

Academy for Lifelong Learning

# **Artificial Intelligence**

## Great Promise, Considerable Threat

Bob Fabian  
30 September 2020

# Presentation Plan

- Definitions: AI and AGI
- Before Expert Systems
- Before Neural Networks
- The Neural Network
  - And powerful computers
  - And back propagation
  - And Big Data
- Delivering impressive results
- *Where should we go from here?*

# Artificial Intelligence

- AI: The theory and development of (computer) systems able to perform tasks that normally require human intelligence, such as chess, visual perception, speech recognition, decision-making, and translation between languages.

Definition from Oxford Languages

- AGI: Artificial general intelligence is the hypothetical intelligence of a machine that has the capacity to understand or learn any intellectual task that a human being can, [Eventually going beyond what would be possible for any human?]

Basic definition from Wikipedia

# Enhance Human Capabilities

- **Mechanical amplification**

- Our physical machines amplify our ability to accomplish physical tasks – we're stronger, faster, more consistent
  - With the right physical technology

- **Conceptual amplification**

- Abstractions allow us to accomplish more conceptually. Logic improves our reasoning. Math improves our calculations.
  - Reasonable to view the mind as evolved

- **AI: Amplify all Human Capabilities**

- Will humans then be free to explore artistic expression and rich human interaction?
  - Will humans be free, ... or just irrelevance?

# Threats

- Mechanical
  - Trained labourer plus machine can do the work of 10 or 100 or ...
    - What are the other 9 or 99 going to contribute?
- Computational
  - Office worker plus computer can do the work of 10 or 100 or ...
    - Draftsmen, law clerks, stenographers – examples of displacement
- Artificial Intelligence
  - Worker with AI able to do more and do it faster and more accurately ...
    - Workers without AI will just not be able to compete. Disturbing vision.

# Open Question

- We can recognize when a goal needs to be scrapped and replaced by a more appropriate goal
- Original Goal
  - Too costly
  - Now inappropriate
  - Unacceptable impact
- Should an Artificial General Intelligence be allowed to replace its given goal with something it concludes would be “better”?
  - Makes me very uncomfortable

# Very Brief History

- pre-1960: Simple routine tasks, but cost-effective (bright promise)
- 1960s: AI programs seen as “the future”, strong research funding
- 1975: Start of first AI Winter
  - Could only solve “toy” problems. Funding drought.
- 1980s: Expert (rule based) systems coped with increasing complexity
- 1990: Second AI Winter
  - The solved problems were too far from the real world
- 2000: Neural networks started to learn from experience
  - Experience and not rules used to guide steps
- NOW: Heavy investments. Active use of neural nets.

# Before Expert Systems

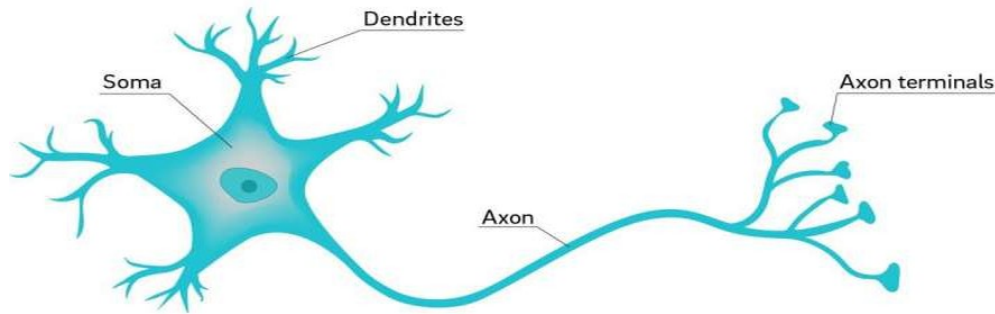
- Much work on the nature of mathematics and proofs. A mathematical system may not be provably consistent, but proof generation would be possible.
  - Early systems were able to generate impressive results.
- Much work on the deep structure of language. To really understand or translate the goal was to uncover the deep structure of the utterance.
  - Some promise, but systems never lived up to that promise.



