Short-Term Scientific Mission Grant
- APPLICATION FORM[1] -

Action number: CA20111

Applicant name: Karol Pąk

Details of the STSM

Title: Automated Translation of Mizar Declarative Proof

Start and end date: 17/04/2023 to 18/05/223

Goals of the STSM

Purpose and summary of the STSM.

(max.200 word)

Declarative proof languages have been developed in many systems to provide more readable and more manipulation proofs style. The two earliest such languages, the Mizar language and Isabelle/Isar represent completely opposite approaches and include a number of incompatible features. Such support for formal proofs in both cases has given rise to two world's largest databases of formalized knowledge, but simultaneously prevent easy migration between them. We reconstructed selected components of the Mizar environment in the Isabelle logical framework as well as we successfully tested the simultaneous use of theory from Isabelle/HOL and the exported part of MML. However, the automatic export must be corrected by hand.

Goals: The aim of the proposed STSM is to improve our automatic export of Mizar scripts by covering the whole system of Mizar types and reducing the number of individual proof steps justifications that are unable to handle. More concretely, we want to extract more background knowledge from the Mizar verification process and to match it to the proof steps to complete their justifications using dedicated heuristics or machine learning techniques.

Working Plan

Description of the work to be carried out by the applicant.

(max.500 word)

During the 4-week visit we plan to:

1. Extract properties of predicates and terms that are implicitly used in the justification.
2. Reconstructing the full list of the first-order logic functors and predicates in the justification of a

[1] This form is part of the application for a grant to visit a host organisation located in a different country than the country of affiliation. It is submitted to the COST Action MC via-e-COST. The Grant Awarding Coordinator coordinates the evaluation on behalf of the Action MC and informs the Grant Holder of the result of the evaluation for issuing the Grant Letter.

given step that refers to a Mizar schema (second-order theorem).
3. Covering structural Mizar types by our automatic export.
4. Discuss and pre-implement selective import of background information.
5. Experimental implementation of solutions that improve the number of acceptable justifications.

Today we have an infrastructure to extract the semantic information with the original Mizar proof script that combines the syntactic information provided by the Mizar parser with the semantic one provided by the Mizar analyzer. Using this solution we were able to automatically translate definitions, notations, typing information, and the declarative proof outlines of 200 articles from the Mizar Mathematical Library (MML) to the Isabelle/Mizar object logic [1], [2]. However, the explicit justifications of the individual Mizar by steps are mostly insufficient for the Isabelle/Mizar automation [3] and have to be often extended by a list of substitutions or by the Mizar background information [1]. The latter we can approximate based on the form of the justified expression (1). Similarly, in the case of the justification by reference to a given Mizar scheme we expect to be able to determine a list of logic functors and predicates, eg. directly from the Mizar schematizer (2), which verifies the correctness of such justifications.

The next and most important step for our automatic translation will be to fulfill the gap between the information about Mizar structural types that we can extract from the semantic representation of Mizar proof scripts, and the information that must be provided for the emulation of the Mizar structure in the Isabelle/Mizar (3). It is important to note that the Mizar structured types are used in almost 80% of the MML. Our experience so far has shown [2], as it is implemented in the Mizar system, that it is necessary to selectively import background information to speed up our Isabelle/Mizar automation (4).

The most effective solution to realize the last stage would be the extraction of substitutions performed in the justification of each step directly from the Mizar verifier. However, all our attempts so far to extract this information have failed. Therefore, we plan further experimental development of proof methods in Isabelle/Mizar, that derive additional information as the Mizar typing information before a standard Isabelle method is called on the goal.

[1] Cezary Kaliszyk, Karol Pąk, Semantics of Mizar as an Isabelle Object Logic, Journal of Automated Reasoning, 2018.

[2] Cezary Kaliszyk, Karol Pąk, Declarative Proof Translation, Tenth International Conference, Interactive Theorem Proving, ITP 2019, Portland, OR, USA, September 8-13, 2019. Proceedings, volume 141 of Leibniz International Proceedings in Informatics (LIPIcs), pages 35:1-35:7. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2019.

[3] Cezary Kaliszyk, Karol Pąk, Isabelle Import Infrastructure for the Mizar Mathematical Library, In F. Rabe and W. M. Farmer and G. O. Passmore and A. Youssef, editors, Intelligent Computer Mathematics - 11th International Conference, CICM 2018, Hagenberg, Austria, August 13-17, 2018, Proceedings, volume 11006 of Lecture Notes in Computer Science, pages 131-146. Springer, 2018.

Expected outputs and contribution to the Action MoU objectives and deliverables.

Main expected results and their contribution to the progress towards the Action objectives (either research coordination and/or capacity building objectives) and deliverables.

(max.500 words)

The main expected result of this mission is the development of a prototype of syntactic-semantic export of the Mizar scripts to provide automatic translation of Mizar scripts that use Mizar structural types. We expect that our prototype will provide script translation where variables of structured type are used, but we will also focus on generating information (with justifications) calculated by the Mizar system based on a given structure definition, including multi-inheritance rules and non-emptyness of the defined type.

We also expect progress in improving the number of acceptable justifications in our translation, especially in the case of justifications that refer to Mizar schemes. We plan to expand the list of

premises indicated explicitly in the justification of a given proof step, by a list of instantiated properties of predicates and functors, that occur in the statement of the proof step. This will bring us closer to independent verification of the MML in the Isabelle system as well as to combine results proved in Tarski-Grothendieck set theory with ones proved in plain higher-order logic. Indeed, we proved that the higher-order Tarski-Grothendieck set theory has a model if a 2-inaccessible cardinal exists [1].

[1] Chad E. Brown, Cezary Kaliszyk, Karol Pąk, Higher-order Tarski Grothendieck as a Foundation for Formal Proof, Tenth International Conference, Interactive Theorem Proving, ITP 2019, Portland, OR, USA, September 8-13, 2019. Proceedings, volume 141 of Leibniz International Proceedings in Informatics (LIPIcs), pages 9:1-9:16. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2019.