Perspectives on Teaching Data Visualization

Jason Dykes* City University London
Daniel F. Keefe† University of Minnesota
Gordon Kindlmann‡ University of Chicago
Tamara Munzner § University of British Columbia
Alark Joshi¶ Yale University

ABSTRACT

We propose to present our perspectives on teaching data visualization to a variety of audiences. The panelists will address issues related to increasing student engagement with class material, ways of dealing with heavy reading load, tailoring course material based on the audience and incorporating an interdisciplinary approach in the course.

Developing and teaching truly interdisciplinary data visualization courses can be challenging. Panelists will present their experiences regarding courses that were successful and address finer issues related to designing assignments for an interdisciplinary class, textbooks, collaboration-based final projects.

1 INTRODUCTION

Teaching a course in data visualization can be daunting. We have a growing collection of research papers coupled with some excellent textbooks and the ever increasing supply of resources on the web. The other conundrum of teaching such a course is the level at which one has to teach it in order to make it engaging and sufficiently challenging.

In this panel, we propose to discuss our experiences with teaching data visualization to a variety of students including non-majors, higher level undergraduate students and graduate students. Our panelists have developed or have been involved with developing novel visualization courses that include students from a wide variety of backgrounds. We address the challenges involved with creating assignments for such students and keeping them engaged throughout the semester by benefiting from their unique expertise.

Our panelists will share their experiences with developing and teaching such courses and will provide some insight into the successful and unsuccessful aspects of their respective courses.

2 DANIEL KEEFE

My experiences with teaching data visualization come from my current Interactive Scientific Visualization course at the University of Minnesota and the course Virtual Reality Design for Science, which I co-developed several years ago with David Laidlaw (Brown University) and Fritz Drury (Rhode Island School of Design). The Brown/RISD course was unusual in that half of the students were from computer science at Brown and half from art, design, and illustration at RISD. Due to these different backgrounds, finding a common ground for assignments and discussion was a challenge. To overcome this, we incorporated a number of traditional art classroom activities, such as critique, into the course, and we also developed a VR tool that artists can use to design visualizations in VR without programming. From this experience and other courses at RISD, I have learned how valuable visual design assignments and group critiques can be to teaching. I now incorporate each of these in my teaching, even in courses that draw entirely computer science students.

My visualization course at the University of Minnesota places an emphasis on interacting with data and draws heavily upon research in virtual reality and human-computer interaction. One strategy that I have used in this course to teach about visual design and critique is to lecture and develop assignments around sketching (prototyping) user experiences, as inspired by Bill Buxtons recent book, which includes a number of smart case studies and rapid design activities, from paper to video prototyping.

My perspective is that visualization courses should reflect the interdisciplinary (often team approach) to discovery that is part of visualization. I try to do this by highlighting the connections between visualization and art/design, but to be successful in visualization students also need to learn how to work with scientists, doctors, and other collaborators who need our tools. The best way that I have found to learn this is to practice. Thus, in my visualization course, students work directly with my current research collaborators. This has some cost associated with it, but it is usually not too difficult to identify a good course project to spin off from our current research efforts, and the interdisciplinary discussions that come from these projects cannot be simulated. Last year, three student projects from this course were presented as posters at VisWeek, collaborations with evolutionary biologists, physical therapists, and archaeologists.

Biosketch Daniel F. Keefe is an Assistant Professor in the Department of Computer Science and Engineering at the University of Minnesota. His research centers on applications of visualization, human-computer interaction, and computer graphics to other disciplines, including orthopaedics, medical device design, surgical training, archaeology, evolutionary biology, fluid mechanics, art, and design. Particular interests include visualization of motion and fluids, natural user interfaces for exploring vast scientific datasets, and supporting creative design processes within data-rich environments. He did post-doctoral work at Brown University jointly with the departments of Ecology and Evolutionary Biology and Computer Science, and he received the Ph.D. in Computer Science in 2007 from Brown University, where his dissertation work was nominated for the ACM Dissertation Award.

3 GORDON KINDLMANN

The Scientific Visualization class I teach at the University of Chicago (CMSC 23710/33710) is organized around a mathematical and algorithmic definition of fundamental visualization methods. In a school that is known more for theoretical inquiry than engineering prowess, I feel it is important to start with the theoretic descriptions, as well as reading widely in the visualization literature to see how methods are described and published, and let students build up their own implementations and working knowledge. The result is (if all goes well) a class of students who can understand and critique visualization techniques from a position of personal insight and experience. Although the Scientific Visualization class is now part of a two-course Scientific Computation sequence, the instruction of the class does not assume any particular background in numeric methods or software engineering. This motivates the use of Python as the language for the programming projects, since this language and the wealth of associated packages makes it easy (and satisfying) to

*e-mail: j.dykes@soi.city.ac.uk
†e-mail:keefe@cs.umn.edu
‡e-mail:glk@uchicago.edu
§e-mail:tmm@cs.ubc.ca
¶e-mail:alark.joshi@yale.edu

start generating images from scratch.

