Discourse Based Automatic Summarization / Diskurtso-egituran oinarrituriko laburpen automatikoa

Proposer(s) / Proposatzailea(k):
Mikel Iruskieta

Contact / Kontaktua:

mikel.iruskieta@ehu.eus

Description / Deskribapena

Students make understanding-exercises after reading texts and they use different summarization techniques intending to show they understand the text. But these techniques are most of the times superficial and they do not develop writing techniques, because the teacher cannot correct all the summaries produced by students.

Once we have collected a corpus of different sources: students: primary education and undergraduate students, and teachers.

An automatic summarizer could be developed and evaluated based in this corpus and the automatic summaries can be offered to students and teachers to observe what corrections they propose with the aim to improve the automatic summaries.

Keywords: summarization, coherence, abstraction, extraction, evaluation, Primary Education

Goals / Helburuak

To develop an automatic summarizer based in discourse structure (under RST or related formalism).

Requirements / Betebeharrak

Linguist or computer scientists

Framework / Esparrua

In collaboration with the Association of "Ikastolas" https://www.ikastola.eus/

Tasks and plan / Atazak eta plana

1. Study different automatic summary strategies.

2. Design of an automatic system for summarization, based in discourse structure

3. Study different evaluation systems to compare automatic summaries outputs (extractive and abstractive) against different gold standards produced by students or teachers in a real scenario (Ikastolas), obtained with COMPRESS-EUS:

http://ixa2.si.ehu.es/clarink/tresnak/compress-eus/.

4. Give the automatic output to students and teachers and describe how they correct the automatic summaries, to improve the system.

References:

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

Atutxa, U. Ansa, O. Iruskieta, 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 TwenteThesis.

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.