2. Review of Some of the Papers of the Conference

2.1. Probability Theory, Possibility Theory and Belief Functions

Using probability theory, possibility theory and belief functions to manage uncertainty was one of the main topics at this conference. A number of papers on this topic was presented. Philippe Smets' paper 'Probability of deductibility and belief functions' gives a fresh interpretation of the Dempster-Shafer theory based on the probability of deductibility. Shenoy's paper 'Information sets in decision making' offers a roll-back procedure for solving game trees using local computation. Wilson's two papers discuss the relation between the Dempster-Shafer theory and decision making respectively default logic. Liu, Bundy and Robertson's paper 'Recovering incidence functions' develops a method for creating a set of medium level elements and its probability assignment when a lower bound of probability on a set of axioms is given. Two papers from the I.R.I.T. center in France led by Dubois and Prade are concerned with 'Possibilistic logic: from nonmonotonicity to logic programming' and 'Uncertainty in constraint satisfaction problems: a probabilistic approach', respectively. Gebhardt and Kruse's paper 'A new approach to semantic aspects of possibilistic reasoning' is about providing well-formed semantics of possibilistic reasoning in knowledge-based systems. Moral's paper 'A formal language for convex sets of probabilities' continues his research on convex sets.

2.2. ATMS and Belief Revision

Assumption-based truth maintenance systems and belief revision was also one of the main topics at the conference. Several papers in the proceedings are concerned with this topic. Among them J. Kohlas's paper 'Symbolic evidence, arguments, supports and valuation networks' presented a method to obtain symbolic supports for hypothesis in assumption-based knowledge base. Then the degree of support can be easily calculated from the corresponding symbolic arguments. This analysis is identical to the discussion in Liu, Bundy and Robertson's paper 'On the relations between incidence calculus and ATMS', apart from that in the latter paper the support arguments are the possible worlds in incidence calculus.

3. Summary

Apart from the topics mentioned above, there were also some other issues like modal logic and handling inconsistency, such as Gabbay and Hunter's 'Restricted access logics for inconsistent information' and 'Making inconsistency respectable: part2-meta-level handling of inconsistent data', default logic, conditional independence, etc.

The emphasis of the conference was on the comparison and integration of different techniques.

The conference was a great success in the sense that it brought most of the active researches in Europe together to present and exchange their recent results.

KB&KS'93: 1st Conference and Workshop on Building and Sharing of Very Large-Scale Knowledge Bases

Tokyo, 1–5 December 1993

N.J.I. Mars
Faculty of Informatics, University of Twente, P.O. Box 217, 7500 AE Enschede, Netherlands

The first Conference and Workshop on building and sharing of very large-scale knowledge bases (KB&KS'93) were held from 1 to 5 December 1993, in Tokyo, Japan. In this note a brief description is given of the presentations at that conference and workshop. It appears that all presented papers

were invited contributions; no open call for submissions has appeared. The conference and workshop attracted some 400 participants, of which about 80% were Japanese.

Opening addresses were given by Eiji Kageyama, President of Japan Information Processing Development Center; Osamu Watanabe, Director General, Machinery and Information Industries Bureau, Ministry of International Trade and Industry, Japan; Yasuhiro Kato, Deputy Director-General, Science and Technology Agency, Japan; and Su-Shing Chen, Program Director, National Science Foundation, U.S.A.

The first keynote address under the title 'From Infancy to Adolescence' was given by Kazuhiro Fuchi, Chairperson of the KB&KS'93 Organizing Committee. Fuchi emphasized that technology develops in steps, not gradually. Even though no technology has developed as rapidly as computer technology, it is still subject to that gradual development. For AI, we are just at the beginning of the development. His dream is that AI will approach the intelligence of humans.

The Japanese Fifth Generation Computer System Project was aimed at enabling technology for AI, namely parallel computers; so is the Electronic Dictionary Research Project. These projects themselves will not lead to the intelligence envisaged by Fuchi. More emphasis should be placed on sharing, rather than on building, very large knowledge bases. These knowledge bases, in Fuchi's opinion, should be freely available to all of mankind!

The second keynote address, entitled 'Very Large-Scale Knowledge Bases Embodying Intelligence Space' was presented by Toshio Yokoi, Chairperson of the KB&KS'93 Steering and Program Committee. Yokoi, who heads the Electronic Dictionary Research Project, gave three reasons for building very large knowledge bases: 1) as a technology in itself, for building a knowledge infrastructure; 2) as an approach to pursuing a breakthrough in Artificial Intelligence, in particular, by making AI into an empirical science; and 3) as a movement towards semantic (as opposed to syntactic) information processing.

