Statement of Conjoint Work

The thesis “Improving Project Health using Machine Learning” by Pârțachi Profir-Petru, the undersigned, is composed of three main papers which were written in collaborations with others. For each paper, I include the list of authors herein and detail exactly my contributions. All authors have contributed significantly to the writing of the papers; therefore, I detail only non-writing work herein. Additionally, all work was done under the careful supervision of Earl Barr, his insight has shaped all my work and our regular meeting were often used to checkpoint and plan out my research tasks.

Aide-mémoire: Improving a Project’s Collective Memory via Pull Request-Issue Links

Authors: Profir-Petru Pârțachi, David R. White, Earl T. Barr.

I have implemented the proposed PR-Issue linker, performed the exploratory data analysis and feature engineering as well as wrote the evaluation and result analysis scripts. The experimental and EDA designs were done in close collaboration with David White who made sure I perform these correctly and guided me patiently.

Flexeme: Untangling Commits Using Lexical Flows

Authors: Profir-Petru Pârțachi, Santanu Kumar Dash, Miltose Allamanis, Earl T. Barr

I have implemented in full the construction of our new structure, though the idea of the structure was from Earl Barr, and the details of the specification were worked in close collaboration. I also implemented the necessary methods to construct the dataset, reproduced previous work in the area and implemented our proposed untangling algorithms as well as their evaluation on the constructed dataset. The RefiNym implementation was provided by Santanu Dash who helped me integrate it with Flexeme. The original PDG extraction implementation for C# code was provided by Miltose Allamanis. Santanu Dash and I performed the manual evaluations.

POSIT: Simultaneously Tagging Natural and Programming Languages

Authors: Profir-Petru Pârțachi, Santanu Kumar Dash, Christoph Treude, Earl T. Barr

I have implemented the preprocessing scripts, the Neural Network that realises POSIT, the adaptation of the previous State-of-the-art to our problem, and the necessary evaluation scripts. The Code Comments corpus was provided in by Santanu Dash. The original implementation of TaskNav was provided by Christoph Treude. The informal proof of context-sensitivity of mixed-text was worked on in close collaboration with Earl Barr without whom the proof would have not been finished in a timely manner. Manual evaluation of POSIT was done together with Santanu Dash, while manual evaluations of TaskNav augmented with POSIT were done together with Christoph Treude.

Date: 01/07/2020 Name: Profir-Petru Pârțachi Signature:

Date: 01/07/2020 Supervisor Name: Earl T. Barr Supervisor Signature: