

Daniyah Arshad

CS4395.001

09 April 2023

## CHATBOT - Report and Evaluation

### **System Description**

This chatbot is intended to be able to hold limited conversation about color theory, to the extent of discussing how colors are perceived and what their generalized meanings are in the context of art and graphic design, color schemes and the ways they are used effectively in art, and finally, general fun facts about color theory, such as the history of how it was developed or what color theory truly is. This chatbot's system was developed using NLP techniques, using a rules based approach to build the dialogue flow of the system.

The chatbot introduces itself, and the user is then prompted to greet the chatbot in turn. Then, the chatbot asks for the user's name. The chatbot then parses through the user's response to find the name. After this, the chatbot creates a file with the users name, to store information about that user. Before that though, it checks if a file already exists for the user. If it does, it asks a question to the user based on their past conversations. Afterwards, the chatbot proceeds as normal.

The chatbot then asks the user what they would like to talk about, offering the three options mentioned before: color meanings, color schemes, and fun facts. If the user says something that is not one of these three options, then the chatbot apologizes, and searches the web and offers a response based on the results.

However, if the user responds with one of the topics mentioned, then the chatbot initiates a conversation going down the chosen route. It asks a question about the user's favorite color if it is the first option, and parses through the user's response to find the "basic" color version of their response. For example, lavender becomes purple, cerulean becomes blue, rose becomes pink, etc. For the color scheme route, it asks for the user's favorite color scheme. In both routes, it then records the user's response down in their file. Afterwards, it responds with either the color meaning or color scheme definition. After the user responds, it runs a sentiment analysis on the user's response and responds based on it. In the case of a fun fact, it randomly chooses a topic and presents a fun fact, and the user is given the opportunity to respond. The chatbot runs a sentiment analysis and responds based on it.

Finally, if the user enters some form of a goodbye in the main loop of the chatbot, the chatbot recognizes this as a decision to exit the chatbot and ends running, leaving the user with a goodbye.

### **Specific NLP Techniques**

The specific NLP techniques utilized in the development of this chatbot are using NLTK tools, parts of speech tagging, TF-IDF, sentiment analysis, synsets, and also using a web scraper to get miscellaneous information for the user. NLTK tools are fundamental to the development of this bot, as it gives the ability to parse through the responses and the data in a way that is manageable. The most used tools are the tokenizer tools, as the data file is tokenized into sentences, and the responses are tokenized into words, and are then further processed by the bot. However, the lemmatizing tool is also important in converting the data to a workable form for the TF-IDF vectorizer.

The parts of speech tagging tool is used primarily in detecting the users name when the chatbot asks for it. Names are categorized as proper nouns, and when responding to a question about one's name, the name should be the only proper noun in the sentence. Thus, the chatbot tokenizes the sentence, tags it with parts of speech tags, and then searches for the proper noun. Once it finds the proper noun, it records it as the person's name. Thus, the parts of speech tagging tool is important in ascertaining the user's identity.

