

# Rule-Based Chatbots

Natalie Parde

UIC CS 421

# 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
  - (`. * [yY][oO][uU] . * [mM][eE]`) → (WHAT MAKES YOU THINK I 3 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

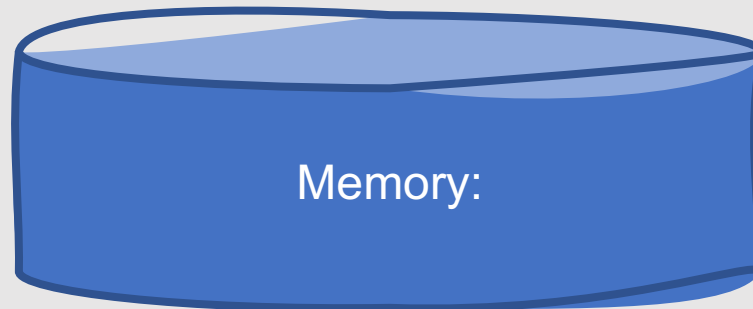
- Alternately, if no keywords are matched in an input, ELIZA can access facts from its **memory**
  - Whenever “my” is the highest-ranked keyword, select a transformation rule, apply it to the input, and store the transformed input on a stack
    - My boyfriend made me come here. → DOES THAT HAVE ANYTHING 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

Men are all alike.



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)

