Readers who have experience either of conventional systems development or of unsuccessful real-world KBS projects are likely to recognize the relevance of KADS and to welcome this book as the first readily available exposition of the methodology for analysis. We must hope that further volumes appear covering all the other stages and issues in KBS development. The publishers are to be congratulated on a well-produced and speedily published (November 1989) book, which has a good index and very few errors.

Reviewed by Ian Neale, ARIES at City, The City University, Northampton Square, London EC1V 0HB.

Abductive inference models for diagnostic problem-solving by Yun Peng and James A. Reggia, Springer-Verlag, Berlin, 1990, pp 268.

For many years diagnostic problem-solving has been, and still is, an active research area. The understanding and formalization of diagnostic reasoning, especially in medical domains, has intrigued researchers since before the advent of knowledge-based systems. This work has in fact significantly contributed to the technology of expert systems; diagnostic reasoning is certainly one of the best documented applications of this technology.

The book by Peng and Reggia is a notable addition to the existing literature on the subject. The authors by discussing abductive inference models for diagnostic problem-solving, to the exclusion of deductive inference models, on the one hand present an incomplete picture of diagnostic reasoning but on the other hand they emphasise the importance of abductive inference in diagnostic reasoning; in many discussions on diagnostic reasoning abduction takes second role to deduction. Abductive inference, or hypothesis generation, is intrinsically related to diagnostic reasoning. The authors informally define abduction as the process of generating plausible explanations for a set of observed manifestations. Deductive reasoning is also intrinsically related to diagnostic reasoning. Often a large number of equally plausible alternative explanations are generated through abduction and deductive reasoning is required to differentiate between these alternatives; hence the use of the term hypothetico-deductive reasoning for diagnostic reasoning. The "parsimonious covering theory" which is the generic name given to the various models presented in the book, does not therefore constitute a complete theory of diagnostic reasoning. This is not to say that it does not make a worthwhile contribution to the field of diagnostic reasoning, as explained below.

In general, there are two main schools of thought regarding diagnostic reasoning. One school advocates the use of formal models, which by and large are founded on probability theory: the other school advocates the use of heuristic models which emulate as closely as possible the elicited (articulated) thought processes of human diagnosticians. The formal models tend to be simpler in symbolic knowledge but richer in probabilistic knowledge. The heuristic models tend to be richer in symbolic knowledge, and often use qualitative rather than numeric models for handling uncertainty; heuristic models can be difficult to formalize. Formal models are domain-independent, applicable in any domain satisfying the model's underlying assumptions. Generic heuristic models normally constitute generalizations of domain specific models. Formal models and heuristic models can certainly be integrated. There are strengths and weaknesses for either type of model. Which approach is the better one (efficient and effective) for a specific diagnostic domain given the particular design objectives can only be proven in practice.

The above book belongs to the formal model school of thought – although some aspects of the models presented are justified by relating them to human diagnostic behaviour. Readers belonging to the heuristic model school may find this book rather heavy reading. It abounds in lemmas, theorems and proofs. Mathematical appendices can be skipped, but still there is much notation which needs to be understood to appreciate the discussion on the models.

Parsimonious covering theory generates plausible multi-element (i.e. multi-disorder) hypotheses for a set of manifestations. The elements of a hypothesis collectively cover all the manifestations. In addition, each hypothesis must satisfy a parsimony criterion. Three such criteria are given: minimality, irredundancy and relevancy. The irredundancy criterion (no subset of the hypothesis elements is itself a plausible hypothesis) is the one preferred by the authors and used throughout the book. The theory therefore deals with the general and far more complicated diagnostic problem of multiple rather than single failures. Contrary to classification theories for diagnostic problem-solving where a hypothesis is selected from amongst a set of pre-specified hypotheses, a hypothesis here is constructed by gradually assembling together a number of hypothesis elements. Multi-fault diagnostic problems are not simple and their solutions are NP-complete. This is true for the algorithms given in the present book.

The basic parsimonious covering theory is discussed in Chapter 3. In the basic theory there are two disjoint sets, the disorders which constitute hypothesis elements and the manifestations which constitute the observations. Each disorder causes certain manifestations and each manifestation is caused by certain disorders. The causal associations have no uncertainties. Algorithm BIPARTITE gives the solutions to a diagnostic problem in *generator form*. An example of a generator is ({d1,d2} {d4,d8}) where d1, d2, d4 and d8 belong to the set of disorders. The generator implicitly defines the four alternative solutions (d1,d4), (d1,d8), (d2,d4) and (d2,d8), i.e. each element of a generator is a set of disorders, where each of these disorders explains the same subset of the present manifestations and hence constitute alternative plausible explanations for the particular subset of present manifestations. Solutions generated by the algorithm can be justified in terms of the parsimony criterion used, e.g., minimality or irredundancy. The basic theory has a limitation, namely that it may generate a large number (say 20 or more) of alternative solutions, each of which is equally plausible according to the coverage and chosen parsimony criterion. The theory does not offer a means for differentiating between such alternatives. In a real life situation this would not be acceptable, especially if each alternative leads to a different treatment.

