On Abstract Logics for Interacting Provers

Alexandra Fikiori[1], Asterios Gkantzounis[1,2], Foivos Skarpelos[1,3], and Petros Stefaneas[1]

[1]Department of Mathematics, School of Applied Mathematical and Physical Sciences, NTUA, Athens
[2]Supported by Papakyriakopoulos PhD Scholarship
[3]Supported by Research Committee PhD Scholarship

January 31, 2025

Useful interactions between Automated theorem proving (ATP) and Machine learning (ML), more specifically Large Language Models, (LLMs) would benefit from an abstract logical framework for proofs and validity. We propose the use of the theory of Institutions, an abstract model-theoretic tool for specification and programming.[8]

Lean and Metamath are proof assistants that rely on distinct mathematical-logical foundations, i.e. specific Institutions. Lean is built on dependent type theory, which enables the construction of highly structured mathematical objects and proofs. On the other hand, Metamath, centers around its main database, the Metamath Proof Explorer, which is based on first-order logic. Additionally, other Metamath databases explore alternative systems, including intuitionistic logic, Quine's New Foundations, and higher-order logic. Metamath verifies proofs using a simple substitution rule by focusing on the correctness of substitutions, without assuming a specific logic system. Hence we need a logic-independent approach to enhance the development of semantics for such tools. While Metamath and Lean are powerful tools, they have certain limitations that highlight the need for a more advanced logic-independent framework, such as Institutions.

Institutions form the core of categorical abstract model theory, which formalizes the notion of a logical system by defining its syntax, semantics, and the satisfaction relation between them. They provide a common framework that moves beyond the details of specific logics, which helps in studying their connections. An Institution consists of four main components: (a) a collection of *signatures* (serving as vocabularies for forming sentences in a logical system) and signature morphisms (b) a set of *sentences* for each signature, which represents the syntax of the logic (c) a set of *models* for each signature, which provides the meaning or semantics of the sentences and (d) a *satisfaction relation* that connects the models with the sentences, indicating which sentences are true in which models. Institutions extends Tarski's semantic definition of truth [7] and generalizes Barwise's Translation Axiom [6]. Institutions highlight the fact that *truth is invariant under change of notation.* The key aspect of any logical system is the satisfaction relationship between its syntax (i.e., its sentences) and its semantics (i.e., its models). While this relationship assumes a fixed vocabulary, Institutions provide the flexibility to work with multiple vocabularies simultaneously. This ability allows for translations between different logical systems, enabling one to be interpreted within another. A main result from Institution Theory gives the conditions under which you can translate proofs from one theorem prover to another.

Systems like Lean and Metamath that demonstrate how LLMs can be utilized to assist in constructing formal proofs, can clearly benefit from the Institutional approach. The main methodological implications for providing a general framework for ML and LLMs are:

- The relativistic view
- The translation of problems and solutions
- The multi-language aspect

In our presentation we will demonstrate how Institutions can be used to approach the semantics of systems like Lean and Metamath and the benefits of such an approach.

References

[1] Peiyang Song, Kaiyu Yang, and Anima Anandkumar. *Towards Large Language Models as Copilots for Theorem Proving in Lean.* Available at: `https://arxiv.org/pdf/2404.12534`, 2024.

[2] Stanislas Polu and Ilya Sutskever. *Generative Language Modeling for Automated Theorem Proving.* Available at: `https://arxiv.org/pdf/2009.03393`, 2020.

[3] The Lean Prover Community. *Dependent Type Theory.* Available at: `https://leanprover.github.io/theorem_proving_in_lean/dependent_type_theory.html`.

[4] Wikipedia contributors. *Metamath.* Available at: `https://en.wikipedia.org/wiki/Metamath`.

[5] Metamath Proof Explorer. *Set Theory Axioms in Metamath Proof Explorer.* Available at: `https://us.metamath.org/mpeuni/mmset.html#axioms`.

[6] Jon Barwise. *Axioms for abstract model theory*. Annals of Mathematical Logic, 7:221-265, 1974. Available at: https://www.sciencedirect.com/science/article/pii/0003484374900163.

[7] Alfred Tarski. *The semantic conception of truth*. Philosophical and Phenomenological Research, 4:13-47, 1944. Available at: https://uh.edu/~garson/Tarski.PDF.

[8] Joseph A. Goguen and Rod Burstall. *INSTITUTIONS: Abstract Model Theory for Specification and Programming*. Available at: https://cseweb.ucsd.edu/~goguen/pps/ins.pdf.