



Dialogue Systems and Chatbots

Natalie Parde, Ph.D.

Department of Computer
Science

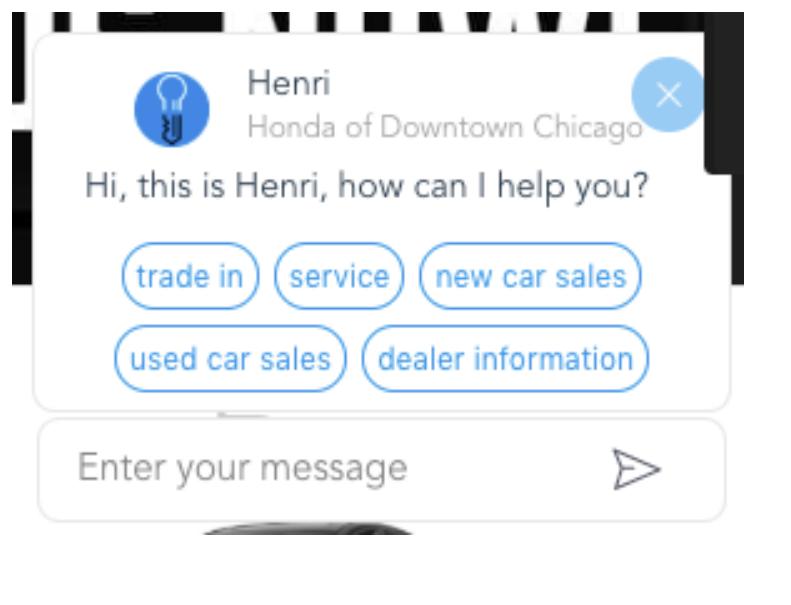
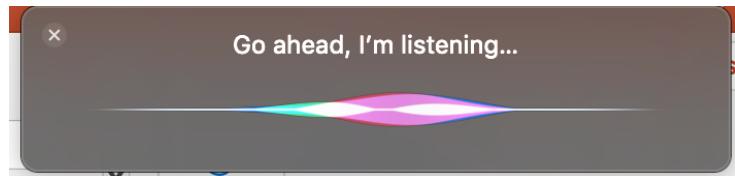
University of Illinois at
Chicago

CS 421: Natural Language
Processing
Fall 2019

Many slides adapted from Jurafsky and Martin
(<https://web.stanford.edu/~jurafsky/slp3/>).

What are dialogue systems?

- Programs capable of communicating with users in natural language via text, speech, or both
- Often referred to as conversational agents



Dialogue systems are everywhere!

Types of Dialogue Systems

Task-Oriented

- Designed to leverage conversational interactions to help users complete tasks

Chatbots

- Designed to carry out extended, unstructured conversations (similar to human chats)

Many dialogue systems contain elements of both of these types!

- Task-oriented agents may seem more natural if they also contain a chatbot component

U: Hey

A: Hi, **how are you?**

U: I'm doing good, how are you?

A: **I'm doing good as well. Would you like me to help you reserve a room for your meeting?**

Designing high-quality conversational agents requires an understanding of how human conversation works.

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn---

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Turns:** Individual contributions to the dialogue
 - Typically a sentence, but may be shorter (e.g., a single word) or longer (e.g., multiple sentences)

Turn

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- Understanding turn structure is very important for spoken dialogue systems!
- Systems must know:
 - When to stop talking
 - Dealing with interruptions
 - When to start talking
 - Detecting when the human user has finished speaking
- Detecting when a user has finished speaking is called **endpoint detection**
 - Challenging due to noise and speech pauses

Turn

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Speech Acts:** Types of actions performed by the speaker
 - Also referred to as **dialogue acts**
- Major dialogue act groups:
 - **Constatives**
 - **Directives**
 - **Commissives**
 - **Acknowledgments**

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Constatives:** Making a statement
 - Answering
 - Claiming
 - Confirming
 - Denying
 - Disagreeing
 - Stating

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Directives:** Attempting to get the addressee to do something

- Advising
- Asking
- Forbidding
- Inviting
- Ordering
- Requesting

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Commissives:** Committing the speaker to a future action
 - Promising
 - Planning
 - Vowing
 - Betting
 - Opposing

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Acknowledgements:**
Expressing the speaker's attitude regarding some social action
 - Apologizing
 - Greeting
 - Thanking
 - Accepting an acknowledgement

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Grounding:** Establishing common ground by acknowledging that the speaker has been heard and/or understood

- Saying “okay”
- Repeating what the other speaker said
- Using implicit signals of understanding like “and” at the beginning of an utterance

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Grounding:** Establishing common ground by acknowledging that the speaker has been heard and/or understood
 - Saying “okay”
 - Repeating what the other speaker said
 - Using implicit signals of understanding like “and” at the beginning of an utterance

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Grounding:** Establishing common ground by acknowledging that the speaker has been heard and/or understood
 - Saying “okay”
 - Repeating what the other speaker said
 - Using implicit signals of understanding like “and” at the beginning of an utterance

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- Conversations have structure
 - Questions set up an expectation for an answer
 - Proposals set up an expectation for an acceptance or rejection
- These dialogue act pairs are called **adjacency pairs**
 - First pair part: Question
 - Second pair part: Answer

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- However, two dialogue acts in an adjacency pair don't always immediately follow one another!
- Adjacency pairs can be separated by **side sequences** or **subdialogues**
 - Interruptions
 - Clarifying questions
 - Corrections
- Some adjacency pairs also have **presequences**
 - Requests may be preceded by questions about a system's capabilities

