

Briefly Noted

Hidden Markov Models for Speech Recognition

X. D. Huang, Y. Ariki, and M. A. Jack
(University of Edinburgh)

Edinburgh: Edinburgh University Press
(Edinburgh Information Technology Series
7, edited by S. Michaelson and
M. Steedman), 1990, x + 276 pp.
Hardbound, ISBN 0-7486-0162-7 (distributed
in North America by Columbia University
Press), \$60.00

The art of automatic speech recognition has advanced remarkably in the past decade. With the advances in accuracy and scope, there has come, for the time being, a strong convergence on a class of statistical methods based on a structure called a hidden Markov model (HMM). HMM-based systems dominate speech recognition, and success in the speech domain has spawned many attempts to extend HMM methods to related pattern-recognition fields such as document recognition and handwriting recognition.

The book under review answers a clear need. It introduces most of the theory and techniques needed to build a complete HMM-based speech recognition system. Huang, Akiri, and Jack use the first 90 pages to cover general methods of pattern recognition, speech-signal processing, and statistical language modeling. The next 90 pages cover signal quantization and the theory of HMMs for speech recognition. The last 60 pages cover practical issues with examples provided and explained.

The style of the book is very dry but clear; it reads like an abstract of an engineering text. The majority of the pages are thick with equations, but the authors seem to use only as much mathematics as will be needed to actually implement the algorithms. There are few derivations and no proofs. The authors present the basic algorithms and the best algorithms (in their judgment), but offer neither historical perspective nor critical review of current research. The book gives algorithm examples written in a high-level English pseudo-code. These examples are concrete and will be helpful for a novice. The tone of the book is decidedly practical, although it is too concise. The book would need substantial expansion to work as a graduate-level text.

In sum, for the reader-cum-system-builder

who is ready to implement an HMM speech recognition system from the bottom up, this is the best available book that I have seen. Huang, Akiri, and Jack have distilled and presented the essential techniques of the trade.—*Jared Bernstein, SRI International*

Conceptual Information Retrieval: A Case Study in Adaptive Partial Parsing

Michael L. Mauldin
(Carnegie Mellon University)

Boston: Kluwer Academic Publishers (The Kluwer International Series in Engineering and Computer Science: Natural Language Processing and Machine Translation, edited by Jaime Carbonell), 1991, xx + 215 pp.
Hardbound, ISBN 0-7923-9214-0, \$62.50, £43.25, Dfl 145.00

Michael Mauldin has attempted the difficult task of building a full-fledged information retrieval system in the traditional design, but one with a language-understanding flavor. His purpose is to demonstrate that use of semantic knowledge will improve our capability to retrieve information. And he demonstrated his achievement by testing his system, FERRET, against a Boolean retrieval prototype of contemporary design. Unfortunately, his effort is marred by the commonly recognized flaws of the IR experimental methodology. However, given the realities of human interaction, he had little choice but to employ that methodology to convince his target audience.

For the computational linguist, the work is of interest because it demonstrates the practicality of recognized techniques. There is little new here; however, in putting the pieces together, the author has provided us with convincing evidence that a high-level parse can be extremely useful for a number of activities requiring analysis of text. Mauldin includes a catalog of suggested uses in his last chapter.

FERRET employs an adapted version of DeJong's (1979) FRUMP to parse articles from UseNet concerning astronomy. The parse is really a skimming of the text. Complex bits are passed over. Then "sketchy scripts" and case frames are compiled from the resulting conceptual dependencies (CD) (Schank et al. 1975). It was the author's purpose to demonstrate that a wide range of textual domains could be successfully treated in such a way, and the resulting representa-

tion reminds one of a topical index expressed in frames. The technique, by his own admission, works best with "paragraph-sized chunks of text".

As well as the text-derived frames, FERRET allows access to *Webster's Seventh New Collegiate Dictionary* in order to add to the available real-world knowledge. A genetic-like learning algorithm further enhances knowledge at search time by offering a capability similar to relevance feedback in vector searching. Test results indicate that although the algorithm offers other capabilities, its most effective use is to generalize concepts when attempts at matching are to be improved upon. Since the matching mechanism is primitive, allowing only yes/no matches at present, 'learning' serves to amplify answers by allowing adaptation of abstracts relevant to the query. All relevance judgments were made by the author.

Queries were solicited by questionnaire from the UseNet group readers. There were some constraints placed on the subject matter in order to channel the questions toward the material contained in the knowledge base, again narrowing the domain of demonstration. And the contributors of the final questions were 23 in number. Their queries were automatically translated into CD graphs and the matching and learning retrieval process was undertaken, comparisons being made with the Boolean-keyword retrieval system built as a control.

FERRET prevailed by proving to do somewhat better than the traditional system. But it was, as its author himself recognized, just a cut above the keyword system. The semantic representation, although extremely robust, was not suitably expressive to impress with the power of representing meaning. It allowed some advantage in dealing with problems of synonymy and ambiguity but was not capable of coping with paraphrase. Both the matcher and the knowledge representation are slated for further investigation. In the meantime, IR people can point to a successful implementation of a conceptual retrieval system as a beacon, and computational linguists can see a clear use for work that is all too often regarded as the result of iconoclastic academic pursuit, thanks to the author's courage in attempting to implement a holistic conceptual retrieval design as a one-man operation.—*Judith Dick, University of Maryland*

References

- DeJong, Gerald F. (1979). *Skimming stories in real time: An experiment in integrated understanding*. Doctoral dissertation, Department of Computer Science, Yale University.
- Shank, Roger C.; Goldman, Neil M.; Rieger, Charles J. III; and Riesbeck, Christopher K. (1975). *Conceptual Information Processing*. (Fundamental Studies in Computer Science, Volume 3). North-Holland.

Research in Humanities Computing 1: Selected Papers from the ALLC/ACH Conference, Toronto, June 1989

Ian Lancashire (editor)
(University of Toronto)

Oxford: Clarendon Press (Research in Humanities Computing Series, edited by Susan Hockey and Nancy Ide), 1991, xvi + 353 pp. Hardbound, ISBN 0-19-824251-4, no price listed

"*Research in Humanities Computing* is an annual publication which represents the state of the art in humanities computing. Each volume contains a selection of papers presented at the joint annual conference of the Association for Computers in the Humanities (ACH) and the Association for Literary and Linguistic Computing (ALLC). ACH and ALLC are the two major associations for the use of computers in scholarly research and teaching in humanities disciplines such as archaeology, art, history, languages, literature, music, and philosophy. This, the first volume, contains twenty papers from the Dynamic Text ACH-ALLC Conference held in Toronto in June 1989."—*From the publisher's announcement*