Scalable Synthesis of Formally Verified Neural Value Function for Hamilton-Jacobi Reachability Analysis (Abstract Reprint)

Yujie Yang[1,2], Hanjiang Hu[2], Tianhao Wei[2], Shengbo Eben Li[1], Changliu Liu[2]

[1]Tsinghua University, China
[2]Carnegie Mellon University, USA

Abstract Reprint. This is an abstract reprint of the journal article by Yang, Hu, Wei, Li, and Liu (2025).

Abstract

Hamilton-Jacobi (HJ) reachability analysis provides a formal method for guaranteeing safety in constrained control problems. It synthesizes a value function to represent a long-term safe set called feasible region. Early synthesis methods based on state space discretization cannot scale to high-dimensional problems, while recent methods that use neural networks to approximate value functions result in unverifiable feasible regions. To achieve both scalability and verifiability, we propose a framework for synthesizing verified neural value functions for HJ reachability analysis. Our framework consists of three stages: pre-training, adversarial training, and verification-guided training. We design three techniques to address three challenges to improve scalability respectively: boundary-guided backtracking (BGB) to improve counterexample search efficiency, entering state regularization (ESR) to enlarge feasible region, and activation pattern alignment (APA) to accelerate neural network verification. We also provide a neural safety certificate synthesis and verification benchmark called Cersyve-9, which includes nine commonly used safe control tasks and supplements existing neural network verification benchmarks. Our framework successfully synthesizes verified neural value functions on all tasks, and our proposed three techniques exhibit superior scalability and efficiency compared with existing methods.

References

Yang, Y.; Hu, H.; Wei, T.; Li, S. E.; and Liu, C. 2025. Scalable Synthesis of Formally Verified Neural Value Function for Hamilton-Jacobi Reachability Analysis. *Journal of Artificial Intelligence Research*, 83: 19:1–19:36.