Overcoming Research Challenges

Three years ago I set a goal for myself: write a PhD on speech recognition and deep learning. To accomplish my goal, I had to learn a lot, fast.

My approach to learning difficult concepts is the following: teach myself as much as I can, and get help when I hit a wall. As a result, I have become a better learner, better communicator, and more confident researcher. I have learned to formulate my questions well, listen to what is being said, and most importantly, I'm not afraid to fail.

The fear of failure held me back from seeking help for the first year. That year was marked with caffeine fueled late nights in the lab, teaching myself linear algebra, probability theory, and backprop. My progress was slow, so I worked harder. Asking for help didn't seem like an option. However, it became obvious that I couldn't learn everything from books, papers, and blog posts alone. I needed directed answers to my questions. I needed to learn from the experts. The only problem was, I had to go find them.

If the mountain won't come to Mohammed...

My dissertation topic is multi-task learning for acoustic modeling in speech recognition for low-resource languages. Of that multifaceted topic, I needed help in one area in particular: neural net acoustic modeling. Jean-Luc Gauvain and Lori Lamel lead a research team at CNRS near Paris, specializing in speech recognition for low-resource languages. I reached out to them, asking if I could learn from them as a visiting scholar. They said yes. Together, we applied for an NSF grant and funding from the French Embassy to the US, and were lucky enough to get both. Along with my NSF Graduate Research Fellowship, I had enough funds to work out of their lab for a year. So I packed my bags and headed for Paris.

However, even after I arrived at the lab, I hesitated to ask for help. Lori got her PhD from MIT in Electrical Engineering, and Jean-Luc wrote *the* paper on MAP adaptation for GMMs. I felt out of my league. However, one day I ran into a problem in training my models, and I couldn't figure it out by myself. The problem was somewhere at the very beginning of the training pipeline, but I didn't understand the procedure well enough to troubleshoot. It was a very fundamental question about Expectation-Maximization of Hidden Markov Models. I understood the iterative algorithm (Baum-Welch) after it got started, but I couldn't wrap my head around the very first step: parameter initialization.

I made sure I could express the problem well, drawing out a detailed diagram, and then went over to ask Lori and Jean-Luc. After some back-and-forth clarifications, I had my answer! Each state of the Markov chain is assigned an equal chunk of the observed data sequence: that's all. The intuition I worked out with is the following: the initial parameters don't actually matter, because the EM algorithm is guaranteed to produce increasingly better models every iteration, approaching the best model given the data. We could just as well randomly initialize the parameters, and given enough iterations of Baum-Welch our model will still be the best possible model. However, by using the data itself as a first guess, our model

will begin training with a head-start, shortening the total number of iterations we need to get the best model.

With this new understanding, I got back to troubleshooting. After combing through my scripts, I realized that because of my small dataset, some parameters were being assigned too few training data examples. To solve the problem, I collapsed some parameters among linguistically similar labels.

Later that day at lunch with the postdocs, I excitedly told them about the answer to the problem and my new intuition. To my surprise they were actually interested to hear the details of Jean-Luc's explanation. The postdocs worked on another section of the speech recognition pipeline (the language model), and didn't know all the intricacies of acoustic modeling. This taught me that it is possible to do good research without being an expert in every single aspect of the system pipeline. From that day on, I spent more time talking with the researchers and postdocs, learning more at the lunch table than from my textbooks.

Since then, I've taken every chance I get to learn one-on-one from machine learning researchers. I spent a week at the University of Edinburgh learning about neural speech synthesis with Simon King, and an afternoon learning about the Kaldi speech toolkit with Dan Povey's lab group at Johns Hopkins University.

My Research Passion & the Google AI Residency

Everyone should have access to speech technology in their native language.

This is the conviction that drives my research. I want the world to be a place where someone born blind doesn't have to learn a second language just to read Moby Dick. I've personally met very smart, dedicated people who couldn't attend university because they are blind, and the audiobooks they need to study don't exist in their native language. Working on technologies for the Kyrgyz language, I've played a small role to make speech synthesis available for free, but there is a hard limit as to what I can accomplish alone.

Google is hands down the world's leader in accessible language technologies, and I want to be a part of the team to make these technologies possible.

The Google AI residency is the ideal first step to accomplish my goals. I want to publish influential work and get more people caring about low-resource languages. With my knowledge of linguistics and passion for machine learning, the Google Brain and AI teams are the perfect place for me to flourish, and share my enthusiasm for research. Collaborating with researchers at Google, I will be able to help create new approaches and algorithms, testing theories on multi-task learning, and working towards better methods for low-resource languages. I have many ideas for multilingual experiments in particular, extracting information from big datasets to transfer knowledge to smaller domains. With the guidance of Google researchers, I will be able to efficiently run more experiments.

I see the list of languages *not* in Google Translate as boxes waiting to be checked off, and I want to help make that happen.

Thank you for your time and consideration.