Is Reinforcement Learning helpful to style consistency in Machine Translation?

Yuan Cao, Hang Zhao, Xuting Zhang

Problem

Large Language Models (LLMs) have shown outstanding performance in the field of Machine Translation (MT) in recent years. Beyond basic translation accuracy, machine translation also requires style consistency between the source and target texts.

Reinforcement Learning has given LLMs strong reasoning abilities and achieved great success in tasks with clearly verifiable answers, such as math reasoning and code generation.

Our goal is to explore whether, in the translation process, a large model can use reinforcement learning to maintain translation accuracy while keeping the writing style consistent.

Model

We plan to use a small open-source model, such as Qwen3-1.7B, as the base model, and apply reinforcement learning on open-source datasets from multiple sources.

Unlike tasks with fixed answers, machine translation allows multiple valid outputs, making evaluation more difficult. To address this, we propose multi-objective reinforcement learning to improve translation quality. The optimization goals include the following aspects:

1. Format Reward: Use regular expressions to ensure that the model follows a structured response format.
2. Semantic Reward: Use BLEU and COMET to measure meaning similarity between the translation and the reference.
3. Style Reward: Use a pretrained style embedding model and a sentiment analysis model to compare the style similarity between the source text and the translation.