Week 4: Chatbot Assignment – Cade Menezes

Step 1: Test the Chatbot

Does it show a good grasp of syntax, semantics, and pragmatics? Can it remember information it was previously told? What is it good at doing, in a conversation, and what are some of its weaknesses? Copy and paste your conversation into the assignment document. The conversation should be about 300 Words.

```
> Hello, I am Eliza. I'll be your therapist today.
* Hello, I'm Cade
> How long have you been Cade?
* For 19 years.
> Tell me more...
* I was born in Toronto, Ontario. Then moved to the middle-east
for a decade. Then came back.
> Would you say that you have psychological problems?
* Yes.
> I see. Let's explore that a bit more.
* I have anxiety and a lack of motivation to do things.
> I'm not sure I understand you fully.
* My psychological problem is that I have anxiety, as a result
of that, I prevent myself from doing new things and finishing my
work.
> Would you say that you have psychological problems?
* Yes.
> Are you sure about that?
* Yes I am sure about that
> Did you come to me because you are sure about that?
* Yes I did.
> I see. Let's explore that a bit more.
* Sure, what would you like to explore
> We were discussing you, not me.
* Okay
> OK... "Sure, what would me like to explore". Tell me more.
* What is my name
> Have you asked such questions before?
* You asked me what my name was
> You're not really talking about me, are you?
* No I am not
> How long have you been not?
* I am Cade
```

> Do you believe it is normal to be Cade?
* I do believe it is normal to be Cade
> I'm not sure I understand you fully.
* What do you do as a therapist
> Oh... Let's move on to something else for a bit. do as a therapist?
* I don't understand.
> Why don't you understand?
* Because your question makes no sense.
> Why not?
* "do as a therapist?" is not a proper question.
> I see. Let's try another topic and we will come back to that issue later.
* What do you know about dogs?
> We were discussing you, not me.
* I like dogs
> Tell me more...
* Dogs are cute and cuddly, and are always your best friend
> What about your own best friend?
* My best friend is really nice too. He's a human though
> Why do you bring up the topic of friends?
* I love my friends. They help a lot with my mental health
> What does that suggest to you?
* That life is better with friends.
> That is an interesting response. Let me make anote of that for next time.
* Okay sounds good.
> Would you say that you have psychological problems?
* Yes, I would.
> I understand.
* Okay.
> OK... "My best friend is really nice too. He's a human though". Tell me more.
* His name is Cam, and we have been through a lot together. He has helped me in my darkest times, and vice versa. He is one of my closest friends.
> No need for names in our session. Continue
* Anyways, I love him so much. He is still in my life, and we are playing games later tonight.
> I think you are blocking what you really want to say.
* There's nothing I really want to say about Cam.

```
> Tell me more...
* goodbye i gotta go
> That is an interesting response. Let me make anote of that for
next time.
```

Step 2: Analyze the Chatbot's Abilities

After chatting with the chatbot, describe its ability with human language. Does it show a good grasp of syntax, semantics, and pragmatics? Can it remember information it was previously told? What is it good at doing, in a conversation, and what are some of its weaknesses? Demonstrate your understanding of the linguistics concepts by answering these questions with explicit reference to the linguistics concepts and providing justification for your answers in 350-500 words.

According to the website, Eliza, built in the 1960s, uses a pattern matching and substitution system rather than understanding actual language. This was how language processing worked in the previous decades. This can be seen within Eliza's use of **syntax**. She is able to examine and mirror the user's grammatical structures, but is not able to understand it. As an example, I said "His name is Cam", and the chatbot replied with "No need for names in our session." This shows that Eliza recognizes name mentions as a cue, but doesn't actually process the conversation to ask meaningful follow-up questions about Cam. When it comes to semantics understanding, Eliza's ability to understand beyond keywords is nonexistent. When I said "I like dogs," she responded with "Tell me more…" and later, "What about your own best friend?" She shifted topics abruptly, mistaking my mentioning of dogs as a cue to discuss personal relationships. She shows a lack of semantic processing, she doesn't understand what a dog is, nor does she understand that my best friend is referred to as a human. In a **pragmatic** sense, Eliza doesn't follow conventional conversational norms. She consistently repeats the same question, "Would you say that you have psychological problems?" which ignore memory and relevance. She doesn't adjust responses based upon the logical or emotional flow of the conversation. When I correct her incomplete question "do as therapist?" she doesn't acknowledge the correction, but pivots ambiguously "Let's try another topic." She demonstrates an inability to repair conversation breakdowns. Eliza also lacks memory. She cannot retain past information in a meaningful way. When I stated my name, she didn't remember when I asked for it later. When I mentioned and gave details about my best friend, she didn't follow up or build upon what I had said. This example shows that Eliza has limited discourse coherence, and her inability to give informed future responses that build upon previous conversation information. Eliza is strong in giving the illusion that she understands what the user is saying. She produces therapeutic responses like "What does that suggest to you?" or " tell me more…" Despite the lack of real understanding and comprehension, she gives the illusion of reflection and intrigue. Eliza's

weakness is clear within her inability to manage contextual replies, truly understand user inputs, and adapt within the conversation. Her responses are often generic, especially within logical and emotionally complex conversations.

Step 3: Predict if your Chatbot Would Pass the Turing Test
How do you think that your chatbot would do at a Turing test? What skills has it demonstrated that would allow it to pass, and what weaknesses or inhuman traits would give it away? Overall, after weighing these traits, do you think that the chatbot would pass or fail? Demonstrate your understanding of the linguistics concepts by answering these questions with explicit reference to the linguistics concepts and providing justification for your answers. Your full answer for step 3 should be 350-500 words long.

The Turing test figures out whether a machine can mimic a human within conversation convincingly. If a judge cannot reliably distinguish a machine from a human, then the machine passes the Turing test. Based around my interaction with Eliza, I believe that she would have failed the Turing test. Based on Eliza's pragmatic competence, Eliza's inability to rely on coherence, contextual replies, and the ability to respond in a meaningful way, she would not be able to produce a human conversation. When I said "My psychological problem is that I have anxiety" Eliza repeats the previous questions at a later time "Would you say that you have psychological problems?" She cannot track the flow of the conversation and shows her inability to respond to new information, which a judge would quickly be able to determine as non-human. Eliza also has limited ability with semantics. She cannot accurately understand the meaning of the user's words and sentences. I asked her "What is my name?" which is a basic test of memory and semantic understanding, she deflects with an unrelated question. A human would either answer with my name or acknowledge the topic. This semantic evasion also occurs when I say "I like dogs" and she immediately shifts the topic to "What about your best friend?" without understanding the difference between pets and humans. Her responses are based on detecting keywords rather than true understanding. Eliza's syntax is her strong area, she can often produce grammatically correct sentences. Though there are some errors that occur during conversation that shows the limits to this chatbots' syntax. When I ask "What do you do as a therapist?" She responds with "do as a therapist?" a sentence that she's ungrammatically replied with. This shows that she doesn't actually understand sentence structure like a human and her syntax is just surface-level pattern recognition. One last major giveaway is her lack of memory. I state my name and personal information and Eliza never references them again. A human would be able to recall and act like they listen to the information. Eliza treats each input as isolated, and breaks that illusion of human-like conversations. Overall, Eliza's reliance on repetition, lack of memory, generic replies, and inability to contextual replies will certainly lead a human judge to the conclusion that they are speaking to a machine rather than a human. Her therapeutic tone is good

enough to appear human briefly, but she cannot sustain the illusion for long. Therefore, Eliza would fail the Turing Test.