

Book Reviews

The Oxford Handbook of Computational Linguistics

Ruslan Mitkov (editor)

(University of Wolverhampton)

Oxford: Oxford University Press, 2003,
xx+784 pp; hardbound, ISBN
0-19-823882-7, £76.00, \$150.00

Reviewed by

Peter Jackson

Thomson Legal & Regulatory

This collection of invited papers covers a lot of ground in its nearly 800 pages, so any review of reasonable length will necessarily be selective. However, there are a number of features that make the book as a whole a comparatively easy and thoroughly rewarding read. Multiauthor compendia of this kind are often disjointed, with very little uniformity from chapter to chapter in terms of breadth, depth, and format. Such is not the case here. Breadth and depth of treatment are surprisingly consistent, with coherent formats that often include both a little history of the field and some thoughts about the future. The volume has a very logical structure in which the chapters flow and follow on from each other in an orderly fashion. There are also many cross-references between chapters, which allow the authors to build upon the foundation of one another's work and eliminate redundancies.

Specifically, the contents consist of 38 survey papers grouped into three parts: Fundamentals; Processes, Methods, and Resources; and Applications. Taken together, they provide both a comprehensive introduction to the field and a useful reference volume. In addition to the usual author and subject matter indices, there is a substantial glossary that students will find invaluable. Each chapter ends with a bibliography, together with tips for further reading and mention of other resources, such as conferences, workshops, and URLs.

Part I covers the full spectrum of linguistic levels of analysis from a largely theoretical point of view, including phonology, morphology, lexicography, syntax, semantics, discourse, and dialogue. The result is a layered approach to the subject matter that allows each new level to take the previous level for granted. However, the authors do not typically restrict themselves to linguistic theory. For example, Hanks's chapter on lexicography characterizes the deficiencies of both hand-built and corpus-based dictionaries, as well as discussing other practical problems, such as how to link meaning and use. The phonology and morphology chapters provide fine introductions to these topics, which tend to receive short shrift in many NLP and AI texts.

Part I ends with two chapters, one on formal grammars and one on complexity, which round out the computational aspect. This is an excellent pairing, with Martín-Vide's thorough treatment of regular and context-free languages leading into Carpenter's masterly survey of problem complexity and practical efficiency.

Part II is more task based, with a focus on such activities as text segmentation, part-of-speech tagging, parsing, word sense disambiguation, anaphora resolution, speech recognition, and text generation. (One might wonder why speech recognition is in

Part II rather than Part III, given that the chapter makes no reference to other areas that have appropriated these techniques and applied them elsewhere.) Some of these chapters make the obvious connections to topics in Part I, but others could have done more in this regard. However, there are many forward references to applications in Part III to which these techniques are pertinent.

The levels of treatment accorded to the topics in Part II are perhaps a little mixed, with some being less introductory than others. Thus Carroll's survey of parsing techniques provides an abstract overview for the researcher who already has a firm grasp of the computational issues and can appreciate the differences in parsing strategy and structural bookkeeping that he describes. Similarly, Karttunen's treatment of finite-state technology is more a summary of transducers for NLP than an introduction to the field. This is probably fine for a handbook of this kind, although it might limit the usefulness of these chapters for some students.

By contrast, Mikheev's chapter on segmentation and Voutilainen's chapter on POS tagging provide more background for the general reader and are comprehensible by nonexperts. Each of these chapters provides enough material to get a student started on the conception and planning phases of a segmentation or tagging project. Mitkov's exposition of anaphora resolution is particularly clear, being full of illuminating (and often entertaining) examples that highlight the distinctions between different kinds of anaphora, such as coreferring and non-coreferring. Kittredge's overview of sublanguages and controlled languages is also well organized and a model of clarity.

