Lecture Notes in Artificial Intelligence 10259

Subseries of Lecture Notes in Computer Science

More information about this series at http://www.springer.com/series/1244

Annette ten Teije · Christian Popow
John H. Holmes · Lucia Sacchi (Eds.)

Artificial Intelligence in Medicine

16th Conference on Artificial Intelligence in Medicine, AIME 2017
Vienna, Austria, June 21–24, 2017
Proceedings

Springer

Editors
Annette ten Teije
Vrije Universiteit Amsterdam
Amsterdam
The Netherlands

John H. Holmes
University of Pennsylvania
Philadelphia, PA
USA

Christian Popow
Medical University of Vienna
Vienna
Austria

Lucia Sacchi
University of Pavia
Pavia
Italy

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-319-59757-7 ISBN 978-3-319-59758-4 (eBook)
DOI 10.1007/978-3-319-59758-4

Library of Congress Control Number: 2017943002

LNCS Sublibrary: SL7 – Artificial Intelligence

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The European Society for Artificial Intelligence in Medicine (AIME) was established in 1986 following a very successful workshop held in Pavia, Italy, the year before. The principal aims of AIME are to foster fundamental and applied research in the application of artificial intelligence (AI) techniques to medical care and medical research, and to provide a forum at biennial conferences for discussing any progress made. For this reason, the main activity of the society is the organization of a series of biennial conferences, which have been held in Marseilles, France (1987), London, UK (1989), Maastricht, The Netherlands (1991), Munich, Germany (1993), Pavia, Italy (1995), Grenoble, France (1997), Aalborg, Denmark (1999), Cascais, Portugal (2001), Protaras, Cyprus (2003), Aberdeen, UK (2005), Amsterdam, The Netherlands (2007), Verona, Italy (2009), Bled, Slovenia (2011), Murcia, Spain (2013), and Pavia, Italy (2015). This volume contains the proceedings of AIME 2017, the 16th Conference on Artificial Intelligence in Medicine, held in Vienna, Austria, June 21–24, 2017.

The AIME 2017 goals were to present and consolidate the international state of the art of AI in biomedical research from the perspectives of theory, methodology, systems, and applications. The conference included two invited lectures, full and short papers, tutorials, workshops, and a doctoral consortium.

In the conference announcement, authors were invited to submit original contributions regarding the development of theory, methods, systems, and applications for solving problems in the biomedical field, including AI approaches in biomedical informatics, molecular medicine, and health-care organizational aspects. Authors of papers addressing theory were requested to describe the properties of novel AI models potentially useful for solving biomedical problems. Authors of papers addressing theory and methods were asked to describe the development or the extension of AI methods, to address the assumptions and limitations of the proposed techniques, and to discuss their novelty with respect to the state of the art. Authors of papers addressing systems and applications were asked to describe the development, implementation, or evaluation of new AI-inspired tools and systems in the biomedical field. They were asked to link their work to underlying theory, and either analyze the potential benefits to solve biomedical problems or present empirical evidence of benefits in clinical practice.

AIME 2017 received 141 abstract submissions; 113 thereof were eventually submitted as complete papers. Submissions came from 35 countries, including 13 outside Europe. All papers were carefully peer-reviewed by experts from the Program Committee with the support of additional reviewers. Each submission was reviewed in most cases by three reviewers, and at least by two reviewers. The reviewers judged the overall quality of the submitted papers, together with their relevance to the AIME conference, originality, impact, technical correctness, methodology, scholarship, and quality of presentation. In addition, the reviewers provided detailed written comments on each paper, and stated their confidence in the subject area.

A small committee consisting of the AIME 2017 scientific chair, Annette ten Teije, the local organization chair, Christian Popow, John H. Holmes, doctoral consortium chair and AIME 2015 scientific chair, and Lucia Sacchi, AIME 2015 local organization co-chair, made the final decisions regarding the AIME 2017 scientific program. This process began with virtual meetings held monthly starting in March 2016. The process ended with a two-day face-to-face meeting of the committee in Vienna to assemble the final program.

As a result, 21 long papers (an acceptance rate of 22%) and 24 short papers (including demo papers) were accepted; one short paper was withdrawn. Each long paper was presented in a 25-minute oral presentation during the conference. Each regular short paper was presented in a five-minute presentation and by a poster. Each demo short paper was presented in a five-minute presentation and by a demo during the demo session. The papers were organized according to their topics in the following main themes: (1) Ontologies/Knowledge Representation (2) Bayesian Methods; (3) Temporal methods; (4) Nature Language Processing; (5) Health Care Processes; (6) Machine Learning; and (7) Demo's.

