Teaching Statement
Matei Zaharia

I really enjoy teaching, so the opportunity to continue doing it is one of the reasons I seek an academic position. I believe that organizing and communicating ideas effectively, both to students and practitioners, is one of the main roles of academia, and is key to maintaining rapid progress in our field. Through my experience at Berkeley, I got the opportunity to teach in both university courses and workshops for professionals, as well as to mentor several younger students in research.

In the classroom, my main teaching experience is from TAing the undergraduate Operating Systems course, where I held weekly discussion sections, supervised group projects, prepared exams, and helped design a new final project. I also helped organize, design and teach the AMP Camp, a two-day training camp where professionals learned large-scale data analysis using software developed at the AMP Lab as part of my graduate work (Spark and Shark).[^1] The camp sold out at 150 on-site attendees, and had over 3000 people watching via live stream. Finally, I gave guest lectures for Jim Demmel, Marti Hearst and John Canny in their courses on data analysis and parallel computing.

As a teaching approach, I strongly like to promote learning by doing. I believe that the best way to learn anything is through hands-on experience, so I want to give students plenty of opportunities to do so. One of the new resources I am excited to leverage here is cloud computing. For example, as a TA for undergraduate Operating Systems, I helped design a new final project where students ran MapReduce computations on Amazon EC2—one of the first such projects at Berkeley. Similar projects are now being used in several other courses. I also helped develop EC2-based exercises for the AMP Camp, where each attendee could work with a Wikipedia dataset on a 4-node cluster. Many attendees stayed past the end of the two-hour exercise session to finish the problems. I believe that apart from offering experience with real systems, these kinds of exercises give students the confidence that they can try out, and learn, a new technology (e.g., Hadoop MapReduce or EC2) on their own, an essential skill in today’s marketplace.

In terms of research mentoring, I was fortunate to work with a number of younger graduate and undergraduate students, including Mosharaf Chowdhury, Tathagata Das, Reynold Xin, Haoyuan Li, Kay Ousterhout, Patrick Wendell, Ankur Dave and Amir Khodaei. My involvement ranged from having younger students help on a project I was leading (e.g., Spark), to taking a guidance and support role for a project led by a junior student (e.g., Orchestra with Mosharaf, or Shark with Reynold). I also guided two undergraduates through highly independent projects with me: Ankur Dave, who built a Pregel implementation over Spark (part of the Spark paper) and a Spark debugger, and Amir Khodaei, who is building a streaming SQL engine. In my advising, I hope to foster the same independence in students that I mentioned with classroom teaching: to give them the confidence that they can learn new technologies, solve problems, communicate about them, and eventually start proposing them on their own.

As a new faculty member, I would be qualified to teach courses in operating systems, networking, databases, data structures, security, and introductory computer science, as well as seminars in large-scale data analysis and cloud computing. I am excited to draw on my previous experience, where applicable, to give students the opportunity to work with real datasets and real distributed systems.

Finally, in addition to teaching course concepts, I also hope to give undergraduates a sense of the excitement of research, and to give interested students a chance to work on research problems. My own research career started from the opportunity to work as a research assistant at Waterloo as early as my first year. I still appreciate this opportunity, and I wish to offer the same to others.

[^1]: [http://ampcamp.berkeley.edu](http://ampcamp.berkeley.edu)