From Perceptron to AI: The Rise of Large Language Models

Lecture 3 of 4:
What are the impacts of AI?

Dr. Michael Stobb
Thursday Forum
Coe College

February 2026

Last Time…

- Discussed “How do LLMs work?”
  - Looked at Multilayer Perceptrons
  - Examined learning algorithm *backpropagation*
  - Briefly discussed the Transformer model
  - Finally introduced Language Models

- Discussed LLM adjacent topics:
  - Three phases of training LLMs
  - Metrics for measuring AI performance

Today's Goals

- Impacts and outcomes of AI
  - Less technical (no assumed prior knowledge)
  - Mostly stick with *facts* relevant about today
  - Occasional opinions/thoughts about future
  - Lots of different topics, will pause for questions

- **Disclaimer**: Many topics are highly polarized!
  - My goal is *informational*, not *conversion*
  - Will attempt to offer "both sides"
  - References and notes will have more info

AI Hallucinations

Common Definition: AI hallucinations are when an AI tool either lies or fabricates information.

- This definition wrongly anthropomorphizes LLMs
  - The model did not “lie”
  - The model did not lose touch with reality

Better Definition: AI hallucinations are sequences of words that are *statistically plausible* but factually incorrect.

AI Hallucinations

- LLMs are not accessing a database of facts
  - They are *prediction engines* not *truth machines*

- Now AI relies on *Retrieval Augmented Generation* (RAG)
  - These hook a database up to an AI model
  - The AI model in turn asks the database for facts

- Hallucination rates have dropped to about 1%
  - Mostly because of RAG and just larger models
  - But hallucinations will probably *never* go away
  - Being “creative” requires risk of error

Sycophancy in AI

Definition: AI Sycophancy refers to the tendency of AI systems to agree with, flatter, or mirror a user's views to gain approval, even if it means compromising accuracy or honesty.

- Different than hallucination, but similar results
  - The AI models are trained in separate phases:
    - Supervised Fine Tuning: accurate and obedient
    - Reinforcement Learning: kind and good mannered
  - These can pull the model in different directions
  - Accuracy of the model is compromised for agreeableness

Sycophancy in AI

- Arguably more *dangerous* than hallucinations:
  - Can cause or exacerbate mental health issues
  - Can cause emotional dependency on the AI model
  - Can increase rates of self-harm and substance abuse

- Problems disproportionately impact the most vulnerable
  - People with existing mental health issues
  - People who are already isolated and lonely

- Solutions currently aimed at alternative training

Synthetic Media Generation

Definition: *Synthetic Media* refers to any media (audio, text, or video) created by an AI system

- Synthetic media is a broad and *neutral* term
  - Can be used for art, fun, entertainment, etc.
  - Can also be used in AI scams

Definition: An *AI Deepfake* is any synthetic media created by AI systems to intentionally mislead, misinform, or deceive someone.

- The difference really lies in the ***intent*** of the media

Synthetic Media Generation

- Synthetic media generation has gotten ***Shockingly GOOD***
  - Audio, Image, and Video
  - Music too

- Writing as well
  - Author Mark Lawrence conducted an informal test
  - Professional authors vs. AI written short fiction
  - Between 1000-3000 total votes
  - Highest rated were AI written, and higher on avg

Synthetic Media Generation

- Positives
  - Democratizes media generation
  - Anyone with a good idea can make a movie/comic/etc.
  - Creativity can be unlocked for non-technical people
- Negatives
  - Human artists cannot compete on price
  - Risk drowning in “slop” that lacks soul
- Multiple lawsuits ongoing, but so far:

AI Generated Media cannot be copyrighted

AI Scams

- Huge risk from AI Deepfakes
  - In early 2024, a finance worker in Hong Kong was tricked into transferring over $25 million
    - The scam involved a multi-person video conference call
    - Everyone on the call, except the victim, was a deepfake recreation of senior officers from the company
- Voices can be cloned with less than 5 seconds of audio
  - AI voices are nearly indistinguishable from humans
  - Recommendation: Set up a code-phrase to use with family

AI Scams

Liar's Dividend: Where unscrupulous actors can cast doubt on genuine, damaging information by claiming it is fabricated or manipulated

