

RULE BASED CHATBOT

Chaitanya Agarwal, 2018114003

Akshat Chhajer, 2018114014

Computational Linguistics 2

Fall 2019

Abstract

The simplest kind of dialogue systems are chatbots, systems that can carry on extended conversations with the goal of mimicking the unstructured conversations or ‘chats’ characteristic of informal human-human interaction. This report talks you through developing a chatbot both in Hindi and English which relies entirely on rules. Inspired from the very first conversational AI, ELIZA, we try to replicate the script of a Rogerian Psychologist and discuss the results in the later sections of the report.

ELIZA

ELIZA (Weizenbaum, 1966) is the most important chatbot dialogue system in the history of the field of **Converational AIs**. ELIZA was designed to simulate a **Rogerian psychologist**, based on a branch of clinical psychology whose methods involve drawing the patient out by reflecting patient's statements back at them. Rogerian psychology is the rare type of conversation in which, as Weizenbaum points out, one can "assume the pose of knowing almost nothing of the real world". If a patient says "I went for a long boat ride" and the psychiatrist says "Tell me about boats", you don't assume she didn't know what a boat is, but rather assume she had some conversational goal. Most chatbots trying to pass the Turing test choose a domain with similar properties. Consider the following conversation where ELIZA tries to give very therapeutic advice, exercepted from a full interaction given in the original paper.

Person: Men are all alike.

ELIZA: IN WHAT WAY

Person: They're always bugging us about something or other.

ELIZA: CAN YOU THINK OF A SPECIFIC EXAMPLE

Person: Well my boyfriend made me come here.

ELIZA: YOUR BOYFRIEND MADE YOU COME HERE

Person: He says I'm depressed much of the time.

ELIZA: I AM SORRY TO HEAR THAT YOU ARE DEPRESSED

..

..

ELIZA: WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU

Person: My father

ELIZA: YOUR FATHER

Person: You are like my father in some ways.

ELIZA: WHAT RESEMBLANCE DO YOU SEE

Person: You are not very aggressive but I think you don't want me to notice that.

ELIZA: WHAT MAKES YOU THINK I AM NOT AGGRESSIVE

Person: You are afraid of me

ELIZA: DOES IT PLEASE YOU TO BELIEVE I'M AFRAID OF YOU

Person: My father is afraid of everybody

ELIZA: WHAT ELSE COMES TO YOUR MIND WHEN YOU THINK OF YOUR FATHER

Person: Bullies.

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

ELIZA works by **pattern/transform rule matching**. Consider the following:

(\$ YOU \$ ME) [pattern] -> (WHAT MAKES YOU THINK I \$ YOU) [transform]

This rule would transfer *You hate me* into *WHAT MAKES YOU THINK I HATE YOU*.

Each ELIZA pattern/rule is linked to a keyword that might occur in a user input. Keywords are associated with a rank, with specific words being more highly ranked, and more general words ranking lower. Consider the following user sentence:

I know everybody laughed at me

Because it has the word "I", this sentence could match the following rule whose keyword is *I*:

(I *) -> (YOU SAY YOU 2)

Producing:

YOU SAY YOU KNOW EVERYBODY LAUGHED AT YOU

If no keyword matches, ELIZA chooses a non-committal response like *PLEASE GO ON*, *THAT'S VERY INTERESTING* or *I SEE*.

The **algorithm** that drives it:

function ELIZA GENERATOR(*user sentence*) **returns** *response*

Find the word *w* in *sentence* that has the highest keyword rank

if *w* exists

 Choose the highest ranked rule *r* for *w* that matches *sentence*

response ← Apply the transform in *r* to *sentence*

if *w* = 'my'

future ← Apply a transformation from the 'memory' rule list to *sentence*

 Push *future* onto memory stack

else (no keyword applies)

 either *response* ← Apply the transform for the NONE keyword to *sentence*

or

response ← Pop the top response from the memory stack

return(*response*)

Abstraction to Hindi

We introduce the following inflections to accommodate gender pronouns, and articles in Hindi.

For the rest of it we follow the same approach as in the original ELIZA script in English.

"हूँ" : "हो",

"था" : "थे",

"मैं" : "आप",

"मेरे पास" : "आपके पास",

"मेरा" : "आपका",

"हो" : "हूँ",

"आपके पास": "मेरे पास",

"आपका" : "मेरा",

"आप" : "मैं",

"मुझे" : "आपको"

Results

Outcome for English

We achieve a conversational AI which totally captures the essence of the language.

Outcome for Hindi

We achieve a conversational AI which partially captures the essence of Hindi. It fails to capture the semantics of the language as it is a free word order language. We also get semantically ambiguous sentences as responses often.

Discussion

If extended to a probabilistic CFG for Hindi, we would get better results as it would encode the syntax and the semantics better.

References

1. ELIZA - A Computer Program For The Study of Natural Language Communication Between Man And Machine, Joseph Weizenbaum.
<https://web.stanford.edu/class/linguist238/p36-weizenbaum.pdf>
2. Natural Language Processing, A Panian Perspective
<http://ltrc.iiit.ac.in/downloads/nlpbook/nlp-panini.pdf>