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Initiative:** Conversational control
- Generally, the speaker asking questions has the **conversational initiative**
- In everyday dialogue, most interactions are **mixed-initiative**
 - Participants sometimes ask questions, and sometimes answer them

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- Although normal in human-human conversations, mixed-initiative dialogue is very difficult for dialogue systems to achieve!
- In question-answering systems (e.g., “Alexa, what’s the weather like right now?”) the initiative lies entirely with the user
 - Systems such as these are called **user-initiative** systems
- Opposite of user-initiative system: **system-initiative** system
 - Can be very frustrating!
Salesperson: Which variety of caramel popcorn would you like?
Natalie: I don’t want caramel popcorn.
Salesperson: Which variety of caramel popcorn would you like?
Natalie: Can I quit?
Salesperson: Which variety of caramel popcorn would you like?
Natalie: Um, regular.

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I’ll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.

Properties of Human Conversation

- **Inference:** Drawing conclusions based on more information than is present in the uttered words
- **Implicature:** The act of implying meaning beyond what is directly communicated

Mentioning the party on Saturday (especially in response to the salesperson's question!) implies that the popcorn will be needed by that time

The salesperson infers that the popcorn should be ready for pickup by Friday

Natalie: Hi, I would like to order thirteen buckets of cheesy popcorn.

Salesperson: Um okay when do you need those?

Natalie: I want to bring them to a party on Saturday.

Salesperson: And what size buckets would you like?

Natalie: Extra large.

Salesperson: Okay, our cheesy popcorn is really popular. Would you be okay with six buckets of cheesy popcorn and seven buckets of caramel popcorn?

Natalie: No.

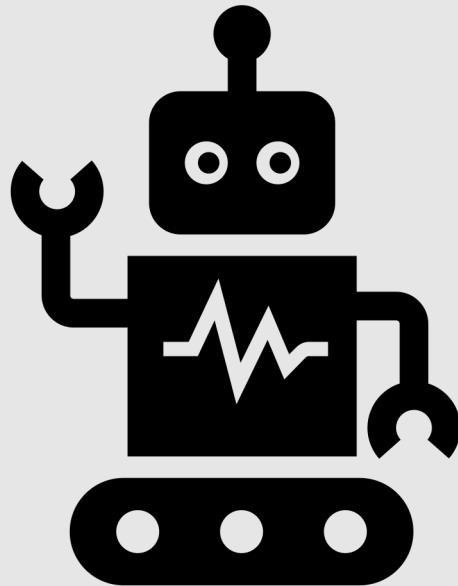
Salesperson: Okay, what about some of our other flavors? We have ranch-flavored popcorn--

Natalie: I'll take that. Eight buckets of ranch-flavored popcorn and five buckets of cheesy popcorn.

Salesperson: Okay.

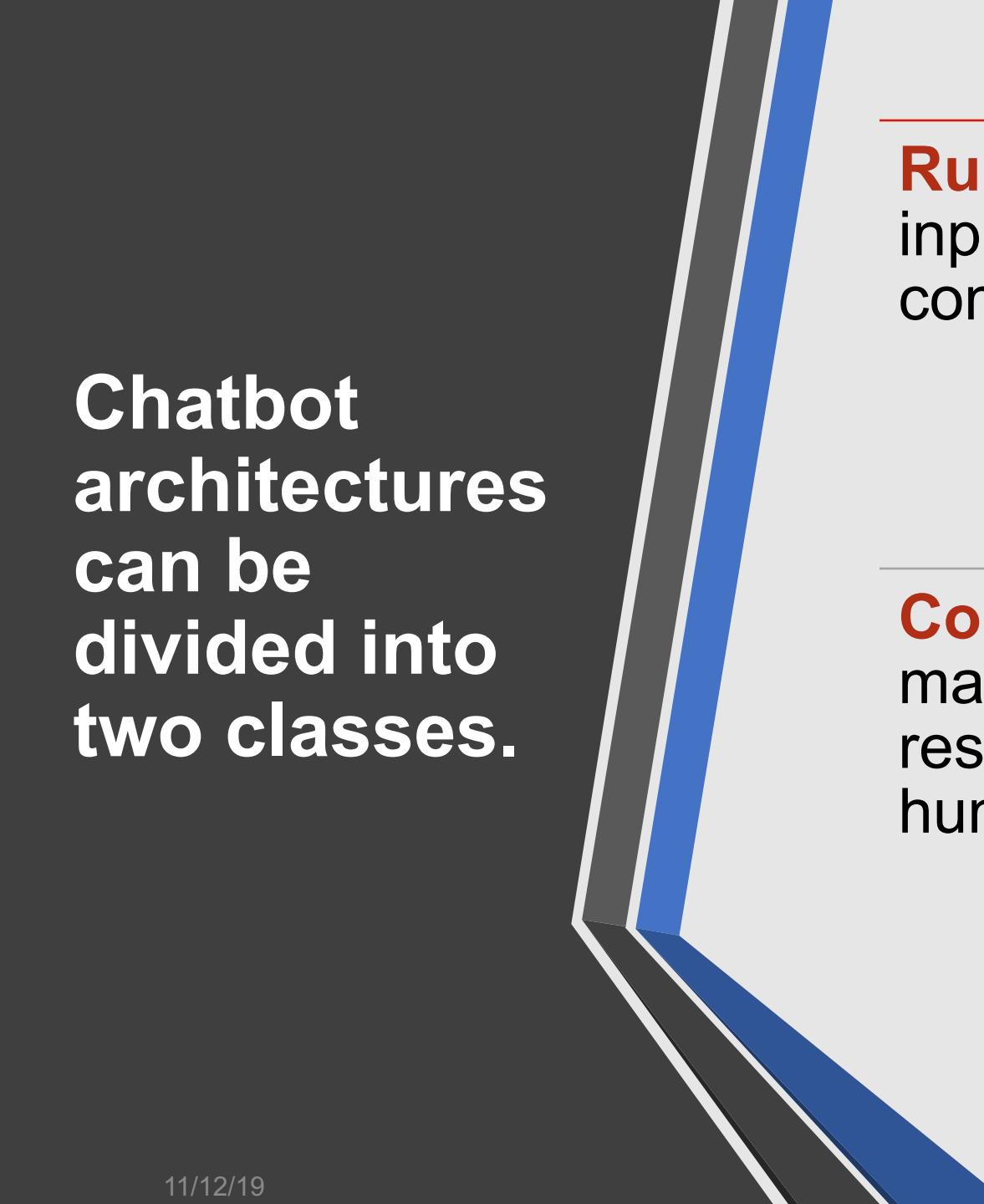
Natalie: Actually, wait. Seven buckets of ranch and six buckets of cheesy popcorn, still all in extra large.