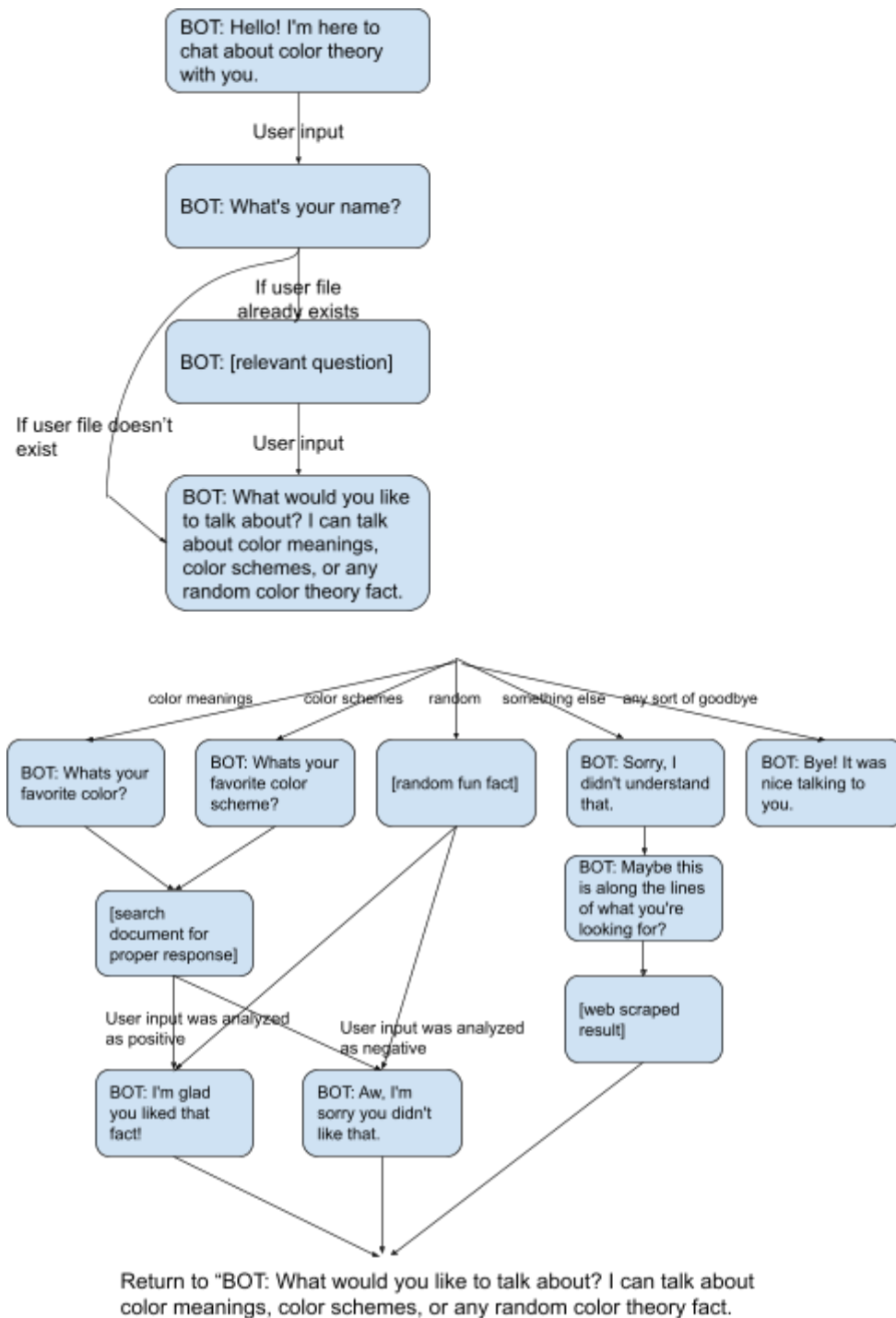
The TF-IDF vectorizer is instrumental in searching the text document for the specific factoid or piece of information requested, as it checks for the term frequency against the remaining document and returns a numeric value based on whether or not that result is available. Should it not be available, then the bot apologizes and lets the user know that it doesn't understand.

The sentiment analyzer tool, using sentiwordnet, analyzes each individual word of the users response after it has been tokenized using the NLTK built in function, and generates a positive and negative score for the response as a whole. Depending on whether the positive or negative score is greater, the bot responds differently.

Synsets are incredibly interesting, providing the placement of a specific word in a very thoroughly developed hierarchy. They are used here to determine whether the color is a general color (pink, blue, red, etc.) or a more niche color (lavender, crimson, cerulean, etc.). If the color is a more niche color, then the database will not have information on that specific color, however, using the synset hypernym function, the chatbot can find the general version of that niche color, and thus give the user an adequate response based on the given input.

Finally, web scraping is used using beautiful soup to parse through a google search query of a user's response to look up and generate a response based on what the user may be looking for, even if it may not be available in the constraints of the chatbot's dataset. The NLTK sentence tokenizer is used to further parse through the scraped web result data to find a legible response to return to the user.

