JOHANNES KEPLER
UNIVERSITÄT LINZ

**o.Univ.-Prof. Dr.**
**Hanspeter Mössenböck**
Institute for System Software

T +43 732 2468 4340
F +43 732 2468 4345
hanspeter.moessenboeck@jku.at

Secretary:
**Karin Gusenbauer**
Ext 4342
karin.gusenbauer@jku.at

Master's Thesis / Project in Software Engineering

**Qualitative Assessment of Compiler Metrics in a Predictive Machine Learning Task**

| | |
|---|---|
| Student: | Wolfgang Kisling<br>k01156093 |
| Supervisors: | Prof. Hanspeter Mössenböck<br>DI Raphael Mosaner |
| Start: | 9. March 2021 |

Graal [1] is a just-in-time compiler for the JVM platform that is itself written in Java. Due to the modular design of Graal the addition of novel optimizations to the existing compiler is easy and straightforward. However, tuning such a compiler optimization for a non-trivial use case is a complex and error-prone task typically done manually by experienced engineers.

In an ongoing work [2], efforts are made to apply machine learning to help improving optimizations in a dynamic compiler. While machine learning has already been used in static (ahead-of-time) compilers [3], it remains to be successfully introduced to dynamic compilers such as Graal [1].

In order to support this ongoing work, the thesis should give a qualitative assessment of feature data and success metrics extracted for particular predictive tasks in a dynamic compiler [2]. Initially, the feature data has to be analyzed and pre-processed, including outlier filtering and looking for inconsistencies. The data should then be used to train different kinds of predictive models and to analyze why specific features or datapoints may or may not be important for a designated optimization decision.

The goals of this thesis are:

- Analyze and assess feature data
- Pre-process the data for the machine learning task
  - Filter outliers and features
  - Look for noisy or faulty data
- Create and train different machine learning models to evaluate their predictive power with respect to the underlying data

The progress of the thesis should be discussed on regular basis with the supervisors. A time schedule with milestones must be presented 3 weeks after the work starts. This schedule will be continuously refined as the work progresses. The final thesis should be submitted not later than 8. September 2021.

[1] http://openjdk.java.net/projects/graal

[2] R. Mosaner, “Machine learning to ease understanding of data driven compiler optimizations”. SPLASH Companion 2020. ACM, pp. 4–6, 2020

[3] Z. Wang and M. O'Boyle, “Machine learning in compiler optimization”. Proceedings of the IEEE, vol. 106, no. 11, pp.1879–1901, 2018

**JOHANNES KEPLER**
**UNIVERSITÄT LINZ**
Altenberger Straße 69
4040 Linz, Österreich
www.jku.at
DVR 0093696