**Assignment 4**

**Chat Bots**

Melissa Hunfalvay

Email: [Melissa.Hunfalvay@gmail.com](mailto:Melissa.Hunfalvay@gmail.com)

Data 650 9040

Summer 2022

Professor Ozan Ozcan

Date: August 10th, 2022

**Table of Contents**

Introduction………………………………………………………………………………………..3

Dialog Implementation and Implementation Discussion………………………………………...13

Integration with Watson Discovery Service……………………………………………………..26

Conversation data analysis experiment…………………………………………………………..34

Analysis and Enhancement………………………………………………………………………43

Conclusion……………………………………………………………………………………….47

References………………………………………………………………………………………..50

Appendix…………………………………………………………………………………………53

Appendix A: CISS Survey……………………………………………………………………….54

**Introduction (12 pts)**

*Define the purpose of the chatbot and the conversation requirements. Who is an end-user? How will an end-user interact with the bot? What are the expected benefits of the chatbot you are going to build?*

A chatbot is a computer program that simulates human conversations. This allows humans to interact with digital devices as if they are communicating with a real person (Oracle, n.d.).

The purpose of chatbots, in general, is to create efficiencies in operations, offer convenience for customers (with 24/7 availability) via added services and interactions with the business. Chatbots can help businesses differentiate themselves by scaling services, personalizing interactions and being proactive in their interactions with customers. Chatbots can allow customers to have an unlimited number of conversations with multiple people at the same time. According to consumer research, 54% of customers prefer chatbots as their method of communicating with a business with the main reason listed as the chatbot providing instant responses and answers to simple questions (Jovic, 2022). For example, banking chatbots save an average of four minutes per inquiry compared to traditional call centers (Oracle, n.d.).

The specific purpose of a chatbot depends on the use case. They are used in banking, customer service (e.g., ordering event ticket sales, checking into hotels), customer success, IT service management such as help with password updates, outage alerts, system status. Businesses are using chatbots for onboarding new employees, helping employees with routine activities and ordering supplies.

The chatbot built for this assignment is based on an industry wide, validated survey called the Convergence Insufficiency Symptom Survey (CISS; Borsting, Rouse, Michell, et al., 2003, Appendix A). The CISS is used by health care professionals, most often optometrists, to assess the visual status of their patient (the end user). The patient interacts with the bot in one of two ways:

1. On the optometrist’s website as a pre-screening to determine if a visit is needed
2. At the optometrists as part of a routine medical exam, in the waiting room prior to seeing the doctor

The benefits of building a CISS specific chatbot are:

1. It is a scientifically validated survey for examining convergence insufficiency (CI; Borsting, Rouse, Mitchell, 2003).
2. CI is an underrepresented issue that often goes untested unless explicit reasoning in an annual optometric exam. Symptoms on the CISS chatbot could trigger a reason to test.
3. By having the chatbot on the optometrist’s website it increases awareness of CI and allows access to all patients to test at any time to determine if they are suspected of having a CI. This is important as a CI is one factor in assessment of a head injury which is an event that can occur at any time causing symptoms that were previously not present to manifest.
4. By having the chatbot available in the waiting room it ensures all patients have taken the survey as part of the pre-screening process, prior to seeing the doctor.

CI is a very serious issue and a chatbot with availability to all could be critical to awareness and follow-on care. The prevalence of persons with CI is up to 15% of the United States population according to Healthline. This is one in 6 school-aged children (Fox Vision Development Center, 2021). Hence, many millions of persons in the United States alone are affected by abnormal CI issues and associated conditions.

Symptoms associated with CI are numerous and varied. Common symptoms include headache, blurry vision, visual fatigue and tearing eyes. Binocular vision is also important for the ability to perceive depth and relationships between objects, such as seeing which object is closer than another.

Secondary symptoms may include overall fatigue, and inability to concentrate, irritability, short attention span, burning of the eyes, motion sickness, vertigo and a general “poor performance” on tasks.

According to the Dizziness and Headache Optometry Center, as the eyes are not working well together in the person with BVD, then there are also compensatory physical behaviors that occur. These include head tilts, face pain, and pain in the upper back, shoulders, and neck. Changes in gait can also be observed (The Dizziness and Headache Optometry Center, 2021).

Specific issues associated with reading may include sleepiness when reading, slowness, word skipping, line skipping, frequent loss of space, omission of words and words moving on the page. Children with undiagnosed CI issues are more often reading below grade level.

Symptoms can tend to get worse towards the end of a day due to fatigue or when using electronic equipment for an extended period.

According to the Lucile Packard Children’s Hospital at Stanford University, symptoms in children can include an inability to focus which can develop into secondary vision loss (Lucile Packard Children’s Hospital at Stanford University, 2021).

Fitful and restless sleep can occur with CI problems. A lack of proper sleep can also exacerbate symptoms creating a vicious cycle of dysfunction according to the Dizziness and Headache Optometry Center ((The Dizziness and Headache Optometry Center, 2021). Such dysfunction can also escalate to panic attacks when driving and even agoraphobia (uneasiness in large places).

These CISS symptoms are serious! And prevalence rates are high. Of the 32.3% of the population with the dysfunction, 56% are symptomatic and are therefore suffering the effects of this disorder (The Dizziness and Headache Optometry Center, 2021).

Access to a chatbot that examines symptoms and provides follow-on steps can be a critical piece in the healthcare of millions of patients. Therefore, the conversation requirements for this chat bot are:

1. To examine if patients have symptoms
2. To determine if those symptoms may be related to a CI
3. If the symptoms are related to a CI to recommend next steps in the medical process
4. If a person has symptoms but they are not related to a CI then provide other recommendations for care

*Use Visio to build the dialog flow diagram. Discuss at least 5 dialog scenarios.*

Figure 1 shows the dialog flow for the CISS chatbot.

