CIIAM - Post Doctoral Position on Contextual Information Inference for Argument Mining

Context

Argument mining is a branch of NLP aimed at extracting and analyzing argumentative structures in texts. It includes identifying argumentative parts, classifying components such as claims and premises, as well as classifying argumentative relations, such as support or attack, between these components. The goal is to provide a structured representation of arguments in a text, thereby facilitating the analysis of relationships and argumentative patterns for various applications. In recent years, each of these tasks has garnered increasing interest from the NLP community, and numerous advances have resulted in the publication of several comparative studies of existing methods [BF20, TD20a, TD20b]. Observations from these studies indicate a predominance of approaches based on neural networks, and more recently, on pre-trained language models. Pre-trained language models, such as those based on the BERT Transformer architecture, have significantly outperformed traditional machine learning models, particularly in the field of argument mining. This includes tasks such as argument classification and prediction of relations between components [MT20]. In parallel, these efforts have multiplied in response to application and contextual constraints, addressing argument mining in domains such as health, politics, and social media [DM17]. Despite these advances, limitations reported in the literature remain numerous. Practically speaking, there are few resources, and the development of datasets for training and evaluating argument mining tasks proves to be very costly (extensive manual annotation by qualified annotators, revision iterations to ensure reliability, and the complexity of creating a representative corpus) [GM22]. Currently, there is no large-scale corpus or benchmark available to objectively evaluate the various tasks related to argument mining. From a computational perspective, automatically understanding all argumentative structures and the linguistic devices that underlie them remains complex. Furthermore, few works explore the formalization and modeling of information from pragmatic analysis, thus limiting the inference of high-level knowledge, such as presupposed knowledge shared between the speaker and the listener, or common sense knowledge [SS19].

Meanwhile, the use of prompts in the NLP field has also seen significant expansion, particularly with the advent of pre-trained language models such as GPT-3 (Generative Pre-trained Transformer 3) and LLAMA [LP23]. Prompting has since become a full-fledged research domain that continues to evolve. Used as stimuli to guide the operation of language models, the use of prompts supports the guidance of these models for various applications, such as text generation, classification, automatic summarization, translation, information retrieval, and many others. In the context of argument mining, and more broadly to enhance

language understanding, prompting techniques can be used to synthesize knowledge. This involves expanding the range of stimuli to instruct language models on how to reason, integrating the necessary knowledge for overall meaning construction and the processes involved in its creation.

Objectives

By adopting text generation-based approaches, particularly those related to the use of advanced prompting techniques, our goal is to develop contextual "frameworks". Through this innovative approach, our goal is to develop argument mining techniques that not only capture the fundamental mechanisms underlying debates and discussions but are also robust against the nuances and subtleties of human communication. By synthesizing contextual information, we aim to enhance the ability to analyze argumentative discourse accurately and nuancedly, integrating linguistic and multimodal dimensions of communication.

Application

Applications are to be sent before june 2014 to the position supervisor Anaïs Ollagnier (ollagnier@i3s.unice.fr).

The application package comprises:
• an up-to-date resume
• a copy of the Ph.D diploma
• recommendation letters or a list of references

Required qualifications. Candidates fit the following profiles:
• Ph.D in computer science
• Experience with ML/DL libraries
• Knowledge of Argument Mining & Prompt engineering
• Proficiency in both written and spoken English

References

[BE22] Bender, E. M., & Lascarides, A. (2022). **Linguistic fundamentals for natural language processing II: 100 essentials from semantics and pragmatics**. Springer Nature.

[BF20] Boltužić, F., & Šnajder, J. (2020). **Structured prediction models for argumentative claim parsing from text**. Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije, 61(3), 361-370.

[CP22] Céline Poudat, Sarra El Ayari. **Annoter, visualiser et spécifier les textes et les genres avec les marqueurs de RDA. Représentation du discours autre et genres de discours : de l'annotation aux enjeux communicatifs et éducatifs**, Mar 2022, Créteil, France.⟨hal-03594460⟩

[DM17] Dusmanu, M., Cabrio, E., and Villata, S. (2017). **Argument mining on twitter: Arguments,**
facts and sources. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2317–2322.

[DU80] Ducrot, O. (1980). **Analyses pragmatiques**. Communications, 32(1), 11-60.

[GM22] Gaku Morio, Hiroaki Ozaki, Terufumi Morishita, Kohsuke Yanai; **End-to-end Argument Mining with Cross-corpora Multi-task Learning**. *Transactions of the Association for Computational Linguistics* 2022; 10 639–658. doi:https://doi.org/10.1162/tacl_a_00481

[GP22] Goffredo, P., Haddadan, S., Vorakitphan, V., Cabrio, E., & Villata, S. (2022, July). **Fallacious argument classification in political debates**. In Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI (pp. 4143-4149).

[LP23] Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). **Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing**. ACM Computing Surveys, 55(9), 1-35.

[LT23] Ludovic Tanguy, Céline Poudat, Mai Ho-Dac. **Etude des comportements et interactions dans les discussions en ligne de la Wikipédia. L'argumentation collaborative : Discussions en ligne et apprentissage de l'esprit critique**, CLLE; LERASS, Oct 2023, Toulouse, France.⟨hal-04338587⟩

[NO23] Nicolas Ocampo, Ekaterina Sviridova, Elena Cabrio, and Serena Villata. 2023. **An in-depth analysis of implicit and subtle hate speech messages**. In Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2023, Dubrovnik, Croatia, May 2-6, 2023, pages 1989–2005.

[MO19] Jacques Moeschler. **Non-lexical Pragmatics: Time, causality and logical words**. T. 23. Wal-
ter de Gruyter GmbH & Co KG, 2019.

[MT20] Mayer, T., Cabrio, E., & Villata, S. (2020). **Transformer-based argument mining for healthcare applications**. In ECAI 2020 (pp. 2108-2115). IOS Press.

[SS19] Storks, S., Gao, Q., & Chai, J. Y. (2019). **Recent advances in natural language inference: A survey of benchmarks, resources, and approaches**. arXiv preprint arXiv:1904.01172.

[TD20a] Trautmann, D. (2020). **Aspect-based argument mining**. arXiv preprint arXiv:2011.00633.