Middle chapters in Part II concentrate upon technologies to solve somewhat higher-level problems, such as natural language generation, speech recognition, and text-to-speech synthesis. Lamel and Gauvain's treatment of speech recognition is again an overview, rather than an introduction to the field, and is unlikely to be accessible to nonexperts. For example, the cepstral transformation and the Mel scale could be better motivated; neither is formally defined or linked to a glossary entry. Many students of NLP will not be familiar with these concepts and will not understand their importance in linear prediction and filterbank analyses using hidden Markov models.

These fairly specialized topics are then followed by useful chapters on subjects of interest to most computational linguists. Samuelsson's chapter on statistical methods does a very efficient job of imparting the basics of probability theory, hidden Markov models, and maximum-entropy models, together with a little dry humor. Mooney concentrates on the induction of symbolic representations of knowledge, such as rules and decision trees, in his chapter on machine learning. This focus avoids overlap with more statistical learning methods, such as naive Bayes, and allows room for covering case-based methods, such as nearest-neighbor algorithms.

Hirschman and Mani's chapter on evaluation represents a valiant attempt to cover, in a few pages, what remains a neglected topic in computational linguistics and natural language processing. It's a sad fact of life that gold standard-based approaches, such as those used in the Message Understanding and Text Retrieval Conferences, only take one so far in proving the effectiveness of research prototypes or final products. Measures such as precision and recall are useful yardsticks, but the real issue is, what value does the system deliver to an end user? More specifically, what does the system enable a knowledge worker to do that he or she could not do before? Academic researchers are typically not well placed to either pose or answer such questions, but any purveyor of natural language software must somehow address them. The section on evaluation of mature output components is the most relevant here.

McEnery provides an able introduction to corpus linguistics, albeit with a primary focus upon English, and briefly summarizes some of the advances that annotated corpora have enabled. Vossen's chapter on ontologies provides many useful pointers

to resources around the globe and makes an explicit attempt to outline areas of NLP in which such resources can and have been used. However, it is clear that the value of ontological approaches has yet to be fully demonstrated and that many of the tools are still in their infancy. Part II ends with a compact and readable overview of lexicalized tree-adjoining grammars by Joshi, which both motivates the formalism and illustrates its power.

Part III provides overviews of important areas such as machine translation, information retrieval, information extraction, question answering, and summarization. These chapters will be particularly attractive to practitioners in these fields, as they provide succinct and realistic overviews of what can and cannot be achieved by current technology. I confess to having read these chapters first. In fact, it might not be a bad strategy for some readers to dive straight into an application area in which they are particularly interested, and then read other chapters as needed, using the cross-references as a guide.

Machine translation is accorded two chapters, one that discusses the earlier, rule-based approaches and one that deals with more recent, empirical approaches based on parallel corpora. Both chapters give the general reader a good feel for the issues, the strengths and limitations of the various methods, and the kinds of tools that are currently available to assist translators. Somers's brief survey of statistical approaches to MT is particularly insightful on the topic of early successes and subsequent lack of improvement.

In the information retrieval chapter, Tzoukerman, Klavans, and Strzalkowski provide a frank assessment of how little impact natural language processing has had upon current search engine technology, beyond the application of tokenization and stemming rules. Whether attempting to apply WordNet to query expansion or seeking to disambiguate query terms, researchers have typically either failed to deliver improvements or failed to scale complex solutions to applications of commercial value. In attempting to elucidate why this is the case, the authors provide good coverage of the experimental literature and related research systems, such as CLARIT and DR-LINK. They conclude that NLP techniques to date have either been too weak to have a measurable impact or too expensive in terms of effort or computation to be cost-effective.

Grishman's information-extraction chapter provides a clear exposition of two problems: identifying proper names and recognizing events. Grishman provides an overview of the work done under the auspices of the Message Understanding Conferences in these areas, as well as an update on machine-learning approaches to the problem of building extraction patterns. Hearst's chapter on text data mining distinguishes this area from information retrieval and text categorization, linking the field to exploratory data analysis. Further chapters of interest include Hovy on summarization, André on multimodal and media systems, and Grefenstette and Segond on multilingual NLP.

