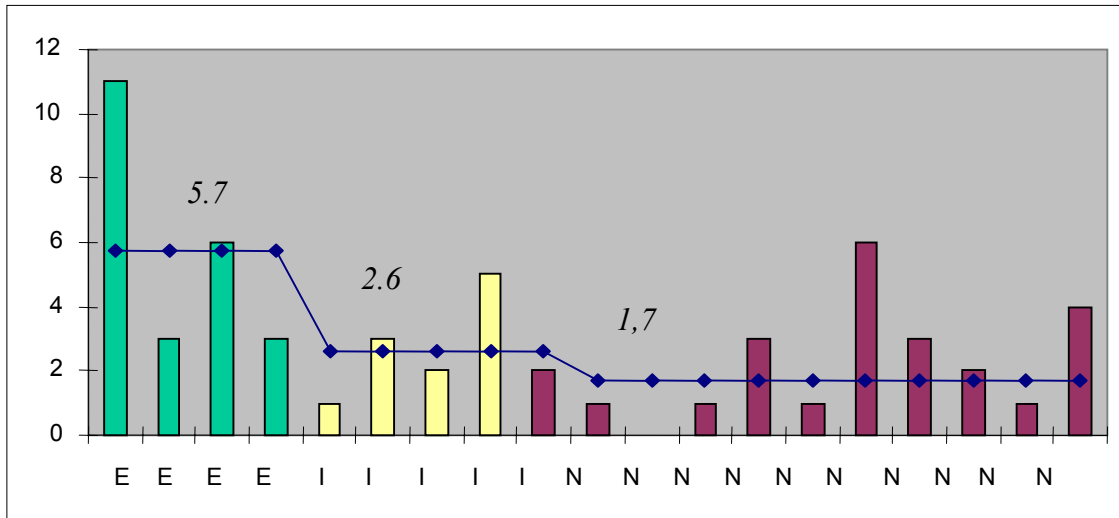


Figure 1 – Relation between the level of expertise and the usage of interpreted features in information extracted from thin-section rock images by geologists in their descriptions of rock samples. The bars indicate the number of interpreted features and the line represents the media for each class (experts (E), intermediate (I) and novices (N)).

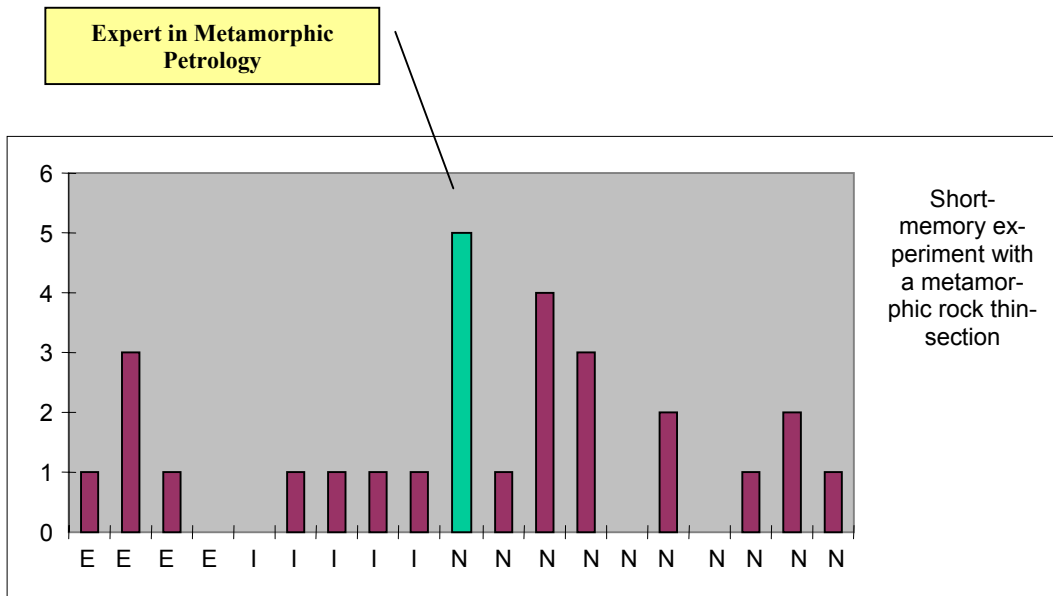


Figure 2 – The same experiment as covered in Figure 1, but using a metamorphic rock sample. The relation between expertise and amount of features are not longer perceived. The highest bar refers to a professor of metamorphic petrology, who exhibits an expert behavior in this experiment.

