

To what extent are Joseph Weizenbaums arguments still relevant today

Prof. Emeritus Arild Jansen,
Sender for rettsinformatikk, IFP, UiO
arildj@jus.uio.no

UNIVERSITETET
I OSLO

Short CV

- Born in Berlin, 1923. Jewish, immigrating to the U.S. in 1936
- Studying mathematics at Waine State University in Detroit from 1941
- Serve in the U.S. Army in as a meteorologist 1942v, (turned down for cryptology work because of his "enemy alien" status.
- Returned to Wayne State, B.S (1948) in Mathematics MS (1950)
- As research assistant, Weizenbaum helped create a digital computer.
- Worked for General Electrics on ERMA, a computer system allowing automated check processing via Magnetic Ink Character Recognition
- Moved to MIT in 1964; gradually he got various academic positions (Harvard, Stanford and many others
- Moved to Berlin in 1986, and was buried at the Weissensee Jewish Cemetery. ¹ A memorial service was held in Berlin on 18 March 2008.

UNIVERSITETET
I OSLO

The main headlines in the book

- ➔ «Doctor & Eliza» as basis for computer models in psychology
- ➔ Science and the Compulsive Programmer
- The computer and natural languages
- ➔ Artificial intelligence and human thinking
- Incomprehensible programs
- ➔ Against the imperialism of Instrumental reason

The intention of this seminar is to present some of Weizenbaum's contributions in computer science and not least recall some of his arguments in the debate about the limitations of computers in imitating human thinking and judgement.

UNIVERSITETET
I OSLO

Some of Weizenbaums basic questions

In order to understand how the computer attained so very much power, both as an actor and as a force on human imagination, we must first discuss:

1. What is it about computers that has brought the *view of man as a machine* to a new level of plausibility?
2. The fact that individuals bind themselves with strong emotional ties to machines ought not in itself to be surprising - but why has man come to yield his own autonomy to a world viewed as machine?
3. Man has ceased to believe in his own autonomy and has begun to rely on autonomous machines. But how can humans explain to himself what these machines can do and even how they do what they do!



'Doctor' and 'ELIZA'

- 1966: Released [ELIZA](#), written in his own "SLIP" program language, performing natural language processing
- The program applied pattern matching rules to statements to figure out its replies. (Similar to many of chatbots to day?)
 - Using a script (DOCTOR), it could engage humans in conversations, resembling a session with an empathic psychologist.
- He got shocked to learn how users opened their hearts to the system:
 - His secretary, when using it, asked Weizenbaum: "would you mind leaving the room please?"
- Many hailed the program as a 'thinking machines', a misguided interpretation that Weizenbaum attempt to correct



How did ELIZA Work


The functioning of ELIZA was rather simple:

- ELIZA searched the text submitted by its conversation partner for relevant keywords. When a keyword or pattern was found, it produced an appropriate response according to specific transformation rules, based on a two-stage process:
 1. The input was first decomposed, broking down the sentence into small segments.
 2. The segments were reassembled, readapted according to appropriate rules – for instance, by substituting the pronoun 'you' with 'I' – and programmed words were added to produce a response.
- In addition, when it was impossible to recognize a keyword, it would employ preconfigured formulas, such as '*I see*' or '*Please go on*', or alternatively create a response through a 'memory' structure that drew from previously inserted inputs ([Weizenbaum, 1966: 37](#))



A typical session : (ELIZA answers in UPPER CASE)

- Men are all alike
 - IN WHAT WAY
- They are 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 must of the time
 - I AM SORRY TO HEAR YOU ARE DEPRESSED
- It is 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
-

 UNIVERSITETET
I OSLO


Eliza as a forerunner of AI
or just a «smart» piece of software ?

ELIZA as an software object was interpreted in two different, even contrasting visions:

1. A behavioural approach: ELIZA (as a chatbot) provides only the appearance of intelligence;
 - capable of conducting convincing conversations, without questioning how this result is reached
2. Capable of actually *replicate intelligence* and *understanding* by artificial means.

However, Weizenbaum contended that ELIZA not exhibited intelligence, only provided *the illusion* of it.

