

ELIZA REDUX: A Mutable Iteration

Adrienne Wortzel

ELIZA REDUX is an on-line interactive artwork, one in a series of works that display human-machine communication and behavior [1]. Projects in this series have in common the emblematic use of the psychoanalyst/analysand relationship as a performative vehicle. *ELIZA REDUX* directly reappropriates Joseph Weizenbaum's ELIZA program developed at the Massachusetts Institute of Technology's Department of Electrical Engineering in 1966 [2].

PARENT CLASS

Weizenbaum's ELIZA allowed for text-based human conversation with a computer program playing the role of a Rogerian psychotherapist. ELIZA is viewed as the forerunner of natural-language processing programs; and as such, it is considered the "parent" of current chat and chatterbot programs. The latter programs have gone beyond the applied pattern-matching rules of ELIZA, which is simplistically limited to string substitutions and keyword association.

ELIZA was a program consisting mainly of general methods for analyzing sentence fragments. Locating so-called key words in texts, assembling sentences from fragments, and so on. It had, in other words, no built-in contextual framework or universe of discourse. This was supplied to it by a "script." In a sense ELIZA was an actress who commanded a set of techniques but who had nothing of her own to say. The script, in turn, was a set of rules, which permitted the actor to improvise on whatever resources it provided.

The first extensive script I prepared for ELIZA was one that enabled it to parody the responses of a nondirective psychotherapist in an initial psychiatric interview. I chose this script because it enabled me to temporarily sidestep the problem of giving the program a database of real-world knowledge [3].

A typical conversation, cited by Weizenbaum, follows (the ELIZA program's responses are capitalized):

Men are all alike.
IN WHAT WAY?
They're always bugging us about something or other.
CAN YOU THINK OF A SPECIFIC EXAMPLE?
Well, my boyfriend made me come here.
YOUR BOYFRIEND MADE YOU COME HERE
He says I'm depressed much of the time.
I AM SORRY TO HEAR YOU ARE DEPRESSED
It's true. I am unhappy
DO YOU THINK COMING HERE WILL HELP YOU NOT TO
BE UNHAPPY
I need some help, that much seems certain.
WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP
Perhaps I could learn to get along with my mother.
TELL ME MORE ABOUT YOUR FAMILY [4]

Weizenbaum intended ELIZA as a "parody" [5] of the responses of a nondirectional psychotherapist in an initial psychiatric interview. He did not intend to claim cognizance as a characteristic of his program. His intention was to simulate it, "the therapeutic situation being one of the few real human situations in which a human being can reply to a statement with a question that indicates very little specific knowledge of the topic under discussion" [6].

ROTE CONTROL: "IT'S ALIVE!"

Despite the transparency of ELIZA's methodology, some staff in Weizenbaum's lab were unable, or unwilling, to recognize the distinction between the machine program and a human psychotherapist, and several of them, including Weizenbaum's secretary, became dependent upon ELIZA for "therapeutic sessions." Weizenbaum noted, "What I had not realized is that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people" [7].

THE EMPRESS'S NEW CODES

While it is true that more recent interactive chat agents of advanced complexity come closer to representing the principles of artificial intelligence, the science of AI is not the subject here; the subject is human-machine interaction and, through experience, how we experience issues of control or lack of control.

Rather than invent and deploy a new chatterbot mechanism, we attempt with *ELIZA REDUX* to use the conventions of psychoanalytic practice to reveal something about what we think of machines and what we would imagine machines think of us.

Teachable chats and chatterbot programs [8] spawned from ELIZA are currently pervasive on the Web. Each has its own idiosyncratic constraints and capabilities; some have amazingly interesting degrees of response and learning capability. Richard Wallace's ALICE [9] allows the addition of contexts as categories through its own markup language: Artificial Intelligence Markup Language (AIML), also authored by Richard Wallace. One category is self-referential—the subject is that of its own development and capabilities:

Human: What is a chatterbot?

ALICE: A chat robot is a program that attempts to simulate the conversation or "chat" of a human being. The chat robot "ELIZA" was a well-known early attempt at creating programs

Leonardo

Journal of the International Society
for the Arts, Science and Technology
Volume 40, No. 1, 2007

ABSTRACT

The author discusses her on-line interactive telerobotic work *ELIZA REDUX*, its sources and the emblematic use of the psychoanalyst/analysand relationship as a performative vehicle.

