

数理・情報のフロンティア
2019 年度採択研究代表者

2020 年度 年次報告書

梶原 智之

愛媛大学 大学院理工学研究科
助教

自然言語処理の真価を引き出す言い換え生成

§ 1. 研究成果の概要

本研究では、機械翻訳への応用を中心に、自然言語処理における入力感度の課題に取り組んでいる。この研究の背景には、「大学を受験する」「大学を受ける」のように、入力文の意味が同じでも、モデルにとって得意な表現の翻訳は成功し、苦手な表現の翻訳は失敗するという問題がある。本研究では、このような些細な表現の変化に影響を受けない、安定して高品質な自然言語処理の実現を目指す。今年度は、入力文とその言い換えを1つの格子グラフとして表現する Lattice-to-Sequence モデルを提案し、英日翻訳の実験によって翻訳品質および安定性の改善を確認した。提案手法では、「受験する」「受ける」のような人間の語彙選択の曖昧性を考慮した機械翻訳を実現できる。

【代表的な原著論文情報】

- 1) 西原大貴, 梶原智之, 荒瀬由紀, 藤田篤. “言い換えラティスを用いたテキスト生成の性能改善”, 言語処理学会第 27 回年次大会, pp.1729-1733, 2021.
- 2) Daiki Nishihara, Tomoyuki Kajiwara. “Word Complexity Estimation for Japanese Lexical Simplification”, In Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020), pp.3107-3113, 2020.
- 3) 梶原智之, 西原大貴, 小平知範, 小町守. “日本語の語彙平易化のための言語資源の整備”, 自然言語処理, Vol.27, No.4, pp.801-824, 2020.
- 4) Akifumi Nakamachi, Hiroki Shimanaka, Tomoyuki Kajiwara, Mamoru Komachi. “TMUOU Submission for WMT20 Quality Estimation Shared Task”, In Proceedings of the Fifth Conference on Machine Translation (WMT 20), pp.1037-1041, 2020.