Ref.: If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA - Simone Natale, 2019 (sagepub.com)

 UNIVERSITETET
I OSLO

Automation of psychotherapy


However, many psychiatrists believed that ELIZA could grow into automated psychotherapy:

"[...] If this method proves beneficial, then it would provide a therapeutic tool which can be made widely available to mental hospitals and psychiatric centers suffering shortage of therapists. Because of the time-sharing capabilities of modern computers, several hundred patients an hour could be handled

[...] The human therapist, involved in the design and operation, would not be replaced, but would become a much more efficient man since his effort would no longer be limited to the one-to-one patient-therapist ratio as now exists

[...] A human therapist can be viewed as an information processor and decision maker with a set of decision rules which are closely linked to short-term and long term goals .. He is guided in these decisions through rough empiric rules telling him what is appropriate to say and not to say in certain context.

Colby, K.M., J.B. Watt and J.P. (1966) : ELIZA – a computer method of psychotherapi; Preliminary Communication : Journal of nervous and Mental Disease, vol 142 (2) 1966

 UNIVERSITETET
I OSLO

Weizenbaums own reflections to these visions

"I had though it essential, as a prerequisite to the very possibility that one person might help another learn to cope with his emotional problems, that the helper himself participate in the other's experience of those problems, and, in large parts by way of his own empathic recognition of them, himself come to understand them." [Weizenbaum, 1976, p 8-9

Further, a comment to Colby et al's view of the therapist as an information processor

"What can the psychiatrist's image of his patient be when he sees himself, not as an engaged human being acting as a healer, but as an information processor following rules!



10

A new mental disorder: The compulsive programmer

- A professional programmer aims at solve a problem, discussing the problems with other and leave the computer when finished
- A compulsive programmer, however sees the problem as an opportunity to interact (or live) with the computer:

"Bright, young men of dishevelled appearance, sunken glowing eyes, rumpled cloths, unshaven faces and uncombed hair; sitting in front of the console for more than twenty ours, having only cokes and sandwiches. [...] They prefers complex programs with grandiose but imprecisely stated goals, they hardly discuss with others or document the programs.

"He are struggling with two opposing facts: He knows he can make the computer do anything he want it to do, while the computer constantly displays undeniable evidence of his failure to them"



11

Science and the compulsive programmer

The reason why Weizenbaum were so interested in the compulsive programmer was that he found a continuity between this pathological behaviour and those of the modern scientist (and technologist in general)

Citing E. Bergler, he see that many scientists act as gamblers:

1. He is certain that he will win
2. He has unbounded faith in his own cleverness
3. He believe that every aspect of life can and nature can finally be explained in exclusively scientific terms:

Any contradiction between a particular scientific notion and the fact of experience will be explained by other scientific notions : there is a ready reserve of possible scientific hypothesis available to explain any conceivable event .. (Polanyi 1964, cited by Weizenbaum)



12

On the AI debate: Computers can not act as humans

An organism is defined, in large parts, by the problem it faces. *Man* faces problems no machine could possibly be made to faces. Thus, computers and men are not species of the same genus.

There are things people come to know only as a consequence of having been *treated as a human beings*, not *least* the context –specific nature of information provided in communication between humans

Intelligence is a vague and “meaningless” concept in and of itself, as it requires a frame of reference, including the cultural and social setting.

The ‘artificial intelligentsia’ argue that *there is no domain of human thought over which machines cannot range*, by e.g. saying: There is nothing man know that cannot (in principle) be made accessible to a computer

Weizenbaum claims: It is NOT obvious that all human knowledge is encoded in “information structures”, e.g. such as intuition and emotional responses to other humans.



UNIVERSITETET
I OSLO

13

Can computer systems be understood as organisms ?

■ Given that e.g. a robot both can sense and affect its environment, Weizenbaum admit that robots can (in a limited sense) be “socialized”, that is being modified by its experience with the world.

Then, if both machines and humans are socializable, what is then the difference ?

