METHODS FOR KNOWLEDGE ACQUISITION AND REFINEMENT IN SECOND GENERATION EXPERT SYSTEMS

Nada Lavrac
Jozef Stefan Institute, Computer Science Department
Jamova 39, 61000 Ljubljana, Yugoslavia

Igor Mozetic
George Mason University
Department of Computer Science and Center for Artificial Intelligence
4400 University Drive, Fairfax, VA 22030

Abstract

First generation expert systems rely on the use of surface knowledge, such as associational or heuristic. Second generation technology is characterized by two additional features: *deep knowledge* and *machine learning*. Three second generation methods for knowledge acquisition are reviewed: learning rules from examples, model-based rule learning, and semi-automatic model acquisition. The man-machine process of acquiring and refining knowledge extends the role of expert systems to *expert support systems*, since both man and machine learn through repeated knowledge refinement cycles. Explanation of solutions and of the knowledge base itself is crucial for this man-machine learning process. An extended expert system shell schema is presented that includes a knowledge acquisition and a knowledge explanation module.

1 Introduction

First generation expert systems rely on *surface*, or shallow, operational knowledge acquired through a process of direct articulation. On the other hand, the trend of the second generation expert systems is to include *deep* knowledge as well, thus capturing the underlying causal structure of the problem domain (Steels 1985), and to at least partially automatize the knowledge acquisition process by using machine learning techniques (Michie *et al.* 1984). We review three methods for knowledge acquisition in expert systems: learning rules from examples, model-based rule learning, and qualitative model acquisition. All the methods were used in the development of a medical expert system KARDIO for the ECG diagnosis of cardiac arrhythmias (Bratko, Mozetic & Lavrac 1988, 1989). The reviewed methods can be considered as second generation knowledge acquisition methods, and could be incorporated into a knowledge acquisition module of a second generation expert system shell.

The process of acquiring knowledge typically requires several 'refinement cycles' in which the machine's performance is compared with the human expert's performance. Also, the machine representation of knowledge is compared with the original expert's knowledge. In the refinement cycles, both man and machine will learn, provided that the machine's representation of knowledge is compact and comprehensible. This man-machine learning process that consists of acquiring and refining knowledge, extends the role of expert systems to *expert support systems* (Luconi, Malone & Scott Morton 1986). Explanation of the solutions and of the knowledge base itself is crucial for this man-machine learning process. Bohanec, Rajkovic & Lavrac (1988) introduce an extended schema of an expert system shell that includes a knowledge acquisition and a knowledge explanation module. The task of the knowledge acquisition module is to support user-friendly encoding of knowledge as required by one of the reviewed methods. The task of the knowledge explanation module is not just to display knowledge in an understandable form, but also to represent knowledge from different viewpoints and at different levels of detail.

2 Second generation methods for knowledge acquisition

Knowledge acquisition is typically a demanding mental process, where a knowledge engineer collaborates with domain experts. In this process the knowledge engineer's objective is to convert human know-how into 'say-how' through a process of articulation. *Direct encoding* of rules, semantic nets, frames, etc., encounters 'the bottleneck problem of applied artificial intelligence' (Feigenbaum 1977) and is named the 'old style knowledge engineer's route map' by Michie (1986). In Figure 1, which shows the flow of knowledge in different knowledge acquisition paradigms, this 'route map' corresponds to the knowledge flow through box A. Recently, there have been a number of interview techniques (Welbank 1983, Hart 1986) and software tools developed, e.g., AQUINAS (Boose & Bradshaw 1987), KADS (Breuker & Wielinga 1987), MORE (Kahn, Nowlan & McDermott 1985). For the most part, they use a conceptual model to interact with the user thus hiding the complexity and unfamiliarity of the model (rules, nets, ets.) upon which the knowledge base is actually constructed.

This research was supported in part by the Slovenian Research Council, the COST 13 project Machine Learning and Pattern Recognition, the Defense Advanced Research Project Agency under Grant No. N00014-87-K-0874, and the Office of Naval Research under Grant No. N00014-88-K-0226.

