Biophysical Journal Volume 66 April 1994 1258–1260

Book Reviews

Artificial Intelligence and Molecular Biology by Lawrence Hunter, Editor

AAAI Press/The MIT Press, Cambridge, Massachusetts, 1993. 470 pages. $39.95

Reviewed by Lynda B. M. Ellis

The title of this book is open to a wide range of interpretations. Besides the usual "What is artificial intelligence?" and "What is molecular biology?," there is the question of which topics to include at the intersection of the two. One approach to these questions of definition is to take a snapshot of the moving target. In this case the camera was focused on the 1990 Spring Symposium on Artificial Intelligence (AI) and Molecular Biology sponsored by the American Association for Artificial Intelligence, and papers presented there represent a good, although necessarily incomplete, overview of one aspect of the field of computational biology. Authors were given the opportunity to update their papers, and there are 1991 and 1992 references. These bibliographies can serve as an introduction to the literature in the field, but these 12 chapters are, for the most part, conference papers not exhaustive reviews of the current literature.

Although not explicitly stated, this book seems to have as its intended audience computer and other physical scientists interested in biological applications of AI techniques and only secondarily biologists interested in the same. This is tacitly acknowledged with an introductory chapter on "Molecular Biology for Computer Scientists" by Lawrence Hunter, but no similar chapter on "Computer Science for Molecular Biologists." The physical scientist without significant biological background should read this introductory chapter with caution. The author (who is also the book's editor) is to be commended for undertaking such a challenge, but there are too many over-simplifications and outright misstatements to permit its unqualified endorsement. Also, although there is mention of organismal variation, experimental error and the innate variability found in data from experiments involving material derived from living organisms are unfortunately omitted. These concepts can be ignored only with great risk by computer scientists who deal with biological data.

In the forward, Joshua Lederberg notes the daunting difficulty of merging the insights gained from each of the book's chapters. The first chapter describes each of the rest of the chapters and attempts to relate them to one another. A seven-page comprehensive index assists the reader who is interested in certain techniques or topics. Techniques include computational linguistics, neural networks, simulation, parallel computation, and pattern matching. These are applied to topics as diverse as protein and nucleic acid structure prediction from primary sequence, bacterial gene regulation, predicting metabolic pathways, and protein structure determination using nuclear magnetic resonance or x-ray crystallography. Because each chapter is written by a different set of authors, few generalities apply. However, chapters usually emphasize theory, techniques, or methodologies rather than results. If results are presented, they are often sketchy and not compared with those of other researchers.

Four of the twelve chapters deal with protein structure prediction. This is the largest subset of chapters devoted to one topic. Holbrook et al. apply neural networks to this problem. They present results for prediction of secondary structure, surface exposure of an amino acid, and the binding state of cysteine (SS or SH), and they are to be commended for opening up the neural network black box and assigning biological meaning to its weights. In a mere 15-page chapter, Zhang and Waltz describe a novel methodology to use the information in protein tertiary structure to assign residues to 23 residue state classes based on primary sequence and use this methodology to predict secondary and supersecondary structure. But no results are shown for their predictions, nor is this technique compared with work in the same area by Eisenberg and others. Lathrop et al. discuss several techniques including pattern matching, parallel computation, and multiple sequence alignment, including results on predicting structural motifs, an interesting section on "Significance, Validity and Pattern Quality," and a very complete bibliography of over 100 references. Hunter, in his second chapter, emphasizes method almost exclusively, but with perhaps better justification than others, inasmuch as he uses techniques of machine learning and discovery theories to develop a cognitive theory of the discovery process in this application area.

Three chapters simulate metabolic pathways. Karp models the bacterial trp operon using the GENSYM expert system. Experiments involving the trp biosynthetic pathway, transcription of the trp operon, and regulation of the entire trp operon were successfully simulated. Mavrovouniotis develops a database of metabolites and bioreactions, including Gibbs free energies, and uses it to predict biosynthetic paths to given products from given reactants such as the biosynthesis of lysine from glucose and ammonia. Galper et al. simulate DNA metabolic pathways, including introductory material for both computer scientists and molecular biologists, and describe not only the methods and results but also the user interface. Their simulation was developed for use by biochemists as an inductive reasoning tool, and thus it attempts "to provide methods for specifying the properties of DNA in as many ways as is natural for a scientist."

