Farid Meziane
Elisabeth Métais (Eds.)

LNCS 3136

Natural Language Processing and Information Systems

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

Springer

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

Lecture Notes in Computer Science 3136

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

Preface

Welcome to NLDB 2004, the 9th International Conference on the Application of Natural Language to Information Systems, held at the University of Salford, UK during June 23–25, 2004. NLDB 2004 followed 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 have become regular tracks in recent NLDB conferences.

NLDB 2004 saw a 50% increase in the number of submissions, and NLDB 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 made the work of the members of the program committee more difficult than usual. Sixty-five papers were submitted from 22 different countries. Twenty-nine 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
- Recognition of information in natural language description
- 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, UK. 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 an event is a collective effort and the result of a team's work. First of all we would like to thank the

此为试读,需要完整PDF请访问: www.ertongbook.com

members of the Program Committee for the time and effort they devoted to the reviewing of the submitted 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 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 in 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

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

Hae-Chang Rim, Korea University, Korea
Tim Ritchings, University of Salford, UK
Hongchi Shi, University of Missouri-Columbia, USA
Niculae Stratica, Concordia University, Canada
Vijay Sugumaran, Oakland University, Rochester, USA
Veda Storey, Georgia State University, USA
Lua Km Teng, National University of Singapore, Singapore
Bernhard Thalheim, Kiel University, Germany
Babis Theodoulidis, UMIST, UK
Sunil Vadera, University of Salford, UK
Ronald Wagner, University of Linz, Austria
Hans Weigand, Tilburg University, The Netherlands
Werner Winiwarter, University of Vienna, Austria
Stanislaw Wrycza, University of Gdansk, Poland

Additional Reviewers

Farida Aoughlis, University of Tizi-Ouzou, Algeria
Andrew Basden, University of Salford, UK
Frances Bell, University of Salford, UK
Maria Bergholtz, Stockholm University, Sweden
Aleksander Binemann-Zdanowicz, Kiel University, Germany
Antonio Ferrández, Universidad de Alicante, Spain
Laura Compoy-Gomez, University of Salford, UK
Fiedler Gunar, Kiel University, Germany
Maria Kutar, University of Salford, UK
Didier Nakache, CNAM, France
Alessandro Oltramari, CNR, Italy
Peggy Schmidt, Kiel University, Germany

Organizing Committee

Sunil Vadera (Chair), University of Salford, UK
Farid Meziane, University of Salford, UK
Samia Nefti, University of Salford, UK
Mohamad Saraee, University of Salford, UK
Edwin Mit, University of Salford, UK

Lecture Notes in Computer Science

For information about Vols. 1–3055

please contact your bookseller or Springer

Vol. 3172: M. Dorigo, M. Birattari, C. Blum, L. M.Gambardella, F. Mondada, T. Stützle (Eds.), Ant Colony, Optimization and Swarm Intelligence. XII, 434 pages. 2004.

Vol. 3158: I. Nikolaidis, M. Barbeau, E. Kranakis (Eds.), Ad-Hoc, Mobile, and Wireless Networks. IX, 344 pages. 2004.

Vol. 3157: C. Zhang, H. W. Guesgen, W.K. Yeap (Eds.), PRICAI 2004: Trends in Artificial Intelligence. XX, 1023 pages. 2004. (Subseries LNAI).

Vol. 3156: M. Joye, J.-J. Quisquater (Eds.), Cryptographic Hardware and Embedded Systems - CHES 2004. XIII, 455 pages. 2004.

Vol. 3153: J. Fiala, V. Koubek, J. Kratochvíl (Eds.), Mathematical Foundations of Computer Science 2004. XIV, 902 pages. 2004.

Vol. 3152: M. Franklin (Ed.), Advances in Cryptology – CRYPTO 2004. XI, 579 pages. 2004.

Vol. 3148: R. Giacobazzi (Ed.), Static Analysis. X, 393 pages. 2004.