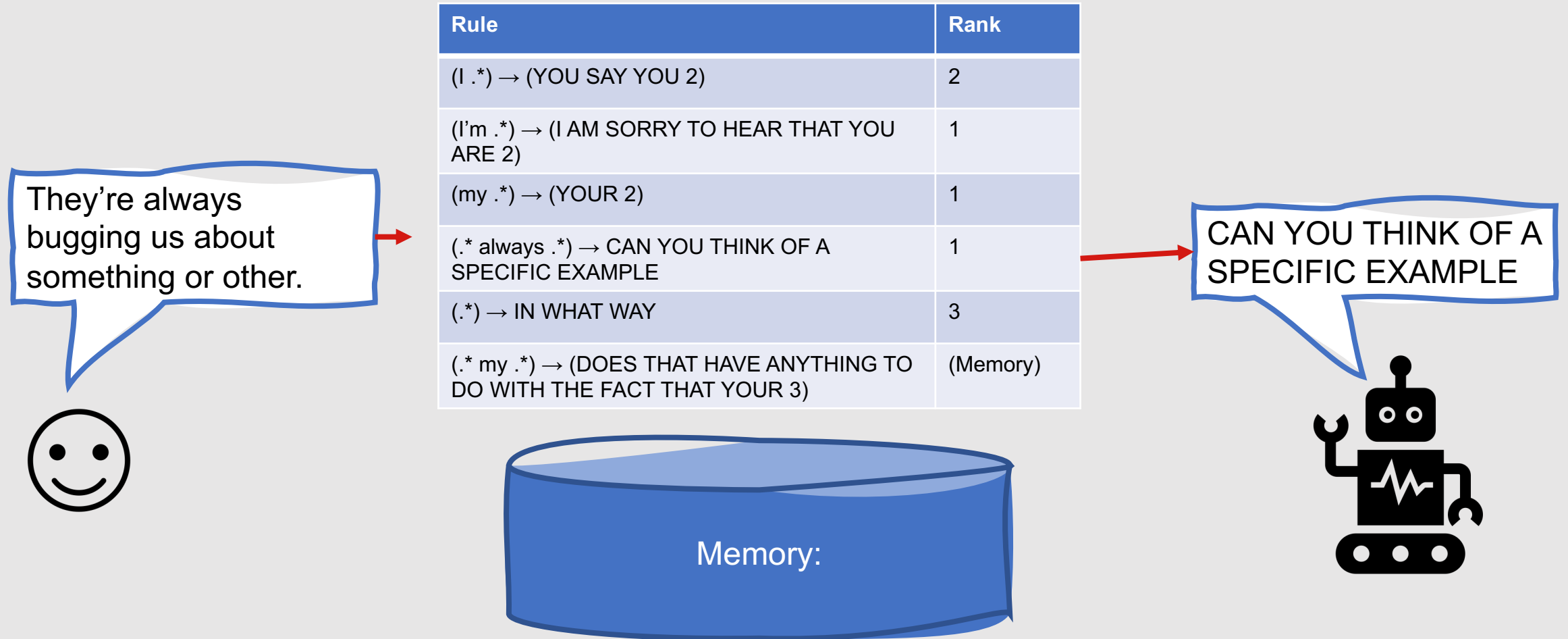


# Example: ELIZA





# Example: ELIZA



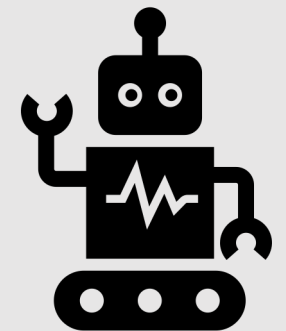
# Example: ELIZA

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)

Memory:



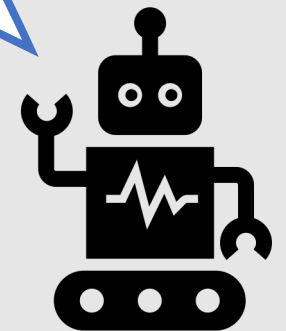
# Example: ELIZA

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

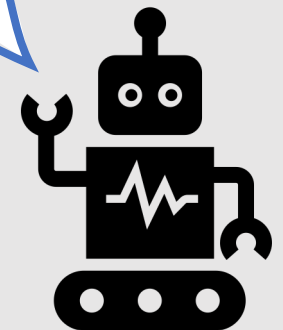
Rule	Rank
(I .*) → (YOU SAY YOU 2)	2
(I'm .*) → (I AM SORRY TO HEAR THAT YOU ARE 2)	1
(my .*) → (YOUR 2)	1
(. * always .*) → CAN YOU THINK OF A SPECIFIC EXAMPLE	1
(.*) → IN WHAT WAY	3
(. * my .*) → (DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR 3)	(Memory)

He says I'm depressed much of the time.



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

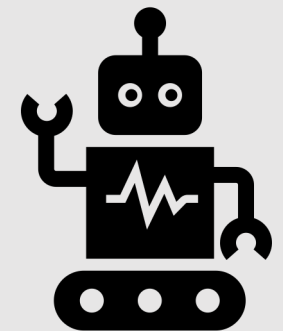
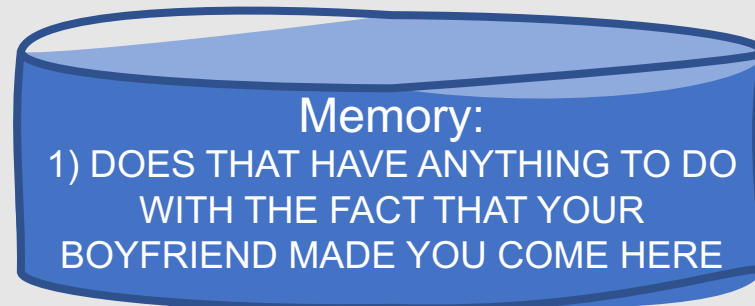


# Example: ELIZA

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)

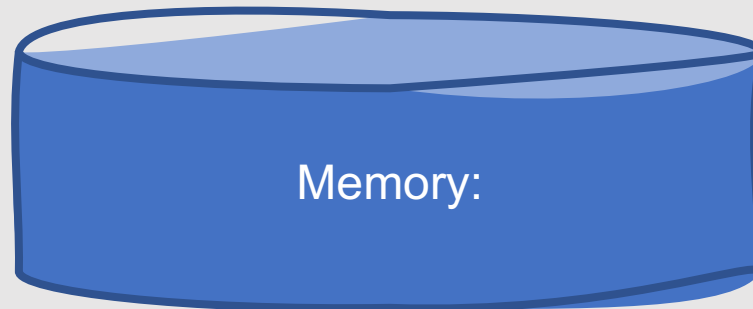


# Example: ELIZA

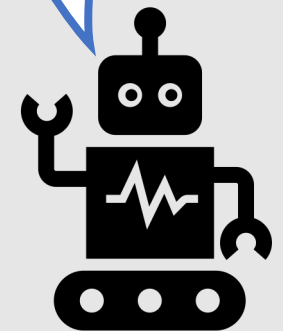
Maybe he is the problem.



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  $w == \text{"my"}:$ 
        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)