“Every organisms is socialized by the process of dealing with problems that confronts it. Every species will (if only for that reason) be socialized differently

The growth of a child, including the separation from the mother is unique, and cannot be modelled by information processing today, and he continues:

There appears to be no prospect whatever mankind will know enough neurophysiology within the next several hundred years to have the intellectual basis for designing such a machine”



UNIVERSITETET
I OSLO

14

What is it to be a human

“The lesson here is rather that the part of the human mind which communicates to us in rational and scientific terms is itself an instrument that disturbs what it observes, particularly its voiceless partner, the unconscious, between which and our conscious selves it mediates”

“The human individual is in a constant state of becoming; the maintenance of that state, of his humanity, indeed of his survival, depends crucially on his seeing himself, and on his being seen by other human being, as a human being [..]”

“Computers can make juridical decisions, computers can make psychiatric judgements. [...] But they ought not be given such tasks. They may even be able to arrive at ‘correct’ decisions in some cases – but always and necessarily on base no being should be willing to accept.”



UNIVERSITETET
I OSLO

15

Weizenbaum asked for a broader debate on role of science in society

- A large group (the 'artificial intelligentsia') do believe computer can, should and will do everything, but also
- those (e.g. Mumford, Ellul, Boulding, Dreufus..) like myself, believing that there are limits to what computers ought to be put to do

The question is whether or not every aspect of human thought is reducible to a logical formalism, that is entirely computable. Mans autonomy and his corresponding responsibility is a central issue; there are important differences between men and machines *as thinkers*:
"Scientific statements can never be certain, they can only be more or less credible. How then can science, which itself surely and ultimately rests on vast arrays of human value judgements, demonstrate that human value judgement are illusory [...] that science has become the sole legitimate form of understanding in the common wisdom

The crucial distinction between deciding and choosing.

- The last chapters displays Weizenbaum's ambivalence towards what tasks computer technology should be used to perform
- Weizenbaum asserts that the definition of tasks and the selection of criteria for their completion is a creative *act* that relies on human values
- *Deciding* is a computational activity, something that can ultimately be programmed.
- *Choice*, however, is the product of judgment, not calculation.
- In deploying computers to make decisions that humans once made, the agent doing so has made a *choice* based on their values that will have particular, non-neutral consequences for the *subjects who will experience the outcomes of the computerized decisions that the agent has instituted.*

Science as 'poison'

Science may not only be seen as *addictive drug* [...] it has been converted to a slow-acting *poison*. By attributing certainty to it we has virtually delegitimized other ways of understanding.

While man in the past viewed arts, especially literature as sources of intellectual nourishment, arts are today perceived as largely entertainment :

[...] *The ancient Greek, the Oriental theatre the Shakespearian, the Ibsen's and Chekhov's stages were the schools and the curricula they taught were in the past vehicles for understanding the societies they represented.*

Today, we can count, but we are rapidly forgetting how to say what is worth counting and why!

His views on the future

In an interview with MIT's *The Tech*, Weizenbaum elaborated on his fears, expanding them beyond the realm of mere artificial intelligence, explaining that his fears for society and the future of society were largely because of the computer itself.

Despite working so closely with computers for many years, Weizenbaum frequently worried about the negative effects they would have on the world, particularly with regards to the military, calling the computer "a child of the military." He made it clear that he did not think of himself as a pacifist, believing that there are certainly times where arms are necessary, but by referring to defence as computer-controlled killings and bombings, humanity as a whole would be less inclined to embrace violent reactions so quickly.



19

Does Weizenbaums thinking provide new arguments today?

The view of computer as both a power and an actor illustrates Weizenbaum's perspective when discussing the role of technology in shaping the society

Not only do computers change society at large, but it do also (through mutual influence) shape man as individuals; what we do, how we reason and judge, and even (to a certain extent) attain our values to automated decisions

His way of reasoning corresponds to the STS/ANT tradition, but it has taken time before being widely accepted

However, I am not in a position to evaluate all his critical arguments to the "achievements" of AI, but I believe I do support many of them.

Some relevant references:

[Joseph Weizenbaum: A Parrhesiastes in the Digital Age \(capurro.de\)](#)

[Joseph Weizenbaum - Wikipedia](#)

[Computer Pioneers - Joseph Weizenbaum](#)

[Joseph Weizenbaum, Professor emeritus of computer science, 85 | MIT News | Massachusetts Institute of Technology](#)

20
