Visualizing Neural Machine Translation Attention and Confidence

Matīss Rikters[a], Mark Fishel[b], Ondřej Bojar[c]
[a]Faculty of Computing, University of Latvia
[b]Institute of Computer Science, University of Tartu
[c]Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics
E-mail: matiss@lielakeda.lv, fishel@ut.ee, bojar@ufal.mff.cuni.cz

Confidence Scores

$$CDP = \frac{1}{J}\sum_{j} \log\left(1 + \left(\sum_{i} \propto_{ji}\right)^{2}\right)$$

$$AP_{out} = -\frac{1}{I}\sum_{i}\sum_{j} \propto_{ji} \cdot log \propto_{ji}$$

$$AP_{in} = -\frac{1}{I}\sum_{j}\sum_{i} \propto_{ij} \cdot log \propto_{ij}$$

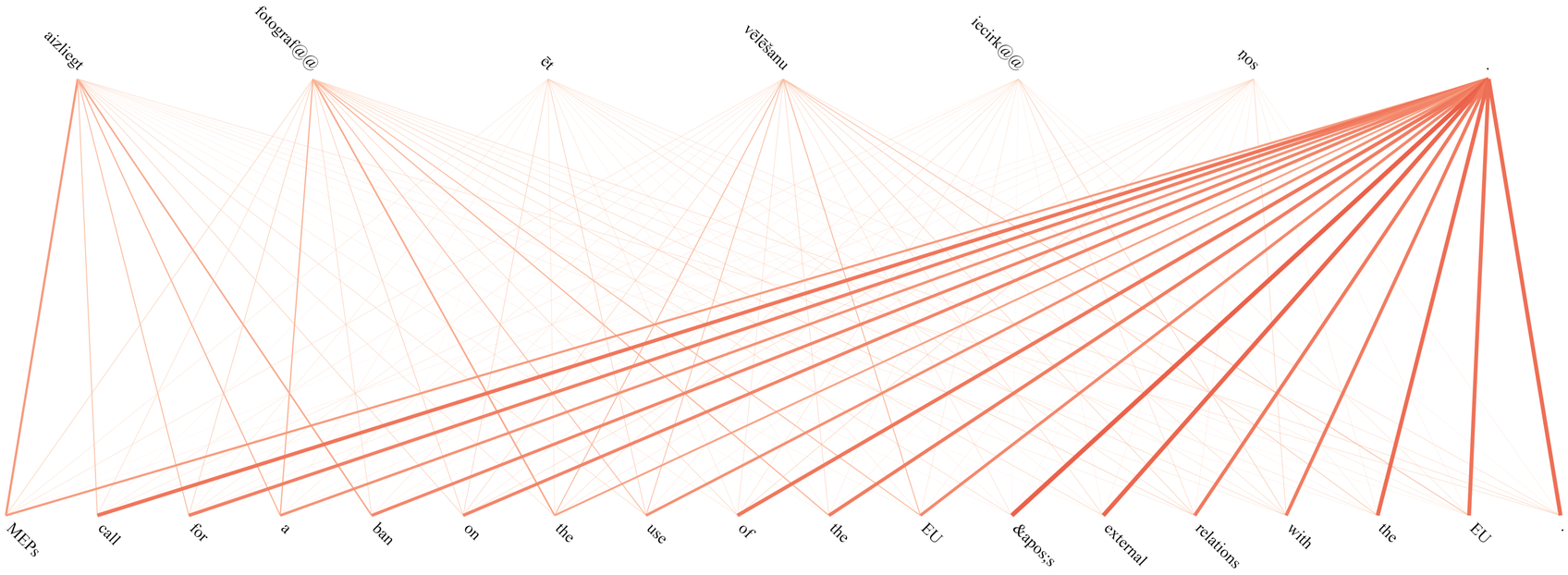
$$confidence = CDP + AP_{out} + AP_{in}$$

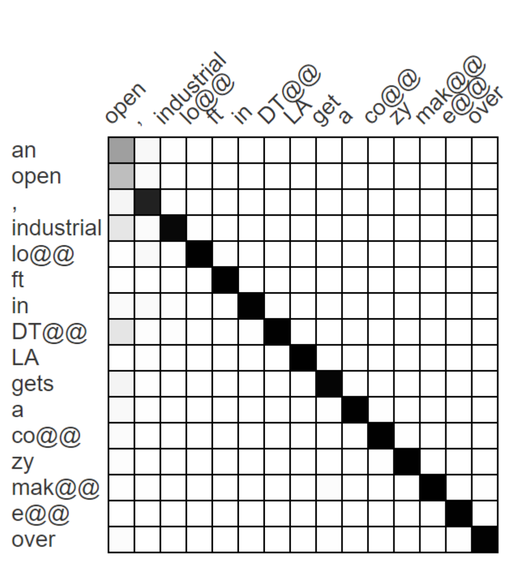
$$\text{percentage} = e^{-C(X^2)}$$

Terminal Visualisations

Lacking Confidence

Excessive Confidence





Features

Works with attention alignment data from

- Nematus
- Neural Monkey
- AmuNMT (fork github.com/barvins/amunmt)

Visualise translations in

- Linux Terminal or Windows PowerShell
- Web browser
- Line form or matrix form
- Save as PNG
- Sort and navigate dataset by confidence scores

GitHub Poster Demo

ej.uz/nmt-github ej.uz/nmt-poster ej.uz/nmt-attention

Acknowledgements

This research was supported by the ICT COST Action IC1207 ParseME: Parsing and multi-word expressions - towards linguistic precision and computational efficiency in natural language processing, the grant H2020-ICT-2014-1-645442 (QT21) and Charles University Research Programme "Progres" Q18+Q48.