

数理・情報のフロンティア
2021 年度採択研究代表者

2022 年度
年次報告書

内藏 理史

情報・システム研究機構 国立情報学研究所
特任研究員

圏論と自動検証による機械学習の仕様保証

研究成果の概要

機械学習のアルゴリズムはしばしば確率的プログラムとして記述される。本研究では圏論的意味論を理論的な道具立てとして確率的プログラムの検証手法を与え、それによって機械学習のアルゴリズムが望ましい性質を持つことを保証することを目指している。

今年度はプログラムの検証において重要な役割を果たす最弱事前条件についての研究を行った。圏論を用いることで様々な計算効果に対する最弱事前条件が統一的に扱える枠組みを与えた先行研究を発展させる形で、高階関数型言語の最弱事前条件が CPS 変換を用いることで構文的に計算できるということを一般的な状況において証明した。さらにこの結果を適切に具体化することで(1)プログラムが出力する文字列についての検証を行う既存研究(2)確率的プログラムの平均コストを検証する既存研究(3)確率的プログラムのコストのモーメントを検証する手法、などが上記の主定理の系として導けることを示した。これらの結果は SYCO 10 や PPL 2023 のような国内外のワークショップにて口頭発表やポスター発表を行った。また現在国際会議論文の執筆も行っている。

また上記とも関連して(1)CPS 変換を用いたエフェクトハンドラーの検証や(2)CPS 変換によって得られた論理式が成り立っているか(妥当性)を判定するアルゴリズムについても検討を行った。(1)のエフェクトハンドラーは Pyro という確率的プログラミング言語でも用いられている言語機構であるが、上記の研究をエフェクトハンドラーに拡張することで CPS 変換を用いた検証が適用できる可能性がある。(2)については妥当性の判定に篩型システムを用いるアプローチが有望であると考え fairness の自動検証を目標に現在も研究を進めている。

【代表的な原著論文情報】

1) Aguirre, A., Katsumata, S., & Kura, S. (2022). Weakest preconditions in fibrations. *Mathematical Structures in Computer Science*, 32(4), 472-510. doi:10.1017/S0960129522000330