MC meeting: WG5 update

David M. Cerna, <u>Adrian De Lon</u>

October 11th 2024

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY
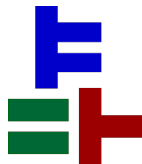
WG5 events (past/future)

- ▶ WG5 meeting at AITP 2022 europroofnet.github.io/wg5-aitp22
- ▶ Prague Workshop 2023 europroofnet.github.io/Prague23
  - ▶ Practical Aspects of Machine Learning in Theorem Proving
  - ▶ Datasets Generation for Data-Deficient Domains
- ▶ Joint WG4-WG5 meeting, colocated with CICM 2023 europroofnet.github.io/cambridge-2023
  - ▶ Workshop on Natural Formal Mathematics (NatFoM)
  - ▶ Workshop on Libraries of Formal Proofs and Natural Mathematical Language
- ▶ Alignment of Proof Systems and Machine Learning europroofnet.github.io/wg5-vienna24
- ▶ Summer School on AI for Reasoning and Processing of Mathematics, Kutaisi, Georgia europroofnet.github.io/Kutaisi24
- ▶ Theorem Proving with LLMs: SoA and Future Perspectives, Edinburgh, planned for March or April 2025
- ▶ Naproche School: Formalization in Controlled Natural Language, Bonn, planned for mid-June 2025

the

First deliverable completed

Learning Guided Automated Reasoning:
A Brief Survey

Lasse Blaauwbroek[1,2], David Cerna[3], Thibault Gauthier[1], Jan Jakubův[1,4],
Cezary Kaliszyk[4], Martin Suda[1], and Josef Urban[1]

[1] Czech Technical University in Prague
[2] Radboud University Nijmegen
[3] Czech Academy of Sciences Institute for Computer Science
[4] University of Innsbruck

Abstract. Automated theorem provers and formal proof assistants are
general reasoning systems that are in theory capable of proving arbitrarily hard theorems, thus solving arbitrary problems reducible to mathematics and logical reasoning. In practice, such systems however face
large combinatorial explosion, and therefore include many heuristics and
choice points that considerably influence their performance. This is an
opportunity for trained machine learning predictors, which can guide the
work of such reasoning systems. Conversely, deductive search supported
by the notion of logically valid proof allows one to train machine learning systems on large reasoning corpora. Such bodies of proof are usually
correct by construction and when combined with more and more precise
trained guidance they can be bootstrapped into very large corpora, with
increasingly long reasoning chains and possibly novel proof ideas.
In this paper we provide an overview of several automated reasoning and
theorem proving domains and the learning and AI methods that have
been so far developed for them. These include premise selection, proof
guidance in several settings, AI systems and feedback loops iterating
between reasoning and learning, and symbolic classification problems.

▶ *D8: Detailed technical report on the evaluation of techniques for
learning proof search guidance and premise selection in automated
theorem provers.*
link.springer.com/chapter/10.1007/978-3-031-61716-4_4

Second deliverable planned for 2025

- *D14: white paper on including restricted natural language proof formats to existing proof libraries*
- some progress during the Trimester Program *Prospects of Formal Mathematics* at the Hausdorff Institute of Mathematics in Bonn (from May to August 2024)
  - Lean (e.g. by Patrick Massot)
  - Naproche
  - Mizar (experimental)
  - GF
- will be worked on at WG5 events in 2025
- completion in summer 2025, after the final WG5 event

Upcoming event: Theorem Proving with LLMs (Edinburgh)

- ▶ two-day event, planned for March or April 2025
- ▶ number of expected people: 20–25
- ▶ WG5 meeting will take place during the event
- ▶ relevant to WG5 goal for 2024/25: *share knowledge and discuss the results and perspectives on the use of large language models in proof systems*
- ▶ deliverable D14 will be discussed and worked on during that meeting.

Upcoming event: Naproche school (Bonn)

▶ three-day event, planned for June 2025
▶ number of expected people: 20–25
▶ mix of local and EPN-funded participants
▶ we will ask HCM for additional funding
▶ also a followup to the HIM Workshop on *"Bridging between informal and formal"*
▶ deliverable D14 will be worked on during that meeting and completed shortly afterwards.