CSCE 240H FINAL PROJECT: DISEASE CHATBOT

CHRISTINE STEEGE, AZARIAH LAULUSA, RAE JONES, LUCAS AUST, JAMEL CHOUARFIA, AND CALEB BRUNSON

CSCE 240H SPRING 2023

5/2/2023

PROGRAM SCOPE

- Diseases: The chatbot can support HIV, Scabies, Rabies, Rubella, Avian Flu, Malaria, Measles, Polio, and Mumps.
- Data Sources: CDC and WebMD articles for each of the diseases.
- Language: Java

CAPABILITIES

- Detects and answers questions related to:
  - The information sections on the CDC and WebMD webpages for the supported diseases
  - Requests for all the information found for a disease
  - Multiple questions in a single prompt
  - Common small talk questions
- Uses context of previous questions to infer disease if no disease name is provided
- Can be downloaded and ran as .jar application on any operating system
- Detects and corrects for simple spelling errors

OBTAINING CONTENT

- HTML parsing is used to detect section headers and section content from the from the WebMD or CDC webpages.
- This content is written to a formatted text file.
  - Add the section header with delimiter to the top of the section
  - Add the section content.
  - Add the source of the data to the bottom of the section.

```
<div>
  <h3 id="whatismeasles">What is measles?</h3>
</div>
<div class="float-md-right col-md-4 cr ml-md-3 pt-2"> … </div>
<p>
  "Measles is a disease caused by a highly contagious virus. People with measles
  spread the virus through the air when they cough, sneeze, or
  breathe.  "
</p>
<p>
  <a href="https://www.cdc.gov/measles/symptoms/signs-symptoms.html">Symptoms</a>
  " of measles include high fever, cough, runny nose, red and watery eyes,
  and rash. Koplik spots (tiny white spots inside the mouth) can appear 2 to 3
  days after symptoms begin. Some people who become sick with measles also get a
  serious lung infection, such as pneumonia. Although severe cases are rare,
  measles can cause swelling of the brain and even death. Measles can be
  especially severe in infants and in people who are malnourished or who have
  weakened immune systems."
```

```
<SECT>What is measles?
Measles is a disease caused by a highly contagious virus. People with measles
spread the virus through the air when they cough, sneeze, or breathe.
(Source: https://wwwnc.cdc.gov/travel/diseases/measles)


<SECT>Symptoms
Symptoms of measles include high fever, cough, runny nose, red and watery eyes,
and rash. Koplik spots (tiny white spots inside the mouth) can appear 2 to 3 days
after symptoms begin. Some people who become sick with measles also get a serious
lung infection, such as pneumonia. Although severe cases are rare, measles can
cause swelling of the brain and even death. Measles can be especially severe in
infants and in people who are malnourished or who have weakened immune systems.
(Source: https://wwwnc.cdc.gov/travel/diseases/measles)
```

DATA INTEGRATION AND OPTIMIZATION

- Using a class called *DiseaseDataProcessor*.
  - Processes data from extracted files, WebMD and CDC, based on the disease and question.
  - Checks if the program has an extracted disease file for the user's disease.
    - Stores prompts and corresponding data into a HashMap.
  - Looks through prompts from stored data and gives the best answer using HashMap.
    - Questions that involve both websites will be given the answer with the most information.
    - If the stored data from the website text files match with the prompt/user's question.