In looking for gaps in the book as a whole, one cannot help noticing that the chapters on ontologies, word senses, and lexical knowledge acquisition (by Matsumoto) are among the few to touch upon semantic information processing. This is in marked contrast to many AI and NLP collections from the 1970s, in which articles on knowledge representation languages and text interpretation schemes abounded. Also absent are connectionist models of speech and language, which were perhaps more popular in the 1980s than they are today. These omissions may reflect a new realism in the field, in which the emphasis is now upon methods that are scalable, less knowledge intensive, and more amenable to empirical evaluation.

Overall, this is an impressive volume that demonstrates just how far the field has progressed in the last decade. During that time, we are fortunate to have seen many

advances in both the theory and practice of computational linguistics research, and one feels that these must be attended by improvements in natural language processing in the near future. When one combines the newer corpus-based approaches with continued advances in algorithms and representations in other areas, and then factors in annual increases in computing power and storage capability, one sees a recipe for further successes on hard problems like speech recognition, machine translation, and broad-coverage parsing.

Peter Jackson is vice-president of research and development at Thomson Legal & Regulatory, where he leads a group that specializes in the retrieval, mining, and classification of legal information. Over the last 20 years, he has published books and papers on expert systems, theorem proving, information extraction, and text categorization. His address is Thomson Legal & Regulatory, D1-N329, 610 Opperman Drive, St. Paul, MN 55123; e-mail: Peter.Jackson@Thomson.Com; URL: <http://members.aol.com/jacksonpe/music1/home.htm>.

Readings in Machine Translation

Sergei Nirenburg, Harold Somers, and Yorick Wilks (editors)

(University of Maryland, Baltimore County, University of Manchester Institute of Science and Technology, and University of Sheffield)

Cambridge, MA: MIT Press, 2003,
xv+413 pp; ISBN 0-262-14074-8, \$55.00,
£36.95

Reviewed by
Elliott Macklovitch
Université de Montréal

This is a wonderful book—and not just for people actively involved in machine translation. Anyone with an interest in the history of computational linguistics will find much to relish and learn from in this weighty collection of articles past. Lest we forget, MT was one of the first nonnumerical applications proposed for the digital computer following the Second World War, and its often tumultuous 50-year history has had a significant impact on the entire field of computational linguistics. Indeed, this very journal can trace its lineage back to the journal whose original title was *Mechanical Translation*.

Though not a proper history of MT, *Readings in Machine Translation* is certainly a historical collection. The editors have sought to bring together in one volume “the ‘classical’ MT papers that researchers and students want, or should be persuaded, to read” (page xi) and that, alas, are often so difficult to find. For this alone, Nirenburg, Somers, and Wilks deserve our gratitude. The volume begins with the famous memorandum that Warren Weaver sent out to some 200 professional acquaintances in 1949, which is generally taken to mark the genesis of machine translation; and the most recent paper included dates back to the fourth MT Summit in 1993. The 36 articles that span the intervening period are said to constitute “MT’s communal inheritance.”

By what criteria were the papers selected? The editors cite three: personal taste; the aforementioned problem of availability, which is certainly a real one, particularly in the case of some of the early classics; and historical significance, which is said to be the main criterion for inclusion in the volume. The articles selected by the editors are supposed to represent “the most important papers from the past 50 years” of MT (p. xi). Well, as criteria go, that certainly sets a high standard! And yet many of these articles seem to meet it with ease. In addition to Weaver’s memorandum and a monumental state of the art published by Bar-Hillel in 1960, both of which absolutely must be read, there is a jewel of a piece written by Victor Yngve in 1957 that argues for what later became known as second-generation MT: systems that analyze the input text into an essentially syntactic intermediate representation that serves as the basis for transfer, rather than applying a bilingual dictionary directly to the input string, as first-generation systems did. There is Martin Kay’s “The Proper Place of Men and Machines in Language Translation”—an MT classic if ever there was one—in which the author derides the pursuit of fully automatic MT, not as a legitimate goal of basic research, but as a strategy promising a short-term solution to the burgeoning demand for translation; in its place, Kay proposes a modest, incremental program of machine aids for human translation: “little steps for little feet.” There is Makoto Nagao’s 1984 paper that launched the entire example-based MT paradigm (still very much alive and kicking), as well as a stellar piece by Jan Landsbergen laying out the advantages for