Many of the Computer Science classes (especially electives) at the University of Chicago are small, so graduate and advanced undergraduates sit together in the same class. This works to the benefit of the undergrads who are surrounded by peers with more experience and who ask more sophisticated questions; it is less clear what benefits there are for the graduate students. Grading in the class is based on programming assignments which require students to implement visualization methods, apply them to given datasets, and describe (in prose) what features in the data is revealed by the method. Graduate students are required to complete more tasks and answer more questions than the undergraduates, and they may opt to use C or C++ instead of Python for their implementations.

While it may seem intimidating, the class has a mathematical flavor derived from vector and tensor analysis. I feel most treatment of linear algebra and tensor analysis is unnecessarily opaque, and that a clear discussion of these basic mathematical tools can empower rather than confuse students. The most important aspect of this is a reliance on direct (rather than index) notation for all vectors and tensors, so that students appreciate vectors and tensors as fundamental objects, with particular relationships and descriptive powers, rather than as collections of numbers in an arbitrary basis. This also helps students build their geometric intuition for the vectors and tensors involved in visualization (for example, the gradient is the direction and rate of fastest change, rather than a list of partial derivatives). With this mathematical background, the students can understand that the scalar, vector, and tensor fields studied in scientific visualization all have scalar, vector, and tensor attributes, which can be ingredients in visualization methods.

The class follows a fairly traditional structure in laying out the fundamentals of scientific visualization. After a review of some aspects of perceptual psychology (emphasizing the fundamental role that luminance plays in our perception of image structure), I introduce the basics of vector and tensor analysis. This uses Taylor's theorem as a context for understanding derivatives (gradient, Jacobian, Hessian), and using the spectral decomposition as a context for understanding eigenvalues and eigenvectors. With this context, we discuss methods for visualizing scalar (colormaps, isocontours, volume rendering), vector (glyphs, streamlines, LIC), and tensor (glyphs, hyperstreamlines, generalized LIC, topology) attributes of scientific data fields.

There are different ways of keeping students engaged with the material. Besides the spontaneous questions and discussion that can arise in a small class, I think that teaching from published research papers is one way of making the material more intriguing. Students see that many of the important papers were written in the last ten or twenty years, that many of them were written by students. It becomes clear that this is a very current area of inquiry, and that new innovations and discoveries are not out of their reach. I am hoping that a future iteration of the class will provide the structure to allow students to pick an area of interest and build towards a research contribution by the end of the class.

Biosketch Gordon Kindlmann is an Assistant Professor in the Department of Computer Science at the University of Chicago. Previously, he was a Postdoctoral Fellow at the Laboratory of Mathematics in Imaging in the Brigham and Women's Hospital at Harvard Medical School. He completed his doctoral work on visualizing diffusion tensor fields at the University of Utah with his advisor, Chris Johnson. He got his Master's and Bachelor's degrees in Computer Science from Cornell University.

4 Tamara Munzner

I have taught an introductory graduate information visualization course seven times since 2002, and contributed to Pat Hanrahan's visualization courses in 2001 and 1996. I will discuss the evolution of my course over the years in terms of both content and structure. These changes include: making the course accessible to nonmajors by adding analysis and survey projects in addition to programming projects; requiring written comments on the reading before lecture, and a series of experiments on how to respond to and how to grade them; and adding more and more milestones to the projects in order to give students more early guidance to a successful outcome.

I lecture on the required core readings in the first part of the course, while the last part is individual student presentations to the group on several papers on a topic of their choice. In some previous iterations, I interleaved my own lectures with student presentations. The advantages of the current structure is that the students have done the bulk of the readings before they must choose project topics, and they are more engaged in talks from their fellow students that cover material they have not seen rather than material they have just read.

The reading is a mix of original papers and chapters from the Ware and Tufte books. The biggest student complaint about the course is the very heavy reading load, with up to five papers/chapters per lecture for two lectures a week for the first six weeks. The problem with teaching from original papers is the intellectual points that I intend for them to illustrate are often quite different from the intent of the authors in writing the papers. I have thus begun writing an infovis textbook with a heavy focus on synthesis and integration of ideas. With such a textbook, I look forward to reworking the course to have less reading and more design exercises before the students embark on a large final project.

