Overview of NLP

Alexis Jennings / CS 4395.001

A. Natural language processing (NLP) is what allows a computer to understand natural human language, as computers can only understand data in terms of 1s and 0s.
B. AI is an umbrella term that NLP falls under. NLP is only one of the many sub-areas that utilize AI technology.
C. Both natural language understanding and natural language generation are parts of NLP. However, NLU focuses on computers' reading comprehension, while NLG focuses on computers' ability to write.
D. Some examples of technology that use NLP are: email filters, smart assistants like Siri and OK Google, as well as spell check and voice text messaging.
E. The three approaches:
    a. The first of the three approaches to emerge was rules-based NLP, which began in the 1960s. As the name implies, language was processed based on a set of rules. If the system was given something not explained in the rules, the system would therefore not be able to derive an understanding of what was being said. The main problem with a rules-based approach, however, is that rules do not scale up to the complexity of human language. Some examples of the rules-based approach are spell check software, context-free grammars, and Eliza chatbot.
    b. The second of the three approaches was based in statistics and probability, and began to emerge in the 1980s. Software using this approach learned from large corpuses of data. Some applications that were developed during this time could determine word frequencies, and also used traditional machine learning algorithms. However, problems arose once again: this approach relied too heavily

on large amounts of data, and needed a large amount of processing power as well. An example that uses this approach would be word counters.

c. The third of the three approaches is deep learning, and it emerged in the 2010s. Deep learning is the current iteration of approaches to NLP, and is a type of machine learning. This approach improved results in language translation, generation, and understanding. Similar to the statistics and probability-based approach, it needs a large amount of data and processing power. Examples using deep learning are the analysis of notes and text in electronic health records, and machine translation.

F. I am interested in NLP because of my fascination with machine translation and speech-to-text software. I want to learn more about how applications can listen to a person speak, understand what they are saying, and be able to translate what they have said. I think that it would be great if I could one day help contribute to increasing the reliability of machine translators.