KNOWLEDGE ACQUISITION TOOLS *for* EXPERT SYSTEMS

KNOWLEDGE-BASED SYSTEMS

VOL. 2

Knowledge Acquisition Tools for Expert Systems

Knowledge-Based Systems
Volume 2

edited by

J. H. Boose
Boeing Computer Services,
P.O. Box 24346,
Seattle, Washington 98124–0346,
U.S.A.

and

B. R. Gaines
Department of Computer Sciences,
University of Calgary,
Calgary, Alberta T2N 1N4
Canada

1988

ACADEMIC PRESS

Harcourt Brace Jovanovich, Publishers
London San Diego New York Berkeley Boston
Sydney Tokyo Toronto

ACADEMIC PRESS LIMITED
24/28 Oval Road
LONDON NW1 7DX

United States Edition published by
ACADEMIC PRESS INC.
San Diego, CA 92101

British Library Cataloguing in Publication Data
Is available

ISBN: 0-12-115920-5

Typeset by The Universities Press (Belfast) Ltd
Printed in Great Britain at the University Press, Cambridge

Knowledge Acquisition Tools for Expert Systems

Knowledge-Based Systems

One of the most successful and engaging initiatives in Artificial Intelligence has been the development of knowledge-based systems (or, expert systems) encoding human expertise in the computer and making it more widely available. Knowledge-based system developments are at the leading edge of the move from information processing to knowledge processing in Fifth Generation Computing.

The Knowledge-Based Systems Book Series publishes the work of leading international scientists addressing themselves to the spectrum of problems associated with the development of knowledge-based systems. The series will be an important source for researchers and advanced students working on knowledge-based systems as well as introducing those embarking on expert systems development to the state-of-the-art.

Volume 2 has been compiled from the International Journal of Man–Machine Studies:

Volume 26 Number 1 January 1987
Volume 26 Number 2 February 1987
Volume 26 Number 4 April 1987

Volume 27 Number 2 August 1987
Volume 27 Number 3 September 1987

Contributors

G. ABRETT *BBN Laboratories, 10 Moulton Street, Cambridge, MA 02238, U.S.A.*
A. M. AGOGINO *Department of Mechanical Engineering, University of California, Berkeley 5136, Etcheverry Hall, Berkeley, CA 94720, U.S.A.*
J. H. ALEXANDER *Computer Research Laboratory, Tektronix Laboratories, P.O. Box 500, Beaverton, OR 97077, U.S.A.*
J. BENTOLILA *Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
J. H. BOOSE *Boeing Computer Services, P.O. Box 24346, Seattle Washington 98124–0346, U.S.A.*
J. M. BRADSHAW *Knowledge Systems Laboratory, Boeing Advanced Technology Center 77–64, Boeing Computer Services, P.O. Box 24346, Seattle, WA 98124, U.S.A.*
G. H. BREAUX *Carnegie Group Inc., 650 Commerce Court, Station Square, Pittsburgh, Pa 15219, U.S.A.*
J. A. BREUKER *University of Amsterdam, Herengracht 196, 1016 BS Amsterdam, The Netherlands*
M. H. BURSTEIN *BBN Laboratories, 10 Moulton Street, Cambridge, MA 02238, U.S.A.*
P. R. COHEN *Experimental Knowledge Systems Laboratory, Department of Computer and Information Science, University of Massachusetts, Amherst, Massachusetts 01003, U.S.A.*
D. M. COMBS *Medical Computer Science Group, Knowledge Systems Laboratory, Departments of Medicine and Computer Science, Stanforth University School of Medicine, Stanforth, California 94305–5479, U.S.A.*
N. M. COOKE *Computing Research Laboratory and Department of Psychology, New Mexico State University, Las Cruces, NM 88003, U.S.A.*
P. DE KLERK *Carnegie Group Inc., 650 Commerce Court, Station Square, Pittsburgh, Pa 15219, U.S.A.*
J. DIEDERICH *Research Division "Expert Systems", Institute for Applied Information Technology, German Research Institute for Mathematics and State Processing, Schloß Birlinghoven, Postfach 1240, D-5205, Sankt Angustin 1, West Germany*
D. EHRET *Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
L. ESHELMAN *Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
L. M. FAGAN *Medical Computer Science Group, Knowledge Systems Laboratory, Departments of Medicine and Computer Science, Stanforth University School of Medicine, Stanforth, California 94305–5479, U.S.A.*
M. J. FREILING *Computer Research Laboratory, Tektronix Laboratories, P.O. Box 500, Beaverton, OR 97077, U.S.A.*
B. R. GAINES *Department of Computer Science, University of Calgary, Calgary, Alberta T2N 1N4, Canada*
W. A. GALE *AT & T. Bell Laboratories 2C278, 600 Mountain Avenue, Murray Hill, NJ 07974, U.S.A.*
C. GARG-JANARDIN *School of Industrial Engineering, Purdue University, West Lafayette, Indiana 47907, U.S.A.*
S. GENETET *Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
M. GRIMES *Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
T. R. GRUBER *Experimental Knowledge Systems Laboratory, Department of Computer and Information Science, University of Massachusetts, Amherst, Massachusetts 01003, U.S.A.*

