Wikipedia

ELIZA effect

The **ELIZA effect**, in computer science, is the tendency to unconsciously assume computer behaviors are analogous to human behaviors, that is anthropomorphisation.

Contents

Overview

In its specific form, the ELIZA effect refers only to "the susceptibility of people to read far more understanding than is warranted into strings of symbols—especially words—strung together by computers".[1] A trivial example of the specific form of the Eliza effect, given by Douglas Hofstadter, involves an automated teller machine which displays the words "THANK YOU" at the end of a transaction. A (very) casual observer might think that the machine is actually expressing gratitude; however, the machine is only printing a preprogrammed string of symbols.[1]

More generally, the ELIZA effect describes any situation[2][3] where, based solely on a system's output, users perceive computer systems as having "intrinsic qualities and abilities which the software controlling the (output) cannot possibly achieve"[4] or "assume that [outputs] reflect a greater causality than they actually do".[5] In both its specific and general forms, the ELIZA effect is notable for occurring even when users of the system are aware of the determinate nature of output produced by the system. From a psychological standpoint, the ELIZA effect is the result of a subtle cognitive dissonance between the user's awareness of programming limitations and their behavior towards the output of the program.[6] The discovery of the ELIZA effect was an important development in artificial intelligence, demonstrating the principle of using social engineering rather than explicit programming to pass a Turing test.[7]

Origin

The effect is named for the 1966 chatterbot ELIZA, developed by MIT computer scientist Joseph Weizenbaum. When executing Weizenbaum's *DOCTOR* script, ELIZA parodied a Rogerian psychotherapist, largely by rephrasing the "patient" 's replies as questions:

> Human: Well, my boyfriend made me come here.
> ELIZA: Your boyfriend made you come here?
> Human: He says I'm depressed much of the time.
> ELIZA: I am sorry to hear you are depressed.

> Human: It's true. I'm unhappy.
> ELIZA: Do you think coming here will help you not to be unhappy?[8]

Though designed strictly as a mechanism to support "natural language conversation" with a computer,[9] ELIZA's *DOCTOR* script was found to be surprisingly successful in eliciting emotional responses from users who, in the course of interacting with the program, began to ascribe understanding and motivation to the program's output.[10] As Weizenbaum later wrote, "I had not realized ... that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people."[11] Indeed, ELIZA's code had not been designed to evoke this reaction in the first place. Upon observation, researchers discovered users unconsciously assuming ELIZA's questions implied interest and emotional involvement in the topics discussed, *even when they consciously knew that ELIZA did not simulate emotion*.[12]

See also

- Turing test
- Loebner Prize
- Semiotics
- Chatbot
- Intentional stance
- Uncanny valley
- Philosophical zombie

Notes

1. Hofstadter, Douglas R. (1996). "Preface 4 The Ineradicable Eliza Effect and Its Dangers, Epilogue" (https://books.google.com/books?id=somvbmHCaOEC&pg=PA157). *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought* (https://books.google.com/books?id=somvbmHCaOEC). Basic Books. p. 157. ISBN 978-0-465-02475-9.
2. Fenton-Kerr, Tom (1999). "GAIA: An Experimental Pedagogical Agent for Exploring Multimodal Interaction". *Computation for Metaphors, Analogy, and Agents*. Springer. p. 156. doi:10.1007/3-540-48834-0_9 (https://doi.org/10.1007%2F3-540-48834-0_9). "*Although Hofstadter is emphasizing the text mode here, the "Eliza effect" can be seen in almost all modes of human/computer interaction.*"
3. Ekbia, Hamid R. (2008). *Artificial Dreams: The Quest for Non-Biological Intelligence* (https://books.google.com/books?id=4wEIlNjO0x0C&pg=PA8). Cambridge University Press. p. 8. ISBN 978-0-521-87867-8.
4. King, W. (1995). *Anthropomorphic Agents: Friend, Foe, or Folly* (Technical report). University of Washington. M-95-1.
5. Rouse, William B.; Boff, Kenneth R. (2005). *Organizational Simulation* (https://books.google.com/books?id=371wV4dI7ckC&pg=PA308). Wiley-IEEE. pp. 308–309. ISBN 978-0-471-73943-2. "*This is a particular problem in digital environments where the "Eliza effect" as it is sometimes called causes interactors to assume that the system is more intelligent than it is, to assume that events reflect a greater causality than they actually do.*"
6. Ekbia, Hamid R. (2008). *Artificial Dreams: The Quest for Non-Biological Intelligence* (https://books.google.com/books?id=4wEIlNjO0x0C&pg=PA156). Cambridge University Press. p. 156. ISBN 978-0-521-87867-8. "*But people **want** to believe that the program is "seeing" a football game at some plausible level of abstraction. The words that (the program) manipulates are so full of associations for readers that they CANNOT be stripped of all their imagery. Collins of course knew that his program didn't deal with anything resembling a two-dimensional world of smoothly moving dots (let alone simplified human bodies), and presumably he thought that his readers, too, would realize this. He couldn't have suspected, however, how powerful the Eliza effect is.*"