The above limitation is addressed by extending the basic theory to include uncertainty (Chapters 4 and 5). Disorders are given prior probabilities and causal links between disorders and manifestations are given causal strengths represented as $P(m_j \mid d_i)$ where m_j is a manifestation and d_i is a disorder. Hypotheses are assigned relative likelihood measures based on the prior probabilities and causal strengths. The relative likelihood measure is a product of three weights, each of which is itself a product:

- a weight reflecting how likely the hypothesis is to cause the present manifestations;
- a weight based on manifestations expected by the hypothesis which are absent;
- a weight based on prior probabilities of the hypothesis elements.

All products involve only probabilistic information local to the hypothesis elements and present manifestations and the effects of hypothesis elements. Problem-solving strategies, both for closed problems (the status of each manifestation, as present or absent, is known in advance) and open problems (some manifestations are known to be present, some are known to be absent and the status of the rest is unknown) are given. The strategies are based on the computation of an upper bound for the relative likelihood measure. The solutions here are not expressed in generator form.

Another extension to the basic theory is to generalize bipartite problems (problems consisting of two disjoint sets, disorders and manifestations) to allow for layered and hyper-bipartite problems (Chapter 6). This extension does not include probabilistic knowledge (the authors say that the incorporation of uncertainties is presently in progress). Chapter 6 would be better placed immediately after Chapter 3 on the basic theory, especially since the reader is advised to go back over the generator form calculus before reading Chapter 6. A layered problem has n disjoint sets (or layers) of entities where $n \geq 2$. The lowest layer is the manifestations, the highest layer is the disorders and the intermediate layers are pathological states. Causal links can only exist between entities in adjoining layers; thus each causal chain must be of the same length which the authors admit is a limitation. A solution to a layered problem is obtained by sequentially applying the BIPARTITE algorithm to adjoining pairs of layers starting from the lowest pair. Prior to each application of the algorithm, except for the first application, the output from the previous application must be suitably converted to constitute the input for the next application. Hyper-bipartite problems are the generalisation of layered problems by removing the constraint of fixed length causal chains. The algorithm for layered problems can be used for hyper-bipartite problems by incorporating pseudo states which convert a general hyper-bipartite causal network to a layered causal network. This conversion is done dynamically (i.e., only to the relevant parts of the network) and is transparent to the user. The authors accept that completely converting a real life causal network to a layered network would not be practicable.

The last topic covered in the book is a diagnostic model based on the connectionist approach (Chapter 7). The connectionist model is compared against the probabilistic models. The conclusion here is that the connectionist model is promising but further investigation is required.

In the concluding chapter the authors briefly discuss other application areas for abductive inference models, namely natural language processing and learning and discovery. The limitations of the parsimonious covering theory which constitute future research directions are also discussed. These are:

- incompleteness, i.e., deduction (information acquisition) is not covered;
- richer structures are required, i.e., neither disorders nor manifestations have internal structure;
- temporal reasoning is missing;
- deeper models, e.g., structure and function models, are missing.

In their final note the authors state (p. 267) that "Any theory is only a hypothetical model of the real world. Although it captures key features in a rigorous manner, parsimonious covering theory as presented in this book is only a hypothetical model of the abductive inference process that humans conduct to solve diagnostic problems. The next step is to demonstrate the validity of this hypothesis by building real systems and compare their behaviours with systems based on other theories as well as human performance." In a sense it is a pity that such experiments could not be conducted prior to the writing of the book. Positive results would have considerably strengthened the validity of the book. In spite of this the book is useful reading for anyone interested in diagnostic reasoning.

Reviewed by Elpida Keravnou, Department of Computer Science, University College London, Gower Street, London WC1E 6BT UK.

HERMES and the golden thinking machine by Alexander Tzonis

This claims to be a book which can be read on several levels: part detective thriller and part introduction to various topics in AI and cognitive science. To write such a book is an ambitious task, since it implicitly promises to stand on its merits at each level.

As a detective story, the book introduces a large cast of characters and a range of exotic Greek locations. Very few of the cast make more than a fleeting appearance, or are sufficiently developed to make plausible suspects. If we exclude those who enter merely to present a single set-piece speech, and those who are killed en route, the search space of possibilities is drastically pruned. The general reader looking for an intriguing thriller is likely to be frustrated by the lack of convincing characterization, and the lengthy pauses in the plot to allow the author to pursue his other aim.

As an introduction to AI subjects, including search methods, analogy and knowledge representation, the book's main technique is the set-piece speeches already mentioned. These are either given by one of the many characters met by the hero, Hermes (a young Harvard professor of archaeology and artificial intelligence) during his adventures; or sometimes by Hermes himself to his assistant Nina. Some of them are quite amusing, if you already know something about the subject in question. There are also a myriad of minor allusions and oblique references, which the interested reader may enjoy spotting.

AI is a sufficiently varied and interesting subject to make many introductory books and articles