

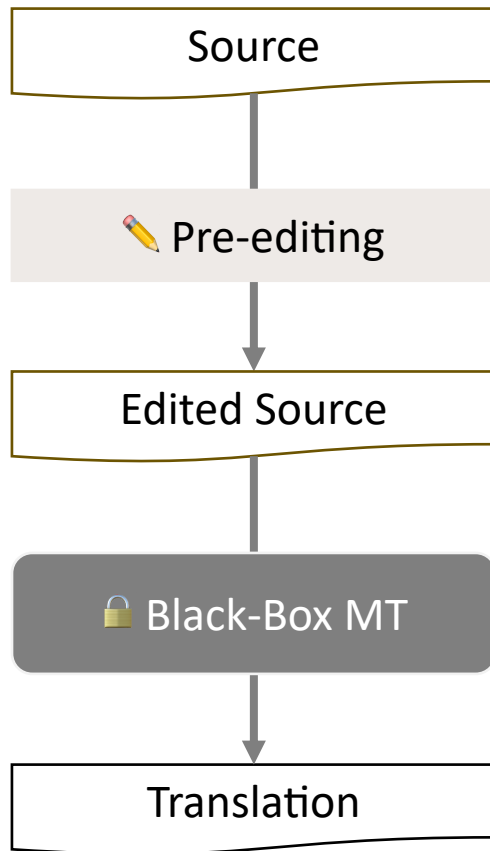
Targeted Source Text Editing for Machine Translation: Exploiting Quality Estimators and Large Language Models

Hyuga Koretaka[†], Atsushi Fujita[‡], Tomoyuki Kajiwara[†]

[†]Ehime University, Japan

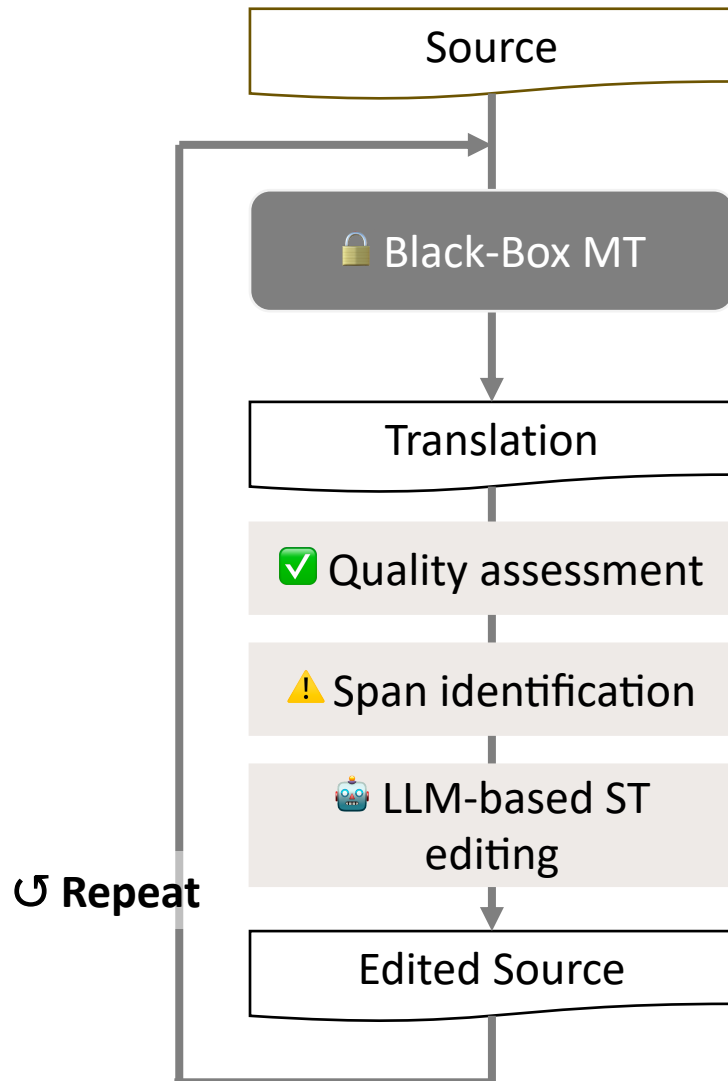
[‡]National Institute of Information and Communications Technology (NICT), Japan

Source Text Editing for MT



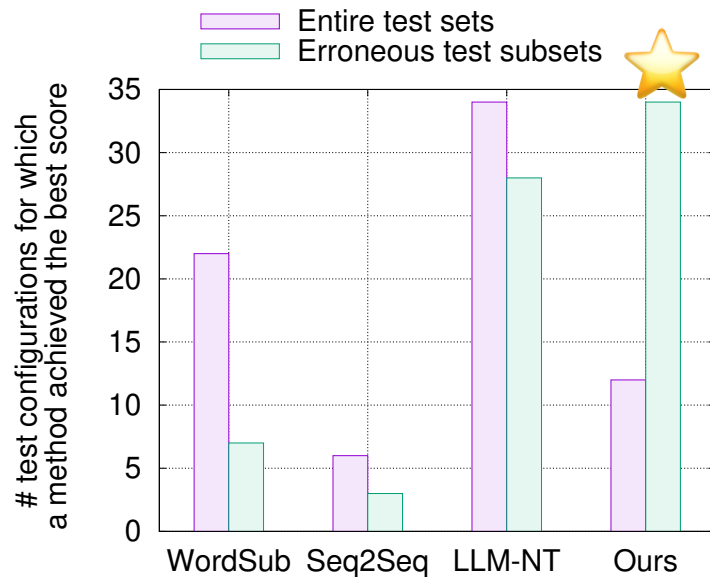
- Editing source texts to be translated so that they can be better translated
- An effective way to exploit black-box MT systems
 - Rule-based ([Shirai+ 93](#), [Kim+ 94](#), [Yamaguchi+ 98](#), [Shirai+ 98](#))
 - Seq2seq ([Sun+ 10](#), [Nanjo+12](#), [Mirkin+13](#), [Menta+ 20](#), [Koretaka+23](#))
 - LLM ([Ki & Carpuat 25](#))
- Existing methods remain two challenges
 - Performing only limited types of edit operations
 - With no reference to actual translation errors

Our Method for Targeted Source Text Editing



- Search for the best translation
 -  Evaluating translation quality
 - Segment-level quality estimator ([XCOMET-XL](#), Q)
- Through iteratively generating multiple hypotheses by
 -  Identifying source text span
 - Span-level quality estimator ([XCOMET-XL](#), E)
 - Word aligner ([OTAlign](#), A)
 -  Editing the identified source span ([Miyata & Fujita 17, 21](#))
 - LLM (e.g., [Llama-3.1-70B-Instruct](#), P) as a paraphraser

Our Method Performed Best for Erroneous Test Subsets



- Our targeted method achieved the best COMET score in the largest number of test configurations (MT system x test data)
 - Non-targeted counterpart (LLM-NT based on [Ki & Carpuat 25](#)) follows it
 - These two were more beneficial than non-LLM-based WordSub and Seq2Seq
- While being developed with Japanese-to-English examples, our method also worked well for English-to-Japanese tasks
 - {Japanese, English}-to-Chinese did not always benefit
 - Better segment-level QE will minimize the risk of quality deterioration