HANDLING USER INTERACTION

Spelling Correction

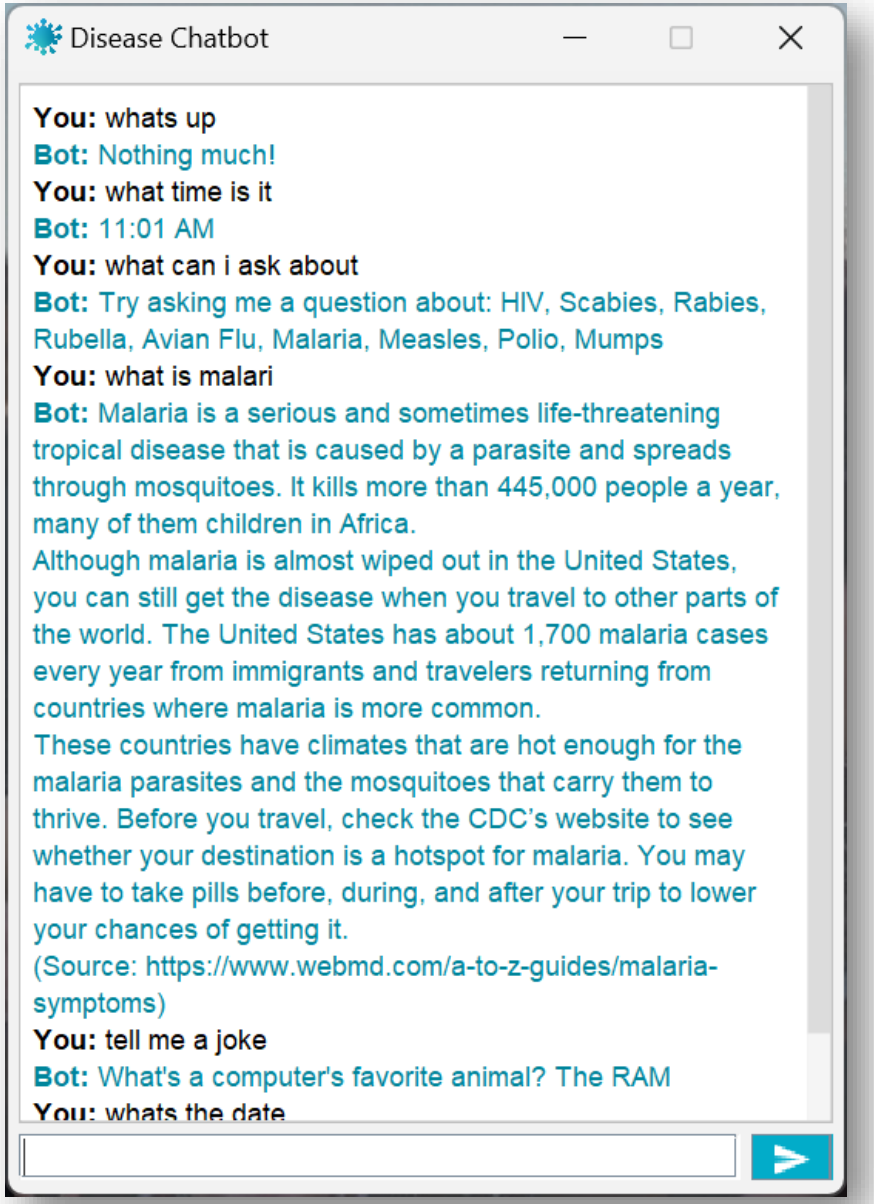
- Using an outside library, input is corrected based on keywords.

Disease Chatbot

- Intent is detected using regex matching for keyword patterns for each prompt and potential small talk questions
- Any detected prompts are answered using the extracted information for the provided disease name (or context of previous questions if no disease was provided).

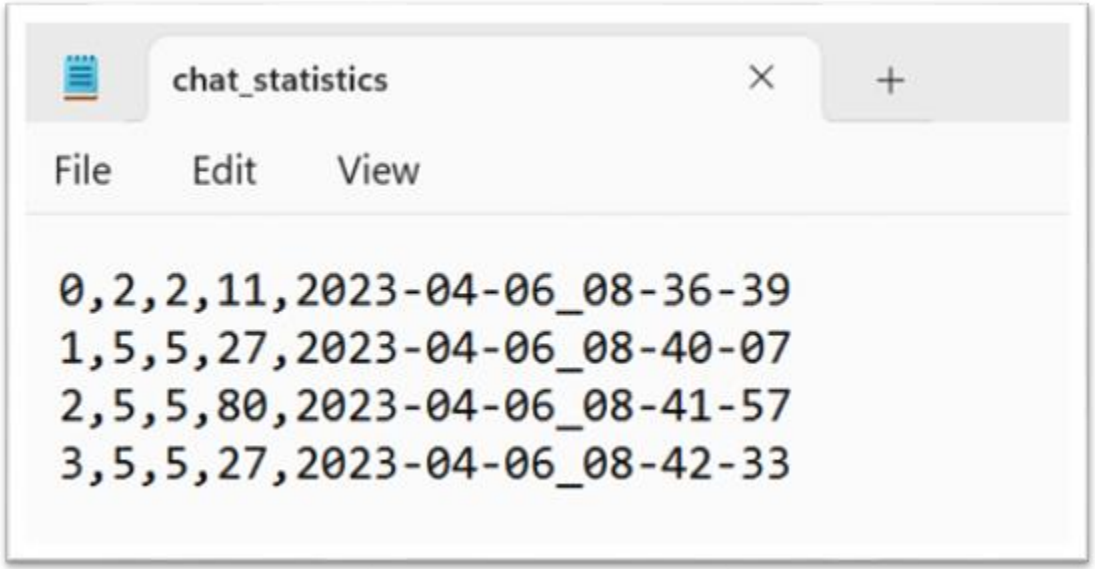
Chatbot GUI

- The user provides their prompt through a simple GUI window which also displays the generated bot response.



