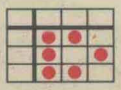


Topics in Information Systems

Cooperative Interfaces to Information Systems

Edited by
L. Bolc and M. Jarke

Springer-Verlag

Cooperative Interfaces to Information Systems

Edited by
L. Bolc and M. Jarke

With 62 Figures

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo

Series Editors

Michael L. Brodie
Computer Corporation of America, Four Cambridge Center
Cambridge, MA 02140/USA

John Mylopoulos
Department of Computer Science, University of Toronto
Toronto, Ontario M5S 1A4/Canada

Joachim W. Schmidt
Fachbereich Informatik, Johann Wolfgang Goethe-Universität
Dantestraße 9, D-6000 Frankfurt a. M. 11/FRG

Volume Editors

Leonard Bolc
Institute of Informatics, Warsaw University
PKiN, pok. 850, 00-901 Warsaw/Poland

Matthias Jarke
Fachbereich Informatik, Johann Wolfgang Goethe-Universität
Dantestraße 9, D-6000 Frankfurt a. M. 11/FRG

ISBN 3-540-16599-1 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-16599-1 Springer-Verlag New York Berlin Heidelberg

Library of Congress Cataloging-in-Publication Data. Cooperative interfaces to information systems. (Topics in information systems). Bibliography: p. Includes index.
1. On-line data processing. 2. Programming languages (Electronic computers)
3. Information storage and retrieval systems. I. Bolc, Leonard, 1934- . II. Jarke, Matthias. III. Series. QA76.55.C66. 1986. 004'.33. 86-13965.
ISBN 0-387-16599-1 (U.S.)

Typsetting, printing and bookbinding: Appl, Wemding
2145/3140-543210

Topics in Information Systems

Editors:

Michael L. Brodie
John Mylopoulos
Joachim W. Schmidt

Topics in Information Systems

Series Description

Dramatic advances in hardware technology have opened the door to a new generation of computer systems. At the same time, the growing demand for information systems of ever-increasing complexity and precision has stimulated the need in every area of Computer Science for more powerful higher-level concepts, techniques, and tools.

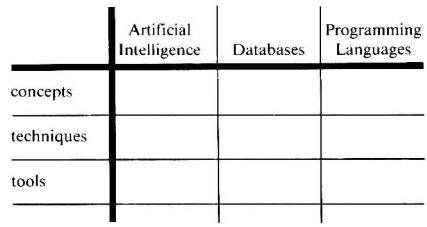
Future information systems will be expected to acquire, maintain, retrieve, manipulate, and present many different kinds of information. These systems will require user-friendly interfaces; powerful reasoning capabilities, and shared access to large information bases. Whereas the needed hardware technology appears to be within reach, the corresponding software technology for building these systems is not. The required dramatic improvements in software productivity will come from advanced application development environments based on powerful new techniques and languages.

The **concepts**, **techniques**, and **tools** necessary for the design, implementation, and use in future information systems are expected to result from the integration of those being developed and used in currently disjoint areas of Computer Science. Several areas bring their unique viewpoints and technologies to existing information processing practice. One key area is **Artificial Intelligence** (AI) which provides knowledge bases grounded on semantic theories of information for correct interpretation. An equally important area is **Databases** which provides means for building and maintaining large, shared databases based on computational theories of information for efficient processing. A third important area is **Programming Languages** which provides a powerful tool kit for the construction of large programs based on linguistic and methodological theories to ensure program correctness. To meet evolving information systems requirements, additional research viewpoints and technologies are or will be required from such areas as **Software Engineering**, **Computer Networks**, **Machine Architectures**, and **Office Automation.**

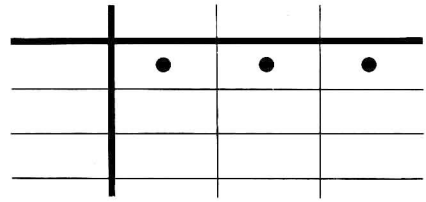
Although some integration of research results has already been achieved, a quantum leap in technological integration is needed to meet the demand for future information systems. This integration is one of the major challenges to Computer Science in the 1980s.

Topics in Information Systems is a series intended to report significant contributions on the integration of concepts, techniques, and tools that advance new technologies for information system construction. The series logo symbolizes the scope of topics to be covered and the basic theme of integration.

