

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

New York University, NY, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Farid Meziane Elisabeth Métais (Eds.)

Natural Language Processing and Information Systems

9th International Conference on Applications
of Natural Language to Information Systems, NLDB 2004
Salford, UK, June 23-25, 2004
Proceedings



Springer

Volume Editors

Farid Meziane

University of Salford, School of Computing, Science and Engineering
Newton Building, Salford M5 4WT, UK

E-mail: f.meziane@salford.ac.uk

Elisabeth Métais

Laboratoire CEDRIC, CNAM

292 rue Saint Martin, 75141 Paris cedex 3, France

E-mail: elsa@cnam.fr

Library of Congress Control Number: 3540225641

CR Subject Classification (1998): H.2, H.3, I.2, F.3-4, C.2

ISSN 0302-9743

ISBN 3-540-22564-1 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin Heidelberg 2004

Printed in Germany

Typesetting: Camera-ready by author, data conversion by PTP-Berlin, Protago-TeX-Production GmbH

Printed on acid-free paper SPIN: 11019275 06/3142 5 4 3 2 1 0

Preface

Welcome to NLDB04, the Ninth International Conference on the Application of Natural Language to Information Systems, held at the University of Salford, UK during June 23-25, 2004. NLDB04 follows on the success of previous conferences held since 1995. Early conferences then known as Application of Natural Language to Databases, hence the acronym NLDB, were used as a forum to discuss and disseminate research on the integration of natural language and databases and were mainly concerned with natural language based queries, database modelling and user interfaces that facilitate access to information. The conference has since moved to encompass all aspects of Information Systems and Software Engineering. Indeed, the use of natural language in systems modelling has greatly improved the development process and benefited both developers and users at all stages of the software development process.

The latest developments in the field of natural language and the emergence of new technologies has seen a shift towards storage of large semantic electronic dictionaries, their exploitation and the advent of what is now known as the semantic web. Information extraction and retrieval, document and content management, ontology development and management and natural language conversational systems are becoming regular tracks in the last NLDB conferences.

NLDB04 has seen a 50% increase in the number of submissions and has established itself as one of the leading conferences in the area of applying natural language to information systems in its broader sense. The quality of the submissions and their diversity have made the members of the program committee work more than usual. 65 papers were submitted from 22 different countries. 29 were accepted as regular papers, while 13 were accepted as short papers. The papers were classified as belonging to one of these themes:

- Natural Language Conversational Systems
- Intelligent Querying
- Linguistic Aspects of Modeling
- Information Retrieval
- Natural Language Text Understanding
- Knowledge Bases
- Natural Language Text Understanding
- Knowledge Management
- Content Management

This year we were honored by the presence of our invited speaker Fabio Ciravegna from the University of Sheffield, United Kingdom. His lecture on “Challenges in Harvesting Information for the Semantic Web” was highly appreciated and initiated vivid discussions.

We are very thankful for the opportunity to serve as Program Chair and Conference Chair for this conference. However, the organization of such event is a collective effort and a team work. First of all we would like to thank the members of the Program Committee for the time and effort they devoted to the reviewing of the submit-

ted articles and to the selection process. My thanks go also to the additional reviewers for their help and support. We would like to take this opportunity to thank the local organizing committee, especially its chairman Sunil Vadera, for their superb work. We would like to thank Nigel Linge the head of the School of Computing Science and Engineering, Tim Ritchings the head of the Computer Science, Multimedia and Telecommunication discipline and Mr. Gary Wright from the External Relations Division for their help and support.

Obviously we thank the authors for their high quality submissions and their participation to this event and their patience during the long reviewing process.

June 2004

Farid Meziane
Elisabeth Métais

Organization

Conference Chair

Elisabeth Métais, Conservatoire National des Arts et Métiers de Paris, France.

Program Committee Chair

Farid Meziane, University of Salford, UK.

Program Committee

1. Diego Mollá Aliod, Macquarie University, Australia
2. Kenji Araki, Hokkaido University, Japan
3. Mokrane Bouzeghoub, PRiSM, Université de Versailles, France
4. Fabio Ciravegna, University of Sheffield, UK
5. Gary A Coen, Boeing, USA
6. Isabelle Comyn-Wattiau, CEDRIC/CNAM, France
7. Antje Düsterhöft, University of Wismar, Germany
8. Günther Fliedl, Universität Klagenfurt, Austria
9. Alexander Gelbukh, Instituto Politecnico Nacional, Mexico
10. Nicola Guarino, CNR, Italy
11. Rafael Muñoz Guillena, Universidad de Alicante, Spain
12. Jon Atle Gulla, Norwegian University of Science and Technology, Norway
13. Harmain Harmain, United Arab Emirates University, UAE
14. Helmut Horacek, Universität des Saarlandes, Germany
15. Paul Johannesson, Stockholm University, Sweden
16. Zoubida Kedad, PRiSM, Université de Versailles, France
17. Leila Kosseim, Concordia University, Canada.
18. Nadira Lammari, CEDRIC/CNAM, France
19. Winfried Lenders, Universität Bonn, Germany
20. Jana Lewerenz, sd&m Düsseldorf, Germany
21. Robert Luk, Hong Kong Polytechnic University, Hong Kong
22. Heinrich C. Mayr, Universität Klagenfurt, Austria
23. Paul McFetridge, Simon Frazer University, Canada
24. Elisabeth Metais, CEDRIC/CNAM , France
25. Farid Meziane, Salford University, UK
26. Luisa Mich, University of Trento, Italy.
27. Ana Maria Moreno, Universidad Politecnica de Madrid, Spain
28. Jian-Yun Nie, Université de Montréal, Canada
29. Manual Palomar, Universidad de Alicante, Spain
30. Odile Piton, Université Paris I Panthéon-Sorbonne, France
31. Reind van de Riet, Vrije Universiteit Amsterdam, The Netherlands