Salesperson: Okay, we will have seven extra-large buckets of ranch-flavored popcorn and six extra-large buckets of cheesy popcorn ready for you to pick up on Friday.



Chatbots

- Dialogue systems designed to carry out extended conversations with the **goal of mimicking unstructured, informal, human-human chats**
- Can be used for **entertainment**
 - Cleverbot:
<https://www.cleverbot.com/>
- Or more **practical** purposes
 - ELIZA



Chatbot
architectures
can be
divided into
two classes.

Rule-based: Interpret and respond to input using a (sometimes extremely complex) set of rules

Corpus-based: Automatically learn to map from a user utterance to a system response based on large datasets of human-human conversations

Rule-based Chatbots

- Revisiting **ELIZA**:
 - Rogerian psychologist (reflect patients' statements right back at them)
 - Takes the stance of knowing almost nothing about the world
 - Common position for chatbots trying to pass the Turing Test ...users may assume that the chatbot is pretending not to know anything to fulfill some type of conversational goal

U: Men are all alike.
E: IN WHAT WAY
U: They're always bugging us about something or other.
E: CAN YOU THINK OF A SPECIFIC EXAMPLE
U: Well my boyfriend made me come here.
E: YOUR BOYFRIEND MADE YOU COME HERE
U: He says I'm depressed much of the time.
E: I AM SORRY TO HEAR YOU ARE DEPRESSED

Rule-based Chatbots

- ELIZA works by **matching patterns** and **transforming slots** from those patterns into responses
 - $(\cdot^* [yY][oO][uU] \cdot^* [mM][eE]) \rightarrow (\text{WHAT MAKES YOU THINK I } 3 \text{ YOU})$
 - You hate me → WHAT MAKES YOU THINK I HATE YOU
- Each pattern used by ELIZA is based on **keyword(s)** that might occur in users' sentences

Rule-based Chatbots

- In ELIZA, keywords are associated with a rank
 - More specific keywords → higher rank
 - More general keywords → lower rank
- Consider two rules:
 - (I .*) → (YOU SAY YOU 2)
 - (EVERYBODY .*) → (WHO IN PARTICULAR ARE YOU THINKING OF)
- Consider an input sentence:
 - “I know everybody made a cooler chatbot.”
 - Since “I” is more general than “everybody,” the system will return **WHO IN PARTICULAR ARE YOU THINKING OF** rather than **YOU SAY YOU KNOW EVERYBODY MADE A COOLER CHATBOT**

Rule-based Chatbots

- What if no keywords are matched in an input?
 - ELIZA defaults to a non-committal response
 - PLEASE GO ON
 - THAT'S VERY INTERESTING
 - I SEE



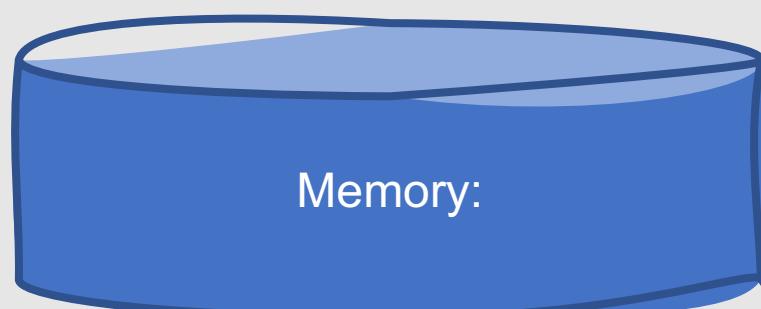
Rule-based Chatbots

- Alternately, if no keywords are matched in an input, ELIZA can access facts from its **memory**
 - Whenever “my” is the highest-ranked keyword, select a transformation rule, apply it to the input, and store the transformed input on a stack
 - My boyfriend made me come here. → DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE
 - Pop the most recent transformed input off the stack if no keywords are matched in an input