# Before Neural Networks

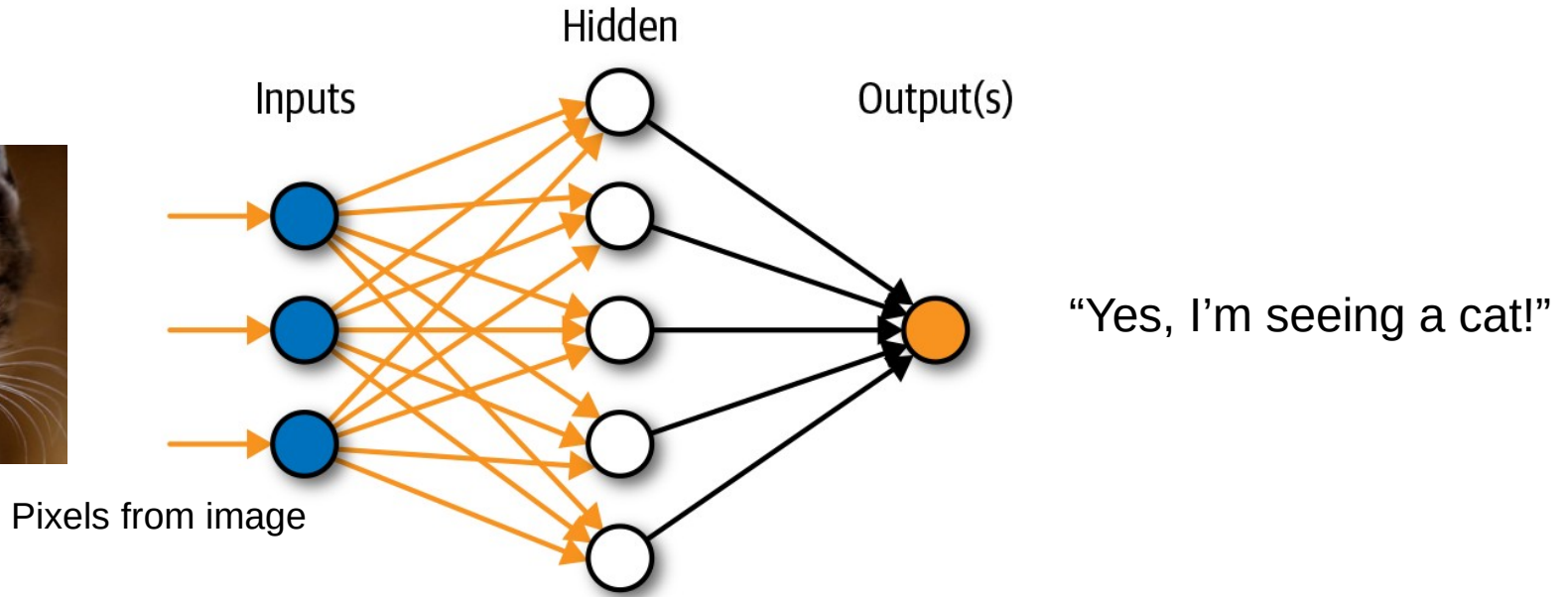
- Classic Story
  - The recently retired senior engineer for a production system was called back to get the line moving. He demanded \$10,000 to do the job. He picked up a hammer and walked down the line to the third dark green machine. He whacked the cover at the base of the machine. The line started up.
  - “Well I never ... We paid \$10,000 for one hammer blow!”
  - “No, that’s for knowing *where* to whack the machine.”

# Human Neural Networks



- Neuron is excited by receiving sufficient charge from its Dendrites. When excited, it sends a signal down its Axon which is transmitted to other Neurons by its Axon terminals. [A frog has about 16,000,000 such neurons in its brain.]