32. Hae-Chang Rim, Korea University, Korea
33. Tim Ritchings, University of Salford, UK
34. Hongchi Shi, University of Missouri-Columbia, USA
35. Niculae Stratica, Concordia University, Canada
36. Vijay Sugumaran, Oakland University Rochester, USA
37. Veda Storey, Georgia State University, USA
38. Lua Km Teng, National University of Singapore, Singapore
39. Bernhard Thalheim, Kiel University, Germany
40. Babis Theodoulidis, UMIST, UK
41. Sunil Vadera, University of Salford, UK
42. Ronald Wagner, University of Linz, Austria
43. Hans Weigand, Tilburg University, The Netherlands
44. Werner Winiwarter, University of Vienna, Austria
45. Stanislaw Wrycza, University of Gdansk, Poland

Additional Reviewers

1. Farida Aoughlis, University of Tizi-Ouzou, Algeria
2. Andrew Basden, University of Salford, UK.
3. Frances Bell, University of Salford, UK
4. Maria Bergholtz, Stockholm University, Sweden
5. Aleksander Binemann-Zdanowicz, Kiel University, Germany
6. Antonio Ferrández, Universidad de Alicante, Spain
7. Laura Compoy-Gomez, University of Salford, UK
8. Fiedler Gunar, Kiel University, Germany
9. Maria Kutar, University of Salford, UK.
10. Didier Nakache, CNAM, France.
11. Alessandro Oltramari, CNR, Italy
12. Peggy Schmidt, Kiel University, Germany

Organizing Committee

1. Sunil Vadera (Chair), University of Salford, UK.
2. Farid Meziane, University of Salford, UK.
3. Samia Nefti, University of Salford, UK.
4. Mohamad Saraee, University of Salford, UK.
5. Edwin Mit, University of Salford, UK.

Organization Website and Contact

The conference was organized by the Computer Science Research Centre and the School of Computing, Science and Engineering at the University of Salford, United Kingdom. The website of the NLDB Conference is www.nldb.org. For all information, contact Farid Meziane at F.Meziane@salford.ac.uk

Table of Contents

Regular Papers

Natural Language Conversational Systems

A Natural Language Model and a System for Managing TV-Anytime Information from Mobile Devices	1
<i>Anastasia Karanastasi, Fotis G. Kazasis, Stavros Christodoulakis</i>	

State- and Object Oriented Specification of Interactive VoiceXML Information Services	13
<i>Thomas Schwanzara-Bennoit, Gunar Fiedler</i>	

Interpreting Semi-formal Utterances in Dialogs about Mathematical Proofs	26
<i>Helmut Horacek, Magdalena Wolska</i>	

Intelligent Querying

Event Ordering Using TERSEO System	39
<i>Estela Saquete, Rafael Muñoz, Patricio Martínez-Barco</i>	

The Role of User Profiles in Context-Aware Query Processing for the Semantic Web	51
<i>Veda C. Storey, Vijayan Sugumaran, Andrew Burton-Jones</i>	

Deriving FrameNet Representations: Towards Meaning-Oriented Question Answering	64
<i>Gerhard Fliedner</i>	

Lightweight Natural Language Database Interfaces	76
<i>In-Su Kang, Seung-Hoon Na, Jong-Hyeok Lee, Gijoo Yang</i>	

Ontology-Driven Question Answering in AquaLog	89
<i>Vanessa Lopez, Enrico Motta</i>	

Schema-Based Natural Language Semantic Mapping	103
<i>Niculae Stratica, Bipin C. Desai</i>	

Avaya Interactive Dashboard (AID): An Interactive Tool for Mining the Avaya Problem Ticket Database	114
<i>Ziyang Wang, Amit Bagga</i>	

Linguistic Aspects of Modeling

Information Modeling: The Process and the Required Competencies of Its Participants	123
<i>P.J.M. Frederiks, T.P. van der Weide</i>	

Experimenting with Linguistic Tools for Conceptual Modelling: Quality of the Models and Critical Features	135
<i>Nadzeja Kiyavitskaya, Nicola Zeni, Luisa Mich, John Mylopoulos</i>	