machine translation of Montague-style semantics (an approach all but moribund now, at least insofar as MT is concerned). One reads these papers today, decades after they were written, and one still cannot help but be impressed.

Needless to say, not all the articles included in *Readings in Machine Translation* come up to this high standard; that would be too much to expect. However, there are a fair number of papers that don't appear to even come close to the editors' stated selection criteria, unless of course one invokes the lame justification of personal taste. I won't bother to name names, out of respect for the elderly and the departed; but most of the papers I have in mind should be fairly obvious to all from a cursory perusal of the table of contents. ("Who is *that*?" is a good first indicator.) In other cases, one wishes the editors had made more liberal use of their prerogative to abridge. There are articles containing long tables filled with obscure codes and idiosyncratic terminology that can't possibly present any interest to the vast majority of contemporary readers.

Another reason for the excessive length of *Readings in Machine Translation* is that the book is divided into three distinct sections, each under the responsibility of one of the editors. The historical section is under Nirenburg's editorship and includes papers up to the late 1960s; Wilks's section is on theoretical and methodological issues; and Somers's is on system design. There are obvious overlaps between these divisions, in the sense that articles included in one section could just as well fit into another. The editors acknowledge this, and in itself it is not very serious. More tiresome, perhaps, is the fact that each section is prefaced by its own introduction (in addition to a common introduction to the entire volume) in which the editors sometimes marshal "their" articles in an attempt to argue for a certain perspective on machine translation. In his introduction, for example, Nirenburg cites numerous, often lengthy passages from the articles by the early MT pioneers that purportedly support his preferred approach to meaning-based MT. Well, maybe they do and maybe they don't; but either way, gathering grist for one's mill has a rather unseemly feel in this context.

A more serious criticism of *Readings in Machine Translation* is that the book is somewhat dated. This is a rather paradoxical charge for a collection of historical articles; what I mean by it is this: By the editors' own admission, the volume took much more time to bring to publication than they had originally anticipated. (In fact, I was sent a preliminary version by the publisher in 1999.) As a result, the editors' assessment of the most significant recent trends in MT is not entirely up to date. In the last few years, for example, there has been an impressive resurgence of activity in machine translation, particularly in the United States, where statistical methods drawn from speech recognition and various techniques borrowed from machine learning have proven remarkably successful. Had the editors been more aware of the profound impact of these new influences on the field, they would perhaps have modified their selection of articles. As it is, only two of the thirty-six papers in the collection explicitly address data-driven or statistical methods in MT: the seminal paper "A Statistical Approach to Machine Translation," published in 1990 by Peter Brown and his colleagues at IBM; and an earlier piece entitled "Stochastic Methods of Mechanical Translation" by Gilbert W. King.

Which brings me to my final criticism of this otherwise wonderful volume. Perhaps you recognized the name of Gil King (as Nirenburg calls him in his introduction), but I confess that I didn't. So I looked him up in John Hutchins's *Machine Translation: Past, Present, Future* and discovered (somewhat to my embarrassment) that King was the director of the project at IBM T. J. Watson Research Center in the late 1950s that eventually produced the Mark I system, later installed at the U.S. Air Force's Foreign Technology Division. Okay. . . . And where did the article included in this collection first appear? To find the answer to that question, and indeed to locate the source of

all thirty-six papers included in *Readings in Machine Translation*, one has to search the source notes that appear at the end of the volume: three dense pages of references whose order doesn't always correspond to that of the articles themselves. It would have been so much easier and more helpful to display this information on the first page of each contribution! Indeed, one wishes the editors had seen fit to include a short introductory note to each article, providing a few words of historical background on the author, or at least his or her affiliation at the time the paper was published.