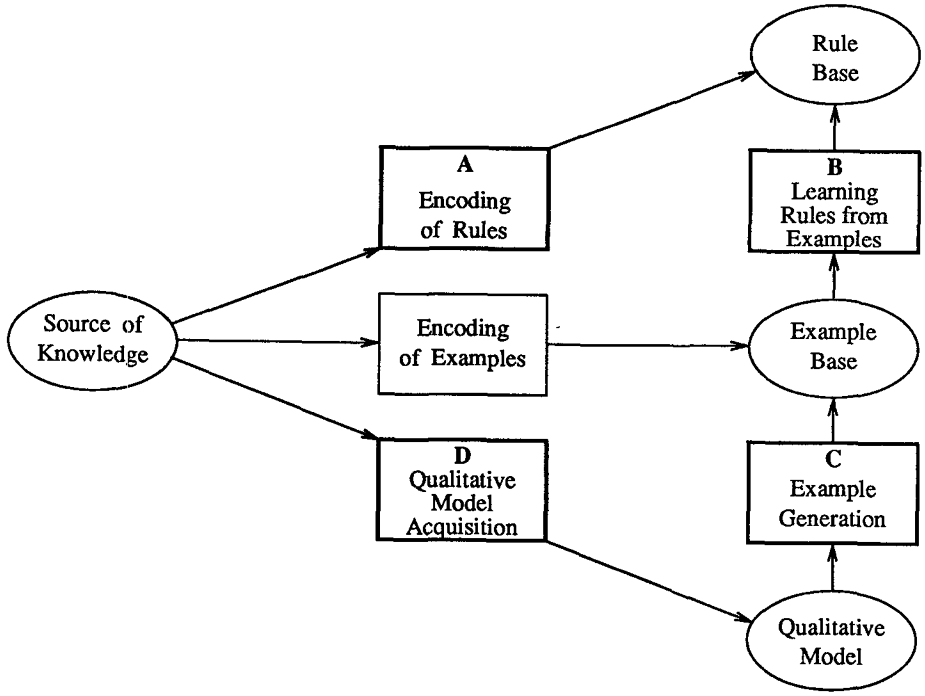

Figure 1. Knowledge flow in different methods for knowledge acquisition.

2.1 Learning rules from examples

Second generation knowledge acquisition methods are characterized by two features: use of deep knowledge, and application of machine learning techniques. The method of *learning rules from examples* is recognized by Michie (1986) as a 'new style knowledge engineer's route map' where rules are elicited from the expert to the machine memory via the language of examples rather than via the language of explicit articulation. Effective algorithms for inductive inference are required. There are a number of inductive learning programs, such as the programs of the TDIDT family (top-down induction of decision trees, e.g., Quinlan 1986), or the AQ family (Michalski *et al.* 1986) that accept tutorial examples and induce knowledge in the form of decision trees or rules, respectively. In Figure 1 this process is represented by box B where the source of knowledge is either an expert, formulating a series of thoroughly chosen examples, or preferably an existing database of examples interpreted and 'cleaned' with the help of an expert. Results of applying machine learning systems, i.e., ASSISTANT (Bratko & Kononenko 1987) and AQ15 (Michalski *et al.* 1986), on several real-life medical domains show that the systems' predictive accuracy is at the level of the best domain experts (Table 1).

Medical domain	ASSISTANT	AQ15	Medical specialists	
Lymphography	77%	82%	85%	estimate
Breast cancer	72%	68%	64%	5
Primary tumor	46%	41%	42%	4

Table 1. Diagnostic accuracy of the learning programs ASSISTANT and AQ15, averaged over 4 experiments, as compared to medical specialists. The last column denotes the number of specialists tested.

2.2 Model-based rule learning

Second generation knowledge acquisition methods are also concerned with the elicitation of *deep* knowledge that captures the underlying causal structure of the problem domain. Such knowledge can be represented in the form of a model that states the 'first principles' or basic 'rules of the game' from which operational decisions can be derived. The prevailing type of knowledge in such a model is *qualitative* (de Kleer & Brown 1984; Forbus 1984; Kuipers 1986; Bratko, Mozetic & Lavrac 1989). This has several advantages over conventional numerical modeling: the qualitative view is often closer to reasoning about the physical or physiological processes being modeled; to execute the model we do not have to know the exact numerical values of the parameters in the model; a qualitative simulation may be computationally less complex than numerical simulation. A qualitative simulation can be used for the explanation of the mechanisms of a system which is being modeled more naturally than in numerical modeling.