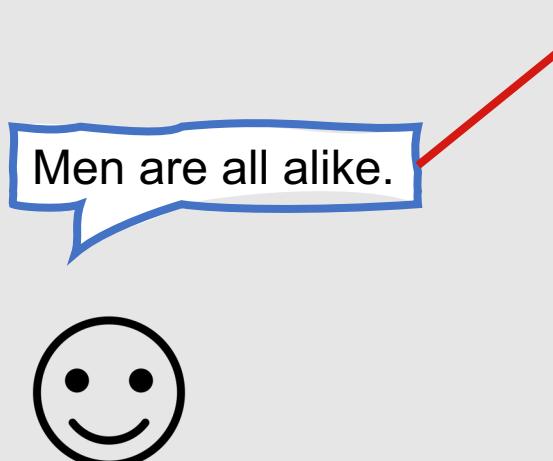
Example: ELIZA

Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(.* always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(.* my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)

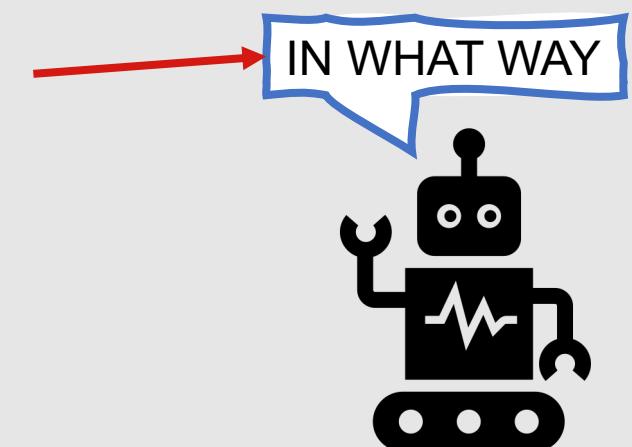
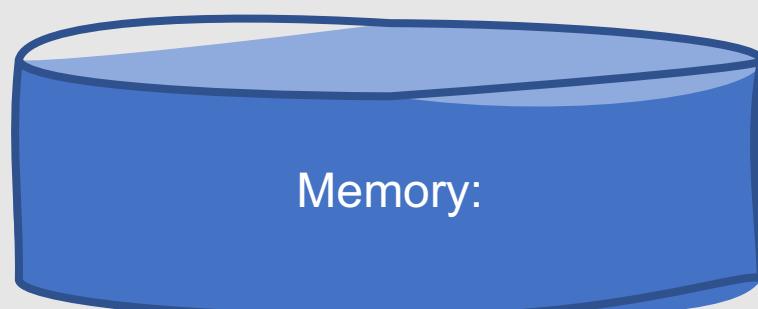
Men are all alike.



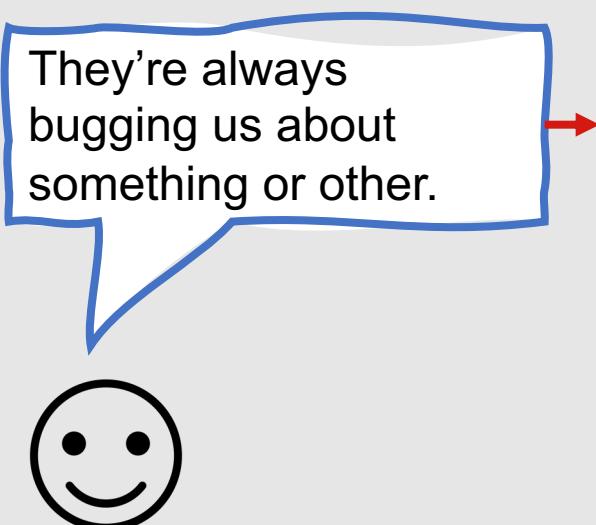
Example: ELIZA



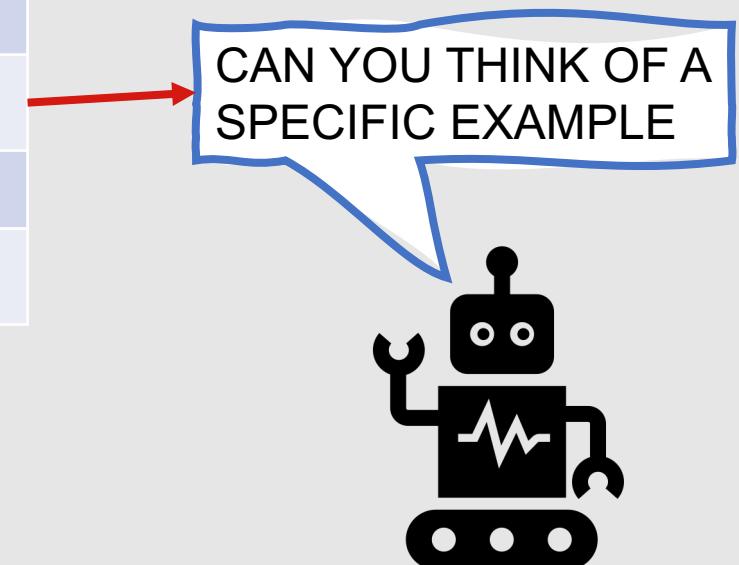
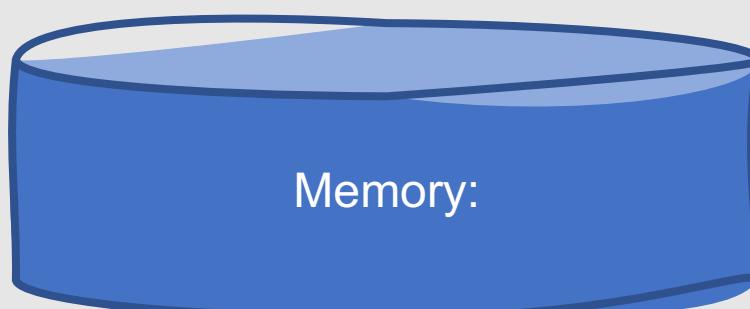
Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(.* always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(.* my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)



