Teaching Statement

Linh Thi Xuan Phan
http://www.cis.upenn.edu/~linhphan/

I believe that teaching is an essential part of a meaningful academic career – if research is to have any significant impact on society, it must be shared with people at all levels, from students to colleagues, from fellow researchers to engineers. For me, teaching is a tremendously rewarding experience: not only does it allow me to share with students the knowledge and problem-solving skills they need to improve the world around them, but it also helps me learn how to become a better researcher and educator. Above all, *teaching is about people*, and I truly believe that only with a caring heart and a desire to make a difference, one can become a great teacher. I have always done my best to help students and to show them just that. I am thrilled to be a recipient of the **Lindback Award for Distinguished Teaching**, the University's highest teaching honor that is given to only eight of Penn's nearly 5,000 faculty members each year.

1 Course development: CIS 5050 Software Systems

I have taught CIS 5050 ("Software Systems") nearly every semester since I joined the tenure track. CIS 5050 is a core systems course in the MSE and PhD programs. Although it is a graduate-level course, it attracts a substantial number of undergraduate students. In Fall 2024, about half of the class this semester are undergraduates.

CIS 5050 teaches fundamental concepts of distributed systems and the design principles for building large-scale computational systems. The class covers key building blocks – such as synchronization primitives, group communication protocols, and replication techniques – that form the foundation of modern distributed systems, such as cloud-computing platforms or the Internet. Besides foundational concepts, it also covers real-world case studies of distributed systems, such as GFS, MapReduce, Spark, and Dynamo.

When I took over the course, I redesigned all the materials: I developed a fresh set of slides, I designed a completely new set of homework assignments from scratch, and I created a new final project with a more ambitious scope. The first assignment is meant to familiarize the students with basic UNIX concepts and to give them some hands-on experience with synchronization and concurrency; this is useful because the course attracts a highly diverse group of students – including not only undergraduates, Master's students, and PhD students in CIS, but also students from other departments – and some of them have never programmed in C/C++ before. The second assignment creates a set of email servers (SMTP and POP3) that work with a real e-mail client; this shows the students how a distributed system works under the hood, it gives them experience with socket programming and the client/server paradigm, and it provides opportunity to work with real-world specifications (RFCs). The third assignment asks the students to build a fully-distributed chat system with multiple servers that supports several different consistency models. This exposes the students to complex ordering protocols, as well as to practical issues such as packet loss and network delay, and the challenges of building a reliable system that can tolerate these issues.

The final project is to build a fault-tolerant cloud platform called PennCloud in teams of four. PennCloud is loosely inspired by Google Apps; the required features include a Gmail-like webmail service and a storage service that is somewhat analogous to Google Drive. This brings together many of the concepts that are discussed in class – such as fault-tolerance and consistency – and it exposes the students to the challenges of building a large software system in a team. The homework assignments are intended as building blocks and can be used as part of the project; thus, despite the limited time, the students can get a large system with a lot of functionality that they can be proud of, and that can later become part of their portfolio, e.g., for job interviews. The project includes a detailed proposal, work-in-progress presentations, a demo, and a final report, all of which provide an opportunity to practice important soft skills. (I always personally attend and evaluate all project demos in my class.)

The feedback from the students has been very positive. The course and instructor ratings went up significantly (with an upward trend). Based on Penn Course Review, I was the first to achieve a rating above three for this course – despite the fact that the course is now more challenging, as reflected by the increased difficult level (previously 2.6, now 3.4 out of 4). Since tenure, my instructor and course ratings have been consistently above 3.51 and 3.47 (out of 4.0), respectively.