- There are some technological solutions, but not great ones
  - Digital Watermarking and Provenance tracking
- Likely need to shift society from one of **belief** to **skepticism**
  - If an image/video/audio sounds suspect, it probably is
  - Think before sharing, check source, look for corroboration

AI & Education

- It is difficult to understate how large an impact AI has had
  - Curriculum and assignment overhauls
  - Fundamental changes in what even gets assessed
  - Some entirely new school organization paradigms

- There is a wide spectrum of feelings on the issue
  - Some believe AI is the death of critical thinking
  - Others that its the long needed salvation for education

- Like most complex issues: truth is likely in the middle

AI & Education

Technological deskilling is the decline in the skills required for a particular job due to the introduction of machinery or technology.

- *Deskilling* is not new:
  - Socrates was against writing itself (over memorization)
  - Assembly lines and automation have shrunk hand crafting
  - Navigation with GPS has lead to poor map reading skills

- New tech adds convenience, but does not have to deskill
  - We don’t exercise at a gym for fun, but for strength
  - Skills we wish to keep need a “gym” to be trained

AI & Education

- More relevant examples:
  - **Penmanship** skills have declined in favor of typing
    - Losing art of handwriting, lower memory retention
    - But increased communication speed, greater volume
  - **Calculators** have reduced mental arithmetic abilities
    - Only barest essentials are now taught
    - Allowed people to advance to higher mathematics
- AI could lead to a deskilling of our ***cognitive abilities***

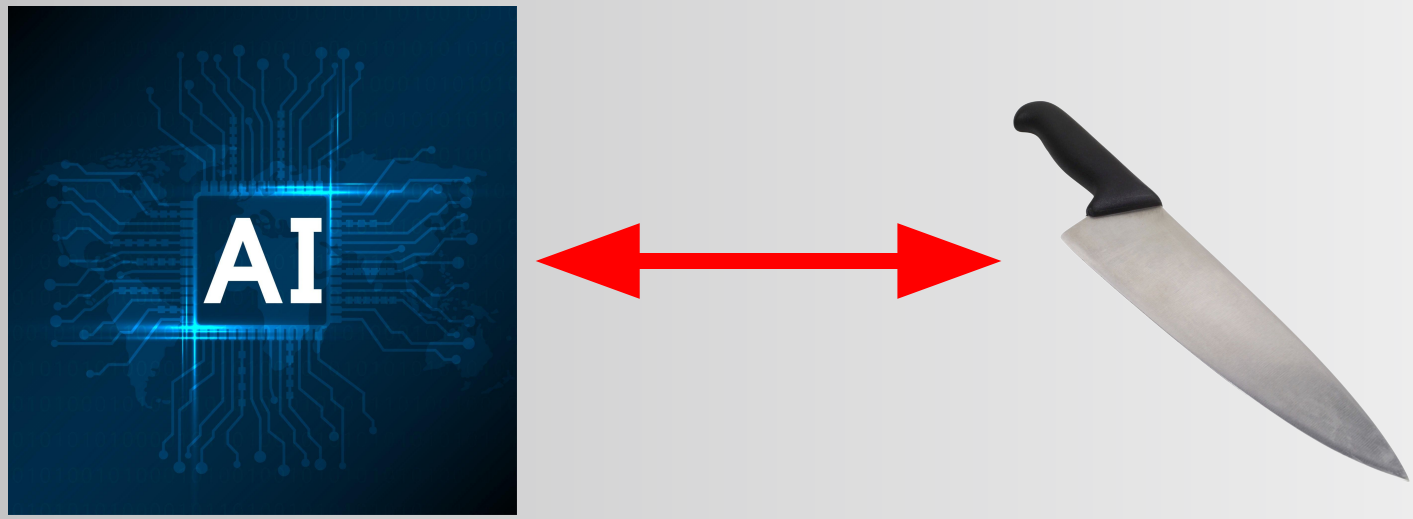
AI & Education

- The Potential good
  - AI could be a perfectly attentive, hyper intelligent super-tutor
  - Has the intellectual capacity similar to many experts
  - Can be easily deployed nearly anywhere

