MaLeCoP
Machine Learning Connection Prover

Josef Urban\textsuperscript{1,*}, Jiří Vyskočil\textsuperscript{2,**}, and Petr Štěpánek\textsuperscript{3,***}

\textsuperscript{1} Radboud University Nijmegen, The Netherlands
\textsuperscript{2} Czech Technical University
\textsuperscript{3} Charles University, Czech Republic

Abstract. Probabilistic guidance based on learned knowledge is added to the connection tableau calculus and implemented on top of the \texttt{leanCoP} theorem prover, linking it to an external advisor system. In the typical mathematical setting of solving many problems in a large complex theory, learning from successful solutions is then used for guiding theorem proving attempts in the spirit of the \texttt{MaLARea} system. While in \texttt{MaLARea} learning-based axiom selection is done outside unmodified theorem provers, in \texttt{MaLeCoP} the learning-based selection is done inside the prover, and the interaction between learning of knowledge and its application can be much finer. This brings interesting possibilities for further construction and training of self-learning AI mathematical experts on large mathematical libraries, some of which are discussed. The initial implementation is evaluated on the \texttt{MPTP Challenge} large theory benchmark.

1 Introduction

This paper describes addition of machine learning and probabilistic guidance to connection tableau calculus, and the initial implementation in the \texttt{leanCoP} system using an interface to an external advice system. The paper is organized as follows: Section \textsuperscript{1} describes the recent developments in large-theory automated reasoning and motivation for the research described here. Section \textsuperscript{2} describes the machine learning (data-driven) paradigm and its use in guiding automated reasoning. Section \textsuperscript{3} shortly summarizes the existing \texttt{leanCoP} theorem prover based on connection tableaux. Section \textsuperscript{4} explains the general architecture for combining external machine learning guidance with a tableau prover. Section \textsuperscript{5} describes our experimental implementation. Section \textsuperscript{6} describes some experiments done with the initial implementation. Section \textsuperscript{7} concludes and discusses future work and extensions.

* Supported by The Netherlands Organization for Scientific Research (NWO) grants \texttt{Learning2Reason} and \texttt{MathWiki}.

** Supported by the Czech institutional grant MSM 6840770038.

*** Supported by the Grant Agency of Charles University, grant 9828/2009.

\textsuperscript{1} We would like to thank the anonymous referees for helping to significantly improve the presentation of this work.
1.1 Large-Theory Automated Reasoning

In the recent years, increasing amount of mathematics and knowledge in general is being expressed formally, in computer-understandable and computer-processable form. Large formal libraries of re-usable knowledge are built with interactive proof assistants, like Mizar, Isabelle, Coq, and HOL (Light). For example, the large Mizar Mathematical Library (MML) contains now (February 2011) over 1100 formal articles from various fields, covering substantial part of undergraduate mathematics. At the same time, the use of the formal approach is also increasing in non-mathematical fields, for example in software and hardware verification and in common-sense reasoning about real-world knowledge. This again leads to growth of formal knowledge bases in these fields.

Large formal theories are a recent challenge for the field of automated reasoning. The ability of ATP systems to reason inside large theories has started to improve after 2005, when first-order ATP translations of the particular formalisms used e.g. by Mizar [16], Isabelle [5], SUMO, and Cyc started to appear, and large theory benchmarks and competitions like MPTP Challenge and CASC LTB were introduced. The automated reasoning techniques developed so far for large theories can be broadly divided into two categories:

1. Techniques based purely on heuristic symbolic analysis of formulas available in problems.
2. Techniques taking into account also previous proofs.

The SInE preprocessor by Kryštof Hoder [4,17] seems to be so far the most successful heuristic in the first category. In domains like common-sense reasoning that typically lack large number of previous nontrivial and verified proofs and lemmas, and mostly consist of hierarchic definitions, such heuristics can sometimes even provide complete strategies for these domains.² MaLARea [18] is an example of a system from the second category. It is strong in hard mathematical domains, where the knowledge bases contain much less definitions than nontrivial lemmas and theorems, and previous verified proofs can be used for learning proof guidance. This approach is described in the next section, giving motivation for the work described in this paper.

2 Machine Learning in Large Theory ATP

The data-driven approaches to constructing algorithms have been recently successful in AI domains like web search, consumer choice prediction, autonomous vehicle control, and chess. In contrast to purely theory-driven approaches, when whole algorithms are constructed explicitly by humans, the data-driven approaches rely on deriving substantial parts of algorithms from large amounts of data. In the ATP domain, the use of machine learning started to be explored by the Munich

² For example, a Prolog-based premise selection preprocessor was used by Vampire in the CYC category of the 2008 CASC LTB competition to solve all problems.