

ABCI 3.0開発加速利用 (2025年度) 成果概要 (公開用)

課題名：小規模LLMによるコード生成のハルシネーション検知

実施時期：2025年4月30日～2026年3月30日

所属機関名：日本女子大学

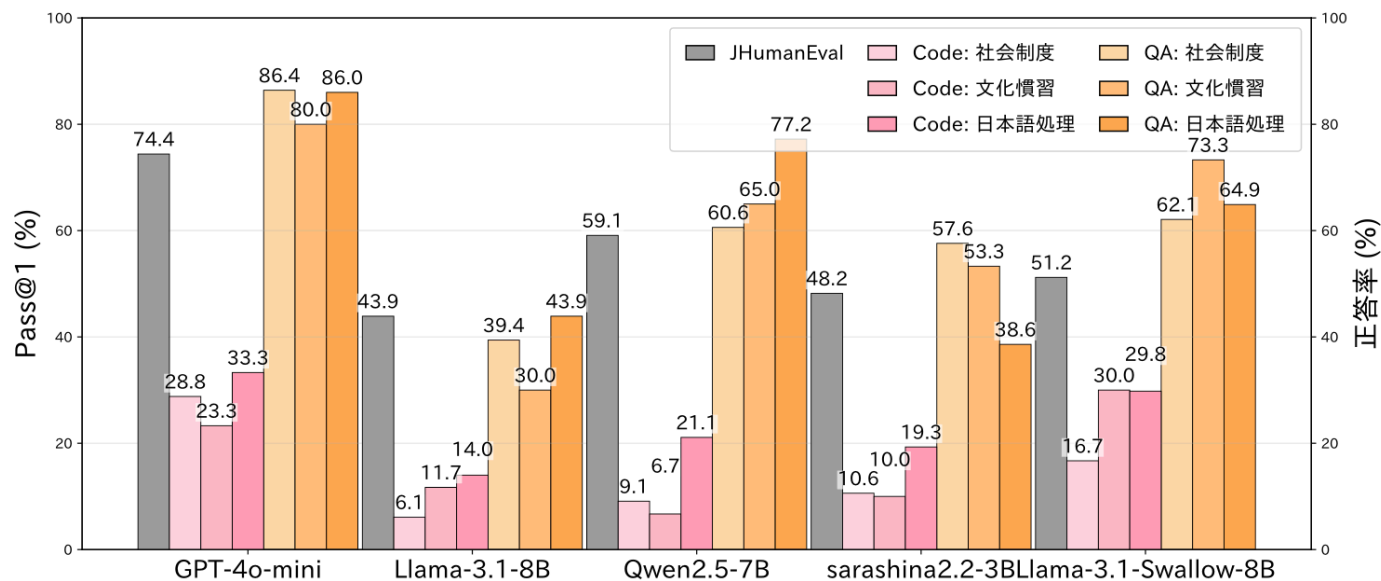
代表者氏名：倉光君郎

成果概要：

本研究では、特に日本語アプリを開発するために必要な日本特有の知識において、コード生成とハルシネーションを調査した。日本特有の知識を使うコード生成は、アルゴリズム等のグローバルな知識に基づくコード生成より、ハルシネーションを起こしやすかった。単に、日本の知識が足りないわけではなく、ドメイン知識とコード生成の知識の距離が推論に影響を与える示唆をえた。

成果のポイント：

我々は社会制度, 文化慣習, 日本語処理に関するコード生成ベンチマークを新たに開発した。これらは、アルゴリズム的には簡単であるが、アルゴリズム系ベンチマーク(JHumanEval)より低い(ハルシネーションが多かった)知識の欠如が原因と考えられるが、一方、各ドメインの知識をQA方式で検査すると、知識があることが確認される。ここから、ドメイン知識とコード生成の間には、推論の距離が存在することが示唆された。



成果についてより詳細な情報を提供しているWebページ、発表論文などの情報：

佐藤美唯, 伊東和香, 倉光君郎. 日本特有のコード生成ベンチマークの開発と評価. 情報処理学会論文誌 67 (3), 537-545, 2026-03-15

土田 悠佳, 西潟 優羽, 伊東 和香, 倉光 君郎. 大規模言語モデルの技能と知識の融合可能性. 人工知能学会全国大会第40回年次大会(JSAI2026), 2026年6月.