Vol. 3146: P. Érdi, A. Esposito, M. Marinaro, S. Scarpetta (Eds.), Computational Neuroscience: Cortical Dynamics. XI, 161 pages. 2004.

Vol. 3144: M. Papatriantafilou, P. Hunel (Eds.), Principles of Distributed Systems. XI, 246 pages. 2004.

Vol. 3143: W. Liu, Y. Shi, Q. Li (Eds.), Advances in Web-Based Learning – ICWL 2004. XIV, 459 pages. 2004.

Vol. 3142: J. Diaz, J. Karhumäki, A. Lepistö, D. Sannella (Eds.), Automata, Languages and Programming. XIX, 1253 pages. 2004.

Vol. 3140: N. Koch, P. Fraternali, M. Wirsing (Eds.), Web Engineering. XXI, 623 pages. 2004.

Vol. 3139: F. Iida, R. Pfeifer, L. Steels, Y. Kuniyoshi (Eds.), Embodied Artificial Intelligence. IX, 331 pages. 2004. (Subseries LNAI).

Vol. 3138: A. Fred, T. Caelli, R.P.W. Duin, A. Campilho, D.d. Ridder (Eds.), Structural, Syntactic, and Statistical Pattern Recognition. XXII, 1168 pages. 2004.

Vol. 3136: F. Meziane, E. Métais (Eds.), Natural Language Processing and Information Systems. XII, 436 pages. 2004.

Vol. 3134: C. Zannier, H. Erdogmus, L. Lindstrom (Eds.), Extreme Programming and Agile Methods - XP/Agile Universe 2004. XIV, 233 pages. 2004.

Vol. 3133: A.D. Pimentel, S. Vassiliadis (Eds.), Computer Systems: Architectures, Modeling, and Simulation. XIII, 562 pages. 2004.

Vol. 3131: V. Torra, Y. Narukawa (Eds.), Modeling Decisions for Artificial Intelligence. XI, 327 pages. 2004. (Subseries LNAI).

Vol. 3130: A. Syropoulos, K. Berry, Y. Haralambous, B. Hughes, S. Peter, J. Plaice (Eds.), TEX, XML, and Digital Typography. VIII, 265 pages. 2004.

Vol. 3129: Q. Li, G. Wang, L. Feng (Eds.), Advances in Web-Age Information Management. XVII, 753 pages. 2004.

Vol. 3128: D. Asonov (Ed.), Querying Databases Privately. IX, 115 pages. 2004.

Vol. 3127: K.E. Wolff, H.D. Pfeiffer, H.S. Delugach (Eds.), Conceptual Structures at Work. XI, 403 pages. 2004. (Subseries LNAI).

Vol. 3126: P. Dini, P. Lorenz, J.N.d. Souza (Eds.), Service Assurance with Partial and Intermittent Resources. XI, 312 pages. 2004.

Vol. 3125: D. Kozen (Ed.), Mathematics of Program Construction. X, 401 pages. 2004.

Vol. 3124: J.N. de Souza, P. Dini, P. Lorenz (Eds.), Telecommunications and Networking - ICT 2004. XXVI, 1390 pages. 2004.

Vol. 3123: A. Belz, R. Evans, P. Piwek (Eds.), Natural Language Generation. X, 219 pages. 2004. (Subseries LNAI).

Vol. 3121: S. Nikoletseas, J.D.P. Rolim (Eds.), Algorithmic Aspects of Wireless Sensor Networks. X, 201 pages. 2004.

Vol. 3120: J. Shawe-Taylor, Y. Singer (Eds.), Learning Theory. X, 648 pages. 2004. (Subseries LNAI).

Vol. 3118: K. Miesenberger, J. Klaus, W. Zagler, D. Burger (Eds.), Computer Helping People with Special Needs. XXIII, 1191 pages. 2004.

Vol. 3116: C. Rattray, S. Maharaj, C. Shankland (Eds.), Algebraic Methodology and Software Technology. XI, 569 pages. 2004.

