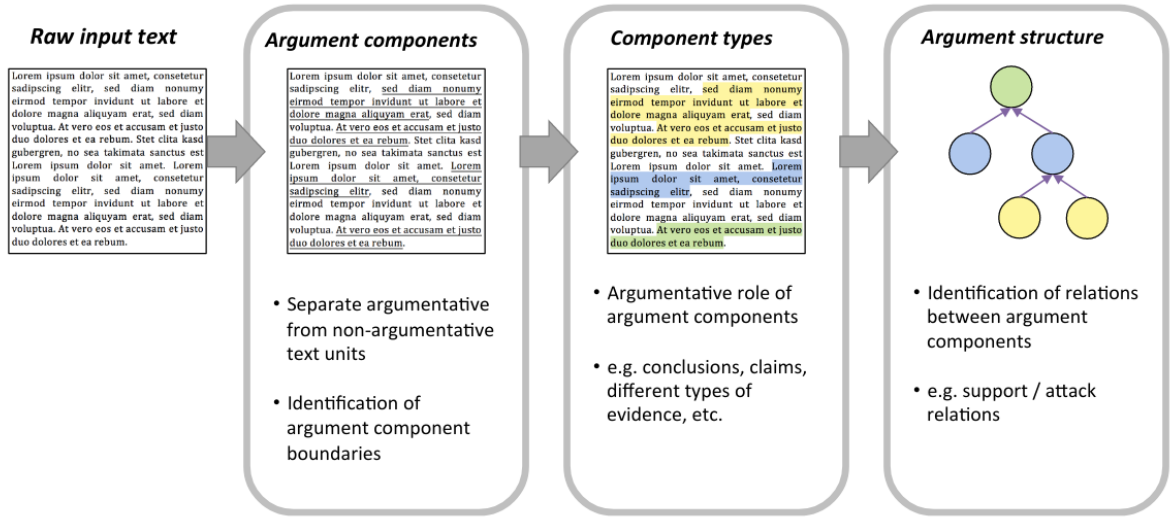
Argument Mining and Generation

Proposers: Rodrigo Agerri

Contact: rodrigo.agerri@ehu.eus

Description

Argument mining aims at extracting natural language arguments and their relations from text (Cabrio and Villata, 2018) Two stages are crucial: (i) argument extraction (e.g., claim, premises) and the identification of their textual boundaries; (ii) argument relation extraction: the prediction of the relations (e.g., *attack* and *support*) holding between the arguments identified in (i). This is depicted in the following Figure.



Most of the work on argument mining has been done in English for the education (Stab and Gurevych 2017) or medical (Mayer et al. 2021) domain, and has been focused on very simple ***support*** and ***attack*** relations. Additionally, the availability of corpora and approaches on Natural Language Generation of arguments is extremely scarce, both in terms of languages and domains (Reisert et al. 2019, Chung et al. 2019).

In this proposal the aim would be to experiment on argumentation (both mining and generation) from a multilingual perspective, trying also to semi-automatically generate resources for languages other than English.

Objectives

The candidate may choose between the following objectives:

1. (Semi-) Automatic development of datasets for argument mining and/or generation in languages other than English.
2. Few-shot learning for cross-lingual argument mining on specific domains such as misinformation.
3. Natural Language Generation to automatically generate domain-expert guided counter-arguments in real time with the aim of counteracting the spread of misinformation in social media, that go beyond the *support/attack* paradigm.

Tasks and Plan

- Month 1: Start of the project, defining the objectives and tasks.
- Month 2: Start experiments. Optionally, it is recommended for the candidates to attend the "Seminar on language technologies. Deep Learning (LAP 18). https://ixa.si.ehu.es/master/programa_html
- Months 3-5: Experiments and final development.
- Final month: Writing up.

References

Chung, Yi-Ling, Elizaveta Kuzmenko, Serra Sinem Tekiroglu and Marco Guerini. "CONAN - COunter NArratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech." *In ACL 2019.*

Mayer, Tobias, Santiago Marro, Elena Cabrio and Serena Villata. "Enhancing evidence-based medicine with natural language argumentative analysis of clinical trials." Artificial intelligence in medicine 118 (2021): 102098 .

Reisert, Paul, Benjamin Heinzerling, Naoya Inoue, Shun Kiyono and Kentaro Inui (2019). "Riposte! A Large Corpus of Counter-Arguments." In *ArXiv.*

Schick, Timo and Hinrich Schütze (2021) "True Few-Shot Learning with Prompts -- A Real-World Perspective.". In arxiv.

Stab, Christian and Iryna Gurevych (2017). "Parsing Argumentation Structures in Persuasive Essays." *Computational Linguistics* 43 (2017): 619-659.