The evidence of the ontological support for memorization was also demonstrated when experts were requested to organize pictures of common objects and rocks. In the classification of rock pictures, the experts utilized evidently interpreted aspects of the rock, such as the quality

of porosity or the kind of cementation. The novices organized the pictures mainly in terms of color, texture or abundance of minerals. The organization of the domain in the expert's mind relies on aspects that make the problem-solving process easier or more effective, rather than the taxonomy commonly used for teaching of students.

4 Knowledge acquisition techniques for visual domains

Giving the importance of interpreted features in the problem-solving process of reservoir evaluation, it is highly desirable to establish particular methods for knowledge acquisition and representation. Although a full methodology for knowledge acquisition in visual domains is beyond the scope of this paper, we can at least outline some rules for developing expert systems for visual domains, on the basis of our own experience.

Rule 1: The ontology of the domain should be extracted from real cases (descriptions of previously-solved problems) to the extent that these are available. Avoid representing an ontology assembled from books or from direct interviews with the expert: in both of those instances the ontology may be useful for transferring and sharing knowledge, but not for real problem-solving.

In this project, we have developed a first ontology from the interviews with the expert. This initial ontology shown itself no relations with the solution process of the expert. Eventually, a second ontology was collected from cases and concurrent interviews during the problem-solving process. This ontology, which supports the *PetroGrapher* system, has 132 concepts (mainly organized in a partonomy and, secondarily, in 2 separated models: graphs and triangles) in that present version. The models are fully described in [Abel, 2001].

Rule 2: Do not believe in a simple, uniform and elegant domain model. Experts apply multiple kinds of concepts, organized in many different ways according to the needs of their current task. It is remarkable and not yet clear how they can integrate and navigate from one to another, but it is an illusion to believe in reproducing expert performance by using simple and uniform models.

This project has evidence that the experts in the domain applied distinct ontologies to solve particular parts of the problem. The cognitive objects applied for visual recognition - the visual chunks - are not the same used for communication (and that is why they can not be collected by interview). The objects applied for rock classification (triangles of composition in our project) are not clearly related with the complex relations that are used for genetic interpretation (organized as chunks and graphs in our project).

Rule 3: Treat interpretation features as another ontology and elicit them during the solution process only. Assume that the expert is not completely aware of them and does not give them names, but can use them nevertheless. Transform your elicitation session into a “giving it a name” session. Do not trust in easy names for intrinsically complex features simply because the expert can recognize them in a glance. Your eventual user cannot do the same.

Extracting the concepts of an ontology in image-based reasoning domains is a challenge since the expert applies visual objects to support inference and describe only partially these objects during communication, if he is aware of them. Novice level users are not trained to recognize expert-level visual features in their problems, as it was demonstrated in our experiment described in session 3, so can not follow the same solution path of the expert. Being aware of this limitation during the knowledge acquisition phase can extend the amount of useful objects on the ontology, simply by help the expert in externalizing them.

Rule 4 : Treat the mapping between the domain ontology and the interpretation feature ontology as a distinct process to follow after the Rule 3 step. Interpreted features indeed have relations with the concepts of the ontology. Most of the time such relation can be described in terms of aggregation of instances of the concepts. Sometimes it is necessary to add new terms to the domain ontology.

We have extended the Common KADS knowledge-model language [Schreiber, Akkermans, 2000] for this purpose by including an inference concept, - the *visual chunk* - further related to the ontological concepts. An enhancement over the *knowledge graphs* tool of Leão [Leão and Rocha, 1990] helped us through the knowledge acquisition process. Both concepts and the detailed mapping of objects during inference is described in [Silva, Abel, 2003].

Rule 5 : Choose the problem-solving method to represent the inference path according to the ontology of interpretation features, not to the domain ontology. Since the inference is carried out at the expertise level (i.e., which models the tacit knowledge of the expert), these will be the necessary objects.

Rule 6 : Take advantage of the domain ontology to define the objects of the user interface, which should be defined at the novice or intermediate level.

In the *PetroGrapher* project, we have developed 3 fully versions of the interface for quantitative evaluation of the mineral constituents of a rock sample, without getting the approval of the user about the interaction model and the interface design. The problem was mainly related with defining what is the “natural” way for the user introducing a large amount of information of different types and formats in the system. The trial proposals were defined according different well-known standard of interfaces, following the intuition that a known model would be easier to be absorbed by the user.

A further deep study of the problem has suggested that the “natural view” of the domain is represented by the ontology, then it should defined the interaction model. The immediate approval of the next interface project lead us to applied the same principals with many other specific-domain project. A model of correspondence between ontological objects described in Common KADS, and interaction objects for graphical interfaces is under development.

