Teaching Statement

Joy Arulraj
Carnegie Mellon University
jarulraj@cs.cmu.edu

I enjoy interacting and researching with students and derive satisfaction in helping them advance in their careers. The opportunity to continue doing it is a key factor in my decision to pursue an academic position.

Teaching Experience

My teaching experience comes from three semesters as a teaching assistant (TA) of Computer Science at the undergraduate and graduate levels. My first exposure to teaching came when I served as a TA for the introductory course on programming at Wisconsin. This role involved running weekly discussions, grading programming assignments and exams, and assisting students during office hours. I found it particularly rewarding to see the interest in programming this class sparked in students.

More recently, I served as the head teaching assistant for the graduate database course at Carnegie Mellon. Since we were offering this course for the first time in over a decade, there was ample opportunity to redesign the course content to better reflect the needs of our student body. I assisted with the restructuring of the syllabus by creating an array of new programming assignments. The most effective way for students to learn is through hands-on experience of full-featured database management systems (DBMSs). This would both strengthen their grasp of the design principles underlying DBMSs and enable them to learn modern software development practices. We needed an open-source DBMS for this purpose. Understanding the internals and making significant changes in state-of-the-art DBMSs takes a lot of time, and we also did not have the required in-house expertise. On the other hand, educational DBMSs, such as SimpleDB, do not allow the student to fully appreciate the architecture of real DBMSs. We concluded that Peloton, a research DBMS that I had been building for a year, would be suitable for this course.

I started preparing for this course by refactoring Peloton and streamlining its installation process. I began compiling a wiki with detailed information on setting up and running benchmarks on Peloton. I documented the core components of the DBMS and its design principles. Next, I wanted the students to make experimental changes in different parts of the system without worrying about breaking existing functionality. To achieve this, I expanded the regression testing suite that validates the functionality of the different components of Peloton. I encouraged the students to make use of our continuous integration server during development. I integrated the programming assignments into an auto-grading server to give immediate feedback to students on their solutions. I also maintained a leaderboard about which students have the fastest (and correct) implementation for each programming assignment. We used these results as a teaching tool to discuss the impact of the trade-offs that the students made in their code. Students commented that this was the most hands-on systems course that they had taken at CMU and appreciated the modern development practices that they learned in the course.

My most recent experience has been with the undergraduate database course. Based on feedback from students who took the graduate course, we decided to revise the undergraduate course content to focus more on the internals of DBMSs. As the head teaching assistant, I was responsible for assisting in the revision of course material, occasional guest lectures, and guiding the other teaching assistants in developing new programming assignments. We decided to situate all the assignments around the development of a new storage engine for the SQLite DBMS. The students learned the fundamentals of DBMSs but did so in a more concrete and applicable format. Overall, these experiences gave me hands-on exposure to the work required to teach at both the graduate and undergraduate levels.

As a faculty member, I would be qualified and excited to teach courses in data management and data science at both the graduate and undergraduate level. I would also be happy to teach classes in operating systems and introductory computer science. At the undergraduate level, I envision an upper-division course on teaching the art of designing software systems. This Art of Designing Systems class would introduce students to the fundamentals of architecting large complex systems that are simple to build and extend. This course would involve significant software implementation. A distinguishing feature of elite programmers is their ability to decompose systems elegantly into smaller loosely-coupled modules. This class will instill this rare skill in students by gradually exposing them to the layers of abstraction present in a DBMS. Although the programming toolchain will inevitably change, I am confident that the skills that the students acquire while exploring such a system will serve them well. In doing so, I hope to expose students to the rewarding craft of architecting systems and present opportunities for collaboration.
At the graduate level, I look forward to teaching Research Topics in Data Management, which would couple readings and discussion of current research with an open-ended, multi-week research project. The readings will focus on emerging topics in data management, such as self-driving DBMSs, hardware-conscious architectures, and data science accelerators. Like the undergraduate class, I hope to attract graduate students from diverse disciplines seeking to leverage principles underlying DBMSs and systems programming to advance their research. Computer science is an exciting and empowering field, and I will seek to convey this sentiment in my teaching.

Mentoring Experience

As the lead graduate student for the Peloton project, I have had the pleasure of interacting and researching with several talented undergraduate and graduate students. I think that new students should initially be given well-scoped tasks so that they get acquainted with the research field and are comfortable with taking on more open-ended problems. After they have gained confidence, I try to understand their interests and help them identify a research problem that excites them. I patiently guide them through the entire research process, starting from strategizing potential solutions to finally presenting their findings. I find it particularly rewarding when this research experience helps them advance in their careers. I will next summarize a few illustrative mentoring experiences.

Matthew Perron, a master’s student, took the graduate database course centered around Peloton and expressed his interest in contributing to the research project. He began by optimizing the logging manager in Peloton and expanding the associated testing infrastructure. After he had gained familiarity with the system, I introduced him to a new write-behind logging algorithm and highlighted related research problems. His hard work formed a central part of our evaluation of write-behind logging, and he was the second co-author on our VLDB paper. This research experience sparked his interest in pursuing a Ph.D. program. He was accepted to many top schools and is currently in the Ph.D. program at MIT.

Another student, Lin Ma, got involved in research when he visited CMU as an undergraduate research intern. He began his internship by optimizing the performance of cold data storage in the H-Store DBMS. I helped guide him through prototyping alternative design decisions, evaluating their performance, and then finally documenting the findings. Lin presented this work at DAMON and was the lead author on the corresponding research paper. He then joined the Ph.D. program at CMU and is a junior Ph.D. student in our group. He is currently pursuing a new research direction on transforming Peloton into a “self-driving” DBMS and I continue to mentor him.

I have similarly worked with another talented undergraduate student, Qian Li, on developing a new query execution engine in Peloton. This engine took advantage of the gather-scatter DRAM, a recently proposed hardware technology. I helped guide her through designing and implementing new query processing operators for accelerating real-time analytics using gather-scatter DRAM. Qian is the lead author on the corresponding research paper that is under preparation. Her formative research experience while working on this project sparked her interest in pursuing a Ph.D. program. She is currently in the Ph.D. program at Stanford.

Diversity

At Carnegie Mellon, I have been fortunate to participate in the BiasBusters workshops that offer voluntary training on recognizing and dealing with unconscious bias. These training sessions helped me develop a better understanding of the subtle behaviors and cues that help create an inclusive environment. This knowledge has resulted in my making tangible transformations, such as gently correcting biases and adopting gender-inclusive language. As a mentor, I have made strides to increase the diversity of my mentees by actively reaching out to underrepresented groups. I am proud to have gender parity among the undergraduate students I have mentored. As a faculty member, I will continue raising awareness about the importance of campus diversity and strive to create an inclusive research environment.