## Diagram of Dialogue Tree



### Sample Dialogue Interactions

BOT: Hello! I'm here to talk to you about color theory.

USER: Heya!

BOT: What's your name?

USER: Dave

BOT: Oh, I remember you! You learned about Newton last time, didn't you?

USER: Yes, I did!

BOT: What would you like to talk about? I can talk about color meanings, color schemes, or any random color theory fact.

USER: Color meanings.

BOT: What's your favorite color?

USER: Cerulean

BOT: Blue is the color of the ocean and the sky; it often symbolizes serenity, stability, inspiration, or wisdom.

USER: Wow, that's cool!

BOT: I'm glad you liked that fact!

-----

BOT: Hello! I'm here to talk to you about color theory.

USER: Hello.

BOT: What's your name?

USER: Bethilda

BOT: What would you like to talk about? I can talk about color meanings, color schemes, or any random color theory fact.

USER: Color schemes.

BOT: What's your favorite color scheme?

USER: Complementary

BOT: A complementary color scheme is when you use opposite color pairs to maximize contrast.

USER: That sucks. I already knew that.

BOT: Aw, I'm sorry you didn't like that fact.

-----

BOT: Hello! I'm here to talk to you about color theory.

USER: Hello.

BOT: What's your name?

USER: Lola

BOT: What would you like to talk about? I can talk about color meanings, color schemes, or any random color theory fact.

USER: A random color fact sounds cool!

BOT: Did you know A color wheel or color circle is an abstract illustrative organization of color hues around a circle, which shows the relationships between primary colors, secondary colors, tertiary colors etc.

USER: Huh. Well, that sure is a fact.

BOT: Aw, I'm sorry you didn't like that fact.

## **Appendix for Knowledge Base**

The knowledge base was composed of information compiled across several websites discussing color theory, specifically using articles from the invisionapp website and the interactive-design website. These two websites had articles that had general, concise information that was helpful in constructing a cherry picked knowledge base of all the information I wanted to have, though much of the information was trimmed down and sometimes rephrased just to keep it more concise and legible. Otherwise, the information would be a bit overwhelming for a simple chatbot.

The information itself were small blurbs of information, discussing various aspects of color theory. Small, digestible facts that were easy to comprehend and condensed enough to not be presented as a large paragraph if a chatbot were to print it.

The specific sources are:

<https://www.invisionapp.com/inside-design/understanding-color-theory-the-color-wheel-and-finding-complementary-colors/>

<https://www.interaction-design.org/literature/topics/color-theory#:~:text=Color%20theory%20is%20the%20collection,%2C%20psychology%2C%20culture%20and%20more.>



### **Appendix for Sample User Models**

The sample user models were personally constructed by myself, using my knowledge of the chatbot to construct different scenarios that could test both the limits and the strengths of the chatbot, helping me further develop the chatbot. The different user models stored information on what colors and color schemes the fake users liked, as well as what the fake users learned about if they chose to learn about a random fun fact. The information is formatted in such a way that if the user decides to chat with the chatbot again, the chatbot can quickly formulate a returning user response based on the most recent information on the user's file.

### **Evaluations of Chatbot and Analysis of Strengths and Weaknesses**

After a thorough evaluation of the chatbot from multiple perspectives, there are a few conclusions. The chatbot is only able to respond based on its rigid framework, which is overall a weakness, however, it is a strength in that it establishes consistency and makes it easier for users to know what to expect out of a response. Additionally, while the chatbot is a rules based chatbot and thus, cannot deviate from its script, it uses more advanced techniques to parse through the user inputs to better determine what the user intended, thus partially making up for its greatest weakness. Finally, the chatbot searches the web for a response if it does not understand what the user asked, which can be frustrating if the user simply made a typo, but can be interesting to experiment with if the user doesn't want to answer within the given prompts.