

Lecture Notes in Artificial Intelligence 5847

Edited by R. Goebel, J. Siekmann, and W. Wahlster

Subseries of Lecture Notes in Computer Science

Sobha Lalitha Devi  
António Branco  
Ruslan Mitkov (Eds.)

# Anaphora Processing and Applications

7th Discourse Anaphora  
and Anaphor Resolution Colloquium, DAARC 2009  
Goa, India, November 5-6, 2009  
Proceedings

## Series Editors

Randy Goebel, University of Alberta, Edmonton, Canada  
Jörg Siekmann, University of Saarland, Saarbrücken, Germany  
Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

## Volume Editors

Sobha Lalitha Devi  
Anna University – K.B. Chandrasekhar Research Centre  
MIT Campus of Anna University, Chromepet, Chennai 600044, India  
E-mail: sobha@au-kbc.org

António Branco  
Universidade de Lisboa  
Faculdade de Ciências  
Departamento de Informática  
Cidade Universitária, 1749-016 Lisboa, Portugal  
E-mail: antonio.branco@di.fc.ul.pt

Ruslan Mitkov  
University of Wolverhampton  
School of Humanities, Languages and Social Studies  
Research Group in Computational Linguistics  
Wolverhampton WV1 1SB, UK  
E-mail: r.mitkov@wlv.ac.uk

Library of Congress Control Number: 2009936009

CR Subject Classification (1998): I.2.7, I.2, I.7, F.4.3, H.5.2, H.3

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743  
ISBN-10 3-642-04974-5 Springer Berlin Heidelberg New York  
ISBN-13 978-3-642-04974-3 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.com

© Springer-Verlag Berlin Heidelberg 2009  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper SPIN: 12775965 06/3180 5 4 3 2 1 0

## Preface

Distribution of anaphora in natural language and the complexity of its resolution have resulted in a wide range of disciplines focusing their research on this grammatical phenomenon. It has emerged as one of the most productive topics of multi- and interdisciplinary research such as cognitive science, artificial intelligence and human language technology, theoretical, cognitive, corpus and computational linguistics, philosophy of language, psycholinguistics and cognitive psychology. Anaphora plays a major role in understanding a language and also accounts for the cohesion of a text. Correct interpretation of anaphora is necessary in all high-level natural language processing applications.

Given the growing importance of the study of anaphora in the last few decades, it has emerged as the frontier area of research. This is evident from the high-quality submissions received for the 7<sup>th</sup> DAARC from where the 10 excellent reports on research findings are selected for this volume. These are the regular papers that were presented at DAARC.

Initiated in 1996 at Lancaster University and taken over in 2002 by the University of Lisbon, and moved out of Europe for the first time in 2009 to Goa, India, the DAARC series established itself as a specialised and competitive forum for the presentation of the latest results on anaphora processing, ranging from theoretical linguistic approaches through psycholinguistic and cognitive work to corpus studies and computational modelling. The series is unique in that it covers this research subject from a wide variety of multidisciplinary perspectives while keeping a strong focus on automatic anaphor resolution and its applications.

The programme of the 7<sup>th</sup> DAARC was selected from 37 initial submissions. It included 19 oral presentations and 8 posters from over 50 authors coming from 14 countries: Belgium, Czech Republic, Denmark, Germany, India, Norway, Portugal, Russia, Romania, Spain, The Netherlands, Taiwan, UK and the USA. The submissions were anonymised and submitted to a selection process by which each received three evaluation reports by experts from the programme committee listed below.

On behalf of the Organising Committee, we would like to thank all the authors who contributed with their papers for the present volume and all the colleagues in the Programme Committee for their generous and kind help in the reviewing process of DAARC, and in particular of the papers included in the present volume. Without them neither this colloquium nor the present volume would have been possible.

Chennai, August 2009

Sobha Lalitha Devi  
António Branco  
Ruslan Mitkov

# Organisation

The 7<sup>th</sup> DAARC colloquium was organized by AU-KBC Research Centre, Anna University Chennai, Computational Linguistics Research Group.

