Advancing Verification Competitions as a Scientific Method
18 – 12 February 2019 Lorentz Center @Snelius

The goal of this Lorentz workshop was to understand and advance various kinds of verification competitions as a scientific method. We aimed to explore the organizational factors that contribute to the scientific impact of competitions. The gained insights will be used for improving existing competitions and for successfully creating new competitions that serve as a driver for the progress in the field of formal verification.

During the meeting, we collected the content for writing a joint white paper on verification competitions. We got a comprehensive overview on the structure and setup of many recent verification competitions and we identified a catalogue of general guidelines. This white paper will serve as a "reference manual" for setting up and running competitions. Ideally, this work becomes the starting point of an exchange platform for competition organizers.

In emerging collaborations, a big topic will be the selection of benchmarks. This is currently a big challenge in all competitions. The discussions on this topic might lead to joint projects between the participating data scientists and competition organizers. By applying techniques from data science, the benchmark selection process might move from a currently quite ad-hoc approach to a scientifically well-founded procedure.

The workshop was an extensive forum to exchange between verification competition organizers of different communities. It was very interesting to learn that even though the tasks given to the competition participants are very heterogeneous in their nature, the challenges occurring in the competition organization are very similar. The exchange of experiences was extremely valuable, because so far there is hardly any documentation on the organization of competitions.

The first day was entirely filled with the introduction of the different competitions. As the presented competitions are affiliated to very different communities, this was crucial for gaining a common understanding. The other four days were dedicated to special topics: benchmarks and industrial applications on Tuesday, data analysis and result aggregation on Wednesday, execution platforms on Thursday, and finally, result validation and exploitation on Friday. In the morning, we started with two 45 keynote presentations related to the topic of the day. The other sessions were used for discussions and for working on the white paper.

We invited participants who did not give a keynote talk, to start the discussion session with a short kickoff presentation and to moderate the discussions. As preparation, we conducted an online survey among the participants, and we provided a summary of the results to the moderators of the respective sessions. This format had the effect that every participant took an active role in the workshop and it also resulted in a vivid exchange between the participants.

We greatly appreciated the hospitality of the Lorentz center. We especially enjoyed the manifold facilities of the venue and the professional support of the Lorentz staff.