Diagram

Description automatically generated

Figure 1: Dialog flow for CISS Chatbot

*Scenario 1:* Mr. Vision introduces himself and prompts for the user’s name. Then asks if the user has any vision symptoms. The user inputs a vision symptom that is recognized by the chatbot and provides a recommendation on what to do next. Then says goodbye. Figure 2 shows this pathway in red.

A specific example of this scenario is: Name = Melissa. Vision problem = Unable to see. Vision symptom = double vision. Recommendation = Go to the emergency room. Goodbye.

*Diagram

Description automatically generated*

Figure 2: Scenario 1 pathway through the flow chart.

*Scenario 2:* Mr. Vision introduces himself and prompts for the user’s name. Then asks if the user has any vision symptoms. The user inputs a vision symptom that is not recognized by the chatbot. The chatbot prompts again for a vision symptom until a recognized symptoms is entered. Then the path continues from scenario 1. Then says goodbye. Figure 3 shows this pathway in red.

A specific example of this scenario is: Name = Melissa. Vision problem = Hurt leg. Chatbot response = one of three random rotating variables, stating the response is not understood. Prompt again for vision problem. Vision problem = Unable to see. Vision symptom = double vision. Recommendation = Go to the emergency room. Goodbye.

Diagram

Description automatically generated

Figure 3: Scenario 2 pathway through the flow chart.

*Scenario 3:* Mr. Vision introduces himself and prompts for the user’s name. Then asks if the user has any vision symptoms. The user does not have any vision symptoms. The chatbot asks if they have seen a doctor for their annual vision screening. User says they have. Then chatbot says goodbye. Figure 4 shows this pathway in red.

A specific example of this scenario is: Name = Melissa. Vision problem = Feeling good. Chatbot response = have you seen your optometrist this year? User = yes. Recommendation = Excellent. Make sure to keep up the annual visits. Goodbye.

*Diagram

Description automatically generated*

Figure 4: Scenario 3 pathway through the flow chart.

*Scenario 4:* Mr. Vision introduces himself and prompts for the user’s name. Then asks if the user has any vision symptoms. The user does not have any vision symptoms. The chatbot asks if they have seen a doctor for their annual vision screening. User says they have not. Then chatbot recommends they get an annual exam. Then chatbot says goodbye. Figure 5 shows this pathway in red.

A specific example of this scenario is: Name = Melissa. Vision problem = Feeling good. Chatbot response = have you seen your optometrist this year? User = no. Recommendation = Your vision is an important part of your health. Make an appointment with your optometrist as soon as you can. Goodbye.

*Diagram

Description automatically generated*

Figure 5: Scenario 4 pathway through the flow chart.

*Scenario 5:* Mr. Vision introduces himself and prompts for the user’s name. Then asks if the user has any vision symptoms. The user does not have any vision symptoms. The chatbot asks if they have seen a doctor for their annual vision screening. User says does not give an understood response. The chatbot repeats the question until an understood response occurs. Then chatbot follows the path from scenario 3 or 4, depending on the user’s response. Then chatbot says goodbye. Figure 6 shows this pathway in red.

A specific example of this scenario is: Name = Melissa. Vision problem = Feeling good. Chatbot response = have you seen your optometrist this year? User = tree. Chatbot states the response is not understood. Chatbot response = have you seen your optometrist this year? The chatbot asks if they have seen a doctor for their annual vision screening. User says they have not. Then chatbot recommends they get an annual exam. Then chatbot says goodbye.*Diagram

Description automatically generated*

Figure 6: Scenario 5 pathway through the flow chart.

*Scenario 6:* Mr. Vision introduces himself and prompts for the user’s name. Then asks if the user has any vision symptoms. The chatbot does not recognize the user’s input and repeats asking if the user has any vision symptoms. The user inputs a vision symptom that is recognized by the chatbot and provides a recommendation on what to do next. Then says goodbye. Figure 7 shows this pathway in red.

A specific example of this scenario is: Name = Melissa. Vision problem = blfIUF. Chatbot = can you please reword. Vision problem = Unable to see. Vision symptom = double vision. Recommendation = Go to the emergency room. Goodbye.

Diagram

Description automatically generated

Figure 7: Scenario 6 pathway through the flow chart.

**Dialog Implementation and Implementation Discussion (25 pts)**

*Overarching development*

Mr. Vision chatbot has five parent nodes and 3 child nodes (see figure 8). The overarching conceptual development of Mr. Vision is based on the CISS scientifically validated survey (Borsting, Rouse, Mitchell, 2003). The dialog component of the conversation, seen in Figure 8, uses intents and entities which are used to gather information to enable Mr. Vision to assist the user.

*Graphical user interface, application

Description automatically generated*

*Graphical user interface, application

Description automatically generated*

Figure 8: Shows the parent and child nodes of Mr. Vision

Mr. Vision has a total of eleven entities with at least three synonyms per entity (see figure 9). An entity is “a term or object that is relevant to the intent” (Azraq, Aziz, Nappe et al., 2017). In other words, an entity provides context for the intent. It answers the question of “how”. How will the user’s goal be achieved? Entities always begin with an @ symbol in IBM Watson.

Graphical user interface, application, table, Excel

Description automatically generated

Figure 9: Mr. Vision *initial* list of entities with at least three synonyms per entity.

System entities are a type of entity that the IBM Watson Assistant Conversation service offers. The purpose of system entities is to extract meaning from common entities for example, time, System entities are toggle switches and always state with @sys-time. See figure 10 for the system entities used by Mr. Vision. See figure 11 for an example of the @system-number entity being used in a conversation.