Example: ELIZA



Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(.* always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(.* my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)

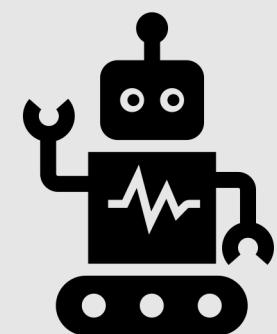
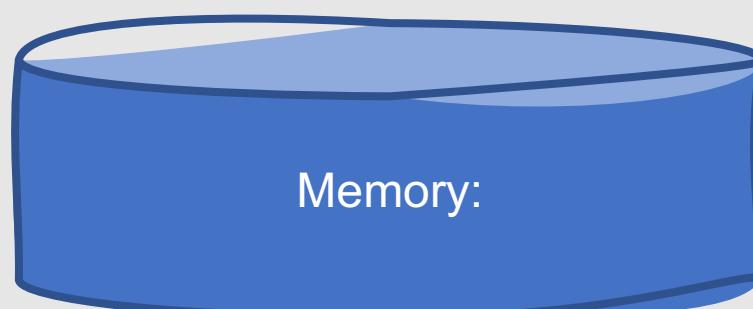


Example: ELIZA

Well my boyfriend
made me come here.



Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(.* always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(.* my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)



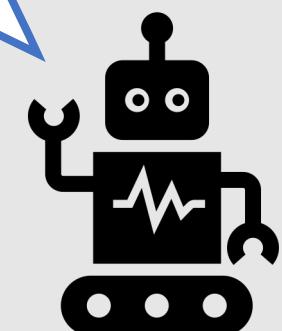
Example: ELIZA

Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(.* always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(.* my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)

Well my boyfriend made me come here.



YOUR BOYFRIEND MADE YOU COME HERE



Memory:
1) DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE

Example: ELIZA

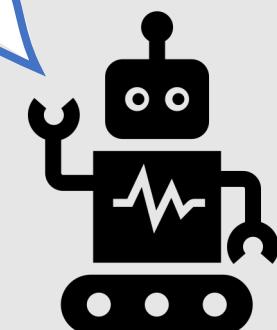
He says I'm depressed
much of the time.



Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(.* always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(.* my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)

Memory:
1) DOES THAT HAVE ANYTHING TO DO
WITH THE FACT THAT YOUR
BOYFRIEND MADE YOU COME HERE

I AM SORRY TO
HEAR THAT YOU ARE
DEPRESSED MUCH
OF THE TIME



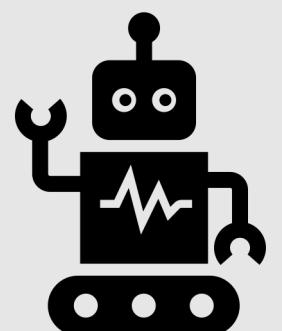
Example: ELIZA

Maybe he is the problem.



Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(.* always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(.* my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)

Memory:
1) DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE

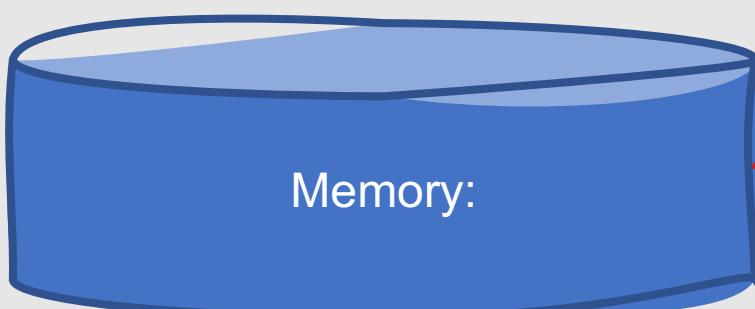


Example: ELIZA

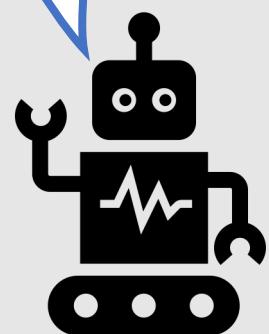
Maybe he is the problem.



Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(.* always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(.* my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)



DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE



Formal Algorithm: ELIZA

```
find the word  $w$  in the sentence  $s$  that has the highest keyword rank
if  $w$  exists:
    choose the highest ranked rule  $r$  for  $w$  that matches  $s$ 
    response  $\leftarrow$  apply the transformation in  $r$  to  $s$ 
    if  $w == "my"$ :
        future  $\leftarrow$  apply a transformation from the memory rule
        list to  $s$ 
        push future onto the memory stack
    else: # No keyword applies!
        either:
            response  $\leftarrow$  apply the transformation for NONE to  $s$ 
        or:
            response  $\leftarrow$  pop the top response from the memory stack
```

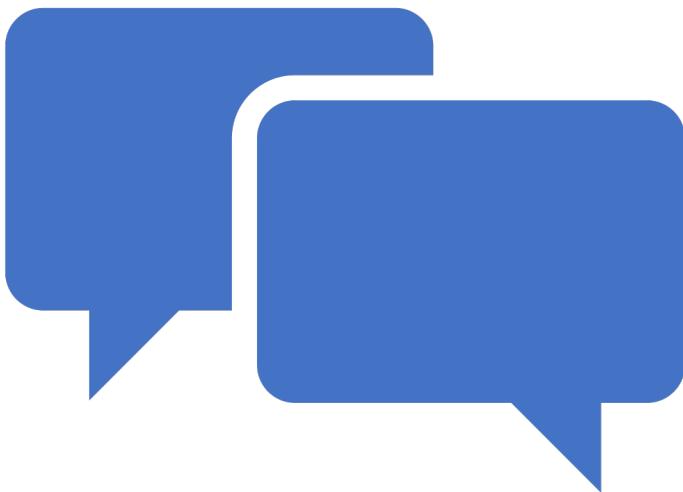
Continued popularity of ELIZA's framework....