# Artificial Neural Network



O.K. Strengthen the connections saying “cat”

# Artificial Neural Network II

- Demonstrated Successes
  - Real-time speech recognition
  - Real-time facial recognition
  - Real-time language translation
- Network will have millions of neurons
- Learning will extend over millions of examples
  - Strengthen paths to correct answer
  - Weaken paths to incorrect answer

# Artificial Neural Networks III

- Almost (?) impossible to explain why an AI neural net system reaches its conclusion
- Mirrors part of what seems to happen in the human brain, but not everything
  - No layering of response – it's all done in one layer
  - No evolution of species's response to its environment
  - Little role for patterns or metaphors
- Benefits large enough there's unlikely to be 3<sup>rd</sup> AI Winter
  - But may need to add one or more somethings to the mix

# Artificial Neural Networks IV

- The computational and storage challenges are massive
  - How to synchronize thousands of parallel processes
  - How to guesstimate most promising parameter values
  - How to apply fuzzy logic on such a massive scale
- How to add in the insights from expert systems
  - Is it a good idea to work to combine approaches
- How to factor in security and privacy concerns

# Marriage of Big Data & AI

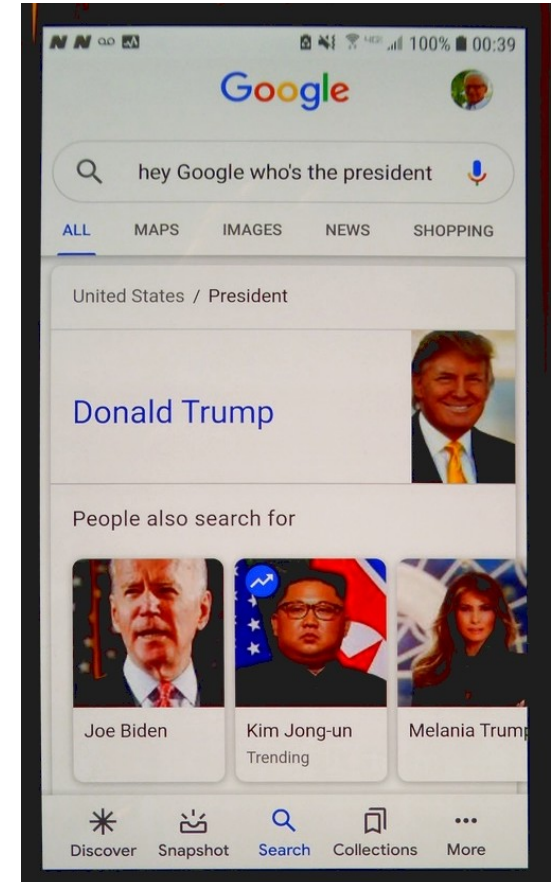
- Parallel business effort to collect and analyze massive amounts of data
  - Develop insights about:
    - Customers
    - Voters
    - Markets
  - It goes deeper than what pundits can deliver
- Obvious: Train AI to extract more valuable insights

# Impressive Results



Google recognized my spoken words, and instantly came up with the correct answer.

**Practically impossible just a very few years ago!**





# Intelligent Machines

- Turing test: Will a knowledgeable observer be able to determine if the respondent is human or machine?
  - What if a system uses tricks to fool the observer?
  - What if the system is just too good to be human?
- Can the system be taught how to reach any conclusions that humans have been able to reach?
- Can the system learn, by itself, how to reach any conclusions that humans have been able to reach?

# Slaughterbots



[https://www.youtube.com/watch?v=HipTO\\_7mUOw](https://www.youtube.com/watch?v=HipTO_7mUOw)

# If AI Really Delivers

- Should AI be used to power autonomous killing machines?
- Assume AI is able, (in a limited sphere), to accomplish much more than a human could ...
  - How will AI's goals be open to challenge by others?
  - Should AI be allowed to proceed, regardless of impact?
  - Should AI be allowed to directly compete with humans?
  - Should AI be allowed to alter goals humans set for it?