Beyond teaching the specifics of visualization, I also want to impart the more general research skills of how to read and critique research papers, and how to present coherently in both speaking and writing. I have developed a rubric for detailed written comments on their written and oral work that addresses both style and content, culminating in feedback for their final project at the level of a paper review.

Biosketch Tamara Munzner is an associate professor at the University of British Columbia Department of Computer Science, where she has been since 2002. She was a research scientist from 2000 to 2002 at the Compaq Systems Research Center in California, earned her PhD from Stanford between 1995 and 2000, and was a technical staff member at the Geometry Center mathematical visualization research group from 1991 to 1995. Tamara was InfoVis Co-Chair in 2003 and 2004 and EuroVis Co-Chair in 2009 and 2010. Her research interests include the development, evaluation, and characterization of information visualization systems and techniques from both user-driven and technique-driven perspectives.

5 Jason Dykes

Jason Dykes is a Senior Lecturer in the giCentre and Department of Information Science, City University London and a National Teaching Fellow of the UK Higher Education Academy.

A Geographic Information Scientist with 20 years experience in teaching and developing innovative learning resources, Jason uses visualization as both subject matter and method, engaging students through constructivist and exploratory approaches to learning: these include group exercises and field work. Author of the cdv, panoraMap and pixelEx software for education, Jason has been involved in a number of funded educational projects, pioneered distance learning at City through the online Masters in Geographic Information and published on pedagogy in the GeoSciences and Information Visualization.

Along with giCentre colleagues, Jason has delivered hands-on software supported tutorials at VisWeek in 2007, 2008 and 2009 and been awarded Best Paper at the GIS Research UK meetings in 2007, 2008, 2009 & 2010.

Most recently, Jason has developed Masters Level learning programmes for students with background in arts, social science and the humanities and has been espousing a constructively aligned approach to course development by collecting a series of structured learning outcomes from the InfoVis community.

Jason will present a curriculum and teaching strategy for a Masters programme under development at City, reflect upon approaches for engaging students and discuss an outcomes-based approach to visualization education.

6 ALARK JOSHI

In the Spring 2010 semester, I developed and taught a new introductory data visualization college seminar at Yale University. The class consisted of students from all levels (freshman to seniors) from a wide variety of majors such as Physics, Economics, Political Sciences, Biology, Computer Science and so on. The aim of the class was to introduce students to the strengths of data visualization and to familiarize them with the visualization techniques being developed by the data visualization community.

One of the primary goals of the course was to instill good data visualization principles as well as the ability to critically evaluate visual representations that students would routinely see in the media or research papers in their field. With that in mind, the reading material was drawn from books by Edward Tufte, Stephen Few and research papers. In order to keep the course material relevant, I focused on topics that students are generally interested in such as the Environment (Global Warming, Carbon Footprint, Natural Disasters), Economy, Social Media (Facebook, Twitter), Security, Privacy and so on.

Since my class consisted of students who were not necessarily proficient with computer programming, I decided to use visualization tools such as Tableau Software and IBM's Many Eyes for most of my assignments. Additionally, I held a programming workshop for students who wished to use *Processing* for their assignments and final projects. Class assignments were always relevant, which kept the students engaged in the material. For example, right after the devastating earthquake in Haiti, I asked the students to create a geographic visualization of the amount of aid that countries were pledging as compared to their gross domestic product(GDP). These kinds of assignments brought about discussions on their blogs and in the classroom.

In order to increase student engagement in the class material, I required students to maintain a blog for the class. They were required to post their reading response on their blog in addition to interesting data visualizations that they encountered in their daily life. On the class website, I maintained a short preview for each of their blogs which led to students reading each others blogs before class. This helped me identify the overall sentiment regarding the class material before I went into the class and helped me highlight or relate to a specific student's comments on their blog in the classroom. Based on the class reviews, the students seemed to have enjoyed this concept of having a blog for the class.

For the final project, I required students to propose their own project. In the project proposal phase, other students and I provided feedback regarding the feasibility of the project and pointed groups to relevant data sources. This ensured student engagement and led to successful outcomes for the students. We are in the process of refining some of the projects for deployment on the web.

Biosketch Alark Joshi is a Postdoctoral Associate in the Department of Diagnostic Radiology at Yale University, where he works on data visualization projects for improved neurosurgical planning and treatment. He is a core member of the BioImage Suite team at Yale, whose mission is to develop and disseminate advanced image analysis and visualization software for widespread use. He received his Doctoral degree in Computer Science from the University of Maryland Baltimore County with Penny Rheingans, his Masters in Computer Science from the Stony Brook University and his Bachelor's degree from the University of Pune, India.