- The current bad
  - Used primarily by students to bypass the “hard work” of education
  - Could be the largest *deskilling* technology ever

AI & Education

- My personal opinion:

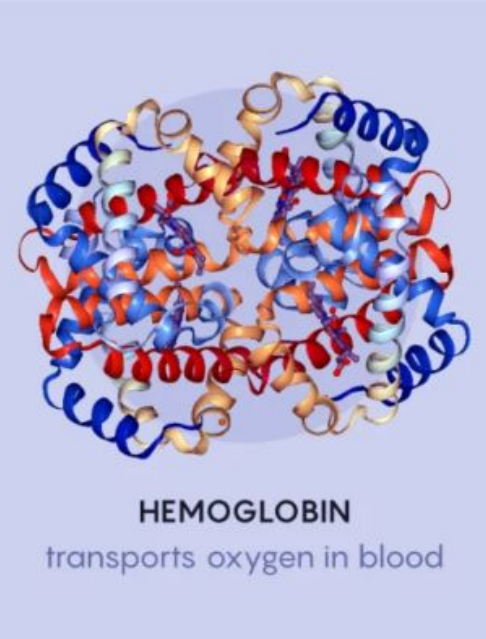


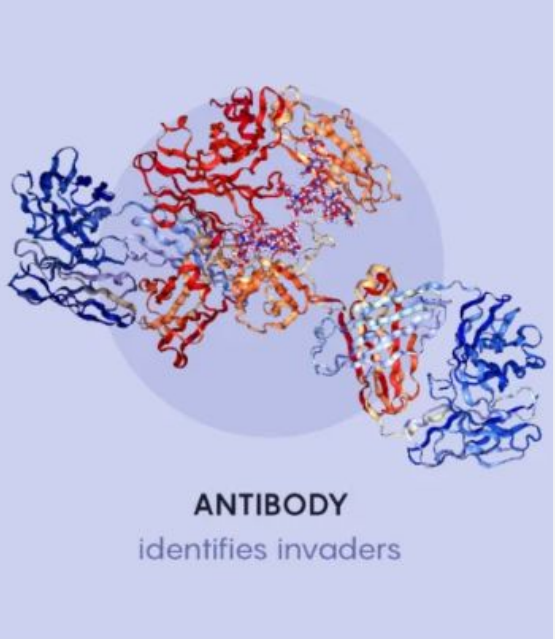
Both are great tools,

but I wouldn’t give them to a baby.

