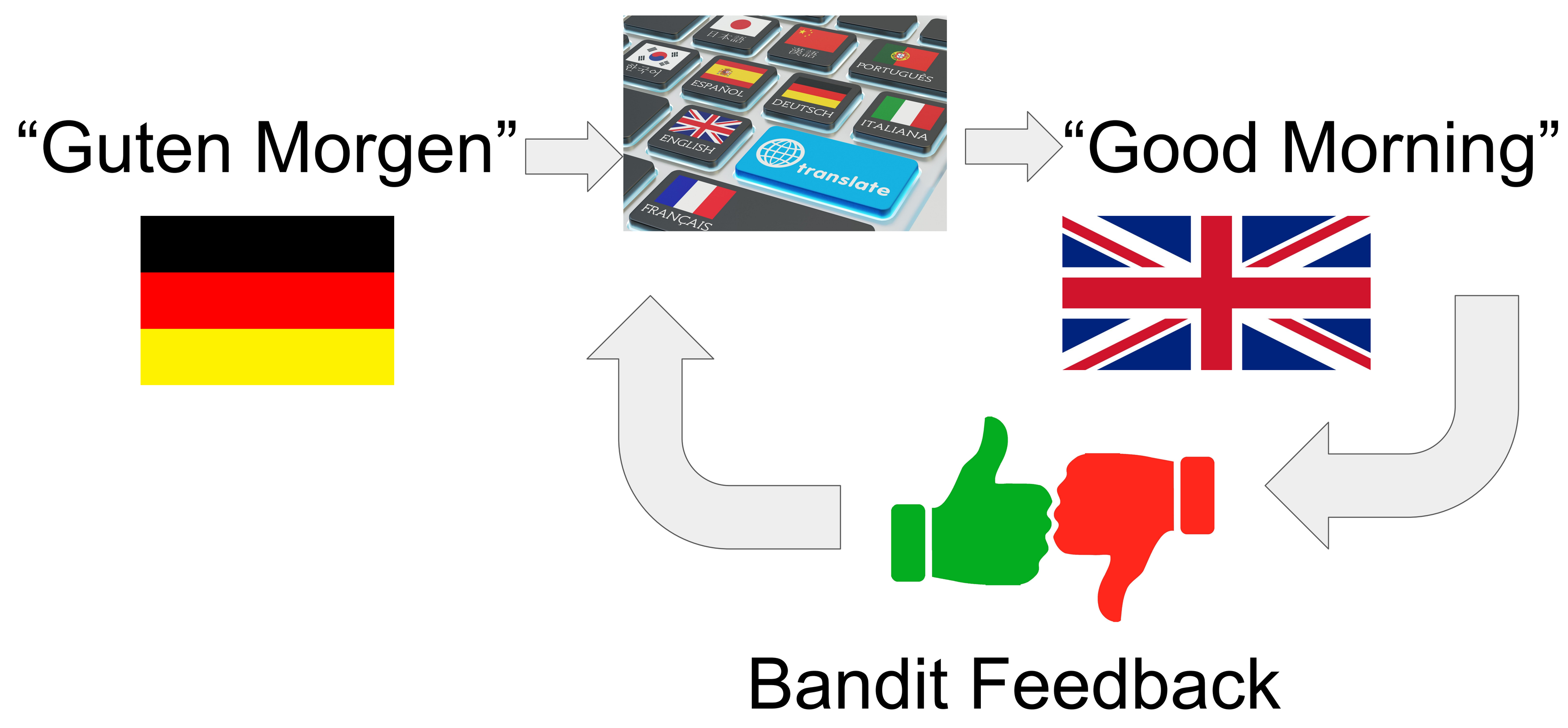
The University of Maryland Bandit Machine Translation System

Amr Sharaf, Shi Feng, Khanh Nguyen, Kianté Brantley, Hal Daumé III
University of Maryland