But these are more or less minor quibbles, and they do not significantly detract from the value of this generous volume: generous in the size of its pages; generous in the quality of the paper and its cloth binding; generous above all in the intellectual caliber of the articles it offers for our delectation. In an age of skyrocketing book prices, MIT Press is to be commended for pricing *Readings in Machine Translation* so affordably.

References

Hutchins, John. 1986. *Machine Translation: Past, Present, Future*. Ellis Horwood, Chichester, England.

Elliott Macklovitch is coordinator of the RALI Laboratory at Université de Montreal and president of the Association for Machine Translation in the Americas (AMTA). His address is RALI-DIRO, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Canada H3C 3J7; e-mail: macklovi@iro.umontreal.ca.

Language Modeling for Information Retrieval

W. Bruce Croft and John Lafferty (editors)

(University of Massachusetts, Amherst, and Carnegie Mellon University)

Dordrecht: Kluwer Academic
Publishers (Kluwer international series
on information retrieval, edited by W.
Bruce Croft), 2003, xiii+245 pp;
hardbound, ISBN 1-4020-1216-0, \$97.00,
£62.00, €99.00

Reviewed by

Paul Thompson

Dartmouth College

The papers in this edited volume are expanded versions of papers originally written for the Workshop on Language Modeling and Information Retrieval held May 31–June 1, 2001, at Carnegie Mellon University in Pittsburgh, Pennsylvania. As the editors note, these papers provide a cross-section of current work on the language-modeling approach to information retrieval, which has become a very active area of research in the past few years. The editors place the papers in this volume in three broad categories: (1) papers addressing the mathematical formulation and interpretation of the language-modeling approach to information retrieval; (2) papers concerned with using the language-modeling approach for ad hoc information retrieval; and (3) papers using the language-modeling approach for other information retrieval application areas, including topic tracking, classification, and summarization. This book provides an excellent introduction to this new field of research by bringing together extended versions of papers from many of the field's leading researchers.

Stated simply, the goal of the language-modeling approach to information retrieval is to predict the probability that the language model of a particular document being considered for retrieval could have generated the user's query. By casting the document retrieval problem in this way, language-modeling techniques that have been developed over many years, particularly for speech recognition, can be applied to document retrieval. On the other hand, at least in this simple statement, the language-modeling approach appears to ignore the concept of relevance. By contrast, the traditional approach to probabilistic information retrieval, as expressed, for example, in the probability-ranking principle, explicitly states that the goal of probabilistic information retrieval is to predict the probability that a document will be judged relevant by a user, taking into account all evidence available to the retrieval system, which then ranks the documents in decreasing order of probability of relevance. As mentioned by the editors in the preface, this concern with the relationship of the language-modeling approach to relevance was one of the underlying themes of the workshop, and the issue is taken up by some of the authors.

The three chapters addressing the mathematical formulation and interpretation of language modeling for information retrieval are Lafferty and Zhai's "Probabilistic Relevance Models Based on Document and Query Generation," Lavrenko and Croft's "Relevance Models in Information Retrieval," and Sparck Jones et al.'s "Language Modeling and Relevance." Taken together, these three chapters illustrate the controversies with respect to the theoretical status of the language-modeling approach. Lafferty and Zhai claim to show an equivalence between the underlying probabilistic semantics