He saw four areas for research in large knowledge bases: 1) development of common sense knowledge bases (as in the Cyc-project); 2) lexical knowledge bases (as in the EDR project); 3) common knowledge bases in engineering fields; and 4) distributed cooperating systems. The knowledge

bases that are to be build should 1) be open; 2) contain all knowledge common in a chosen field; 3) be applicable to various representation methods; 4) be able to evolve; and 5) be able to grow to a very large scale.

The knowledge infrastructure which would result from the availability of these very large knowledge bases can be characterized by embodying all knowledge, consist of many interconnected distributed knowledge bases, and be owned in various ways, from proprietary to public domain. Realisation of this ideal will take the better part of the twenty-first century, according to Yokoi.

Very large knowledge bases obviously should function as repositories of knowledge. Additionally, they should support knowledge acquisition, and support knowledge base usage (presumably by containing meta-knowledge). Knowledge representation media should (ultimately) range from natural language, through formal representation languages, to pictures and sound.

In the short run, the research areas identified by Yokoi as essential for large-scale knowledge acquisition were: natural language processing, corpus-based language processing, case-based knowledge processing, and large-scale ontologies.

The first series of regular presentations was grouped under the title 'Academic and Social Demands for KB&KS'. In this category, two papers were presented. 'KB&KS as a New Economic and Social Infrastructure' was presented by Ken-ichi Imai, Director of Research, Stanford Japan Center, Japan, who took an economic perspective on large knowledge bases. He compared the speed of acceptance of canals, railways, and roads, and predicted on that basis that we are only at the dawn of the information age.

'On the Nature of Human Knowledge' was the title of the lecture (in the literal sense) by Norio Fujisawa, Director, Kyoto National Museum, and Professor Emeritus, Kyoto University, Japan. It provided a philosophical perspective on the concept 'knowledge', all the way from Plato and Aristotle.

The session on 'Language Technology and Science' contained five papers. The talk 'Current Status and Future Trends of Natural Language Processing' by Makoto Nagao, Professor, Kyoto University, Japan, provided some insight into work done by Nagao and his group, in particular on the analysis of Japanese.

The paper 'Analysis and Generation Technologies' by Hozumi Tanaka, Professor, Tokyo Institute of Technology, Japan, discussed efficiency aspects of various parsing techniques; the analysis of ill-formed sentences (using standard bottom-up analysis followed by top-down error recovery, following Mellish); probabilistic parsing, by attaching probabilities to the grammar rules or to the actions in the LR-table; and language generation.

A very stimulating talk 'Towards Automated Knowledge Acquisition' was given by by Yorick Wilks, Professor, University of Sheffield, U.K. He proposed more work be done in automating knowledge acquisition by extending existing techniques for natural language analysis. In particular, he recommended more work in lexicon construction from machine-readable dictionaries, and information extraction into preformed templates from corpora.

His thought-provoking talk addressed questions like 'Is all knowledge template-like?' and 'Can recall and precision of template-filling knowledge extraction be made commercially attractive?' His most controversial issue was his recommendation of natural language, in a restricted and sanitised form, as a knowledge representation formalism. To achieve that goal, people from various linguistic backgrounds should get together in the development of ontologies, which so far are too language and culture dependent. Another fascinating proposal was the automatic extraction of ontologies from large text corpora.

In the talk 'Acquisition and Exploitation of Textual Resources for NLP' Susan Armstrong-Warwick, Professor, University of Geneva, Switzerland, presented the paradigm shift that has been occurring over the last few years in computational linguistics into corpus-based work. She discussed some of the goals of corpus-based computational linguistics, tools available, and the problems of the limited availability of corpora.

Finally, in the talk 'Knowledge Based Processing in Machine Translation' Jun-ichi Tsujii, Professor, The University of Manchester Institute of Science and Technology, U.K., discussed the use of knowledge in machine translation.

In the session 'Knowledge Technology and Science', five talks were given. Under the title 'How Can People Share Large Knowledge Bases' Setsuo Ohsuga, Professor at the University of Tokyo, Japan discussed the physical organization of cooperating knowledge bases, illustrated for the case of chemicals design.