S. A. HAYWARD *STC Technology Ltd., Six Hills House, London Road, Stevenage, Herts., SG1 1YB, U.K.*
R. L. JOSEPH *Carnegie Group Inc., 650 Commerce Court, Station Square, Pittsburgh, Pa 15219, U.S.A.*
G. S. KAHN *Carnegie Group Inc., 650 Commerce Court, Station Square, Pittsburgh, Pa 15219, U.S.A.*
C. M. KITTO *Knowledge Systems Laboratory, Boeing Advanced Technology Center, Boeing Computer Services, P.O. Box 24346, Seattle, WA 98124, U.S.A.*
G. KLINKER *Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
S. MARCUS *Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
M. MAY *Research Division "Expert Systems", Institute for Applied Information Technology, German Research Institute for Mathematics and Data Processing, Schloß Birlinghoven, Postfach 1240, D-5205, Sankt Angustin 1, West Germany*
J. McDERMOTT *Department of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
J. E. McDONALD *Computing Research Laboratory and Department of Psychology, New Mexico State University, Las Cruces, NM 88003, U.S.A.*
S. L. MESSICK *Computer Research Laboratory, Tektronix Laboratories, P.O. Box 500, Beaverton, OR 97077, U.S.A.*
E. A. MOORE *Applicon/Schlumberger, 4251 Plymouth Road, Ann Arbor, MI 48105, U.S.A.*
K. MORIK *Technical University Berlin, Institute for Applied Computer Science, Computer-based Informations Systems, Project KIT-Leaner, Sekr. FR 5–8, Frankslinstr. 28/29, D-1000, Berlin 10, West Germany*
M. A. MUSEN *Medical Computer Science Group, Knowledge Systems Laboratory, Departments of Medicine and Computer Science, Stanforth University School of Medicine, Stanforth, California 94305–5479, U.S.A.*
E. P. O. PLANTINGA *Department of Computer Science, Redeemer College, Ancaster, Ontario, L9*
A. RAPPAPORT *Neuron Data, 444 High Street, Palo Alto, CA 94301, U.S.A.*
S. REGOCZEI *Computer Studies, Trent University, Peterborough, Ontario, K9J 7B8, Canada*
S. REHFUSS *Computer Research Laboratory, Tektronix Laboratories, P.O. Box 500, Beaverton, OR 97077, U.S.A.*
I. RUHMANN *Research Division "Expert Systems", Institute for Applied Information Technology, German Research Institute for Mathematics and Data Processing, Schloß Birlinghoven, Postfach 1240, D-5205, Sankt Angustin 1, West Germany*
G. SALVENDY *School of Industrial Engineering, Purdue University, West Lafayette, Indiana 47907, U.S.A.*
M. L. G. SHAW *Department of Computer Science, University of Calgary, Alberta, Canada T2N 1N4*
E. H. SHORTLIFFE *Medical Computer Science Group, Knowledge Systems Laboratory, Departments of Medicine and Computer Science, Stanforth University School of Medicine, Stanforth, California 94305–5479, U.S.A.*
S. J. SHULMAN *Computer Research Laboratory, Tektronix Laboratories, P.O. Box 500, Beaverton, OR 97077, U.S.A.*
M. TAN *Department of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.*
B. J. WIELINGA *University of Amsterdam, Herengracht 196, 1016 BS Amsterdam, The Netherlands*

Preface

The initial success of expert system developments and the development of a number of reasonably domain-independent software support systems for the encoding and application of knowledge have opened up the possibility of widespread usage of expert systems. In particular, Fifth Generation Computing System development programs worldwide assume this will happen and are targeted on knowledge processing rather than information processing. However, what Feigenbaum has termed *knowledge engineering,* the reduction of a large body of knowledge to a precise set of facts and rules, has already become a major bottleneck impeding the application of expert systems in new domains. We need to understand more about the nature of expertise in itself and to be able to apply this knowledge to the elicitation of expertise in specific domains.