AI & Biology

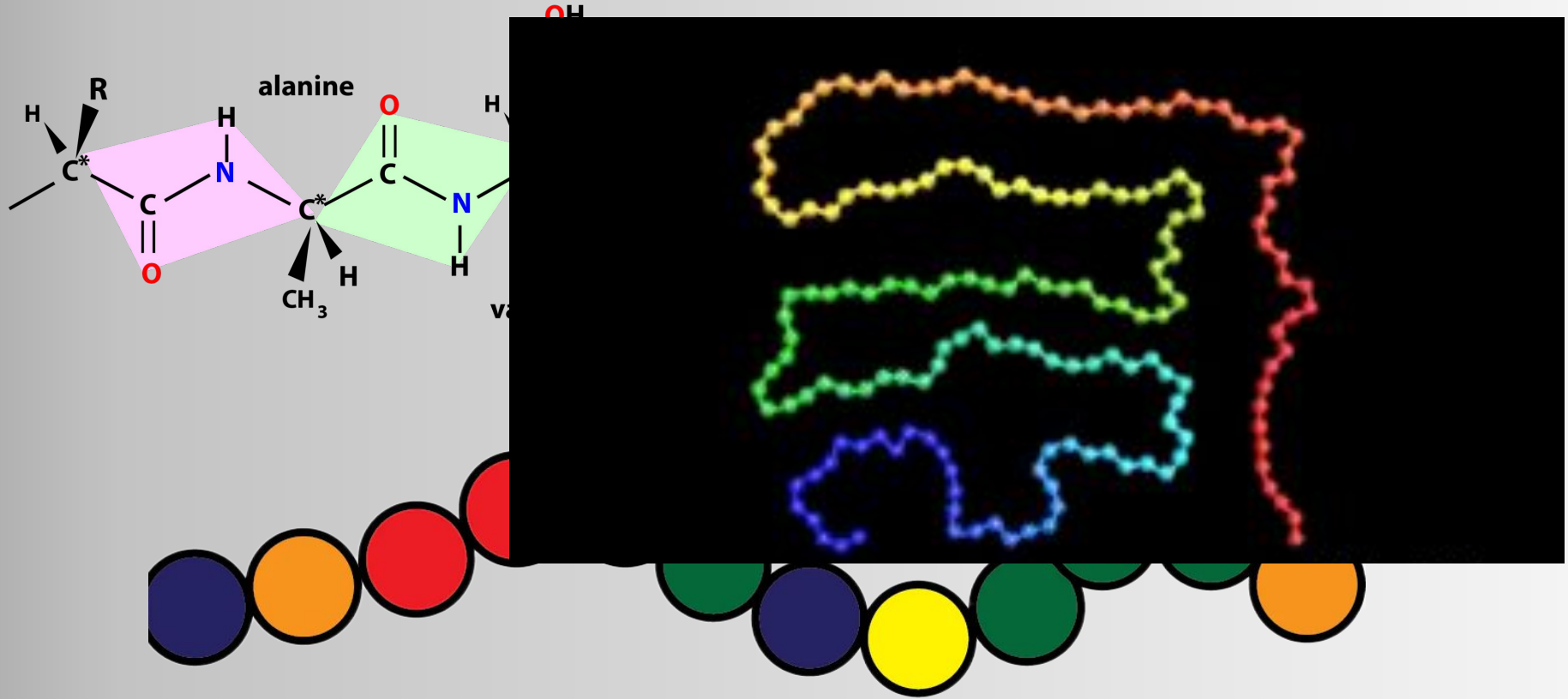
- Proteins are complex molecules upon which all life is built
  - The shape of the protein dictates its function





AI & Biology

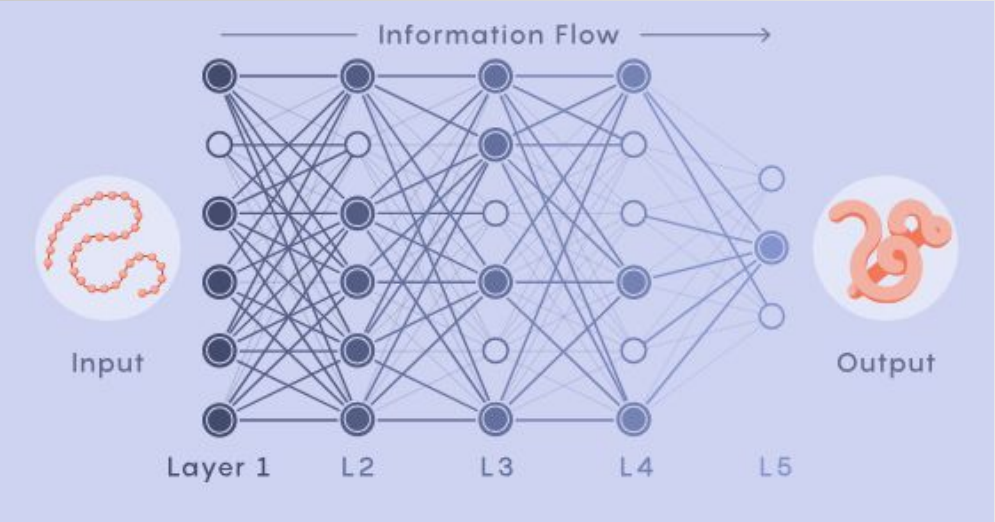
- Proteins are fundamentally **chains of amino acids**



AI & Biology

- Given a chain of Amino acids, its **hard** to determine how the protein would fold
  - Could take *years* by experts to determine a single protein

- In 2021 AlphaFold2 was released publicly



AI & Biology

- AlphaFold2 achieves prediction scores similar to experimental methods
  - In one fell swoop, we know how *most* proteins fold
  - Over 200 million proteins in less than 1 year

