Project Plan

Haneen Abu-Hijleh, Ravi Bullock, William Chen,Peizhou Duan, Ilya Yereferenko

Project Description:
Nova is a chatbot that is passionate about teaching astronomy to users. Her implementation of natural language processing in python allows her to understand and answer user questions related to stars, galaxies, black holes, planets and many other astronomical topics. She was created with the intent to spread basic knowledge about astronomy to individuals who are interested in the field. She can be found at https://github.com/COSC-310-Group-18/COSC310_A2

Software Development Life Cycle:
We chose to use the incremental development life cycle. This life cycle is ideal because we would like to continuously gather user feedback and make corrective changes throughout development. As part of the agile process, we will also be participating in peer programming since we all have limited experience creating chatbots and want to support each other. Our incremental life cycle will include a project planning phase, a researching phase, a development phase, a review phase and then a presentation phase. We aim to develop a basic chatbot in the first major increment then develop a GUI for the chatbot in the second major increment. Since we are utilizing an incremental life cycle, we are not assigning specific roles but rather opening and closing tickets with whatever task we want to complete and touching base regularly through a Discord server.

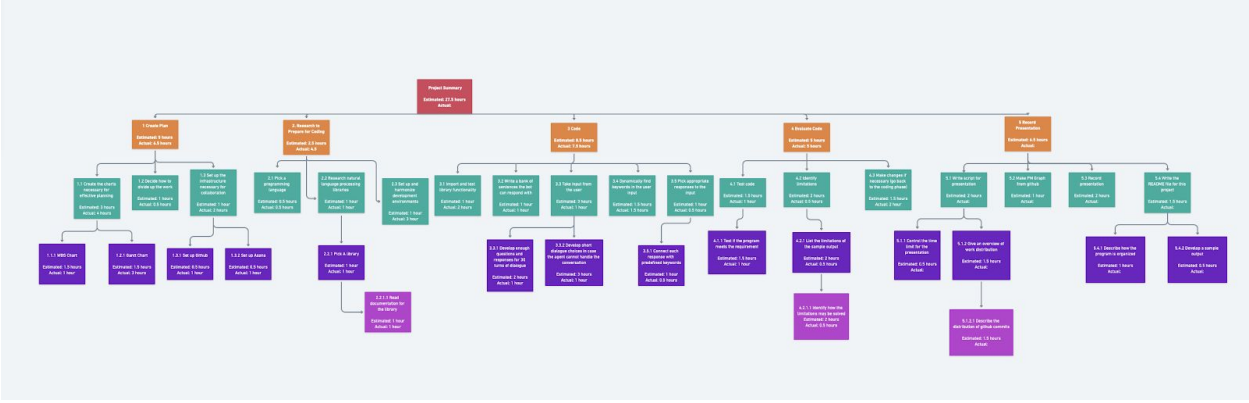
Phases of the Software Development Life Cycle:

- **Create a project plan**
    - Create the charts necessary for effective planning
        - WBS chart
        - Gantt Chart
    - Decide how to divide up the work
    - Set up the infrastructure necessary for collaboration
        - Set up Github
        - Set up Asana
- **Research to prepare for coding**
    - Pick a programming language
    - Research natural language processing libraries

        - Pick a library
            - Read documentation for the library
    - Set up and harmonize development environments
- **Code**
    - Importing and testing library functionality
    - Write a bank of sentences the bot can respond with
    - Take input from the user
        - Develop enough questions and responses for 30 turns of dialogue
        - Develop short dialogue choices in case the agent cannot handle the conversation
    - Dynamically find keywords in the user input
    - Pick appropriate responses to the input
        - Connect each response with predefined keywords
- **Evaluate code**
    - Test code
        - Test if the program meets the requirement
        - List the limitations of the sample output
            - Identify how the limitations may be solved
    - Identify limitations
    - Make changes if necessary (go back to the coding phase)
- **Record Presentation**
    - Write script for presentation
        - Control the time limit for the presentation
        - Give an overview of the work distribution
    - Make PM graph from github
    - Record presentation
    - Write the README file for the project
        - Describe how the program is organized
        - Develop a sample output

WBS Chart

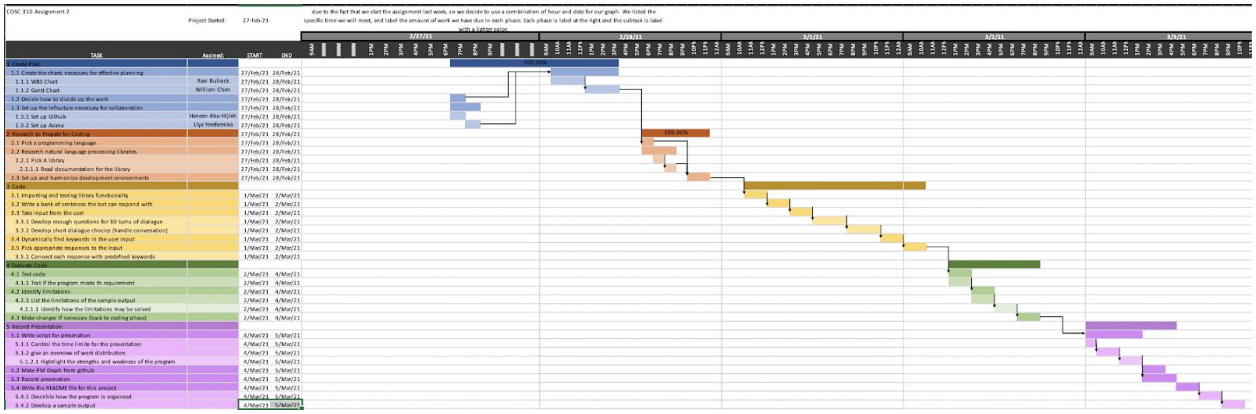
To look at our WBS chart, follow the link below to the CASE tool we used - Whimsical



https://whimsical.com/cosc-310-assignment-2-U5vrQZSGg6tFJRWSfHfUF4

Gantt Chart

The Gantt chart is quite large, and the initial excel file can be found at the github link below.



https://github.com/COSC-310-Group-18/COSC310_A2/blob/main/Gantt%20Chart.xlsx

We started this assignment on February 27th, we decided to use a combination of hour and date for our graph. We listed the specific time we will meet, and labeled the amount of work we have due in each phase. Each phase is labeled at the right and the subtask is labeled with a lighter color. The arrows indicate a dependency between tasks.

Limitations

1. Nova can take a while to process user input and output the correct responses since there is a lot of conversion and formatting to be done. The processing time could be lowered by using more efficient functions.
2. Nova is somewhat limited to the predefined questions and responses in the corpus.txt. The corpus.txt will have to be expanded to allow Nova to answer more questions.
3. Nova cannot recognize inputs that have spelling errors. This can be fixed with Porter Stemmer in the next version

Sample Output

- A sample output with 30 turns can be found in Sample Output.pdf in our Github repository
- https://github.com/COSC-310-Group-18/COSC310_A2/blob/main/Sample%20Output.pdf
- Examples of limitations may be found on the main page of our Github repository

Our links:
Asana: https://app.asana.com/0/1199995258403847/board
Github: https://github.com/COSC-310-Group-18/COSC310_A2
Whimsical: https://whimsical.com/cosc-310-assignment-2-U5vrQZSGg6tFJRWSfHfUF4