

Static Analysis for Software Engineering (version 3)

ECE750-T5 (LEC 003), Winter 2022

Patrick Lam

Update, January 2022. Stay safe, and hopefully we can meet in person by the end of January.

Still “interesting” times. The University would like to project the “everything is back to normal” vibe. Everything isn’t back to normal, but hopefully we’re getting there (Jan 2022: oops). I’ve been out of country for the pandemic, so please bear with me as I get used to being back in Canada. I’m here to support each of you.

Logistics

First, let’s talk about the mechanics.

Class meetings. We are scheduled to meet 4pm-5:20pm on Tuesdays and Wednesdays in EIT 3141.

- *Tuesdays* will be optional video watching and (positive, supportive) discussion about presentation style. The video watching is required, but watching it in EIT 3141 with me on Tuesday afternoon is not. I do especially recommend that the week’s presenters come on Tuesdays for constructive feedback.
- *Wednesdays* will be mandatory discussion about paper content. I am expecting you to be prepared to discuss the week’s papers on Wednesday.

Every week, we will discuss 2 papers, which you’ve signed up to present. Last year was online, and students prepared recorded presentations. The feedback from students was that they appreciated the opportunity to watch the presentations at their own pace, so we’ll do that again.

Please upload and share your presentations with the class for Tuesday 4pm; we will then all view the presentations before the Wednesday meeting, either in EIT 3141 or on your own time.

On Wednesdays, we’ll talk about each of the papers, including positives and negatives, and then we’ll compare and contrast the papers.

If we can’t meet in person, we will use this Zoom link:

<https://uwaterloo.zoom.us/j/91335036247?pwd=eksxYzdmdmQzRTRwUTZDd3pFUlJjZz09>

I am also available for individual and group meetings. Just let me know and we’ll work out a time that works for all participants. It has become fairly clear that walking outside is fairly low risk, especially masked, and I encourage you to set up a meeting with me to walk and talk about this course, your project, or anything you’d like to talk about.

COVID backup plans. If the epidemiological situation changes for the worse and the university goes distance again, we will cancel the Tuesday group viewings and we will discuss papers online on Wednesdays. (It’s worse to discuss papers online, but getting COVID is even worse.)

If you get COVID, you can submit a Verification of Illness and we’ll shuffle things around to make things work.

If I get COVID, you will still upload videos and have class discussions without me. (I think we all hope I don’t get COVID.)

Content

Code review is a key technique for ensuring software quality. However, human reviewers have limitations: limited time, limited attention spans, and a limited understanding of the implications of a software modification (due to interactions between parts of a potentially vast codebase).

Computers can successfully carry out many tasks at which humans fail. A goal of my research is to find classes of properties which are amenable to automatic verification, using program analysis techniques (which trace their roots to optimizing compilers). Many challenges exist. The most notorious problems include the undecidability of the halting problem (which we overcome using approximations) and the need to deal with unknown input values. More recently, large codebases and plugin-based software architectures pose additional challenges to static analysis.

Nevertheless, a number of program analysis techniques for software engineering have been proposed and even deployed in commercial development environments. Facebook's Infer tool, Coverity, and Microsoft's Static Driver Verifier are three of the most successful examples of static analysis in practice.

In this seminar course, we will first briefly explore the strengths and limitations of program analysis techniques. These techniques traditionally come from the compiler research community. The bulk of this course, however, will consist of a discussion of current research papers in the field of software verification using program analysis techniques.

Objectives. After this course, you will:

- understand the strengths and limitations of static and dynamic analysis techniques;
- be familiar with the research literature on static and dynamic analysis for software verification.
- have carried out a small research project implementing a program analysis and evaluating its efficacy at solving a software engineering problem.

General Information

Course Web Page: <https://patricklam.ca/sase-2022>

Instructor:

Prof. Patrick Lam
Office Hours: By appointment
Email: patrick.lam@uwaterloo.ca

Course Description

Overview of techniques used in static and dynamic analysis, including dataflow analysis, type systems, and pointer analysis. Typestate properties. Applications to software engineering (notably concurrency and security).

Here is a brief week-by-week schedule:

- 1 Organization
- 2 Empirical Studies I
- 3 Empirical Studies II
- 4 Initialization Analyses
- 5 Dataflow Analysis
- 6 Pointer Analysis
- 7 Fuzzing
- 8 Testing
- 9 Repair
- 10 Domain-Specific Languages
- 11 Libraries

Reference Material

The reference material for this course consists of the research papers that we'll discuss every week. Most of the papers have been posted on the authors' webpages, and I've included those links when possible. You may need to use the campus proxy to download the ACM links to a couple of papers.