The problems of knowledge engineering have been stated clearly:

> "Knowledge acquisition is a bottleneck in the construction of expert systems. The knowledge engineer's job is to act as a go-between to help an expert build a system. Since the knowledge engineer has far less knowledge of the domain than the expert, however, communication problems impede the process of transferring expertise into a program. The vocabulary initially used by the expert to talk about the domain with a novice is often inadequate for problem-solving; thus the knowledge engineer and expert must work together to extend and refine it. One of the most difficult aspects of the knowledge engineer's task is helping the expert to structure the domain knowledge, to identify and formalize the domain concepts." (Hayes-Roth, Waterman & Lenat 1983)

The knowledge acquisition bottleneck has become the major impediment to the development and application of effective knowledge-based systems. Many research groups around the world have been working on knowledge acquisition methodologies, techniques and tools to overcome this problem. In 1985, members of a number of these groups realized that there had been rapid progress in knowledge acquisition research and application. However there was subtantial duplication of effort and limited communication between researchers, and therefore it would be valuable for a workshop to be held that would encourage the sharing of results and experience.

The American Association for Artificial Intelligence agreed to sponsor such a workshop. John Boose of Boeing Advanced Technology Centre and Brian Gaines of the Knowledge Science Institute at the University of Calgary agreed to organize it. Other researchers agreed to contribute effort to the organization and refereeing of papers, resulting in a program and local arrangements committee of: Jeffrey Bradshaw, Boeing Advanced Technology Centre, William Clancey, Stanford University, Catherine Kitto, Boeing Advanced Technology Centre, Janusz Kowalik, Boeing Advanced Technology Centre, John McDermott, Carnegie-Mellon University, Ryszard Michalski, University of Illinois, Art Nagai, Boeing Advanced Technology Centre, Gavriel Salvendy, Purdue University, and Mildred Shaw, University of Calgary.

The response to the call for papers for the Workshop on Knowledge Acquisition for Knowledge-Based Systems (KAW) was overwhelming. The intention was to hold a discussion-intensive meeting of some 35 highly involved researchers. In practice over 120 papers were submitted and some 500 applications to attend were received from about 30 countries. Apart from increasing the refereeing and organizational problems beyond all expected bounds, this response indicated the magnitude and impact of the knowledge acquisition bottleneck and the worldwide interest.

These submissions resulted in 60 people attending the first *Knowledge Acquisition for Knowledge-Based Systems Workshop* (*KAW*) from November 3–7, 1986, at the Banff Centre in Banff, Canada. Each of the 120 papers submitted was refereed by five to seven referees and 42 papers were finally selected. Much of the rejected material was of high-quality and it would have been possible to base a major conference on the material and requests to attend. However, it was decided that the priority at that stage should remain that of establishing in-depth communication between research groups.

It was also clear that it was important to disseminate the workshop material as widely as possible, and arrangements were made to publish revised versions of the papers in the *International Journal of Man–Machine Studies* after the Workshop. These papers have now been gathered together as the first two volumes of the *Knowledge-Based Systems* series.

The table below shows the format of the first KAW. It was very effective in establishing a network linking the community of knowledge acquisition researchers

Knowledge Acquisition for Knowledge-Based Systems *AAAI Workshop, Banff, November 1986*	
Structure	Residential workshop Accomodation, meals and sessions together Attendance limited to 60 (originally 35) 120 papers submitted, 43 accepted Several hundred requests to attend
Overview/ Summary Papers	Gaines—Overview of Knowledge Acquisition Clancey—Cognition and Expertise McDermott—Interactive Interviewing Tools I Boose—Interactive Interviewing Tools II Salvendy—Analysis of Knowledge Structures Michalski—Learning
Mini-Conference	Cognition & Expertise 6, Learning 8 Analysis of Knowledge Structures 7 Interactive Interviewing Tools 16
Workshops on Major Issues	Cognition & Expertise Interactive Interviewing Tools Learning Knowledge Representation
Panels on other Issues	Knowledge Acquisition Methodology/Training Reasoning with Uncertainty
Papers and Books	Preprint volume of all papers to attendees Four special issues of IJMMS in 1987 Two books, Academic Press 1988

worldwide. It resulted in two further KAWs in 1987, a second one at Banff again sponsored by the American Association for Artificial Intelligence (AAAI), and the first European KAW (EKAW) in London and Reading, England, sponsored by the Institute of Electrical Engineers. Papers from these workshops have again been published in the *International Journal of Man–Machine Studies* and constitute the third volume of the *Knowledge-Based Systems* series.

