The Inflation of Artificial Intelligence (AI)

A presentation by
Theo Pavlidis

Jefferson's Ferry, August 24, 2023

1

1

Artificial Intelligence (abbreviated to AI) refers to methods that replicate human **cognitive** processes: visual and auditory perceptions, reasoning, speaking, etc.
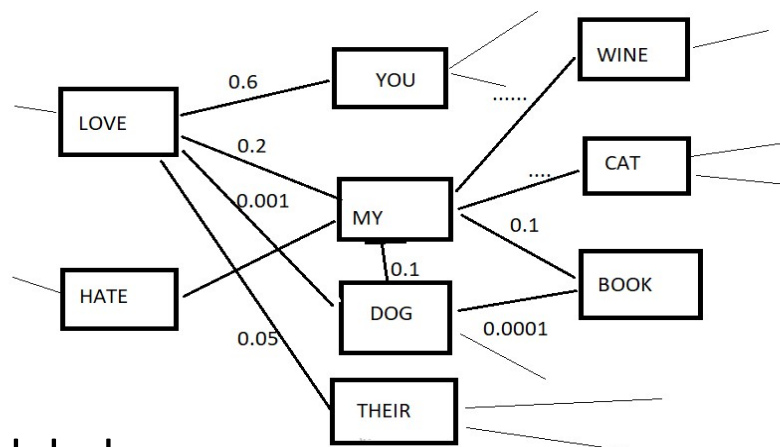
2

2

Generative AI refers to methods that replicate human **creative** processes, such as generating text or images.

For text, **Generative AI** is a **Large Language Model (LLM)** that can generate probabilities of a series of words.

3

3

Part of an LLM


```
                                    ┌──────┐
                             0.6    │ YOU  │        ┌──────┐
┌──────┐                     ───────└──────┘ ......  │ WINE │
│ LOVE │          0.2                                └──────┘
└──────┘       ──────                   ┌──────┐
              0.001      ┌──────┐   ....  │ CAT  │
                         │  MY  │        └──────┘
                         └──────┘  0.1
┌──────┐          0.1       │         ┌──────┐
│ HATE │                    │         │ BOOK │
└──────┘       ┌──────┐     └──────┐  └──────┘
          0.05 │ DOG  │   0.0001
               └──────┘
               ┌───────┐
               │ THEIR │
               └───────┘
```

The branch labels are probabilities between each pair of words.

4

4

2

We will use the term Chatbot 🤖 🔲
for applications that rely on an LLM to
produce text in response to a user request
such as "write a love letter," "who is G. R.
Emlin?," etc.

5

5

Similar methods can be used to create pictures.

| playing soccer | playing football | playing rugby (two answers) |

Output of DALL-E in response to requests to draw: "A bunch of monkeys ..."

Note the many errors.
We will skip picture creation for now.

6

6

An LLM may contain billions of parameters that are determined during **training** over large volumes of text, a process that requires enormous computer resources that became available only in the last few years.

7

7

Thus, the recent advances are due to improvements in overall computer technology rather than to breakthroughs in Generative 𝔸𝕀 methods.

8

8

Important (and often forgotten) Fact

An LLM has no information about facts.
It only knows which words come together, so it will produce a linguistically correct output that may or may not be factual.

9

9

Unfortunately, proponents of 𝔸𝕀 make claims (based on flawed evidence) that an LLM produces factually correct output.

10

10

Example of part of a training set

How many months are in a year? …………………. Twelve
How many inches are in a foot? ……………………Twelve
How many gods were in mount Olympus? …….Twelve
How many congressional districts are in
Georgia? .... Fourteen
How many congressional districts are in New
Jersey? …………………………………………………………. Twelve
How many congressional districts are in Ohio? Fifteen

11

11

How would 🤖 respond to the question
"How many feet are in a yard?"

If the only instances of "How many" are those
shown, it will probably say "Twelve" or, possibly,
"I do not know."