William R. Swartout, Professor, University of Southern California, U.S.A. discussed the well-known USA Knowledge Sharing Effort, with its four subgroups (Interlingua, KRSS, KQML and Sharable Reusable Knowledge Bases). Bob J. Wielinga, Professor, University of Amsterdam, the Netherlands discussed the European (or at least: Amsterdam) activities in the same area. By creatively reinterpreting prior work in the new light of the knowledge sharing and reuse activities, he was able to make the gap between the activities in the USA and Europe seem less dramatic.

In the talk 'Viewing Data through a Knowledge Representation Lens' Ronald J. Brachman, Department Head, AT&T Bell Laboratories, U.S.A. emphasized the need for the AI community not to neglect the vast amounts of data (as opposed to knowledge) available in the commercial and scientific world. Interactive exploration of large data repositories were described as promising areas of research, illustrated through AT&T's work in the IMACS project.

In the presentation 'Knowledge Acquisition and Ontology' Riichiro Mizoguchi, Professor, Osaka University, Japan, discussed Japanese work on building ontologies.

The final session 'Sharable Knowledge Sources' united a number of loosely-coupled topics. 'The Future of Academic Information Service' by Hisao Yamada, Professor and Director of Research and Development Department at the National Center for Science Information Systems, Japan described (mainly) bibliographic data bases available to scholars in Japan, and their ever increasing importance.

The talk 'International Activities Aiming at Sharing Linguistic Resources' by Antonio Zampolli, Professor, University of Pisa, Italy provided an overview of European activities aimed at providing researchers access to text corpora, including lexica, while the companion talk 'The Role of the Text Encoding Initiative in Creating, Maintaining and Using Well-Documented and Multi-Purpose Electronic Resources' by Susan Hockey, Professor and Director of the Center for Electronic Texts in the Humanities, U.S.A. emphasized the importance of the coding and documentation schemas proposed in the Text Encoding Initiative.

Finally, in his presentation 'Cyc: Priming the

Knowledge Sharing Pump' Douglas B. Lenat, Director of the Cyc Project, Microelectronics and Computer Technology Corporation, U.S.A. discussed the recent progress in this ambitious effort to codify a large fraction of human 'consensus' common sense. Currently, the Cyc knowledge base stands at some 50.000 concepts, and 2 million assertions.

The Workshop associated with the Conference (with the same participants) was also organized in sections. The first section was 'Knowledge Sharing'. It contained four papers. In 'Toward the Knowledgeable Community' Toyoaki Nishida and Hideaki Takeda (Advanced Institute of Science and Technology, Nara, Japan) discussed the advantages of knowledge sharing over knowledge accumulation. To empirically verify their intuitions on the desirable architecture of a distributed system of knowledge-based (and other) agents, the authors have implemented a prototype distributed information system, containing knowledge about touristic and geographic aspects of a region of Japan. Their ultimate application is to be in engineering design.

The talk 'Knowledge Sharing in Integrated User Support Environments: Applications, Frameworks, and Infrastructure' Robert Neches (University of Southern California, U.S.A.) inevitably overlapped considerably with Swartout's talk at the main conference in describing the USA Knowledge Sharing Effort. Additionally, Neches showed some of his work (with Georgia Tech) on developing a sharable knowledge base in the area of human-computer interface design.

'Context: A Real Problem for Large and Sharable Knowledge Bases' by Bob Jansen (CSIRO, Australia) was an amateur's attempt at summarizing the philosophical and linguistic literature on context, and thus bound to fail.

Dimitris Karagiannis (University of Vienna, Austria) and Knut Hinkelmann (DFKI, Germany) discussed ways to increase the efficiency of large rule-based system by selectively creating subsets of the rules to be sued in their paper 'Sharing of Very Large-Scale Knowledge Bases: A Rule-Selection Approach'. Because they lack a sizeable knowledge base for experiments, questions about the scalability of their approach to realistically sized knowledge bases remained open.

In the paper 'A New Framework of Very Large Knowledge-Bases' by Kazumasa Yokota and

Akira Aiba (Institute for New Generation Computer Technology, Japan) a highly abstract architecture for cooperating heterogeneous knowledge-based systems was presented. Finally, in his paper 'I Like Mokum' R.P. van de Riet (Vrije University, the Netherlands) discussed the benefits computer science may have from work done in linguistics, in particular in the area of disambiguating specifications given in natural language.

A feature of this workshop that was a welcome change from the conventional AI meetings was the attendance of several representatives of the database community, who contributed their experience in the problems of large-scale data storage. John Mylopoulos and Vinay K. Chaudhri of the University of Toronto, Canada discussed the topic 'Adapting Database Implementation Techniques to Manage Very Large Knowledge Bases' and Jiawei Han (Simon Fraser University, Canada) et al. 'Knowledge Discovery in Object-Oriented and Active Databases'.

