large knowledge-based systems. Currently there are no tools available for this task, with the exception of primitive tools such as knowledge-base editors, e.g., KREME, etc.

This book is divided into two parts, and investigates two important aspects of practical knowledge representation systems:

- *reasoning* with represented knowledge; and
- *revising* represented knowledge.

The first half of the book deals with the dynamic aspects of knowledge representation. It contains a comprehensive survey of knowledge representation formalisms, the architecture of various services of knowledge based systems, and their management. The emphasis is however on *hybrid* systems, including semantic networks, frame systems, structural inheritance networks, etc. The book demonstrates a deep understanding of these systems and of the underlying theories. The analysis is further constrained to a typical member of the family of hybrid representation systems based on KL-ONE.

The second half of the book is devoted to the study of the problems associated with the revision of knowledge bases. It is often assumed, when discussing inferences, that knowledge bases grow monotonically. This excludes the possibility that a fact or definition has to be retracted or modified once entered. Here, instead of proposing a solution, the book takes a more general view on the problem of revision, which is then used as a yardstick for the effectiveness of previous solutions such as network editing approaches, knowledge base editing, deletion and additions of definitions systems, etc.

The target audience of the book falls into three groups: teachers/students, researchers, and other practitioners in the field. Overall the book is a popularized version of the author's PhD thesis. As such it contains a considerable amount of theory and formalism which is of limited value to the first and last groups. It is however valuable reading for prospective PhD students and most serious researchers in the field. And it is still a good PhD thesis.

Review by Dr. Paul Refenes.

KARDIO: a study in deep and qualitative knowledge for expert systems by Ivan Bratko, Igor Mozetic and Nada Lavrac, The MIT Press, Cambridge, MA, 1989, pp 260.

Disorders in the electrical control system of the heart are called cardiac arrhythmias. There are about 30 basic arrhythmias, each of which causes some characteristic change in the normal ECG. Simple arrhythmias combine to give complex combinations of multiple arrhythmias. Although the ECG features corresponding to simple arrhythmias are well documented in the medical literature, there is no systematic description of ECG features for multiple arrhythmias. This knowledge simply does not exist to render the application of a dialogue-based knowledge acquisition technique a possibility. If the expertise does not exist, the only viable knowledge acquisition technique is to "learn" the expertise from the domain first principles. This was the motivation behind the KARDIO project which started in 1982. The book by Bratko, Mozetic and Lavrac is an account of the research conducted in the period 1982–1987. The initial intention of KARDIO was ECG diagnosis, directly from the actual ECG curve; however, in the present system the ECGs are manually translated to a symbolic description. Prediction is also provided, since it is intrinsically related to the automatic derivation of the relevant expertise: "Given a particular legal combination of simple arrhythmias what are the possible ECG descriptions?"

The terms "deep" and "qualitative" included in the book title are currently very much in vogue within the knowledge engineering community. Through its title the book should therefore have an instant appeal to many knowledge engineers, who will benefit if they actually decide to invest in the book. From the title, however, one may not immediately guess that a central theme in KARDIO is automatic knowledge acquisition through rule induction based on techniques which have been established in the early 1980s. The success of the KARDIO project and the wide dissemination of its results through the publication of the book will undoubtedly inject a substantial new dose of interest in rule induction.

The authors do not attempt to give an in-depth analysis of what a deep model is; instead, their "definitions" are restricted to the usual four-liners: a deep model enables the application of domain first-principles; a deep model is not a shallow (surface) model which is one that directly associates problem specifications to problem solutions; a deep model derives a solution whilst a shallow model retrieves a solution; a shallow model is more effective, but a deep model can answer questions which the shallow model cannot. These statements are effectively illustrated initially through a simple circuit example, and subsequently through the actual deep model of the heart used in the KARDIO system. The circuit example brings out very naturally the intrinsic relativeness in the concept of deepness – "this model is deep relative to that model".

Deep versus shallow is one dimension along which domain models can be analyzed. Qualitative versus quantitative is another dimension. A qualitative model is informally introduced as one where the model variables take symbolic rather than numeric values, and where the behaviour of the model is derived from the behaviour of its components. Again, these statements are succinctly illustrated through the pedagogical example of the circuit, and of course in a realistic setting through the model of the heart.

The electrical model of the heart developed in KARDIO is both deep and qualitative. Conceptually, the model is a network of components and links, where components represent impulse generators, conduction pathways, impulse summators and ECG description generators. The behaviour of these components is described by a set of 35 global rules which are partitioned under the following classes corresponding to the four types of component: impulse generation; impulse conduction; summation of impulses; producing ECG. Given some initial conditions, i.e., a set of arrhythmias, the rules are applied in a forward chaining fashion until no more rules are triggered, and then all the ECG descriptions thus generated are collated together to give the overall ECG picture corresponding to the given set of arrhythmias. A simple arrhythmia defines a particular abnormal state for a single heart component. Thus from these initial abnormal states impulses are generated, conducted along pathways, summated and translated to ECG features. Impulses are of two types: permanent and occasional. Each global rule has the format "Cause ⇒ Consequence where Condition". Where there are a number of possible instantiations for a Cause, the Condition is used to decide which is the correct choice. A valid Condition instantiation must unify with a local rule. Local rules are assertions about the heart components (impulse generators and summators, and conduction pathways).

