Science 1000: Part #6 (Wareham):

Where We Are:
Artificial Intelligence

WHAT IS ARTIFICIAL INTELLIGENCE (AI)?

AI IN SOCIETY

RELATING WITH AI

What is Artificial Intelligence (AI)?

- Computer-based AI research officially started in 1956.
- 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.
- Original goal in 1956 was Strong AI, which turned out to be very hard to achieve; is now most often Weak AI in which heuristics (*e.g.*, Google search) or massive 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, robot motion / driving.

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.
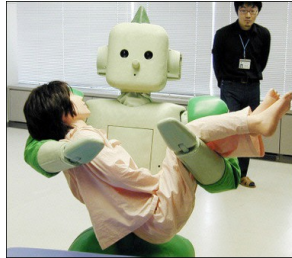
The Perils of Artificial Intelligence

- 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)
- Lowering of human valuation of humanity

Case Study: ELIZA and Other Chatbots

- Created by Joseph 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 can exist when dealing with modern chatbots, *e.g.*, falling in love with a chatbot.

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

Talk to Eliza

```
> Hello, I am Eliza.
* I have many problems.
> Don't you ever say Hello?
* No. That's one of my problems. Rudeness.
> You are being a bit negative.
* Yes, I'm very negative most of the time. It runs in my
family.
> Did you come to me because you are very negative most
of the time. It runs in your family?
* I'm amazed by your insight. Yes, that's it. You're
amazing.
> Perhaps you would like to be amazing?
```

Input:

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, trust in and assign responsibility to the creators of AI systems, not the systems themselves.
- Do not over- or under-estimate the abilities or value of human beings – we may be only biological mechanisms, but we are beautiful and powerful mechanisms.

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