The use of learning systems as a tool for the elicitation of Knowledge

Dick J. Bierman

Psych. Lab., University of Amsterdam
Weesperplein 8, 1018 XA Amsterdam
The Netherlands
UUCP: .. mcvax!uvapsy!bierman

Subfield: Knowledge Acquisition and Learning, Knowledge Representation

Keywords: Knowledge Elicitation methods, Categorization , Discrimination.

ABSTRACT

A methodology for the elicitation of global and in-depth knowledge which is based upon the use of old cases presented to learning systems is described. The basic idea is that knowledge that has been used by an Expert in the past is embedded in the database of old (classification) problems that he solved. It has been shown that some of this knowledge can be extracted from this database by presenting the old data in an appropriate format to a learning system which constructs a decision tree. However, human experts which are confronted with these trees, often can identify little familiar material. This is a major obstacle for the use of these trees as a framework in the Knowledge elicitation procedure.
In the present study 2 learning systems are used. Apart from a system that builds a decision tree, a system is used which is able to learn 'categories' from examples. This latter system is an implementation of Rosch's model of Categorization. Here, the representation of the accumulated knowledge takes the form of a 'prototype'. It is hypothesized that, since this representation is supposed to be a valid psychological one, it will be usable by the knowledge engineer as a framework in the Knowledge elicitation procedure. In the proposed procedure special emphasis is given to (old) cases for which the classification by the trained system differs from the classification by the human Expert. These 'pathological' cases are presented again to the human expert, which has to solve them while thinking aloud. The protocol is then analyzed in terms of the 'prototype' and the trace of the pathological case in the decision tree.

1. Introduction

It is generally acknowledged that the elicitation of knowledge which is needed to drive expert systems is a 'bottleneck problem' (Feigenbaum, 1981). The traditional interview approach is said to yield only a few rules per man-day (Quinlan,1986). Although this figure applies to the In-depth knowledge which is elicited near the end of the procedure, this was reason enough to stimulate the research in machine learning methods as a means of explicating knowledge (Michie, 1983). On the other hand, methodologies (eg. the KADS-methodology, Wielinga & Breuker, 1984) have been proposed to structure the knowledge elicitation process in order to make it more transparent and thereby more efficient. In those methodologies emphasis is given to the necessity of a well specified framework for interpretation of the verbal material, be it interviews with, or thinking aloud protocols produced by, the expert. The analysis generally focusses on those parts of the protocols that appear to be uninterpretable, where the expert seems to 'jump' to conclusions. In practice each protocol has this type of uninterpretable 'black holes'. But in the end very often these instances have a trivial explanation and no new knowledge emerges from the analysis. In the present study the structured knowledge elicitation procedure (model driven protocol analysis) is combined with the use of learning systems. The learning systems are used to produce information wich enables the identification of those cases wich are suitable to elicit new knowledge thus avoiding a lot of time consuming analysis of uninteresting cases.

2.General Procedure

The elicitation procedure consists of three major parts:
1: Learn
2: Pathology-detection
3: Confrontation
In the first phase a training set of old cases is selected to be presented to a learning system. After the training the systems are able to classify other cases from the old databases. The trained system has become a (first order) model of the Expert.

In the second phase the old database is presented to the 'trained' system for classification. If the classification by the system differs (considerably) from that made in the past by the Expert we call this a 'pathological case'.

In the third phase the Expert is confronted with the set of pathologies. The knowledge Engineer might directly ask the Expert why he deviated from the model or give him the cases to solve again while thinking aloud. Analysis of the thinking aloud protocol should occur in terms of deviations from the model and thus produce additions to the knowledge base.

A schematic comparison between the traditional methodology and the proposed methodology is given in fig.1.

