Multilingual natural language processing: Towards universal translation with neural approaches

Author: Marta R. Costa-jussà, Universitat Politècnica de Catalunya, Barcelona

Standfirst: State of the art neural network approaches enable massive multilingual translation. How close are we to universal translation in any language?

Main text: The quality of machine translation is approaching human parity for several language pairs [10]. The field is progressing rapidly further with recent advances in natural language processing and highly multilingual systems can be learned with a shared system for all languages [**2**]. As *Tang* and colleagues demonstrate, we are getting close to universal translation *[9]*. They combine the advantages of previous research (English or non-English centric) in shared [2] and massive multilingual models [3] with unsupervised techniques [4] and show that simultaneous, multi-lingual fine tuning can improve the training of translations significantly, especially for languages that do not offer large bilingual text collections. Works such as [2,3,4] bring us closer to the goal of universal communication, which could enable world-wide conversations.
Even though it might be difficult to agree on what this generic and ambitious goal entails. We may agree that it involves the capability of transparently communicating in all (or a great proportion of) languages (no matter if they are spoken, signed or written). In any case, a system that allows for "universal" translation either can translate from/to any language or it can be extended to any language. There remain open challenges which include increasing the number of covered languages, dealing with non-written languages and producing robust and fair outputs.

Historically, multilingual machine translation was addressed from a pairwise perspective, where a system is built for each translation pair, or from a pivot perspective, where an intermediate language is used as a "hub", e.g. translating source languages into English and then into the target languages. These alternatives offer the advantage that new languages can be deployed without retraining any of the languages in the system. The prominent neural machine translation approach [1] soon moved to a new shared proposal [2]. Neural machine translation relies on the power of current deep learning algorithms and it uses an architecture of encoder-decoder, meaning that it codifies a text in the source language into an intermediate representation, which is decodified into the target language. In addition to the pairwise and pivot alternative, multilingual machine translation involves a shared system that acts as a "unique" translator, which means that all languages are translated with the same system. This is a desired situation because in production you only have to maintain one since system. Figure 1 reflects the pairwise vs the shared system architectures.

Costa-jussà, M.R. Multilingual natural language processing: Towards universal translation. "Nature machine intelligence", Maig 2021, vol. 3, núm. 5, p. 376-377.
https://doi.org/10.1038/s42256-021-00346-7

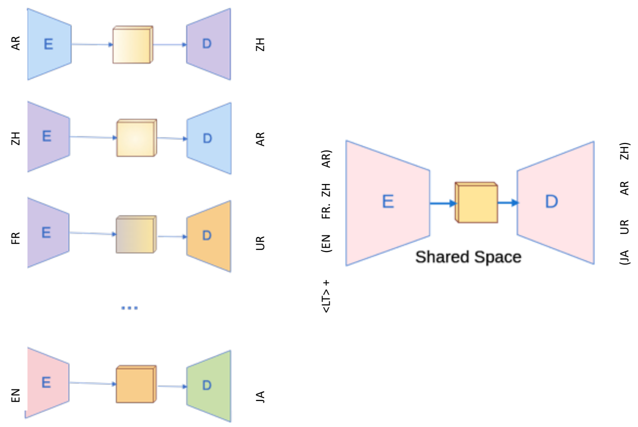

Figure 1: Pairwise system (left) vs Shared system (right)

These shared architectures have evolved to massive multilingual ones [3], currently covering up to 100 language pairs (50 in the particular case of *Tang et al.*). One of the big limitations of previous works, which is solved by *Tang et al*, was that they did not allow for incrementally introducing new languages. Finetuning on bilingual text collections allows to solve this issue.

Without a doubt, Tang et al. is the culmination of previous research, e.g. the advent of neural architectures in machine translation [1,11], sharing modules [2], massively training [3] and combining pretraining strategies [9]. However, there are still some challenges to solve, which we discuss as follows:

