JAIST Repository
https://dspace.jaist.ac.jp/

Title	エンティティ認識型機械翻訳を改善するための多面的強化学習アプローチ
Author(s)	TRIEU, HOANG AN
Citation	
Issue Date	2026-06
Type	Thesis or Dissertation
Text version	ETD
URL	https://hdl.handle.net/10119/20625
Rights	
Description	Supervisor: NGUYEN, Minh Le

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

氏　　名	Trieu Hoang An		
学位の種類	博士（情報科学）		
学位記番号	博情第 586 号		
学位授与年月日	令和 8 年 6 月 24 日		
論文題目	A Multi-Aspect Reinforcement Learning Approach to Improving Entity-Aware Machine Translation		
論文審査委員	Nguyen Le Minh	JAIST	Professor
	Shinobu Hasegawa	JAIST	Professor
	Masaharu Mizumoto	JAIST	Assoc. Prof.
	Ashwin Ittoo	Liege University	Professor
	Masnizah Mohd	UKM University	Assoc. Prof.

論文の内容の要旨

Language barriers have long hindered effective knowledge sharing and the advancement of society. In recent years, the rapid development of machine translation, one of the most complex tasks in natural language processing, has contributed significantly to overcoming this obstacle. The impressive performance of modern translation systems demonstrates their capability in handling general-domain text with high accuracy. However, when it comes to terminology-sensitive domains such as medicine or law, these systems often fail to accurately translate domain-specific terms. This limitation poses a serious issue: even a single mistranslated term can alter the meaning of the original text and lead to severe misinterpretations, especially in high-stakes contexts. Motivated by this challenge, the goal of this thesis is to develop a machine translation system capable of reliably and correctly translating terminologies in specialized domains.

The first challenge lies in ensuring that the system maintains strong general-domain translation performance after domain adaptation. Fine-tuning a model for a specific field may cause it to lose previously learned knowledge, resulting in degraded performance on general text. To address this, we propose a multiple-aspect reinforcement learning approach that evaluates and optimizes the model using multiple complementary criteria. This technique allows us to control knowledge retention inside the model, preventing catastrophic forgetting while simultaneously enhancing its ability to correctly translate specialized terminology. We further introduce clear definitions of the proposed aspects and describe how they can be incorporated into the training process to improve overall system performance.

The second challenge pertains to the limited availability of high-quality data for assessing and improving terminology translation accuracy. Creating datasets that precisely identify and evaluate domain-specific terminology usage is costly and resource-intensive. To overcome this, we propose a data generation framework for producing a high-quality silver-standard dataset to support training and fine-tuning. Additionally, we outline criteria for constructing a smaller but reliable gold test dataset used for evaluation and testing.

In summary, this thesis aims to enhance machine translation systems by improving their ability to translate specialized terminology while preserving strong performance in general-domain translation. Through a combination of reinforcement learning and efficient data generation methods, our approach provides a promising direction for advancing terminology-aware machine translation.

Keywords: Machine Translation, Large Language Models, Entity-Aware Machine Translation, Reinforcement Learning, Dataset Construction.

論文審査の結果の要旨

Mr. Trieu Hoang An has completed his Ph.D. research on utilizing multi-aspect reinforcement learning (MRL) and multi-task learning for entity-aware machine translation. The proposed techniques demonstrate how LLMs can work together with advanced MRL methods to enhance the performance of machine translation in specific domains, such as the medical and legal domains.

The thesis addresses an important problem in machine translation: improving the accuracy of domain-specific terminology while maintaining strong general translation quality. Its motivation is clear, especially for high-stakes fields such as medicine and law, where mistranslated terms can cause serious misunderstandings.

The proposed multi-aspect reinforcement learning approach is meaningful because it aims to improve terminology translation while reducing catastrophic forgetting after domain adaptation. The data generation framework for silver-standard training data and gold-standard evaluation data also adds practical value, especially in domains where high-quality terminology resources are limited. In addition, the candidate developed a named entity-aware machine translation corpus for the medical domain, which is also a valuable contribution to the research community.

Overall, the thesis is well motivated and relevant to current NLP research. For the final version, it would be useful to clearly define the proposed aspects, explain the evaluation criteria for terminology accuracy, and provide strong experimental comparisons with existing methods.

This is an excellent dissertation, and we approve awarding a doctoral degree to Mr. Trieu Hoang An.