Coreference and Its Applications

Proceedings of the Workshop

Edited by
Amit Bagga, Breck Baldwin, and Sara Shelton

22 June 1999
University of Maryland
College Park, Maryland, USA

Published by the Association for Computational Linguistics

Order additional copies from:

Association for Computational Linguistics
75 Paterson Street, Suite 9
New Brunswick, NJ 08901 USA
+1-732-342-9100 phone
+1-732-342-9339 fax
acl@aclweb.org

PREFACE

Coreference is in some sense nature's own hyperlink. It conveys how individual statements are connected within documents, across documents and across bodies of human knowledge. Consequently coreference resolution algorithms are at the core of Natural Language Processing. Most of the work done on coreference deals with a single language and a single text document (usually newswire).

As NLP research matures into "application" phases (as opposed to theory-development), NLP systems are moving beyond traditional research sources to document sets which reflect a more natural, research-oriented mix. This shift can be seen in both the document sets and tasks used in recent HUB, MET, and TDT evaluations. The new sources consist of documents in several different languages, documents with data from noisy sources, and documents containing multimedia. In order for NLP systems to make a successful transition to these new sources, it is critical for coreference resolution systems to also work on these new sources.

The Workshop on Coreference and Its Applications, held on June 22nd, 1999 at the University of Maryland, College Park, Maryland, USA, was organized around the goals of discussing, promoting, and presenting new research results (positive and negative) regarding the theory, design and evaluation of coreference resolution systems that deal with non-traditional data sources. In particular, the goal of the workshop was to focus on systems dealing with the following types of coreference:

- Cross-document coreference
- Coreference resolution in languages other than English
- Coreference resolution on noisy data
- Coreference resolution on non-text data (example: human speech)
- Coreference resolution on multimedia data

In addition, the workshop also focused on innovative NLP applications that rely heavily on coreference resolution systems. The workshop was sponsored by the Association for Computational Linguistics (ACL).

We would like to thank all authors who showed their interest by submitting papers to the workshop. We would also like to thank the members of the program committee: Branimir Boguraev (IBM Research), Ed Hovy (USC Information Sciences Institute), Mark T. Maybury (MITRE), and Ruslan Mitkov (University of Wolverhampton).

Amit Bagga
Breck Baldwin
Sara Shelton

Workshop Program

9:00--9:05	***Opening***	
Session: Cross-Document Coreference		
9:05--9:30	Amit Bagga and Breck Baldwin, *General Electric CRD and University of Pennsylvania*	Cross-Document Event Coreference: Annotations, Experiments, and Observations
9:30--9:55	Yael Ravin and Zunaid Kazi, *IBM T.J. Watson Research Center*	Is Hillary Rodham Clinton the President? Disambiguating Names Across Documents
Session: Multimedia Coreference		
9:55--10:20	Koichi Yamada, Kazunari Sugiyama, Yasunori Yonamine and Hiroshi Nakagawa, *Yokohama National University*	Identification of Coreference Between Names and Faces
10:20--10:50	***Coffee Break***	
10:50--11:15	Utiyama Masao and Hasida Koiti, *Shinshu University and Electrotechnical Laboratory (ETL)*	Automatic Slide Presentation from Semantically Annotated Documents
Session: Coreference in Non-English Languages		
11:15--11:40	Masaki Murata, Hitoshi Isahara and Makoto Nagao, *Communications Research Laboratory, and Kyoto University*	Resolution of Indirect Anaphora in Japanese Sentences Using Examples "X *no* Y (Y OF X)"
11:40--12:05	Masaki Murata, Hitoshi Isahara and Makoto Nagao, *Communications Research Laboratory, and Kyoto University*	Pronoun Resolution in Japanese Sentences Using Surface Expressions and Examples
12:05--13:30	***Lunch Break***	
13:30--13:55	Michael Paul, Kazuhide Yamamoto and Eiichiro Sumita, *ATR Interpreting Telecommunications Research Laboratory*	Corpus-Based Anaphora Resolution Towards Antecedent Preference
13:55--14:20	Marco Rocha, *Universidade Federal de Santa Catarina*	Coreference resolution in dialogues in English and Portuguese
14:20--14:45	Pedro Amo, Francisco L. Ferreras, Fernando Cruz and Saturnino Maldonado, *Universidad de Alcala*	Orthographic Co-Reference Resolution Between Proper Nouns Through the Calculation of the Relation of "Replicancia"
14:45--15:10	Jesus Peral, Manuel Palomar and Antonio Ferrandez, *University of Alicante*	Coreference-oriented Interlingual Slot Structure & Machine Translation
15:10--15:45	***Coffee Break***	
Session: Applications of Coreference		
15:45--16:10	Saliha Azzam, Kevin Humphreys and Robert Gaizauskas, *University of Sheffield*	Using Coreference Chains for Text Summarization

16:10--16:35	Thomas S. Morton, *University of Pennsylvania*	Using Coreference for Question Answering
Session: Coreference Annotation		
16:35--17:00	Kees van Deemter and Rodger Kibble, *University of Brighton*	What is coreference, and what should coreference annotation be?

Table of Contents