Understanding Complex Systems: The Need for an Open Source Based Homework Assignment in CS 3140 at UVA

A Technical Report submitted to the Department of Computer Science

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Johnathan Eftink

Spring, 2025

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Rosanne Vrugtman, Department of Computer Science

Understanding Complex Systems: The Need for an Open Source Based Homework Assignment in CS 3140 at UVA

CS4991 Capstone Report, 2024

Johnathan Eftink
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia USA
cvdwa@virginia.edu

ABSTRACT

In the real world, most computer-based projects are not a few java classes, but a large collection of interlocking parts that require a deep knowledge of the code base, which CS students at UVA frequently lack. However, using an open source project will enable students to have this experience. CS 3140 currently has six homework assignments that range from starting a project from scratch to working on a project that is only missing some information. Adding an additional homework with an even larger code base will teach the students how to understand which portions of code are useful or unnecessary for them to read when time is crucial. Since students will have learned how to discern useful code and developed skills that currently only come from real world experience, they will be able to walk into professional work with an ability to be onboarded faster. Polling the students on whether this skill distinguished them from other interns/professionals will allow the department to refine the homework over time.

1. INTRODUCTION

In the fall semester of 2022, UVA started rolling out its new Computer Science curriculum. With this new curriculum came many new classes, such as CS 3140; Software Development Essentials. The goal of this new class was to give students a starting point in their software development education that later classes would be able to expand upon. As a software development course, it focused on larger projects that would take about two weeks to complete.

The curriculum for CS 3140 was primarily made by one professor, which also means that the starter code was made by one person. While it is beneficial that they created them, as it allows them to know every bit of the code, it also means that the students only learn what code developed by one person looks like. Every programmer has a unique set of design choices that comes with them. Exposing the students to code developed by other people would help them expand their understanding of common software development practices.

An open source-related homework would allow the students to see how other developers code, while also letting them learn more about the world of open source code. The homework should be limited however, since these students would not have the coding abilities necessary to make changes to an open source project. Overall, having a assignment analyze an open source project would allow students to understand open source software better and allow them to see how other software developers work in the real world.

2. RELATED WORKS

Many researchers have investigated how to use open source projects in the classroom and tried to find their advantages. One of these,

Conde, et al. (2021), noticed that many new graduates did not have the skills that much of the professional world required. They also found that working with open source projects as a student can help elevate them to industry expectations. These led Conde, et al. to create a blueprint for how to include open source projects in one's curriculum.

Research by Salerno, et al. (2023), discussed how participating in open source work improves students' belief in their own programming ability. This is important as many students might not have the confidence in their code that is needed to perform well in the professional world. Salerno, et al. (2023) also recognized the potential difficulties likely to occur if an open source project were deployed. Lack of documentation, technical hurdles, and issues related to tools can slow down the progress of the class and cause issues for the students.

Other important difficulties would need to be understood, as well. Morgan and Jensen (2014) found that at first a few students in a course would struggle with completing meaningful contributions. However, in the second year of the course, students were able to contribute more to the projects and student evaluations even indicated that the students found it as one of their most worhtwhile courses. This shows that getting students to contribute to an open source will likely prompt issues at first, but once the difficulties are better understood, the changes pay off.

3. PROPOSAL DESIGN

One of the key parts of CS 3140 is that students must learn how to create a Unified Modeling Language (UML) diagram. UML diagrams can be used to describe interactions between classes or even to suggest a design. In CS 3140, it is used to show the interactions/relationships between classes. This helps students understand UML and similar diagrams that they might experience in the corporate world. Currently, the homework assignment that has the students work on a UML diagram uses a project that they are already working on, so it is limited in scope.

A UML diagram-based homework assignment is a great place to use open source code, as students will be able to work on code that has a large scope and is able to have a large enough diagram. It is also important to note that is also highly scalable in scope, so if the homework needs to be harder or easier, it can be easily scaled. Scalable in scope means the assignment can either have students look at a small subset of all files or look at large directories of files.

3.1 Homework Design

The homework will use a Java-based open source project. With the current UML diagram assignment, students are given a collection of fewer than twelve Java classes for which they create a UML diagram. This diagram maps how classes interact with each other, and any cardinality for these interactions. For instance, if there is a student class that contains a list of courses class, then there is a line drawn between the student and courses, that indicates the presence of a list.

As described, the homework assignment can use an open source project to increase the students' understanding of open source projects and UML diagrams. A complex open source project for which students would be able to write a partial UML diagram might be "Open LaTeX Studio" (Brudzinski, 2018). Many students will end up having to write in LaTeX for later courses such as CS 3120 so, while not immediately relevant to the student, the course content will become more relatable.

The "Open LaTeX Studio" open source project contains 54 Java files that the assignment could include (Brudzinski, 2018). Using all 54 would be much more than is expected for CS 3140 students, so a more doable exercise would be to work on a smaller set of files (within the /editor/src/main/java/latexstudio/editor/

directory of the project). If the new assignment is seen as too easy, more directories can be added, increasing the difficulty.

