

creates the desired, realistic impression that these are not to be separated in business environment. The analysis of business is shown to be necessary before implementation of information systems and application development can take place.

At times this seems a strongly business oriented book for those unfamiliar with that world. The problems of implementing applications are highlighted and research solutions are offered from published analyses of strategies and corporate structures. An important distinction is made for the data-oriented reader between the issues and the tasks involved in implementing a database system; the task is computing but the issues are frequently people and their jobs.

The business analysis chapter (Chapter 2) deals sensitively with the people and politics of business, balancing people and technology. It stresses that this is essential to smooth and efficient implementation of information systems. The principles of business beyond the profit motive are explained for computer scientists and the role of well-managed data as a company resource explained for top and middle management as well as operational staff. Both the information system and the quality of database administration are vital to good performance in implementation and usage, in areas such as security, privacy and integrity of data, and there is a need to recognize this in appointment of associated personnel.

Building on the basics of earlier sections on databases and database environments, newer technologies are described. These include distributed databases along with their requirements, principles and design, and fourth generation tools, workbenches and data dictionaries. Sound advice is offered to business people on evaluating tools, with descriptions, explanations and cost benefit analyses, and including criticism of inflated claims of suppliers and writers. The book finishes positively with two examples of successfully implemented systems, which it calls "the approach in action". These bring together the principles and pragmatism described so far, to complete a useful introduction and summary of a rapidly developing, interdisciplinary subject.

Reviewed by Catherine Hearne, Advanced Computation Laboratory, Imperial Cancer Research Fund, London, UK

Rough sets—theoretical aspects of reasoning about data by Zdzisław Pawlak, Kluwer Academic, The Netherlands, 1991, pp 229, £56.00, ISBN 0-792-31472-7.

When a well-established author, especially one who has been working in a single field for as long as Professor Pawlak has, sets out to write a monograph on the subject that he invented, the resulting book is usually well worth reading. Books such as Judea Pearl's *Probabilistic reasoning in intelligent systems* and Bob Kowalski's *Logic for problem solving*, to take two examples which I can find directly to hand, are cases in point. Ideas which were previously spread around a number of papers are brought together in a single volume, and together are greater than the sum of the parts, giving a view of the author's area of expertise which is both broad and deep, and which suggests both applications and related avenues of research. Since I am a strong believer in the usefulness of rough sets I was greatly disappointed to find that this volume is not such a book. Indeed, the collection of all the ideas that Professor Pawlak has previously published on rough sets has almost exactly the opposite effect to that which I expected. Having read the book I found myself wondering quite where all the ideas and insights that I had previously found in the various papers on rough sets had gone to, and I am still not quite sure. However, I think I can, to some extent, explain what has happened.

The first thing to understand about rough sets is that the idea behind them is extremely simple. This is in no way intended as a criticism of them, since I firmly believe that ideas do not have to be complex to be good. The idea is the following. Consider a set of old-fashioned children's building blocks, made of wood and painted bright colours—red cubes, red pyramids and yellow pyramids, blue spheres and yellow spheres. Now, consider dividing the blocks into groups by their colours alone. The set of red blocks will contain cubes and pyramids, that is cubes and pyramids cannot be distinguished by colour alone. Alternatively, one can consider describing the colour of the blocks in

terms of their shapes, so that it is possible to categorise them when blindfold. From this point of view any cube is known to be red, but there are red objects that are not cubes, so that to say that the only red objects are cubes is to underestimate the size of the set of red objects. Conversely, all red objects are cubes or pyramids, but to say that any cube or pyramid is red is an overestimation since the set of cubes and pyramids includes yellow pyramids. However, this pair of lower and upper approximations to the concept *red* {*cube*} and {*cube*, *pyramid*} is the best description of the concept that may be attained for the set of blocks when they are described using the shape of the blocks alone. This pair is also the rough set of this description, and that is all that there is to the idea of a rough set—it is a pair of approximations to a concept using a given language.

The second thing to understand about rough sets is that formalising this simple idea seems to be far from simple. Of the whole monograph, 116 pages are devoted to precise mathematical formulation of rough sets and attendant concepts. Now, in general I'm not opposed to precise mathematical formulations, since this often seems to be the best way to convey ideas precisely and concisely. However, there are limits and these have to do with both the complexity of the idea being expressed, since if it is very simple then a mathematical formulation is often longwinded or confusing, and the apparent relevance of the idea being expounded. The latter is particularly important, since I find that ideas are much less exciting when expressed formally and so, unless properly motivated by the author I tend to find it difficult to keep ploughing through expression after impenetrable expression. Unfortunately, this book seems to me to exceed both limits. Firstly many of the ideas are sufficiently simple, as the examples show, that they could be conveyed far more easily in text, and secondly the work is woefully motivated. Even the examples, which are thankfully plentiful, don't help the motivation since they are all such small nuts that they merely stress the size of the mathematical steamhammer being applied to them.

The third thing that should be understood about rough sets is that the applications of the theory are currently rather scarce, and they lack variety. The five different applications discussed in the second half of this book are all roughly the same. In each case, the idea is that a number of combinations of attribute values are assessed to determine a minimal consistent set. Depending upon the particular instance this set might be the logical combination of inputs necessary to determine a given output from a digital circuit, the critical set of demands that a particular country might make of a Middle East peace settlement, or the conditions under which the setting of a cement clinker kiln should be adjusted. In this latter case, a number of "decision rules" are extracted and these are rather confusingly referred to as a "decision algorithm"—confusingly because there is nothing algorithmic about them. The form of this analysis is exactly the same in each case, each attribute in turn is tested to see if it is really necessary and duplicate or equivalent sets of attributes are removed. Although extensions and connections with other methods are hinted at, nothing concrete is stated.

Thus the book fails to be a great monograph because it does not cover enough ground, and does not tie the subject up in such a way that the reader winds up thinking "Wow, there are so many things that it is possible to do with rough sets that I understand why people are excited about them", which is the kind of reaction that is provoked by better books. Instead, what is presented is the basis of a theory. This basis is presented in a highly formal way without a great vision for its application. It is therefore a theory which still needs a lot of work to be done on it and with it, and this is perhaps the way that the book should be approached—as a summary of what has been achieved to date, and as a source book of results that should act as a spur to new research. Viewed in this way the book deserves to be heartily recommended as an essential purchase to all those interested in pursuing work on rough sets. However, as in the case of the companion volume, which I reviewed in the last issue of *The Knowledge Engineering Review*, the price makes it the kind of book that will be bought by libraries rather than individuals, and so I feel I can only recommend that it is borrowed from a library that has a particularly large amount of money to spend.

Reviewed by Simon Parsons, Advanced Computation Laboratory, Imperial Cancer Research Fund, London, UK