of the language-modeling approach and the standard probabilistic model of information retrieval. In particular they disagree with Sparck Jones et al.'s contention that the language-modeling approach, as usually stated, makes sense only if there is a single relevant document that generates a user's query. Sparck Jones et al., rejecting the simple formulation of the language-modeling approach, take pains to informally sketch what a sounder theoretical construction might be like. The result, although incomplete, is complex. It is unclear whether there are any theoretical advantages in abandoning the standard probabilistic model. Lavrenko and Croft also take up the issue of relevance, introducing the concept of a relevance model for information retrieval, that is, a language model reflecting word frequencies in the class of documents relevant to a particular information need. One of their motivations for introducing relevance models is to overcome one of the major disadvantages of the language-modeling framework: its difficulty in incorporating user interaction. They present many experimental results supporting the retrieval effectiveness of their formal model.

Four chapters discuss language modeling in the context of ad hoc information retrieval: Greiff and Morgan's "Contributions of Language Modeling to the Theory and Practice of IR," Xu and Weischedel's "A Probabilistic Approach to Term Translation for Cross-Lingual Retrieval," Manmatha's "Applications of Score Distributions in Information Retrieval," and Zhang and Callan's "An Unbiased Generative Model for Setting Dissemination Thresholds." Greiff and Morgan argue that statistical estimation in language modeling addresses the bias-variance trade-off and that the importance of language-modeling for information retrieval is that it focuses the attention of information retrieval research on this issue. Thus the good experimental results for the language modeling approach reported throughout this book may be due more to its improved statistical estimation techniques than to the use of language modeling as a theoretical framework.

The remaining three chapters describe applications of language modeling to related information retrieval tasks (i.e., topic tracking, text classification, and summarization): Kraaij and Spitters's "Language Models for Topic Tracking," Teahan and Harper's "Using Compression-Based Language Models for Text Categorization," and Mittal and Witbrock's "Language Modeling Experiments in Non-extractive Summarization." These chapters support the book's premise that the language-modeling approach shows promise not only for ad hoc retrieval, but also for other related information retrieval activities.

Information retrieval and computational linguistics have been more or less closely related fields since the 1950s. Statistical approaches have always played an important role in information retrieval. Within the field of computational linguistics, corpus linguistics has emerged over the past 20 years as an increasingly influential approach. Language-modeling techniques originally developed to support speech recognition are now transforming the field of probabilistic information retrieval. However, as alluded to by Sparck Jones et al., the well-established role of language modeling in speech processing and its exploration in machine translation and summarization are activities fundamentally different from document retrieval. It is to be expected that other areas of human-language technology, such as dialogue modeling and mixed-initiative interaction, might also find more application in information retrieval research.

Paul Thompson is a senior research engineer and lecturer at the Thayer School of Engineering at Dartmouth College. His research interests include probabilistic information retrieval and natural language processing. Thompson's address is Institute for Security Technology Studies, Dartmouth College, 45 Lyme Road, Suite 200, Hanover, NH 03755; e-mail: Paul.Thompson@dartmouth.edu.

Probabilistic Linguistics

Rens Bod, Jennifer Hay, and Stefanie Jannedy (editors)

(University of Amsterdam, University of Canterbury, and Lucent Technologies)

Cambridge, MA: The MIT Press, 2003,
xii+451 pp; hardbound, ISBN
0-262-02536-1, \$85.00; paperbound ISBN
0-262-52338-8, \$35.00

Reviewed by
Patrick Juola
Duquesne University

For nearly 40 years, the generally accepted view among linguists has been that “language” is a categorial structure, with probability relegated to perhaps explaining errors of human “performance.” In the past 10 years, theoretical developments, including example- and statistical-based parsing, the mature development of connectionist natural language processing, and stochastic versions of traditional theories (such as stochastic optimality theory), have amassed much evidence that this view is limiting and that replacing categorial structures with probabilistic reasoning will result in better performance.

