Computer Science 1400: Part #8:

Where We Are:
Artificial Intelligence

WHAT IS ARTIFICIAL INTELLIGENCE (AI)?

AI IN SOCIETY

RELATING WITH AI

What is Artificial Intelligence (AI)?

Artificial Intelligence (Merriam-Webster):
1. a branch of computer science dealing with the simulation of intelligent behavior by computers.
2. the capability of a machine to imitate intelligent human behavior.

- Two flavors of AI:
  - **Strong AI**: Design computer systems that demonstrate full human-level intelligence using "same" mechanisms.
  - **Weak AI**: Design computer systems that demonstrate human-like abilities using any mechanisms.

Artificial Intelligence: Beginnings

Town Clock
(Munich; 1500s)

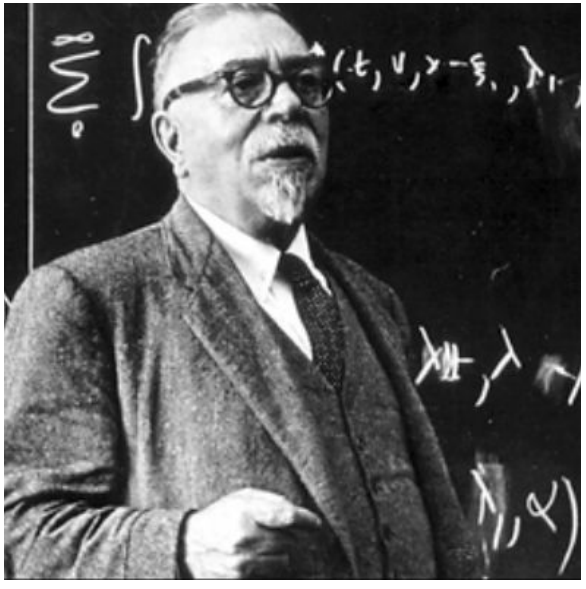
Mechanical Duck
(1739)

Mechanical Turk
(1770)

- First AI artifacts are mechanical automata which simulate various intelligent processes, *e.g.*, movement, reasoning.

Artificial Intelligence: The 1940s

Warren McCulloch and Water Pitts
(1898–1969 / 1923-1969)



Norbert Wiener
(1894–1964)

- Initial focus on natural models of neural (McCulloch-Pitts) and homeostatic (Wiener) processes.

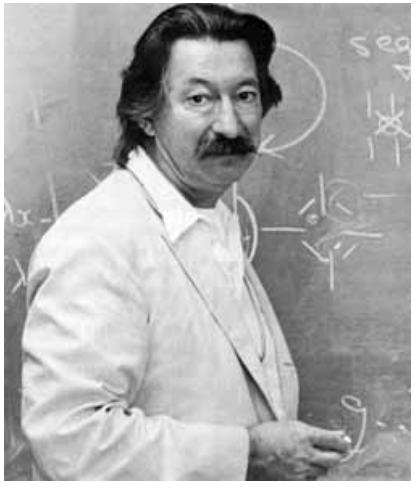
Artificial Intelligence: The 1950s

John McCarthy
(1927–2011)

Allen Newell and Herb Simon
(1927–1992 / 1916-2001)

- AI born at Dartmouth Conference in 1956 (McCarthy).
- Focus shifts to abstract information-processing models (*e.g.*, General Problem Solver (GPS) (Newell-Simon)).

Artificial Intelligence: The 1960s



Joe Weizenbaum
(1923–2008)

Marvin Minsky and Seymour Papert
(1927–2016 / 1928–)

- Information-processing-based AI systems proliferate (*e.g.*, ELIZA (Weizenbaum)); first rule-based expert system created (*e.g.*, MYCIN); first-generation neural network research killed off by Minsky and Papert.

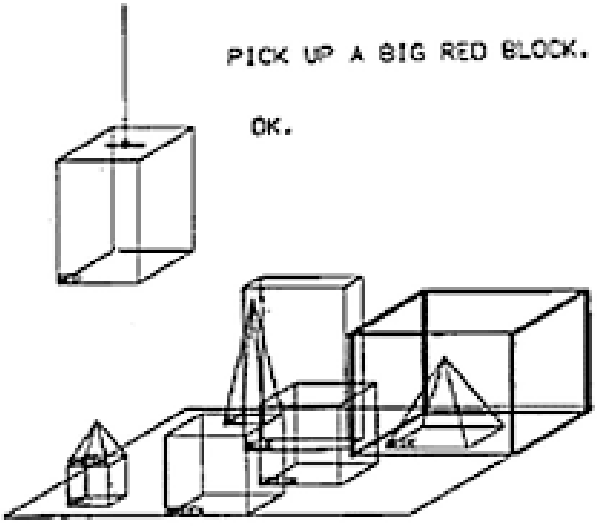
Artificial Intelligence: The 1960s (Cont'd)

