Academic Statement of Purpose
Danniell Hu
Computer Science PhD, Rackham
########

I never thought my worst enemy would be someone's hair color. When I joined my undergraduate software engineering research lab, the last thing I expected was that I'd become a part-time hairdresser. Yet, there I was, wielding crochet hooks to braid hair during our neuroimaging experiment on program comprehension. Amidst challenges between brain activation signal strength and participant comfort, I stumbled upon an unconventional solution—hair braiding. This experience not only showcases my adaptability but also highlights my **creativity, communication, and motivation in the face of difficulty**—traits that have consistently propelled me forward in my research journey.

My research journey began at the University of Michigan. I sought out a position as an undergraduate researcher in professor Westley Weimer's lab, where my research efforts resulted in a **first author submission to the peer reviewed Transactions in Software Engineering Journal, a top-tier venue in software engineering.**

In addition to my background in undergraduate research, my aspiration to pursue a PhD in computer science is further fueled by my current role as an R&D software engineer at Stryker, a medical device company. At Stryker, I develop embedded systems applications for medical equipment (hospital beds, cots, mattresses, etc.). Beyond the developmental aspect of R&D, I actively participate in VOC (voice of customer) visits to hospitals nationwide for research purposes. During these visits, I immerse myself in the workflow of nurses to gain firsthand insights into how they interact with our products. This allows me to identify their pain points and pinpoint issues that need resolution to streamline their workday. We then bring these challenges back to the lab, where we brainstorm solutions, aiming to enhance healthcare through product innovation. However, I realize that there are underlying issues that are root causes for these hardships. For instance, while the creation of a hospital bed automating aspects of a nurse's workflow can alleviate some challenges, the crux of their struggle often lies in the overwhelming patient-to-caregiver ratio. Industry grapples with limitations in addressing these foundational problems due to a lack of capacity and resources for comprehensive research. I believe that investigating the deeper reasons behind a problem is an important component of solving it, going beyond surface-level solutions to create lasting impacts—thus influencing my interest to pursue research.

Outside of my work at Stryker, my research has primarily involved the study of debugging and the cognition behind it. Debugging involves the identification and resolution of code defects. Despite debugging's importance, prevailing debugging techniques lack efficiency, efficacy, and scalability. Additionally, the cognitive processes of debugging remain poorly understood. Debugging requires logical reasoning, pattern recognition, and problem-solving skills, but what happens in the brain while identifying and fixing software issues is still unknown. If we were to close this knowledge gap, we could develop more tools and techniques to improve the debugging process.

My **creativity** was essential when pursuing this research topic. Firstly, while debugging is a holistic, interactive process, prior research has predominantly focused on dissecting individual aspects of debugging. Before my work, there was no investigation of the interactive debugging process as a whole at the cognitive level. I hypothesized that the interactive debugging process could be decomposed into discrete segments that correspond to different patterns of activity in the brain, otherwise known as distinct neural stages. To investigate my hypothesis, I co-designed an experiment that employs functional near-infrared spectroscopy (fNIRS), a lightweight technique used to assess brain activity by measuring blood oxygen levels. We realized that this approach allowed for non-invasive measurements of participant cognitive processes during debugging tasks. While other medical imaging techniques have been used in previous studies, we opted for fNIRS both due to its ecological validity and research indicating its comparability to other methods.

I constructed a set of 16 distinct stimuli, each consisting of a problem description, a buggy solution, and a set of test cases. To enable a controlled neuroimaging experiment, I strategically selected and modified interview questions from LeetCode to construct stimuli mirroring real-world debugging challenges. Although

previous medical imaging studies had used static stimuli in non-interactive editors, I advocated for the adoption of VSCode as the code editor due to its familiarity among participants, high ecological validity, and novelty in medical imaging studies. Additionally, I adapted existing extensions to capture behavioral data (keystrokes, time spent on each file, etc.), which were used in over 200 program snippets across participants.

