2023 年度年次報告書
次世代 AI を築く数理・情報科学の革新
2023 年度採択研究代表者

渡邊 知樹

総合研究大学院大学 複合科学研究科
大学院生

Compositional なモデル検査によるニューロシンボリック AI の安全性保証

研究成果の概要

本研究では、compositional なモデル検査によって、ニューロシンボリック AI の安全性保証を実現することを目指している。本年度は compositional なモデル検査の理論と応用の発展に関して研究成果が得られた。まず、研究項目の一つに挙げていた多目的最適化による compositional なモデル検査の高速化に関して成果を挙げ、モデル検査の最高峰国際会議である CAV、トップ国際会議の一つである TACAS にそれぞれ筆頭著者として論文が採択された。この一連の研究は、ニューロシンボリック AI に代表されるシステムの安全性保証に貢献する見込みが高い。また、別の方向性として、compositional な双模倣関係を理論的に整備した。この仕事は、理論計算機科学の最高峰国際会議である LICS に共著論文が採択された。モデル検査と密接に関係している双模倣関係の compositional なアプローチは、異なる尺度の安全性保証技術として発展する見込みがある。

【代表的な原著論文情報】

- 1) Kori, M., Watanabe, K., Rot, J., Katsumata, S. (2024). Composing Codensity Bisimulations. In: Dal Lago, U., Esparza, J. (eds) Logic in Computer Science. LICS 2024. To Appear.
- 2) Watanabe, K., van der Vegt, M., Junges, S., Hasuo, I. (2024). Compositional Value Iteration with Pareto Caching. In: Gurfinkel, A., Ganesh, V. (eds) Computer Aided Verification. CAV 2024. To Appear.
- 3) Watanabe, K., van der Vegt, M., Hasuo, I., Rot, J., Junges, S. (2024). Pareto Curves for Compositionally Model Checking String Diagrams of MDPs. In: Finkbeiner, B., Kovács, L. (eds) Tools and Algorithms for the Construction and Analysis of Systems. TACAS 2024. Lecture Notes in Computer Science, vol 14571, pp. 279–298. Springer, Cham.