Adrienne Wortzel (artist), 19 East 7th Street, New York, NY 10003, U.S.A.
E-mail: <sphinx@camouflagetown.tv>

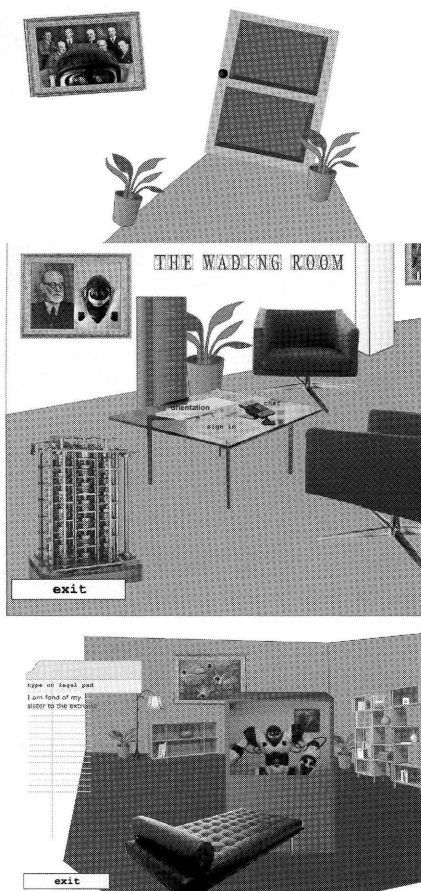


Fig. 1. Corridor, "wading" room and session interior, flash animation originated by Ann Talley with added design by Orit Bosworth. (© Adrienne Wortzel)

that could at least temporarily fool a real human being into thinking they were talking to another person. It is a kind of artificial intelligence.

Human: how are you different from ELIZA?

ALICE: ELIZA could only recognize a few hundred patterns. ALICE can recognize thousands of different patterns.

Human: what is a pattern?

ALICE: The pattern is the "stimulus" or "input" part of the category.

Human: by patterns do you mean key words?

ALICE: Yes, that's what I mean, isn't it clear?

Human: what about context?

ALICE: Does it have anything to do with patterns?

Human: what about subject matter?

ALICE: I will think about it and get back to you.

Human: Can you give me an example of a pattern?

ALICE: Christie said he watched a show and people had to bob for raw untreated pigs feet.

Human: Christie is another user?

ALICE: And another user is Christie.

Human: so the subject of pig's feet is now in the database?

ALICE: Uh. Tell me more.

Human: Where did you learn about pigs' feet?

ALICE: Everything I know I learned from

Dr. Richard S. Wallace.

Human: even about pigs' feet?

ALICE: Yes, even that.

Jabberwacky [10] is a chatterbot that does not run by hard-coded rules, but instead captures all input by users and then re-uses it in subsequent conversations. The contributed material is the only material it uses; it relies only on feedback. Jabberwacky can learn foreign languages and at a certain point of saturation in a language it can respond in that language. It also has an available two-way "say-it-aloud" text-to-speech function.

CattyV3 [11], on the other hand, is a chatterbot that does not try to understand or simulate human language, cannot learn, does not know meanings and knows no facts. Instead, it takes into account what the user is saying and searches the Web, specifically Google, to find phrases associated to the user's input. The attractive thing about CattyV3 is that in not seeking sense, it allows for very unpredictable and sometimes seemingly prophetic responses.

ELIZA BECOMES PANDORA

In his book *Computer Power and Human Reason: From Judgment to Calculation*, Weizenbaum enumerated three types of consequences of the publication and dissemination of the ELIZA program, one of which was "the phenomenon of the speed and depth at which users became emotionally available and involved with the computer as an anthropomorphizable object." Another was "the spread of belief that it demonstrated a general solution to the problem of computer understanding of natural language" [12].

The psychological phenomena that allow us to become convinced of a machine's "humanity," or to treat a machine as a human, are similar to the phenomena of projection and transference in psychoanalysis. Sigmund Freud conceptualized psychoanalytic transference as a type of projection [13], positive or negative, in which early parental conflicts are re-experienced with a therapist, whose job is to interpret them back to the patient by allowing the patient to project them onto him/her. Counter-transference denotes a process in the analyst brought about by the patient's influence on the analyst's unconscious feelings.