The LOGO Turtle (1969)

Artificial Intelligence: The 1960s (Cont'd)

Shakey (1969)

Artificial Intelligence: The 1970s



SHRDLU (1970)

Hubert Dreyfus (1929–)

- Retreat to “toy” micro-world systems (*e.g.*, SHRDLU); emergence of AI critics into popular culture (*What Computers Can’t Do* (1972) (Dreyfus); *Computer Power and Human Reason* (1976) (Weizenbaum)).

Artificial Intelligence: The 1980s

Rodney Brooks (1954–)

Genghis (1989)

- Second-generation neural network research begins; rise of reactive systems (*e.g.*, Genghis (Brooks)); massive governmental (Fifth Generation Project (MITI: Japan) / Strategic Computing Initiative (DARPA: USA)) and industrial start-up funding
- Over-selling leads to crash and late 1980s “AI Winter”.

Artificial Intelligence: The 1990s

Gary Kasparov vs. IBM's Deep Blue (1997)

Artificial Intelligence: The State of the Art

- Three types of techniques:
  1. Search-based: GPS, Theorem Proving
  2. Knowledge-based: Rule-based Expert Systems, Automated Reasoning
  3. World-based: Reactive Robotics, Machine Learning
- Original goal in 1956 was Strong AI, which is very hard; is now usually Weak AI in which heuristics (*e.g.*, Google search) and/or brute-force processing (*e.g.*, IBM's Watson) are used to get human-level speed if not accuracy.
- Notable successes wrt particular domains, *e.g.*, natural language processing, autonomous vehicles.

Artificial Intelligence: The State of the Art (Cont'd)

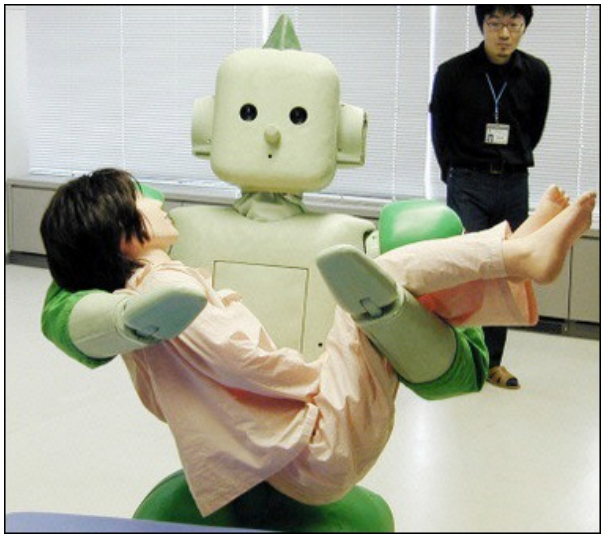
IBM's Watson wins *Jeopardy* (2011)

Artificial Intelligence: The State of the Art (Cont'd)

Google's self-driving car (2016)

The Joys of Artificial Intelligence

- Easier / more natural interaction with computers on focused topics (*e.g.*, psychological / medical advice).
- Replacement of humans with computers in physically demanding / dangerous / non-rewarding situations (*e.g.*, battlefield, child / elder care).
- Long-overdue re-assessment of the nature of humanity.