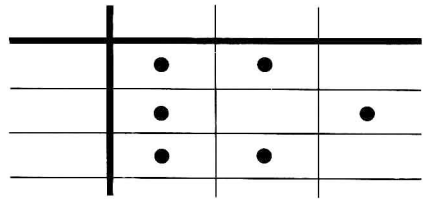
The logo will appear on each book to indicate the topics addressed.



The first book of the series, "On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases and Programming Languages", Michael L. Brodie, John Mylopoulos, and Joachim W. Schmidt (Eds.), February 1984, which deals with concepts in the three areas, has the logo:



The book at hand, "Cooperative Interfaces to Information Systems", Leonard Bolc and Matthias Jarke (Eds.), September 1986, which deals with very high level communication with computerized information systems, has the logo:



In addition, the series comprises the following volumes:

"Query Processing in Database Systems", Won Kim, David S. Reiner, and Donald S. Batory (Eds.), March 1985,

"Office Automation", Dionysios C. Tsichritzis (Ed.), March 1985.

"On knowledge Base Management System: Integrating Artificial Intelligence and Database Technologies", Michael L. Brodie and John Mylopoulos (Eds.), June 1986.

Future books in the series will provide timely accounts of ongoing research efforts to reshape technologies intended for information system development.

September, 1986

Michael L. Brodie
John Mylopoulos
Joachim W. Schmidt

Preface

Information systems are large repositories of factual and inferential knowledge intended to be queried and maintained by a wide variety of users with different backgrounds and work tasks. The community of potential information system users is growing rapidly with advances in hardware and software technology that permit computer/communications support for more and more application areas.

Unfortunately, it is often felt that progress in user interface technology has not quite matched that of other areas. Technical solutions such as computer graphics, natural language processing, or man-machine-man communications in office systems are not enough by themselves. They should be complemented by system features that ensure cooperative behavior of the interfaces, thus reducing the training and usage effort required for successful interaction.

In analogy to a human dialog partner, we call an interface *cooperative* if it does not just accept user requests passively or answer them literally, but actively attempts to understand the users' intentions and to help them solve their application problems. This leads to the central question addressed by this book:

What makes an information systems interface cooperative, and how do we provide capabilities leading to cooperative interfaces?

Many answers are possible. A first aspect concerns the *formulation and acceptance* of user requests. Many researchers assume that such requests should be formulated in natural language. In this area, a central problem of cooperativity is the comprehension of fragmentary and seemingly ambiguous problem statements; if the system could understand incomplete requests, it would save the user a lot of repetitious work. Sometimes, general linguistic analysis leads to the completion of gaps in a query expression, e.g., caused by elliptical references. Often, however, a clear understanding of a request can only be accomplished by taking into account application-specific knowledge; knowledge not only about the domain of discourse, but also about the capabilities, tasks, and intentions of the user.

A second aspect concerns the *presentation* of the answer. A cooperative interface should adapt its output presentation to the cognitive limitations and objectives of the user. If the answer consists only of a few data, the user may desire a

natural language sentence rather than an incomprehensible table as an answer. If the answer is uninformative, there may even be a need to over-answer the initial question. For example, few people would be happy with the answer "yes" to the question: "do you know what time it is?" Conversely, if the answer contains a lot of detailed information, the user may wish to process it further before looking at the result on-line. One effective means of condensing complex output data is their presentation in graphical form.

A third aspect of cooperativity is related to the system's *reaction to exceptional situations*. Trivial errors (e. g., spelling) should be detected and corrected largely automatically. If the system returns an unexpected answer (especially no answer where the user expected one), it should explain the problem and guide the user in its correction. Long delays of an answer should also be explained, in order to avoid misleading the user.

The third aspect could be considered a special case of a fourth one: the problem of supporting the user in *navigating through complex dialogs*. Dialogs may have a complex structure, consisting of different modes of interaction, cross-references, etc. A cooperative interface should maintain a trail of the dialog, sufficient to understand references to previous results and to explain the status of the dialog, for instance, when the user resumes the dialog after an interruption.

Finally, a fifth aspect is the *adaptability* of the interface to new applications, databases, or even different natural languages. Cooperative interfaces should reduce the need for calling expensive external specialists - and the delays caused by their involvement - by providing tools that allow end users to develop, or at least extend, tailored applications interfaces.