In principle, a qualitative model can be used for problem solving directly. It can answer prediction, diagnostic, and control type of questions. The *prediction* task is to find the observable results of applying some input to the system, given a functional state of the system. The *diagnostic* task is: given the inputs to the system and the observable manifestations, find the system's functional state (normal or faulty, which components are failed). The *control* task is to determine the input control to the system, assuming its state, in order to achieve a desired output.

However, a model is primarily designed for simulation and prediction. Using it to solve diagnostic and control tasks might be computationally expensive. On the other hand, by

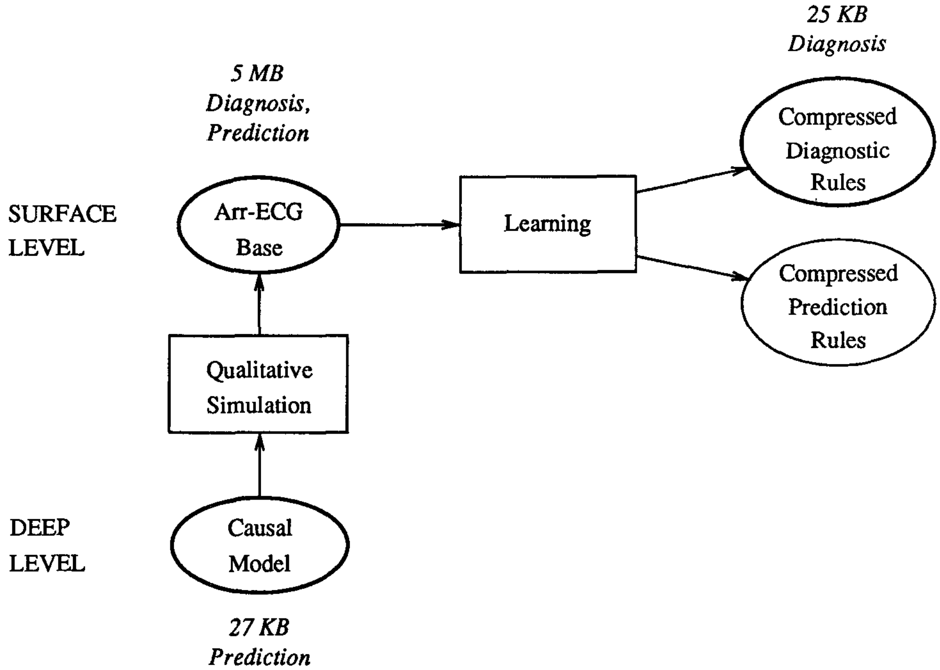

Figure 2. Deep and surface levels of cardiological knowledge, and transformations between representations.

exhaustive qualitative simulation, the model can be used to automatically generate examples of any possible behavior. From such an exhaustive set of examples operational decision rules can be generated by inductive learning methods. So, what we call *model-based rule learning* consists of two steps: example generation using a qualitative model (box C in Figure 1), and learning rules from the automatically generated examples (box B in Figure 1). This knowledge acquisition paradigm has been used in KARDIO to generate compressed diagnostic and prediction rules as shown in Figure 2. We have developed a deep qualitative model of the electrical activity of the heart, and have used it for the automatic synthesis (through simulation) of the surface knowledge about the ECG interpretation. This has the form of pairs (Arrhythmia, ECG description) relating one of the 2,419 possible combined arrhythmias to the corresponding ECG patterns (there are altogether 140,966 ECG patterns). The surface representation facilitates fast ECG diagnosis, but is rather complex in terms of memory space (over 5 MB, stored as text file). This motivated the compression of the surface knowledge by means of an inductive learning program of the AQ family into a compact and diagnostically efficient representation (Mozetic 1986). In Figure 2 the representations are arranged as to emphasize the distinction between the deep and surface levels of knowledge.