Evaluation

You will present one paper to the class and complete a course project, which includes a short presentation on the last day of class. There will also be a take-home final examination where you will demonstrate your understanding of the papers we've discussed as well as self-reflection on your course project.

Course project	45%	due last day of class
Project presentation	5%	during last class
Paper presentation	10%	throughout term
Participation in class discussions	10%	throughout term
Final exam	30%	final exam period

I'm going to mark participation based on my perception of whether you've engaged with the papers under discussion each week. You don't have to pipe in on every paper, but I expect you to contribute something substantial to at least one paper every two weeks. A contribution can be either a thoughtful question or a remark about a paper, and can be either live or on the discussion forum in the week of the paper presentation. I may also ask you questions.

Presentation. You will present one of the papers that we are discussing in this class by signing up on the Google Doc that I'll share. Be prepared to answer questions on the papers.

I will provide submarks for both delivery and content of your paper and project presentations, and I'll send you timely feedback and suggestions on how to improve your presentations.

"Discussing" means that you need to be actively involved in the discussions of the papers. Typically classes like this one are small, so each of your contributions is important.

Course project. The course project gives you an opportunity to work on a particular application of program analysis techniques to software engineering issues. My default position is going to be individual projects, but we'll see how many people sign up for the course and whether you have good reasons to have pairs.

I expect that most projects will include an implementation component. To help you stay on track, I will expect a project proposal by the third week of class and a mid-term project update (1 page) by the end of week 8. (Start early!) I will post a list of suggested projects on the webpage. Please come see me to talk about possible projects!

Final examination. The open-notes final examination will ask you to synthesize the information you've seen throughout the semester. Potential questions include summarizing key points from selected papers, or sketching out how the techniques from one paper might apply to the problems that another paper addresses, as well as some self-reflection on your project.

Lateness. This is a graduate seminar, so I will offer some flexibility on due dates. However, this flexibility is often not in your best interest. You must consult with me before you hand in something late; otherwise, you will lose 5% on the project mark for a late project proposal or mid-term project report. The default mark for a late project submission is 0.

Fair Contingencies for Emergency Remote Teaching. We are facing unusual and challenging times. The course outline presents the instructor's intentions for course assessments, their weights, and due dates in Winter 2022. As best as possible, we will keep to the specified assessments, weights, and dates. To provide contingency for unforeseen circumstances, the instructor reserves the right to modify course topics and/or assessments and/or weight and/or deadlines with due and fair notice to students. In the event of such challenges, the instructor will work with the Department/Faculty to find reasonable and fair solutions that respect rights and workloads of students, staff, and faculty.

Required inclusions

Academic Integrity: In order to maintain a culture of academic integrity, members of the University of Waterloo community are expected to promote honesty, trust, fairness, respect and responsibility. [Check www.uwaterloo.ca/academicintegrity/ for more information.]

Grievance: A student who believes that a decision affecting some aspect of his/her university life has been unfair or unreasonable may have grounds for initiating a grievance. Read Policy 70, Student Petitions and Grievances, Section 4, www.adm.uwaterloo.ca/infosec/Policies/policy70.htm. When in doubt please be certain to contact the department's administrative assistant who will provide further assistance.

Discipline: A student is expected to know what constitutes academic integrity [check www.uwaterloo.ca/academicintegrity/] to avoid committing an academic offence, and to take responsibility for his/her actions. A student who is unsure whether an action constitutes an offence, or who needs help in learning how to avoid offences (e.g., plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course instructor, academic advisor, or the undergraduate Associate Dean. For information on categories of offences and types of penalties, students should refer to Policy 71, Student Discipline, www.adm.uwaterloo.ca/infosec/Policies/policy71.htm. For typical penalties check Guidelines for the Assessment of Penalties, www.adm.uwaterloo.ca/infosec/guidelines/penaltyguidelines.htm.

Appeals: A decision made or penalty imposed under Policy 70 (Student Petitions and Grievances) (other than a petition) or Policy 71 (Student Discipline) may be appealed if there is a ground. A student who believes he/she has a ground for an appeal should refer to Policy 72 (Student Appeals) www.adm.uwaterloo.ca/infosec/Policies/policy72.htm. Note for Students with Disabilities: The Office for Persons with Disabilities (OPD), located in Needles Hall, Room 1132, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with the OPD at the beginning of each academic term.