Vol. 3115: P. Enser, Y. Kompatsiaris, N.E. O'Connor, A.F. Smeaton, A.W.M. Smeulders (Eds.), Image and Video Retrieval. XVII, 679 pages. 2004.

Vol. 3114: R. Alur, D.A. Peled (Eds.), Computer Aided Verification. XII, 536 pages. 2004.

Vol. 3113: J. Karhumäki, H. Maurer, G. Paun, G. Rozenberg (Eds.), Theory Is Forever. X, 283 pages. 2004.

Vol. 3112: H. Williams, L. MacKinnon (Eds.), Key Technologies for Data Management. XII, 265 pages. 2004.

Vol. 3111: T. Hagerup, J. Katajainen (Eds.), Algorithm Theory - SWAT 2004. XI, 506 pages. 2004.

Vol. 3110: A. Juels (Ed.), Financial Cryptography. XI, 281 pages. 2004.

Vol. 3109: S.C. Sahinalp, S. Muthukrishnan, U. Dogrusoz (Eds.), Combinatorial Pattern Matching. XII, 486 pages. 2004.

Vol. 3108: H. Wang, J. Pieprzyk, V. Varadharajan (Eds.), Information Security and Privacy. XII, 494 pages. 2004.

Vol. 3107: J. Bosch, C. Krueger (Eds.), Software Reuse: Methods, Techniques and Tools. XI, 339 pages. 2004.

Vol. 3106: K.-Y. Chwa, J.I. Munro (Eds.), Computing and Combinatorics. XIII, 474 pages. 2004.

Vol. 3105: S. Göbel, U. Spierling, A. Hoffmann, I. Iurgel, O. Schneider, J. Dechau, A. Feix (Eds.), Technologies for Interactive Digital Storytelling and Entertainment. XVI, 304 pages. 2004.

Vol. 3104: R. Kralovic, O. Sykora (Eds.), Structural Information and Communication Complexity. X, 303 pages. 2004.

Vol. 3103: K. Deb, e. al. (Eds.), Genetic and Evolutionary Computation – GECCO 2004. XLIX, 1439 pages. 2004.

Vol. 3102: K. Deb, e. al. (Eds.), Genetic and Evolutionary Computation – GECCO 2004. L, 1445 pages. 2004.

Vol. 3101: M. Masoodian, S. Jones, B. Rogers (Eds.), Computer Human Interaction. XIV, 694 pages. 2004.

Vol. 3100: J.F. Peters, A. Skowron, J.W. Grzymała-Busse, B. Kostek, R.W. Świniarski, M.S. Szczuka (Eds.), Transactions on Rough Sets I. X, 405 pages. 2004.

Vol. 3099: J. Cortadella, W. Reisig (Eds.), Applications and Theory of Petri Nets 2004. XI, 505 pages. 2004.

Vol. 3098: J. Desel, W. Reisig, G. Rozenberg (Eds.), Lectures on Concurrency and Petri Nets. VIII, 849 pages. 2004.

Vol. 3097: D. Basin, M. Rusinowitch (Eds.), Automated Reasoning. XII, 493 pages. 2004. (Subseries LNAI).

Vol. 3096: G. Melnik, H. Holz (Eds.), Advances in Learning Software Organizations. X, 173 pages. 2004.

Vol. 3095: C. Bussler, D. Fensel, M.E. Orlowska, J. Yang (Eds.), Web Services, E-Business, and the Semantic Web. X, 147 pages. 2004.

Vol. 3094: A. Nürnberger, M. Detyniecki (Eds.), Adaptive Multimedia Retrieval. VIII, 229 pages. 2004.

Vol. 3093: S.K. Katsikas, S. Gritzalis, J. Lopez (Eds.), Public Key Infrastructure. XIII, 380 pages. 2004.

Vol. 3092: J. Eckstein, H. Baumeister (Eds.), Extreme Programming and Agile Processes in Software Engineering. XVI, 358 pages. 2004.

Vol. 3091: V. van Oostrom (Ed.), Rewriting Techniques and Applications. X, 313 pages. 2004.