The course evaluations suggest that the students enjoy the course. Example comments include:

- "Wonderful course. Change nothing. Loved every minute of it"
- "The final project has led me to an offer from Google, thanks this course very much, learned a lot!"
- "Perfect balance of theoretical and practical experience. Leaves you with a feeling of extreme accomplishment when you are able to complete all the assignments and specially the project (which is by far the best project that

I have done). Very hard, leads to a lot of sleepless nights but totally worth it. One of the best courses at Penn"
- "Best CIS instructor. Enthusiastic about her research and teaching. Always stayed after class and answered everyone's questions. Assignment specs were detailed and clear. Piazza questions were answered timely and in great details."
- "The course is amazing. I think it is a course that is fundamentally important in computer science discipline and would strongly recommend everyone to take this course."
- "Awesome, one of the best computer science courses I've taken at Penn!"
- "Professor Phan cares a lot about students getting something out of the course – she is a very enthusiastic, energetic, and approachable instructor."
- "Never seen such a devoted instructor in my life."
- "Hands down the best CS course I have ever taken at Penn."
- "Dr. Linh is THE MOST accessible and responsible professor I've ever had at Penn. She definitely has a passion in teaching and she cares a lot about her students. I feel so lucky to have her as my professor and I really appreciate that she has provided me such a great learning experience."
- "Linh is one of the best profs I have ever had in my life!! If you are a CS major or if you are into systems, definitely consider taking courses with Linh at least once! She is very charismatic and you could tell that she loves what she teaches. She is also super accessible on Edstem and replies to students' posts almost 24-7. Kudos to Linh!"

Even after students have graduated, I frequently received thank you emails from them saying how the course had helped them land jobs or how it has been crucial to their current projects after Penn.

2 CIS 4550/5550: Internet and Web Systems

CIS 4550/5550 is an advanced course that is open to both undergraduate and graduate students. It is a project elective for our undergraduates and a core systems course in our Master's and Ph.D. curriculum. The course focuses on large-scale Internet and Web systems, and it covers many general distributed-systems topics, such as scalability, interoperability, consistency, replication, fault tolerance, and security. An interesting feature of this course is that it is designed around a case study of a modern web system, Google's search engine, and it teaches students all the key elements to build such a system. As such, the class also covers many modern web technologies and web search algorithms such as crawling, indexing, and ranking. Through this course, students gain a comprehensive understanding of how a massive-scale web system really works and hands-on experience in building one.

As with CIS 5050, the course includes a series of individual programming assignments and a final project that builds a Google-style search engine in teams of four. The assignments each develop a basic version of a component that the student will need for the final project, including e.g., a web and application server, a scalable web crawler, a storage server, and an indexer. These assignments, which are built on top of one another, expose students to various aspects of a web system and provide them with hands-on experience in building a large-scale system from the ground up. The final project gives students opportunities to work with real-world (messy) data from the Web and expose them to practical issues associated with the processing such data and deploying their system in a real cloud environment. Through the project, students also learn to develop creative algorithmic and design solutions to address these challenging issues to achieve a high-quality large-scale system.

So far I have taught CIS 4550/5550 once, in Fall 2023. The initial course evaluations were very positive (instructor rating of 3.51 and course rating of 3.39). The students wrote that they enjoyed the course and learned a lot from it ("Amazing and well designed course", "Loved the course! Really learned a lot.", "Absolutely amazing professor! Professor Phan is so passionate and knowledgeable about the material. It really shows how dedicated she is to her students and to the course.", "I think she is the best systems course instructor I have ever met", "Professor Linh was fantastic! She was extremely accessible and was energetic during lecture. She was also super responsive on Ed to answer any questions we had, and her timely and detailed responses were deeply appreciated."). Nevertheless, I identified some aspects of the course that I want to improve further – for instance, the handouts and the automatic test cases. I plan to work on these the next time I teach the course.

3 Co-Director of the Master's in Data Science (DATS) program

I have served as co-director of the Master's program in Data Science (DATS) since 2019. The DATS program is run jointly by the CIS and ESE departments, and it was among the first Master's programs of its kind. The first class we admitted had 25 students; the most recent class is nearly 10 times as large, and DATS has become one of the most

popular and competitive Master's programs at Penn Engineering. As a co-director, I was able to help shape the DATS curriculum, develop ways that help students in their studies and careers, and build a strong community among the students. I participate in admissions and teach two elective courses (CIS 5050 & CIS 4550/5550) for DATS, and I also serve as an academic advisor for many students in the program (e.g., nearly 100 students in 2022-2023).