These five aspects, although by no means complete, should serve as an initial circumscription of the idea of cooperative interfaces. Recent developments in the field of Artificial Intelligence - in particular, in natural language processing - have begun to offer promising solutions to some of these questions. This book presents a selection of detailed descriptions and empirical evaluations of natural language systems that have made substantial contributions to the idea of cooperative interfaces to information systems.

The contributions in this book approach the problem of cooperativity from rather different perspectives: further developments in the formal aspects of natural language or graphics technology, functional and empirical evaluation of existing systems, and utilization of domain-specific knowledge about the information system and its users. All of these seem to be feasible strategies to promote our understanding of the requirements and solution strategies for designing cooperative interfaces.

The book is divided into three parts. **Part I** addresses general issues in natural language comprehension and (graphical) output presentation. It describes two important systems that attempt to provide application-independent services.

Jane R. Robinson (SRI International, USA) presents "DIAGRAM: A Grammar for Dialogues", a large and complex grammar to be used by AI systems for interpreting English language dialog. DIAGRAM is an augmented phrase-struc-

ture grammer with rule procedures that cause phrases to inherit attributes from their constituents, and from larger phrases in which they themselves are constituents. DIAGRAM does not only analyze the basic kinds of phrases but also complex ones; its application is not restricted to a particular domain.

Chapter 2, by Frank Zdybel, Jr. (Xerox Palo Alto Research Center, USA) describes „An Engine for Intelligent Graphics", an Advanced Information Presentation System (AIPS) that constructs graphic displays according to incomplete declarative specifications of their content, structure, and appearance. A knowledge-based approach using the knowledge representation language, KL-ONE, provides non-programmers with the capability of viewing complex collections of information in arbitrary ways. Similar to DIAGRAM which is intended as an application-independent *language understanding* tool, AIPS offers a centralized cooperative *graphics output* facility which can present information from a wide variety of sources.

The remainder of the book deals with systems that have the explicit goal of providing cooperative natural language access to different kinds of information systems: relational or other commercial databases, information retrieval systems, image processing systems, and expert systems.

Many early systems have tried to deal with this problem on the basis of purely linguistic analysis. The application semantics of such systems is ideally provided by the users themselves through the definition of an application-specific vocabulary and grammar subset. Some of these systems have now reached a fair degree of maturity which allows them to be subjected to stringent empirical evaluation, thus revealing very clearly the requirements of cooperative natural language query systems, but also the achievements and shortcomings of existing technology.

Part II describes three such systems that were or still are considered for commerciae product development. Each chapter stresses one important aspect: functional requirements (Chap. 3), general desirability and feasibility (Chap. 4), and application development (Chap. 5).

Marjorie Templeton and John Burger (System Development Corporation, USA) discuss in Chap. 5 the functional requirements of cooperative natural language interfaces to commercial databases. Their proposal results from their experience with EUFID, a transportable natural language interface which has been used with database management systems (DBMS) as diverse as the relational INGRES system and Honeywell's World-Wide Data Management System. The chapter "Considerations for the Development of Natural-Language Interfaces to Database Management Systems" elaborates cooperativity requirements for application development, language processing, and underlying DBMS, and concludes with a comprehensive checklist of necessary functional capabilities.

Matthias Jarke, Jürgen Krause und Yannis Vassiliou (New York University, USA and University of Regensburg, West Germany) emphasize the need for a formal empirical evaluation of information systems interfaces in Chap. 6. They develop a framework for the design of such studies and describe its application

to a domain-independent natural language query system. The paper "Studies in the Evaluation of a Domain-Independent Natural Language Query System" presents the combined results of three stages of evaluation with the German and English versions of the system, using about 100 experimental subjects and more than 12000 queries. The results demonstrate that natural language interfaces allow more concise query formulation when compared to formal query languages, and require only vocabulary subsets of manageable size. But they fail to prove the superiority of natural language interfaces in terms of problem-solving quality and efficiency.

Fred Damerau (IBM Yorktown Heights, USA) addresses the important problem of user-developed natural language interfaces. In Chap. 5 he presents "An Interactive Customization Program for a Natural Language Database Query System" which is based on the TQA/REQUEST natural language interface to relational databases. The customization program is viewed as a replacement for a general domain-independent natural language database interface. It employs a system-driven dialog to help end users define their own vocabulary and grammatical rules. These users need not be linguists but should be familiar with the database to be accessed.

Part III of the book describes three major projects in knowledge-based natural language processing.