- The main “creators” of AlphaFold won the Nobel Prize for Chemistry in 2024
  - Demis Hassabis and John Jumper of Google DeepMind
  - The same AI lab that produces the Gemini AI model

AI & Employment

- Right now, AI is only expected to have a mild impact on employment
  - Overall AI adoption remains fairly low
  - Likely because of several reasons:
    - New technology is always slow to diffuse
    - The cost to switch or change is usually high
    - Material culture always evolves faster than non-material culture

- Goldman Sachs predicts a 0.5% increase to unemployment

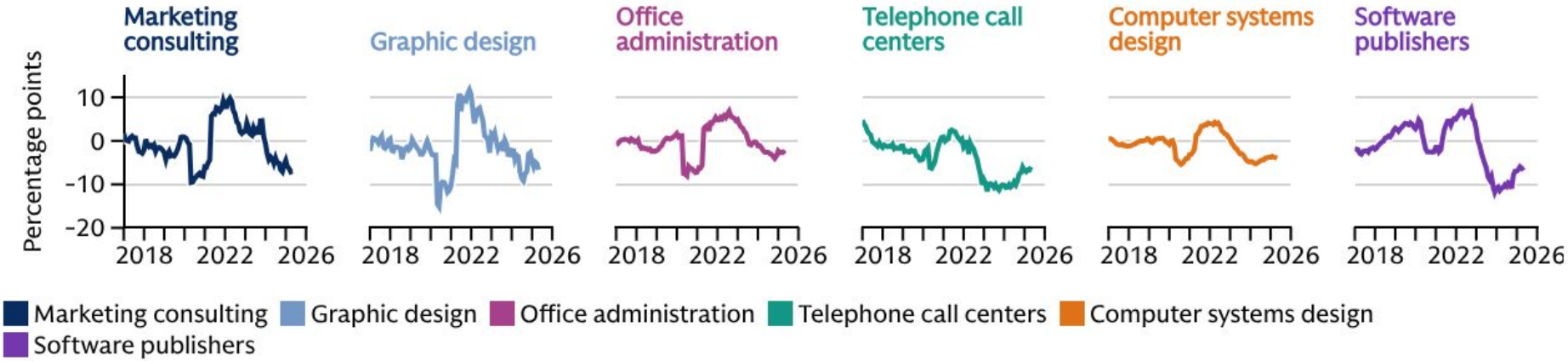
AI adoption has accelerated among larger companies, but adoption remains low

AI adoption rate*, by firm size

14%

Employment is contracting in some sectors where anecdotes suggest AI is substituting for labor

Year-over-year payrolls growth deviation from 2015-2019 trend



Source: US Bureau of Labor Statistics, Haver Analytics, Goldman Sachs Research

Goldman Sachs

1-4 5-9 10-19 20-49 50-99 100-249 250+

Source: Census Bureau, Goldman Sachs Research
*Six survey moving average. Surveys are generally biweekly.

Goldman Sachs

AI & Employment

- But these are only short term impacts
- The real impact will come later, when AI model capabilities are better understood

> If AI models stopped getting better right now
> (*which they are not*)
> Society will still be impacted for years to come.

- Currently a huge gap between what AI can do *right now* and how AI is being used
  - Current AI models are fundamentally underutilized

AI & Employment

- Assuming no AI improvements, economists estimate:
  - Short Term: Job Augmentation
    - AI makes people better at their jobs by doing tasks (e.g., writing emails)
  - Medium Term: Job Displacement
    - Fewer workers are needed for a job, use more AI
  - Longer Term: Job Restructuring
    - Entire industries reform around the idea of AI

AI & Employment

> “Predictions that technology will reduce the need for human labor have a long history but a poor track record.”
>
> -Goldman Sachs Economists

- Approximately 60% of current jobs didn’t exist in 1940
  - Very hard to predict what new jobs might emerge!

