

CoNLL 2015

**The Nineteenth Conference on Computational Natural
Language Learning**

Proceedings of the Shared Task

July 30-31, 2015
Beijing, China

Production and Manufacturing by
Taberg Media Group AB
Box 94, 562 02 Taberg
Sweden

©2015 The Association for Computational Linguistics

Order print-on-demand copies from:

Curran Associates
57 Morehouse Lane
Red Hook, New York 12571
USA
Tel: +1-845-758-0400
Fax: +1-845-758-2633
curran@proceedings.com

ISBN 978-1-932432-66-4 / 1-932432-66-3 (Volume 1)
ISBN 978-1-932432-67-1 / 1-932432-67-1 (Volume 2)

Introduction

This volume contains papers describing the CoNLL-2015 Shared Task and the participating systems. This year, we continue the tradition of the Conference on Computational Natural Language Learning (CoNLL) of having a high profile shared task in Natural Language Processing (NLP), focusing on Shallow Discourse Parsing, which involves identifying individual discourse relations that are present in a natural language text. A discourse relation can be expressed explicitly or implicitly, and takes two arguments realized as sentences, clauses, or in some rare cases, phrases. Shallow Discourse Parsing is a fundamental NLP task and can potentially benefit a range of natural language applications such as Information Extraction, Text Summarization, Question Answering, Machine Translation, and Sentiment Analysis.

A total of sixteen teams from three continents participated in this task, and fourteen of them submitted system description papers. Many different approaches were adopted by the participants, and we hope that these approaches help to advance the state of the art in Shallow Discourse Parsing. The training, development, and test sets were adapted from the Penn Discourse TreeBank (PDTB). In addition, we also annotated a blind test set following the PDTB guidelines solely for the shared task. The results on the blind test set were used to rank the participating systems. The evaluation scorer, also developed for this shared task, adopts an F1 based metric that takes into account the accuracy of identifying the senses and arguments of discourse relations as well as explicit discourse connectives. We hope that the data sets and the scorer, which are freely available upon the completion of the shared task, will be a useful resource for researchers interested in discourse parsing.

For the first time in the history of the CoNLL shared tasks, participating teams, instead of running their systems on the test set and submitting the output, were asked to deploy their systems on a remote virtual machine and use a web-based evaluation platform to run their systems on the test set. This meant they were unable to actually see the data set, thus preserving its integrity and ensuring its replicability. We hope that the successful implementation of this new evaluation protocol in the shared task will encourage its adoption in future NLP evaluation campaigns.

Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Rashmi Prasad, Christopher Bryant, and Attapol Rutherford

Organizers of the CoNLL-2015 Shared Task
July 2015

Organizers:

Nianwen Xue, Brandeis University
Hwee Tou Ng, National University of Singapore
Sameer Pradhan, Boulder Language Technology
Rashmi Prasad, University of Wisconsin-Milwaukee
Christopher Bryant, National University of Singapore
Attapol Rutherford, Brandeis University

Program Committee:

Christopher Bryant, National University of Singapore
Jacob Eisenstein, Georgia Institute of Technology
Graeme Hirst, University of Toronto
Fang Kong, Soochow University
Man Lan, East China Normal University
Junyi Jessy Li, University of Pennsylvania
Annie Louis, University of Edinburgh
Hwee Tou Ng, National University of Singapore
Vincent Ng, University of Texas at Dallas
Sameer Pradhan, Boulder Language Technologies
Rashmi Prasad, University of Wisconsin-Milwaukee
Attapol Rutherford, Brandeis University
Mark Sammons, University of Illinois at Urbana-Champaign
Evgeny Stepanov, University of Trento
Yannick Versley, Universität Heidelberg
Bonnie Webber, University of Edinburgh
Nianwen Xue, Brandeis University

