Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts

Proceedings of a Workshop
Sponsored by the
Association for Computational Linguistics

Edited by
Ruslan Mitkov
and
Branimir Boguraev

11 July 1997
Universidad Nacional de Educación a Distancia
Madrid, Spain

Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts

Proceedings of a Workshop
Sponsored by the
Association for Computational Linguistics

Edited by
Ruslan Mitkov
and
Branimir Boguraev

11 July 1997
Universidad Nacional de Educación a Distancia
Madrid, Spain

© 1997, Association for Computational Linguistics

Order additional copies from:

ACL
P.O. Box 6090
Somerset, NJ, 08875 USA
+1-908-873-3898
acl@bellcore.com

WORKSHOP ORGANISERS

Ruslan Mitkov (University of Wolverhampton)

Branimir Boguraev (Apple Computer, Cupertino)

WORKSHOP PROGRAMME COMMITTEE

Breck Baldwin (University of Pennsylvania, Philadelphia)

Branimir Boguraev (Apple Computer, Cupertino)

David Carter (SRI, Cambridge)

Megumi Kameyama (SRI, Menlo Park)

Christopher Kennedy (University of California, Santa Cruz)

Shalom Lappin (University of London)

Susan LuperFoy (MITRE Corporation, McLean)

Tony McEnery (Lancaster University)

Ruslan Mitkov (University of Wolverhampton)

Celia Rico Pérez (Universidad Europea de Madrid)

Frederique Segond (Rank Xerox Research Centre, Grenoble)

Sandra Williams (BT Research Labs, Ipswich)

PREFACE

In recent years we have witnessed an increased interest in anaphora resolution. We have also seen a tendency towards the development of robust and effective methods to meet the demand of practical NLP systems and to enhance further the automatic processing of growing language resources. This clear trend towards inexpensive, relatively knowledge-poor, corpus-based methods has been evident in recent proposals which address the anaphora resolution problem by deliberately limiting the extent to which they rely on domain and/or linguistic knowledge, and by moving away from the traditional domain/sublanguage restriction.

We believe that the time is right for the ACL'97/EACL'97 workshop "Operational factors in practical, robust anaphora resolution for unrestricted texts". It presents the latest work on the practical requirements of operational and reliable anaphora resolution components. It also investigates the role of, and interactions among, the various factors in anaphora resolution - in particular those that scale well, or that translate easily to knowledge-poor environments.

We are pleased to report a high level of interest in the workshop, which resulted in a large number of quality submissions. Regrettably, this translated into some hard choices concerning the final programme make-up. There can only be a limited number of presentations in a one day workshop; thus, following a review process where every submission was reviewed by three members of the Programme Committee, some good papers did not find a slot in the workshop programme.

We trust that the workshop will be a timely forum for fruitful discussion and that it will give a strong impetus to further advances in the field.

Ruslan Mitkov and Branimir Boguraev

May, 1997

TABLE OF CONTENTS

Massimo POESIO, Renata VIEIRA, Simone TEUFEL -
Resolving bridging references in unrestricted text 1

Tatsunori MORI, Mamoru MATSUO, Hiroshi NAKAGAWA -
Constraints and defaults of zero pronouns in Japanese instruction manuals 7

Ruslan MITKOV -
Factors in anaphora resolution: they are not the only things that matter 14
A case study based on two different approaches.

Hiromi NAKAIWA -
Automatic extraction of rules for anaphora resolution of Japanese 22
zero pronouns from aligned sentence pairs

Roland STUCKARDT -
Resolving anaphoric references on deficient syntactic descriptions 30

Breck BALDWIN -
CogNIAC: high precision coreference with limited knowledge and linguistic resources 38

Megumi KAMEYAMA -
Recognizing referential links: an information extraction perspective 46

Marco ROCHA -
Supporting anaphor resolution in dialogues with a corpus-based probabilistic model 54

Sabine BERGLER -
Towards reliable partial anaphora resolution 62

Tony McENERY, Izumi TANAKA, Simon BOTLEY -
Corpus annotation and reference resolution 67

Kevin HUMPHREYS, Robert GAIZAUSKAS, Saliha AZZAM -
Event coreference for information extraction 75

Ruslan MITKOV -
How far are we from (semi-)automatic annotation of anaphoric links in corpora? 82

Kaja BORTHEN, Thorstein FRETHEIM, Jeanette K. GUNDEL -
What brings a higher-order entity into focus of attention? 88
Sentential pronouns in English and Norwegian.

Andrei POPESCU-BELIS, Isabelle ROBBA -
Cooperation between pronoun and reference resolution for unrestricted texts 94