- **Coverage of languages.** According to Ethnologue, there are 7117 languages spoken today, which means that *Tang et al.* translation system covers less than 1% of the total languages. However, 40% of languages have less than 1,000 speakers remaining and 200 languages account for more than 88% the world's population. With this huge amount of languages, the main question is if the "universal" translation will arrive soon enough to safe the endangered languages or if current systems, while only focusing on highly spoken languages, will contribute to the disappearance of minor ones.
- **Non-written languages.** Machine translation relies heavily on collections and processing of digitized text, but there are languages that are only spoken, meaning that they do not have a written version. In addition, sign languages are encoded in images more than in text. For these cases, and for standard languages with both spoken and textual versions, multimodality of the machine translation system is required. Different research approaches use images to support the translation of texts [5]. The main limitations of these approaches are that images alone allow for low improvement of text translation. In parallel, there are already prototypes that allow for end-to-end speech-to-speech translation systems [12].
- **Robustness, Fairness and Transparency.** Current systems highly depend on the quality of the text to be translated and it is not robust to typos or vocabulary alternatives [6]**.** Stereotype association is an issue in current models. Unbalanced data with a high presence of males over females, among other social unbalances, are guiding our systems which have been shown to amplify stereotypes [8]. There have been several proposals towards facing the fairness problem in several directions, but most of them have been deeply criticized by the social sciences community for being superficial and neither studying nor solving any social harm

[7]. Overall, there is an urgent need for deep transparency in our systems, meaning that we need to explain their performance.

Machine translation is evolving at an incredible speed. Researchers in this area are formulating and answering complex language modeling questions, which may soon reach a system that qualifies as a "universal" translator. The social impact of a tool offering instant translation in any pair of languages is huge. This tool will be able to reduce international conflicts and improve global communication and international relationships, in general. Approaches such as the one presented by Tang et al. show that progress towards this ambitious goal is possible.

References:

[1] Cho, K. van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y. (2014) *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation*. In Proc. EMNLP.

[2] M. Johnson, M. Schuster, Q. V Le, M. Krikun, Y. Wu, Z. Chen, N. Thorat, F. Viégas, M. Wattenberg, G. Corrado, 2017. Google's multilingual neural machine translation system: Enabling zero-shot translation. *Transactions of the Association for Computational Linguistics*, 5:339–351 (2017)

[3] Angela Fan, Shruti Bhosale, Holger Schwenk, Zhiyi Ma, Ahmed El-Kishky, Siddharth Goyal, Mandeep Baines, Onur Celebi, Guillaume Wenzek, Vishrav Chaudhary, Naman Goyal, Tom Birch, Vitaliy Liptchinsky, Sergey Edunov, Edouard Grave, Michael Auli, Armand Joulin, Beyond English-Centric Multilingual Machine Translation, arxiv, 2020

[4] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2020). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 7871–7880).

[5] L. Barrault, F. Bougares, L. Specia, C. Lala, D. Elliott, S. Frank, Findings of the Third Shared Task on Multimodal Machine Translation, *Proceedings of the Third Conference on Machine Translation: Shared Task Papers* (2018)

[6] Shuyan Zhou, Xiangkai Zeng, Yingqi Zhou, Antonios Anastasopoulos, Graham Neubig, Improving Robustness of Neural Machine Translation with Multi-task Learning

[7] S. Blodgett, S. Barocas, H. Daumé III, H. Wallach, Language (Technology) is Power: A Critical Survey of "Bias" in NLP, Proc. of ACL 2020

[8] M.R. Costa-jussà, An Analysis of Gender Bias studies in Natural Language Processing. Nature Machine Intelligence, Vol 1, Num 11, pages 495-496, November 2019

[9] Tang, Y., Tran, C., Li, X., Chen, P., Goyal, N., Chaudhary, V., Gu, J., & Fan, A. (2020). Multilingual Translation with Extensible Multilingual Pretraining and Finetuning.

[10] Hassan, H., Aue, A., Chen, C., Chowdhary, V., Clark, J., Federmann, C., Huang, X., Junczys-Dowmunt, M., Lewis, W., Li, M., Liu, S., Liu, T., Luo, R., Menezes, A., Qin, T., Seide, F., Tan, X., Tian, F., Wu, L., Wu, S., Xia, Y., Zhang, D., Zhang, Z., & Zhou, M. (2018). Achieving Human Parity on Automatic Chinese to English News Translation. ArXiv, abs/1803.05567.

[11] A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, Ll., Gomez, A., Kaiser

L. and, Polosukhin, I. (2017) Attention is all you need, CoRR, abs/1706.03762
[12] Ye Jia Ron J. Weiss Fadi Biadsy Wolfgang Macherey Melvin Johnson Zhifeng Chen Yonghui Wu, Direct speech-to-speech translation with a sequence-to-sequence model Interspeech (2019)