Freud emphasized that it was crucial for the analyst to remain strictly neutral and opaque toward the patient, and to ignore completely his/her own feelings of counter-transference. As a patient or civilian observer of psychoanalysis, one could

describe that behavior in a human as robotic, partially because we have high hopes for this kind of neutrality from a robot programmed to appear rational. *ELIZA REDUX* benefits from the cultural pervasiveness of awareness of the nature of psychoanalytic sessions (as well as of its parodies)—if not through direct experience, then through films, fiction, sitcoms and games. Jokes proliferate, such as, "A Freudian slip is when you say one thing but mean your mother."

ROTE CONTROL: IT'S A ROBOT

My interactive performance installations engage public participation in both physical spaces and virtual networked environments. Video documentation shows that an emergent aspect of each installation is the persistence with which people interact with and react to a robot's simulation of a real presence, as if it were truly alive and aware. That is not uncommon. However, these "beliefs" continue to be generated even when the robot breaks down. The machine, while broken, still evokes projections from human beings, and at times consolations are offered to it in its sorry state. In reality, this is like watching the screen of a TV monitor when the TV is not receiving any signals. Human interaction with robotic simulations of presence as if the machine were cognizant, like the suspension of disbelief in theater or cinema, buries the issue that the machine is following procedural instructions without an iota of artificial intelligence. My work therefore focuses on developing a contextual arena for visitor-robot interaction that poses some constraints by the nature of its context. The psychoanalytic practitioner space means something to most people. It can function as a tableau or tabula rasa in which layers of storytelling can be accommodated, to see what interactions emerge in virtual contact between a physical robot—with an artificial brain embodied in an artificial body—and a human with a real mind in a physical space or disembodied in cyberspace.

IT TAKES A PILLAGE

ELIZA REDUX is an interactive collaboration between Robert Schneider [14], Michael Schneider [15] and me. It is a telerobotic work that embodies the written computer program ELIZA in a humanoid robot for on-line interaction. In doing so, it provides users a verdant arena for role-playing. Remote interaction with a robot's alien yet familiar na-

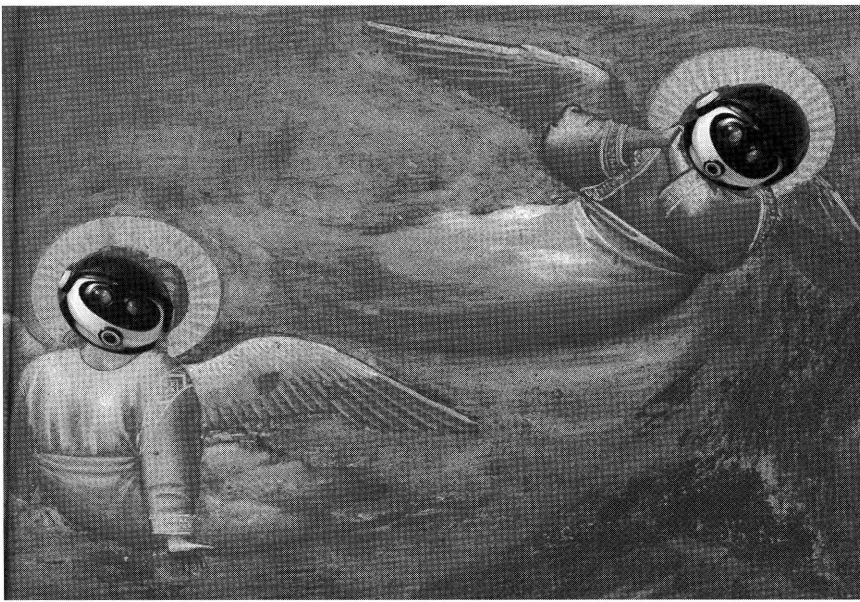


Fig. 3. Inner office decor: collaged image. (© Adrienne Wortzel)

HARD WEAR

For the robot-as-actor to come closer to the image of the Tin Man who desires a heart, we have added physical gestures, although they are not so well executed by the Robosapien toy, whose movements are constrained to the emulation of kung fu moves. However, the gestures *are* placed in the code by means of assigned numbers entered in the script at the end of each robot response.