Vol. 3089: M. Jakobsson, M. Yung, J. Zhou (Eds.), Applied Cryptography and Network Security. XIV, 510 pages. 2004.

Vol. 3087: D. Maltoni, A.K. Jain (Eds.), Biometric Authentication. XIII, 343 pages. 2004.

Vol. 3086: M. Odersky (Ed.), ECOOP 2004 – Object-Oriented Programming. XIII, 611 pages. 2004.

Vol. 3085: S. Berardi, M. Coppo, F. Damiani (Eds.), Types for Proofs and Programs. X, 409 pages. 2004.

Vol. 3084: A. Persson, J. Stirna (Eds.), Advanced Information Systems Engineering. XIV, 596 pages. 2004.

Vol. 3083: W. Emmerich, A.L. Wolf (Eds.), Component Deployment. X, 249 pages. 2004.

Vol. 3080: J. Desel, B. Pernici, M. Weske (Eds.), Business Process Management. X, 307 pages. 2004.

Vol. 3079: Z. Mammeri, P. Lorenz (Eds.), High Speed Networks and Multimedia Communications. XVIII, 1103 pages. 2004.

Vol. 3078: S. Cotin, D.N. Metaxas (Eds.), Medical Simulation. XVI, 296 pages. 2004.

Vol. 3077: F. Roli, J. Kittler, T. Windeatt (Eds.), Multiple Classifier Systems. XII, 386 pages. 2004.

Vol. 3076: D. Buell (Ed.), Algorithmic Number Theory. XI, 451 pages. 2004.

Vol. 3074: B. Kuijpers, P. Revesz (Eds.), Constraint Databases and Applications. XII, 181 pages. 2004.

Vol. 3073: H. Chen, R. Moore, D.D. Zeng, J. Leavitt (Eds.), Intelligence and Security Informatics. XV, 536 pages. 2004.

Vol. 3072: D. Zhang, A.K. Jain (Eds.), Biometric Authentication. XVII, 800 pages. 2004.

Vol. 3071: A. Omicini, P. Petta, J. Pitt (Eds.), Engineering Societies in the Agents World. XIII, 409 pages. 2004. (Subseries LNAI).

Vol. 3070: L. Rutkowski, J. Siekmann, R. Tadeusiewicz, L.A. Zadeh (Eds.), Artificial Intelligence and Soft Computing - ICAISC 2004. XXV, 1208 pages. 2004. (Subseries LNAI).

Vol. 3068: E. André, L. Dybkjær, W. Minker, P. Heisterkamp (Eds.), Affective Dialogue Systems. XII, 324 pages. 2004. (Subseries LNAI).

Vol. 3067: M. Dastani, J. Dix, A. El Fallah-Seghrouchni (Eds.), Programming Multi-Agent Systems. X, 221 pages. 2004. (Subseries LNAI).

Vol. 3066: S. Tsumoto, R. Słowiński, J. Komorowski, J.W. Grzymała-Busse (Eds.), Rough Sets and Current Trends in Computing. XX, 853 pages. 2004. (Subseries LNAI).

Vol. 3065: A. Lomuscio, D. Nute (Eds.), Deontic Logic in Computer Science. X, 275 pages. 2004. (Subseries LNAI).

Vol. 3064: D. Bienstock, G. Nemhauser (Eds.), Integer Programming and Combinatorial Optimization. XI, 445 pages. 2004.

Vol. 3063: A. Llamosí, A. Strohmeier (Eds.), Reliable Software Technologies - Ada-Europe 2004. XIII, 333 pages. 2004.

Vol. 3062: J.L. Pfaltz, M. Nagl, B. Böhlen (Eds.), Applications of Graph Transformations with Industrial Relevance. XV, 500 pages. 2004.

Vol. 3061: F.F. Ramos, H. Unger, V. Larios (Eds.), Advanced Distributed Systems. VIII, 285 pages. 2004.