In KARDIO, we demonstrated how the qualitative modeling approach and the machine learning technology can be used to construct knowledge bases with complexity far beyond the capability of traditional dialogue-based techniques for knowledge acquisition. The KARDIO knowledge acquisition paradigm, described in detail in (Bratko, Mozetic & Lavrac 1989) may become a standard technique in the development of practical expert systems. It has already been used in the development of a satellite power supply fault diagnosis system by Pearce (1988).

2.3 Qualitative model acquisition

The model design process can be at least partially automated by means of machine learning. A Qualitative Model Acquisition System (QuMAS), described in (Mozetic 1987a, b; Bratko, Mozetic & Lavrac 1989), supports the construction of a deep model, and the representation of a model at different levels of detail. In QuMAS, partial knowledge about the model and examples of its behavior are provided by the user, and the complete model is automatically constructed and incrementally refined until the desired behavior is achieved. This knowledge acquisition paradigm is represented by box D in Figure 1.

In this approach, we restrict ourselves to functional qualitative models, where a model is defined by its structure (a set of components and their connections) and functions of the individual components. QuMAS consists of three subsystems: a *learner* that hypothesizes functions of components from examples of their behavior, an *interpreter* that can use the hypothesized model to derive its behavior, and a *debugger* that locates faulty functions of components and proposes how to correct them. It is assumed that partial knowledge about the model is given - its structure. Further, examples of the behavior of the model and its constituent components are provided from which the learning part of the system hypothesizes functions of the components. The interpreter of the model is then able to derive its behavior. The user may test the model and compare it with the intended behavior. When a difference between the derived and the intended behavior of the model occurs, a debugger is invoked. The debugger locates faulty hypotheses defining functions of components, proposes examples of behavior that guarantee the intended behavior of the model, and invokes the learner that incrementally refines the hypotheses. The cycle of deriving the behavior of the model, debugging the model, and incremental learning is repeated

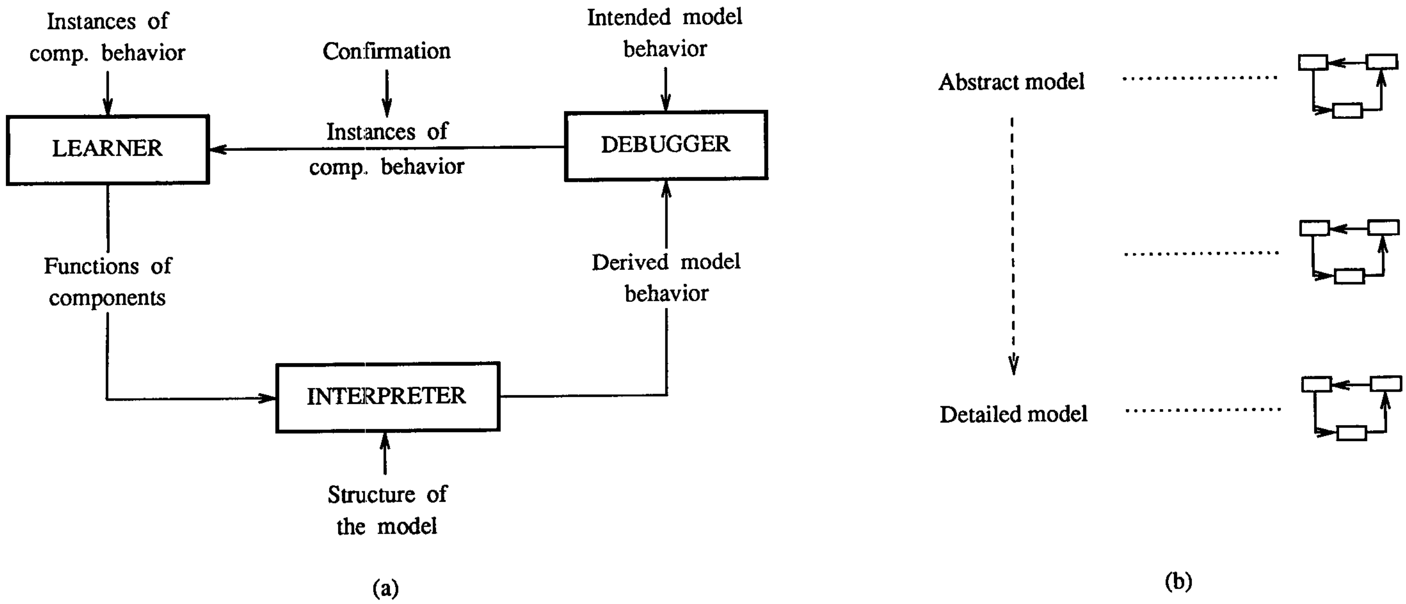

Figure 3. An overview of the qualitative model acquisition system (a), and the top-down model construction method (b).