7. Trappl, Robert; Petta, Paolo; Payr, Sabine (2002). *Emotions in Humans and Artifacts* (https://books.google.com/books?id=jTgMIhy6YZMC&pg=PA353). Cambridge, Mass.: MIT Press. p. 353. ISBN 978-0-262-20142-1. "*The "Eliza effect" — the tendency for people to treat programs that respond to them as if they had more intelligence than they really do (Weizenbaum 1966) is one of the most powerful tools available to the creators of virtual characters.*"
8. Güzeldere, Güven; Franchi, Stefano. "dialogues with colorful personalities of early ai" (http://www.stanford.edu/group/SHR/4-2/text/dialogues.html). Archived (https://web.archive.org/web/20110425191843/http://www.stanford.edu/group/SHR/4-2/text/dialogues.html) from the original on 2011-04-25. Retrieved 2007-07-30.
9. Weizenbaum, Joseph (January 1966). "ELIZA--A Computer Program For the Study of Natural Language Communication Between Man and Machine" (http://www.csee.umbc.edu/courses/331/papers/eliza.html). *Communications of the ACM*. Massachusetts Institute of Technology. **9**: 36. doi:10.1145/365153.365168 (https://doi.org/10.1145%2F365153.365168). Retrieved 2008-06-17.
10. Suchman, Lucy A. (1987). *Plans and Situated Actions: The problem of human-machine communication* (https://books.google.com/books?id=AJ_eBJtHxmsC&dq=Suchman+Plans+and+Situated+Actions). Cambridge University Press. p. 24. ISBN 978-0-521-33739-7. Retrieved 2008-06-17.
11. Weizenbaum, Joseph (1976). *Computer power and human reason: from judgment to calculation* (https://archive.org/details/computerpowerhum0000weiz). W. H. Freeman. p. 7 (https://archive.org/details/computerpowerhum0000weiz/page/7).
12. Billings, Lee (2007-07-16). "Rise of Roboethics" (http://www.seedmagazine.com/news/2007/07/rise_of_roboethics.php). Seed. Archived (https://web.archive.org/web/20090228092414/http://www.seedmagazine.com/news/2007/07/rise_of_roboethics.php) from the original on 2009-02-28. "(Joseph) Weizenbaum had unexpectedly discovered that, even if fully aware that they are talking to a simple computer program, people will nonetheless treat it as if it were a real, thinking being that cared about their problems – a phenomenon now known as the 'Eliza Effect'."

References

- Hofstadter, Douglas. *Preface 4: The Ineradicable Eliza Effect and Its Dangers.* (from *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*, Basic Books: New York, 1995)
- Turkle, S., Eliza Effect: tendency to accept computer responses as more intelligent than they really are (from *Life on the screen- Identity in the Age of the Internet*, Phoenix Paperback: London, 1997)
- ELIZA effect (http://www.catb.org/~esr/jargon/html/E/ELIZA-effect.html), from the Jargon File, version 4.4.7. Accessed 8 October 2006.

This article is based on material taken from the *Free On-line Dictionary of Computing* prior to 1 November 2008 and incorporated under the "relicensing" terms of the GFDL, version 1.3 or later.

Retrieved from "https://en.wikipedia.org/w/index.php?title=ELIZA_effect&oldid=926293618"

This page was last edited on 15 November 2019, at 11:48.