- Still used in many applications today!
 - Modern chatbot systems like ALICE are based on updated versions of ELIZA's pattern/action architecture
- PARRY
 - Another clinical psychology chatbot developed several years after ELIZA
 - Included a model of its own mental state, with affect variables for its levels of fear and anger
 - High anger → hostile output
 - High fear → sequence of deluded statements
 - First known system to pass the Turing Test (psychiatrists were unable to distinguish transcripts from PARRY with transcripts from people with real paranoia)

Corpus-based Chatbots

- Do not use hand-built rules
- Instead, learn to map inputs to outputs based on large human-human conversation corpora
- Very data-intensive!
 - Training modern corpus-based chatbots requires hundreds of millions, or even billions, of words





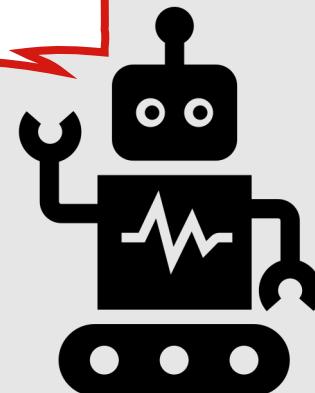
What kind of corpora are used to train corpus-based chatbots?

- Large spoken conversational corpora
 - Switchboard corpus of American English telephone conversations:
<https://catalog.ldc.upenn.edu/LDC97S62>
 - Movie dialogue
 - Text from microblogging sites (e.g., Twitter)
 - Collections of crowdsourced conversations
 - Topical-Chat:
<https://github.com/alexa/alexa-prize-topical-chat-dataset>

Possible responses can also be extracted from non-dialogue corpora.

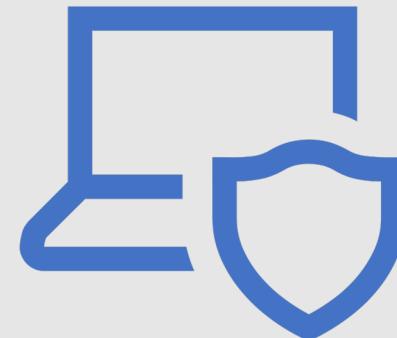
- Possible sources:
 - News
 - Online knowledge repositories (e.g., Wikipedia)
- This allows the chatbot to tell stories or mention facts acquired from non-conversational sources

Did you know that Illinois has the 25th largest land area of all U.S. states?



As humans interact with a chatbot, their own utterances can be used as additional training data as well.

- This allows a chatbot's quality to gradually improve over time
- Some privacy concerns can emerge when using this strategy (it's crucial to remove **personally identifiable information!**)



Corpus-based Chatbots

- Two main architectures:
 - **Information retrieval**
 - **Machine learned sequence transduction**
- Most corpus-based chatbots do (surprisingly!) very little modeling of conversational context
- The focus?
 - Generate a single response turn that is appropriate given the user's immediately previous utterance(s)

Corpus-based Chatbots

- Since they tend to rely only on very recent utterances, many corpus-based chatbots are viewed more as **response generation** systems
- This makes them similar to **question answering systems**:
 - Focus on single responses
 - Ignore context or larger conversational goals

Information Retrieval-based Chatbots

- Respond to a user's turn by **repeating some appropriate turn from a corpus** of natural human conversational text
- Any information retrieval algorithm can be used to **choose the appropriate response**
- Two simple methods:
 - Return the response to the most similar turn
 - Return the most similar turn

How can we return the response to the most similar turn?

- Look for a turn that resembles the user's turn, and return the human response to that turn
- More formally, given:
 - A user query, q
 - A conversational corpus, C
 - Find the turn t in C that is most similar to q (e.g., using cosine similarity) and return the human response to t
 - $r = \text{response}(\underset{t \in C}{\operatorname{argmax}} \frac{q^T t}{\|q\| \|t\|})$

How can we return the most similar turn?

- Directly match the user's query, q , with turns from C , since a good response will often share words or semantic patterns with the prior turn
- More formally:

$$\bullet \quad r = \operatorname{argmax}_{t \in C} \frac{q^T t}{\|q\| \|t\|}$$

Which of these methods works better?



Depends on the application



More often, returning the most similar turn seems to work better

Slightly less noise from the alternate method of selecting the response to the most similar turn

Various techniques can be used to improve performance with IR-based chatbots.

- Possible additional features:
 - **Entire conversation** with the user so far
 - Particularly useful when dealing with short user queries, e.g., “yes”
 - **User-specific** information
 - **Sentiment**
 - Information from **external knowledge sources**

Encoder-Decoder Chatbots

- **Machine learned sequence transduction:** System learns from a corpus to **transduce a question to an answer**
 - Machine learning version of ELIZA
- Intuition borrowed from **phrase-based machine translation**
 - Learn to convert one phrase of text into another
- Key difference?
 - In phrase-based machine translation, words or phrases in the source and target sentences tend to align well with one another
 - In response generation, a **user's input might share no words or phrases with a coherent, relevant response**

How does a chatbot learn to perform this transduction?

