Report on the outcomes of a Short-Term Scientific Mission[1]

Action number: CA20111

Grantee name: Ján Perháč

<u>Details of the STSM</u>

Title: Proof Theory for Programmers: Promoting Output of Theorem Provers to Formative Feedback for Computer Science Students

Start and end date:

01/06/2024 to 09/06/2024

<u>Description of the work carried out during the STSM</u>

Description of the activities carried out during the STSM. Any deviations from the initial working plan shall also be described in this section.

Thanks to this STSM, I visited Michael Kirkedal Thomsen, an associate professor at the Department of Informatics, University of Oslo (UiO). Together, with Joachim Tilsted Kristensen (Michael's PhD student) we have carried out the following activities:

- We have continued with the development of our theorem prover for teaching: OnlineProver. We have extended OnlineProver with definitions of new languages such as the imperative toy language called While, simply typed lambda calculus, together with a set of exercises. Currently, deployed the version supports natural deduction proofs in propositional logic with 16 exercises.

- We have discussed the results of the first experiment. We have used OnlineProver in the course Logic for Informaticians at the Technical University of Košice in the spring semester of 2024. Based on the first experiment, we have submitted one extended abstract to the Types conference [1] and one to the ThEdu'24 [2]. We have discussed the results of the experiment and analyzed received (extensive) reviews from both conferences. Based on that, we have prepared outlines for two full papers for the conferences' postproceedings.

- We have designed future experiments for OnlineProver. We have extended our network with another colleague from the UiO, where we plan to use our tool in his course on logic in the next semester. We have also extended our network with people from the University of Copenhagen

---

**Funded by
the European Union**

and IT University of Copenhagen.

- We have designed an architecture for the extension of our tool, that allows us to use generative AI to provide formative feedback. Our design allows us to do the behavioral analysis of a user proof, such as checking how many steps and what kind of paths, a user tried until they managed to close a proof. Based on such data, together with a set of "error messages" we will be able to use a language model to generate appropriate formative feedback.

Description of the STSM main achievements and planned follow-up activities

Description and assessment of whether the STSM achieved its planned goals and expected outcomes, including specific contribution to Action objective and deliverables, or publications resulting from the STSM. Agreed plans for future follow-up collaborations shall also be described in this section.

We have extended OnlineProver with support for the simply typed lambda calculus and the simple imperative language While. After a discussion with a new member of our network at UIO, we plan to extend our tool with support for sequent calculus in first-order logic. We also plan to use our tool in courses at the University of Copenhagen and the IT University of Copenhagen.

We have designed a new experiment for students in the next semester, as well as one for experts in the field. We will present our tool at the 13th International Workshop on Theorem Proving Components for Educational Software [2], a workshop at CADE30, part of IJCAR 2024, in Nancy, France, from July 1-6, 2024. We plan to conduct the mentioned experiment with experts and professionals during the event. We will analyze the collected feedback and use it in the future development of our tool. We hope that we will find new collaborators at the mentioned event.

We plan to implement our design (created during the STSM) for the use of generative AI to provide formative feedback. In the autumn semester, we plan to design and conduct new experiments.

Based on the outlines prepared during the STSM, we plan to submit two full papers for the post-proceedings of the Types conference [1], as well as the International Workshop on Theorem Proving Components for Educational Software [2].

[1] https://types2024.itu.dk/

[2] https://www.uc.pt/en/congressos/thedu/ThEdu24