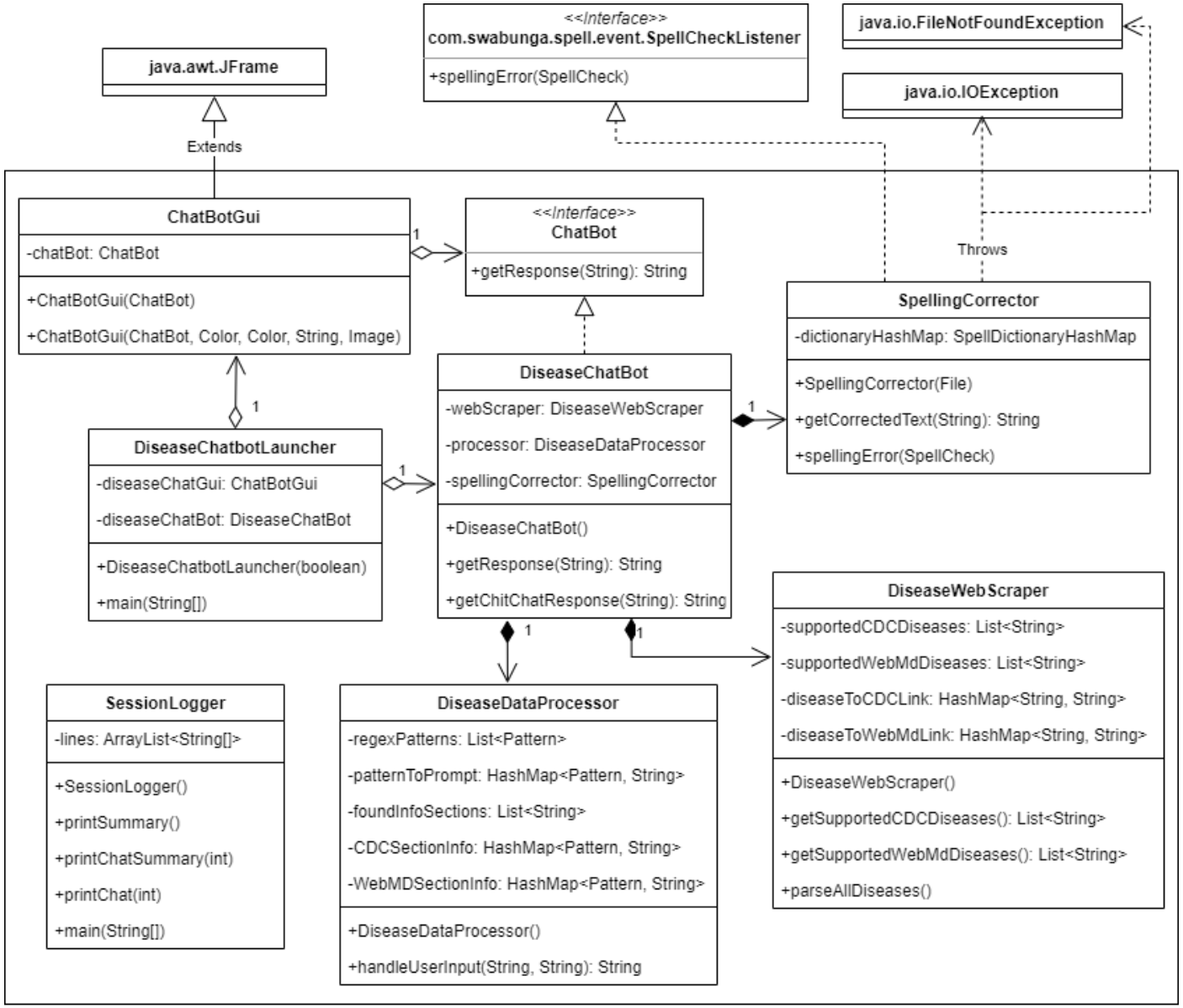
RECORDING SESSION STATISTICS

- Stores user/system utterances and duration during chat session
- Imports info into CSV file
  - Uses columns of CSV to calculate
  - Individual chat summary data
  - Total system/user utterances and duration
  - Overall chat summary data



INTEGRATION

- Integration was straightforward as we decided to use Christine's chatbot as the base
- Our plan at the beginning was to combine separate parts of everyone's program
- However, we decided to just stick with and test one of the bots already made since it already had most of the core functionality



EXPERIENCE DEVELOPING THE CHATBOT

Lucas

- This project challenged my critical thinking, I had to think about problems in a multitude of ways. There were many ways to solve the same problem and identifying the best one was a challenge for me.

Rae

- I enjoyed collaborating with the others to combine elements from everyone. Being able to see different methods to achieve similar products broadened my coding horizon.

Azariah

- This project challenged me and helped me think about computer science in a fun and innovative way. I love the aspect of learning and applying class concepts to projects like this.

Christine

- This project was an exciting opportunity to incorporate outside resources into my programs and taught me how to learn and troubleshoot new tools much more effectively. This project also allowed me to gain experience more formally documenting my code and taught me how to effectively communicate code design with other people.

Jamel

- This project was fun since I got to see how I've improved over the semester. I can now do stuff that I had zero clue on how to do before.

GITHUB LINK

HTTPS://GITHUB.COM/JAMELCHOUARFIA/CSCE240FINALPROJECT