Vol. 3060: A.Y. Tawfik, S.D. Goodwin (Eds.), Advances in Artificial Intelligence. XIII, 582 pages. 2004. (Subseries LNAI).

Vol. 3059: C.C. Ribeiro, S.L. Martins (Eds.), Experimental and Efficient Algorithms. X, 586 pages. 2004.

Vol. 3058: N. Sebe, M.S. Lew, T.S. Huang (Eds.), Computer Vision in Human-Computer Interaction. X, 233 pages. 2004.

Vol. 3057: B. Jayaraman (Ed.), Practical Aspects of Declarative Languages. VIII, 255 pages. 2004.

Vol. 3056: H. Dai, R. Srikant, C. Zhang (Eds.), Advances in Knowledge Discovery and Data Mining. XIX, 713 pages. 2004. (Subseries LNAI).

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
Anastasia Karanastasi, Fotis G. Kazasis, Stavros Christodoulakis

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

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

Intelligent Querying

Event Ordering Using TERSEO System 39
Estela Saquete, Rafael Muñoz, Patricio Martínez-Barco

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

Deriving FrameNet Representations: Towards Meaning-Oriented Question Answering 64
Gerhard Fliedner

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

Ontology-Driven Question Answering in AquaLog 89
Vanessa Lopez, Enrico Motta

Schema-Based Natural Language Semantic Mapping 103
Niculae Stratica, Bipin C. Desai

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

此为试读,需要完整PDF请访问: www.ertongbook.com

Linguistic Aspects of Modeling

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

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

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

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

Information Retrieval

Effectiveness of Index Expressions 171
F.A. Grootjen, T.P. van der Weide

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

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

Natural Language Text Understanding

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

Semantic Enrichment for Ontology Mapping 217
Xiaomeng Su, Jon Atle Gulla

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

Knowledge Bases

Language Modeling for Effective Construction of Domain Specific Thesauri 242
Libo Chen, Ulrich Thiel

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

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

Natural Language Text Understanding

Improving Web Searching Using Descriptive Graphs 276
Alain Couchot

An Unsupervised WSD Algorithm for a NLP System 288
Iulia Nica, Andrés Montoyo, Sonia Vázquez, Mª Antònia Martí

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

Synonymous Paraphrasing Using WordNet and Internet 312
Igor A. Bolshakov, Alexander Gelbukh

Knowledge Management

Automatic Report Generation from Ontologies: The MIAKT Approach 324
Kalina Bontcheva, Yorick Wilks

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

Short Papers

Content Management

Towards Linguistic Foundations of Content Management 348
Gunar Fiedler, Bernhard Thalheim

Constructing Natural Knowledge Ontologies to Implement Semantic Organizational Memory 354
Laura Campoy-Gomez

Improving the Naming Process for Web Site Reverse Engineering 362
Sélima Besbes Essanaa, Nadira Lammari

On Embedding Machine-Processable Semantics into Documents 368
Krishnaprasad Thirunarayan

Information Retrieval

Using IR Techniques to Improve Automated Text Classification 374
Teresa Gonçalves, Paulo Quaresma

Architecture of a Medical Information Extraction System 380
Dalila Bekhouche, Yann Pollet, Bruno Grilheres, Xavier Denis

Improving Information Retrieval in MEDLINE by Modulating MeSH Term Weights 388
Kwangcheol Shin, Sang-Yong Han

Identification of Composite Named Entities in a Spanish Textual Database 395
Sofía N. Galicia-Haro, Alexander Gelbukh, Igor A. Bolshakov

Intelligent Querying

ORAKEL: A Natural Language Interface to an F-Logic Knowledge Base 401
Philipp Cimiano

Accessing an Information System by Chatting 407
Bayan Abu Shawar, Eric Atwell

Ontology-Based Question Answering in a Federation of University Sites: The MOSES Case Study 413
P. Atzeni, R. Basili, D.H. Hansen, P. Missier, Patrizia Paggio, Maria Teresa Pazienza, Fabio Massimo Zanzotto

