

Tagging Text with Lexical Semantics

Why, What, and How?

Proceedings of the Workshop

4-5 April 1997

Washington, D.C., USA

Sponsored by the

Association for Computational Linguistics

Special Interest Group on the Lexicon

Published by the Association for Computational Linguistics

Tagging Text with Lexical Semantics

Why, What, and How?

Proceedings of the Workshop

4-5 April 1997

Washington, D.C., USA

Sponsored by the

Association for Computational Linguistics

Special Interest Group on the Lexicon

Published by the Association for Computational Linguistics

©1997, Association for Computational Linguistics

Order additional copies from:

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

PREFACE

The general topic of this volume is tagging word tokens with lexical semantic information. The first six papers discuss what information should be tagged and how human tagging (hand-tagging) can be performed so as to maximize accuracy. Hand-tagging is important because it can be used to create standard corpora on which to train and evaluate automated methods for tagging lexical semantic information. The next seven papers concentrate on such automated methods. Automatic tagging should prove useful in a variety of areas including machine translation, information retrieval, information extraction, and dialog understanding.

Although the general topic is tagging lexical semantic information, the majority of the papers concentrate on a more specific problem: word sense disambiguation. However, particular word senses often entail general lexical semantic information. For example, if the sense of *bank* used refers to something that lies along a river, then one knows that it is a natural formation, inanimate, etc. Consider also the verb *bounce* which in *she bounced a check* is a telic and in *she bounced a ball* is an activity. Such entailments are especially transparent when using a structured lexicon such as WordNet, which is the case for many of the papers in the volume.

The workshop itself and these proceedings are the result of many people's efforts. The people who submitted articles provided high quality papers to choose from, the program committee provided in depth reviews and did so punctually, the presenters and workshop leaders provided polished final documents in a timely fashion, and Priscilla Rasmussen answered many questions regarding money matters, hotel reservations, etc. To all these people, thanks! In addition, a very special thanks to Martha Palmer who both served on the program committee and acted as a ghost co-chair, providing general guidance and answering questions on a myriad of topics.

The members of the program committee should not go unnamed: Martha Evens, Helmut Feldweg, Michael Johnston, Douglas Jones, Kevin Knight, Boyan Onyshkevych, Martha Palmer, Philip Resnik, Evelyne Viegas, David Yarowsky, and Annie Zaenen.

Finally, I would like to gratefully acknowledge the US Government inter-agency TIPSTER Text Program which provided partial funding for the workshop.

Marc Light
Program Chair
March, 1997
University of Tübingen, Germany

TABLE OF CONTENTS

Presentations

<i>Getting serious about word sense disambiguation (invited talk)</i> Hwee Tou Ng	1
<i>Experience in WordNet sense tagging in the Wall Street Journal</i> Janyce Wiebe, Julie Maples, Lei Duan, and Rebecca Bruce	8
<i>Desiderata for tagging with WordNet synsets or MCCA categories</i> Kenneth C. Litkowski	12
<i>A frame-semantic approach to semantic annotation</i> John B. Lowe, Collin F. Baker, and Charles J. Fillmore	18
<i>A lexicon for underspecified semantic tagging</i> Paul Buitelaar	25
<i>Analysis of a hand-tagging task (invited talk)</i> Christiane Fellbaum, Joachim Grabowski, and Shari Landes	34
<i>Measuring semantic entropy</i> I. Dan Melamed	41
<i>Sense tagging: semantic tagging with a lexicon</i> Yorick Wilks and Mark Stevenson	47
<i>Selectional preference and sense disambiguation</i> Philip Resnik	52
<i>Investigating complementary methods for verb sense pruning</i> Hongyan Jing, Vasileios Hatzivassiloglou, Rebecca Passonneau, and Kathleen McKeown	58
<i>Towards a bootstrapping framework for corpus semantic tagging</i> Roberto Basili, Michelangelo Della Rocca, and Maria Teresa Pazienza	66
<i>Sense tagging in action</i> Andrew Harley and Dominic Glennon	74
<i>A perspective on word sense disambiguation methods and their evaluation</i> Philip Resnik and David Yarowsky	79

Working Sessions

<i>Writing annotation instructions</i> Janyce Wiebe	87
<i>Combining knowledge sources for automatic semantic tagging</i> Douglas Jones and Boyan Onyshkevych	88
<i>Semantic tagging and NLP applications</i> Thierry Declerck and Judith Klein	90
<i>Evaluating automatic semantic taggers</i> Philip Resnik and David Yarowsky	91
<i>Constructing semantic tagsets</i> Alain Polguère	92
<i>Structured lexicons and semantic tagging</i> Bonnie J. Dorr and Mari Broman Olsen	93