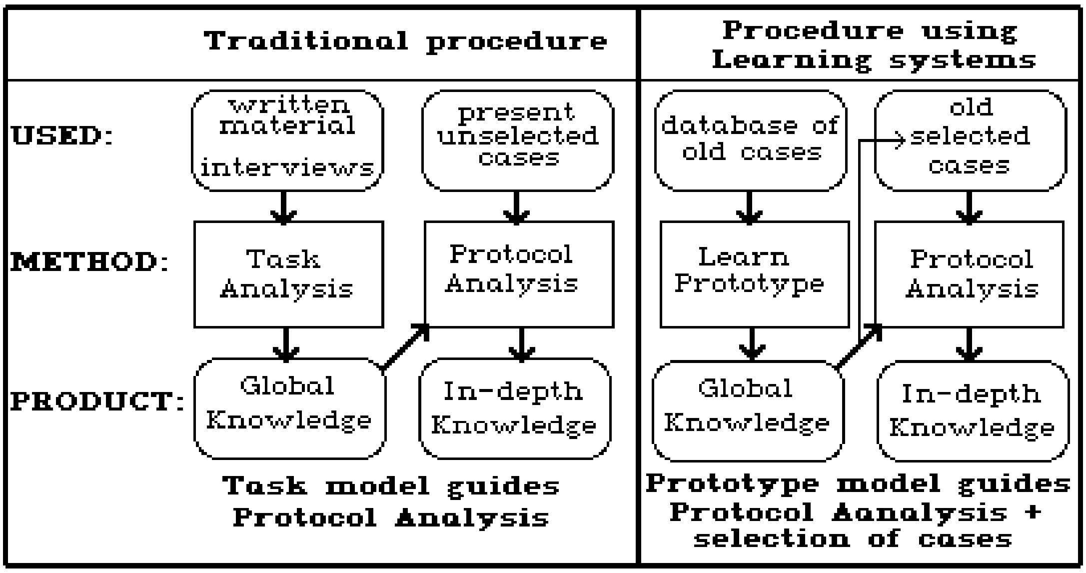

Figure. 1

3.The concrete study

To test the feasibility of this general approach a specific study was started in 1986 and will be completed in 1987.

3.1 Task domain

The task which we choose for this study is that of a selection Psychologist who has to classify potential candidates for a specific job into a few classes like 'very well suitable' and 'not suitable'. It has been shown that this type of task is characteristic for a large set of procedural tasks. Even activities such as robot planning can be recast as a classification problem (Dechter & Michie, 1985). In this task the 'objects' to classify have a number of attributes such as scores on Personality and Intelligence Tests. On the basis of these attributes the 'object' is classified by the Expert selection Psychologist.

3.2 The learning systems

Previous work which tried to apply learning systems in the process of knowledge acquisition used systems like ACLS, which construct a decision tree from examples. However it was found that, although the resulting decision trees were able to classify new cases properly, these trees, which represent the knowledge of the human expert, very often were hardly recognized by the same expert. This representation offered therefore not a fruitful framework for the Knowledge Engineer to base his further interviews on. This situation is not very different from a representation by linear regression models which have shown to have considerable predictive power (eg. Schmidt & Hunter, 1981). However the linear regression formula does not make a lot of sense to the human expert. Therefore it was decided not only to use a ACLS type of learning system but also to use a learning system that is supposed to create a psychological valid representation of the human expert's knowledge.

3.2.1 The Prototype-model

The 'prototype' model has been developed by Rosch (Rosch, 1978). The 'protoype' model allows for non-monotonic relations between the values of the attributes and the class-determination. During the learn-phase a training-set of old cases consisting of the values of the attributes and the resulting classification are offered to the system. The system learns which attributes contribute to which degree to the final classification decision (Bierman & Akkerman, 1986). After the learning phase new cases can be offered to the system which will calculate a overlap-scoree of the new instance with the 'protoype' of a class.

3.2.2 The decision-tree model

We used a system similar to ACLS. The choise of the root-attribute (and subsequent layers) which is one of the crucial operations in this type of models, was done on the basis of maximizing informational contents (through a chi-2 test for stochastic independence). If all attributes have the same number of potential values, this is an appropriate choise (Kononkenko et al, 1984). We did not use any of the mechanism which are thought to prevent noise from entering the data. This was done on purpose since one of our goals is to detect 'pathological' cases. The distinction between noise and a pathological cases can not be made by any of these noise reduction mechanisms.

