

I have studied and worked on computational linguistics for quite some times. There is a glooming fear among the Natural Language Processing (NLP) community that nowadays people can solve NLP problem without any linguistic knowledge, i.e. word embeddings. On the other hand, linguistic knowledge is sometimes important in other fields, such as data mining or computer vision, because natural language text is part of any “big data”. I rather see this phenomenon as a positive trend of crossing between various topics. In fact, I am interested in the opportunity that I can work in a various fields with my background of mathematics, machine learning and linguistics.

Instead of a formal introduction of my research interests, I offer a summary of my master thesis, which earned me excellent score at Charles University in Prague. The thesis was carried out in 12 months under supervision of Dr. Ondrej Bojar (without course works), providing me an insight into the life of a PhD student. Working with Dr. Bojar was indeed a motivation for me to develop myself more. I grew more and more interested in computational linguistics in general. He gave me the full freedom to propose my plan, implement necessary tools, conduct experiments and write report while simultaneously giving me advices and monitoring my project.

We approached the problem of statistical machine translation between Czech and Vietnamese, which had been proved to be difficult, and no one had ever worked on it. After discussing, we decided to alleviate the problem by pivoting, i.e using a third language. The project offered a realistic condition that we had to prepare our own training data, test data and implementation. Sometimes, I could not help feeling overwhelmed with the huge amount of work as we started from scratch. The project taught me how to maintain a schedule in a long period, especially when some experiments failed.

After graduation, I remained wary about committing myself prematurely to a PhD program or a specific research topic. I opted to take a gap year away from school, hoping to gain some perspective, work on a different topic and strengthen my background before embarking on a long period of PhD. I applied and became a research assistant at the National University of Singapore, which is a bit closer to my home. I moved to Singapore to work under supervision of Professor Hwee Tou Ng.

Upon joining Prof. Ng’s research group. I was assigned the task of grammatical error correction. After discussion, we agreed to focus on the statistical machine translation approach, i.e. building a system to translate bad English to good English. I devised a method to post-process the output of standard SMT systems, resulting in a significant improvement ( $p\text{-value} < 0.01$ ) over a state-of-the-art baseline. I subsequently wrote a paper (as first author) on these results and submitted it to the IJCAI 2016.

Time being a research assistant allows me to reread my past research project from a more critical stance. It also gives me sufficient time to search and research for doctorate program that best suit my needs. University of Edinbergn is always in my mind. Not only it is home to the famous Moses system that I have been using for statistical machine

translation, it has an outstanding record of research.

Though I cannot predict how different the environment there is from my past experience, I believe that I can adapt to the new environment. My past experience has taught me that I am no genius, but I can get things done through hard work. It took me a month to cope with and thrive at Charles University in Prague. It took me much less time to familiarise with the research environment here in Singapore. I believe that I can do it again.