5 Conclusion

The study of expert skills in the geological domain has demonstrated that experts develop a wide variety of representations and hierarchies - which, moreover, differ from what is found in the domain literature. Experts also retain a very large number of symbolic abstractions of images, which we call visual chunks. The chunks hold links with the internal hierarchical arrangement of knowledge, and play an important role in guiding the inference process. These cognitive resources integrate the collection of tacit expert knowledge in geological reservoir characterization. An association of knowledge graphs and case analysis has turned out to be an effective tool in helping to make explicit and acquire the declarative knowledge and causal relations of the domain, which were not evident in elicitation sessions conducted on a traditional knowledge-acquisition basis.

The domain model, in this work, expresses knowledge over two levels: the externalization level, which describes the concepts at an intermediate (between novice and expert) stage of expertise; and the expertise level, which has the meaning given above. We have introduced the concept of visual chunk as a primitive for representation. We believe that should be included in the armory of knowledge acquisition, e.g. in the CommonKADS library.

Acknowledgements

This work is supported by the Brazilian Project Financing Agency – FINEP, with the special program CTPETRO.

References

- [Abel, 2001] Mara Abel. Estudo da perícia em petrografia sedimentar e sua importância para a engenharia de conhecimento. *Doctoral Thesis*, in Programa de PG Computação, UFRGS: Porto Alegre, Brazil, 2001.
- [Abel, Castilho, 1998] Mara Abel, José M. V. Castilho, and John Campbell. Analysis of expertise for implementing geological expert systems. in *World Conference in Expert Systems*. Mexico City: Cognizant Communication Offices, 1998.
- [Benjamins and Fensel, 1998] Richard Benjamins and Dieter Fensel. Editorial: problem-solving methods. *International Journal of Human-Computer Studies*, 49(4): p. 305-313, 1998.
- [Borst, 1997] W. N. Borst. *Construction of engineering ontologies*, in University of Twente, University of Twente: Enschede 1997.
- [Duda, Hart, 1979] R.O. Duda, P.E. Hart, P. Barret, J. Gaschnig, K. Konolige, R. Reboh, and J. Slocum. *Development of the PROSPECTOR consultation system for mineral exploration.*, Stanford Research Institute International: Menlo Park, 1979.
- [Ericsson and Smith, 1991] K. Anders Ericsson and Jacqui Smith. *Toward a general theory of expertise: prospects and limits*. New York: Cambridge University Press, 1991.
- [Gaines and Shaw, 2003] Brian R. Gaines and Mildred I.G. Shaw. *Personal Construct Psychology and the Cognitive Revolution*, Knowledge Science Institute, University of Calgary, 2003.
- [Gappa and Puppe, 1998] Ute Gappa and Frank Puppe. A study of knowledge acquisition - experiences from the SISYPHUS III experiment for rock classification. In *Workshop on Knowledge Acquisition, Modeling and Management*. Alberta, Canada: Voyager Inn 1998.
- [Leão and Rocha, 1990] Beatriz F. Leão and A. F. Rocha. Proposed methodology for knowledge acquisition: a study on congenital heart disease diagnosis. *Methods of Information in Medicine*, (29): p. 30-40, 1990.
- [Nonaka, Takeuchi, 1995] Ikujiro Nonaka, Hiro Takeuchi, and Hirotaka Takeuchi. *The knowledge-creating company: how Japanese companies create the dynamics of innovation*, New York: Oxford University Press, 1995.
- [Schreiber, Akkermans, 2000] Guus Schreiber, Hans Akkermans, Anjo Anjewierden, Robert de Hoog, Nigel Shadbolt, Walter van de Velde, and Bob Wielinga. *Knowledge engineering and management the CommonKADS methodology*. Cambridge: The MIT Press 2000.
- [Schultz, Fang, 1988] A. W. Schultz, J. H. Fang, M. R. Burston, H. C. Chen, and S. Reynolds. *XEOD: an expert system for determining clastic depositional environments*, in *Geobyte*: [S.I.], 22-32. 1988.
- [Shadbolt, O'hara, 1999] Nigel Shadbolt, Kieron O'hara, and Louise Crow. The experimental evaluation of knowledge acquisition techniques and methods: history, problems and new directions. *International Journal of Human-Computer Studies*, 51: p. 729-755, 1999.
- [Silva, Abel, 2003] Luís A. L. Silva, Mara Abel, Luiz F. De Ros, John A. Campbell, and Clesio S. Santos. An image-based reasoning model for rock interpretation. in *Proceedings of IJCAI-99 Workshop in Intelligent Computing in the Petroleum Industry ICIP2003*. Acapulco, México, August 2003.
- [Sowa, 1984] J.F. Sowa. *Conceptual structures: information processing in mind and machine*, Reading: Addison Wesley 1984.
- [Wagner, Chung, 2003] William P. Wagner, Q.B. Chung, and M.K. Najdawi. The impact of problem domains and knowledge acquisition techniques: A content analysis of P/OM expert system case studies. *Expert Systems with Applications*, 24:79-86 January 2003.