Linguistic Aspects of Modelling

Semantic Tagging and Chunk-Parsing in Dynamic Modeling 421
Günther Fliedl, Christian Kop, Heinrich C. Mayr, Christian Winkler, Georg Weber, Alexander Salbrechter

Semantic Filtering of Textual Requirements Descriptions 427
Jorge J. García Flores

Author Index 435

A Natural Language Model and a System for Managing TV-Anytime Information from Mobile Devices

Anastasia Karanastasi, Fotis G. Kazasis, and Stavros Christodoulakis

Lab. of Distributed Multimedia Information Systems / Technical University of Crete (MUSIC/TUC)
University Campus, Kounoupidiana, Chania, Greece
{allegra,fotis,stavros}@ced.tuc.gr

Abstract. The TV-Anytime standard describes structures of categories of digital TV program metadata, as well as User Profile metadata for TV programs. In this case study we describe a natural language model and a system for the users to interact with the metadata and preview TV programs stored in remote databases, from their mobile devices contrary to their limited configurations. By the use of the TV-Anytime metadata specifications the system limits greatly the possibility for ambiguities. The interaction model deals with ambiguities by using the TV-Anytime user profiles and metadata information concerning digital TV to rank the possible answers. The interaction between the user and the system is done by the use of a PDA and a mobile phone with metadata information stored on a database on a remote TV-Anytime compatible TV set.

1 Introduction

The number of digital TV channels has increased dramatically the last few years, and several industrial sectors and content producing sectors are active in defining the environment in which the TVs of the future will operate.

The TV-Anytime Forum is an association of organizations which seeks to develop specifications to enable audio-visual and other services based on mass-market high volume digital storage in consumer platforms - simply referred to as local storage [1]. These specifications target interoperable and integrated systems, from content creators/providers, through service providers, to the consumers and aim to enable applications to exploit the storage capabilities in consumer platforms. The basic architectural unit is an expanded TV set (known as a Personal Digital Recorder – PDR) capable of capturing digital satellite broadcasts according to user interests as they are described in his profile and storing them into large storage devices. The current TV-Anytime standard specifications define the structures for the metadata that can be used to describe TV programs and broadcasts, as well as for the metadata that can be used to describe the user profile. Expanded versions of the TV-Anytime architecture foresee also last mile TV-Anytime servers, Internet connection of the TV set and mobility aspects. Mobile devices (mobile phones, PDAs, etc.) in the TV-Anytime architecture can be used by a user to communicate with the home TV set not only for viewing TV programs, but also for managing the contents of the TV set (like previewing its con-

F. Meziane and E. Métais (Eds.): NLDB 2004, LNCS 3136, pp. 1–12, 2004.
© Springer-Verlag Berlin Heidelberg 2004

tents, searching for content, deleting content that has been recorded for him by the TV set, etc.) and for managing his profile preferences [2].

There is a strong need for new interface paradigms that allow the interaction of naïve users with the future TV sets in order to better satisfy their dynamic preferences and access information. The usual pc-based interfaces are not appropriate to interact with mobile devices (like mobile phones or PDAs) or with TV sets. Natural language interfaces (NLIs) are more appropriate interface styles for naïve users, and they can also support voice-based interactions for mobile devices.

The appeal of natural language interfaces to databases has been explored since the beginning of the '80s [6], [7]. Significant advances have been made in dialogue management [3], [4], [5], but the problem of reliable understanding a single sentence has not been solved. In comparison to the efforts made several years ago to enrich the databases with NLIs which faced the prohibitive cost of dialogues to fully clarify the query [3], our environment is more concrete than general purpose interfaces to database systems, since the structure imposed by the TV-Anytime specifications for the metadata greatly limit the possibilities for ambiguities.

