Invited Talk: Using SMT and Abstraction-Refinement for Neural Network Verification

Guy Katz¹, Yizhak Yisrael Elboher²

¹The Hebrew University of Jerusalem, Jerusalem, Israel ²The Hebrew University of Jerusalem, Jerusalem, Israel

Abstract

Deep neural networks are increasingly being used as controllers for safety-critical systems. Because neural networks are opaque, certifying their correctness is a significant challenge. To address this issue, several neural network verification approaches have recently been proposed, many of them based on SMT solving. However, these approaches afford limited scalability, and applying them to large networks can be challenging. In this talk we will discuss a framework that can enhance neural network verification techniques by using over-approximation to reduce the size of the network — thus making it more amenable to verification. This approximation is performed such that if the property holds for the smaller (abstract) network, it holds for the original as well. The over-approximation may be too coarse, in which case the underlying verification tool might return a spurious counterexample. Under such conditions, we can perform counterexample-guided refinement to adjust the approximation, and then repeat the process. This approach is orthogonal to, and can be integrated with, many existing verification techniques. For evaluation purposes, we integrate it with the recently proposed Marabou framework, and observe a significant improvement in Marabou's performance. Our experiments demonstrate the great potential of abstraction-refinement for verifying larger neural networks.

SMT'21: 19th International Workshop on Satisfiability Modulo Theories, July 18-19, 2021, Online

^{© 0 2021} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)