Graphical user interface, application, Teams

Description automatically generated

Figure 10: System entities used by Mr. Vision

Graphical user interface, application, Teams

Description automatically generated

Figure 11 Shows the use of the @sys-number system entity during a conversation

Mr. Vision has four intents with at least five examples per intent (see figure 12 for a portion of the initial intents in Mr. Vision). An intent “represents the purpose of a user’s input” (Azraq, Aziz, Nappe et al., 2017). In other words, intents represent the “what” in a conversation. What is the user looking to achieve by engaging with Mr. Vision?

A screenshot of a computer

Description automatically generated with low confidence

Figure 12: A portion of Mr. Vision’s intents.

Synonyms are used with both intents and entities for the purpose of providing an understanding of natural language by Mr. Vision. For instance, in English there are many ways to say “goodbye”, such as “bye”, “later”, “bye-bye” and even in the United States to said “adios” a Spanish goodbye (see Figure 12). All these synonyms are important for Mr. Vision to recognize to provide a larger context of understanding and meaning to the conversation.

Mr. Vision also contains context variables. A context variable enables greater meaning in natural language as the context in which a sentiment occurs matters to the interpretation of meaning. Context variables allow Mr. Vision to store information and continue passing it across different dialog nodes. One example of a context variable in Mr. Vision is the users name ($name). Using a name is personal and allows Mr. Vision to personalize the conversation in the beginning and when saying goodbye.

Slots are used to customize the information Mr. Vision needs to response to a user within a single node. Mr. Vision uses slots to check for CISS symptoms in entities and save them as context variables.

Handlers are used within slots to manage input that does not meet the slot check for condition. Mr. Vision uses handlers in the “Have a vision problem” node (see Figure 13) to stop prompting for the vision symptom if #goodbye intent is recognized in the input. These are also regular expressions are used to parse user input.

Graphical user interface, application

Description automatically generated

Figure 13: Example of handler and logic for Mr. Vision

Condition responses are used throughout Mr. Vision. A condition response triggers a nodes execution. For instance, in Mr. Vision the welcome condition triggers the dialog to start from this node. The response triggers a message to the user. Responses may vary to include text, images, pauses, options or more.

The response to each parent node in Mr. Vision has at least three variations, this provides some depth to Mr. Visions conversation. Variations can take the form of text, images, options, or pauses. These will be shown in the more detailed explanation of the nodes, in the next section.

*Dialogue Scenarios*

The following section will take a more detailed look at Mr. Vision’s dialog scenarios to include screen shots of try it out panel to illustrate the input and training the bot.

Mr. Vision introducing himself and his mission in the welcome (parent) node (Figure 14).

Graphical user interface, text, application, email

Description automatically generated

Figure 14: Welcome parent node to Mr. Vision showing options for responses.

Within the welcome node there are three different text responses that are randomly set so that Mr. Vision adapts and shows different messages each time. This is seen in the “try it out panel” (Figure 15 and 16). Furthermore, an image of convergence insufficiency is part of the welcome message.

Graphical user interface, application, Teams

Description automatically generatedGraphical user interface, application, Teams

Description automatically generated

Figures 15 and 16 show two different greeting messages from the welcome parent node

In scenario 3 (figure 4 flow diagram) Mr. Vision follow the following path: Name = Melissa. Vision problem = Feeling good. Chatbot response = have you seen your optometrist this year? User = yes. Recommendation = Excellent. Goodbye node = Vision is an important part of life. Have a good day. This flow is seen in Figure 17 (after prompting for the user’s name). This dialog flow demonstrates the use of context variables ($name), option responses (yes/no), intents, and entities.

Graphical user interface, application, Teams

Description automatically generated

Figure 17: Scenario 3 dialog walk through as seen in the “Try it out panel”.

Should a user put in a symptom that is not understood by Mr. Vision it will direct the dialog to the “anything else” parent node (see Figure 18). The anything else node is designed to allow Mr. Vision to always respond to the user, even if they do not understand what to do. In this case, adding a symptom called “leg” shows a response of “Can you reword your statement? I am not understanding.” A pause of 10,000 milliseconds was added to this section to allow time for Mr. Vision to search for answers via the Web Crawl (Figure 18).

Graphical user interface, text, application, email

Description automatically generated

Figure 18: shows the pause used in the vision problems node along with text options.

Once this is tried a few times I wanted to exit Mr. Vision to stop the circular conversation, and so I typed “bye”. Mr. vision recognized “bye” as a synonym from the “goodbye” parent node and took the action to prompt that node to give a goodbye message and end the conversation (figure 19). This is an example of scenario 6 (figure 7 above). This dialog flow demonstrates the use of intents, entities, synonyms, conditional responses, images, options, and text response options.

Graphical user interface, application, Teams

Description automatically generated

Figure 19: Shows scenario 6 whereby a response is not understood by Mr. Vision and eventually the user escapes by saying bye

Should a person have recognizable vision symptoms from the CISS survey, then there are various options and next steps for healthcare. Table 1 shows the CISS symptoms and follow-on options.

Table 1: CISS Symptom and Next Step in the Healthcare Process

|  |  |
| --- | --- |
| Symptom | Next step |
| Double vision | Consider going to the optometrist as soon as possible |
| Hurt, headache, cannot focus, slow | Take over the counter eye drops to moisturize your eyes. Consider seeing a doctor. |
| Tired, fatigue | Sounds like you need some rest. Consider going to bed early tonight and see how you feel tomorrow. |
| Tugging, cannot remember | You might have a convergence problem. Try getting some rest and if the symptoms persist make an appointment with your optometrist. |

Figure 19 (a, b, c) shows several of these examples as tested out in the try it option panel. These are examples of the scenario 1 dialog. This dialog flow demonstrates the use of intents, entities, slots, (to customize the information Mr. Vision needs to response to a user within a single node) and handlers (used within slots to manage input that does not meet the slot check for condition).

