

TEACHING STATEMENT

Yu Su (ysu@cs.ucsb.edu)

A conceptual framework is a coherent system of concepts based on which new concepts are recognized, interpreted, and integrated. I believe learning is the life-long procedure of creating, developing, and crafting one's conceptual framework of the world. Accordingly, teaching is to understand, and to help understand and develop, each student's conceptual framework, both of a discipline and of the whole world. Through research we expand the boundary of knowledge, and through teaching we pass on the knowledge, and the ability to seek and expand knowledge, to the next generation, both of which I find extremely enjoyable and rewarding.

I gradually realized the role and importance of conceptual framework through years of learning and teaching, but it had never been more clear before my first summer in UCSB. When I was changing my PhD research area from computer networking to data mining, I was faced with an all-out confrontation between my old scientific conceptual framework, comprising concepts like network protocols, and the new concepts of stochasticity and empiricism that I had to learn and integrate into my conceptual framework. It was a painful process, but my advisor helped me sort out the commonalities and differences, and finally develop a coherent conceptual framework for the new area. It has benefited me enormously in my whole PhD study, and will continue to benefit in the years to come. Through experiences like this I shape my own way of teaching: (1) One's conceptual framework largely determines how one learns. Understanding it is the first step of effective teaching. (2) No conceptual framework is the same. Acknowledge the differences and tailor accordingly. (3) Do not feed knowledge. Instead, help students develop their own conceptual framework by guiding them to learn by themselves. A new concept is truly learned only when it becomes a coherent part of a conceptual framework.

As the lead student researcher in several funded projects, an essential part of my role is to mentor junior graduate students. When mentoring, I always work closely with students to help develop their conceptual framework of research. One of the students I mentored is Hanwen Zha, who is now a second year PhD student at UCSB. When he first entered the program, I found that although he had some experience in data mining research, he lacked a solid conceptual framework of the area as a whole. One direct consequence was that he was confused about how to do research, and started to question the whole point of PhD study. I shared with him the founding story of the area, like how and why data mining developed out of a workshop in SIGMOD, and walked him through how several classic papers evolved into independent subareas of data mining, and the philosophy of the area reflected in the evolution. Gradually he connected the previously isolated concepts and became an independent data mining researcher with a solid conceptual framework of the area. Together we published a paper in the SIAM Data Mining Conference 2018 (to appear), and he is now independently leading his own project on text mining in scientific literature.

I was teaching assistant for multiple undergraduate-level core-CS courses such as CS32 (Object-Oriented Design and Implementation) and CS176A (Introduction to Computer Communication Networks). Also, I was teaching assistant and guest lecturer for graduate-level courses including CS290D/291K (Advanced Data Mining) and CS273 (Data and Knowledge Bases), and I gave a 1.5-hour tutorial on Knowledge Base Construction and Querying to a group of over 40 graduate students in the ACM Conference on Information and Knowledge Management 2017. When lecturing, I find it most effective with an organic combination of electronic materials and chalkboard. For example, in the Word Embedding lecture for CS291K, I used PowerPoint presentations to show intriguing illustrations of word embedding geometry, and used chalkboard, in a step-by-step and interactive way, to derive the formulae for optimizing word embedding models. In this case handwriting on chalkboard inspired students to think more actively and ask more questions, and encouraged active learning in the classroom.

Through these teaching experiences I also learned the differences between teaching undergraduates and graduates. Undergraduates are in a critical stage of creating their first scientific conceptual framework of a discipline. It is of the outmost importance to help them create a solid one, by both giving comprehensive coverage of the course subject, and teaching them methodologies to examine new concepts and develop their own conceptual framework. In contrast, to varying degrees, graduates have already developed a conceptual framework. Teaching is then more

about guiding each of them to further explore the concepts that trigger their most interests, or debate on how and why some concepts conflict with their conceptual framework, both of which can encourage self-exploration and fruitful, long-lasting outcome.

As a first-generation college student growing up in an under-developed town in China, I benefit enormously from equal education. I deeply understand the irreplaceable role of education as ladder for social class transition, and I believe all people should have an equal right to education regardless of their gender, ethnic group, and social class. As an advocate for data science democratization, I am strongly interested in collaborating with other STEM or non-STEM faculty members to develop interdisciplinary courses on data science. I would particularly encourage students of underrepresented groups to join the courses, and will also actively seek to recruit students of underrepresented groups to join my research group. In addition, I would love to participate in outreach programs and stimulate students' interests in data science and STEM early in their education.

When I teach more, I learn more. I am excited to continue teaching and mentoring and refining my teaching philosophy. Trained in multiple fields including Data Mining, Natural Language Processing, and Machine Learning, I would be delighted to teach both undergraduate-level and graduate-level courses on these subjects, including related courses like Databases and Data Management, Artificial Intelligence, and Information Retrieval. I would also welcome the opportunity to teach other core computer science courses in the curriculum to undergraduates. Apart from such regular courses, I also look forward to teaching or developing new graduate-level courses on more specialized topics like Natural Language Understanding, Text Mining, Knowledge Base Construction, and Deep Learning.