ACL 2016

The 54th Annual Meeting of the Association for Computational Linguistics

Proceedings of the SIGNLL Conference on Computational Natural Language Learning: Shared Task

August 7-12, 2016
Berlin, Germany

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

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-2016 Shared Task and the participating systems. The 2016 shared task is on multilingual Shallow Discourse Parsing (SDP), and is a follow-on to the 2015 shared task. The languages covered in this shared task are English and Chinese. The SDP task 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 24 teams from three continents participated in this task, and 20 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 for English and Chinese were adapted from the Penn Discourse TreeBank (PDTB) and the Chinese Discourse TreeBank (CDTB) respectively. In addition, we also annotated a blind test set for each language following the PDTB and CDTB guidelines solely for the shared task. The results on the blind test sets 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.

As with the 2015 CoNLL shared task on SDP, participants did not each run their systems locally on the test set. Instead, they were asked to deploy their systems on a remote virtual machine and use a web-based evaluation platform called TIRA to run their systems on the test set. This kept them from seeing the data set, thus preserving its integrity and ensuring its replicability.

We would like to thank all the participants of the 2016 Shared Task, as well as the program committee for helping us review the system description papers. Special thanks are due to the SIGNLL board members, Xavier Carreras and Julia Hockenmaier, for their support of the shared task over the last two years. We would also like to thank the PDTB team and CDTB team for providing annotated data for the shared task, and the Linguistic Data Consortium for their help with releasing the data to the participants. Special thanks go to Martin Potthast and the TIRA team, who provided their computing resources and more importantly their time in assisting teams to run their systems.

Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Attapol Rutherford, Bonnie Webber, Chuan Wang, and Hongmin Wang
Organizers of the CoNLL-2016 Shared Task
July 2016

Organizers:

Nianwen Xue, Brandeis University
Hwee Tou Ng, National University of Singapore
Sameer Pradhan, Cemantix.org
Attapol Rutherford, Brandeis University
Bonnie Webber, University of Edinburgh
Chuan Wang, Brandeis University
Hongmin Wang, National University of Singapore

Program Committee:

Christian Chiarcos, Goethe-Universität Frankfurt am Main
Dan Goldwasser, Purdue University
Rohit Jain, IIIT , Hyderabad
Yangfeng Ji, Georgia Institute of Technology
Yusuke Kido, The University of Tokyo
Fang Kong, Soochow University
Majid Laali, Concordia University
Man Lan, East China Normal University
Zhenghua Li, Soochow University
Todor Mihaylov, Heidelberg University
Yusuke Miyao, National Institute of Informatics, Japan
Hwee Tou Ng, National University of Singapore
Stephan Oepen, Universitetet i Oslo
Sameer Pradhan, cemantix.org
Lianhui Qin, Shanghai Jiao Tong University
Attapol Rutherford, Brandeis University
Niko Schenk, Goethe University Frankfurt am Main
Manfred Stede, University of Potsdam
Evgeny Stepanov, University of Trento
Yannick Versley, University of Heidelberg
Chuan Wang, Brandeis University
Bonnie Webber, University of Edinburgh
Gregor Weiss, University of Ljubljana
Nianwen Xue, Brandeis University
Jiajun Zhang, Chinese Academy of Sciences
Hai Zhao, Shanghai Jiao Tong University

Table of Contents

Conference Program

Friday, Aug. 12

14:15–15:45 Oral Presentations

CoNLL 2016 Shared Task on Multilingual Shallow Discourse Parsing
Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Attapol Rutherford, Bonnie Webber, Chuan Wang and Hongmin Wang

OPT: Oslo–Potsdam–Teesside. Pipelining Rules, Rankers, and Classifier Ensembles for Shallow Discourse Parsing
Stephan Oepen, Jonathon Read, Tatjana Scheffler, Uladzimir Sidarenka, Manfred Stede, Erik Velldal and Lilja Øvrelid

An End-to-End Chinese Discourse Parser with Adaptation to Explicit and Non-explicit Relation Recognition
Xiaomian Kang, Haoran Li, Long Zhou, Jiajun Zhang and Chengqing Zong

Two End-to-end Shallow Discourse Parsers for English and Chinese in CoNLL-2016 Shared Task
Jianxiang Wang and Man Lan

Do We Really Need All Those Rich Linguistic Features? A Neural Network-Based Approach to Implicit Sense Labeling
Niko Schenk, Christian Chiarcos, Kathrin Donandt, Samuel Rönnqvist, Evgeny Stepanov and Giuseppe Riccardi

Discourse Sense Classification from Scratch using Focused RNNs
Gregor Weiss and Marko Bajec

17:00–18:00 Poster Presentations

Robust Non-Explicit Neural Discourse Parser in English and Chinese
Attapol Rutherford and Nianwen Xue

A Constituent Syntactic Parse Tree Based Discourse Parser
Zhongyi Li, Hai Zhao, Chenxi Pang, Lili Wang and Huan Wang

SoNLP-DP System for ConLL-2016 English Shallow Discourse Parsing
Fang Kong, Sheng Li, Junhui Li, Muhua Zhu and Guodong Zhou

Shallow Discourse Parsing Using Convolutional Neural Network
Lianhui Qin, Zhisong Zhang and Hai Zhao

Friday, Aug. 12 (continued)

SoNLP-DP System for ConLL-2016 Chinese Shallow Discourse Parsing
Junhui Li, Fang Kong, Sheng Li, Muhua Zhu and Guodong Zhou

UniTN End-to-End Discourse Parser for CoNLL 2016 Shared Task
Evgeny Stepanov and Giuseppe Riccardi

The CLaC Discourse Parser at CoNLL-2016
Majid Laali, Andre Cianflone and Leila Kosseim

Discourse Relation Sense Classification Using Cross-argument Semantic Similarity Based on Word Embeddings
Todor Mihaylov and Anette Frank

IIT (BHU) Submission on the CoNLL-2016 Shared Task: Shallow Discourse Parsing using Semantic Lexicons
Manpreet Kaur, Nishu Kumari, Anil Kumar Singh and Rajeev Sangal

The Virginia Tech System at CoNLL-2016 Shared Task on Shallow Discourse Parsing
Prashant Chandrasekar, Xuan Zhang, Saurabh Chakravarty, Arijit Ray, John Krulick and Alla Rozovskaya

DA-IICT Submission for PDTB-styled Discourse Parser
Devanshu Jain and Prasenjit Majumder

Discourse Relation Sense Classification with Two-Step Classifiers
Yusuke Kido and Akiko Aizawa

Adapting Event Embedding for Implicit Discourse Relation Recognition
Maria Leonor Pacheco, I-Ta Lee, Xiao Zhang, Abdullah Khan Zehady, Pranjal Daga, Di Jin, Ayush Parolia and Dan Goldwasser

SDP-JAIST: A Shallow Discourse Parsing system @ CoNLL 2016 Shared Task
Minh Nguyen

Finding Arguments as Sequence Labeling in Discourse Parsing
Ziwei Fan, Zhenghua Li and Min Zhang

Discourse Relation Sense Classification Systems for CoNLL-2016 Shared Task
Ping Jian, Xiaohan She, Chenwei Zhang, Pengcheng Zhang and Jian Feng