until the intended behavior of the model is achieved, i.e., the user believes that the model is correct and complete with respect to the actual system being modeled. An overview of QuMAS is in Figure 3a.

QuMAS embodies two types of learning. Initial data-driven learning generalizes examples of components' behavior into rules on the basis of similarities and differences and does not require any user interaction. The second type of learning is model-driven where the debugger actively constructs examples of components' behavior which satisfy the intended model behavior, and then queries the user for confirmation. QuMAS therefore offers a tradeoff between the initial amount of knowledge provided by the user, and the time one is willing to spend on debugging the model. QuMAS is used interactively by the model designer, and takes advantage of the hierarchical model representation to speed up the automatic learning of the model (Figure 3b). The hierarchy also has a role in generating good and concise explanations on points selected by the user. A substantial submodel of the KARDIO heart model was reconstructed semi-automatically using QuMAS (Mozetic 1987a, b; Bratko, Mozetic & Lavrac 1989).

3 The role of refinement cycles in knowledge acquisition

Development of an expert system is essentially an iterative process that typically needs several *refinement cycles*. In each cycle an expert and a knowledge engineer refine the knowledge base by comparing the human performance with the machine performance, and the original human knowledge with the generated machine representation. The refinement cycles need to be repeated until the intended performance of the system is achieved.

Expert system shells have to integrate a variety of tools that allow for acquisition, explanation and utilization of complex domain knowledge. The classical expert system schema (Figure 4a) cannot cover all the required functions of the system. In this schema, an expert system consists of a domain dependent knowledge base, and of a domain independent expert system shell which incorporates an inference engine and a user interface. In Figure 4b we propose a new schema of an expert system shell. Here the structure of the shell is extended to incorporate: the *knowledge acquisition module* which provides tools for acquiring and editing the knowledge base, the *knowledge explanation module* allowing for different representations of the knowledge base, and the *knowledge utilization module* which applies the knowledge base to find solutions to a problem. It also has to provide explanations of particular solutions, and enable the knowledge base validation.

The new schema covers all functional requirements of expert systems as recognized by Gaines (1987), namely, apply, explain, acquire, display, edit, and validate the knowledge base. According to our view, an expert system shell is not aimed just to support problem solving, but should also actively support knowledge acquisition and refinement of the acquired knowledge. Therefore two new modules are introduced: the knowledge acquisition and the knowledge explanation module.

The knowledge acquisition module is to support one or more methods for knowledge acquisition described in the previous section. It has to incorporate a machine learning system and/or a system that supports direct or semi-automatic construction of qualitative models. The knowledge explanation module is aimed at generating different explanations - from displaying the knowledge base in a compact and comprehensible form, to representing the knowledge from different viewpoints and at different levels of detail. This may be achieved by varying the syntax and the semantic of the representation and by varying the level of detail (Bohanec, Rajkovic & Lavrac 1988). By a 'different syntax' we mean the representation of the same knowledge in a different

