New Interactions for Users of Machine Translation

M2 Internship

November 18, 2024

Keywords: Multilingual Large Language Models, Post-Edition, Confidence Estimation

Research Context and Questions

Machine Translation (MT) technologies are routinely used in a multitude of contexts for a variety of tasks. In most cases, MT users are amateur translators, who have not been specifically trained to use MT tools, are mostly unaware of their strengths and weaknesses, and who, contrarily to professional translators, may only have a partial command (in the extreme case no command at all) of the source or the target language. In most cases, the translation is not the final product, rather a convenient way to achieve some communicative task – the translation need not be perfect, but should be fit for its intended purpose. Yet, existing online translation interfaces provide little help to achieve such goals: in most cases, it consists of two windows, one where the source text is entered, and the second where the target text is generated.

In this research we aim to explore ways to provide novel visualisations and interaction techniques for lay users of MT, focusing on ways to assist writers to identify the parts that are reliably translated, and to act on the others so as to improve the translation. We plan, for instance, to investigate the following set-ups:

> **Pre-edition** Alice writes a message M in his mother tongue L_A, and automatically translates it before sending $T_{A\to B}(M)$ to Bob (in a language L_B that she does not necessarily read). In this situation, translating $T_{A\to B}(M)$ backwards into L_A may help Alice to spot in $T_{B\to A}(T_{A\to B}(M))$ fragments whose translation are suspicious. By iteratively pre-editing (Miyata and Fujita, 2021) the message M so that $T_{B\to A}(T_{A\to B}(M))$ gets closer to M, Alice may eventually get more confident that what Bob gets actually correspond to the meaning of M (Zouhar and Bojar, 2020). Can we guide Alice to perform this task more effectively, e.g. by highlighting the differences, by also using automatic confidence measures (Specia et al., 2017; Rei et al., 2022) or explanation generation techniques (Stahlberg et al., 2018; Vamvas and Sennrich, 2022; Briakou et al., 2023; Sennrich et al., 2024)? Assuming that Alice wants some parts of the

translation (e.g. terms, or places, or dates, or amounts) to be 100% correct, can we leverage existing resources to increase her confidence in the translation?

[**Post-edition**] Bob conversely needs to produce a publishable message in L_2 (a language he masters), based on an initial version in L_1 (a language that he may not know), that he received from Alice. After computing an automatic translation, Bob can revise the text to improve the style or content. Again, can we help Bob so that he can focus on fixing the most problematic parts of the translation, e.g. with confidence estimation Guerreiro et al. (2024)? Can we provide Bob additional interactions, such as the ability to correct multiple repeated errors in just one edit? Can we help Bob perform time vs. quality tradeoffs, so that the translation effort is used optimally?

This internship will explore topics at the intersection of natural language processing (NLP) and human-computer interaction (HCI). In the first stage, the intern will assimilate the state of the art in this nascent field. They will then focus on one setup (pre- or post-edition), develop an experimental platform, and run small-scale experiments with actual MT users.

Internship conditions

The internship will be jointly supervised by Marine Carpuat[1] and François Yvon[2], with the help of Gilles Bailly (ISIR). The internship will take place in the MLIA team of ISIR[3]. ISIR is under the dual supervision of Sorbonne Université, a world-class multidisciplinary university, and the French National Centre for Scientific Research (CNRS), one of the most important research institutions in the world. ISIR includes 6 research teams and 226 people. The intern will be located at *4, place Jussieu, 75005 Paris*.

- Remuneration: around 600€ along with the refund of half the Navigo (public transport) card.
- Starting date: the internship is expected to start in February or March 2025.
- Duration: 5-6 months.

Requirements

We are looking for a second-year Master's student with a strong background in Natural Language Processing, Machine Learning or Human-Computing Interaction. The

[1] https://cs.umd.edu/~marine
[2] https://fyvo.github.io/
[3] https://www.isir.upmc.fr/teams/mlia/presentation-mlia/?lang=en

intern is expected to be proficient in programming, especially in the Python language, and to have already worked under Linux. They should also have experience with a deep learning framework, preferably PyTorch.

Application

Please send a resume along with a cover letter (in French or English) and grade transcripts for the last two years to François Yvon at `yvon@isir.upmc.fr`. A list of pointers to example projects (e.g., via GitHub) is a plus.

References

Eleftheria Briakou, Navita Goyal, and Marine Carpuat. 2023. Explaining with contrastive phrasal highlighting: A case study in assisting humans to detect translation differences. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 11220–11237, Singapore. Association for Computational Linguistics.

Nuno M. Guerreiro, Ricardo Rei, Daan van Stigt, Luisa Coheur, Pierre Colombo, and André F. T. Martins. 2024. xcomet: Transparent machine translation evaluation through fine-grained error detection. *Transactions of the Association for Computational Linguistics*, 12:979–995.

Rei Miyata and Atsushi Fujita. 2021. Understanding pre-editing for black-box neural machine translation. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 1539–1550, Online. Association for Computational Linguistics.

Ricardo Rei, Marcos Treviso, Nuno M. Guerreiro, Chrysoula Zerva, Ana C Farinha, Christine Maroti, José G. C. de Souza, Taisiya Glushkova, Duarte Alves, Luisa Coheur, Alon Lavie, and André F. T. Martins. 2022. CometKiwi: IST-unbabel 2022 submission for the quality estimation shared task. In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 634–645, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.

Rico Sennrich, Jannis Vamvas, and Alireza Mohammadshahi. 2024. Mitigating hallucinations and off-target machine translation with source-contrastive and language-contrastive decoding. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 21–33, St. Julian's, Malta. Association for Computational Linguistics.

Lucia Specia, Carolina Scarton, and Gustavo Henrique Paetzold. 2017. *Quality estimation for Machine Translation*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.

Felix Stahlberg, Danielle Saunders, and Bill Byrne. 2018. An operation sequence model for explainable neural machine translation. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 175–186, Brussels, Belgium. Association for Computational Linguistics.

Jannis Vamvas and Rico Sennrich. 2022. As little as possible, as much as necessary: Detecting over- and undertranslations with contrastive conditioning. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 490–500, Dublin, Ireland. Association for Computational Linguistics.

Vilém Zouhar and Ondřej Bojar. 2020. Outbound translation user interface ptakopět: A pilot study. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 6967–6975, Marseille, France. European Language Resources Association.