Karlsruher Institut für Technologie

**Application-oriented Formal Verification Research Group**
**Institute of Information Security and Dependability (KASTEL)**
Am Fasanengarten 5
76131 Karlsruhe
https://formal.kastel.kit.edu

# Modular Verification of Neural Networks

**Topic for ”Praxis der Forschung”**

**Philipp Kern, Samuel Teuber**

## Background I: Neural Network Verification

- **Question:** Does NN $g$ respect specification $\psi(x, y)$?
- **Problem:** Verification is **NP-complete** even for piece-wise linear feed-forward NNs
- Current approaches do not scale to State-of-the-Art networks

## Background II: Traditional Program Verification

- Break down **large program** into **smaller components**
- Verify properties on components
- **Compose properties** to prove correctness of complete program

## Idea: Modularize the verification process for NNs

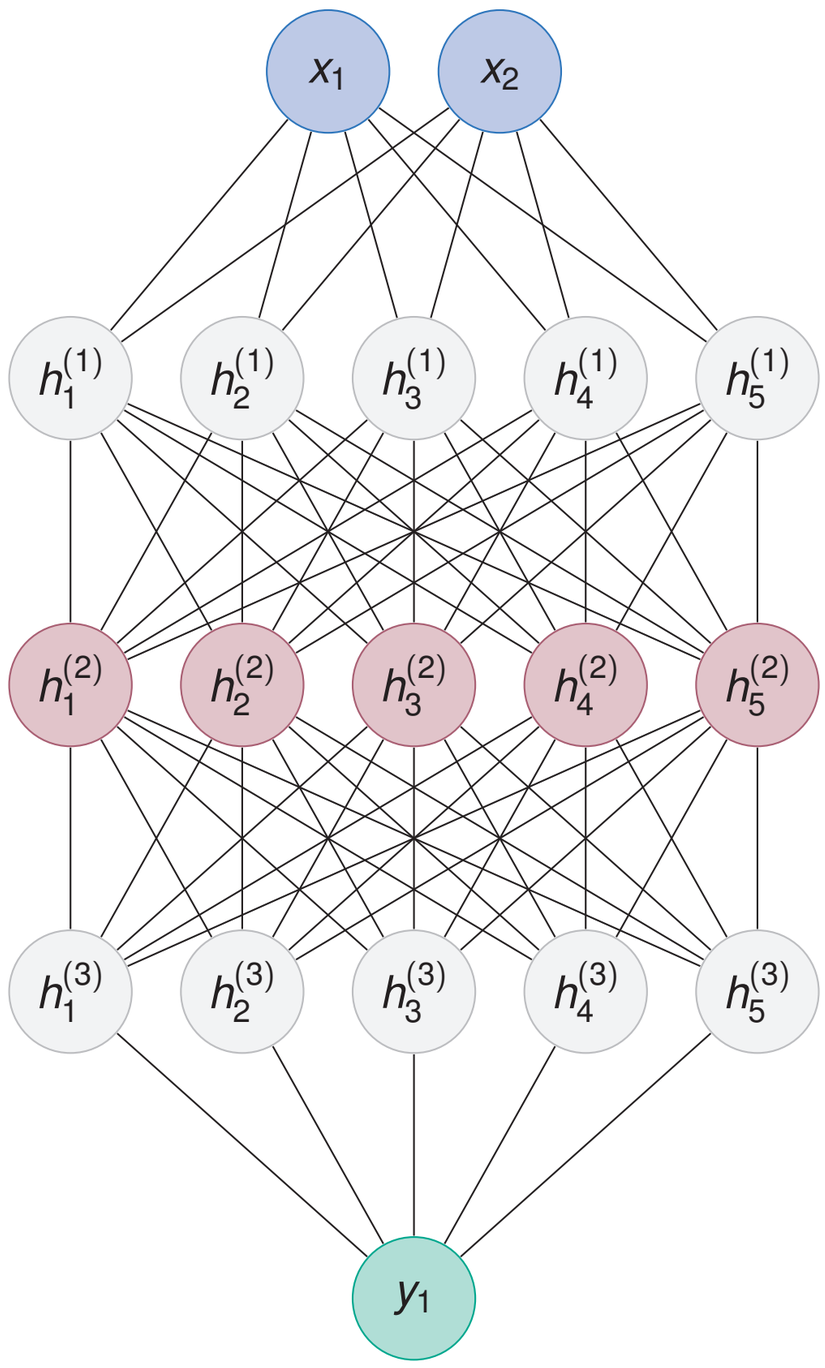
**Infer properties $\phi_1, \phi_2$ such that:**

$$\phi_1\left(x, h^{(2)}\right) \text{ holds on } x \to h^{(2)}$$

$$\phi_2\left(h^{(2)}, y\right) \text{ holds on } h^{(2)} \to y$$

$$\phi_1\left(x, h^{(2)}\right) \wedge \phi_2\left(h^{(2)}, y\right) \implies \psi(x, y)$$

**In particular**: Application to recursive neural networks (LSTMs)




**Relevant Literature:**

- Data-Driven Invariants [2, 4, 1]
- NN Invariants [5, 6, 3]
- Modular Bounded Model Checking [7]

[1] Jialu Bao et al. “Data-Driven Invariant Learning for Probabilistic Programs”. In: *Computer Aided Verification*. Ed. by Sharon Shoham et al. Cham: Springer International Publishing, 2022, pp. 33–54.

[2] Michael D. Ernst et al. “The Daikon system for dynamic detection of likely invariants”. In: *Science of Computer Programming* 69.1 (2007). Special issue on Experimental Software and Toolkits, pp. 35–45. ISSN: 0167-6423.

[3] Marc Fischer et al. “Shared Certificates for Neural Network Verification”. In: *Computer Aided Verification - 34th International Conference, CAV 2022, Haifa, Israel, August 7-10, 2022, Proceedings, Part I*. Ed. by Sharon Shoham et al. Vol. 13371. Lecture Notes in Computer Science. Springer, 2022, pp. 127–148.

[4] Cormac Flanagan et al. “Houdini, an Annotation Assistant for ESC/Java”. In: *FME 2001: Formal Methods for Increasing Software Productivity*. Ed. by José Nuno Oliveira et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 500–517.

[5] Divya Gopinath et al. “Finding Invariants in Deep Neural Networks”. In: *ArXiv* abs/1904.13215 (2019).

[6] Yuval Jacoby et al. “Verifying Recurrent Neural Networks Using Invariant Inference”. In: *Automated Technology for Verification and Analysis - 18th International Symposium, ATVA 2020, Hanoi, Vietnam, October 19-23, 2020, Proceedings*. Ed. by Dang Van Hung et al. Vol. 12302. Lecture Notes in Computer Science. Springer, 2020, pp. 57–74.

[7] Marko Kleine Büning et al. “Automatic Modularization of Large Programs for Bounded Model Checking”. In: *Formal Methods and Software Engineering*. Ed. by Yamine Ait-Ameur et al. Cham: Springer International Publishing, 2019, pp. 186–202.