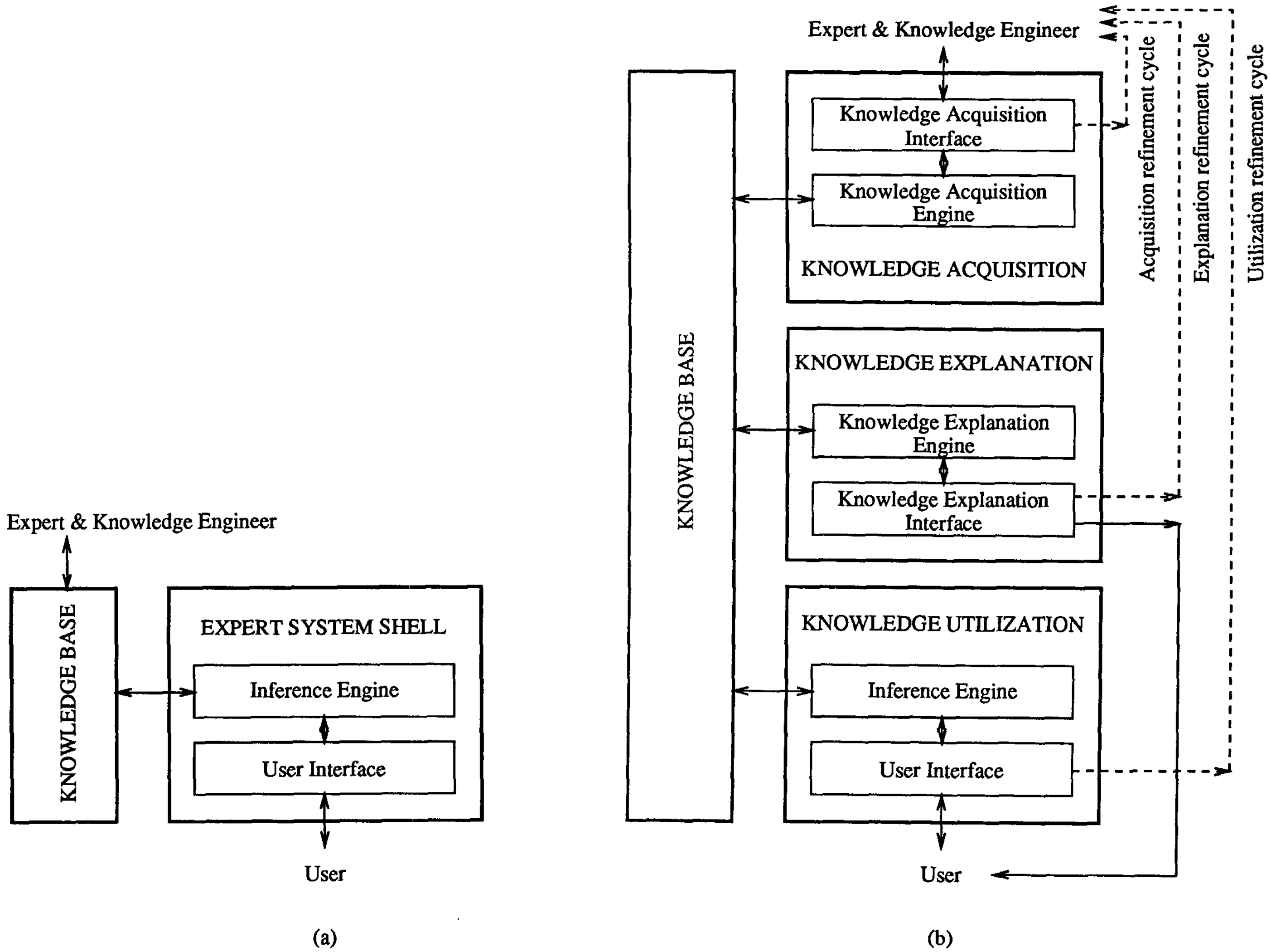

Figure 4. A classical (a) and an extended schema (b) of an expert system shell.

language, e.g., a graphical or tabular representation. By a 'different semantic' we understand either the reorganization of knowledge (e.g., grouping rules, or expressing one set of rules with another), or the representation of additional information derived from the original knowledge base (e.g., different statistics, Bayesian probabilities, informativity of attributes, etc.). By a 'different level of detail' we mean knowledge representation by using attributes at a chosen level of the taxonomic hierarchy, or by eliminating too specific knowledge, e.g., presenting only the most important rules. Some of these features are provided in the machine learning systems ASSISTANT and AQ15, mentioned in the previous section.

According to our expert system shell schema, we distinguish between the *acquisition*, *utilization* and *explanation refinement cycle* (represented by dashed lines in Figure 4b). These refinement cycles have to be supported by appropriate development tools.

