TimeNorm for ixa-pipes

Proposer(s) / Proposatzailea(k): names / izenak

Rodrigo Agerri and German Rigau

Contact / Kontaktua: german.rigau@ehu.eus

Description / Deskribapena

IXA pipeline (Agerri et al. 2014) is a modular set of Natural Language Processing tools (or pipes) which provide easy access to NLP technology. We are now developing ixa-pipe-time, a new module for detecting and normalizing temporal expressions. ixa-pipe-time uses TimeNorm library (Bethard 2013) for converting natural language expressions of dates and times to their normalized form. However, only English and Italian TimeNorm grammars exist (Mirza and Minard 2014).

Goals / Helburuak

To contribute on the new ixa-pipe-time module by developing new TimeNorm grammars for Spanish, Basque, etc. (https://github.com/ixa-ehu/ixa-pipe-time)

Requirements / Betebeharrak

Basic knowledge of Linux command-line interface:

- execution of programs through the command line
- handling of text files

Framework / Esparrua

IXA pipeline (Agerri et al. 2014) is a modular set of Natural Language Processing tools (or pipes) which provide easy access to NLP technology. TimeNorm library (Bethard 2013) for converting natural language expressions of dates and times to their normalized form. English and Italian TimeNorm grammars exist (Mirza and Minard 2014). TempEval3 datasets (UzZaman et al. 2013).

Tasks and plan / Atazak eta plana

- Develop new Spanish, Basque, etc. grammars for TimeNorm
- Test ixa-pipe-time with the English and new grammars on TempEval-like data

References

Agerri R., Bermudez J. and Rigau G. *IXA pipeline: Efficient and Ready to Use Multilingual NLP tools*. Proceedings of the 9th Language Resources and Evaluation Conference (LREC'14). Reykjavik, Iceland. 2014.

Steven Bethard. 2013. A Synchronous Context Free Grammar for Time Normalization. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pp. 821-826.

Mirza, Paramita, and Anne-Lyse Minard. FBK-HLT-time: a complete Italian Temporal Processing system for EVENTI-EVALITA 2014. In *Fourth International Workshop EVALITA 2014*, pp. 44-49. 2014.

UzZaman, Naushad, Hector Llorens, Leon Derczynski, James Allen, Marc Verhagen, and James Pustejovsky. Semeval-2013 task 1: Tempeval-3: Evaluating time expressions, events, and temporal relations. In *Second Joint Conference on Lexical and Computational Semantics (* SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, vol. 2, pp. 1-9. 2013.