

## ELIZA/DOCTOR PROGRAM

### Introduction

ELIZA was a computer program written by Joseph Weizenbaum [2] of MIT University in the late 60s which is considered to be the first chatterbot, i.e. a program that can partially mimic a human in a conversation with a human. In many ways ELIZA has provided insights not just into what a serious NLP system should achieve but also has provided a lot of insight into human reactions to computer systems which look like “intelligent” systems but are not so. ELIZA was not meant to be an AI system, it was meant to be a toy or a parody system. As Weizenbaum mentions ELIZA was supposed "to play (or rather, parody) the role of a Rogerian psychotherapist engaged in an initial interview with a patient" [1]. However, the reaction to the ELIZA program by the test subjects was completely unexpected and many formed an emotional attachment to the “therapist”. Weizenbaum was very disturbed by the unexpected reaction to ELIZA by lay and that was one of the reasons why he wrote the book “Computer Power and Human Reason” an attempt to clarify his position on computer science, AI, and its relations to human society.

### Working

ELIZA ran many scripts, each of which processed the responses of the users by using simple parsing and by substitution of them into pre-made templates and using hardcoded phrases, seemingly continued the conversation. The most popular script used in the ELIZA program was DOCTOR which was supposed to be a parody of a Rogerian Psychotherapist, i.e. it would ask the questions which are like those favoured by the 60s therapists who indulged in what is known as “Person Centered Therapy” in which the therapist would often return a question asked by the patient to themselves to increase empathy. This choice of a therapy setting was deliberate as questions like “what is your favourite colour” can be returned to the human conversant with replies like “why do you ask?” in such a setting. This kind of setting doesn’t need the system to actually understand the questions and the answer templates are pretty much finite.

Thus, Weizenbaum wanted to create something simple enough which will not have requirement to know real world data and thus sidestep issues of natural language processing or any kind of semantics. Nor did Weizenbaum have any intention to create an NLP system, the ELIZA project was supposed to be something very simple, almost a toy program, or in Weizenbaum’s own words, a parody. What DOCTOR does is use a lot of data from the statements the human provides in order to compose responses by using simple templates, these templates having the flavour of stereotypical therapy statements. Thus, DOCTOR does not even attempt to answer the questions in the traditional sense but follows an if-else pattern of using a certain template for a certain trigger. There is no attempt to comprehend the semantics of the statements of the human.

A sample transcript of the ELIZA program when the user keeps the conversation as it was originally intended would look like this:

I am the psychotherapist. Please, describe your problems. Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

Can you elaborate on that?

Too much work.

Why do you say too much work?

I am working a lot.

Earlier you said you felt unhappy?

Yes.

I would appreciate it if you would continue.

Well, I don't know what to do.

Maybe your plans have something to do with this.

No they don't.

Go on, don't be afraid.

I'm depressed.

What causes you to be depressed?

A sample transcript of the ELIZA program when the conversation is not in the “therapist” setting would look like this:

I am the psychotherapist. Please, describe your problems. Each time you are finished talking, type RET twice.

I wish to know about you.

Why do you say i might want to know me?

To write my report

What makes you believe to write your report?

I do not understand your question, please clarify

Earlier you said i would like to know me?

yes

Can you elaborate on that?

I wish to know ELIZA

How does it feel to want?

Want what?

Maybe your plans have something to do with this.

Er, what plans?

Is it because of your life that you say er what plans?

It is because of you I say that.

I would appreciate it if you would continue.

I will.

As can be seen in this ‘conversation’ with ELIZA it has a habit of parsing the responses of the user, i.e. me, to frame its next question. I continued the conversation and after a while the templates it uses to set these responses become predictable. It also has a few set phrases for conversation bits when it can’t find a response in its script, like “Maybe your plans have something to do with this.” While ELIZA (and its modern, more advanced brethren) may seem convincing for a while to someone who does not know that they are conversing with a chatterbot, at least for a while, these chatterbots will not stand sustained inquiry from a sceptical subject who is actively trying to do a limited version of the Turing test.

ELIZA was first implemented in the SLIP language (Symmetric List Processor), a language incensed by Weizenbaum himself as an extension to FORTRAN but with better functionality to process doubly linked lists. For its time, ELIZA was revolutionary in many aspects, as interfaces were not really common in the computers of the late 60s due to the absence of personal computing and thus the idea of interactive computation had not arisen yet or entered into popular fancy. Even though the perceived intelligence of it was an illusion (and a very bad illusion) the fact remains that it was the first genuine human machine interface (pretending to be a human – intelligent machine interface) attempting to use natural language. However, what was more interesting about ELIZA was not so much its own implementation or characteristics which were simple enough but that of the reaction of the lay public to it which first illustrated the fascination with “intelligent” systems that became a theme of how people react to computers later.

## Reactions to ELIZA and the “ELIZA Effect”

As Weizenbaum discovered, many subjects who experimented with ELIZA got emotionally attached to it. Many did so despite Weizenbaum’s informing them that there is no intelligence involved and that in fact ELIZA is not ‘answering’ them but only regurgitating a hardcoded script. In the article I have referenced below [1] Weizenbaum goes into some lengths that his subjects at times will not believe him and would for example; ask him to leave the room when they were talking to the “therapist” to maintain “privacy”. Weizenbaum realized what has been happening, that the subjects want ELIZA to be answering them.

This phenomenon is now known as the “ELIZA Effect” which is the tendency of humans to mistakenly confuse computer behaviours having intent behind them and draw analogies to human behaviours subconsciously. Like ELIZA, humans start attributing emotion and understanding to a programs output where none exist [3] or is supposed to exist. It is also important to note that ELIZA’s code was not supposed to elicit such a reaction so this was not a pre-calculated reaction. As mentioned in the introduction this disturbed Weizenbaum to a large extent and he spends a considerable time explaining why in his article referenced. Weizenbaum explains the dangers of humans leaving decisions to computer program or AI systems, even genuine AI systems, even assuming such systems were possible in the future, as he maintains that there is a difference between a computational “decision” and a “choice” and the non-equivalency of the former to the latter. While the human can choose to decide against something which is computationally the right decision by using non-quantifiable factors like judgement and emotions, and alter these decisions as they see fit, a computational system cannot do so. While a human can choose a different decision for the same input parameters multiple times, a computational system will not be able to do so. This article of Weizenbaum is significant as it touches upon issues which were seriously considered much later.

There were other important reactions to ELIZA as well. It influenced a lot of things from interface design to game design with many later developers creating similar programs in a variety of settings. ELIZA like programs found their places in some of the first computer games. Chatterbots are components of several different types of systems nowadays, like company websites and online help desks. Programs like these have implications for things like art and artificial creativity as well, starting discussions challenging programmers to define what creativity is.

## Conclusion

ELIZA, while itself a relatively simple system, is important from the point of view of understanding human intelligence and hypothetical machine intelligence as it provokes us to inquire what intelligence actually is and what can be considered genuine intelligence or “original” intentionality in a machine. It also makes us consider the amount of what is perceived as intelligent behaviour of an agent lies in the eyes of the observer. The worry Weizenbam expresses in his article is genuine and is needed to be considered carefully from both an AI and ethical point of view.

## References

- [1] Weizenbaum, Joseph. *Computer Power and Human Reason* (New York: Freeman, 1976)
- [2] Weizenbaum, Joseph. "ELIZA - A Computer Program for the Study of Natural Language Communication between Man and Machine," *Communications of the Association for Computing Machinery* 9 (1966): 36-45.
- [3] Suchman, Lucy A. *Plans and Situated Actions: The problem of human-machine communication* (Cambridge University Press, 1987)