community there must be user participation and education, which could well begin with a text such as this.

The quality and potential of the chapters may distract from hard questions about their real contribution to biological science or to advances in AI. The work described here is still largely that of AI applications research. The emphasis in most chapters is on tool development for biological research in challenging areas of interpretation and rationalization of observed phenomena. These areas require the intuition and experience of domain experts, but they also offer the opportunity for expert systems and automation. Unfortunately, there is as yet little evidence of successful laboratory use of this.

Hunter concedes that computation in biology has so far proceeded without AI, despite its use elsewhere. He believes that AI must now make its own new discoveries, rather than repeat results from conventional means, if it is to gain widespread and practical acceptance. This book may reinforce the view that AI techniques have contributed nothing new to biology, but demonstrate more potential than progress. This publication also undoubtedly represents the influence of a new opportunity for AI; the Human Genome Project, which may allow AI techniques to prove themselves in an area whose needs cannot otherwise be met. This global research endeavour to determine the molecular meaning of life is generating an abundance of data with a complexity which exceeds the capacity of conventional computational techniques. These data raise many problems of interpretation which cannot be easily addressed, perhaps most noticeably in the prediction of the protein structure from DNA sequence. The need for symbolic manipulation and reasoning argues strongly for the use of artificial intelligence techniques in these areas, as is clearly demonstrated by the contributors to this book. Thus a new opportunity is being provided by molecular biology for AI which only requires that the latter is sufficiently mature to begin a lively and productive relationship.

Reviewed by Catherine Hearne, ACL, ICRF, London, UK

Artificial intelligence and molecular biology by L. Hunter (Ed.), AAAI Menlo Park, CA and MIT Press, Cambridge MA, 1993, pp 470, £35.95, ISBN 0-262-581159.

Spare a thought for those who will soon have to tackle the biggest cryptographic puzzle of all time—deciphering the 3000 million "bits" of data found in the genetic blueprint of a human being. The majority of the complete genetic blueprint of a human—the genome—is stored on the strands of DNA found in the chromosomes of the cells of our body. The massive international program of the Human Genome Project plans to read the billions of "bits" of data recorded on the DNA strands. These bits are in the form of one of four chemicals (abbreviated to A, T, G or C) which are strung together to make a DNA strand. Some regions of DNA strands—called genes—direct the manufacture of proteins.

Proteins are essentially the molecules which give the cell structure and the ability to do the specialized jobs that different cells do. Proteins are made of amino acids which are strung together in an order directed by the gene. By obtaining the sequence of the entire genome, scientists will have access to the genes and proteins that regulate the processes of embryogenesis, death and the bit in between. Given the DNA sequence of a gene, it is easy to predict the order in which the amino acids are built into a protein. However, prediction of the actual 3D structure of this protein is incredibly difficult because the amino acids of the protein interact with themselves and the surrounding water in a complex manner. In water, proteins are not long strings of amino acids, but are folded into intricate structures. Thus, by reading the sequence of a gene, we cannot easily predict the structure—and the function—of a protein.

This book aims to introduce readers to the various approaches used to examine biological systems through computational techniques. It attempts—and succeeds—to draw together a range of reviews covering biology-oriented aspects of information processing and artificial intelligence.

The chapters contain meaty explanations of the history, application and direction of the techniques used to study DNA sequences, protein structure and to model complex biological systems. As this field is in its infancy, Hunter has carefully balanced the book to present the reader with a number of possible approaches to the problems faced.

Newcomers to biology are introduced to the subject of biological molecules in an outstanding overview by the Nobelist Joshua Lederberg. His discussion of molecular biology in this book must rank as one of the clearest pieces of popular scientific writing ever. The most stimulating discovery for me in this book was the emerging application of computational linguistics to DNA sequence analysis and protein structure prediction. The introductory chapter by Searls on this topic deserves special praise.

Artificial intelligence approaches discussed here include neural net learning and knowledge based simulation which can detect patterns or model complex systems of biological molecules. The role of these models is discussed by Lederberg in his concluding chapter. He emphasizes that any model or interpretation of a system is only as reliable as the rules we use to define it, and lists no less than 13 major biological dogmas which have proven to be partly or wholly incomplete. The constant watchword is "vigilance" against inductive reasoning and incorrect rules which are incorporated into the algorithms used in these calculations. Nevertheless, the potential for discovery in this field is immense. For example, over 90% of the DNA in the genome does not support the formation of proteins. There is much speculation about a possible regulatory role for this "junk" DNA. Perhaps through deductive approaches using linguistics and neural net pattern recognition, patterns may be detected in the organization of the genome itself. Biologists cannot yet do experiments on the regulatory role of large scale DNA organization, although there is clear evidence that it occurs.

The only criticisms of the book are minor. Given the range of difficult subject matter covered, it would have been useful if there was a glossary or, at least, an indication in the index of the most descriptive entry of a term. Hunter makes one minor exaggeration in ascribing a role for quantum mechanics in protein structure, where it (fortunately) has no role to play. One omission on DNA regulatory regions are chromosome attachment sites—points where DNA strands are attached to the chromosome skeleton and which can drastically influence DNA function.

I found this book a stimulating and balanced introduction to a subject which will be the computational and biological equivalent of the Apollo Program in the 1960s and 1970s. Everyone with an interest in artificial intelligence or information theory—or even a curious mind—should skim a copy and savour the wealth of ideas described. Anyone working in this field will find it an excellent source of reference and definitive reviews. By examining the very nature of life, computational approaches to biology will potentially deliver more than a biro which can write upside down. This excellent book has started the countdown.

Reviewed by B.R. Clark

Case-based reasoning by Janet Kolodner, Morgan-Kaufmann, CA, USA, 1993, pp 668, £45.95, ISBN 1-55860-237-2.

Does case-based reasoning deserve to be considered as a radical new idea, or does it simply reconstitute a number of traditional approaches in a new guise? Janet Kolodner has contributed to research in this area from its early days in the late 1970s, and has produced a wide ranging account of the field. Unlike earlier texts (such as Riesbeck and Schank's *Inside Case-Based Reasoning*, she has attempted to provide more than a collection of examples of systems—much of her effort has been devoted to providing a rationalized summary of the methodology. This should provide us with a window onto this area of activity, and allow us to place it in the wider context of KBS research.

To provide a balanced account of case-based reasoning, it is necessary to accommodate both the "cognitive modelling" view, in which psychological plausibility is a dominant theme, and the