## Organising Committee

António Branco	University of Lisbon
Sobha Lalitha Devi	Anna University Chennai
Ruslan Mitkov	University of Wolverhampton

## Programme Committee

Alfons Maes	Tilburg University
Andrej Kibrik	Russian Academy of Sciences
Andrew Kehler	University of California San Diego
Anke Holler	University of Goettingen
António Branco	University of Lisbon
Christer Johansson	Bergen University
Claire Cardie	Cornell University
Constantin Orasan	University of Wolverhampton
Costanza Navarretta	University of Copenhagen
Dan Cristea	University of Iasi, Romania
Elsi Kaiser	University of Southern California
Eric Reuland	Utrecht Institute of Linguistics
Francis Cornish	University of Toulouse-Le Mirail
Georgiana Puscasu	University of Wolverhampton
Graeme Hirst	University of Toronto
Iris Hendrickx	Antwerp University
Jeanette Gundel	University of Minnesota
Jeffrey Runner	University of Rochester
Joel Tetreault	Education Testing Service, USA
Jos Van Berkum	Max Planck Institute for Psycholinguistics
José Augusto Leitão	University of Coimbra
Kavi Narayana Murthy	Hyderabad Central University
Klaus Von Heusinger	Konstanz University
Lars Hellan	Norwegian University of Science and Technology
Maria Mercedes Piñango	Yale University
Marta Recasens	University of Barcelona
Martin Everaert	Utrecht Institute of Linguistics
Massimo Poesio	University of Essex
Patricio Martinez Barco	University of Alicante

Peter Bosch	University of Osnabrueck
Petra Schumacher	University of Mainz
Renata Vieira	PUC Rio Grande do Sul
Richard Evans	University of Wolverhampton
Robert Dale	Macquarie University
Roland Stuckardt	University of Frankfurt am Main
Ruslan Mitkov	University of Wolverhampton
Sergey Avrutin	Utrecht Institute of Linguistics
Shalom Lappin	King's College, London
Sivaji Bandopadhyaya	Jadavpur University
Sobha Lalitha Devi	Anna University Chennai
Tony Sanford	Glasgow University
Veronique Hoste	Gent University
Vincent Ng	University of Texas at Dallas
Yan Huang	University of Auckland

# Table of Contents

## Resolution Methodology

Why Would a Robot Make Use of Pronouns? An Evolutionary Investigation of the Emergence of Pronominal Anaphora . . . . .	1
<i>Dan Cristea, Emanuel Dima, and Corina Dima</i>	
Automatic Recognition of the Function of Singular Neuter Pronouns in Texts and Spoken Data . . . . .	15
<i>Costanza Navarretta</i>	
A Deeper Look into Features for Coreference Resolution . . . . .	29
<i>Marta Recasens and Eduard Hovy</i>	

## Computational Applications

Coreference Resolution on Blogs and Commented News . . . . .	43
<i>Iris Hendrickx and Veronique Hoste</i>	
Identification of Similar Documents Using Coherent Chunks . . . . .	54
<i>Sobha Lalitha Devi, Sankar Kuppan, Kavitha Venkataswamy, and Patabhi R.K. Rao</i>	

## Language Analysis

Binding without Identity: Towards a Unified Semantics for Bound and Exempt Anaphors . . . . .	69
<i>Eric Reuland and Yoad Winter</i>	
The Doubly Marked Reflexive in Chinese . . . . .	80
<i>Alexis Dimitriadis and Min Que</i>	

## Human Processing

Definiteness Marking Shows Late Effects during Discourse Processing: Evidence from ERPs . . . . .	91
<i>Petra B. Schumacher</i>	
Pronoun Resolution to Commanders and Recessors: A View from Event-Related Brain Potentials . . . . .	107
<i>José Augusto Leitão, António Branco, Maria Mercedes Piñango, and Luís Pires</i>	

Effects of Anaphoric Dependencies and Semantic Representations on  
Pronoun Interpretation ..... 121  
    *Elsi Kaiser*

**Author Index** ..... 131