The model of the heart supports a predictive task ("what ECG features correspond to a given set of arrhythmias?") but the format of the global behavioural rules does not accommodate a diagnostic task effectively ("what arrhythmias could be the cause of a given ECG description?"). To derive the ECG features corresponding to a set of arrhythmias the global rules are applied in a forward-chaining breadth-first fashion. To derive an explanation for an ECG description the rules need to be applied in the reverse order, but backward chaining yields a much higher branching factor for this approach to be practicable. It may therefore be said that the heart model supports prediction but not diagnosis, at least not in a practicable way. The fact that this model supports the task of prediction much better than the task of diagnosis may be interpreted by one school of thought, in second-generation systems, as being constructed with a specific bias regarding its use and hence the model is not deep. There is another school of thought however which advocates that an effective deep model is one that is strongly coupled with its usage and in this respect the KARDIO heart model is deep; after all the model is really only being used in the context of a predictive task.

Paradoxically, the main function of the KARDIO system is diagnosis not prediction, although the behavioural rules in its model relating causes to their effects support the latter rather than the former. KARDIO simply uses this deep model as a stepping stone to getting to the shallow diagnostic knowledge. The deep model is used to derive a surface knowledge-base consisting of (⟨set of arrhythmias⟩ ⟨ECG description⟩) associations, by performing a model simulation for every legal combination of arrhythmias. This knowledge base is too big and unwieldy, and hence operational surface knowledge is obtained by applying a rule induction algorithm (of the AQ class) on a carefully selected subset of the original surface knowledge base. Two separate operational knowledge bases are induced, one for a predictive task and one for a diagnostic task. In either case, the subset of the original surface knowledge base has to be appropriately transformed before the induction algorithm can be applied. This further illustrates that effective operational knowledge is derived within a specific operational bias.

As a side point, an issue which is very briefly touched in the book is that the generated surface rules, relating arrhythmia combinations and ECG features, are often more complicated than the corresponding rules given in the medical literature because they make explicit conditions which are essentially background domain knowledge which would be difficult to encode and maintain. I share the authors' view that explicating background knowledge is difficult, not least because of the unconstrained nature of the elicitation taks, but in my particular experiences the explicit representation of even a fraction of the background knowledge can have advantages.

To me the significant contribution of the KARDIO project to knowledge engineering is a second-generation methodology for knowledge acquisition based on a deep qualitative model of the domain. The deep model is a stepping stone leading to shallow operational models for specific tasks; it is not a run-time model that dynamically provides run-time support to the specific tasks. In this respect, KARDIO does not provide a second-generation architecture for expert systems; the operational systems in KARDIO are traditional first-generation systems. In Steels (1986) original specification, a second-generation expert system has two models, a deep model and a shallow model, each with its own knowledge-base and inference engine. The two models can support each other during consultations. The shallow model may provide the answer to a problem efficiently but it is the deep model that provides adequate explanations. The shallow model may fail to solve a problem which is then referred to the deep model; the derived solution is used to upgrade the shallow model. The shallow model is therefore gradually constructed from the deep model in the context of actual problems. The KARDIO components can of course be used in this fashion at least for a predictive task.

The book consists of five chapters and three appendices. Chapter 1 introduces the reader to all the relevant concepts (deep model, shallow model, qualitative model, induction), and gives an excellent overview of the KARDIO methodology. Chapter 2 gives a detailed description of the model of the heart, listing all the global and local behavioural rules. Chapter 3 gives the model simulation algorithm and discusses the derivation of the surface knowledge-base. Chapter 4 discusses the derivation of the operation rule bases through induction. Chapter 5 discusses the semi-automatic construction of a deep qualitative model by a process of top-to-bottom refinement. The three appendices respectively give: part of the surface arrhythmia-ECG base; operational diagnostic rules; and operational prediction rules.

The book is well written with a coherent theme running throughout. A fair amount of the presented material is technical (Prolog clauses). The knowledge engineering methodology which is illustrated through the KARDIO project is generic; making the book very useful reading for all knowledge engineers, and not just those interested in medical knowledge engineering. I have certainly benefited much from reading the book.

References

Steels, L, 1986. "Second-generation expert systems" In: *Research and development in expert systems III*, Bramer MA (Ed.), Cambridge: Cambridge University Press.

Reviewed by Elpida Keravnou, Department of Computer Science, University College London, Gower Street, London WC1E 0BT UK.