The importance of natural language interfaces to databases has increased rapidly the last few years due to the introduction of new user devices (including mobile devices such as PDAs and mobile phones) for which traditional mouse based interfaces are unacceptable. Research has been published in the area of NLIs to interactive TV based information systems [8], [9]. A well-known problem with the NLIs is that user interactions may be ambiguous. Ambiguity in the NLIs is a serious problem and most systems proposed in the literature often lead to lengthy clarification dialogues with the user to resolve ambiguities [14]. These dialogues systems face the problem that the users often do not know the answers to questions asked by the system. Unlike the previous systems we do not resolve the remaining ambiguities with clarification. Instead we can take advantage of the TV-Anytime user profile specifications in order to rank the possible interpretations and present to the user at the top position the one with the highest ranking.

In this paper we present a model for natural language interactions with a TV set in an environment that follows the TV Anytime specifications, both for the TV program metadata as well as for the user profile metadata. The metadata are stored in databases with last mile connections. The natural language interactions are used to preview programs or summaries of programs as well as to completely manage the metadata and the programs that the TV set keeps for the user. In addition we describe an implementation of this TV-Anytime compatible natural language interaction model that works on a PDA and a mobile phone, which communicates with the TV-Anytime TV set for managing its programs and metadata and also allowing the previewing of TV programs from the mobile device.

The best-known dialogue systems that have been developed for digital TV and mobile environments are related to the MIINA project [11] and the Program Guide Information System of NOKIA [12]. In the context of MIINA project, a system has been developed for information retrieval from the set-top-box Mediaterminal of NOKIA. The user is allowed to insert queries for TV programs, channels, program categories and broadcast time, using a natural language. However, the natural language interaction in this model is rather simple since it is only related to the information provided by a traditional TV-Guide. The Program Guide Information System is an electronic call-in demo application offering information about television programs

over the phone by allowing the user to converse with the system in natural language sentences. This system is not based on TV-Anytime metadata structures for describing the programs or the user profiles. The scope of the interaction does not include any management of the stored content except retrieval or the user profiles. The main differences between those systems and the one described in this paper is that the present system uses the TV-Anytime content and consumer metadata specifications for a complete management of TV programs and user profiles, and that the system uses additional information that exists in the TV-Anytime User Profile in order to avoid length clarification dialogues and help the user to get the most relevant answers at the top of the result list.

In section 2 of this paper the natural language model for digital TV environment is presented, along with the functionality provided and the representation of the information that the system collects from the user's input. In section 3 we present the algorithm for resolving the ambiguities instead of using clarification dialogues. In section 4 there is the analysis of the system architecture and of the modules that constitute it. Section 5 presents the implementation environment of the system and of the applications from the client side. In section 6 we present an example of a user's utterance and the actions taken by the system in order to satisfy the user's request. Finally section 7 presents the results of the system's evaluation based on user experiments and section 8 concludes by summarizing the content of this paper.

2 The Natural Language Model for the Digital TV Environment

The proposed Natural Language Model allows a user to determine the rules of management of digital TV data (programs and metadata), retrieve TV program content based on any information of its metadata description, express his preferences for the types of TV programs that will be stored, manage his selection list (i.e. programs that have been selected by the PDR or the user himself as candidates for recording), by creating his profile and modify any of the above.

The user's utterance is constituted by a combination of sub-phrase. The categories of these sub-phrases are Introduction phrases, to define the functionality, Search phrases, to define the TV-Anytime information, Target phrases, to define where each of the functions is targeting, Temporal phrases , to define phrases about date and time and Summary phrases, to define summaries with audio/visual content.

The structure that represents the information gathered by the user's utterance is shown in figure 1. This structure consists of three parts namely Element, Element Type and Element Value. The first structure part (Element) is used to differentiate the TV-Anytime metadata information (modeled as *TVA-properties*) from the information that directs the system to the correct management of the user's input (modeled as *flags*). The TV-Anytime information about date and time is modeled as *temporal* Elements. The second structure part (Element Type) is used in order to further specialize the aforementioned information and to obtain its corresponding Element Value (the third structure part), from the user's utterance. When a user inserts an utterance into the system, it generates a feature structure [10] that follows the structure of the model.

此为试读,需要完整PDF请访问: www.ertongbook.com