Depending on the type of the system's knowledge base, the *acquisition refinement cycle* in Figure 4b consists of a feedback loop from any type of acquired knowledge back to the source of knowledge (see Figure 1). In the case when a knowledge base is induced from a set of tutorial examples this corresponds to Michie's 'first refinement cycle' of his 'new style knowledge engineer's route map' (Michie 1986). His 'second refinement cycle' corresponds to what we call here the *utilization refinement cycle* in which the system's performance is compared with the human expert's performance.

The knowledge explanation module is aimed to provide different representations of the knowledge base. This motivates the expert and the knowledge engineer to further check and elaborate the knowledge base as new ideas are triggered that may lead to finding inadequate or missing knowledge. In the process of knowledge acquisition we call this the *explanation refinement cycle.* In this cycle, the generated machine representation of knowledge is compared with the original expert's knowledge. Differences between the two can result from either an error in the machine representation, or a slip in the original human codification of knowledge. In the former case, the error is a consequence of incorrectly encoded rules, errors in learning examples, or an error in the deep causal model, depending on the type of the system's knowledge base. In the latter case, the refinement cycles may expose blemishes in the existing, original expert formulations, and may thus help to improve them.

In the man-machine dialogue through several refinement cycles both man and machine learn (Chambers & Michie

1969). This enables considering the process of knowledge elicitation as a man-machine learning process consisting of the stepwise acquisition and refinement of knowledge. The level of human and machine knowledge grows. From this point of view, the role of an expert system is not only in solving problems and explaining the solutions, but also in building (new) human-type knowledge through refinement cycles. By acquiring this refined knowledge the machine knowledge is improved and allows for better performance. This man-machine learning process that consists of acquiring and refining knowledge, extends the role of expert systems to *expert support systems* (Luconi, Malone & Scott Morton 1986). In contrast to expert systems, where main goal is to simulate experts' performance in problem solving, expert support systems stimulate human mental processes in the process of knowledge acquisition, and therefore supporting human learning as well.

4 Conclusion

The development of an expert system is typically an iterative process. It consists of refinement cycles that are repeated until the intended performance of the system is achieved. In the utilization refinement cycle, a machine's performance is compared with the human expert's performance. In the acquisition and in the explanation refinement cycle, the generated machine representation of knowledge is compared with the original expert's knowledge. The three refinement cycles are named according to the three modules proposed in the extended expert system shell schema. We emphasize the importance of the knowledge acquisition, and the knowledge explanation cycle. They allow seeing the development of an expert system as a process in which both man and machine learn, and in which (new) human-type knowledge is generated. The reviewed methods for knowledge acquisition and refinement, all actually used in KARDIO to model a substantial real-life problem, emphasize the role of machine learning techniques in the development of second generation expert systems.

Acknowledgements

All the described methods for knowledge acquisition were used (and two of them developed) when building KARDIO, a medical expert system for the ECG diagnosis of cardiac arrhythmias, in collaboration with Ivan Bratko. The extended expert system shell schema was proposed in collaboration with Marko Bohanec and Vladislav Rajkovic. We are grateful to Donald Michie for his valuable comments on an earlier draft of the paper.

References

Bohanec, M., Rajkovic, V., & Lavrac, N. (1988) Knowledge explanation in expert systems: A decision support system and machine learning view. *Proc. IASTED Intl. Conference on Expert Systems*, Geneve, Switzerland.

Boose, J., & Bradshaw, J.M (1987) AQUINAS: A knowledge acquisition workbench for building knowledge-based systems. *Proc. First European Workshop on Knowledge Acquisition for Knowledge-based Systems*, Reading University.

Bratko, I., & Kononenko, I. (1987) Learning diagnostic rules from incomplete and noisy data. In: Phelps, B. (ed.) *AI Methods in Statistics*, London: Gower Technical Press.

Bratko, I., Mozetic, I., & Lavrac, N. (1988) Automatic synthesis and compression of cardiological knowledge. In: Hayes, J., Michie, D., & Richards, J. (eds.) *Machine Intelligence 11*, Oxford: Oxford University Press. Also in: Michie, D., & Bratko, I. (eds.) *Expert systems: Automating knowledge acquisition*, Addison-Wesley, 1986.