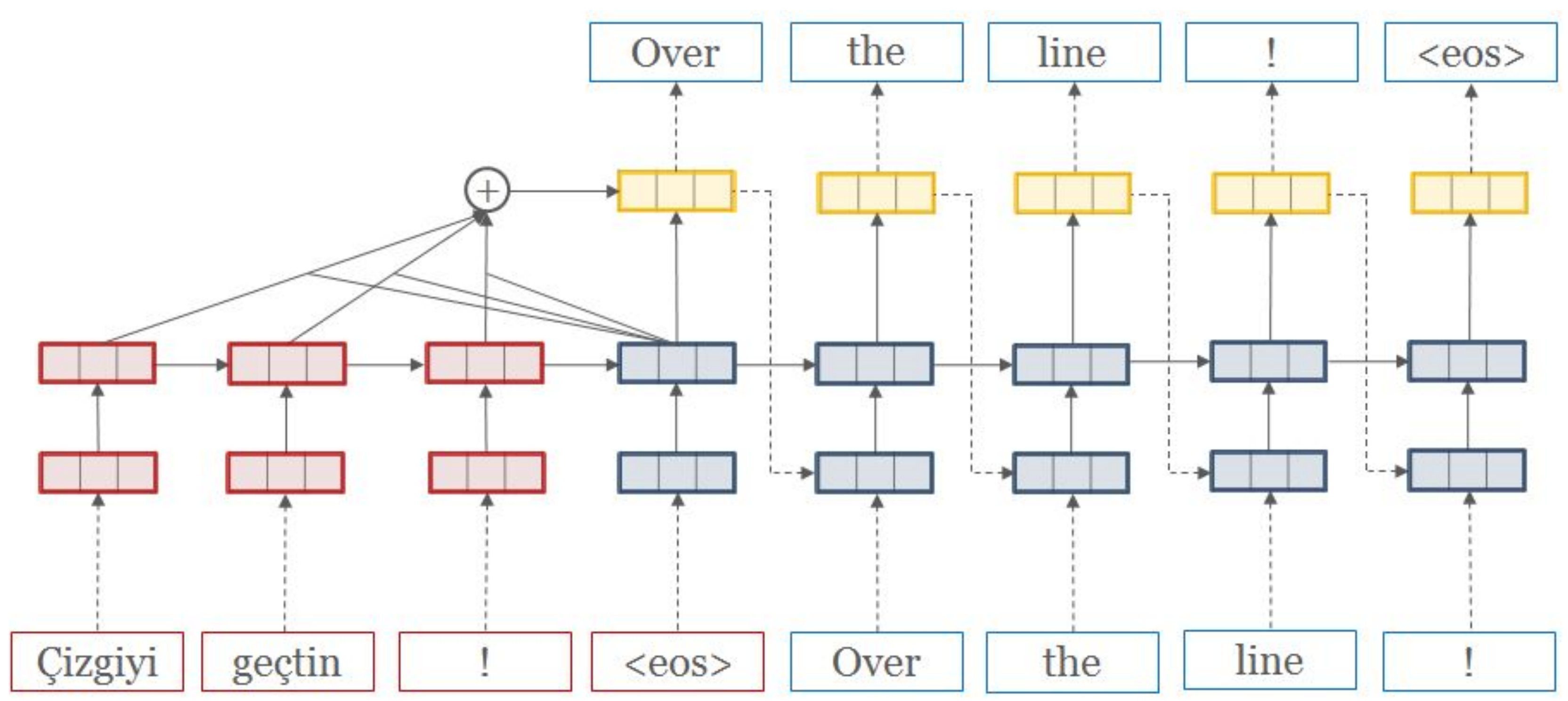
What is bandit machine translation?



The UMD Neural Bandit MT Systems

- Adapts a translation system to a new domain;
- Receives a German sentence to translate, produces an English sentence, and only gets a scalar score as feedback;
- Extends a standard Neural MT system in two ways:
  a. Robust reinforcement learning techniques to learn effectively from the bandit feedback;
  b. Domain adaptation using data selection from a large corpus of parallel data.

