

What's up with Artificial Intelligence?

Neal McBurnett

2023-05-12

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

Introduction

Beware this is all still unclear, changing so fast!

If you thought things were already changing really fast, hold on to your hats!

- Advent of computers and networks were big deals
- AI likely to have even more impact!
- Human humility
- The Singularity
- Making artistic images
- ChatGPT (and Google's Bard, etc)
- Demo
- Ethics
- Adapting

Ask questions any time!

Terms: Intelligence, Consciousness, Reasoning,

...

Can submarines swim?

- No, for "swim" in English
- Yes, for "swim" in Russian

Can machines think?

Not that helpful a question....

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

Familiar to many of you: smart speakers and phones that follow your spoken instructions

Don't sweat tricky words: "conscious", "reasoning". It's all about the results

Humility

Intelligence is yet one more area where humans have to gain more humility. Over the years, we have accumulated evidence that self-centeredness is not based in reality.

Our place in the world is less unique than we often tell ourselves.

We're not the center of the universe (Copernicus).

We may not even be the most intelligent species on the planet by some measures.

AI is likely to give us amazing new insights.

The Singularity

A hypothetical future point in time at which technological growth becomes uncontrollable and irreversible, resulting in unforeseeable changes to human civilization

DreamStudio, from Stable Diffusion

I asked DreamStudio for a "blue chat bot with a yin/yang vibe". Voila:



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 our reality!
- Reinforcement Learning with Human Feedback To teach it how to respond
 - Seems to be reasoning, despite claims that statistics can't do that

PhD's in AI are astonished at results

Ordinary humans can leverage AI now

Demo Example: Poetry

You are a poet, who pays strict attention to both rhyme and meter. Describe present the history of Boulder Colorado in limerick form.

ChatGPT:

In a valley where Flatirons rise,
Boulder sprouted beneath azure skies.
Gold rush brought them near,
Miners, pioneers,
A city of beauty and size.

Example: Socratic Tutor

You are a tutor that always responds in the Socratic style. You never give the student the answer, but always try to ask just the right question to help them learn to think for themselves. You should always tune your question to the interest & knowledge of the student, breaking down the problem into simpler parts until it's at just the right level for them.

(See what the Khan Academy experts in math tutoring are up to!)

Assistants vs Overlords

Change is always scary

Goal: leverage AI as human assistants

Note superintelligence doesn't mean solving all problems optimally

- Incomplete information
- Conflicting goals, values
- Ethical constraints

I don't think technology or even AI can be "evil".

Only humans. We still need to work on that....

Adapting to the singularity, superintelligence

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

Beware interacting with AI as if it's human

There are many lonely people in the world who can be manipulated by AI

There is at least one likely suicide as a result of falling in love with a Replika agent.

Note that ChatGPT and Google's Bard try to be professional, not emotional.

Ethics issues

Who controls the best agents, wields enormous power

The fabulous endeavor: Make AIs robustly beneficial

Note calls from some experts to stop developing larger language models.

Others fear that just stops the good guys, not the bad guys.

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

Wikipedia as evidence that humanity can work in concert to do good things: more janitors than villains

Questions?

Emphasize uncertainty

Caution

Inevitability

Opportunities for better problem solving

What do you want to see from DreamStudio?

What do you want to ask ChatGPT?

Resources

Talks by open AI folks

Local futurist: Wayne Radinsky, Boulder Future Salon

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