Encoder-decoder models

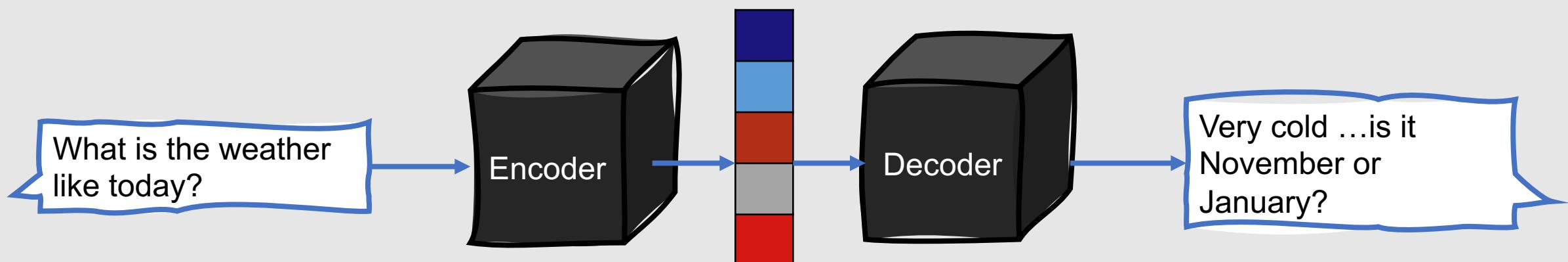
- Accept sequential information as input, and return different sequential information as output

Also recently used in:

- Machine translation
- Question answering
- Summarization

How do encoder-decoder models work?

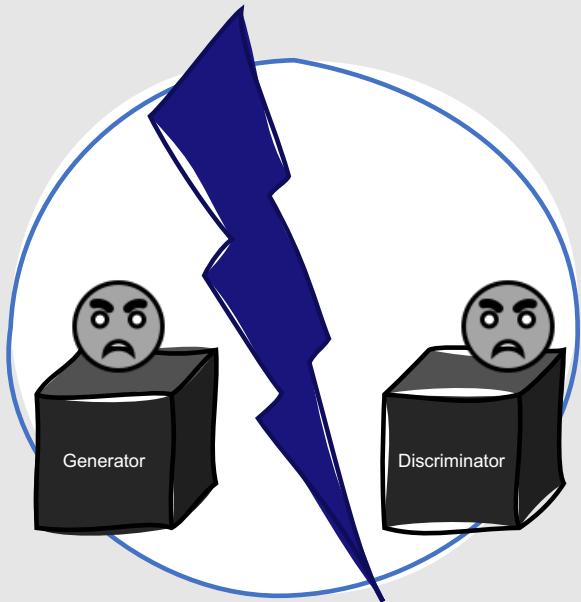
- In NLP applications, encoders and decoders are often some type of **RNN**
- Encoders take sequential input and generate an **encoded representation** of it
 - This representation is comprised of the outputs from the last hidden state of the encoder network ...it is undecipherable to casual observers!
- Decoders take this representation as input and generate a sequential (interpretable) output



Encoder- Decoder Chatbots

- Basic encoder-decoder models tend to produce repetitive (and therefore boring) responses that don't encourage further conversation
 - "I'm okay"
 - "I don't know"
- To avoid this, it is important to incentivize response diversity
 - Mutual information objective function
 - Beam search

Encoder-Decoder Chatbots



- Other challenges?
 - Inability to model prior context
 - Can be solved by using a **hierarchical model** that summarizes information over multiple turns
 - Often poor multi-turn coherence
 - Can be addressed to some extent using **reinforcement learning** or **adversarial networks** to learn to choose responses that make the overall conversation more natural

How are chatbots evaluated?

Currently best evaluation technique?

- Collect human ratings!

Automated metrics correlate poorly with human judgements

- Typically measure slot-filling accuracy or word overlap with a gold standard
- These metrics work best when there are few possible responses and the responses lexically overlap with one another

Task-based Dialogue Systems

- Dialogue system has a specific **goal** (generally, helping a user solve some task)
 - Making a reservation
 - Purchasing a product
- Task-based dialogue systems are **frame-based**
 - Assume a set of user **intentions**, each of which consist of a collection of **slots** that can be filled by possible **values**
 - Set of **frames** is sometimes called a **domain ontology**

Frame-based Dialogue Systems

- Simple (but very long-lived!) frame-based dialogue system: GUS
 - Genial Understanter System
 - Originally used for travel planning (Bobrow et al., 1977)

GUS: Hello. My name is GUS. I can help you plan a trip by air. Where do you want to go?

Client: I want to go to San Diego on May 28

GUS: What time do you want to leave?

Client: I must be in San Diego before 10 am

GUS: Would you like P.S.A. flight 102 that arrives at 9:15 am?

Client: When is the next flight?

GUS: Air California flight 310 that leaves at 8:30am and arrives at 10:00am

Client: I'll take the first one.

GUS: In what name should I make a reservation, please?

Client: Dan Bobrow

GUS: I have confirmed the following flight: P.S.A. flight 102 on Wednesday May 28 from San Jose to San Diego that leaves at 7:30 am and arrives at 9:15 am.