To foster the DATS community, I founded the DATS Student Mentoring program, in which selected groups of second-year students serve as mentors for first-year, accelerated Master's, transfer, and dual-degree students. The goal of the program is to introduce new students to the structure and environment of DATS, and to enable mentors and mentees to work together, network, develop new skills, and help one another. The program has become quite successful, with 22 students volunteering as mentors last year, and is highly appreciated by our students. It is also frequently mentioned in applications from prospective students.

4 Mentoring

I believe that engaging students in research is a valuable teaching tool. In the past several years, I have actively involved undergraduates and Master's students in research – not only to provide them with experience in developing creative solutions for interesting problems, but also to encourage them to explore their potential and to pursue advanced degrees. Through my class, many students have developed deeper interests in distributed systems, and they have further pursued research projects with me on distributed systems and cyber-physical systems. As an advisor, I guide them towards understanding both the theoretical foundations and the challenges in designing and implementing practical systems, while also teaching them other aspects, such as literature survey and research writing. Some of these projects also led to successful publications; for instance, we turned Tianyang Chen's master's thesis into an RTAS 2018 paper that received the Outstanding Paper Award, and a paper based on Jiaqi Lian's thesis project has been accepted at IROS 2024.

I have been fortunate to be able to advise several talented doctoral students. So far, five of my students have graduated; one of them (Jaewoo Lee) has become a faculty member, and the others have started successful careers in industry. These students had very solid track record – for instance, Meng had five RTAS papers, as well as papers at RTSS, ICCPS, EMSOFT, DAC, CLOUD and RTNS; one of his papers received the Best Paper Award at RTAS 2019 and one was nominated for the best paper award at RTSS 2013. Neeraj Gandhi, who recently graduated, also has many papers at top venues, such as RTAS (3), EuroSys (2), ICRA (2) and IROS; one of his papers won the Best Student Paper Award at RTAS 2019 and one was nominated for the Best Paper Award in Multi-Robot Systems at ICRA 2021. He is also a recipient of the Jonathan M. Smith Fellowship for his systems work at Penn.

5 Diversity and outreach

I believe that computer science educators have a crucial role to play in increasing the enrollment of underrepresented groups in computer science. I am committed to improving the representation of women in computer science and engineering, as well as to helping female students in their studies and research. At Penn, I have also participated in informal panels, sponsored by the Penn Graduate Women in Science and Engineering, to share my experiences with female graduate students, as well as giving talks, e.g., at WiCS High School Day for Girls. I have also actively (and successfully) recruited several female students as teaching assistants for my class.

Outside Penn, I have served in different capacities to foster diversity. I have served as a mentor for female students in my community. As a member of the Diversity subcommittee within the IEEE Technical Committee on Real-Time Systems (TCRTS), I developed initiatives and policies that aim to increase transparency and to advance the representation of women and underrepresented groups within the real-time systems research community at large and in the leadership roles (such as PC members and PC chairs). Several policies that I co-developed were implemented by the IEEE TCRTS, such as keeping a large percentage of women and underrepresented researchers in the Technical Program Committee of RTSS and RTAS, no deadlines on weekends, child care at conferences for attendees with small children, travel support for female and underrepresented students, mentorship program for female students who attend the conferences for the first time, awards dedicated to women who made significant contributions to the real-time community, etc. I have been continuing to push these efforts in other capacities, e.g., as a Steering Committee member for EMSOFT and RTNS, and as a PC chair (co-chair) of conferences. Through our effort, we have begun to see changes in the perception within the community and in the leadership. I plan to continue working on these issues to make long-term impactful changes to the real-time, embedded and cyber-physical systems community.