Elayne Delgado

Clemmie Malley

Midterm – Project Proposal

Due: July 15, 2015

**SUMO On Protégé**

**Project Proposal**

We are proposing to complete the SUMO on Protégé project topic. The objective of this project is to use the owl file provided by Adam Pease and the Protégé application to repair the owl code so that it can be able to be used unanimously with any application. Currently the owl file does not load correctly in Protégé. For example: The MISO entries are at the highest level with Entity. Therefore there is a need to update the owl file in order for it to be read correctly through any application moving forward. Our goal is make sure that the SUMO code matches the owl code after inputting the code into python and adding the missing classes into Protégé. This will require us using python and other tools to input the missing or corrupted code.

**What is needed to complete this project?**

We will use the Protégé application to read the SUMO owl file using the url provided on Adam Pease’s website (http://www.adampease.org/OP/SUMO). This will be a project that focuses on ontology. We will need to analyze the owl file and determine the correct hierarchies, classes, instances, etc. that the unreadable data need to be placed and updated in. We may attempt using python to automate the corrections once they have been identified or we may need to use python to read instance of the SUMO owl file to allow pieces of the file at a time as it is very large. In conclusion we would need the SUMO code, Protégé, and either python notebook or wingIDE to complete this project. We will also need research and help from the internet in order to complete this.

**Preliminary Exploration**

As of now, we have opened the owl file in Protégé by accessing the url above. We found that there is a hierarchy present along with object properties and data properties. Moving forward, we plan to research online to determine if there is a simple solution to fixing the read error. If no solution is found, we plan to divide the owl file and validate each line is correctly transferred to Protégé. If a line is missing or written incorrectly we will make these adjustments in protégé and save to the owl file.

Problems that may arise may be the length of the SUMO code as it is very long. This project can be time consuming and we only have a few weeks left to complete. If we are not able to code a faster process for reading in python, length of the code could lead to an incomplete readable file.

**Project Significance**

By adjusting the SUMO to owl code our hope is this will allow for universal use of the owl file. Currently when the SUMO owl file is read in Protégé it is missing some of the SUMO code and therefore the hierarchy is coming through incorrectly. The end result of this project will be an owl file uploaded to GitHub that is corrected and accessible for other people throughout the internet to use. Additionally, we will compile a technical paper documenting our actions and conclusions from the assignment.

**Project Challenges**

We will face many challenges when adjusting the SUMO code to the owl code, because of our limited experience in programing. However, because we have been working with Protégé since the beginning of the semester and have been gaining knowledge about python little by little, we should be able to adjust the code so that everyone can be able to upload it into Protégé without a problem.

The other issue we might face is our lack of knowledge of SUMO, we will have to do extensive research on SUMO in order to complete this. The SUMO file has many lines of code and it could be very challenging to get through all of it depending on its complexity.

Also, as stated above already, time is a challenge for this project. As the SUMO file is so large it will take time to find the exact issue. Although we may be able to use Python to read through the code faster, this may not be a viable solution once we begin our project. We hope that we will be able to achieve a completed working file by the end of the semester. We will keep the professor updated.

**Project Development**

In order to start this project we have to do research on SUMO. Since it is an unfamiliar language to us, we have to know how it will work with Protégé and what we can add to make sure the owl file runs properly. Additionally, we will need to research and see if there is a known solution for solving the hierarchy issue with the owl file.

Once we become familiar with SUMO and if we have not found another solution, we will start reading through the code to see where the error is located. The url reads into Protégé easily but it is clear that there are hierarchy issues.

After loading the url into Protégé and analyzing the SUMO code we will need decipher through the hierarchies, classes, objects, etc. to see if it is clear which are missing in Protégé that should be found in the SUMO code. We will likely need to use python to read portions of the owl file at a time to identify where the errors begin and end as the file is extremely large and could take weeks to read through completely if the error is not evident.