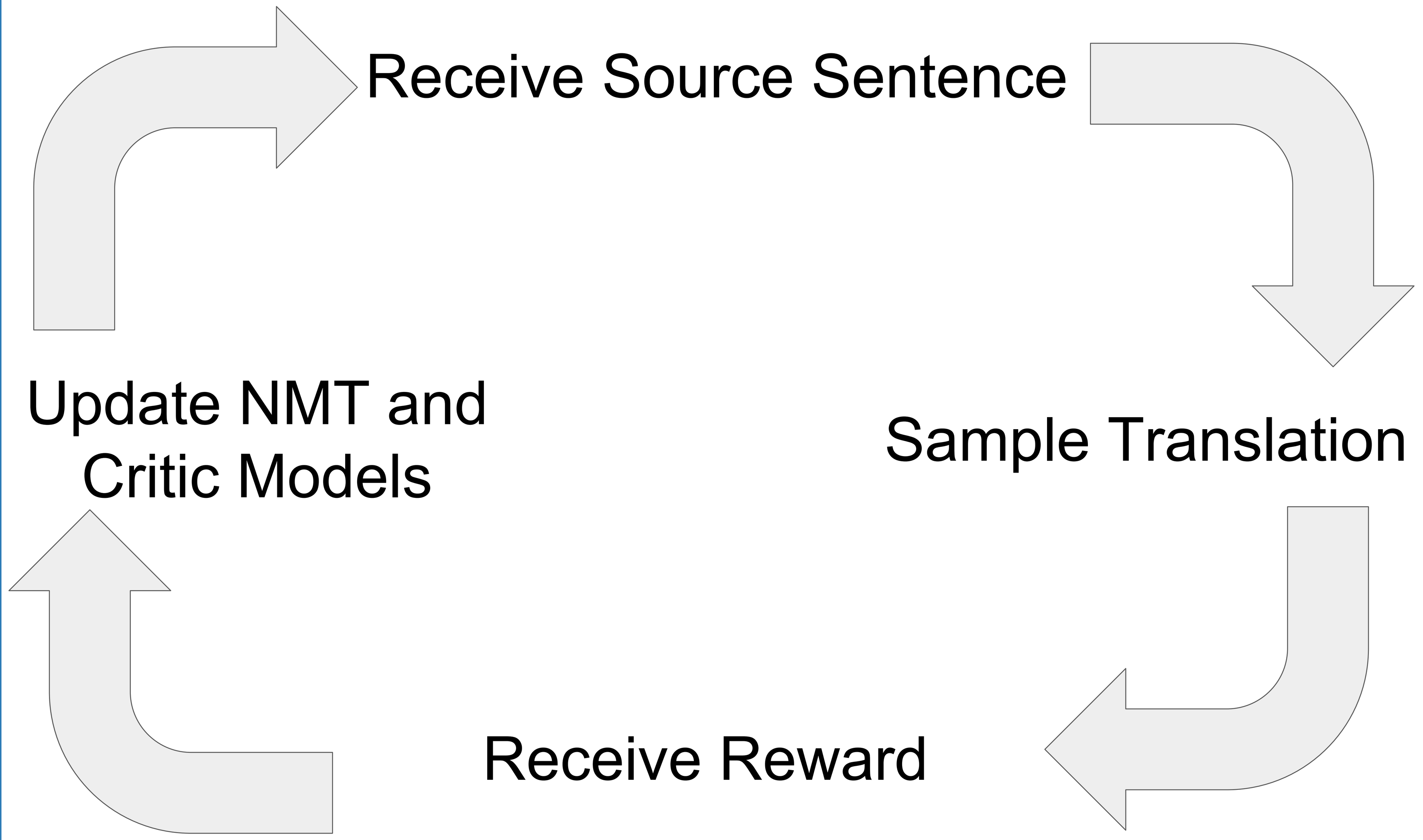
Neural MT Architecture



Schematic view of the OpenNMT neural machine translation model. (Credit: Klein et al., 2017)

Reinforcement Learning

- The NMT model can be viewed as a Markov decision process operating on a continuous state space:
  a. **States:** the hidden vectors hdec generated by the decoder.
  b. **Action space:** the target language’s vocabulary.
- Advantage Actor-Critic (A2C) algorithm:



Domain Adaptation

- **Approach:** Choose the best out-of-domain parallel data for training using Moore and Lewis (2010) cross-entropy based data selection technique.
- **Cross-Entropy Difference:** uses the cross-entropy difference $H_I(s) - H_O(s)$ for scoring a given sentence s, based on an in-domain language model LM_I and an out-of-domain language model LM_O
- **Cross-Entropy:**

