The MIT Press

Journals

Books   Journals   Digital   Resources   About   Contact

Journal Resources

Editorial Info
Abstracting and Indexing
Release Schedule
Advertising Info

Author Resources

Submission Guidelines
Publication Agreement

CODRA: A Novel Discriminative Framework for Rhetorical Analysis

Shafiq Joty, Giuseppe Carenini and Raymond T. Ng

👁 **Download Options** ＞

Abstract   Full Text   Authors

Clauses and sentences rarely stand on their own in an actual discourse; rather, the relationship between them carries important information that allows the discourse to express a meaning as a

Author Reprints

Reader Resources

Rights and Permissions
Most Read
Most Cited

More About Computational Linguistics ⌄

Metrics ⌄

Open Access ⌄

Computational Linguistics Computational Linguistics is Open Access. All content is freely available in electronic format (Full text HTML, PDF, and PDF Plus) to readers across the

whole beyond the sum of its individual parts. Rhetorical analysis seeks to uncover this coherence structure. In this article, we present CODRA — a COmplete probabilistic Discriminative framework for performing Rhetorical Analysis in accordance with Rhetorical Structure Theory, which posits a tree representation of a discourse.

CODRA comprises a discourse segmenter and a discourse parser. First, the discourse segmenter, which is based on a binary classifier, identifies the elementary discourse units in a given text. Then the discourse parser builds a discourse tree by applying an optimal parsing algorithm to probabilities inferred from two Conditional Random Fields: one for intra-sentential parsing and the other for multi-sentential parsing. We present two approaches to combine these two stages of parsing effectively. By conducting a series of empirical evaluations over two different data sets, we demonstrate that CODRA significantly outperforms the state-of-the-art, often by a wide margin. We also show that a reranking of the k-best parse hypotheses generated by CODRA can potentially improve the accuracy even further.

Forthcoming

Most Read   See More

Lexicon-Based Methods for Sentiment Analysis (13965 times)
Maite Taboada et al.
Computational Linguistics
Volume: 37, Issue: 2, pp. 267-307

Computational Linguistics and Deep Learning (10500 times)
Christopher D. Manning
Computational Linguistics
Volume: 41, Issue: 4, pp. 701-707

Near-Synonymy and Lexical Choice (3653 times)
Philip Edmonds et al.
Computational Linguistics
Volume: 28, Issue: 2, pp. 105-144

(Note that the Most Read numbers are based on the number of full text downloads over the last 12 months.)

Most Cited   See More

Support OA at MITP

Lexicon-Based Methods for Sentiment Analysis (436 times)
Maite Taboada et al.
Computational Linguistics
Volume: 37, Issue: 2, pp. 267-307

A Systematic Comparison of Various Statistical Alignment Models (174 times)
Franz Josef Och et al.
Computational Linguistics
Volume: 29, Issue: 1, pp. 19-51

Opinion Word Expansion and Target Extraction through Double Propagation (147 times)
Guang Qiu et al.
Computational Linguistics
Volume: 37, Issue: 1, pp. 9-27

(Note that the Most Cited numbers are based on Crossref's Cited-by service and reflect citation information for the past 24 months.)
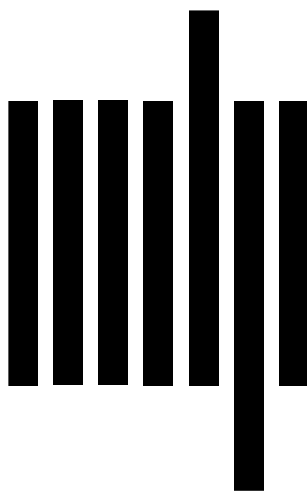
⊚ Download Options ›

Favorite ♡

Sign up for Alerts 🔔

Download Citation ↓

RSS TOC

RSS Citation

Submit your article

Support OA at MITP 🔒

Journals

Books

Terms & Conditions

Privacy Statement

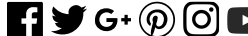Contact Us

US

One Rogers Street Cambridge MA 02142-1209

UK

Suite 2, 1 Duchess Street London, W1W 6AN, UK

Connect

f 🐦 G+ ⓟ 📷 ▶

© 2018 The MIT Press
Technology Partner: Atypon Systems, Inc.
CrossRef Member
COUNTER Member
The MIT Press colophon is

registered in the U.S. Patent and

U.S. Patent and
Trademark Office.
Site Help