Evaluating Software Developers' Code Comprehension Leveraging Low-Cost Psychophysiological Measurements

Master's Thesis Proposal

Have you ever read some really hard to understand program code? Have you imagined how much easier it would have been to understand such program code if it had been written in a more comprehensible way?

Code comprehension i.e. interpreting code is one of the major tasks done by software developers throughout their professional life [1]. The comprehensibility of code influences software quality [2], work performance [3], and ultimately mental health [4]. You may have experienced this yourself in programming-intensive projects and courses like SWA (software architecture).

There exist several approaches to empirically study how understandable a program is. Traditionally, different methods such as think-aloud protocols are applied [5]. However, more direct methods using physiological sensors that allow for psychophysiological measurements are increasingly often used in the field and promise novel insights. In your master thesis, you will leverage this potential to study the field of code comprehension. An overview of relevant concepts in this regard is given in Figure 1.

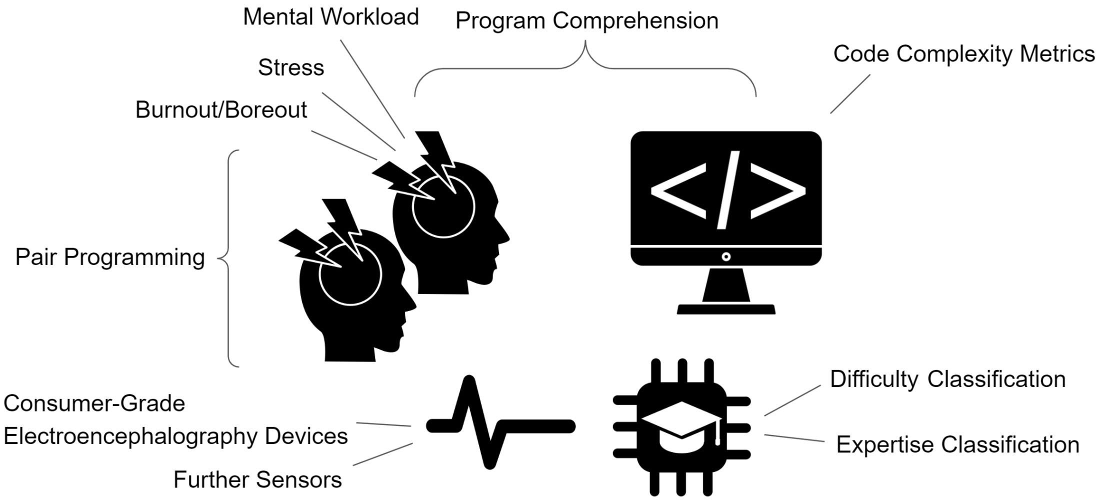

Figure 1: An overview of relevant concepts that can play a role in your master thesis in the field of evaluating code comprehension using psychophysiological measurements.

Recently, code comprehension (also program comprehension) is being studied by researchers using noninvasive neuroimaging techniques. Several such techniques have been developed in the past and allowed to study the reaction of the human brain to stimuli. Three of the most popular ones are Functional Magnetic Resonance Imaging (fMRI), Functional Near-Infrared Spectroscopy (fNIRS), and Electroencephalography (EEG). EEG measures the electrical activity of the brain with the help of electrodes that are positioned at certain locations of the scalp. EEG has several advantages in comparison to the other two mentioned techniques that are, for example, listed by Kosti et al. in [6]. One of the most important advantages is that recently, more and more wearable, low-cost devices capable of measuring EEG data conveniently are entering the market.

Possible advantages of these devices in comparison to those that are used for medical gold standard EEG measurements are an easier use and shorter preparation time, or the possibility to obtain EEG measurements more often, over longer periods, or in more diverse and realistic environments. Challenges exist for example regarding the quality of the data provided by such wearable devices. Although often marketed for use in realistic environments, they perform measurements that are quite susceptible to movements [6]. Therefore, such realistic environments should rather require little movement. This is the case for some of the major tasks conducted in software engineering such as code comprehension.

Depending on your interests, you will work on a concrete current challenge in the field such as: researching possibilities to link code artifacts with brain imaging data and with further datasets, such as eye-tracking or mouse position data; evaluating and extending tool support for multimodal experiments; or evaluating the relationship of brain imaging measurements and measurements of other, less intrusive wearable sensors [7, 8].

You will have access to recent physiological sensors with a focus on wearable EEG devices.

Requirements and Expectations

While you will gather new insights throughout the process of writing your thesis and will be supported in doing so, there are some requirements to ensure that a great master thesis can be achieved. It would be good if some of the following points apply to you. You have:

- Experience with data analysis/machine learning in Python or R (you already applied classification and/or clustering techniques and know that accuracy is not always the best metric to choose when evaluating results)
- A background in software development (university courses or having work experience)
- An interest in body sensors with a focus on wearable EEG and eye-tracking devices
- An interest in data collection and neural time series data processing and analysis
- An interest in neuroscientific theories and concepts
- An interest in publishing your work for the research community to build upon it

Contact

If you are interested or would like to get more information please contact:

Fabian Stolp
Fabian.Stolp@hpi.de
+49 (0) 331 5509-3453
G-2.1.11, Campus 3

Bert Arnrich
Bert.Arnrich@hpi.de
+49 (0) 331 5509-4851
G-2.1.14, Campus 3

Further Resources

Introduction to the brain and possibilities to measure it (Video):
`https://www.tele-task.de/lecture/video/8815/`
Introduction to neuroscience (Video and Text):
`https://www.emotiv.com/neuroscience-guide/`
Introduction to EEG (Video and Text):
`https://www.emotiv.com/eeg-guide/`
Interruptibility and psychophysiological measurements (Video and Text):
`https://dl.acm.org/doi/10.1145/2702123.2702593`
Code reviews and psychophysiological measurements (Video and Text):
`https://dl.acm.org/doi/10.1145/3368089.3409681`
Neural signal processing and analysis (Video):
`https://www.udemy.com/course/solved-challenges-ants/`

References

[1] Janet Siegmund and Jana Schumann. "Confounding parameters on program comprehension: a literature survey". In: *Empirical Software Engineering* 20.4 (2015), pp. 1159–1192.

[2] Andrea Schankin et al. "Descriptive compound identifier names improve source code comprehension". In: *2018 IEEE/ACM 26th International Conference on Program Comprehension (ICPC)*. IEEE. 2018, pp. 31–40.

[3] Janet Siegmund et al. "Toward measuring program comprehension with functional magnetic resonance imaging". In: *Proceedings of the ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering*. 2012, pp. 1–4.

[4] Ramyashilpa D Nayak. "Anxiety and mental health of software professionals and mechanical professionals". In: *International Journal of Humanities and Social Science Invention* 3.2 (2014), pp. 52–56.

[5] Janet Siegmund et al. "Experience from measuring program comprehension-toward a general framework". In: *Software Engineering 2013* (2013), pp. 239–257.

[6] Makrina Viola Kosti et al. "Towards an affordable brain computer interface for the assessment of programmers' mental workload". In: *International Journal of Human-Computer Studies* 115 (2018), pp. 52–66.

[7] Barbara Weber, Thomas Fischer, and René Riedl. "Brain and autonomic nervous system activity measurement in software engineering: A systematic literature review". In: *Journal of Systems and Software* 178 (2021), p. 110946.

[8] Júlio Medeiros et al. "Can EEG Be Adopted as a Neuroscience Reference for Assessing Software Programmers' Cognitive Load?" In: *Sensors* 21.7 (2021), p. 2338.