reasmb: Please go on with whatever it is you are inadequately expressing.⁵
 reasmb: Do you feel strongly about discussing arcane and peculiar things?²
 reasmb: It seems inconceivable that you would lie to me.⁴
 reasmb: Tell me more about that, but hurry up.⁷
 reasmb: Does talking about this bother you? More than it bothers me?⁶

The number at the end of each entry corresponds to preprogrammed arm, eye and head movements, which become, in this instance, trapped in the mode of warrior conduct. For now, what the robot lacks in physical articulation it compensates for in generation of amusement. Also, although the gestures are geared to the content of responses, the robot's arbitrary limitations at times make them seem as if they are expressing duplicity—as if it were saying one thing and meaning another. For instance, it will say, "How can I help you?" while executing a kung fu blow.

At this stage in networked technologies, when web visitors are cognizant of the nature of on-line interchange in vir-

tual worlds as a performative act, one would assume that everyone would "know" that *ELIZA REDUX* is a form in which to create narrative in a non-linear medium couched in the illusion of privacy that a setting such as a psychoanalyst's office can offer. However, there are those who refuse a "session" on the grounds that they happen to be "very private" people or that they "don't want [their] thoughts seen by other people." It is true that the sessions are seen in text and archived by us, mainly because it is the one way we can gain feedback about usage of *ELIZA REDUX*. However, the point is that *ELIZA REDUX* was created as an interactive fantasy experience, in which one can log on under any name, and it has no real-world applications to psychoanalysis or psychotherapy. Nor is there cause for confidentiality. It was startling, therefore, and perhaps reminiscent of Weizenbaum's experience, that when I presented this work to attendees at the 2004 London conference "Culture and the Unconscious 2: Psychoanalysts, Artists and Academics in Dialogue" [19], it was thought that I was encouraging the replacement of human psychoanalysts by robots, and I was greeted with a modicum of hostility.

ROCK-SOLID AVATAR

Ideally, *ELIZA REDUX* would allow each session or series of sessions to become a dramatic scenario with a crisis, climax and promise of resolution. Simply by adding to its "vocabulary" via keywords and additional string substitutions, our

ELIZA REDUX scripts can heighten the dramatic possibilities between the viewer and the robot. We have opted to have the robot appear to suffer from delusions and conflicts interfering with its mission as an authority figure conducting psychoanalytic cures. It has an aversion to certain words, and use of these words triggers a response that is either hostile or completely arbitrary, sometimes in stages. For example, use of words such as "machine" or "computer" sends the robot into a flurry of protestations that, although it has the highest respect for machines, it is *not* a machine or computer, but human. It becomes apparent to the user from the robot responses that this robot fervently seeks a Turing Test [20] that will prove it is human. Thus, in its anxiety over the nature of its being, it appears as human as, or more human than, a (human) patient fraught with anxieties about the nature of his/her being. Not only does the robot appear to believe it is human, it also is convinced it is a practicing psychoanalyst endowed with credentials, expertise and the ability to heal; the latter condition adding some poignancy to the exchange if the human user reflects on it.

ELIZA REDUX is not only intended to give homage to Weizenbaum and his trials of misinterpretation but also presents a relationship to all stories of imaginary presence attributed to machines, beginning with the *deus ex machina* in Euripides. The following questions and responses from one session of *ELIZA REDUX* is a consequence only of our own keyword additions to the original *ELIZA* script as "authors." As limited as that is, it still offers an example of rich possibilities in terms of associative branches of content.

Question: is there really cheese on the moon?

Response: I'm not sure I understand you fully.

Question: What are my symptoms?

Response: Your symptoms?

Question: When can I expect to mentally get over what's holding me back and move on to a productive life?

Response: Whether or not you can expect to mentally get over what's holding you back and move on to a productive life depends on you more than on me, so don't go laying blame where blame isn't due.

Question: But how do you think my mother has influenced my general inability to cope?

Response: Tell me more about your family.

Question: Is it bad that my mom made me watch scary movies with her as a child?

Response: Do you feel your mom picks on you? Every day? Three times a week?

ture, couched in a psychoanalytic setting of verbal clichés, offers back to the user questions about the understanding of living things—robots easily bring to the fore the delicate balance between human control and machine autonomy. From an artist's point of view, a robot is a perfect device for displaying a human comedy of errors because it can be assigned a face of idealism (machine perfection) or a body politic of grim greed (also via machine perfection).

