STATEMENT OF PURPOSE

Shang-Yi Chuang

Carnegie Mellon University ⋄ Machine Learning and Computer Science ⋄ PhD 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 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 as well. I hold a bachelor's degree from National Taiwan University, majoring in mechanical engineering and minoring in electrical engineering. I have also studied robotics and conducted research about humanoid robots with Prof. Tomomichi Sugihara when I spent a year at Osaka University, Japan as an exchange student.

Motivations

I am extremely interested in computational linguistics due to my curiosity about natural languages and my engineering background. My strong interests in linguistics and communications grew back in the year when I was doing research on humanoid robots and experiencing cultural inspirations as an exchange student 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 translation agent with manner adjustment, so I am determined to work on cross-culture communication support to help the people who confront the same problem as I do. Since I majored in engineering in college, combining linguistics and engineering becomes an obvious choice for my future research field to fix this problem. Pursuing a PhD degree can significantly help me master the ability I need, e.g. interdisciplinary adaptability and analytic skills, to contribute to the manner-aware agent. Since Carnegie Mellon University has one of the most renowned machine learning and computer science programs, I would like to apply for it with my best efforts.

Recent Researches

Belows are three projects I have been working on recently: (1) a lite audio-visual speech enhancement system (LAVSE), (2) a multimodal audio-articulatory-movement speech enhancement system (AAMSE), and (3) a cross-lingual QA system. Hence I have hands-on experience in various domains in artificial intelligence, including speech, computer vision, and NLP.

I have published LAVSE with Prof. Yu Tsao and Prof. Hsin-Min Wang at the conference of INTERSPEECH. The proposed LAVSE system addresses the problem of additional processing costs and the privacy issue of facial images of general audio-visual multimodal systems. 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, and they greatly improve the applicability of DL-based models on embedded systems. From these audio-visual projects, I have learned fundamental concepts and the know-how of deep learning, while absorbing domain knowledge of both speech and visual signal processing.

To investigate other potential useful features for speech enhancement and synthesis, our team has incorporated the articulatory movement information into multimodal deep-learning based models and proposed AAMSE. We successfully used the Electromagnetic Midsagittal Articulography (EMMA) dataset to improve the performance of existing enhancement systems with different fusion strategies, and the work was submitted to ICASSP.

In the meantime apart from signal processing, I started an NLP project about cross-lingual transfer learning on a QA system with Prof. Keh-Yih Su. So far, most of the available datasets are in dominant languages, and it is undeniable that technologies mainly serve people who use these languages and lead to unfavorable inequality. The cross-lingual transfer learning project can help reduce the bias in technology with limited data in minority languages. Moreover, this project involves multilingual processing which serves as a core technique of cross-culture communication support, and paves the way to my next step of fulfilling my self-expectation by the machine learning or computer science program at Carnegie Mellon University.

Academic Interests

My ultimate research goal is to build a translation agent which can help 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 not only translate the literal meanings between different languages, but the manners among them as well.

To realize my manner-aware agent, starting by working on implications or offensiveness within corpora, 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 for the special languages existing in different kinds of situations. For example in Japanese, people should not talk to their teachers in the way they talk to their classmates. The evaluation system has to consider the chosen vocabulary in different situations, and it needs to take care of the tone as well. In addition, the level of politeness varies from cultures to cultures, even when the cultures share the same language, e.g. British English and Indian English. Keeping the consistency between evaluations and target languages while avoiding stereotypes and biases can be quite a challenge. Next for the second direction, the monolingual manner-aware agent, some research about affective conversational systems already exist. The techniques have the potential to be extended to the third direction, the multilingual systems, with some state-of-the-art machine translation systems.

Research about manners in the NLP field is not yet pretty common, so I would love the opportunity to learn under the tutelage of Prof. Graham Neubig, Prof. Ruslan Salakhutdinov, or Prof. Scott E. Fahlman. As the artificial intelligence industry prospers with the technology of machine learning becoming increasingly popular, a fair and reasonable translator is naturally expected. I am confident that the manner-aware agent has great potential to enhance cross-cultural communication.

Future Planning

My intention is to accomplish the manner-aware agent no matter where I go. Obviously, the manner-aware agent is a long-term project which I have to keep working on after completing the PhD program. Any of the three directions is a great topic for continuing research, and has the strong potential to help develop politeness-related applications. I will consider embarking on a postdoc study, but I am open-minded to take a job in industry as well. The advantage of remaining in academia is the freedom of conducting cutting-edge research, and the ideas can be more diverse and creative. However, it takes decades to develop new technology, and not to mention the implementation of products. The feedback from customers to researchers is usually pretty slow, or even unheard. The pros and cons of academy or industry are different, but as long as I can continue devoting myself to cross-culture communication support and improve my ability in this domain, my attitude is positive. A lot of research institutes and companies are already working on text to text machine translation. Due to my speech processing background, I am also interested in speech to speech translation which can preserve the emotion within speech signals and is a prospective candidate for further improvement in my manner-aware agent. Due to the globalization trend, these innovative technologies are getting more essential than ever before, and they certainly can benefit people from all around the world.

Carnegie Mellon University has a top-notch Natural Language Processing group. I would love the opportunity to use the experience at 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 PhD degree.