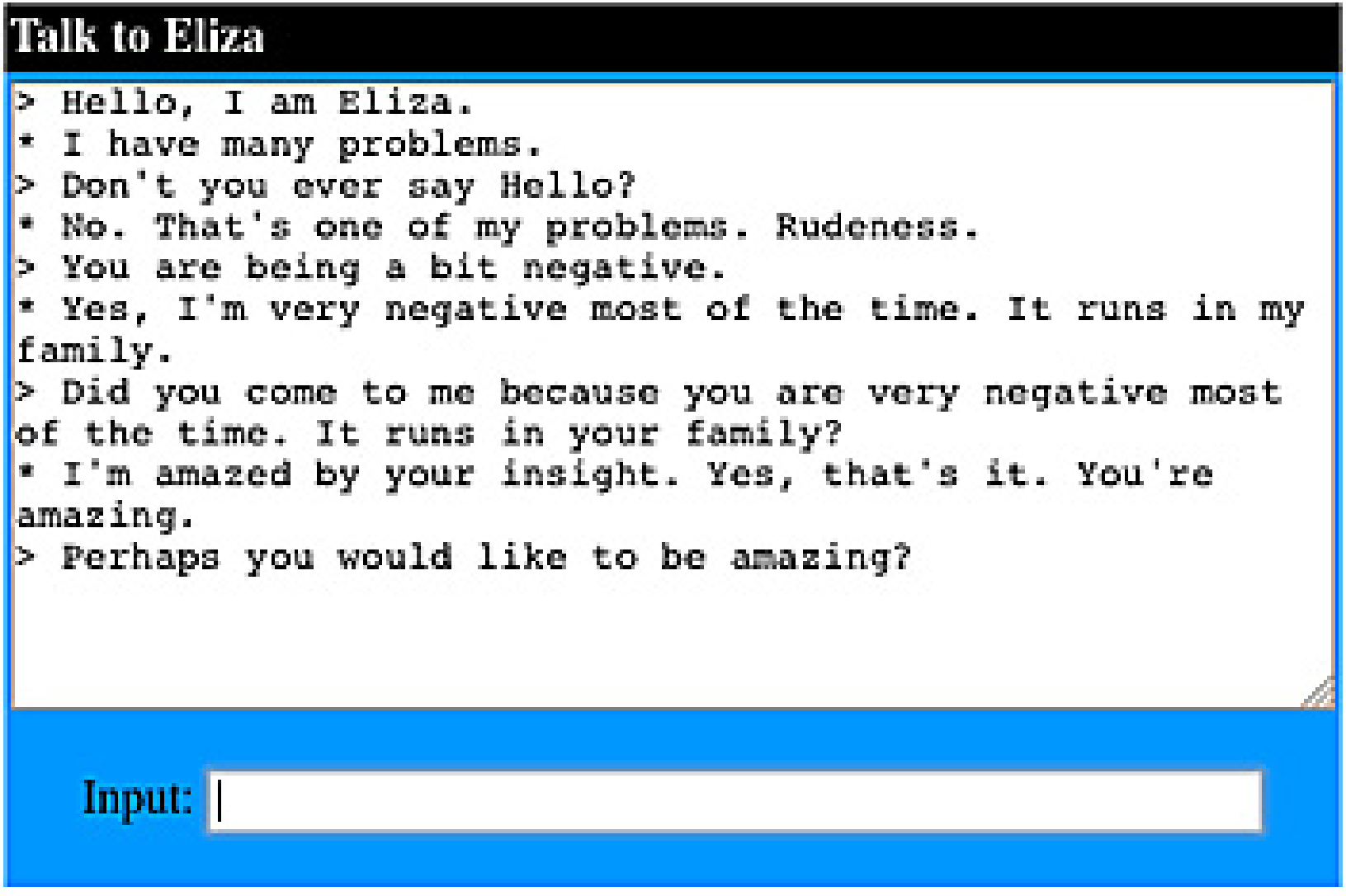
The Perils of Artificial Intelligence

- Psychological or physical trauma from assumption of intelligence and/or understanding where none is present (*e.g.*, chatbots, battlefield robots).
- Lowering of human standards for treatment of other humans (*e.g.*, child / elder care)

Case Study: ELIZA and Other Chatbots

- Created by Joe Weizenbaum in the mid-1960's as a simulation of a Rogerian psychotherapist.
- Simulates intelligent conversation using pattern-matching and response frames (same mechanisms used by modern chatbots).
- Many human beings trust and confide in ELIZA, even when they know ELIZA does not and cannot understand them or their problems.
- Similar confusion when dealing with modern chatbots, *e.g.*, falling in love with a chatbot (Epstein (2007)).

Case Study: ELIZA and Other Chatbots (Cont'd)



Surviving and Thriving with Artificial Intelligence

- Know actual (and do not over- or under-estimate) capabilities of AI systems.
- Beware of exaggerated claims of AI system abilities.
- Until AI systems are actually sentient and capable of being responsible for their actions, assign responsibility to the creators of these systems, not the systems themselves.
- Do not over- or under-estimate the abilities or value of human beings – we may only be mechanisms, but we are beautiful and powerful mechanisms worthy of respect.

"Don't Panic" – *The Hitchhiker's Guide to the Galaxy*
"Let's be careful out there" – *Hill Street Blues*