Graphical user interface, application, Teams

Description automatically generated Graphical user interface, application, Teams

Description automatically generated

Figure 19 (a) shows tired as a symptom and the response is recommended rest. Figure 19 (b) shows headache as the symptom and eye drops as the next step.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 19 (c) shows tugging as the symptom and the response is a possible convergence issue.

To ensure Mr. Vision could accept email addresses from users’ email was added as a new entity (Figure 20). This allowed the users name or email to be stored in the welcome message (see figure 21).

Graphical user interface

Description automatically generated

Figure 20: Parsing email as a new entity with a pattern and synonyms

Graphical user interface, text

Description automatically generated

Figure 21: Shows recognition of an email pattern and

To train Mr. Vision, several methods were employed. First, while using the try it panel should a response be classified as irrelevant the drop-down menu next to irrelevant was uses to select the appropriate intent. Watson then trained on this message and when training was complete the same process was redone to ensure the correct intent was associated with the user’s input. This can be seen in figures 22 (a, b, c) using the input word “tugging” as a symptom.

Graphical user interface, application

Description automatically generatedGraphical user interface

Description automatically generated with low confidence

Figure 22 (a) shows that the input tugging was originally unable to be classified by Mr. Vision. A drop down-menu was used to connect the word with the vision problem intent. Figure 22 (b) shows how that vision intent was used to correctly associate “tugging” with a #vision\_problem and @ciss\_symptom.

Graphical user interface, text, application

Description automatically generated

Figure 22 (c) shows the next time tugging was used in the try it out panel, Watson recognized this word as being associated automatically with #vision\_problem and @ciss\_symptom.

A second method for training Mr. Vision was more manual and included the leading and training question to prevent null pointer exceptions when obtaining user input on their name. Using JSON, the name node was customized the capture the $name variable.

A third method of training included adding more synonyms for intents and entities. This can be done manually or when checking recommended synonyms in Watson Assistant (see Figure 23).

Graphical user interface, application

Description automatically generated

Figure 23 shows a list of synonyms for the goodbye node. User examples can be added int eh text box at the top of the screen or inputted via a csv file.

**Integration with Watson Discovery Service (10 pts)**

* *Follow the provided instructions and the Chatbot session recording to build Web Crawl documents repository in Watson Discovery service and to create the search skill that retrieves the relevant documents from the repository.*

Several sites chosen for the Web Crawl were:

1. The Mayo clinic convergence insufficiency webpage: <https://www.mayoclinic.org/diseases-conditions/convergence-insufficiency/symptoms-causes/syc-20352735>
2. The Mayo clinic home page: <https://www.mayoclinic.org/>
3. Lakeline vision which is a clinic specifically for CI issues: <https://www.lakelinevision.com/what-are-the-signs-of-binocular-vision-dysfunction/>
4. Vision Source blog on the symptoms of Binocular Vision disorders of which CI falls under: <https://www.mayoclinic.org/diseases-conditions/convergence-insufficiency/symptoms-causes/syc-20352735>
5. An academic site that trains providers on how to assess for CI issues: <https://study.com/academy/practice/quiz-worksheet-convergence-insufficiency.html>
6. The American Optometric Association (AOA): <https://www.aoa.org/healthy-eyes/eye-and-vision-conditions/convergence-insufficiency?sso=y>
7. Optometry Association: <https://www.optometrists.org/vision-therapy/vision-therapy-for-children/convergence-insufficiency-2/do-i-have-convergence-insufficiency/>
8. Cedars Sinai Hospital system: <https://www.cedars-sinai.org/health-library/diseases-and-conditions/c/convergence-insufficiency.html>
9. Ferris State University page on CI issues: <https://www.ferris.edu/forms/mco-ciss/index.html>

Figure 24 shows a portion of these websites in the Web Crawl application.

Graphical user interface, text, application, email, Teams

Description automatically generated

Figure 24: Portion of websites added to the Web Crawl repository for Mr. Vision

Figure 25 shows the results after the configuration of the Web Crawl connector was completed.

Graphical user interface, application, Teams

Description automatically generated

Figure 25: Results of the configuration of the Web Crawl

The synchronization settings show 1029 documents were added. These documents were all forms of text. Three enrichments were added with entity extractions from Mayo clinic and concept tagging in healthcare as examples. The category classification was in health and fitness to disease which is a correct classification. Some documents failed (n = 60) when viewing these failures all were from the Mayo Clinic and included irrelevant search skill inputs (e.g., maps of Ohio) and no further action was taken on the document fails.

* *Integrate the search skill into the chatbot dialog implemented in part 2.*
* *Create a virtual assistant. Add the search skill and the dialog skill to the assistant.*
* *Use preview link integration to test the search skill configuration*

The public link generated for testing was:

<https://web-chat.global.assistant.watson.appdomain.cloud/preview.html?region=au-syd&integrationID=eeea12e5-f8bf-4bd7-9a97-748b1586b6eb&serviceInstanceID=5a8de2b0-318a-4570-a621-ac68c4804ef0>

* *The purpose of the search skill and how the user input is processed to retrieve the documents via Web Crawl connector*

The purpose of the search skill is to assist Watson when it does not have an explicit solution to a problem. The process includes routing the user’s question to a search to find an answer from the internet via a Web Crawl connector. Watson Assistant and Watson Discovery interact to extract relevant information for Mr. Vision. Figure 26 show the end-to-end integration flow of search skills between these two Watson services (Shrivastava, 2020).

Diagram

Description automatically generated

Figure 26: End-to-End workflow for Watson Assistant and Watson Discovery

The user input is processed to retrieve the documents via the Web Crawl connector from the Anything Else node. This node searches an external data source for an answer when none of the dialog nodes can answer the question. One example of this from Mr. Vision is when Bill responded “ok” to “How are you feeling today?” Figures 27 and 28 show the interaction between the message ok to the dialog, where no response which triggered the Anything else node which lead to the search skill via the Web Crawler and a response from the internet.