Bratko, I., Mozetic, I., & Lavrac, N. (1989) *KARDIO: A study in deep and qualitative knowledge for expert systems.* Boston: The MIT Press (in press).

Breuker, J., & Wielinga, B. (1987) Knowledge acquisition as modeling expertise: The KADS methodology. *Proc. First European Workshop on Knowledge Acquisition for Knowledge-based Systems*, Reading University.

Chambers, R.A., & Michie, D. (1969) Man-machine cooperation on a learning task. In: Parslow, R.D., Prowse, R.W. & Elliot-Green, R. (eds.) *Computer Graphics: Techniques and Applications*, London: Plenum Press.

de Kleer, J., & Brown, J.S. (1984) A qualitative physics based on confluences. *Artificial Intelligence 24 (1-3)*, pp. 7-83.

Feigenbaum, E.A. (1977) The art of artificial intelligence: Themes and case studies of knowledge engineering. Pub. No. STAN-SC-77-621, Stanford, CA: Stanford University, Department of Computer Science.

Forbus, K.D. (1984) Qualitative process theory. *Artificial Intelligence 24 (1-3)*, pp. 85-168.

Gaines, B.R. (1987) Knowledge acquisition for expert systems. *Proc. First European Workshop on Knowledge Acquisition for Knowledge-based Systems*, Reading University.

Hart, A. (1986) *Knowledge acquisition for expert systems.* London: Kogan Page.

Kahn, G., Nowlan, S., & McDermott, J. (1985) MORE: An intelligent knowledge acquisition tool. *Proc. Ninth Intl. Joint Conference on Artificial Intelligence*, Los Angeles, CA: Morgan Kaufmann.

Kuipers, B. (1986) Qualitative simulation. *Artificial Intelligence 29*, pp. 289-338.

Luconi, F.L., Malone T.W., & Scott Morton, M.S. (1986) Expert Systems: The next challenge for managers. Sloan Management Review.

Michalski, R.S., Mozetic, I., Hong, J., & Lavrac, N. (1986) The multi-purpose incremental learning system AQ15 and its testing application on three medical domains. *Proc. Natl. Conference on Artificial Intelligence, AAAI-86*, pp. 1041-1045, Philadelphia: Morgan Kaufmann.

Michie, D. (1986) Machine learning and knowledge acquisition. In: Michie, D., & Bratko, I. (eds.) *Expert systems: Automating knowledge acquisition*, Addison-Wesley. Also published as Current developments in expert systems. In: Quinlan, J.R. (ed.) *Applications of Expert Systems*, Addison-Wesley, 1987.

Michie, D., Muggleton, S., Riese, C. & Zubrick, S. (1984) RULEMASTER: A second-generation knowledge-acquisition facility. *Proc. First Conference on Artificial Intelligence Applications*, pp. 591-597, Denver.

Mozetic, I. (1986) Knowledge extraction through learning from examples. In: Mitchell, T.M., Carbonell, J.G., & Michalski, R.S. (eds.) *Machine Learning: A Guide to Current Research*, pp. 227-231, Boston: Kluwer Academic Publishers.

Mozetic, I. (1987a) Learning of qualitative models. In: Bratko, I., & Lavrac, N. (eds.) *Progress in machine learning*, pp. 201-217, Wilmslow, UK: Sigma Press.

Mozetic, I. (1987b) The role of abstractions in learning qualitative models. *Proc. Fourth Intl. Workshop on Machine Learning*, pp. 242-255, Irvine, CA: Morgan Kaufmann.

Pearce, D.A. (1988) The induction of fault diagnosis systems from qualitative models. *Proc. Natl. Conference on Artificial Intelligence, AAAI-88*, pp. 353-357, Saint Paul, MN: Morgan Kaufmann.

Quinlan, J.R. (1986) Induction of decision trees. *Machine Learning 1 (1)*, pp. 81-106.

Steels, L. (1985) Second generation expert systems. *Future Generation Expert Systems 1 (4)*, pp. 213-221.

Welbank, M. (1983) A review of knowledge acquisition techniques for expert systems. Ipswich: Martlesham Consultancy Services.