AIME 2017 had the privilege of hosting two invited speakers: Stefan Schulz, from the University of Graz, Austria, and Kenneth J. Barker, from T.J. Watson Research Center, IBM Research, New York, USA. In his keynote entitled "SNOMED CT: The Thorny Way Towards Interoperability of Clinical Routine Data" Stefan Schulz discussed the crucial role of the quality of the vocabularies and the annotation process for achieving data interoperability. The quality of terminology-annotated clinical data should be considered with realism, and the automated annotation approaches have to take into account human inter-annotator disagreement.

Ken Barker's keynote focused on intelligent question answer (QA) systems to support professionals in medicine and health care to explore the medical literature. In their approach the three main dimensions are context analysis, content management, and answer management. Furthermore, the collaborative setting plays a role in the learning capabilities of the adaptable QA system.

The doctoral consortium provided an opportunity for six PhD students to present their research goals, proposed methods, and preliminary results. A scientific panel consisting of experienced researchers in the field (Riccardo Bellazzi, Mor Peleg, David Riaño, Lucia Sacchi, Yuval Shahar, and Allan Tucker) provided constructive feedback to the students in an informal atmosphere. The doctoral consortium was chaired by John H. Holmes.

Four workshops were organized after the AIME 2017 main conference. These included the 9th International Workshop on Knowledge Representation for Health Care (KRH4C) and the 10th International Workshop on Process-Oriented Information Systems in Health Care (ProHealth), joined together for the second time at AIME. This workshop was chaired by David Riaño, Richard Lenz, Mor Peleg, and Manfred Reichert. A second full-day workshop was the Second Workshop on Extracting and Processing of Rich Semantics from Medical Texts, chaired by Kerstin Denecke, Yihan Deng, Thierry Declerck, and Frank van Harmelen. The third workshop was the Second Workshop on Artificial Intelligence for Diabetes, chaired by Clare Martin, Beatriz López, and Pau Herrero Vinas. The fourth workshop was the Workshop on Advanced

Predictive Models in Health Care organized by Niels Peek, Gregor Štiglic, Nophar Geifman, Petra Povalej Brzan, and Matthew Sperrin.

In addition to the workshops, five interactive half-day tutorials were presented prior to the AIME 2017 main conference:

(1) Natural Language Processing for Clinical Information Extraction (Stéphane Meystre, Meliha Yetisgen, Scott DuVall, Hua Xu); (2) Latest Speech and Signal Processing for Affective and Behavioral Computing in mHealth, (Bjorn Schuller, Bodgan Vlasenko, Hesam Sagha), (3) Evaluation of Prediction Models in Medicine (Ameen Abu-Hanna); (4) Medical Decision Analysis with Probabilistic Graphical Models (Francisco Javier Diez, Manuel Luque); (5) Clinical Fuzzy Control Systems and Fuzzy Automata with HL7's Clinical Decision Support Standard: The Fuzzy Arden Syntax (Jeroen de Bruin, Klaus-Peter Adlassnig).

We would like to thank everyone who contributed to AIME 2017. First of all, we would like to thank the authors of the papers submitted and the members of the Program Committee together with the additional reviewers. Thanks are also due to the invited speakers as well as to the organizers of the workshops, the tutorials and doctoral consortium. Many thanks go to the local Organizing Committee, who managed all the work making this conference possible. The free EasyChair conference system (http://www.easychair.org/) was an important tool supporting us in the management of submissions, reviews, selection of accepted papers, and preparation of the overall material for the final proceedings. We would like to thank our sponsors, who so generously supported the conference: the American Medical Informatics Association (AMIA/KDDM), the European Association for Artificial Intelligence (EurAI), and Springer. We thank IMIA for the recent endorsement of the AIME conference. Finally, we thank the Springer team for helping us in the final preparation of this LNAI book.