Stimulated by the success of expert systems, most natural language projects started in the last five years stress the role of domain knowledge in interpreting natural language requests cooperatively. By this approach, a natural language interface should be closely related to the semantic meta-information stored with many modern databases, especially with the so-called deductive databases. A number of different knowledge bases can be provided, for example, on the domain, on the technical information system structure, or on the information system's user. Of course, linguistic knowledge bases remain necessary in this approach as well.

Chapter 6, by Jürgen Janas (Hochschule der Bundeswehr, Munich, West Germany), is a theoretical analysis of "The Semantics-Based Natural Language Interface to Relational Databases" which may serve as a formal introduction to an approach that is less interested in the surface structure of natural language but relies on the structural semantics of the underlying relational database. In particular, the problems of ambiguity, conjunction, and ellipsis are elucidated from the semantics point of view. As a cautioning remark, the author emphasizes that this approach is only suitable for serious users who need data from a database they know fairly well.

Wolfgang Höppner, Katharina Morik and Heinz Marburger (University of Hamburg, West Germany) discuss in Chap. 7, "Talking it over: The Natural Language Dialog System HAM-ANS", the role of language-oriented research in the field of Artificial Intelligence. HAM-ANS attempts to provide cooperative natural language access to a wide variety of information systems: a database for fishery research, an image processing system for scene analysis, and an expert system for hotel reservations. Detailed examples of all three applications are given,

and the overall organization of the system and its knowledge bases is described using the hotel reservation example. The authors emphasize the need for user modelling and present suggestions for the representation and application of user models.

The final Chap. 8, by Giorgio Brajnik, Giovanni Guida, and Carlo Tasso (Universita di Udine, Italy) describes an expert interface to yet another kind of information system: the information retrieval interface IR-NLI. In the chapter "An Expert Interface for Effective Man-Machine Interaction", the authors argue for the use of rule-based systems for interface support. They present a task concept for the respresentation and utilization of meta-knowledge, and discuss the use of learning techniques as a method to facilitate cooperative behavior.

In contrast to many short conference reports, the contributions in this book provide fairly detailed descriptions and results for selected systems and studies at the leading edge of cooperative natural language systems technology. The authors are well-known scientists from several of the major natural language research centers. This organization of the book is intended to offer the reader – student, practitioner, or researcher – a comprehensive view of the problems related not only to the design but also to the evaluation of cooperative interfaces to information systems. The book demonstrates that this area of research is still in its early phases, although some impressive results have been achieved. We look forward to future work on cooperative interfaces, hopefully some of it stimulated by the collection of papers in this book.

Warszawa and Frankfurt
November, 1985

Leonard Bolc
Matthias Jarke

Office Automation

Concepts and Tools

Editor: **D. Tsichritzis**

1985. 86 figures. XII, 441 pages. (Topics in Information Systems). ISBN 3-540-15129-X

Query Processing in Database Systems

Editors: **W. Kim, D. S. Reiner, D. S. Batory**

1985. 127 figures. XIV, 365 pages. (Topics in Information Systems). ISBN 3-540-13831-5

Springer-Verlag
Berlin Heidelberg
New York Tokyo

On Conceptual Modelling:

Perspectives from Artificial Intelligence, Databases, and Programming Languages

Editors: **M. L. Brodie, J. Mylopoulos, J. W. Schmidt**

1984. 25 figures. XI, 510 pages. (Topics in Information Systems). ISBN 3-540-90842-0

Conceptual modelling relates to all areas of computer science, but especially to artificial intelligence, databases, and programming languages. Here is the first published collection of state-of-the-art research papers in these domains. Its purpose is to consider conceptual modelling as a topic in its own right rather than as an aspect of data modelling, and to present and compare research on knowledge representation, semantic data models, and data abstraction in this context.
The contributions consist of overviews and reports, each chapter having been written and edited for readers in all three areas. Also included are transcripts of symposium discussions which took place among the contributors during a workshop on conceptual modelling at Intervale; these interdisciplinary discussions of each paper clarify many aspects which might otherwise remain obscure to non-specialists. Key features of the book include introductions to pertinent concepts, and the integration of recent results; focus on twelve research projects, involving specific applications such as database design; and challenging suggestions for further research, especially in the concluding comments by leading experts in the three main fields of inquiry.

Springer-Verlag
Berlin Heidelberg
New York Tokyo

Table of Contents

Part I

Tools for Cooperative Man-Machine Interaction