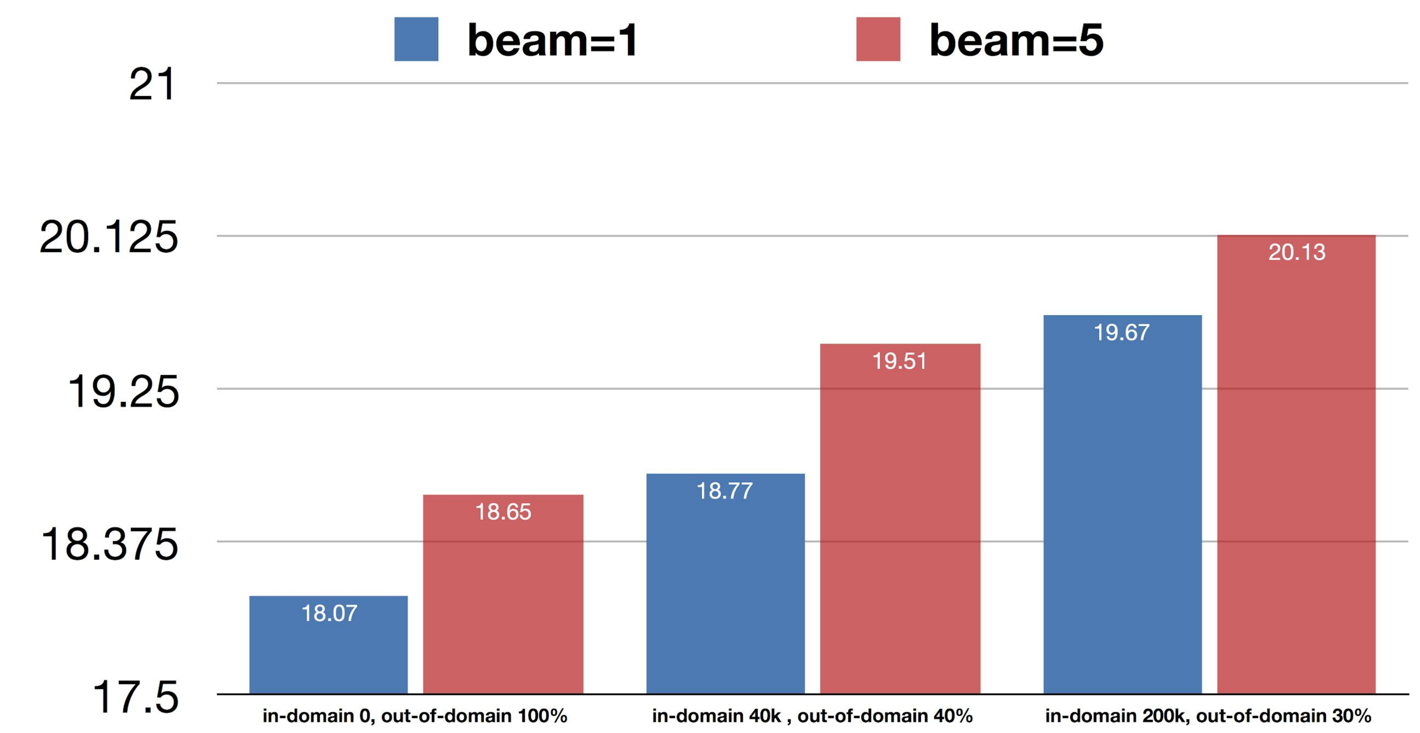
$$H(W) = -\frac{1}{n}\sum_{i=1}^{n} \log P_{LM}(w_i|w_1, \cdots, w_{i-1})$$

Experiments

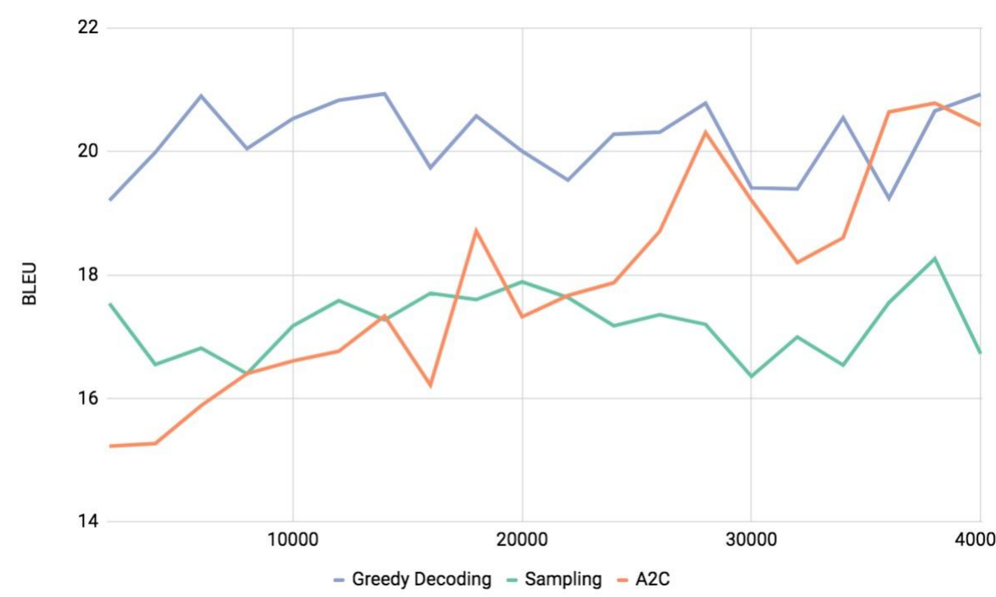
The results support the following conclusions:

1. **Domain adaptation:** data selection for domain adaptation alone improves translation quality by about **1.5 BLEU** points.

2. **Reinforcement Learning:** on top of the domain adaptation, reinforcement learning (which requires exploration) leads to an initial degradation of about 3 BLEU points, which is recovered (on development data) after approximately 40k sentences of bandit feedback.



Average BLEU scores of domain adaptation systems on the training server



Comparing sampling, greedy decoding, and the A2C algorithm on the development data.

References

1. Guillaume Klein, Yoon Kim, Yuntian Deng, Jean Senellart, and Alexander M Rush. 2017. Opennmt: Open-source toolkit for neural machine translation. arXiv preprint arXiv:1701.02810 .
2. Robert C Moore and William Lewis. 2010. Intelligent selection of language model training data. In Pro- ceedings of the ACL 2010 conference short papers. Association for Computational Linguistics, pages 220–224.