In 1988, the third AAAI-KAW was held at Banff in November with a theme of integration of methodologies, and the second EKAW was held at Bonn, Germany, in June with sponsorship from the Gesselschaft für Mathematik and the German Chapter of the ACM. A specialist workshop on the Integration of Knowledge Acquisition and Performance tools was held at the AAAI Annual Conference in St Paul in August. Sessions and tutorials on knowledge acquisition have become prevalent at a wide variety of conferences concerned with knowledge-based systems worldwide.

These two volumes based on the first AAAI-KAW at Banff contain a wide range of material representing foundational work in knowledge acquisition problems, methodologies, techniques and tools from the major research groups worldwide. All those contributing hope that access to this material will enable other researchers and practitioners to expedite their own developments through the shared knowledge and experience documented here.

Knowledge acquisition research is still in its early stages and there are many fundamental problems to be solved, new perspectives to be generated, tools to be developed, refined and disseminated, and so on—the work seems endless. Like many modern technologies, knowledge acquisition requires a large-scale cooperative international effort. It is virtually impossible for one research and dvelopment group to have world-class expertise in all the issues, technologies and experience necessary to develop integrated knowledge acquisition tools for a wide range of knowledge-based systems.

We wish to thank the many people who have been involved in organizing these workshops and the organizations that have given them sponsorship and publicity. We have a fundamental debt to those who put in place the computer communication networks worldwide, such as UseNet, that have made the world a global village and enable networks such as ours to operate effectively. We are particularly grateful to the AAAI for its role in sponsoring the North American Workshops and for providing such effective means of disseminating information to the massive community of those now involved in knowledge-based systems research.

We sometimes wonder how we have become so involved in the bureaucracy of organizing workshops and networks when our personal priorities are hacking new knowledge acquisition tools. However, the stimulation of discussions with colleagues at the workshops and across the networks is vital to the direction of our own research. We hope the books will make this stimulation widely available and bring a new generation of researchers into the knowledge acquisition network.

We have attempted to structure the material by dividing it between two books. However, we must make it very clear that the division between the books and into sections in this volume is our own. It is somewhat arbitrary in places, and was not discussed with the contributors. There are many cross-connections between papers in different sections. There is tool-oriented material in this volume and fundamental material in the other. The reader will find it worthwhile to browse through both volumes to get a feel for the many different perspectives present and interactions possible.

John Boose and Brian Gaines

Knowledge Acquisition Tools for Expert Systems

B. R. GAINES
Department of Computer Science, University of Calgary, Calgary, Alberta T2N 1N4, Canada

J. H. BOOSE
Boeing Computer Sciences, P.O. Box 24346, Seattle, Washington 98124-0346, U.S.A.

1. Introduction

This volume contains the papers concerned with tools for knowledge acquisition from the AAAI Knowledge Acquisition for Knowledge-Based System Workshop in November 1986, in Banff, Canada. We have not split them into sections because of the diversity of topics and techniques covered by many of the papers. This is inevitable when integrated systems of tools are being developed that draw on many techniques. There were also keynote addresses, panels and group discussions at the workshop that addressed major themes but did not result in published papers. This paper attempts to capture the essential issues raised in these other presentations.

2. Plenary Papers

Plenary talks were given on the first day by members of the program committee. Each speaker was asked to summarize papers in their area and give an overview of their views on the area. Topics covered relevant to this volume were:

2.1 INTERACTIVE KNOWLEDGE ACQUISITION TOOLS I, JOHN McDERMOTT, CARNEGIE–MELLON UNIVERSITY

Why build intelligent interactive knowledge acquisition tools? 'Smart' means being able to get more mileage out of a few pieces of information than anyone would think you could.

Users of interactive knowledge acquisition tools include AI programmers, programmers, and domain experts. The user may see the structure of the representation (e.g., TEIRESIAS), the problem-solving strategy (e.g., Roget, SALT), or the domain model (e.g., OPAL).

