

# Artificial Intelligence Update

Neal McBurnett

Boulder Freedom Forum

2023-04-27

# Neal McBurnett

- Consultant on Big Data, Election Integrity, Voting Methods
- Software developer, Ex-Bell Labs
- Student of AI since 1978
- Always convinced that we would invent computers smarter than we are
- Got to learn about/teach undergrad course in AI at CU, 2012 and 2013
- Speaking for myself

Subject to change

Just beginning to organize my thoughts

# Introduction

- Terms
- Human humility
- Singularity
- Early advances in games, speech, vision
- ChatGPT (and Bard, etc)
- Demo?
- Ethics
- Adapting

Questions any time

# Terms: Intelligence, Consciousness, Reasoning,

...

Peter Norvig, 2015: **WHAT DO YOU THINK ABOUT MACHINES THAT THINK?**

In 1950, Alan Turing suggested we should ask not "Can Machines Think" but rather "What Can Machines Do?" Edsger Dijkstra got it right in 1984 when he said the question of Can Machines Think "is about as relevant as the question of whether Submarines Can Swim." By that he meant that both are questions in sociolinguistics: how do we choose to use words such as "think"? In English, submarines do not swim, but in Russian, they do. This is irrelevant to the capabilities of submarines. So let's explore what it is that machines can do, and whether we should fear their capabilities.

"Intelligence measures an agent's ability to achieve goals in a wide range of environments."

Don't sweat tricky words: "conscious" "reasoning" It's all

# Hype, "AI Winter"

Yes, we have been hearing hype over AI for decades, as with all technologies

Again, track the actual capabilities of AI systems

# Humility

Intelligence is yet one more area where humans have to gain a little humility. over the years. We have accumulated evidence that self-centeredness is not based in reality.

We're not the center of the universe, not necessarily the most advanced or successful species on the planet, we may not even be the most intelligent species on the planet, and Our abilities in many realms are less unique than we often tell ourselves.

# Kurzweil - the Singularity

Pioneer futurist

**Games: Chess, StarCraft, Go, AlphaGo, Hanabi**

# Speech recognition, Image recognition, Image generation

Amazing, specialized tools for experts and general public

# OpenAI's ChatGPT: game-changer

Amazing combination of:

- Deep neural network trained on statistics of big data to guess the next word
  - GPT-4 model may already know everything it needs to know about the underlying reality!
- Reinforcement Learning with Human Feedback To teach it how to respond
  - Seems to be reasoning, despite claims that statistics can't do that

Disprove Chomsky's theories

NLP researchers astonished at results

Ordinary humans can leverage AI now

# Demo OpenAI's ChatGPT, Google's Bard

example writing code example political theory: Defining the Political Center: How would Robert Bork define the center which he says does not hold in modern politics

# Assistants vs Overlords

Change is always scary

Goal: leverage AI as human assistants

Human augmentation - e.g. glasses, bicycles, calculators

Note superintelligence doesn't mean solving all problems optimally

Lots of real-world challenges

- Provably unsolvable problems - "NP-hard"
- Incomplete information
- Conflicting goals, values
- Ethical constraints

# Adapting to the singularity, superintelligence

- Security
- Balance of power in the world
- The Economy
- Jobs
- Copyright, licensing
- Ethics
- Fairness, Bias
- Human dignity

# Explanations of results

AXAI, or Algorithmic eXplanatory Artificial Intelligence

Personally I'm guessing AI might be too good at being able to "explain" it's results to people in believable ways

Humans are easy to hack....

Transparency is still very important - check results via other AIs

# Ethics issues

Who controls the best agents? wield enormous power

The fabulous endeavor: Make AIs robustly beneficial book etc.  
todo projects \*Le Nguyen Hoang EPFI

Note push back on developing large language models. TODO  
Note that it's beginning to look like it doesn't actually require  
huge resources to develop. Highly capable AI.

natural concerns over bias, but contrast with bias in current  
human decision-making processes

Implications for politics, governance, social contract:  
[https://en.wikipedia.org/wiki/Universal\\_basic\\_income](https://en.wikipedia.org/wiki/Universal_basic_income)

# Questions?

Emphasize uncertainty

Caution

Inevitability

Opportunities for better problem solving

## Resources

Talks by open AI folks TODO

Local futurist: Wayne Radinsky, Boulder Future Salon

These slides: <https://bcn.boulder.co.us/~neal/talks/ai/>