A picture containing graphical user interface

Description automatically generated

Figure 27: Dialog with Bill

Graphical user interface, text, application

Description automatically generated

Figure 28: Web Crawler response

* + *The URL(s) specify in the connector synchronization setting and why*

The URL’s specified above in the connector synchronization setting were chosen for several reasons. First, they are highly reputable places of clinical care (e.g., Mayo Clinic and Cedars Sinai Hospital systems). Second, because they address convergence insufficiency in some way. For instance, the American Optometry Association is the governing vision group in the United States and has both credibility in matters related specifically to vision and addresses CI on their website. Third, they are chosen because they would expand the connections made between CI issues and Mr. Visions ability to understand and respond effectively to input. For instance, Mayo Clinic has a section on children and CI. Symptoms of CI for children may be limited to reading and learning issues and less clarified in the child’s symptoms response. This helps Mr. Vision to draw from other concepts and “help” the child to understand and explain themselves. In turn, this assists Mr. Vision in providing accurate responses and provides value back to the patient.

* + *The chatbot response and the response relevancy*
  + *The approaches to customize the search skill to improve the relevancy*

Figures 27 and 28 show the correct response process for the search skill. To improve the relevancy of the response Mr. Vision needed training to know how to interpret the response “ok”. To train Mr. Vision in the Analytics tab in User Conversations the ‘Ok” response had no entities assigned and the intent was classified as “irrelevant”. “ok” was given a new entity @ciss\_symptom:ok (Figure 29). Then a new intent #Have\_problem (Figure 30). Finally, these changes can be seen in Figure 31.

Graphical user interface, application, Teams

Description automatically generated

Figure 29: Mr. Vision is trained to assign a response of “ok” to the new entity.

Graphical user interface, application, Teams

Description automatically generated

Figure 30: Mr. Vision is assigned the intent of #Have\_problem from a response of “ok”

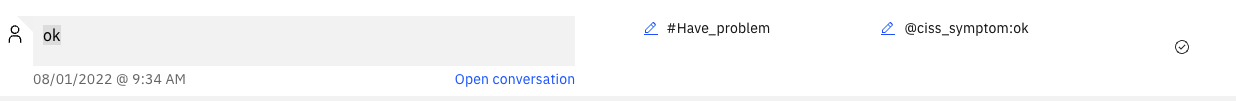


Figure 31: Changes to the response of “ok” are now reflected in the User Conversations

The prior example is a form of relevancy training for the purpose of improving Mr. Visions search results. The result moved from being classified as “irrelevant” to being assigned an “entity” (IBM, n.d).

Other forms of optimizing the search include Query expansion which expands the scope of a query beyond exact word matches. This may include synonyms, antonyms, and even typical misspellings of words. It may also include using stop words to be filtered out of queries as they add little value (Shrivastava, 2020).

A final way to optimize search is via Smart Document Understanding (SDU) which allows Mr. Vision to show a particularly relevant section of a document as a response to a search query (Shrivastava, 2020).

**Conversation data analysis experiment (15 pts)**

* *Share the public link for your assistant with 2-5 friends and/or family members and/or with the course TA. Include the link in your paper.*

The public link generated for testing was:

<https://web-chat.global.assistant.watson.appdomain.cloud/preview.html?region=au-syd&integrationID=eeea12e5-f8bf-4bd7-9a97-748b1586b6eb&serviceInstanceID=5a8de2b0-318a-4570-a621-ac68c4804ef0>

* *Day 1* 
  + *Ask each participant to run 3-4 conversations.*
  + *At the end of the day, take screenshots of the conversation analytics overview tab. Include the screenshots in your paper and interpret the metrics and plots.*

The first testing session was on August 1st, 2022, and included Bill, Terri, and Melissa for a total of 9 conversations (see Figure 32, Scorecard). The average message per conversation is the total messages received during the selected time divided by the total conversations during that same time. In the first day of testing this was (rounded up) five. The maximum number of conversations for a single data point within the selected time is nine. Thirteen messages had a weak understanding.

*Graphical user interface, application

Description automatically generated*

Figure 32: Scorecard of the first day’s Mr. Vision chatbot conversations.

Coverage includes the number of conversations in which Mr. Vision is confident that he can address the customer request. 82.9% of the conversations were confidently covered (n = 34; Figure 33) and 7 were considered not confidently covered. No containment data was recorded as no human agent was logged.

Chart, line chart

Description automatically generatedDiagram

Description automatically generated

Figure 33: Coverage of conversations with Mr. Vision after Day 1 of testing

The analytics also provide the most frequently used intents and entities (Figure 34). #vision\_problem is the most frequently used intent and @annual\_doc\_appt is the most frequently used entity.

Graphical user interface, application

Description automatically generated

Figure 34: Analytics of the most frequently used intents and entities for Mr. Vision.

* *Check the interactions between user and chatbot on the user conversation tab and identify the messages with unrecognized intents and with misclassified entities. Adjust the classification as needed and discuss the adjustments in your paper.*

A total of 44 interactions were displayed in the User Conversation tab after the first day of testing. Some interactions needed intents to be classified, others needed entities, some needed both (see Figure 35).

