Overview of NLP

a. Define NLP in your own words.

Natural Language Processing (NLP) is a branch of artificial intelligence that deals with processing language. It is identifying the characteristics of human language with the end goal of providing some convenience to the user. For instance, NLP can be used to flag certain words in an email that are characteristically found in spam emails to sort out the useful emails from the useless ones. Another instance is taking the words in reviews and determining whether it is a positive or negative review based on that. There are three approaches to NLP when it comes to analyzing language in its text form: rules-based approaches, statistical and probabilistic approaches, and deep learning.

b. Describe the relationship between AI and NLP.

The relationship between AI and NLP is that NLP is a branch of AI, but not everything in AI is completely NLP related. However, a complex NLP project tends to have elements of pure NLP, machine learning (a different branch of AI) and other elements that go into the general realm of AI. Therefore, AI is a general term that refers to everything that involves AI concepts, related to AI, while NLP is just one branch of AI.

c. Write a sentence or two comparing and contrasting natural language understanding and natural language generation.

Natural language understanding is the understanding of something that another person said. However, natural language generation is the creation of a response in human language. Both are needed to communicate properly, since a person must first understand the human language before they can create a response, and in order for someone to understand a response, a response needs to first be created.

d. List some examples of modern NLP applications.

A few examples of modern NLP applications are chat bots or automated messages that can help direct you to the correct person or solve an issue that you may have; Google translate, which takes a user input and can convert it to another language; and the suggestions in ecommerce websites like Amazon, which take key words from previous searches or purchases and suggests different products that you may be interested in.

e. Write 3 paragraphs describing each of the 3 main approaches to NLP, and list examples of each approach.

Rules-based approaches to NLP is the oldest approach, and also the most limited approach of the three main approaches. The main idea is to create a set of rules that the program could look at and create a correct sentence syntactically or determine if a sentence was correct grammatically. As you would need to create more and more rules to increase the

scale the program, a rules-based approach is good for simpler cases of sentence understanding and generation as it is quite fast, but would be extremely tedious and tougher to scale for a more complex case. A few examples of the rules-based approach include using regular expressions or context free grammar. Examples of applications using a rules-based approach are Eliza (a program that would "talk" to the user through a set of regular expressions) and filters that, by using rules, can allow someone to take a set of data about candidates and filter to contain only the ones with the appropriate credentials.

Statistical and probabilistic approaches are more math based and involve performing calculations on data sets. Using these calculations, they can create language models that can better determine the meaning of human language and suggest words to a sentence that make sense grammatically. Machine learning also falls into these approaches, and their performance is based on the size of the data sets provided to it. A few examples of statistical and probabilistic approaches include classic machine learning algorithms such as small neural networks or decision trees. Examples of applications of these approaches include predictive text (predicts the next likely word based on what was previously typed) or being able to determine the actual meaning of a word in the context of the rest of the sentence rather than using its literal meaning.

Deep learning involves having extremely large data sets and are a descendent of the smaller neural networks found in statistical and probabilistic approaches. Of the 3 approaches, deep learning is the most powerful and is the approach that aims to achieve more human-like conversations. However, deep learning requires a lot of resources and may not be the best choice for smaller scale applications. A few examples of deep learning approaches include recurrent neural networks and convolutional neural networks. Examples of applications of deep learning include assistants like Siri or collecting new stories and determining whether they are true or not.

f. Write a paragraph describing your personal interest in NLP and whether/how you would like to learn more about NLP for personal projects and/or professional application.

Personally, I have been very impressed with the several new applications of NLP being released and developed. However, I had never really considered it to be something that I was interested in learning in depth. Upon reflection though, I think that it is worthwhile to at least get a good introduction to it so that I can become a more well-rounded programmer, since the applications of NLP are extremely expansive and can potentially be of use to me in later projects. At the moment, as I have an interest in app/web development, I can see where NLP could be used in those projects to provide a more immersive experience. Learning how I can apply NLP in those circumstances would be fairly interesting and useful for me so I could add more utility to an application.