June 2017

Annette Ten Teije
Christian Popow
John H. Holmes
Lucia Sacchi

Organization

AIME Organization Team

Annette ten Teije	Vrije Universiteit Amsterdam, The Netherlands (Chair)
Christian Popow	Medical University of Vienna, Austria (Local chair)
John H. Holmes	University of Pennsylvania, USA (Doctoral Consortium Chair)
Lucia Sacchi	University of Pavia, Italy (Co-chair)

Artificial Intelligence in Medicine Board

Amar Das	The Dartmouth Institute, USA
Stefan Darmoni	University of Rouen, France
Milos Hauskrecht	University of Pittsburgh, USA
John Holmes	University of Pennsylvania, USA
Jose M. Juarez	University of Murcia, Spain
Mar Marcos	Universitat Jaume I, Castellón, Spain
Roque Marín Morales	University of Murcia, Spain
Stefania Montani	Università del Piemonte Orientale, Italy
Barbara Oliboni	University of Verona, Italy
Niels Peek	The University of Manchester, UK (Chair)
Mor Peleg	University of Haifa, Israel
Christian Popow	Medical University of Vienna, Austria
David Riaño	Universitat Rovira i Virgili, Spain
Lucia Sacchi	University of Pavia, Italy
Annette Ten Teije	Vrije Universiteit Amsterdam, The Netherlands
Paolo Terenziani	Università del Piemonte Orientale, Italy
Samson Tu	Stanford University, USA
Allan Tucker	Brunel University London, UK
Szymon Wilk	Poznan University of Technology, Poland
Blaz Zupan	University of Ljubljana, Slovenia

Program Committee

Ameen Abu-Hanna	AMC-UvA, The Netherlands
Klaus-Peter Adlassnig	Medical University of Vienna, Austria
Laura Barnes	University of Virginia, USA
Riccardo Bellazzi	University of Pavia, Italy
Henrik Boström	Stockholm University, Sweden
Carlo Combi	Università degli Studi di Verona, Italy
Arianna Dagliati	University of Pavia, Italy
Stefan Darmoni	University of Rouen, France

Kerstin Denecke	Bern University of Applied Sciences, Switzerland
Barbara Di Camillo	University of Padova, Italy
Michel Dojat	INSERM, France
Georg Dorffner	Medical University Vienna, Austria
Paulo Felix	USC, Spain
Jesualdo Tomás Fernández-Breis	Universidad de Murcia, Spain
Catherine Garbay	CNRS, LIG, France
Natalia Grabar	STL CNRS, Université Lille 3, France
Adela Grando	Arizona State University, USA
Milos Hauskrecht	University of Pittsburgh, USA
Zhe He	Florida State University, USA
Pedro Henriques Abreu	FCTUC-DEI/CISUC, Spain
John Holmes	University of Pennsylvania, USA (Doctoral Consortium Chair)
Arjen Hommersom	Open University of the Netherlands, The Netherlands
Jose M. Juarez	University of Murcia, Spain
Charles Kahn	University of Pennsylvania, USA
Eleni Kaldoudi	Democritus University of Thrace, Greece
Elpida Keravnou-Papailiou	University of Cyprus, Cyprus
John Kinsella	University of Glasgow, UK
Haridimos Kondylakis	Institute of Computer Science, FORTH, Greece
Pedro Larranaga	University of Madrid, Spain
Nada Lavrač	Jozef Stefan Institute, Slovenia
Michael Liebman	IPQ Analytics, LLC, USA
Helena Lindgren	Umeå University, Sweden
Peter Lucas	Radboud University Nijmegen, The Netherlands
Beatriz López	University of Girona, Spain
Mar Marcos	Universitat Jaume I, Castellón, Spain
Michael Marschollek	Peter L. Reichertz Institute for Medical Informatics, Germany
Roque Marín Morales	University of Murcia, Spain
Paola Mello	University of Bologna, Italy
Silvia Miksch	Vienna University of Technology, Austria
Diego Molla	Macquarie University, Australia
Stefania Montani	Università del Piemonte Orientale, Italy
Robert Moskovitch	Deutsche Telekom Laboratories at Ben-Gurion University, Israel
Laura Moss	University of Aberdeen, UK
Fleur Mougin	ERIAS, INSERM U1219, Université de Bordeaux, France
Anthony Nguyen	The Australian e-Health Research Centre, Australia
Øystein Nytrø	Norwegian University of Science and Technology, Norway
Barbara Oliboni	University of Verona, Italy
Enea Parimbelli	University of Pavia, Italy