Graphical user interface, application

Description automatically generated

Figure 35: Sample of interactions from the User Conversation tab under Analytics

After training Mr. Vision these interactions were reduced (see Figure 36). Adjustments included:

1. Assigning entities for example assigning “can’t focus” to the @ciss\_symptom: can not focus (Figure 36)
2. Assigning intents for example assigning “I am blind as a bat” as a #vision\_problem (Figure 36)
3. Creating new entities for example creating the entity @ciss\_symptom: don’t feel well to interaction “I don’t feel joyful” (Figure 36)

Graphical user interface, application

Description automatically generated

Figure 36: Shows changes made to train Mr. Vision and assign appropriate entities and intents

* *Day 2* 
  + *Ask each participant to run the same 3-4 conversations again.*
  + *At the end of the day, take screenshots of the conversation analytics overview tab. Make sure that the time period includes day 1 and day 2.*
  + *Include and discuss the screenshots of the overview page in your paper. Discuss the changes in top entities and top intents during day 2.*

The second testing session was on August 3rd, 2022, and included Bill, Tp, Mary and Roger for a total of 17 conversations (see Figure 37, Scorecard). The average message per conversation is the total messages received during the selected time divided by the total conversations during that same time. For both days of testing this was 4.04. The maximum number of conversations for a single data point within the selected time is 17. Thirty-one messages had a weak understanding.

Graphical user interface, application

Description automatically generated

Figure 37: Scorecard of both day’s Mr. Vision chatbot conversations.

Figure 38 is the scoreboard for the second day of testing only.

Graphical user interface, application

Description automatically generated

Figure 38: Scorecard for the second day of Mr. Vision chatbot conversations.

Coverage includes the number of conversations in which Mr. Vision is confident that he can address the customer request. 82.9% of the conversations were confidently covered (n = 87; Figure 39) and 18 were considered not confidently covered. No containment data was recorded as no human agent was logged.

Chart, line chart

Description automatically generated

Figure 39: Coverage of conversations with Mr. Vision after Day 1 and 2 of testing

Figure 40 shows similar values for the second day of testing when compared to day 1 and to the overall cumulative results.

Chart

Description automatically generated

Figure 40: Coverage of conversations with Mr. Vision after Day 2 of testing

The analytics also provide the most frequently used intents and entities (Figure 41 & 42). #vision\_problem continued to be the most frequently used intent and @ciss\_symptom is the most frequently used entity.

Graphical user interface, application

Description automatically generated

Figure 41: Analytics of the most frequently used intents and entities for Mr. Vision cumulatively

Graphical user interface

Description automatically generated with medium confidence

Figure 42: Analytics of the most frequently used intents and entities for Mr. Vision on Day 2

* *Check the interactions between user and chatbot on the user conversation tab and identify the messages with unrecognized intents and with misclassified entities. If you were going to run this experiment for a third day, what additional entity and intent classification adjustments would you make?*

A cumulative total of 105 interactions were displayed in the User Conversation tab after testing was complete. If a third day of testing was undertaken, the following intent and entity classification adjustments would be made:

1. Assigning entities where they were unassigned
2. Assigning intents where they were irrelevant
3. Creating new entities if needed
4. Creating new intents if needed
5. Consider the autolearning product which observes options that customers click when a disambiguation list is shown
6. Consider adding new URL’s to the search skill
7. Alburger (2021) suggests creating very specific intents that serve on defined purpose. In healthcare, often symptoms overlap, therefore, as Mr. Vision learns I would expect more intents to be established with a narrower, more specific defined purpose. Specifically trying to differentiate between vision symptoms and other symptoms not related to vision but in which a person also feels unwell, for example a sore leg.

**Analysis and Enhancement (20 pts)**

* *What approaches did you use to personalize the interaction with the chatbot?*

Several approaches were used to personalize Mr. Vision. First, Mr. Vision is given a name that relates to his purpose. Second, Mr. Vision introduces himself (Figure 43).

Graphical user interface, application, Teams

Description automatically generated

Figure 43: shows Mr. Vision introducing himself

A third personalization technique is to vary the introduction and vary the responses throughout the conversation (dialog). Fourth, Mr. Vision interacts with the user by using their name via a context variable $name.

A fifth technique for personalization is to allow the user to correct errors in inputs without having to restart the entire conversation (Zheng, 2017). Mr. Vision does these using slots and the anything else node. When Mr. Vision encounters something that is not understood, perhaps it is a misspelling, it will tell the user it does not understand and give the user a chance to redo the input. In a similar vein should the user input inappropriate language Mr. Vision handles this by using a special node condition called “irrelevant” and a text response asking the user to redo their response. In a paid version of Mr. Vision, the chatbot could also transfer the user to a human.

A sixth technique Mr. Vision uses to personalize the conversation is the use of images. In the welcome message the image is carefully chosen to show what CI looks like. The image in the goodbye node waves goodbye to the user making Mr. Vision seem more “down to earth” and real.

Finally, a seventh way Mr. Vision is personalized is by giving him an overarching personality (Ursu, 2019). As humans have a strong tendency to anthropomorphize names have meaning and are powerful indictors of personality, character, and identity. Mr. Vision is a male; he is here to help the user with discovering a possible vision issue. Mr. Vision cares about the user’s health by reminding them of the importance of vision as a component of the users overall healthcare. In these ways Mr. Vision comes alive, has a personality and therefore is also more personalized to the user.

* *Choose 2 Watson services from the list below and discuss how would you integrate them to enhance the chatbot’s understanding of user’s input and to improve the relevancy of responses.*

Two of the services I would use to enhance Mr. Vision chatbot are: speech to text, and Text to Speech. These services either convert the human voice into the written word for voice control (speech to text) or vice versa. The reason these two services would be extremely useful for Mr. Vision is because the people who are most likely to benefit from Mr. Vision are people who have vision problems and may not be able to see the written word. These services remove this barrier so that those with vision problems do not need to rely on vision to interact with Mr. Vision.

Another Watson service that would be useful for Mr. Vision is language translator. One of the main reasons people do not access health care services, including vision care, is because most healthcare institutions have poor access to interpreter services or have no services at all (Shamsi, 2020). A potentially huge benefit of chatbots in general, and Mr. Vision specifically, is the ability to provide access to healthcare. However, this is only truly beneficial to all if the chatbot can bridge the gaps to use, including both language and vision blockers. Therefore, both translation of the written word via speech to text or text to speech and language translation would enhance the chatbot’s understanding of user’s input and in turn improve the relevancy of responses.

