

An Isabelle Formalization of the Expressiveness of Deep Learning (Extended Abstract)

Alexander Bentkamp¹, Jasmin Christian Blanchette^{2,3}, and Dietrich Klakow⁴

¹ Universität des Saarlandes, Saarbrücken, Germany

`s8albent@stud.uni-saarland.de`

² Inria Nancy – Grand Est & LORIA, Villers-lès-Nancy, France

`jasmin.blanchette@inria.fr`

³ Max-Planck-Institut für Informatik, Saarbrücken, Germany

⁴ Spoken Language Systems (LSV), Universität des Saarlandes, Saarbrücken, Germany

`dietrich.klakow@lsv.uni-saarland.de`

Deep learning has had a profound impact on computer science in recent years, with applications to search engines, image recognition and language processing, bioinformatics, and more. However, on the theoretical side only little is known about the reasons why these deep learning algorithms work so well. Recently, Cohen et al. [4] presented a theoretical approach using tensor theory that can explain the power of one particular deep learning algorithm called convolutional arithmetic circuits.

We present a formalization [1, 2] of their mathematical proof using the Isabelle/HOL proof assistant. This formalization simplifies and generalizes the original proof, while working around the limitations of the Isabelle type system. To support the formalization, we developed and extended reusable libraries of formalized mathematics.

The expressiveness of deep convolutional arithmetic circuits. Sum-product networks (SPNs) are a deep learning architecture [8], also known as arithmetic circuits. SPNs consist of a rooted directed acyclic graph with input variables as leaf nodes and two types of interior nodes: sum nodes and product nodes. The incoming edges of sum nodes are labeled with real weights, which have to be learned during training.

Convolutional arithmetic circuits (CACs) impose the structure of the frequently used convolutional neural networks (ConvNets) on SPNs. CACs consist of alternating convolutional and pooling layers, which are realized as collections of sum nodes and product nodes, respectively. CACs are equivalent to SimNets, which have been demonstrated to perform as well as these state of the art networks, even outperform them when computational resources are limited [3]. Moreover, Cohen et al.’s analysis of CACs allows to deduce properties of ConvNets with rectified linear unit (ReLU) activation [5].

The formalized theorem states that CACs enjoy complete depth efficiency, i.e., except for a negligible set S (a Lebesgue zero set in the weight space of the network), a deep network of polynomial size implements functions that require a network of super-polynomial size when the network is constrained to be shallow.

To simplify the formalization work, we restructured the original proof to obtain a more modular version, which generalizes the result as follows: The set S is not only a Lebesgue null set, but in particular the zero set of a multivariate polynomial $\neq 0$. This stronger theorem gives a clearer picture of how superior deep CACs are opposed to shallow ones in terms of expressiveness.

Formalization of the result in Isabelle/HOL. Isabelle/HOL is a natural choice for the formalization of this result because its strength lies in the high level of automation it provides. Moreover, it has an active community and includes a large library of formalized mathematical theories including measure theory, linear algebra, and polynomial algebra.

Our formalization provides a formal proof of the fundamental theorem from Cohen et al.’s original work for a network with non-shared weights. The formalization does not rely on any axioms beyond those that are built into Isabelle/HOL, and has an approximate total size of 7000 lines.

This project led to the development of general libraries that can be used in future formalizations about possibly completely different topics. Most prominently, we developed a library for tensors and their properties including the tensor product, tensor rank, and matricization. Moreover, we extended several libraries: We added the matrix rank and its properties to Thiemann and Yamada’s matrix library [9], adapted the definition of the Lebesgue measure by Hölzl and Himmelmann [7] to our purposes, and extended Lochbihler and Haftmann’s polynomial library [6] by various lemmas, including the theorem that zero sets of multivariate polynomials $\neq 0$ are Lebesgue null sets.

This formalization is a case study of applying proof assistants to a field where they have been barely used before, namely machine learning. It shows that the functionality and libraries of state of the art proof assistants such as Isabelle/HOL are up to the task. Admittedly, even the formalization of such relatively short proofs is labor-intensive. On the other hand, the process can not only lead to a computer verification of the result, but can also reveal new ideas and results. The generalization and simplifications we found demonstrate how formal proof development can also benefit the research outside the world of formal proofs.

References

- [1] Alexander Bentkamp. Expressiveness of deep learning. *Archive of Formal Proofs*, November 2016. http://isa-afp.org/entries/Deep_Learning_shtml, Formal proof development.
- [2] Alexander Bentkamp. An Isabelle formalization of the expressiveness of deep learning. Master’s thesis, Universität des Saarlandes, 2016. http://matryoshka.gforge.inria.fr/bentkamp_msc_thesis.pdf.
- [3] Nadav Cohen, Or Sharir, and Amnon Shashua. Deep simnets. *arXiv preprint arXiv:1506.03059*, 2015.
- [4] Nadav Cohen, Or Sharir, and Amnon Shashua. On the expressive power of deep learning: A tensor analysis. *CoRR*, abs/1509.05009, 2015.
- [5] Nadav Cohen and Amnon Shashua. Convolutional rectifier networks as generalized tensor decompositions. *arXiv preprint arXiv:1603.00162*, 2016.
- [6] Florian Haftmann, Andreas Lochbihler, and Wolfgang Schreiner. Towards abstract and executable multivariate polynomials in Isabelle. In *Isabelle Workshop*, volume 201, 2014.
- [7] Johannes Hölzl and Armin Heller. Three chapters of measure theory in Isabelle/HOL. In *Interactive Theorem Proving*, pages 135–151. Springer, 2011.
- [8] Hoifung Poon and Pedro Domingos. Sum-product networks: A new deep architecture. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 689–690. IEEE, 2011.
- [9] René Thiemann and Akihisa Yamada. Matrices, Jordan normal forms, and spectral radius theory. *Archive of Formal Proofs*, August 2015. http://isa-afp.org/entries/Jordan_Normal_Form_shtml, Formal proof development.