Components of Frame- based Dialogue Systems

Slots: Specify what the system needs to know

- **Slot fillers** are constrained to values of specific semantic types
 - In the travel domain, these could be *cities*, *dates*, *airlines*, or *times*

These types may have hierarchical structures

- In GUS, *date* is a frame itself, with slots for *month*, *year*, *day*, and *weekday*

Components of Frame-based Dialogue Systems

Slot	Type	Question Template
ORIGIN CITY	city	“From what city are you leaving?”
DESTINATION CITY	city	“Where are you going?”
DEPARTURE TIME	time	“When would you like to leave?”
DEPARTURE DATE	date	“What date would you like to leave?”
ARRIVAL TIME	time	“When do you want to arrive?”
ARRIVAL DATE	date	“What day would you like to arrive?”

Control Structure for Frame-based Dialogue

The control architecture for frame-based dialogue systems is also designed around the frame

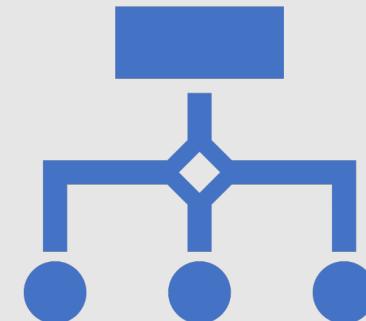
- Goal:
 1. Fill the slots in the frame with the fillers the user intends
 2. Perform the relevant action for the user

The system achieves its goal by asking questions of a user

- Typically these questions are constructed using pre-specified question templates associated with each slot of each frame

Control Structure for Frame-based Dialogue

- The system continues questioning the user until it is able to fill all slots needed to perform the desired task
- GUS attaches **condition-action rules** to slots to reduce monotony
 - If a user has specified a flight destination city, it may automatically fill the hotel destination slot with that value as well



Control Structure for Frame-based Dialogue

- Many domains require multiple frames!
- Dialogue systems must be able to **disambiguate** which slot of which frame a given input is supposed to fill, and then switch dialogue control to that frame
- This can be done using **production rules**
 - Different types of inputs and recent dialogue history match different frames
 - Control is switched to the matched frame
- Once the system has enough information, it performs the desired task (e.g., querying a database of flights) and returns the result to the user



Natural Language Understanding in Frame- based Dialogue Systems

- In a frame-based dialogue system, natural language understanding is necessary for performing three tasks:
 - **Domain classification**
 - **Intent determination**
 - **Slot filling**

Natural Language Understanding in Frame-based Dialogue Systems

Domain Classification: What is the user talking about?

Booking a flight

Setting an alarm

Managing a calendar

Intent Determination: What task is the user trying to accomplish?

Retrieve all flights in a given time window

Delete a calendar appointment

Slot Filling: What particular slots and fillers does the user intend the system to understand from their utterance, with respect to their intent?

Natural Language Understanding for Slot Filling in Frame-based Dialogue Systems

Show me the morning flights from Chicago to Dallas on Thursday.

Domain: AIR-TRAVEL
Intent: SHOW FLIGHTS
Origin-City: Chicago
Origin-Date: Thursday
Origin-Time: morning
Destination-City: Dallas

Wake me tomorrow at 6

Domain: ALARM-CLOCK
Intent: SET-ALARM
Time: 2019-11-13 0600

Natural Language Understanding for Slot Filling in Frame-based Dialogue Systems



In GUS, and in many commercial applications, slots are filled using handwritten rules

wake me (up)? | set (the|an) alarm
| get me up → Intent: SET-ALARM



Rule-based systems often include large quantities (thousands!) of rules structured as semantic grammars

Semantic Grammar: A context-free grammar in which the left-hand side of each rule corresponds to the semantic entities (slot names) being expressed

Semantic grammars can be parsed using any CFG parsing algorithm



Other systems use supervised learning for slot filling

Semantic Grammar

SHOW → show me | i want | can i see

DEPART_TIME_RANGE → (after | around | before) HOUR |
morning | afternoon | evening

HOUR → one | two | three | four | ... | twelve (AM|PM)

FLIGHTS → (a) flight | flights

AMPM → am | pm

ORIGIN → from CITY

DESTINATION → to CITY

CITY → Chicago | Dallas | Denver | Phoenix

Other Components of Frame-based Dialogue Systems

- **Automated Speech Recognition:** Converts audio input to a string of text
 - May or may not be constrained based on the current dialogue domain and/or intent
- **Natural Language Generation:** Produces the utterances that the system outputs to the user
 - Frame-based systems typically use **template-based generation**
 - What time do you want to leave ORIGIN-CITY?
 - Will you return to ORIGIN-CITY from DESTINATION-CITY?
- **Text to Speech Synthesis:** Converts a string of text to an audio output
 - May be done at runtime or using prerecorded statements or phrases

Advantages and Disadvantages of GUS

Why use a simple, rule-based architecture for task-oriented dialogue systems?

- High precision
- Given a narrow domain and sufficient expertise, can provide sufficient coverage

Why explore other options?

- Handwritten rules or grammars can be expensive and slow to create
- Low recall

Summary: Dialogue Systems and Chatbots (Part 1)

- **Dialogue systems** are programs capable of communicating with users in natural language
- Two types of dialogue systems:
 - **Conversational (chatbot)**
 - **Task-oriented**
- Dialogue systems must understand a variety of communicative functions, including **turn-taking**, **dialogue acts**, **grounding**, **conversational structure**, **initiative**, and **implicature**
- Chatbots can be **rule-based** or **corpus-based**
- Corpus-based chatbots can use **information retrieval** or **sequence transduction** methods
- Task-based dialogue systems are **frame-based**
- Frame-based dialogue systems assume a set of **intentions**, each of which consist of a set of **slots** that can be filled by possible **values**
- Slots can be filled using **rule-based** or **machine learning** approaches