* *What messaging platform would you use to deploy the chatbot? Explain the rationale for your platform choice.*

As previously discussed, Mr. Vision should be available in two ways, on a doctor’s website and during check-in at the doctor’s office. After some research across healthcare chatbots and what messaging platforms the most popular healthcare chatbots use it was decided that Sensely would be a good choice (see https://sensely.com/). Sensely is used by the Mayo Clinic and enables patients track their health symptoms using both text and speech. The chatbot platform provides multilingual services while “enhancing access and building trust.” Furthermore, Sensely is an empathy-driven conversational platform with diverse content and allow for many different user input variations.

Given that Mr. Vision is also a healthcare chatbot, and that Mayo Clinic, a world leading healthcare provider uses Sensely, it seems a very promising choice as a messaging platform for Mr. Vision.

* *What approaches would you use to further train the chatbot after deployment*

Several approaches can be taken to train Mr. Vision after deployment. The first, and most important, is to further build out synonyms, intents, and entities to ensure a more diverse dialog can occur. Adding depth to the conversation will also allow the understanding of Mr. Vision to classify if there is a CI problem to be more accurate. Often CI problems are diverse in symptomology. Hence the need for more robust synonyms, intents, and entities. To accomplish this goal, there are two main options for training:

1. Add more websites to the Web Crawl
2. Consider careful manual review of results from the Conversation Data Analysis to manually add or classify language that is undefined.

A second way to train Mr. Vision, which is also an enhancement to the chatbot is to provide some different options for symptom input. For example, a checklist whereby the user simply clicks on what symptoms they are experiencing from a list. Another modification could be to provide dichotomous (yes/no) answering to symptoms. Both these techniques provide different ways for users to interact with Mr. Vision within confined parameters. These parameters can be used to help accurately classify the outcomes.

A third approach to training Mr. Vision would be to use the improve component of the conversation service (Azraq, Aziz, Nappe et al., 2017). This technique is also called “self-feeding” (Hancock, Bordes, Mazare, Weston, 2019). The improve component provides a history of conversations with users which can be used to improve Mr. Visions understanding of user inputs. This can allow for Mr. Vision to find an unrecognized entity and train the workspace to recognize it.

A fourth way to train Mr. Vision after deployment is after learning and self-feeding to review the intents and make sure they are distinct. Alburger (2021) suggests creating very specific intents that serve on defined purpose. In healthcare, often symptoms overlap, therefore, as Mr. Vision learns I would expect more intents to be established with a narrower, more specific defined purpose. One advantage of Mr. Vision is that the intents can be unlimited, and this can help provide success in identifying potential vision issues and next steps for healthcare.

A fifth way to improve and train Mr. Vision once deployed is to create a diverse team to handle the bot. Diversity in a team allows for variation in dialog and therefore allows Mr. Vision to successfully generalize the conversation.

A sixth way to improve and train Mr. Vision once deployed is to not just rely on text (or even images). Text to speech and speech to text, translations to different languages, and never stopping training will all help Mr. Vision improve after deployment to reach and help more people.

**Conclusion (10 pts)**

* *Discuss 3 chatbot implementation challenges and at least 3 recommendations to build the “smart” chatbots.*

Several challenges were experienced in the process of implementing Mr. Vision. These can be explained in various categories, first, challenges with the CISS survey, second, challenges with IBM Watson, third challenges with the process of testing in production.

The CISS survey, although scientifically validated, is a healthcare symptom survey. Symptoms surveys are challenging because they rely on user “opinion” as to whether a symptom is present and occurring at a level in which it could be considered “clinically relevant.” Given the time to complete the project, Mr. Vision simply asks what symptoms are occurring? A smart chatbot would enhance this user input by providing a grade from 1-10, whereby 1 is very low occurrence of symptom and 10 is constantly or very prominent occurrence of the symptom. Furthermore, a smart chatbot could try to determine if the symptom occurred in response to an action. For example, a headache in response to a short period of reading that occurs every time the user reads are a strong indicator of a CI issue. However, a headache that occurs because the user hit their head while on a on-time work site accident is not an indicator for a CI issue. A smart bot could learn to connect and differentiate symptoms with actions and with severity. This would allow Mr. Vision to become more accurate and more intuitive which is seen as one of the most common barriers to adoption of chatbots today according to Michiels (n.d.).

A second implementation challenge was with IBM Watson. When using the Try it out Panel there was a significant amount of initial training needed to connect intents and entities, so they were correctly recognized as being “relevant” to a specific category rather than being irrelevant. This was a manual process. It would be a significant enhancement if IBM Watson could have a much larger cloud-based repository of words with synonyms already attached to intents and entities. In this way it would help with the implementation processes and significantly help with the training process both before and after the bot went into production, helping Mr. Vision to generalize his understanding and become a smart bot sooner. This would also help with another barrier for adoption of chatbots stated by Michiels (n.d.) which is a lack of a simple way to find and begin developing smart chatbots.

A third implementation challenge is the nature of the process for experimentation. Several limitations existed here. The first challenge was time, which related to time to develop and train as well as time to understand the concepts. It is clear there is a lot to understand in Natural Language Processing (NLP) and this requires significant time and effort to understand anything beyond surface level understanding. A second implementation challenge was in only having one person (myself) add to the intents, entities, and synonyms. NLP is not my core set of competencies and although the Web Crawler helps a smart bot could be used to seek out specific sites rather than having the developer (myself) try and figure out the best options. More websites are clearly needed for Mr. Vision to generalize his conversation. Which leads to a fundamental concern about Mr. Vision: could Mr. Vision cause damaging blunders? Such a concerns is a barrier for using chatbots according to Michiels (n.d.). This is especially significant with healthcare challenges. If Mr. Vision was to tell a person who had a significant CI issue that they simply needed to rest instead of going to see a doctor this would be a catastrophic blunder and is an example of a false, negative response. To overcome this response a smart bot may know when to “refer” to a human for greater clarification. Perhaps the smart bot would learn to recognize confusion or conflicting responses. For instance, a persistent headache while reading is highly correlated to a persistent lack of focus also while reading. These are highly predictive indictors of a CI issue. Therefore, if a person entered that they had a headache on a scale of 10 yet had no issues focusing this would be conflicting and may require a referral to a human to sort out. Such a referral would be a smart implementation of a chatbot as the most important goal of Mr. Vision is to provide accurate information and next steps.