The idea of this homework assignment would be to act as a supplement to help explain diagrams that would exist in the corporate world. It would not completely replace the existing homework that includes the creation of a UML diagram, but, instead, come as a smaller assignment after creating a first UML diagram. Since most real-world systems are significantly more complex than what is seen in class, this larger UML diagram can help students see how larger systems interact. At the same time, it will provide practice in diagram-building and understanding and comprehending large systems.

3.2 Learning Objectives

One of the objectives of CS 3140 is to learn how to design functional object-oriented systems. This is aided by UML diagrams which allow the student to gain a deeper understanding of how objects interact with each other. The class currently does not have a project as large as the "Open LaTeX Studio" project, so having students work with a portion of the project will allow them to understand how a multitude of objects can interact with each other. Having to write down class interactions for a larger project also helps with understanding object-oriented systems because they are exposed to more code that is object-oriented. Even though they are not writing more object-oriented code, they still read and attempt to understand parts of it in order to complete the assignment.

Another learning objective of CS 3140 is to understand object-oriented design patterns. Creating a UML diagram for a larger project allows students to see how certain design patterns may be used. Students are easily able to see the "single-responsibility principle" (the idea that a java class should have one purpose) in the DbxAutoSync.java file since the class only has one responsibility and one reason to change (Brudzinski, 2018; Feathers, 2020, p.246). These design patterns are visible throughout the LaTeX project and allow students to see what design principles look like in an application that is usable by the public.

4. ANTICIPATED RESULTS

Conde, et al. (2021) found that working on open source projects allowed students to do a better job of meeting employer expectations. While this new homework does not have the students directly editing an open source project, it still has them looking at open source code and understanding it, which is still a worthwhile skill to employers. Since it is only a UML diagram, many of the issues related to open source projects in schools will not occur. As seen by Hu, et al. (2018), many students ran into issues like dealing with merging in git.

There will be some issues likely as seen by Morgan and Jensen (2014). For instance, the larger breadth of the project will likely cause some issues for students. Most of the projects in the class only deal with modifying a few java classes, so having to potentially look at dozens of classes can cause students to feel overwhelmed. This will likely lead to changes being needed in the number of classes used in the homework, if the UML diagram becomes too complex.

5. CONCLUSION

Incorporating open source projects into the UVA CS 3140 curriculum, particularly through an assignment focused on creating a UML diagram, offers significant benefits to students. By engaging with a larger, real-world codebase, students gain practical experience in understanding and analyzing complex systems. Additionally, working with open source code helps students stand out to potential employers, as it demonstrates their ability to navigate and contribute to existing codebases. The use of a large open source project in homework design also allows for scalable difficulty, making it easier to adjust

the challenge level in future iterations of the course. Using an open source project not only enriches the learning experience but also bridges the gap between academic knowledge and real-world application, better preparing students for careers in software development.

6. FUTURE WORK

To use this new proposed homework, the open source project would first have to be finalized. I suggested "Open LaTeX Studio" since it is relatively easy to understand and within the objectives of CS 3140. If it is found that there is a better open source project, (e.g. more understandable or more complex) then the homework can be adjusted. After the semester, polling can be done to understand how students feel about using an open source project for their homework and its difficulty. Students would also be polled after finishing an internship or starting a full-time job. These polls would be to see if the students found working with an open source project to be beneficial in their job search and start of their career. Their feedback would allow for adjustments.

These homework adjustments could relate to the difficulty. Adjusting the difficulty would only require an adjusted set of homework instructions and a new grading key. The homework could also be adjusted to include new elements such as having the students read and understand the documentation of the open source project. This would give students more experience with open source projects and give them more to mention to a recruiter. This shows that the homework can easily be changed to fit the needs of the students in the ever-changing job market.

REFERENCES

Brudzinski, S. 2018. *Open LaTeX Studio*. Release 1.5, https://github.com/sebbrudzinski/Open-LaTeX-Studio

Conde, J., López-Pernas, S., Pozo, A., Munoz-Arcentales, A., Huecas, G., & Alonso, Á. (2021). Bridging the gap between academia and industry through students' contributions to the FIWARE European Open-Source Initiative: A pilot study. *Electronics*, *10*(13), 1523. https://doi.org/10.3390/electronics10131523

Feathers, M. (2020). *Working effectively with legacy code.* Pearson. https://www.oreilly.com/library/view/working-effectively-with/0131177052

Hu, ZW., Song, Y., Gehringer, EF. (2018). Open-Source Software in Class: Students' Common Mistakes. *Proceedings of the 40th IEEE/ACM International Conference on Software Engineering—Software Engineering Education and Training (ICSE-SEET). (40-48) https://doi.org/10.1145/3183377.3183394*

Morgan, B., Jensen, C. (2014). Lessons Learned from Teaching Open Source Software Development. Corral, L, Sillitti, A, Vlasenko, J, Wasserman, Al. *Proceedings of the 10th IFIP WG 2.13 International Conference on Open Source Systems (OSS). (133-142) Springer-Verlag Berlin https://link.springer.com/chapter/10.1007/978-3-642-55128-4_18*

Salerno, L., de França Tonhão, S., Steinmacher, I., & Treude, C. (2023). Barriers and self-efficacy: A large-scale study on the impact of OSS courses on student perceptions. *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, (320–326). https://doi.org/10.1145/3587102.3588789