

A framework for crowd rating quality evaluation in the statistical machine translation

Dmytro Chasovskyi, Ilya Verenich

Institute of Computer Science, University of Tartu, Estonia

{chasovsk, ilyav}@ut.ee



Introduction

Is this a good translation?

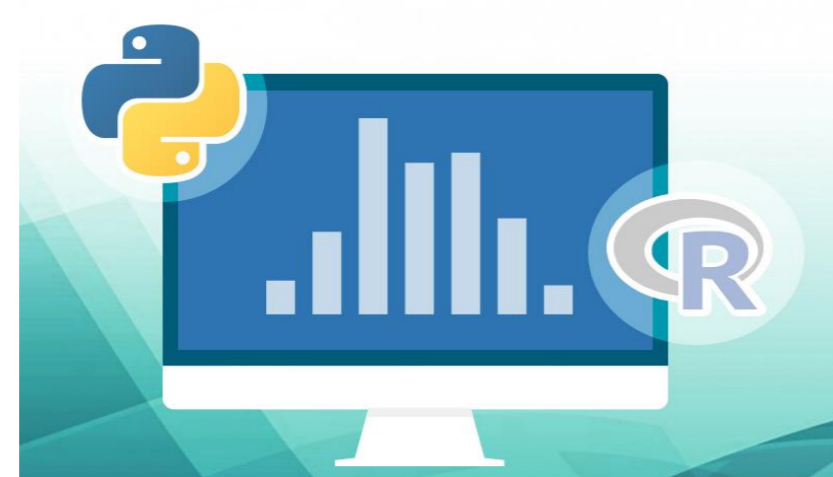


How to assess the translation quality?



Human quality evaluation
Hire professional linguists

+ High quality [1]
- Expensive [1]



Automated quality evaluation
Use metrics such as BLEU, TER,...

+ Cheap, fast, easy to implement
- Suboptimal quality [1]



Crowdrating quality evaluation
Users rate the machine translations or offer their own translation

+ Cheap, better quality [2, 3]
- No available implementation

Contribution

- Implement a framework for crowd rating quality evaluation
- Improve the existing translation algorithms by using the obtained dataset of user-picked translations (work in progress)

Proposed approach

Following are the key steps of the approach:

- User inputs a piece of text to translate
- The backend requests translations from translation providers
- User is prompted to choose the best translation among the returned results
- The best translation, along with the original text, are stored in the database

