

Analysis and Generation of natural language timelines

Proposers / proposatzaileak: Begoña Altuna, Maxux Aranzabe, German Rigau

Contact / kontaktua: begona.altuna@ehu.eus

Description / deskribapena:

Automatically generated timelines can offer a quick insight on the chronology of a list of events, but for that they have to be easily understandable. Temporal information extraction systems extract the information that is displayed in timelines, but that information is not available in a human-friendly format. In this project, the candidate will have to develop a system that will generate natural language sentences for timelines based on the temporal information and argument structures.

Goals / Helburuak:

Development of a system that will analyse texts and/or generate natural language sentences from argument structures, paying special attention to temporal information.

Requirements / Betebeharrak:

- Basic linguistic background
- Basic programming skills

Framework / Esparrua:

Temporal information processing and natural language generation.

Tasks and plan / Atazak eta plana:

To carry out the project, these tasks have been foreseen:

- Analysing the available datasets (e.g. https://en.wikipedia.org/wiki/List_of_timelines)
- Analysing the linguistic features of the input information and the output sentences
- Choosing the best systems (rule-based, machine learning, deep learning, ...)
- Developing and evaluating a system

References / Bibliografia:

- Tim Althoff, Xin Luna Dong, Kevin Murphy, Safa Alai, Van Dang, and Wei Zhang. 2015. TimeMachine: Timeline Generation for Knowledge-Base Entities. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15)*. Association for Computing Machinery, New York, NY, USA, 19–28. DOI:<https://doi.org/10.1145/2783258.2783325>
- Begoña Altuna, María Jesús Aranzabe, and Arantza Díaz de Ilarraza. 2019. EusTimeBank-TL corpusa: denbora-informaziodun testuetatik denbora-lerroetara. In Olatz Arbelaitz, Urtzi Etxeberria, Ainhoa Latatu and Miren Josu Omaetxebarria, editors, *III. IkerGazte Nazioarteko Ikerketa Euskaraz, Giza zientziak eta artea*. (pp. 83–90), Baiona, Euskal Herria.
- Harsimran Bedi, Sangameshwar Patil, Swapnil Hingmire and Girish K. Palshikar. 2017. Event Timeline Generation from History Textbooks. Proceedings of the 4th Workshop on Natural Language Processing Techniques for Educational Applications, pages 69–77, Taipei, Taiwan. <https://www.aclweb.org/anthology/W17-5912.pdf>
- Laparra E., Aldabe I., Agerri R. and Rigau G. Multi-lingual and Cross-lingual TimeLine Extraction. Knowledge-Based Systems, Volume 133, pages 77-89, ISSN 0950-7051, 2017. <https://doi.org/10.1016/j.knosys.2017.07.002>.