Aspirational AI
or "Extreme" AI Ideas

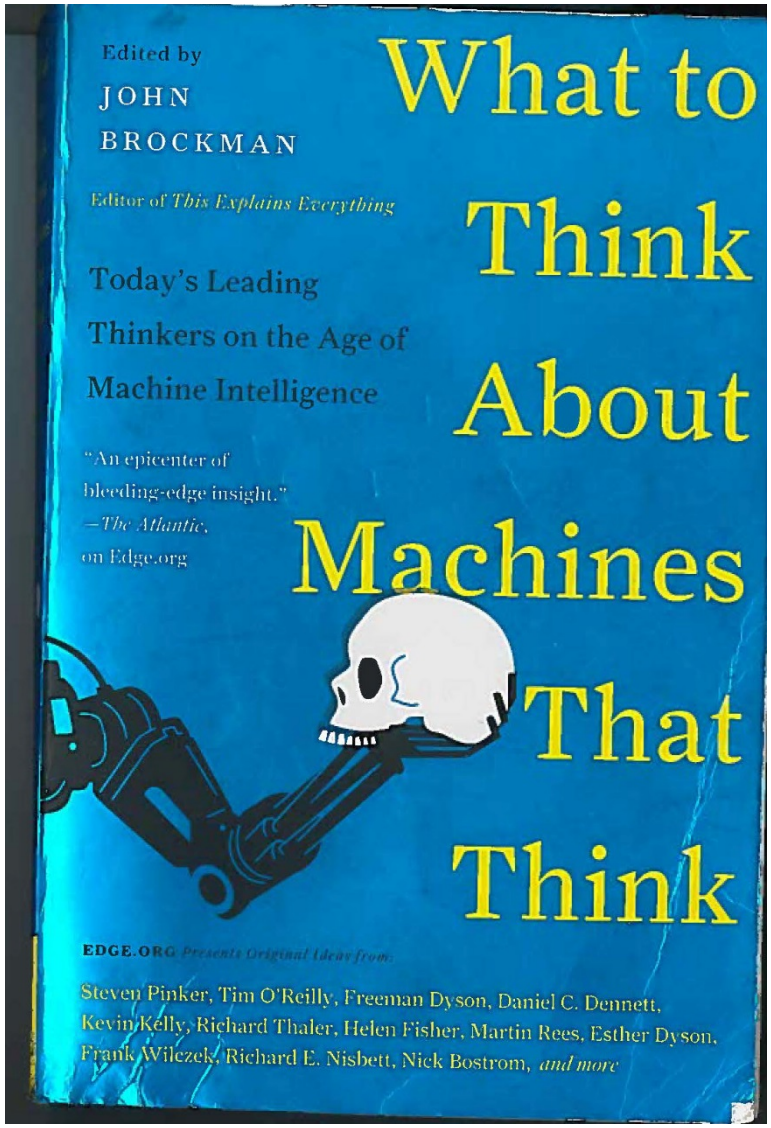
Prof. Neil C. Rowe

Computer Science Department
Naval Postgraduate School

http://faculty.nps.edu/ncrowe

Fall 2019

A book with some wild ideas (2015)



This contains 186 blog entries (limited to 1000 words) from edge.org, on the question "What do you think about machines that think?" Only a few (10%) of authors had training in AI, so many blog entries were uninformed.

Extreme ideas about AI

- Ideas from over-optimists:
  - AI is going to revolutionize the world.
  - AI will cause eventually reach a “singularity” where human society becomes fundamentally different.
- Ideas from over-pessimists:
  - AI will take over the world and reduce humans to servants.
  - AI has fundamental limits because some aspects of thinking are uniquely human.