Several interactive knowledge acquisition tools were evaluated in terms of systems built, the system's inference engine, the intended user, and sources of strength. The tools included KREME (Abrett and Burstein), TKAW (Kahn, Breaux, Joseph, and De Klerk), KNACK (Klinker, Bentolila, Genetet, Grimes, and McDermott), STUDENT (Gale), SALT (Marcus), OPAL (Musen, Fagan, Combs, and Shortliffe), and MOLE (Eshelman, Ehret, McDermott, and Tan).

2.2 INTERACTIVE KNOWLEDGE ACQUISITION TOOLS II, JOHN BOOSE, BOEING AI CENTER

Another set of interactive knowledge acquisition tools was evaluated in terms of the problem the tool was addressing, the approach, and the tool's feature set. The tools included KITTEN (Shaw and Gaines), NEXPERT (Rappaport), KRITON (Diedrich, Ruhman, and May), INFORM (Moore and Agogino), FIS (De Jong), MUM (Gruber and Cohen), SMEE (Garg), and AQUINAS (Boose, Bradshaw, and Kitto).

To describe the state-of-the-art, tool features and rates of inclusion were listed: domain modelling, eight tools; interviewing methods, seven tools; use of Personal Construct Psychology and repertory grids, five tools; induction methods, five tools; ability to handle multiple experts, two tools; multiple uncertainty representations, two tools; protocol analysis, two tools; text analysis, two tools; learning (simple), two tools; decision analysis, one tool; copy-and-edit, almost all tools.

To show his committment to knowledge acquisition tools, John Boose produced a repertory grid giving his construing 20 tools and an 'ideal tool', Figure 1 shows this