Language Resources and Tools for Supporting the System Engineering Process	147
<i>V.O. Onditi, P. Rayson, B.Ransom, D. Ramduny, Ian Sommerville, A. Dix</i>	

A Linguistics-Based Approach for Use Case Driven Analysis Using Goal and Scenario Authoring	159
<i>Jintae Kim, Sooyong Park, Vijayan Sugumaran</i>	

Information Retrieval

Effectiveness of Index Expressions	171
<i>F.A. Grootjen, T.P. van der Weide</i>	

Concept Similarity Measures the Understanding Between Two Agents ...	182
<i>Jesus M. Olivares-Ceja, Adolfo Guzman-Arenas</i>	

Concept Indexing for Automated Text Categorization	195
<i>José María Gómez, José Carlos Cortizo, Enrique Puertas, Miguel Ruiz</i>	

Natural Language Text Understanding

Acquiring Selectional Preferences from Untagged Text for Prepositional Phrase Attachment Disambiguation	207
<i>Hiram Calvo, Alexander Gelbukh</i>	

Semantic Enrichment for Ontology Mapping	217
<i>Xiaomeng Su, Jon Atle Gulla</i>	

Testing Word Similarity: Language Independent Approach with Examples from Romance	229
<i>Mikhail Alexandrov, Xavier Blanco, Pavel Makagonov</i>	

Knowledge Bases

Language Modeling for Effective Construction of Domain Specific Thesauri	242
<i>Libo Chen, Ulrich Thiel</i>	

Populating a Database from Parallel Texts Using Ontology-Based Information Extraction	254
<i>M.M. Wood, S.J. Lydon, V. Tablan, D. Maynard, H. Cunningham</i>	

A Generic Coordination Model for Pervasive Computing Based on Semantic Web Languages	265
<i>Amine Tafat, Michele Courant, Beat Hirsbrunner</i>	

Natural Language Text Understanding

Improving Web Searching Using Descriptive Graphs	276
<i>Alain Couchot</i>	

An Unsupervised WSD Algorithm for a NLP System	288
<i>Iulia Nica, Andrés Montoyo, Sonia Vázquez, M^a Antònia Martí</i>	

Enhanced Email Classification Based on Feature Space Enriching	299
<i>Yunming Ye, Fanyuan Ma, Hongqiang Rong, Joshua Huang</i>	

Synonymous Paraphrasing Using WordNet and Internet	312
<i>Igor A. Bolshakov, Alexander Gelbukh</i>	

Knowledge Management

Automatic Report Generation from Ontologies: The MIAKT Approach	324
<i>Kalina Bontcheva, Yorick Wilks</i>	

A Flexible Workbench for Document Analysis and Text Mining	336
<i>Jon Atle Gulla, Terje Brasethvik, Harald Kaada</i>	

Short Papers

Content Management

Towards Linguistic Foundations of Content Management	348
<i>Gunar Fiedler, Bernhard Thalheim</i>	

Constructing Natural Knowledge Ontologies to Implement Semantic Organizational Memory	354
<i>Laura Campoy-Gomez</i>	

Improving the Naming Process for Web Site Reverse Engineering	362
<i>Sélina Besbes Essanaa, Nadira Lammari</i>	

On Embedding Machine-Processable Semantics into Documents	368
<i>Krishnaprasad Thirunarayan</i>	

Information Retrieval

Using IR Techniques to Improve Automated Text Classification	374
<i>Teresa Gonçalves, Paulo Quaresma</i>	
Architecture of a Medical Information Extraction System	380
<i>Dalila Bekhouche, Yann Pollet, Bruno Grilheres, Xavier Denis</i>	
Improving Information Retrieval in MEDLINE by Modulating MeSH Term Weights	388
<i>Kwangcheol Shin, Sang-Yong Han</i>	
Identification of Composite Named Entities in a Spanish Textual Database	395
<i>Sofía N. Galicia-Haro, Alexander Gelbukh, Igor A. Bolshakov</i>	

Intelligent Querying

ORAKEL: A Natural Language Interface to an F-Logic Knowledge Base	401
<i>Philipp Cimiano</i>	
Accessing an Information System by Chatting	407
<i>Bayan Abu Shawar, Eric Atwell</i>	
Ontology-Based Question Answering in a Federation of University Sites: The MOSES Case Study	413
<i>P. Atzeni, R. Basili, D.H. Hansen, P. Missier, Patrizia Paggio, Maria Teresa Pazienza, Fabio Massimo Zanzotto</i>	
Semantic Tagging and Chunk-Parsing in Dynamic Modeling	421
<i>Günther Fliedl, Christian Kop, Heinrich C. Mayr, Christian Winkler, Georg Weber, Alexander Salbrechter</i>	
Semantic Filtering of Textual Requirements Descriptions	427
<i>Jorge J. García Flores</i>	
Author Index	435