Journal Report 14
01/13/20-01/20/20
Bryan Lu
Computer Systems Research Lab
Period 2, White

Daily Log

Monday, January 13

Figured out the correct arguments to pass through to the parser, got code to work up until a request to a local server, presumably a Stanford NLP server.

Tuesday, January 14

Consulted Dr. White with what to do about the Stanford server, installed the server on my computer, and successfully obtained a connection to the server and received data from it.

Thursday, January 16

Played with Stanford server with small sentences, figured out the right request to reliably return data, but not the data that I needed for the code.

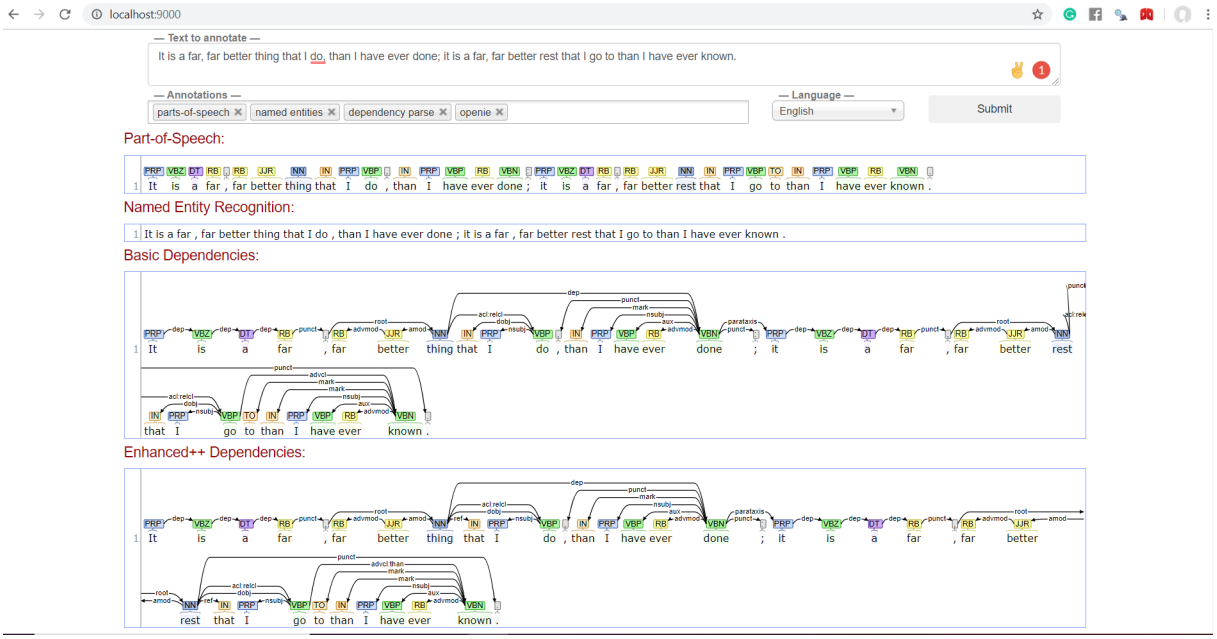
Timeline

Date	Goal	Met
1/6	Figure out what the best next steps are for moving forward, based on not quite reaching my winter goal.	I think I have a fairly concrete plan for what to work on moving forward, but I'm not sure if it's super feasible...need to discuss still.
1/13	Get a script heavily based on their code to be runnable, pending annotations.	Almost, I just need one last push to get the last piece of the code in the syntax parsing to work.
1/20	Adapt the question input for their code to access local files instead of a database, try running their code with olympiad problems.	N/A
1/27	Pending annotations, get everything up until the annotations step to work with olympiad problems, and work out other cosmetic details.	N/A

Reflection

I felt I was pretty productive this week – I'm currently working on running a snippet of code to process one of the questions in their provided testcases set to figure out what sort of structure their "syntax parses" are for each question, i.e. the data about words in each of the sentences for each question.

I finally had to configure the Stanford NLP parsing server that the paper mentioned, and I think I got it to work pretty successfully! As a proof of concept, here's an example of the Stanford parser server running on an actual sentence. This one is a bit longer than the one I've been using all week, but it's to show that it can handle long sentences just fine. This is the final quote from Dickens' *A Tale of Two Cities*, and the parser server returns part-of-speech tags and the dependency relations between different words:



This is a good step as eventually I'm going to need this information when I deal with individual questions.

Obviously, the server does other things, but I've not been patient enough/haven't properly figured out what they are and what they do. What's left for me to figure out is what function is being run on the server so that the last bit of data processing works, because in these two lines of code:

```
score = tree_data['score']
tuples = tree_data['tuples']
```

they seem to be accessing `score` and `tuples` keys in the returned dictionary, and I'm not currently aware of how these keys can be made.