Niels Peek	The University of Manchester, UK
Mor Peleg	University of Haifa, Israel
Christian Popow	Medical University of Vienna, Austria (Local Chair)
Cédric Pruski	Luxembourg Institute of Science and Technology, Luxembourg
Silvana Quaglini	University of Pavia, Italy
David Riaño	Universitat Rovira i Virgili, Spain
Pedro Pereira Rodrigues	University of Porto, Portugal
Lucia Sacchi	University of Pavia, Italy (Co-chair)
Aleksander Sadikov	University of Ljubljana, Slovenia
Stefan Schulz	Medical University of Graz, Austria
Brigitte Seroussi	Assistance Publique, Hôpitaux de Paris, France
Yuval Shahar	Ben Gurion University, Israel
Erez Shalom	Ben Gurion University, Israel
Constantine Spyropoulos	NCSR Demokritos, Greece
Gregor Štiglic	University of Maribor, Slovenia
Annette Ten Teije	Vrije Universiteit Amsterdam, The Netherlands (Chair)
Paolo Terenziani	Università del Piemonte Orientale, Italy
Allan Third	The Open University, UK
Samson Tu	Stanford University, USA
Allan Tucker	Brunel University London, UK
Ryan Urbanowicz	University of Pennsylvania, USA
Frank Van Harmelen	Vrije Universiteit Amsterdam, The Netherlands
Alfredo Vellido	Universitat Politècnica de Catalunya, Spain
Szymon Wilk	Poznan University of Technology, Poland
Blaz Zupan	University of Ljubljana, Slovenia
Pierre Zweigenbaum	LIMSI, CNRS, Université Paris-Saclay, France

Additional Reviewers

Elias Alevizos
Mawulolo Ameko
Luca Anselma
Michael Barrowman
Elena Bellodi
Miguel Belmonte
Asim Bhuta
Alessio Bottrighi
Marcos Luiz de Paula Bueno
Joana Ferreira
Paolo Fraccaro
Vida Groznik
Tianyong Hao
Xiang Ji
Aida Kamisalic
Lefteris Koumakis
Thomas Kupka
Siqi Liu
Daniela Loreti
Begoña Martinez-Salvador
Dragana Miljkovic
Christopher Ochs
Bruno Oliveira
Pierpaolo Palumbo
Luca Piovesan
Vid Podpecan
Vassiliki Rentoumi
Fabrizio Riguzzi
Carla Rognoni
Stelios Sfakianakis
Matthew Sperrin
Grigorios Tzortzis
Natalia Viani
Ute von Jan
Yonghui Wu
Jinghe Zhang

Doctoral Consortium Committee

Riccardo Bellazzi	University of Pavia, Italy
John Holmes	University of Pennsylvania, USA (Chair)
Mor Peleg	University of Haifa, Israel
David Riaño	Universitat Rovira i Virgili, Spain
Lucia Sacchi	University of Pavia, Italy
Yuval Shahar	Ben Gurion University, Israel
Allan Tucker	Brunel University London, UK

Workshops

9th International Workshop on Knowledge Representation for Health Care (KRH4C) and the 10th International Workshop on Process-oriented Information Systems in Health Care (ProHealth)

Co-chairs

David Riaño	Universitat Rovira i Virgili, Spain
Richard Lenz	University of Erlangen-Nuremberg, Germany
Mor Peleg	University of Haifa, Israel
Manfred Reichert	University of Ulm, Germany

Second International Workshop on Extraction and Processing of Rich Semantics from Medical Texts

Co-chairs

Kerstin Denecke	Bern University of Applied Sciences, Switzerland
Yihan Deng	Bern University of Applied Sciences, Switzerland
Thierry Declerck	Saarland University and German Research Center for Artificial Intelligence, Germany
Frank van Harmelen	Vrije Universiteit Amsterdam, The Netherlands

Second Workshop on Artificial Intelligence for Diabetes

Co-chairs

Clare Martin	Oxford Brookes University, UK
Beatriz López	University of Girona, Spain
Pau Herrero Vinas	Imperial College London, UK

Workshop on Advanced Predictive Models in Health Care

Co-chairs

Niels Peek	The University of Manchester, UK
Gregor Štiglic	University of Maribor, Slovenia
Nophar Geifman	The University of Manchester, UK
Petra Povalej Brzan	University of Maribor, Slovenia
Matthew Sperrin	The University of Manchester, UK

Tutorials

Natural Language Processing for Clinical Information Extraction

Stéphane Meystre	Medical University of South Carolina, USA
Meliha Yetisgen	University of Washington, USA
Scott DuVall	University of Utah and Department of Veterans Affairs Salt Lake City Health Care System, USA
Hua Xu	University of Texas, USA

Latest Speech and Signal Processing for Affective and Behavioral Computing in mHealth

Björn Schuller	Imperial College London, UK
Bodgan Vlasenko	University of Passau, Germany
Hesam Sagha	audEERING GmbH, Germany

Evaluation of Prediction Models in Medicine

Ameen Abu-Hanna	University of Amsterdam, The Netherlands

Medical Decision Analysis with Probabilistic Graphical Models

Francisco Javier Díez	UNED, Spain
Manuel Luque	UNED, Spain

Clinical Fuzzy Control Systems and Fuzzy Automata with HL7's Clinical Decision Support Standard: The Fuzzy Arden Syntax

Jeroen de Bruin	Medical University of Vienna, Austria
Klaus-Peter Adlassnig	Medical University of Vienna, Austria

Sponsors


AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.