* *Would the chatbot that you implemented replace the interaction with the live person? Why or why not?*

No, Mr. Vision would certainly not replace a health care professional. Chatbots in general do not replace live persons however, this is especially important when dealing with health-related information. One reason for this is that the CISS survey is only one component of the entire health picture of a patient. When the patient is with an optometrist, the doctor has the patients’ entire medical records. Furthermore, the doctor can do a medical examination of the vision in real time while observing clinical indications. S/he can examine the records in context with the CISS outcomes to determine if other health components should be considered.

Mr. Vision certainly has a place in healthcare as a prescreening tool that can be used to heighten awareness of CI and to efficiency progress patients through a practice flow. However, Mr. Vision does not, and should not replace the healthcare provider.

The biggest experimental take-away I got from this assignment was the complexity of language. I knew there were many synonyms for words, but I learned that within a category of words there is also a lot of overlap. For example, “feeling unwell” could have nothing to do with vision. Therefore, directed questions may be more appropriate for Mr. Vision. I expected the conversation to have to be highly structured to start with, however, I was also pleasantly surprised by the search skill in helping the conversation move beyond what I could provide.

I appreciate how chatbots can be used to improve efficiencies but, on the flip side, I could also see people getting frustrated with them quickly if they do not work the way they expect. I would anticipate that over time, as NLP gets better and better that there will be less cases of frustration.

In conclusion, the purpose of Mr. Vision is noble. With much more dedicated time and effort I think such a concept would be beneficial in allowing access to a validated questionnaire that could encourage more people to obtain appropriate healthcare for CI issues.

**References**

Alburger, J. (2021). How to Train a Chatbot: 8 Effective Tips for Training AI. Retrieved in July 31st, 2022 from: <https://www.hubtype.com/blog/how-to-train-a-chatbot-2>

Azraq, A., Aziz, H., Nappe, N., Bravo, C.R. & Sri, L. (2017). *Building Cognitive Applications with IBM Watson Services: Volume 2 Conversation.* Redbooks, IBM.

Borsting, E.J., Rouse, M.W., Mitchell, G.L., et all and the CITT group. Validity and reliability of the revised convergence insufficiency symptom survey in children. *Optometry and Visual Science,* 80(12): 832-838.

Fox Vision Development Center (2021). *Vision Therapy for Convergence Insufficiency.* Retrieved from: <https://www.foxvisiondevelopment.com/vision-therapy-optometrist/vision-therapy-for-convergence-insufficiency/>

Hancock, B., Bordes, A., Mazare, P-E., Weston, J. (2019). *Learning from Dialogue after Deployment: Feed Yourself, Chatbot!* Retrieved on July 31st, 2022, from: <https://arxiv.org/abs/1901.05415>

IBM (n.d.). Improving result relevance with the tooling. Retrieved on August 3rd, 2022, from: <https://cloud.ibm.com/docs/discovery?topic=discovery-improving-result-relevance-with-the-tooling>

Jovic, D. (2022). The Future is Now – 37 Fascinating Chatbot Statistics. Retrieved on July 29th, 2022, from: <https://www.smallbizgenius.net/by-the-numbers/chatbot-statistics/#gref>

Michiels, E. (n.d.). *Modelling Chatbots with a Cognitive System Allows for a Differentiating user Experience.* Retrieved on July 31st, 2022, from: <http://merode.econ.kuleuven.be/events/2017/poem/PoEM_2017_IndustryTrack4_Chatbots.pdf>

Oracle (n.d.). How do chatbots work? Retrieved on July 29th, 2022 from: <https://www.oracle.com/chatbots/what-is-a-chatbot/#link2>

Shamsi, H.A., Almutairi, A.G., Mashrafi, S.A. & Kalbani, T.A. (2020). Implications of language barriers for healthcare: A systematic review. *Oman Medical Journal, 35*(2): e122.

Shrivastava, R. (2020). *Watson Assistant Search Skill: Taking Virtual Assistant to the Next Level.* Retrieved on August 3rd, 2022, from: <https://medium.com/ibm-watson/watson-assistant-search-skill-taking-virtual-assistant-to-the-next-level-3c122aa8a110>

Stanford Children’s Health. Lucile Packard Children’s Hospital, Stanford University Medical (2021). *Crossed-Eyes (Strabismus).* Retrieved from: <https://www.stanfordchildrens.org/en/topic/default?id=crossed-eyes-strabismus-90-P02109>

The Dizziness and Headache Optometry Center (2021). *What causes binocular vision dysfunction?* Retrieved from: <https://www.dizzinessandheadache.com/blog/what-causes-binocular-vision-dysfunction.html>

Ursu, S. (2019). *How to Design a Chatbot Personality.* Retrieved on July 31st, 2022 from: https://medium.com/the-chatbot-guru/how-to-design-a-chatbot-personality-8f89b5a6b549

Zheng, J. (2017). *6 Tips to Improve the User Experience of Your Customer Support Chatbot.* Retrieved on July 31st, 2022 from: <https://medium.com/ibm-watson/6-tips-to-improve-the-user-experience-of-your-customer-support-chatbot-50baf67b02b7>

**Appendix**

**Appendix A**

**Table

Description automatically generated**