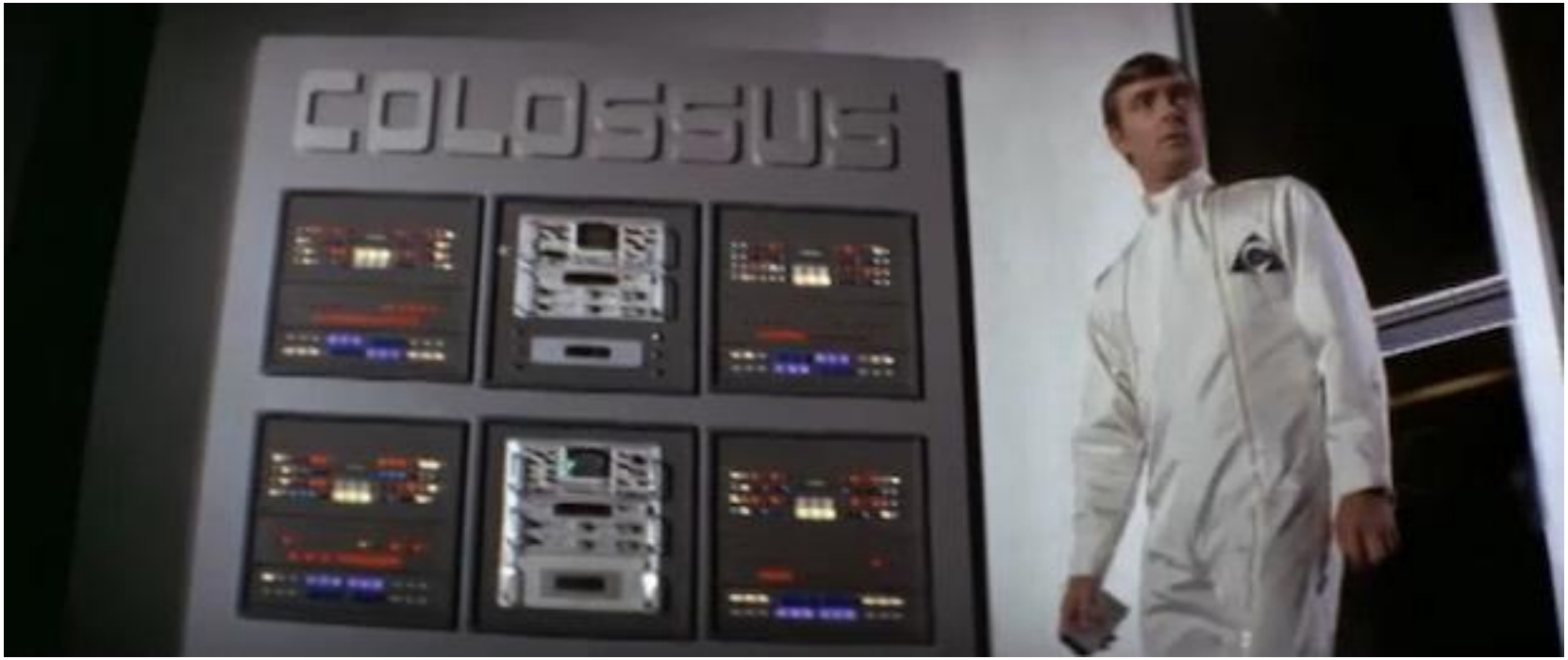
Over-optimism about AI progress

- AI is another automation technology. Most automations have increased effectiveness of technology but not made major changes to society:
  - Portable telephones
  - Aircraft
  - Refrigeration
- Most technological innovations are bad, but bad things tend to disappear. So most AI is bad.
- Having many AI entities on the planet isn't much different than having many human organizations. Organizations often have their own objectives different from those of their human members, and we have rules and laws to control them.

The AI "singularity"

- As AI steadily gets more powerful, it won't be long before artificial intelligence overtakes human intelligence in most areas.
- Some fear that the pace of technological change could increase greatly with machine learning.
- Will this make a big difference? Not necessarily:
  - When engine-powered vehicles overtook humans in speed, travel became easier. But walking is still important exercise.
  - AI intelligence tends to focus on tasks humans cannot do well (like math) because that's where the biggest payoffs are. So AI and humans compete in different arenas.

Evil artificial intelligences in the movies (1)



Evil artificial intelligences in the movies (2)

AI takeover

- Some people are worry about the power of AI technology.
- AI is designed to create obedient servants. So we can train it to be ethical if we like. That's different than nuclear weapons which don't check what is present where they explode.
- Unethical AI can be turned off, unlike some work in recombinant DNA gene modifications.
- So AI systems aren't likely to take over the world.
- Still, any powerful technology raises ethical issues. More about ethics in a subsequent lecture.

Limits of AI (1)

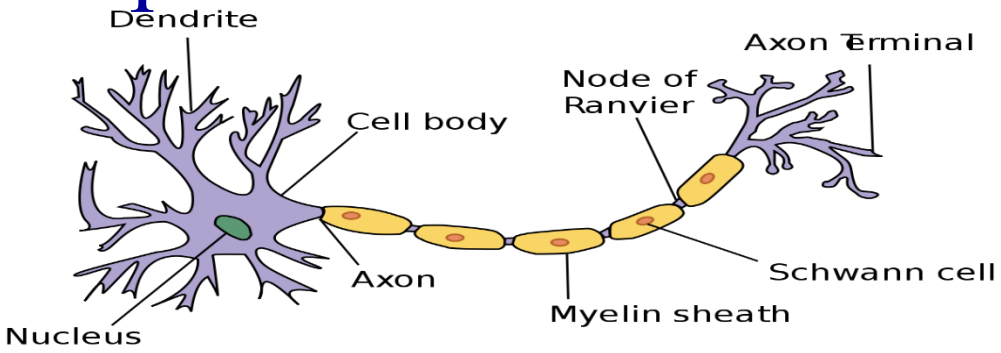
- There's a long history of claims about what AI cannot do. For instance:
  - Beat humans at chess
  - Carry on conversations in English
  - Play the stock market
  - Learn context for decisions (done with neural networks)
  - Communicate with other intelligences (done with speech acts)
  - Have emotions
  - Think about itself (done by meta-reasoning)
- One by one, such tasks have been accomplished by AI. So be skeptical of claims of limits to AI.

Animals can behave very like humans

Limits of AI (2)

- Similar "species-ist" arguments have been made why animals can't think (to justify their abuse). But animal brains are very similar to human brains.
- Theorems in computer science give computational limits, e.g. whether a computer program will stop eventually. But these limits so rarely impede real problems that they can usually be ignored.
- Some argue that AI only simulates intelligence. However, if it walks like a duck and quacks like a duck, isn't it a duck? If it can pass every intelligence test you can think of, it must be intelligent. Computer science shows hardware is not critical to capabilities.

Counter to pessimism: Brains are machines



- Brains are made of neurons.
- Neurons compute a weighted sum of the excitation of their input dendrites, encode as a frequency of firing, transmit it along an axon, and excite other neurons across synaptic membranes.
- About 30 neurotransmitters act as global variables affecting transmission of particular neurons.
- That's all there is – it's very mechanical.
- It should be possible to completely duplicate brain machines by engineered machines.