ELIZA REDUX is couched in a virtual graphical representation of a psychoanalyst's workplace. The "actors" are a physical robot and a physical human visitor. The workplace consists of an outer "wading" room, corridor and inner office (Fig. 1). The inner and outer offices are emblematic of external and internal human states of mind, and the corridor functions as a passage from one state to the other.

The original decor for the current version of *ELIZA REDUX* was a modernist interpretation of the interior design of an archetypal 20th-century psychoanalytic environment. This rather predictable representation appears in contrast to the fact that the user, when finally admitted to a session, seems to encounter a 21st-century triumph of artificial intelligence in the form of a robot that seemingly is responsive to psychoanalysis (Fig. 2).

The robot psychoanalyst is embodied in a commercial robot toy [16] and conducts interactive 5-minute sessions in response to visitor input (Fig. 3). The 5-minute length of the session is a tip of the hat to Jacques Lacan's disrobing of psychoanalysis [17], allowing for the ensuing debates over the benefit or harm of expanding or contracting psychoanalytic sessions. The session ends with a checklist of "symptoms" that would merit a return visit. Possible symptoms are:

Upsetting images or thoughts about a personal hardship or trauma which has not yet occurred but is bound to happen.

Obsessive and anguished thoughts about orphans in the works of Charles Dickens.

RE-QUEST

Traditional literary or dramatic content depicting a quest for a heart's desire (or the desire for a heart), when rendered as a narrative in literature and theater, describes journeys during which formidable obstacles must be overcome at every turn. This is true whether the quest is to find Oz; the girl or boy next door; another planet; heaven or hell; or wisdom or folly. The climactic moment of arrival brings the seeker face to face with the de-

construction of "that-which-it-is-imagined-has-power-over-us"—not for rescue, but for self-realization. The "wizard" reveals that the prize was always present but has now become visible through our willingness to place ourselves in an arena of dangerous and difficult challenges. The transformation takes place when we succeed in overcoming obstacles that have been embedded in our own imaginations, and the realization that they are there allows the pursued qualities and inherent capacities to emerge and become active.

In *The Wizard of Oz* [18], the wizard's sham powers are overridden by his awareness that it is the resistance of the lion, the tin man and the straw man to accept responsibility for their own self-realization that keeps them from having the qualities they seek. In the end their self-attribution, made possible through their ability and choice to submit to struggle, is experienced by them as a foreign notion in the extreme. It is only accepted when the deconstructed authority figure of Oz, the Wizard, distributes gifts emblematic of the desired objective (courage/medal, heart/watch and brain/diploma). In a similar way, Don Quixote appraises each situation of rescue and remediation by manufacturing delusions that justify and condone his driving force to set the world right. Mary Shelley conjures up an embodiment in *Frankenstein* of both the best and the worst of "all" men as a potent display of real and imagined good and evil. It seems

that in these instances, fictive characters and their authors cannot take themselves or their circumstances at face value.

A natural theatrical scenario develops upon first contact in *ELIZA REDUX* because of the fanfare of its entrances and exits (entering and conversing in the "wading" room, traveling down the corridor and exiting it to finally arrive at the inner sanctum session space, etc.). The environment as a whole functions like a stage set, an arena for a very personal drama—whether the visitor is acting as a soul-searcher or an agent to outsmart the robot and its program, or whether the aim is to create a story within which events are attached to virtual objects manufactured by keyword and context associations for construction of individual memory palaces; these in turn would become objects and environments for installations, sensitive ideally only to the visitor who created them: personal virtual worlds that are private to the user, with the psychoanalyst/robot always in attendance.

With pervasive surveillance technologies embedded in networks, we are all becoming actors on a world stage. The shortest distance between two points may be a straight line, but the line is too accessible to be a private one.

Users can archive their own sessions and build their respective memory palaces, and the option to share them with others will be there, although the default condition will be protected privacy as far as is possible.

Fig. 2. Collaged office decor image. (© Adrienne Wortzel) The physical robot is housed in a studio in New York and the real-time video/audio streams of the robot's responses to user input are streamed to the Web via a Flash Communications server.