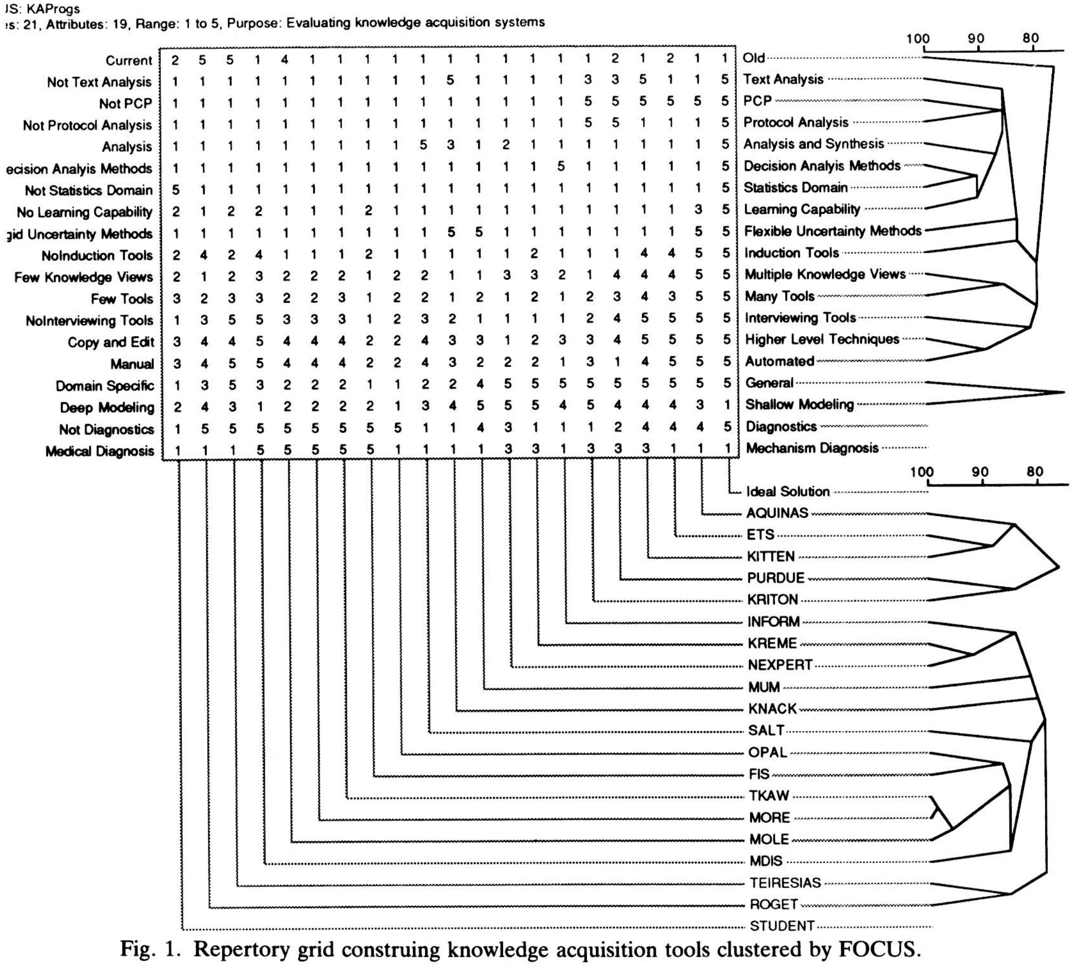

Fig. 1. Repertory grid construing knowledge acquisition tools clustered by FOCUS.

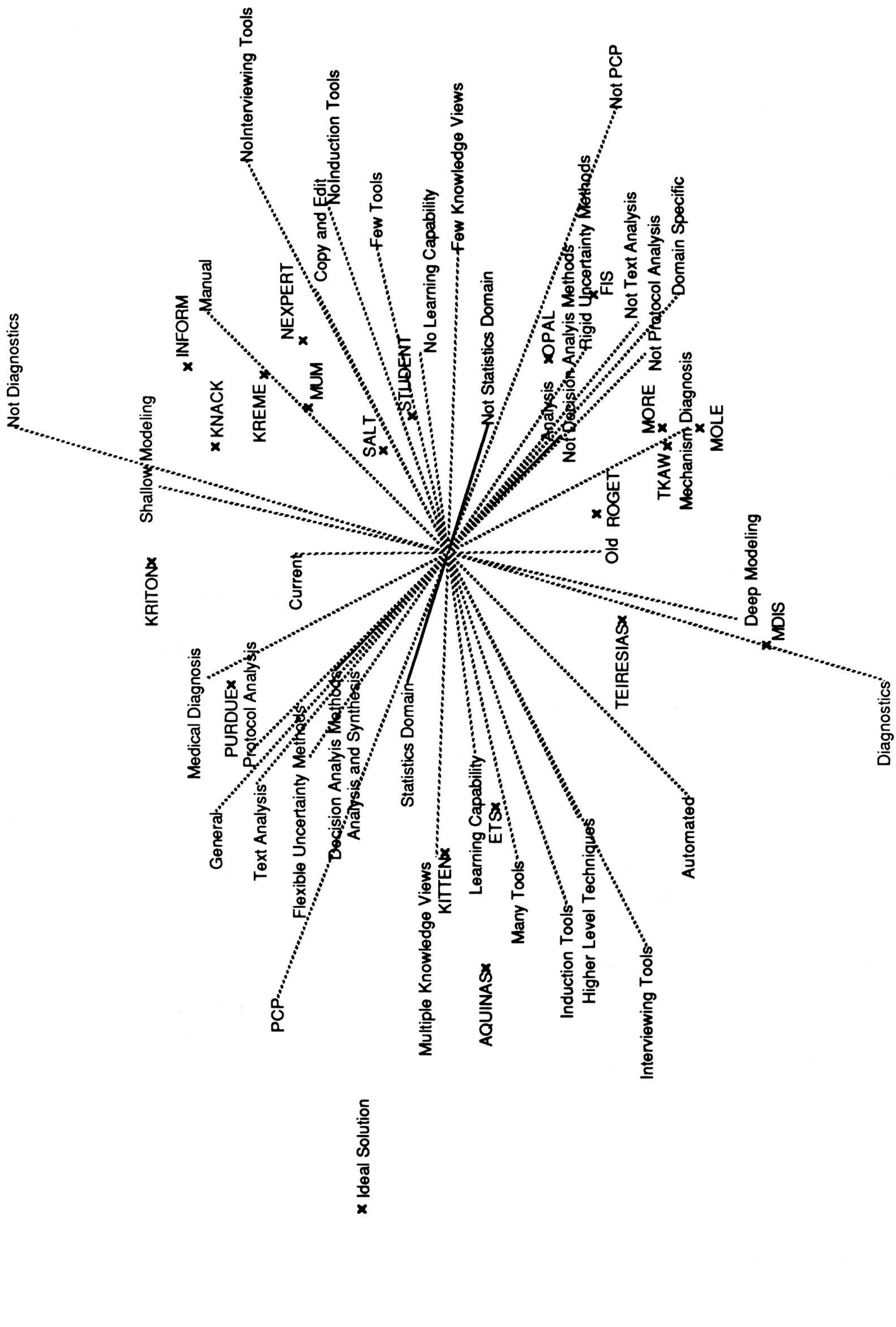

Fig. 2. The tools grid mapped onto its first two principal components.