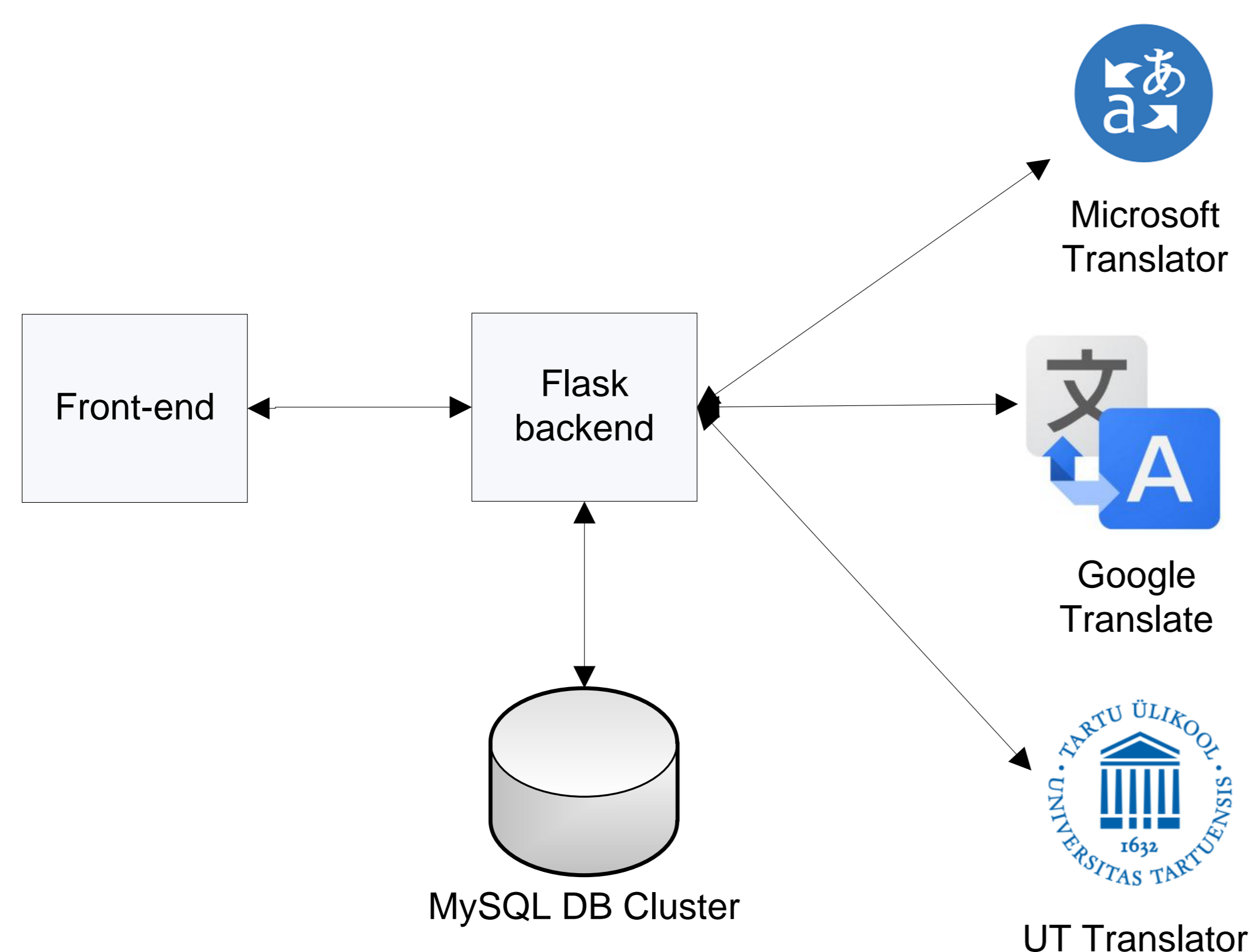
Translation providers

- Microsoft (Bing) Translator
- Google Translate
- "UT Translator" from *Kama* project of the Natural Language Processing Research Group – statistical and neural network-based

Implementation

The source code of the implementation is available at <https://github.com/ChameleonTartu/masintolge>. The implementation is live at masintolge.ut.ee.

Architecture of the proposed framework



Start page with three translation variants

Uni. Tartu Machine Translation



Please choose the best translation:



See väike Mölder jõuab rongile hüpata laisk koer



Kiire pruun rebane... hüppab üle laisa koera



The quick brown fox jumps over the lazy dog

Acknowledgment of the submitted translation

Uni. Tartu Machine Translation



Translation submitted!



See väike Mölder jõuab rongile hüpata laisk koer



Kiire pruun rebane... hüppab üle laisa koera



The quick brown fox jumps over the lazy dog

Future work

- Facilitating the UT translator with the crowd rated translations
- Gamification – users are awarded points for their contribution

Acknowledgment

The studies of Dmytro Chasovskyi are supported by the Estonian Foreign Ministry's Development Cooperation and Humanitarian Aid funds.

References

- CBojar, O., Chatterjee, R., Federmann, C., Graham, Y., Haddow, B., Huck (2016). Findings of the 2016 conference on machine translation (WMT16). In *Proceedings of the First Conference on Machine Translation (WMT)*. Vol. 2, pp. 131-198.
- Goto, S., Lin, D., & Ishida, T. (2014). Crowdsourcing for Evaluating Machine Translation Quality. In *LREC* (pp. 3456-3463).
- Zaidan, O. F., & Callison-Burch, C. (2011, June). Crowdsourcing translation: Professional quality from non-professionals. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*. Vol. 1, pp. 1220-1229.