EurAi


IMIA


Springer

Invited Talks

SNOMED CT: The Thorny Way Towards Interoperability of Clinical Routine Data

Stefan Schulz

Institute of Medical Informatics, Statistics and Documentation,
Medical University of Graz, Graz, Austria
stefan.schulz@medunigraz.at

Abstract. To achieve the goal of data interoperability, the quality of the vocabularies used as well as the coding/annotation processes involved are crucial. Inter-coder agreement is a major quality criterion. In the project ASSESS-CT, the clinical terminology SNOMED CT was assessed for its usefulness regarding manual and machine annotations of clinical texts. Due to the low inter-annotator agreement obtained a manual analysis of disagreements was done.

Keywords: Biomedical ontologies • SNOMED CT • Semantic annotation

The EU support action ASSESS CT (Assessing SNOMED CT for Large Scale eHealth Deployments in the EU) [1] aimed at collecting empirical evidence for the fitness of SNOMED CT, compared to other terminology scenarios. A series of manual and machine annotation experiments was performed in order to measure terminology coverage [2] in clinical models and narratives. One striking result was a generally low inter-annotator agreement when clinical text samples from different domains and languages were annotated by domain experts. Although detailed annotation guidelines had been elaborated before, Krippendorf's Alpha was only 37% when annotating English texts with SNOMED CT.

An in-depth analysis of disagreements between annotators yielded several factors that affect consistent. In the following, frequent reasons of disagreement are listed:

- **Human factors**: ignoring the guideline, slips because of haste or carelessness;
- **Tooling**: Retrieval problems due to poor tool support or lacking entry terms;
- **Lack of textual definitions**. This is a typical problem in SNOMED CT, especially when concepts are semantically close (e.g. *Significant* vs. *Severe*, or *Worried* vs. *Anxiety*);
- **Lack of formal definitions**. Fully defined SNOMED CT concepts could support ex-post reconciliation of annotation disagreements. However, definitional axioms were often found missing, e.g. in the primitive concept *Diabetic monitoring (regime/therapy)*, despite the existence of *Diabetes mellitus (disorder)* and *Monitoring - action (qualifier value)*;
- **Logical polysemy** [3], e.g. *Malignant lymphoma (disorder)* vs. *Malignant lymphoma (morphological abnormality)* or *Liver structure (body structure)* vs. *Liver (body structure);*

- **Navigational concepts**, e.g. *Finding of measures of palpebral fissure (finding)* is a common parent of the pre-coordinated findings, whereas *Measure of palpebral fissure (observable entity)* is to be completed by a qualifier.

These factors can be mitigated by better annotator training, refinement of the guideline, completion of SNOMED CT by textual and formal definitions as well as by flagging of concepts with purely navigational character. Retrieval performance of term browsers can be enhanced, e.g. by fuzzy match and more synonyms. Logical polysemy can be dealt with by postprocessing (addition of inferred concepts).

More difficult are those factors that derive from the characteristics of clinical language, especially as a consequence of its high compactness:

- **Context** (deixis, anaphora) such as in "These ailments have substantially increased since October". Here, "increased" belongs to something in an earlier sentence, which require some reasoning by the reader. The same applies to
- **Co-ordination and negation** such as in "normal factors 5, 9, 10, and 11", or "no tremor, rigidity or bradykinesia";
- **Lexical ambiguity**: One text sample started with "IV:", which could mean the fourth item on a list, the abbreviation for "intravenous" or the 4th cerebral nerve. Here also, the understanding of the context is fundamental;
- **Scope**: in the previous example, even after disambiguation, there was disagreement regarding *Trochlear nerve structure (body structure)* vs. *Exploration of trochlear nerve (IV) (procedure).*

These latter factors require a real understanding of the text, which constitutes an ongoing challenge for automated (NLP-based) annotation. Regarding the scope of annotation, an enhanced guideline could recommend, e.g., that anatomical entities, devices and organisms are, if possible, always seen as modifiers of a "head" concept, normally a procedure or a finding, which however, requires an advance parsing not only of the whole discourse structure.