A specific counter criticism of the quantitative practices in natural language processing is that the kinds of reasoning employed are unnatural and in many cases ignore important observations about linguistics; examples of this include the use of Markov processes or n -gram statistics to describe phonology or syntax, instead of the more linguistically plausible (and traditional) grammar-based formalisms. In this book, Bod, Hay, and Jannedy present some of the evidence in favor of a role for probability in linguistic reasoning. Moreover, they also argue that incorporating probability theory into linguistic theory is not only legitimate but necessary. Probability theory should be compatible with, instead of a substitute for, more traditional linguistic observations.

The book itself is an outgrowth of a symposium organized by the editors at the Linguistic Society of America 2001 annual meeting held in Washington, D.C. Much of the book is thus dedicated to presenting background material specifically for non-computational linguists, who might not have a strong mathematical and statistical background. Chapter 2, an introduction to the theory of probability that starts more or less at the very beginning, is thus either entirely useless or the most important chapter in the book, depending upon the reader’s background. Fortunately, it is well organized, well written (for those that need it), and easily skippable (for those that don’t).

The rest of the book is taken up by individual chapters by noted experts in the various subfields of psycholinguistics, sociolinguistics, historical linguistics and language change, phonology, morphology, syntax, and semantics. The experts are well-chosen and include leading lights such as Baayen, Jurafsky, Manning, and Pierrehumbert, as well as the editors themselves. Most of the individual chapters are similarly structured; the author(s) observe that categorial theories do not neatly account for observed variation and then present a probabilistic model that accounts for both categorial and the observed continuous data. These models are fairly specific to the problems studied by each expert, and together they make an interesting collection of different ways to solve linguistic problems from a variety of standpoints; students of modeling could do much worse than to simply page through and look at each model in turn to see whether it could be adapted to their own studies.

From a broader perspective, a key aspect of the models is that they better account for the observed behavior of real people and real languages. Although “frequency” has long been known to be a very important variable in much of psychology and in psycholinguistics in particular, these models provide specific mechanisms that illustrate how frequency can produce the known effects. Examples of this include Pierrehumbert’s discussion of category discriminability among phonemes, Jurafsky’s discussion of final-consonant deletion and latency, Mendoza-Denton et al.’s discussion of the sociophonetics of /ay/, and Baayen’s discussion of Dutch linking elements. In each case, the effects of frequency are directly modeled in a probabilistic framework and an appropriate causal role is assigned.

One specific point of interest from a more theoretical viewpoint is Jurafsky’s discussion of the psychological reality of probabilistic models (“Surely you don’t believe that people have little symbolic Bayesian equations in their heads?”). He addresses many of the challenges that traditional linguistics might present to probabilists. The bibliography is extensive, and there is a useful glossary of probabilistic terms to help readers keep definitions in mind.

From the standpoint of computational linguistics, the book is slightly disappointing in not discussing computational processes more, as the reader is (usually) left to infer the exact mechanisms and algorithms used to implement the equations discussed in the book. Another weakness is the lack of discussion of statistical inferences, which are often necessary to interpret the probabilistic models themselves (is it reasonable, for example, to expect readers who actually need chapter 2 to understand generalized linear modeling?). The most significant omission is a discussion of how these individual models interact, either with each other or with more traditional categorial models of other linguistic subfields. These are minor weaknesses in an otherwise significant work. For people who already believe that “probabilistic linguistics” is a sensible and meaningful phrase, this book may be useful as a catalog of how probability theory can be and has been applied to the various linguistic subdisciplines. For those who hold to categorial theories of language, this book will at least provide a single source for some major arguments, evidence, and theories supporting probabilistic processes underlying linguistic competence. And for those who believe that probability is really only a description (a quantification of our ignorance, if you will), this volume neatly summarizes ways in which probability may play a key role in human language processing.

Patrick Juola has been working in computational and statistical applications of psycholinguistics since the early 1990s as a Ph.D. student at the University of Colorado. He is currently an assistant professor of computer science at Duquesne University, 600 Forbes Avenue, Pittsburgh, PA 15282; e-mail: juola@mathcs.duq.edu.