CoNLL97

Computational Natural Language Learning

Proceedings of the 1997 Meeting of the ACL
Special Interest Group in Natural Language Learning

Editor: T. Mark Ellison

Madrid, July 11, 1997
Held in conjunction with the 1997 meeting of
the Association for Computational Linguistics.

PREFACE

The field of computational natural language learning (NLL) is not a new one; research in it has been pursued for more than forty years. The last seven years, however, have seen a growth in interest and, correspondingly, in meetings addressing this topic. These have been held under the auspices of: COLING (The Unfinished Language, 90), DARPA (90/91), AAAI (MLNLO/CNLP, 91/93), IJCAI (NLL, 91), ECML (Machine Learning and Text Analysis, 93), the European Networks of Excellence ELSNET and MLNET (MLNLS, 94), and ESSLLI (96).

This year, however, is the first time that the ACL's special interest group in natural language learning have organised a meeting in conjunction with an (E)ACL conference. The papers contained in this volume are those which have been accepted to the conference.

In this meeting, we have attempted to cover as broad a range of topics within the field as possible. The range extends from message understanding, through word categorization, ambiguity resolution, learner modelling and text segmentation to the application of neural neworks for learning speech and phonology.

The combination of this vibrant field, with the occasion of joint EACL/ACL meeting make the studies collected in this volume an exciting and stimulating representation of the field.

I would like to take this opportunity to thank my fellows on the program committee, and the other reviewers who helped contributed towards this workshop.

T. Mark Ellison, *University of Edinburgh*

Program Committee

Walter Daelemans
Computational Linguistics, Tilburg University
PO Box 90153, 5000 LE Tilburg, The Netherlands
walter@kub.nl

T. Mark Ellison
Centre for Cognitive Science, University of Edinburgh
2 Buccleuch Place, Edinburgh Eh8 9LW, Scotland.
marke@cogsci.ed.ac.uk

David Powers
Artificial Intelligence Laboratory
Department of Computer Science
The Flinders University of South Australia
powers@acm.org

The program committee would like to thank the following further reviewers:

Steve Abney
Joe Allen
Aude Billard
Alan Black
Michael Brent
Chris Brew
Michael Collins
James Cussens
Gert Durieux
Steve Finch
Steven Gillis
John Goldsmith
Julia Hockenmaier
Jim Hurford
Mehmet Kayaalp
Laurence Malloy
Andrei Mikheev
Jeffrey M. Siskind
Jakup Zavrel

CoNLL97
Proceedings of the 1997 Meeting of the ACL Special Interest Group in Natural Language Learning

Amit Bagga and Joyce Yue Chai
A Trainable Message Understanding System 1-8

Mary Elaine Califf and Raymond J. Mooney
Relational Learning of Pattern-Match Rules for Information Extraction 9-15

Wide R. Hogenhout and Yuji Matsumoto
A Preliminary Study of Word Clustering Based on Syntactic Behavior 16-24

Ji Donghong, He Jun and Huang Changning
Learning New Compositions from Given Ones 25-32

Mehmet Kayaalp, Ted Pedersen and Rebecca Bruce
A Statistical Decision Making Method: A Case Study on Prepositional Phrase Attachment 33-42

Emin Erkan Korkmaz and Göktürk Üçoluk
A Method for Improving Automatic Word Categorization 43-49

Montse Maritxalar, Arantza Díaz de Ilarraza and Maite Oronoz
From Psycholinguistic Modelling of Interlanguage in Second Language Acquisition to a Computational Model 50-59

Laura Mayfield Tomokiyo and Klaus Ries
What makes a word: Learning base units in Japanese for speech recognition 60-69

Ramin Charles Nakisa and Kim Plunkett
Evolution of a Rapidly Learned Representation for Speech 70-79

Miles Osborne and Ted Briscoe
Learning Stochastic Categorial Grammars 80-87

David M. W. Powers
Learning and Application of Differential Grammars 88-96

Jennifer Rodd
Recurrent Neural-Network Learning of Phonological Regularities in Turkish 97-106

Khalil Sima'an
Explanation-Based Learning of Data-Oriented Parsing 107-116

Christoph Tillmann and Hermann Ney
Word Triggers and the EM Algorithm 117-124

Werner Winiwarter and Yahiko Kambayashi
A Comparative Study of the Application of Different Learning Techniques to Natural Language Interfaces 125-135

Jakub Zavrel, Walter Daelemans and Jorn Veenstra
Resolving PP attachment Ambiguities with Memory-Based Learning 136-144

Timetable

09:00-09:15	**Opening**
09:15-09:40	*Montse Maritxalar, Arantza Díaz de Ilarraza and Maite Oronoz* **From Psycholinguistic Modelling of Interlanguage in Second Language Acquisition to a Computational Model**
09:40-10:05	*Werner Winiwarter and Yahiko Kambayashi* **A Comparative Study of the Application of Different Learning Techniques to Natural Language Interfaces**
10:05-10:30	*Christoph Tillmann and Hermann Ney* **Word Triggers and the EM Algorithm**
10:30-10:50	MORNING BREAK
10:50-11:15	*Mehmet Kayaalp, Ted Pedersen and Rebecca Bruce* **A Statistical Decision Making Method: A Case Study on Prepositional Phrase Attachment**
11:15-11:40	*Jakub Zavrel, Walter Daelemans and Jorn Veenstra* **Resolving PP attachment Ambiguities with Memory-Based Learning**
11:40-12:05	*Wide R. Hogenhout and Yuji Matsumoto* **A Preliminary Study of Word Clustering Based on Syntactic Behavior**
12:05-12:30	*Miles Osborne and Ted Briscoe* **Learning Stochastic Categorial Grammars**
12:30-12:55	*Khalil Sima'an* **Explanation-Based Learning of Data-Oriented Parsing**
12:55-14:30	LUNCH
14:30-15:15	**SIGNLL Meeting**
15:15-15:40	*Jennifer Rodd* **Recurrent Neural-Network Learning of Phonological Regularities in Turkish**
15:40-16:05	*Ramin Charles Nakisa and Kim Plunkett* **Evolution of a Rapidly Learned Representation for Speech**
16:05-16:30	*Emin Erkan Korkmaz and Göktürk Üçoluk* **A Method for Improving Automatic Word Categorization**
16:30-16:50	AFTERNOON BREAK
16:50-17:15	*Amit Bagga and Joyce Yue Chai* **A Trainable Message Understanding System**
17:15-17:40	*Mary Elaine Califf and Raymond J. Mooney* **Relational Learning of Pattern-Match Rules for Information Extraction**
17:40-18:05	*Ji Donghong, He Jun and Huang Changning* **Learning New Compositions from Given Ones**
18:05-18:30	*Laura Mayfield Tomokiyo and Klaus Ries* **What makes a word: Learning base units in Japanese for speech recognition**
18:30-18:55	*David M. W. Powers* **Learning and Application of Differential Grammars**
18:55-19:00	**Closing**