Table of Contents

<i>The CoNLL-2015 Shared Task on Shallow Discourse Parsing</i> Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Rashmi Prasad, Christopher Bryant and Attapol Rutherford	1
<i>A Refined End-to-End Discourse Parser</i> Jianxiang Wang and Man Lan	17
<i>The UniTN Discourse Parser in CoNLL 2015 Shared Task: Token-level Sequence Labeling with Argument-specific Models</i> Evgeny Stepanov, Giuseppe Riccardi and Ali Orkan Bayer	25
<i>The SoNLP-DP System in the CoNLL-2015 shared Task</i> Fang Kong, Sheng Li and Guodong Zhou	32
<i>Shallow Discourse Parsing Using Constituent Parsing Tree</i> Changge Chen, Peilu Wang and Hai Zhao	37
<i>A Minimalist Approach to Shallow Discourse Parsing and Implicit Relation Recognition</i> Christian Chiarcos and Niko Schenk	42
<i>A Hybrid Discourse Relation Parser in CoNLL 2015</i> Sobha Lalitha Devi, Sindhuja Gopalan, Lakshmi S, Pattabhi RK Rao, Vijay Sundar Ram and Malarkodi C.S.	50
<i>The CLaC Discourse Parser at CoNLL-2015</i> Majid Laali, Elnaz Davoodi and Leila Kosseim	56
<i>Shallow Discourse Parsing with Syntactic and (a Few) Semantic Features</i> Shubham Mukherjee, Abhishek Tiwari, Mohit Gupta and Anil Kumar Singh	61
<i>JAIST: A two-phase machine learning approach for identifying discourse relations in newswire texts</i> Son Nguyen, Quoc Ho and Minh Nguyen	66
<i>The DCU Discourse Parser: A Sense Classification Task</i> Tsuyoshi Okita, Longyue Wang and Qun Liu	71
<i>Improving a Pipeline Architecture for Shallow Discourse Parsing</i> Yangqiu Song, Haoruo Peng, Parisa Kordjamshidi, Mark Sammons and Dan Roth	78
<i>A Shallow Discourse Parsing System Based On Maximum Entropy Model</i> Jia Sun, Peijia Li, Weiqun Xu and Yonghong Yan	84
<i>The DCU Discourse Parser for Connective, Argument Identification and Explicit Sense Classification</i> Longyue Wang, Chris Hokamp, Tsuyoshi Okita, Xiaojun Zhang and Qun Liu	89
<i>Hybrid Approach to PDTB-styled Discourse Parsing for CoNLL-2015</i> Yasuhisa Yoshida, Katsuhiko Hayashi, Tsutomu Hirao and Masaaki Nagata	95

Conference Program

Thursday, July 30, 2015

Session 3: 14:00–15:30 Shared Task oral presentations

The CoNLL-2015 Shared Task on Shallow Discourse Parsing

Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Rashmi Prasad, Christopher Bryant and Attapol Rutherford

A Refined End-to-End Discourse Parser

Jianxiang Wang and Man Lan

The UniTN Discourse Parser in CoNLL 2015 Shared Task: Token-level Sequence Labeling with Argument-specific Models

Evgeny Stepanov, Giuseppe Riccardi and Ali Orkan Bayer

The SoNLP-DP System in the CoNLL-2015 shared Task

Fang Kong, Sheng Li and Guodong Zhou

Friday, July 31, 2015

Session 8.a: 16:00–17:30 Shared Task poster presentations

Shallow Discourse Parsing Using Constituent Parsing Tree

Changge Chen, Peilu Wang and Hai Zhao

A Minimalist Approach to Shallow Discourse Parsing and Implicit Relation Recognition

Christian Chiarcos and Niko Schenk

A Hybrid Discourse Relation Parser in CoNLL 2015

Sobha Lalitha Devi, Sindhuja Gopalan, Lakshmi S, Pattabhi RK Rao, Vijay Sundar Ram and Malarkodi C.S.

The CLaC Discourse Parser at CoNLL-2015

Majid Laali, Elnaz Davoodi and Leila Kosseim

Shallow Discourse Parsing with Syntactic and (a Few) Semantic Features

Shubham Mukherjee, Abhishek Tiwari, Mohit Gupta and Anil Kumar Singh

JAIST: A two-phase machine learning approach for identifying discourse relations in newswire texts

Son Nguyen, Quoc Ho and Minh Nguyen

Friday, July 31, 2015 (continued)

The DCU Discourse Parser: A Sense Classification Task

Tsuyoshi Okita, Longyue Wang and Qun Liu

Improving a Pipeline Architecture for Shallow Discourse Parsing

Yangqiu Song, Haoruo Peng, Parisa Kordjamshidi, Mark Sammons and Dan Roth

A Shallow Discourse Parsing System Based On Maximum Entropy Model

Jia Sun, Peijia Li, Weiqun Xu and Yonghong Yan

The DCU Discourse Parser for Connective, Argument Identification and Explicit Sense Classification

Longyue Wang, Chris Hokamp, Tsuyoshi Okita, Xiaojun Zhang and Qun Liu

Hybrid Approach to PDTB-styled Discourse Parsing for CoNLL-2015

Yasuhisa Yoshida, Katsuhiko Hayashi, Tsutomu Hirao and Masaaki Nagata