

I'm 26. I have studied and worked on computational linguistics since I was an undergraduate. There is a glooming fear among the Natural Language Processing (NLP) community that nowadays people can solve NLP problem without linguistic knowledge, i.e. word embeddings. On the other hand, I witness NLP applications being used 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 among different topics. In fact, I am interested in the opportunity that I can work in many 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 (without course works) under the supervision of Dr. Ondřej Bojar, 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. He gave me the freedom to propose my thesis plan and follow the plan, while closely monitoring the project. Sometimes, I could not help feeling overwhelmed by the workloads and the pressure, yet I grew more and more interested in the topics.

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. We aimed to alleviate the problem by pivoting, i.e using a third language. We started the project from scratch, leading to a huge workload of experiments including data preparation, test set preparation and method implementation. However, it showed me the importance of conduct every step with a standard procedure. The project taught me how to maintain a schedule in a long period of time, especially when things don't go according to plan (which tend to happen frequently).

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. 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 Prof. 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. At this moment, I am working as a corresponding author in a project which aims to an upcoming conference.

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 Edinburgh is always in my mind. Not only is it home to the famous KenLM toolkit and Moses system that I have been using for statistical machine translation, but also it has an outstanding record of research.

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. Although I cannot predict how the PhD programme would be, I believe that I can adapt to the new role in

a new environment. Indeed, I am proud of my background in statistics and linguistics, and so as my past experience. I am able to I have proved to be capable of working on different topics with different people.

I should also stress I have a more general interest in NLP related topics, especially statistical machine translation.

To date, I have studied the difference of languages. I believe that I would be valuable in your research group, who can work in a diversified environments and