

Impact of Language and Technology 2

LIN 313: Language and Computers

The University of Texas at Austin

Jessy Li

With material from Katrin Erk

Self-perception

- What are the differences between human language and animal communication systems?
 - Arbitrary form-meaning relationships
 - Cultural transmission
 - Innate language faculty (not proven)
 - Productivity and compositionality
 - Displacement
 - Metalinguistics

Self-perception

- Language is tightly connected to what it means to be human. So what does it mean for the way we see ourselves when computers use language to interact with us?
- Joseph Weizenbaum, ELIZA inventor, “Computer Power and Human Reason”

Three surprises for Weizenbaum

- From practicing psychiatrists, e.g., Colby:
 - *“Further work must be done before the program will be ready for clinical use. If the method proves beneficial, then it would provide a therapeutic tool which can be widely available to mental hospitals and psychiatric centers offering a shortage of therapists. Because of the time-sharing capabilities of modern and future computers, several hundred patients an hour could be handled by a computer system designed for this purpose. The human therapist, involved in the design and operation of this system, would not be replaced, but would become a much more efficient man since his efforts would no longer be limited to the one-to-one patient-therapist ratio as now exists.”*
- Weizenbaum states that there are some things that computers ought not to do, like take on the role of psychiatrists. Would you agree?

Three surprises for Weizenbaum

- People quickly and deeply anthropomorphized ELIZA:
 - *“Once my secretary, who had watched me work on the program for many months and therefore surely knew it would be merely a computer program, started conversing with it. After only a few interchanges with it, she asked me to leave the room. ...clear evidence that people were conversing with the computer as if it were a person who could be appropriately and usefully addressed in intimate terms.”*
- If the secretary understood perfectly well how Eliza was made, would she still converse with Eliza like a real psychotherapist?
- Should people aim to understand all technology that they use? Would it make a difference?
- How can we make machine learning systems more explainable, technically?
 - https://en.wikipedia.org/wiki/Explainable_Artificial_Intelligence

Three surprises for Weizenbaum

- *“Another widespread, and to me surprising, reaction to the ELIZA program was the spread of a belief that it demonstrated a general solution to the problem of computer understanding of natural language. In my paper, I had tried to say that no general solution to that problem was possible, i.e., that language is understood only in contextual frameworks, that even these can be shared by people to only a limited extent, and that consequently even people are not embodiments of any such general solution.”*
- Will computers ever understand language?
- Do people not understand language?

Weizenbaum's questions

- What is it about the computer that has brought the view of man as a machine to a new level of plausibility?
 - Back then, artificial intelligence systems were rule-based. Weizenbaum argues against a view of the world, and of people, that is "logic-based", basically against a rule-based view of humans. Nowadays, they are machine learning based.
- Is there a new view of humans that is still based on a "humans are like computers" metaphor, but more similar to machine learning systems?
- Do you find an "intelligent" system more plausible when it is based on machine learning than completely rule-based?

Weizenbaum's questions

- “Why did you do this”?
 - Robot answer: “Because I was programmed to do so”.
 - Human answer: “Because I chose to”.
- What makes us human?
- Science and technology brought much knowledge and reason, and there has been a wave of rationalism. How has this impacted society?

Weizenbaum's questions

- Weizenbaum lists speech recognition as something we should not do, because it also involves natural language understanding. He asks what can be the use of that, and the most reasonable thing that occurs to him is dictation of medical notes. But he warns that the government may use such technology to spy on us all.
- What do you think of this argument? Weizenbaum's reasoning seems a bit quaint now — but let us ask anyway, is natural language understanding something we should not attempt to do?

Three surprises for Weizenbaum

- “... This reaction to ELIZA showed me more vividly than anything I had seen hitherto the enormously exaggerated attributions an even well-educated audience is capable of making, even strives to make, to a technology it does not understand.”
- What does this entail on the society as a whole?
- What are the scientists’ responsibilities to making his work public? And to whom (or what) is the scientist responsible?

Conflicting tensions

- Are computational linguists to bend to the will of industrial needs or should a “higher” goal of science be kept in mind?
- Should we focus on tasks that make life more convenient or on tasks that are “interesting”?
- Is it better to focus on getting the job done, or on doing things “right”?

The algorithm did it!

- Algorithms and the data that drive them are designed and created by people -- There is always a human ultimately responsible for decisions made or informed by an algorithm.
- "The algorithm did it" is not an acceptable excuse if algorithmic systems make mistakes or have undesired consequences, including from machine-learning processes.
- Discussion: what are accountability criteria needed for algorithmic systems?

Accountability criteria

- **Responsibility:** Make available externally visible avenues of redress for adverse individual or societal effects of an algorithmic decision system, and designate an internal role for the person who is responsible for the timely remedy of such issues.
- **Explainability:** Ensure that algorithmic decisions as well as any data driving those decisions can be explained to end-users and other stakeholders in non-technical terms.
- **Accuracy:** Identify, log, and articulate sources of error and uncertainty throughout the algorithm and its data sources so that expected and worst case implications can be understood and inform mitigation procedures.
- **Auditability:** Enable interested third parties to probe, understand, and review the behavior of the algorithm through disclosure of information that enables monitoring, checking, or criticism, including through provision of detailed documentation, technically suitable APIs, and permissive terms of use.
- **Fairness:** Ensure that algorithmic decisions do not create discriminatory or unjust impacts when comparing across different demographics (e.g. race, sex, etc).