base abstraction. Universität Passau Technical Report MIP-8723 (1987).
26. S. Richardson *et al.*, Experiences of developing a large-scale natural language text processing system: CRITIQUE. IBM Research Report RC 13644 (1987).
27. G. Ruge, Experiments in linguistically based term associations. In *Proceedings of RIAO '91*, pp. 528–545. Barcelona, Spain (1991).
28. R. Sacks-Davis *et al.*, Using syntactic analysis in a document retrieval system that uses signature files. In *Proceedings of the 13th International SIGIR Conference on Research and Development in Information Retrieval*, edited J.-L. Vidick, pp. 179–192. Brussels, Belgium (1990).
29. G. Salton, *Text Processing: The Transformation, Analysis and Retrieval of Information by Computer*. Addison-Wesley (1989).
30. G. Salton *et al.*, On the application of syntactic methodologies in automatic text analysis. *Information Processing and Management* **26** (1), 73–92 (1990).
31. C. Schwarz, Content based text handling. *Information Processing and Management* **26** (2), 219–226 (1990).
32. P. Sheridan and A. F. Smeaton, The application of morpho-syntactic language processing to effective phrase matching. *Information Processing and Management* **28** (2).
33. K. Sparck Jones, What exactly should we look to AI, and NLP especially, for? In *Working Notes from the AAAI Spring Symposium on Text Based Intelligent Systems* (1990).
34. M. Wood and I. Sommerville, An information retrieval system for software components. *ACM SIGIR Forum* **22** (3/4), 11–28 (1988).
35. G. P. Zarri, An outline of the representation and use of temporal data in the RESEDA system. *Information Technology: Research and Development* **2** 89–108 (1983).
36. U. Zernik, On 'Diesel Train1 Engines' and 'To Train2 Airline Pilots': tagging word senses in corpus", In *Proceedings of RIAO '91*, pp. 567–585. Barcelona, Spain (1991).

Book Review

MICHAEL L. MAULDIN
Conceptual Information Retrieval – A Case Study in Adaptive Partial Parsing.
Kluwer Academic Publishers, Boston. ISBN 0-7293-9214-0. £43.25.

This book is about a conceptual information retrieval system called FERRET which was created by the book's author. It is stated that the motivation behind building FERRET was to get away from the limitations of traditional keyword-based IR systems and create a system that 'understands' the content of the documents it is retrieving. Using techniques based on the FRUMP news-skimming system (by DeJong), FERRET parses documents and the user's query, trying to fit them to one of a number of predefined scripts. Once a script has been found a case frame is built, and matching is done on these frames. Hence FERRET is restricted to parsing documents it has scripts for, limiting it to a small domain. In FERRET, the author has attempted to improve on FRUMP by using an on-line dictionary to try to give clues about words not in the system lexicon. He has also incorporated learning algorithms to improve the parsing scripts. Based on user reactions to the results of a query, two kinds of algorithm start up, one which tries to generalise a script so that more documents fit a query; and the other which attempts to mutate a script to allow it to understand new concepts. The latter is the most interesting part of FERRET, because if it is genuinely capable of learning to change its parser this approach to IR would have great potential.

The book is made up of seven chapters describing the motivations behind FERRET, a review of the literature, the system itself, testing, conclusions and future work. According to the publisher's comments on the back of the book, chapters 2 and 3 'serve as an excellent reference in the fields of NLP, IR and AI'. I am not really qualified to comment on the NLP and AI review, but I found the review of IR work, especially that of keyword systems, to be rather thin, summing up weighted keyword retrieval research in just under three pages. The description of the FERRET system in chapters 4 and 5 is well done, with many diagrams and examples to explain how everything works. Chapter 6 outlines the testing that was performed on FERRET. The centre piece of the testing is a comparative study between FERRET and a keyword-based IR system. Surprisingly, the author chose to pit FERRET against a Boolean system. Perhaps this is because each system's retrieval output is an unordered set of documents, thus making comparisons easier. However, it is generally accepted that weighted-keyword IR systems with a good stemming algorithm and relevance feedback outperform Boolean. So the significance of the result (FERRET winning easily), is lessened in the light of that knowledge. The more impressive result was the improvement in recall that the learning algorithms produced, raising it by 30%, although all this improvement came from the generalisation algorithm, with the mutation algorithm contributing nothing. Whether this was due to a failing in the algorithm or just an insufficient amount of training data is not clear.

The style, layout and order of the book are all good, making it very readable. Certainly if you want to know about FERRET, this is the book for you. As a book about conceptual information retrieval then, remembering the reservations outlined above, I would say that for someone relatively new to NLP and AI (like me) it provides a good introduction to the subject.

MARK SANDERSON
Glasgow

Announcement

7–9 OCTOBER 1992

11th International Conference on the Entity Relationship Approach, Karlsruhe, Germany

The Entity Relationship (ER) approach is extensively used in many database and information system design methodologies and has become a *de facto* standard of most manual and computerised design tools. Continuing its tradition, the 11th conference will provide an international and interdisciplinary forum in which researchers and practitioners can share novel research, tool developments and management experiences. The conference will consist of presented papers, invited papers, tutorials, tool demonstrations, and panel sessions.

Topics:

- Conceptual modelling and database design
- Federated information bases
- Innovative applications of the ER approach
- Security and integrity techniques
- Practical issues in database development
- Process modelling, characterisation and implementation
- User interfaces and multimedia databases
- Re-engineering of databases
- Quality control of database design aspects
- Automated design of information systems
- Functional design
- Query languages

For further information contact:

Professor Dr W. Stucky, University of Karlsruhe, AIFB, P.O. Box 6980, D-W-7500 Karlsruhe, Germany. Tel.: +49-721-6083-812. Fax: +49-721-6937-17. email: WST@AIFB.UNI-KARLSRUHE.DE.

Downloaded from https://academic.oup.com/comjnl/article/35/3/278-a/525676 by guest on 11 April 2024