3.3 Preliminary results

For our dataset of old cases we use 87 classifications made by an experienced selection psychologist in the period '81-'85. Each case consists of 28 attribute values and the subjective classification in one of 5 classes. The number of old cases is rather small because we restricted ourselves to a single Psychologist and because it concerned one specific type of job. For the training set for the 'prototypical' model we used only 6 cases which were classified as 'very suitable candidate'. The resulting 'prototype' was compared with 17 other cases from the dataset. This yielding values of overlap with the 'very suitable prototype'. These values correlated satisfactorily with the classifications made by the selection Psychologist (Bierman & Akkerman, 1986). In one case an extreme difference of opinion between the classification by the system and the classification by the Psychologist was found. We will use this case in the third 'confrontation' phase of the Knowledge elicitation procedure (see **2**.).

The training set of the ACLS program consisted of 50 cases including the set used for the 'prototype' model. A decision tree was constructed yielding generally a classification within 5 attributes. This indicates that a lot of the attributes in the majority of cases is not used. This holds especially for classification in the group of 'not suitable' candidates. The singular case which was detected by the comparison of the Psychologist classification and that made by the 'prototype' model showed a significant long branch in the decisison tree, indicating that what otherwise could have been considered to be noise reflects, in fact, in-depth knowledge.

4. Conclusions

The ultimate test for this approach will occur if the 'pathological' cases are discussed with the Psychologist. This should result in the efficient detection of in-depth knowledge thus increasing the rule-dicovery rate significantly.
So far it appears that both learning systems yield useful relationships present in the data. It should be remarked that the 'prototype' model is not able to capture any interaction between attributes and also, compared with the decision tree, it yields less information since, although 'pathological' cases can be detected, it does not tell what the pathology is in terms of attributes. The decisision tree model does allow for interactions and also gives a full account of the pecularity of a 'pathological' case. This difference in informational contents is reflected in the number of cases that are necessary to reach a stable model. Thus the 'prototype' model becomes stable after 10 cases or so while it takes much longer for the 'decision tree' model.

References

Bierman & Akkerman (1986). The development of Knowledge Based Psycho Diagnostic Systems. In: *Proc. of the 2nd Workshop on Computers in Psychology*. May 1986, Nijmegen, The Netherlands. Swets & Zeitlinger (Lisse), In Press.

Dechter, R. & Michie, D. (1985). Sructured induction of plans and programs (Technical Report). IBM Scientific Center, Los Angeles, CA.

Feigenbaum, E.A. (1981). Expert systems in the 1980s. In A.Bond (Ed.), *State of the art report on machine intelligenc*. Maidenhead: Pergamon-Infotech.

Kononkenko, I., Bratko, I., & Roskar, E. (1984). Experiments in automated learning of medical diagnostic rules (Technical Report). Jozef Stephan Institute, Ljubljana, Yugoslavia.

Michie, D. (1983). Inductive rule generation in the context of the Fifth Generation. *Proc. of the 2nd Int. Machine learning workshop*. Univ. of Illinois at Urbana-Champaign.

Quinlan, J.R. (1986). Induction of Decision Trees. *Macine Learning 1:* p.81-106.

Rosch, E. (1978). Principles of Categorization. In: E.Rosch & B.B. Lloyd (Eds.): *Cognition and Categorization,* Hillsdale N.J.: Lawrence Erlbaum, 1978.

Schmidt, F.L., Hunter, J.E. (1981). Employment Testing, old theories and new research findings. *American Psychologist, Vol.36-10,*p.1128-1137.

Wielinga,B.J. & Breuker, J. (1984). Interpretation of verbal data for knowledge acquisition. *Proc. of ECAI-84,* 1984, p.41-50.