If there are many more instances of "How
many" it might give the right answer.

12

12

The training set must not only contain the "right answer" but must have it in a context that can be reached by 🤖 .

13

13

The company *OpenAI* reported that one of its 🤖 scored highly on standardized exams.

It turns out that this happened only when the training set contained material like that of the exam. *Cheater!*

14

14

Sources for more on the subject
- Melanie Mitchell, **How do we know how smart AI systems are?** *Science,* July 13, 2023 issue*.*
- S. Kapoor and A. Narayanan, *AI Snake Oil* (//aisnakeoil.substack.com).

15

15

We look next into another paper that focuses on misconceptions caused by the use of metaphors to describe AI systems or their actions.

16

16

The June 22, 2023 issue of the *Economist* has an article titled

Talking about AI in human terms is natural—but wrong

Subtitled **When it comes to artificial intelligence, metaphors are often misleading**

17

17

The article starts with the following observation.

The ability of 𝔸𝕀 systems to reproduce what seems like the most human function of all, namely language, has ever more observers writing about them.

18

18

When they do, they are tempted by an obvious (but obviously wrong) metaphor, which portrays 𝔸𝕀 programs as conscious and even intentional agents. After all, the only other creatures which can use language are other conscious agents—that is, humans.

19

19

Take the well-known problem of factual mistakes in biographies, the likes of which a 🤖 churns out in seconds. Incorrect birthplaces, books never written; one journalist at *The Economist* was alarmed to learn that he had recently died. In the jargon of 𝔸𝕀 engineers, these are "hallucinations". In the parlance of critics, they are "lies".

20

20

"Hallucinations" might be thought of as a forgiving euphemism. For the "lies" crowd, though, the humanizing metaphor is even more profound: the AI is not only thinking, but it also has desires and intentions.

21

21

The all-too-human desire to yell at a malfunctioning device comes from an ingrained instinct to see intentionality everywhere. It is an instinct, however, that should be overridden when writing about 𝔸𝕀.
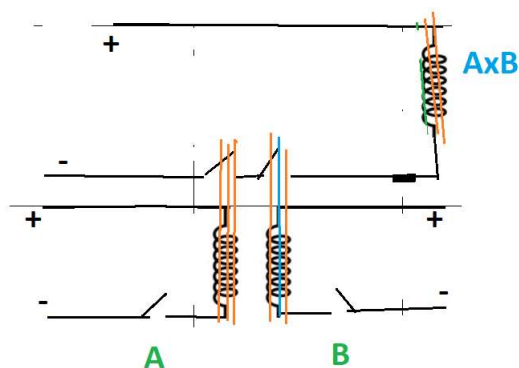
22

22

We will take a quick look at the "insides" of a computer to see why assigning intentions and desires to 𝔸𝕀 is absurd.

23

23

Part of the circuitry of a WW-II computer

AxB

A   B

The key feature is that closing a switch in one circuit causes a switch in another circuit to open or close

24

24

12

Very soon electromechanical relays were replaced by vacuum tubes (triodes), then by transistors (around 1960), then by integrated circuits (mid 1970s). But their logical function remained the same.

25

25

The basic hardware of a computer is very simple, and all the complexity is in the program. The program instructions are converted into 0's and 1's before getting into the hardware.

26

26

• The 0's and 1's of the result are eventually converted into words, giving a frightening— but false—sense that there is a ghost in the machine.

27

27

Extravagant claims seem to be a tradition in AI and, invariably, do not materialize.

Mitchell quotes a prediction made by Marvin Minsky (one of the founders of AI) in 1967:

28

28

"Within a generation the problem
of creating 'artificial intelligence'
will be substantially solved."

Since a generation is about 30 years,
it is clear that the prediction has not
materialized.

29

29

Why such wild statements by people
who might know better?

By claiming that AI can be dangerous,
they invite government regulation
and discourage competition.

30

30

Of course, evil agents can always use technology to cause harm.

31

31