

Virtual Undergraduate Town Hall FAQ
Enabling Science Breakthroughs Using Computer Science
Speaker: Deb Agarwal, Host: Gail Murphy
10/13 at 7pm ET

How much science do you actually have to know?

- I have always loved science, but I chose computer science to major in. I mostly have a high school education, and a little bit of college education in these other sciences. The scientists I work with are desperate for help; they are at a point where they need computer science to make their next science advance. They are happy to teach us their science, so we can help them. I love learning new science areas and I immerse myself in the science in each project. Even Nobel laureates are really patient in teaching us their science.

What is the hardest technical challenge that you have had to tackle from a computer science point of view?

- The hardest technical challenge is figuring out what the scientist really needs. They have a difficulty time translating their needs into computer science, for what they need. We had to change our conversation/approach, we have incorporated Human Computer Interaction techniques into how we engage. It has given us an approach to thoroughly understanding the problem, the context, and what is needed before we try to solve it.

Are there certain tools that you use to be able to work in all these different fields?

- The most important tool you need to use is your brain. Working across disciplines is difficult. In terms of CS tools, moderate scale problems usually require a database and an indexing technique. Machine learning is becoming more prevalent, a great example is predicting the future frequency and magnitude of extreme storms, the predictions use climate simulations of the future. The way you identify the storms is by looking at model grid nodes, each node has wind and direction for each point in time, identification of a storm uses machine learning to put the grid points together to recognize large-scale structures that are a storm.

How does college leadership help an internship application at the lab?

- To apply for an internship at the lab, certainly the more you can reference your ability to work in teams and on team projects the better. Leadership and diversity of activities in college helps you stand out. If you are interested in interning at a lab, look up people you might work with and e-mail them a personal note expressing your interest. Input from participant: My thought was that the 'diversity of activities' would indicate a well-rounded individual, capable of working in a variety of groups focused on a myriad of goals and topics, not just research related ones. So, participation in campus organizations that speak to your other interests would be a good idea. (Deb expressed that she agreed with this input).

What advice could you give on the application process?

- Look at the human resources pages at the place you want to apply to understand what internship opportunities exist. Then look at groups and areas that you would be interested in working with. Apply to the programs you are interested in. If you find groups and individuals you are particularly interested in working with, contact them and ask about internships. In the note indicate why you think you would be a good fit for an internship with them in particular.

How much of what you learned in college have you used in practice?

- I did my undergraduate degree in mechanical engineering. My engineering courses have given me a good basis that I often use to understand the science domains I work with. None of the courses directly help but the general understanding of how things work often comes in handy. In my graduate work I learned computer science and computer engineering. I often use the insights from those classes to understand and solve the challenges I work on. For instance my distributed systems classes help me to understand how cloud based technologies work and how to make effective use of distributed computational resources. During the PhD I learned how to understand a topic deeply and become a world expert in the topic. This experience gave me the confidence that I can do this same thing for other topics and has given me a roadmap for how to do it.

For a research project, is good data or a good methodology more important and contribute more to the success of the project?

- Both are very important but I would say good methodology is more important because it is very difficult to have good data without good methodology. It is really easy to produce bad data if you are not methodical in setting up the data collection.

Are there other ways to become assertive other than theatre?

- In terms of how to be assertive without theater, there are many ways. There are several tricks you can use such as where you sit at the table and talking to people before a meeting to line up advocates for an idea or point you want to get across at the meeting. In order to break into the conversation in a meeting, I often just raise my hand and lean forward. People will stop to give you a chance to speak. Being assertive is aided by preparation - know the points you want to make, align allies and be confident!

What are the challenges of using a Computer Science background to do research in a non-CS field? How much time does it take to make impact?

- Making impact in a science field can typically occur within 3-6 months and should. The key is to take an agile approach to solving the most important priorities first. The key is to start with the most simple and pressing problem the scientist is facing. This allows you to build trust and get them to recognize you care about helping. The priority might be as simple as a data format translator. It is better not to approach cross-disciplinary work with an assumption that your work will lead to a computer science breakthrough. Many of the problems are mundane from a CS perspective but the solution is revolutionary in the science. I am not trying to do research in a non-CS field. Instead I am trying to apply CS to advance a non-CS field. This work often leads to CS research challenges.

You mentioned taking "an agile approach" - what are some ways to quickly come up with a rough hypothesis and developing a relevant problem statement that can be cracked in 3-6months?

- We use an approach based on User-Centered Design to work with the scientist to understand how they work and the challenges they face. We use discussion of the interface and end behavior to build our understanding of what is needed. We do not end up with a hypothesis but instead a concrete description of the thing the scientist needs. This allows us to achieve results in a short timeframe.

What was your first research project/topic?

- My first research project was working with an experiment at the Advanced Light Source to enable remote experimentation (they have thousands of users a year coming from all over the world). The intent was to allow them to do their experiment remotely.

I wonder how would one balance the time and energy for core skills development vs on networking/socializing?

- Balancing core skills versus networking is mostly a matter of segmenting your time into focused work and interaction periods. Both are very important to do and the amount of each needed is relatively situation dependent. Early in my career I spent more time on skills development and at this point in my career, I spend more time on networking.

Is it necessary to obtain a PhD before doing CS research? Can meaningful research be done with just a BS? What would you recommend?

- You can be involved in research with a BS about 1/2 to 2/3 of my group is BS and MS level people. An MS is preferred. The advantage to a PhD is that you are in charge of directing the overall research direction. However, the advantage to being a BS or MS in the environment is that you get to concentrate on building capabilities and not having to write proposals.

Why have some fields of science been so slow in using computer science to help further research?

- In terms of why are fields slow to adopt computing - the key driver is reaching a point where the scientist's existing tools no longer allow them to do the science; biology and earth science reached this point much later than physics and cosmology for instance. The development of large-scale detectors has often driven the need for computer science. However, in earth science they do not really

work with large-scale detectors.

**Is there still a good chance to become a (data science)researcher at a company without going through the PhD track?
What's your view on research job experience at industry vs at academia?**

- Yes, you can become a data scientist with BS/MS/PhD any of the levels are enough to do data science. Each level of education allows you to do different jobs.