

Methods and applications in discourse parsing

Keywords

Coherence, machine learning, parsing, question answering, sentiment analysis

Proposer(s) / Proposatzailea(k)

Kepa Bengoetxea, Mikel Iruskieta

Contact / Kontaktua

kepa.bengoetxea@ehu.es, mikel.iruskieta@ehu.eus

Description / Deskribapena

Discourse Parsing is nowadays a very useful task. Several discourse parser have been developed for many languages. Some of them are multilingual (Braud et al. 2017).

Firstly, the student will review the state of the art.

Secondly, an automatic discourse parser or partial parser could be developed and evaluated intrinsically.

And, finally, the parser could be tested extrinsically in a sound NLP task such as sentiment analysis, summarization, question answering.

Goals / Helburuak

To develop an automatic discourse parser under Rhetorical Structure Theory (RST) formalism and explore possible applications.

Requirements / Betebeharrak

Profile

Computer scientists

Tasks and plan / Atazak eta plana

1. Study the state of the art of discourse parsing
2. Design of an automatic discourse parser, based in Rhetorical Structure Theory.
3. Test the parser in several corpus and evaluate it.
4. Explore possible applications of the parser in different NLP tasks: question answering, sentiment analysis, summarization, machine translation...

References:

Alami, Nabil, Mohammed Meknassi & Nouredine Rais. 2015. Automatic texts summarization: Current state of the art. Journal of Asian Scientific Research 5(1). 1-15.

Atutxa, U. Ansa, O. Iruskietta, M. & Molina, A. 2017. Compress-EUS: i(ra)kasleen laburpenak lortzeko tresna. EUDIA-6 WORKSHOP. Linguistic variation in the Basque language and Education. Euskararen bariazioa eta bariazioaren irakaskuntza. Leioa

Bokaei, Mohammad Hadi, Hossein Sameti & Yang Liu. 2015. Extractive summarization of multi-party meetings through discourse segmentation. *Natural Language Engineering*. in press.

Bosma, Wauter E. 2008. Discourse oriented summarization. Enschede: University of Twente Thesis.

Chengcheng, Li. 2010. Automatic text summarization based on Rhetorical Structure Theory. Proceedings of International Conference on Computer Application and System Modeling (ICCASM 2010). (pp. V13-595-598). Taiyuan, China.

Cohan, Arman & Nazli Goharian. 2015. Scientific article summarization using citation-context and article's discourse structure. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. (pp. 390-400). Lisbon, Portugal.

Molina, A., Torres-Moreno, J. M., SanJuan, E., Da Cunha, I., & Martínez, G. E. S. 2013. Discursive sentence compression. In International Conference on Intelligent Text Processing and Computational Linguistics (pp. 394-407). Springer Berlin Heidelberg.

Molina, A. 2013. Compresión automática de frases: un estudio hacia la generación de resúmenes en espanol. *Inteligencia Artificial*, 16(51), 41-62.

Uzêda, Vinícius Rodrigues de, Thiago Alexandre Salgueiro Pardo & Maria das Graças Volpe Nunes. 2009. A comprehensive summary informativeness evaluation for RST-based summarization methods. *International Journal of Computer Information Systems and Industrial Management Applications - IJCISIM* 1. 188-196.

Uzêda, Vinícius Rodrigues de, Thiago Alexandre Salgueiro Pardo & Maria das Graças Volpe Nunes. 2010. A comprehensive comparative evaluation of RST-based summarization methods. *ACM Transactions on Speech and Language Processing* 6(1-20).

Zahri, N. Adilah Hanin, Fumiyo Fukumoto, Matsyoshi Suguru & Ong Bi Lynn. 2015. Exploiting rhetorical relations to multiple documents text summarization. *International Journal of Network Security and its Applications* 7(2)

Zipitria, I. Arruarte, A. Elorriaga, J. 2013. Discourse measures for Basque summary grading. *Interactive Learning Environments*, 21(6), 528-547.