grid clustered using Shaw's FOCUS algorithm which sorts the original grid and hierarchically clusters it. Note that the 'ideal' tool clusters with no others which is a pretty accurate reflection of the state-of-the-art—we have a long way to go. Figure 2 shows this grid mapped onto its first two principal components. Note also that most of the tool makers were insulated by John's evaluations of their tools which is also a pretty accurate reflection of the state-of-the-art—experts always seem to differ on the most fundamental questions!

3. Working Group Summaries

Attendees participated in several small working groups that attempted to define the aims, objectives, problems, state-of-the-art, and future directions in their areas. The following summaries were presented at the workshop.

3.1 INTERACTIVE KNOWLEDGE ACQUISITION TOOLS, JOHN McDERMOTT, SPOKESPERSON

Aim:

- Facilitate the production of knowledge-based application programs.

Objectives:

- Invent or define a way of organizing or classifying tasks and methods at various levels of abstraction.
- Find a mapping between tasks and methods.
- Develop a technology that, given some collection of tasks within the idealized task hierarchy, finds the associated method collection within the idealized method hierarchy, integrates and possibly specializes the method collection, and defines what kinds of knowledge are needed to perform the tasks and how that knowledge is to be represented.
- Build programs (knowledge acquisition tools) that elicit the required knowledge and represent it and the methods that use it as knowledge-based application programs.

Problems:

- Tasks are ill-defined (i.e. 'configuration, design, interpretation. . . .').
- Methods are sparse (heuristic classification, Newell's weak methods).
- Mapping is poorly understood.

State-of-the-art and future directions ('research bets'):

- Find and clarify a task/method, apply the method to another new task.
- Pick a task, find a method that allows problem-solving in the task domain, apply the method to another new task. Understand methods in the neighborhood and generalize.
- Develop languages for defining task and method hierarchies.
- Look for data to ground ('prove') the task and method hierarchies.
- Build knowledge-based editors.

Slogan:

'Task + method = knowledge-based application'

Contents

The KREME knowledge editing environment

GLENN ABRETT AND MARK H. BURSTEIN
BBN Laboratories, 10 Moulton Street, Cambridge, MA 02238, U.S.A.

One of the major bottlenecks in large-scale expert-system development is the problem of knowledge acquisition: the construction, maintenance, and testing of large knowledge bases. This paper provides an overview of the current state of development of the KREME Knowledge Representation Editing and Modeling Environment. KREME is an extensible experimental environment for developing and editing large knowledge bases in a variety of representation styles. It provides tools for effective viewing and browsing in each kind of representational base, automatic consistency checking, macro-editing facilities to reduce the burdens of large-scale knowledge-base revision and some experimental automatic generalization and acquisition facilities.

1. Introduction

1.1. THE KNOWLEDGE-ACQUISITION PROBLEM

The creation of the large and detailed bodies of knowledge needed to improve the performance of expert systems substantially has proven to be an extremely difficult task. We have identified several factors which make building these very large knowledge bases impractical using current technology:

Knowledge comes in many forms
Human knowledge about the world comes in many forms. Squeezing all the knowledge that an expert system needs into one, or at best two, representational formalisms (e.g. rules and frames) is difficult, time-consuming, and usually an inadequate solution to the task at hand.

Managing large knowledge bases is difficult
As knowledge bases grow in size and complexity they strain the capacities of software tools for knowledge editing, maintenance, and validity checking. Viewpoints at the right level of detail are hard to construct, consistency checking takes up more and more time, and global reorganizations and modifications become virtually impossible to accomplish.

Previously encoded knowledge is not re-used
It is customary to start building a new expert system with an empty knowledge base, even though the completed knowledge base will contain at least some general knowledge about the world. To make matters worse, this general world knowledge is usually entered in a fragmentary and sketchy manner. If general knowledge about the world could be transferred across systems, the gradual accumulation of detail and precision would tremendously enhance the performance of most expert systems.

KNOWLEDGE-BASED SYSTEMS Vol. 2
ISBN 0-12-115920-5