One conclusion of this work is that the quality of terminology-annotated clinical data should be considered with realism, especially if they trigger intelligent systems. Another conclusion is that the creation of annotation gold standards is inherently difficult. The assessment of the performance of automated (NLP) annotation approaches must therefore always take human inter-annotator disagreement into account.

References

1. SNOMED International. What is SNOMED CT? (2017). http://www.snomed.org/snomed-ct/what-is-snomed-ct
2. Kalra, D., Schulz, S., Karlsson, D., et al.: ASSESS CT Recommendations (2016). http://assess-ct.eu/fileadmin/assess_ct/final_brochure/assessct_final_brochure.pdf
3. Arapinis, A., Vieu, L.: A plea for complex categories in ontologies. Appl. Ontology, **10**(3–4), 285–296 (2015)

Collaborative, Exploratory Question Answering Against Medical Literature

Ken Barker

IBM Research, Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA
kjbarker@us.ibm.com

Background

The amount of text being generated in technical domains such as medicine and healthcare has skyrocketed, making broad, individual, independent mastery of such domains impossible. More than ever, intelligent information gathering tools are essential to support any task that requires consuming and acting on such knowledge. Decision makers require more intelligent Question Answering (QA) systems to support decision making. In particular, professionals in medicine and healthcare could benefit from advanced tools to help explore literature intelligently in the context of specific, complex cases. Existing QA systems suffer from several weaknesses:

- They assume that the decision problem can be directly formulated as a question and that the answer can directly inform the decision to be made. But users may need to supply much context and explore relevant information before even knowing what questions to ask.
- They assume questions are independent. But the decision making process may require a series of questions, influenced by the answers to prior questions. Users' questions and the problem state may change through interacting with the system.
- They assume that questions have definitive answers in a corpus. But the question may be ill-specified, requiring the system to generate passages of text. These are broader than a simple phrase (the focus of factoid-based QA systems) but narrower than an entire document (the focus of web search).
- They do not expose their interpretation of the question or how the QA system arrives at a particular answer. Without this information, the decision maker is left to guess where the QA system may have gone wrong. This guessing at opaque system behavior leads to a painful trial-and-error approach to information gathering.
- There is no way to guide their behavior directly. Even if the interpretation were transparent, there is no way for the user to change or correct the interpretation.
- They do not adapt to the decision maker's needs or learn from interactions. QA systems are generally trained prior to deployment using sets of questions and corresponding answers (ground truth).

To address these weaknesses, we conduct research on technology for QA systems that are *Contextual*, *Collaborative*, *Transparent*, *Guidable*, *Adaptable*, and *Proactive*. More simply, we are investigating "Collaborative, Exploratory Question Answering".

Research in Collaborative, Exploratory Question Answering

Our approach to Question Answering involves research in several traditional areas of automated language analysis. We group the research into three main dimensions: *Context Analysis*, *Content Management*, and *Answer Management*. Furthermore, in a collaborative setting, adaptable systems should use all user utterances, gestures and feedback as instances for learning.

Context Analysis

We expect users of our Collaborative, Exploratory QA systems to supply background context and questions in natural language. Context analysis includes applying Natural Language Processing and Learning to understand the user's utterances, determine appropriate actions to take, and formulate the queries needed by the agents who will carry out those actions. Specific research projects include utterance classification, semantic interpretation, concept clustering, dialog, and hypothesis generation.

Content Management

Content management involves finding appropriate unstructured text and structured knowledge resources for the information goals that users have. It includes defining models of the content, crafting appropriate ingestion and indexing strategies, and search. Our research projects in Content Management include combining structured and unstructured sources, term and concept weighting, conceptual query expansion, and lexical, syntactic and semantic indexing and search,

Answer Management

Non-factoid Question Answering poses unique challenges for Answer Management: creating relevant, coherent units from search results, aggregating and visualizing them, and providing answer justification. We conduct specific research in passage segmentation, passage similarity, results clustering, emergent concept induction, and learning from user manipulation of result sets.

An Interactive Platform for Exploratory QA

We have built an interactive Exploratory Question Answering system to serve as a platform for investigating the research challenges described above. The chat-based system allows users to specify background context, to examine and edit the results of interpretation, to state hypotheses and ask questions, and to explore and interact with the results of question answering.

Contents

Demo's