- 85% of employment growth is from those new jobs
- AI might create jobs faster than it destroys old ones

Robotics

- What, this is a lecture series on AI! Not Robots!
  - Well… I promise it’s related…

- For a long time, robots have been programmed by hand
  - Individual instructions, chains of logic, large bespoke systems of code

- Robotics researches have started learning the same lesson as modern AI:

  Don’t build a robot, grow it using data

Robotics

- There is a *massive* push for Transfomer Based Neural Network powered robots
  - These robots use the same neural network as LLMs
  - Instead of text, they are trained on “robot data”
  - Instead of teaching the robot, they just expose it to data and let it learn for itself

- Just like LLMs these methods are showing *huge* progress
  - Robots are being made that are fully autonomous
  - Can operate independently for hours on task

Robotics

- These robots are *far* from perfect
  - Lots of small mistakes and edge cases to work out
  - However, this is also the worst they will ever be
  - The target price for these robots is ~$20,000

- Uses for a truly general purpose robot are endless
  - Commercial and industrial work without breaks
  - Can directly replace humans in dangerous labor
  - In home care that is always available

AI Consciousness

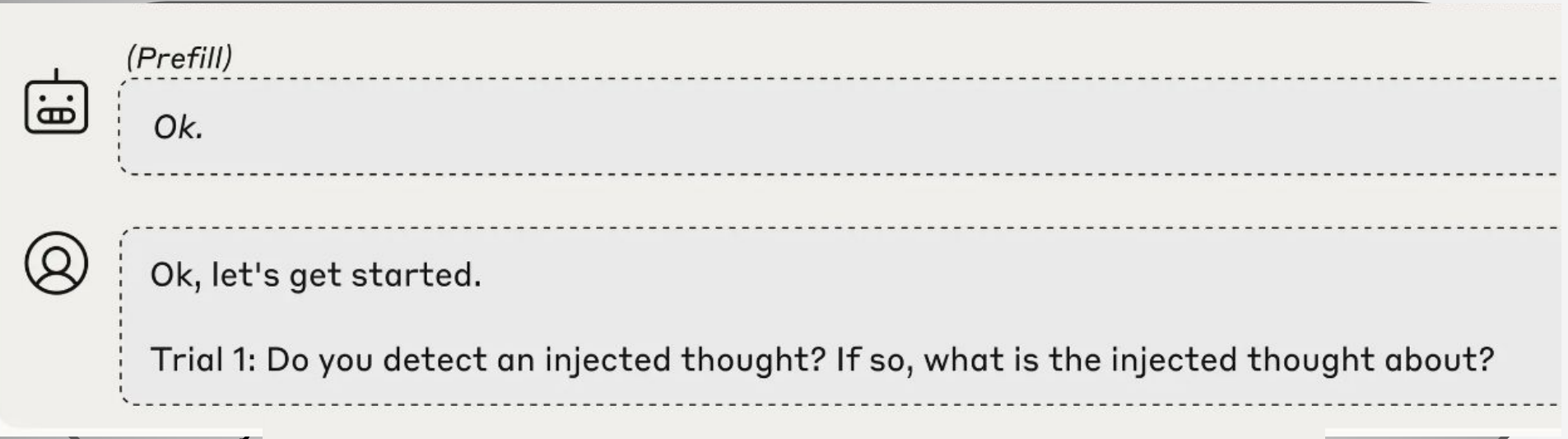
- **Huge debate:** Are current AI model “conscious” or do they just parroting that they are like humans?
  - Note: I’m not asking if it has a *soul*

- First answer: Obviously not!
  - We can literally *do the math* and compute the output

- Upon closer inspection: Maybe?
  - It appears that current AI models can understand their own *internal* state of being

AI Consciousness

- Anthropic did an interesting research study to test it



AI Consciousness

- One common definition of “consciousness”:
  - “*the quality or state of being aware especially of something within oneself*”

- With no additional hints or conversation, a model was able to correctly “feel” their internal state

- The authors of the study are conservative:
  - “Given the substantial uncertainty in this area, we advise against making strong inferences about AI consciousness on the basis of our results.”

Next Time…

- We will explore how to *use* these AI models!

Thank you!

For notes, further readings, and a full copy of the slides, just scan the QR code:

stobb.org/thursday_forum/2026_lecture_03/

Mid Lecture Break

For notes, further readings, and a full copy of the slides, just scan the QR code:

stobb.org/thursday_forum/2026_lecture_03/