Automated Theorem Proving by HyperTree Proof Search with Retrieval-Augmented Tactic Generator

Sho Sonoda[1,5]*, Naoto Onda[5]†, Kei Tsukamoto[2,5]‡, Fumiya Uchiyama[2,5]§,
Akiyoshi Sannnai[3,5]¶ and Wataru Kumagai[4,5]‖

[1] RIKEN AIP, Tokyo, Japan
[2] The University of Tokyo, Tokyo, Japan
[3] Kyoto University, Kyoto, Japan
[4] OMRON SINIC X Corporation, Tokyo, Japan
[5] AutoRes

Abstract

We propose an automated theorem prover (ATP) based on the HyperTree Proof Search (HTPS) enhanced by retrieval-augmented tactic generator (ReProver).

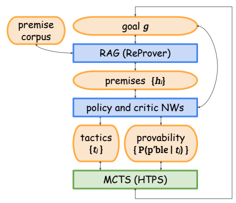
1 Introduction

Large Language Models (LLMs) are known to have limited mathematical abilities. For example, even in a simple task of integer addition, they tend to make mistakes when the digit number increases. Similarly, in mathematical reasoning, they tend to hallucinate when the length and complexity of reasoning increase. In order to reduce those hallucinations, we need a mechanism to automatically verify and correct logical inference by LLMs. Interactive Theorem Provers (ITPs)—such as Lean, Isabelle, Metamath, Coq, etc.——have drawn attention as such an automatic verification system for mathematical reasoning (see e.g. Lu et al., 2023; Li et al., 2024). In this study, we propose an automated theorem prover (ATP) based on the HyperTree Proof Search (HTPS) (Lample et al., 2022) enhanced by retrieval-augmented tactic generator (ReProver) (Yang et al., 2023).

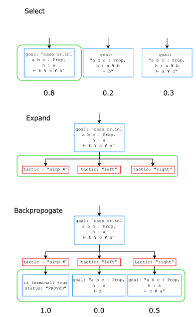
The problem of automated theorem proving with an ITP can be formulated as a reinforcement learning problem where the environment is the ITP, agents are computer algorithms such as LLMs, states are theorem statements, actions are tactics, and the reword is obtained when the agent either prove or disprove the initial statement. Based on the current statement (called the goal or a subgoal), LLM generates a tactic, and when that tactic is input to the ITP, the goal is rewritten to subgoal(s). If the rewriting of the statement is repeated and it reaches the goal state, the proof is complete. The difficulty of this problem lies in the fact that both the state space and the action space are discrete and infinitely large. Therefore, methods like brute force cannot be feasible, and methods using random walks are pessimistic.

The HyperTree Proof Search (HTPS) (Lample et al., 2022) is a Monte Carlo Tree Search (MCTS) algorithm tailored for theorem-proving problems inspired by AlphaZero (Silver et al., 2018). Specifically, they prepare a policy network that generates tactics from proof states, and a critic network that outputs proof possibilities from tactics, and then perform the three phases

*Corresponding Author (sho.sonoda@riken.jp)
†n.onda.go20@gmail.com
‡milanotukamoto@g.ecc.u-tokyo.ac.jp
§uchiyama-fumiya014@g.ecc.u-tokyo.ac.jp
¶sannai.akiyoshi.7z@kyoto-u.ac.jp
‖wkumagai1001@gmail.com



(a) Tactic generator (policy network) and provability estimator (value network) with RAG



(b) Select, Expand, Backpropagate phases in HTPS

Figure 1: Schematic Diagrams of Proposed ATP System

of MCTS: Selection, Expansion, and Backpropagation. Obviously, the performance of HTPS depends on the quality of the policy network. Unfortunately, however, the trained model is not public. Also, the authors reported that the training took more than 1000 days on A100 GPU time, making it difficult to reproduce.

The Retrieval-Augmented Prover (ReProver), on the other hand, is a premise selector and a tactic generator based on Retrieval-Augmented Generation (RAG) developed as a part of LeanDojo (Yang et al., 2023). Specifically, it searches mathlib (mathlib community, 2020) for premises that are similar to the target statement. By connecting to an encoder-decoder model in the latter part of (strict) ReProver, it can also generate tactics. The authors reported that the training took just five days. Also, the trained model has been made available online. However, it only suggests next (top-k) tactics, and does not support efficient search algorithm such as MCTS. In our proposed network, we adopt ReProver as the HTPS policy network (or tactic generator).

2 Experimental Results

We used the Mathlib dataset (mathlib community, 2020) for training, and the MiniF2F dataset (Zheng et al., 2022) for validation. Since MiniF2F was prepared (not in Lean 4 but) in Lean 3, we employed Mathlib3, or the Lean 3 version of Mathlib. The training was carried out on nVidia A100-SXM4-80GB for 24 hours. The timeout for each run was set to 150 seconds, and the number of premise selection was set to 20. The input to the network is the statement of the current goal (i.e., the proposition) written in Lean 3, and the outputs of the generator (policy network) and estimator (critic network) are a set of candidate tactics and the probability values of the provability given each tactic, respectively.

On the MiniF2F-test dataset, our system marked 26.2% for pass@1. As an ablation study, we have conducted an experiment with ReProver without HTPS on the same dataset (miniF2F-test), which marked 25.2% for pass@1, suggesting the improvement by HTPS.

Acknowledgments

SS was supported by JST PRESTO JPMJPR2125. WK was partially supported by the JSPS KAKENHI Grant Numbers 19H04071 and 23H04974.

References

Pan Lu, Liang Qiu, Wenhao Yu, Sean Welleck, and Kai-Wei Chang. A Survey of Deep Learning for Mathematical Reasoning. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14605–14631, 2023.

Zhaoyu Li, Jialiang Sun, Logan Murphy, Qidong Su, Zenan Li, Xian Zhang, Kaiyu Yang, and Xujie Si. A Survey on Deep Learning for Theorem Proving. *arXiv preprint: 2404.09939*, 2024.

Guillaume Lample, Timothee Lacroix, Marie-Anne Lachaux, Aurelien Rodriguez, Amaury Hayat, Thibaut Lavril, Gabriel Ebner, and Xavier Martinet. HyperTree Proof Search for Neural Theorem Proving. In *Advances in Neural Information Processing Systems*, volume 35, pages 26337–26349, 2022.

Kaiyu Yang, Aidan M Swope, Alex Gu, Rahul Chalamala, Peiyang Song, Shixing Yu, Saad Godil, Ryan Prenger, and Anima Anandkumar. LeanDojo: Theorem Proving with Retrieval-Augmented Language Models. In *Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2023.

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, and Demis Hassabis. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 362(6419):1140–1144, 2018.

The mathlib community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020, New Orleans, LA, USA, January 20-21, 2020*, pages 367–381, 2020.

Kunhao Zheng, Jesse Michael Han, and Stanislas Polu. MiniF2F: a cross-system benchmark for formal Olympiad-level mathematics. In *International Conference on Learning Representations*, 2022.