My **effective communication** also allowed us to advance our research goals. My exploration of existing literature on software engineering maintenance revealed an emphasis on reading skills over writing skills. Recognizing this trend, I postulated that participants with diverse levels of reading ability may demonstrate variations in both cognitive and behavioral aspects of debugging. However, I had minimal prior experience with the cognition of reading or reading disabilities. I reached out to a group of psychology researchers specializing in language and reading proficiency, which led us to incorporate the psychological concept of *morphology* into our experiment. To aid in communication with our psychology collaborators, my team and I created a proposal for an experimental design aimed at exploring the cognition of the debugging process, and at the possible disparities between individuals with and without reading disabilities. We then presented our proposal to the psychology group, described our hypotheses, and sought their advice. Throughout the project, our communication with the psychology group remained consistent, fostering a positive and evolving relationship that continues to thrive—notably, a member of the psychology group has become a co-author on my submitted paper.

Communication was also fundamental when I ran pilot studies, recruited participants, and ran experimental sessions. During our first pilot studies, we encountered challenges with the fNIRS cap. fNIRS technology measures brain activity through nodes on a cap fixed to a participant's head. It was challenging to strike a balance between signal strength and participant comfort since tighter caps yielded better signal but could be more painful. This equilibrium was especially demanding for those with darker or thicker hair, as their hair type necessitated a tighter cap fit and exacerbated their discomfort. I took the reins and independently brainstormed solutions to this problem while the rest of my team and advisor were in Australia for a conference. Ultimately, I discovered that putting participants' hair into ponytails increased both access to the scalp and signal quality, and could be easily undone after the experiment. Additionally, the use of crochet hooks to braid hair away from troublesome nodes moved signal quality from poor to excellent. Following testing of various approaches, I devised a hair braiding and signal retention protocol. I conveyed this protocol to the team through a combination of hands-on demonstrations and detailed written guidelines. My initiative and clear communication with the team resulted in successful data collection for the remaining 30 participants.

I have also found that **motivation** is an indispensable trait in research. In my debugging project, this was particularly important during the data analysis phase. Following data collection, I faced additional challenges in the analysis process, particularly due to my limited experience with the required statistical techniques. I was determined to become more familiar with statistical analysis and delved into learning by reading various papers, experimenting with previous datasets to replicate findings, and seeking one-on-one guidance on statistical concepts. I performed an ANCOVA analysis multiple times incorrectly before achieving accuracy. I also had the task of translating pre-formulated research questions into measurable variables, where selecting the most appropriate variables for analysis upfront was a nuanced process that offered many different avenues of inquiry. An additional challenge stemmed from the novelty of investigating reading disabilities in the context of software engineering, which gave the research an exploratory component. To discern emerging trends from our data and avoid coming to false conclusions, I implemented corrections for multiple comparisons using the Benjamini–Hochberg method. Upon completing the data analysis, I found myself more knowledgeable about statistical techniques, albeit with more to learn. This understanding was a direct result of my enduring motivation and perseverance throughout the process. I hope to continue honing my statistical analysis skills as a PhD student.

As a PhD student in computer science, I am interested in continuing research in the intersections of **software engineering, systems, HCI, education, and healthcare**. After graduation, I am interested in becoming a professor or an industrial researcher. My previous experiences in research, industry, interdisciplinary subjects, and proficiency in qualitative and quantitative methods positions me as a well-prepared candidate. I am excited to bring my creativity, effective communication skills, and unwavering motivation to contribute meaningfully to the academic community. My excitement is particularly pronounced in the prospect of collaborating with professors **Elizabeth Bondi-Kelly** (AI for social impact, human collaboration), **Maggie Makar** (ML, health care)**, Rada Mihalcea** (NLP, social applications), **Xinyu Wang** (PL, SE), and **Max New** (PL). I believe my interests align well with these professors and feel that I would be a good fit in the Computer Science department at the University of Michigan.