The third session was entitled 'Knowledge Representation' although only one paper really fell into that category. Hideyuki Nakashima (ETL, Japan) described the importance of context to increase efficiency of machine reasoning in his talk 'Context Reflection for Flexible Knowledge Representation'.

In 'The Role of Ontologies in Structuring Large Knowledge Bases' Nicolaas J.I. Mars (University of Twente, the Netherlands) discussed three case studies in ontology development, and the lessons learned from them.

Finally, Tim Finin (University of Maryland, USA) described the KQML formalism in detail in this presentation 'A Knowledge Query and Manipulation Language for Intelligent Agent Interoperability'.

The fourth session of the workshop was devoted to the topic 'Natural Language Processing and Lexical Knowledge'. Four papers were presented here. The paper 'Computational Science, Cognitive Science, and Conceptual Science: Exploiting Constraints for Multi-lingual Knowledge-based Systems' by Robert C. Berwick (Massachusetts Institute of Technology, U.S.A.) was a rather speculative approach to exploit similarities in the deep structure of disparate languages like English and Japanese, presented in the worst tradition of AI: look what I have done for the following two examples of word pairs; now you, reader, do the

rest of the trick for the other 200000 word pairs...

Hiroshi Yasuhara (EDR, Japan) described in the paper 'Text Compiler and Concept-Tagged Corpus' work done at the Electronic Dictionary Research Project to use a large corpus of disambiguated natural language texts as a knowledge base.

Nancy Ide (Vassar College, U.S.A.) and and Jean Veronis (CNRS and University of Provence, France) discussed the feasibility of extraction useful knowledge from machine dictionaries, and concluded that the original ideas in this area were somewhat naive.

Finally, 'Production of Machine Translation Dictionary with Frequency Information' by Tatsuo Ashizaki (The Japan Information Center of Science and Technology, Japan) showed how to add frequency information to translation dictionaries.

The final session of the workshop had three papers. James A. Hendler and William A. Andersen (University of Maryland, U.S.A.) described 'Massively Parallel Matching of Knowledge Structures', in particular on the Connection Machine.

In 'Toward Knowledge Intensive Engineering' Tetsuo Tomiyama, Takashi Kiriyama and Yasushi Umeda (University of Tokyo, Japan) showed the advantages of bringing large amounts of engineering knowledge to bear on design tasks, and dis-cussed approaches to acquire such large knowledge bases, through machine observation of design processes carried out by human designers.

Finally, in the paper 'Building and Sharing Large Knowledge Bases in Molecular Genetics' Francois Rechenmann (INRIA, France) discussed one of the very few very large knowledge bases with a real practical use. The knowledge base described by him plays an essential role in the integration of experimental results obtained in a large number of research institutes involved in genetics studies.

In summary, the KB&KS'93 conference and workshop showed that, at long last, attention is being paid to scaling issues in knowledge bases; that a carefully thought-out set of concepts (an ontology) seems to be a requirement for building well-structured knowledge bases; that interaction with the database community may be fruitful; and that as yet very few large size knowledge bases exist.

The organisers of the conference hope that this conference marks the start of a series of conferences, but so far, no decision on the date and location of a successor has been taken.

The proceedings of the conference and the workshop are:
Knowledge Building and Knowledge Sharing, edited by K. Fuchi and T. Yokoi, 1994, IOS Press, Amsterdam, ISBN 90 5199 184 3, US$99.

Report on the 1993 Bolzano International School in Philosophy and Artificial Intelligence: NLP and Multilingualism

Sabine Geldof

Vrije Universiteit Brussel, Laboratory for Artificial Intelligence, Pleinlaan 2, B-1050 Brussels, Belgium. Email: sabine@arti.vub.ac.be

We will present the important topics dealt with by the lecturers, relate them to each other, we will expand on the aspects of these topics that relate to linguistics and artificial intelligence and eventually relate them to past or ongoing work at the VUB Artificial Intelligence Laboratory. We are particularly interested in the interaction between artificial intelligence and natural language processing: to what extent are similar problems encountered and how are they tackled in each field?

1. Introduction

The 1993 Bolzano International School in Philosophy and Artificial Intelligence: Natural Language Processing and Multilingualism[1] was directed by Oliviero Stock. Its subtitle 'The Perspec-

---

[1]December 6–10, 1993.