STATEMENT OF PURPOSE

Shang-Yi Chuang
Carnegie Mellon University ⋄ Machine Learning and Computer Science ⋄ Ph.D. Applicant

---

Personal Statement

I am Shang-Yi Chuang, a research assistant working on artificial intelligence about speech processing and computer vision with Prof. Yu Tsao at Academia Sinica, the national academy of Taiwan. I am also a member of a natural language processing (NLP) group collaborating with Prof. Hsin-Min Wang, Prof. Keh-Yih Su, and other research fellows at Academia Sinica. My bachelor's degree is from National Taiwan University, majoring in mechanical engineering and minoring in electrical engineering. And when I was an exchange student at Osaka University in Japan, I studied robotics and conducted research about humanoid robots with Prof. Tomomichi Sugihara as well.

Motivations

Computational linguistics is my passion, and I would like to leverage both my linguistic skills and engineering ability to build a manner-aware agent–a translation agent able to adjust manners. My strong interests in natural languages and communications started to grow in the year when I was doing research on humanoid robots and experiencing cultural inspirations at Osaka University. I got lost in translation–I was able to communicate in Japanese but failed to understand the implications behind the words, and therefore undermined my local social network because of my culturally inappropriate vocabulary choices then. Even now, I still think it would be great if I had a manner-aware agent, and ergo I am determined to devote myself to this cross-culture communication support to help people who face the same problem as I do.

Carnegie Mellon University (CMU) has one of the most renowned machine learning and computer science programs. As an individual pivoting between multiple fields, I particularly admire the extraordinary interdisciplinary environment in CMU. From my own experience in both engineering and cultural domains, I realized a principle that all the technologies, the systems, and research should be human-centered, which CMU takes well care of while a lot of technological educational institutes usually forget. In addition, at the Language Technologies Institute at CMU, Madaan et al. proposed a method of politeness transfer in the paper of "Politeness Transfer: A Tag and Generate Approach", which is exactly a monolingual version of the mentioned manner-aware agent. For the above reasons, I would like to apply for the Ph.D. programs in machine learning and computer science at CMU with my best efforts.

Recent Researches

Belows are two of my projects I have been working on recently: a lite audio-visual speech enhancement system (LAVSE) and a cross-lingual QA system. With these projects, I have gained hands-on experience in various domains in artificial intelligence, including speech, computer vision, and NLP.

I have published LAVSE with Prof. Tsao and Prof. Wang at the conference of INTERSPEECH. The proposed LAVSE system addresses the problems of general audio-visual multimodal systems: additional processing costs and the privacy issue of facial images. The extension journal version of LAVSE is called improved LAVSE (iLAVSE) and it is submitted to IEEE Transactions on Audio, Speech, and Language Processing. iLAVSE further considers the data compression on image domain, the asynchronization in multimodal data, and the problem of low-quality data. These works are robust against unfavorable conditions in real-world scenarios like insufficient hardware or inferior sensors, and hence greatly improve the applicability of DL-based models on embedded systems. From these audio-visual projects, I have

mastered fundamental concepts and the know-how of deep learning, while absorbing domain knowledge of both speech and visual signal processing.

In the meantime, apart from signal processing, I started an NLP project about cross-lingual transfer learning on a QA system with Prof. Su. So far, most of the available datasets are in dominant languages, and it is undeniable that technologies mainly serve people using these languages, leading to unfavorable inequality. The cross-lingual transfer learning project can reduce the bias caused by limited data in minority languages. Moreover, it involves multilingual processing which serves as a core technique of cross-culture communication support. As a result, this project paves the way to my next step of realizing my manner-aware agent by the machine learning or computer science program at CMU.

Academic Interests

My ultimate research goal is to build a translation agent which can adjust the manners within speech. To be more specific, manners can be very different from language to language. Thus, the agent has to be able to translate not only the literal meanings between different languages but the manners among them as well.

To accomplish my manner-aware agent, the work can be broken into three directions: (1) evaluation standards in different cultures, (2) the monolingual manner-aware agent, and (3) the multilingual manner-aware agent. First of all, the existing manner evaluation methods are not yet comprehensive enough when it comes to sociolinguistics. Take Japanese as an example: students should talk to their teachers with honorifics, but students are allowed to drop honorifics when talking to their classmates. Evaluation systems have to consider specific vocabulary in different situations while taking care of conversational tone as well. Besides, the level of politeness varies from culture to culture, even when the cultures share the same language such as British English and Indian English. A conceivable challenge is to keep the consistency between evaluations and target languages, while avoiding stereotypes and biases. Next, for the second direction, the monolingual manner-aware agent, some research about affective conversational systems already exist. The systems have the potential to be extended to the third direction, the multilingual systems, with some help from the state-of-the-art machine translation models and the technique of transfer learning. Any of the three directions is a great topic for future research, and has the strong potential to improve the development of politeness-related applications. With my speech processing background, I am also interested in speech to speech translation which is able to preserve the emotion within speech signals and it is a prospective candidate for further improvement in my manner-aware agent. Led by the unstoppable globalization trend, these innovative technologies are getting more crucial than ever before, and they can certainly benefit people all around the world.

Research about manners in the NLP field is not yet common. After scrutinizing the personal webpages and publications, I found that my research interests coincide with that of Prof. Graham Neubig, Prof. Ruslan Salakhutdinov, and Prof. Scott E. Fahlman, so I am excited about the opportunity to work with them. As the artificial intelligence industry prospers, a fair and reasonable translator is naturally expected to be available in the future. I am confident that the manner-aware agent has great potential to enhance cross-cultural communication.

Conclusion

Pursuing a Ph.D. degree can significantly help me become proficient in the ability I need, e.g. interdisciplinary adaptability and analytic skills, and more capable of contributing to the manner-aware agent. For the reasons that CMU has (1) a top-notch natural language processing group, (2) the amazing collaborative environment, and (3) remarkable professors who share the same interests as I do, I would love the opportunity to use my experience at Osaka University and Academia Sinica to offer my skills, enthusiasm, and unique perspective to your prestigious machine learning or computer science program. Studying in your graduate program is my top choice for a Ph.D. degree.