Four chapters each cover a different, more specialized topic. Searls discusses application of computational linguistic techniques to many problems in nucleic acid and protein sequence analysis. Steeg applies neural networks to RNA structure prediction. Edwards et al. interpret NMR spectra of proteins, and Glasgow et al. apply scene analysis to the phase-shift problem in protein x-ray crystallography.

As mentioned at the beginning of this review, the book is a snapshot of a moving target. Joshua Lederberg predicts that it will be viewed from the future as providing a "pivotal beginning" for understanding of the complete genome. Be that as it may, it would be worthy of purchase by those physical scientists desiring an introduction to AI applications to molecular biology, but it would be of slightly lesser use to the biologist desiring the same. At the end of the first chapter Hunter states: "In order to work with biologists, AI researchers must understand a good deal about the domain and find ways to bridge the gap between these rather different scientific cultures." This may be the most important concept a computer scientist could gain from this book.

Thermodynamics of Membrane Receptors and Channels by Meyer B. Jackson, Editor

CRC Press, Boca Raton, Florida, 1993. 439 pages. $95.00

Reviewed by Derek Marsh, Max-Planck-Institut für biophysikalische Chemie, D-37077 Göttingen, Germany

The title is intriguing and, as such, must inevitably, at this stage of the game, contain some disappointment. Nevertheless, this is a very creditable and wide ranging collection of reviews on membrane thermodynamics, channels, and receptors. Only in a restricted number of cases, understandably in the chapters by the editor, is there an attempt at a real marriage of thermodynamics with channel and receptor function.

The book opens with a short general chapter by Yeates on the structure of membrane proteins and the thermodynamic factors affecting their stability. This is followed by a treatment by Jordan of the interactions of ions with channel proteins that concentrates on model channel potential profiles and the effects of membrane surface potential. Roise contributes a chapter on the incorporation of peptides into membranes, mostly the electrostatics and binding, with consideration of channel-forming toxins, peptide hormone-mediated activation, and targeting sequences in protein translocation. Thermodynamic models for lipid-protein interactions, based principally on hydrophobic mismatch, are covered extensively by Mouritsen and Sperotto, with summary examples ranging from amphiphilic polypeptides to both receptors and other membrane proteins. Abney and Scalettar deal with statistical mechanical models and simulations of protein-protein interactions; protein crowding is a fascinating aspect of membranes and is particularly relevant to the formation of gap junctions.

At this point, we are approximately halfway through the book, and there is a change of gear with the three chapters authored by the editor. The first of these chapters considers the energetics of membrane receptor activation, based on allosteric models, and ends with a short section: "If Not Allostery, Then What?" (fractals, of course!). The two following chapters then deal with examples, taking separately receptors directly gating channels (mostly nicotinic acetylcholine, but also GABA, glutamate, and cyclic nucleotide), and those coupled to G-proteins and protein kinases. These chapters are necessarily more descriptive, but the first contains sections on the energetics of ACh binding and its transduction, and each chapter ends with a consideration of the appropriateness of allosteric models (directly ligand-gated: very; G-protein and kinase-coupled: maybe).

Then follows a chapter on mechanosensitive channels by Martinac, including simple thermodynamic models for the tension-induced activation. Ebrey gives a detailed and critical account of the photocycle of bacteriorhodopsin with reference also to structure and proton pumping, but relatively brief mention of the energetics of transduction. Parsegian and Zimmerberg discuss the thermodynamics of channels under osmotic stress, which they have shown can be used to measure the change in aqueous channel volume on opening. The VDAC channel, for instance, can be gated by osmotic pressure. Finally, the book closes with a chapter by Heinemann and Sigworth on the study of channel fluctuations by high resolution analysis of open channel noise. This illuminating account is unashamedly kinetic and illustrates the unique information that can be obtained by single channel recordings.

This is not a book to be read at a single sitting. No one individual is familiar on a daily basis with the whole of pair distribution functions, Landau theory, picosecond spectroscopy, power spectra, site-specific mutations, the intricacies of receptor pharmacology, and even perhaps thermodynamics. However, each chapter brings its rewards. The book is well produced (my favorite misprint is "which should strong-lycine favor", but this is not representative) by a wide range of experts, contains a mine of information, is actually very readable, and provokes the imagination. Not least, one is prompted to ask what is the thermodynamic basis of receptor and channel function? Maybe this book is just the beginning and a good indication of how complex the question really is.