Don't Try to Get Answers from a Stochastic Parrot

Emily M. Bender
University of Washington
@emilymbender@bsky.social

Harvey Mudd College
Dr. Bruce J Nelson '74 Distinguished Speaker Series
November 12, 2024

https://bit.ly/EMB-HMC-24

Chatbots 101

- Historically, chatbots were built to handle a specific range of "user intents"

  - with increasing ranges of possible expressions of those intents

  - based on mappings to specific appropriate actions

- Pros: Affordances (what the system can do) fairly clear to users

- Cons: Extending such systems time-consuming and brittle

- What if instead the computer could reply to any input, with text that was engaging and on-topic?

https://bit.ly/EMB-HMC-24

Large language models (LLMs) 101

- A language model is a model of the distribution of the spellings of (parts of) words in some collection of text

- For large language models this collection is enormous (trillions of words) … and mostly undocumented

- Languages are symbolic systems, pairings of form and meaning

- The only thing an LLM 'learns' is what sequence of letters and punctuation is likely to come next

For more, see Bender 2023

Stochastic parrots

- Bender, Gebru et al 2021:

An LM is a system for haphazardly stitching together linguistic forms from its vast training data, without any reference to meaning: a *stochastic parrot*.

So why do ChatGPT et al *seem* to understand?

- Model: What words are likely to come next?

- System design: Output word sequences that are plausible continuations of the input and likely to get high scores from human raters

- User interface: Chat, turning plausible next words into what looks like a conversation

- Users: As skilled language users, we can't help but make sense of text that we encounter.

  - We do this reflexively, by imagining a mind behind the text

For more, see Bender 2024

Chatbots as information access systems

- Google's "AI overviews"

- Microsoft's incorporation of ChatGPT into Bing

- OpenAI's "ChatGPT search"

- Earlier: Meta's Galactica

- San Francisco Chronicle's Kamala Harris "news assistant"

- NYC's regulations chatbot

- …

Chatbots are not a good replacement for search

- When the output is correct, this is just by chance

- RAG (Retrieval Augmented Generation) doesn't fix this: papier-mâché of relevant documents/good data is still papier-mâché

- Effectively checking system output is more time-consuming than using alternative methods in the first place

- A hypothetical chatbot search system that is correct 95% of the time is arguably more dangerous than one that is correct 50% of the time

Chatbots are not a good replacement for search

- But even a hypothetical 100% correct searchbot would be a bad information access technology

- Setting things up so you get "the answer" to your question cuts off your ability to do sense-making

  - Refining your question

  - Understanding how different sources speak to your question

  - Locating each source within the information landscape

Situating information sources:
"How can you treat club foot?"

Mayo Clinic
https://www.mayoclinic.org › ... › Diseases & Conditions ⋮

Clubfoot - Diagnosis and treatment

Jun 28, 2019 — Stretching and casting (Ponseti method) This is the most **clubfoot**. Your doctor will: Move your baby's foot into a correct ...

Diagnosis · Treatment · Stretching And Casting...

WebMD
https://www.webmd.com › what-is-clubfoot ⋮

Clubfoot: Why It Happens & How Doctors Treat It

May 2, 2023 — **Treatment**. Your doctor will begin to correct your baby's **clubfo** they're born. Babies don't use their feet until they learn to ...

Discussions and forums

What is the treatment for clubfoot in adults, and what i

Q www.quora.com · Jun 10, 2023 · 2 posts ⋮

My son was born 7 weeks premature with a mild clubfo

Q www.quora.com · Oct 21, 2021 ⋮

My son will be born with a clubfoot. : r/Mommit - Reddit

🔴 www.reddit.com · Apr 27, 2023 ⋮

Pinterest · tedlcamp12
90+ followers ⋮

Club Foot Care

... **Remedies**, **Herbal Remedies**, Natural **Remedies**,. Get ... Bathing **Clubfoot** Baby | Bath **Clubfoot** | **Talipes** Bathing **Club Foot** Baby, Baby Club, ... Nursing a **Clubfoot** Baby ...

*This last one required adding "herbs -hospital" to the query

See more →

Information literacy practices

- The chatbots as search paradigm encourages us to accept answers as given

  - … especially when stated with friendly authoritativeness

- Especially now we need to level-up our information access practices

  - … and hold high expectations regarding provenance

For more, see Shah & Bender 2022a
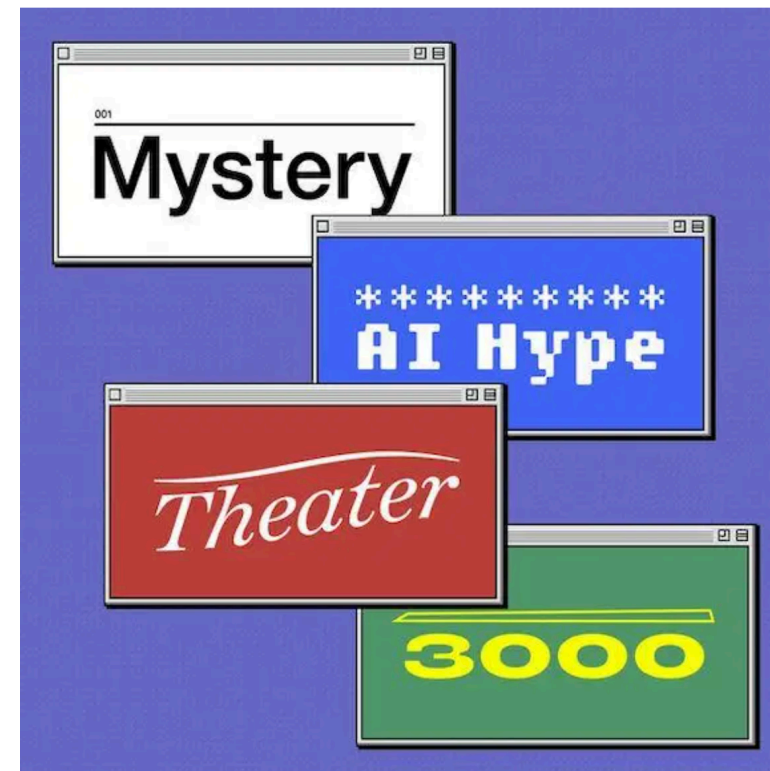Shah & Bender 2022b
Shah & Bender 2024

A word on AI hype

- The very phrase *artificial intelligence* is a kind of hype

- ChatGPT, AlphaFold, and DALL-E aren't the same thing — but talking about them as if they were makes it harder to understand how they actually function

- The ersatz fluency of ChatGPT is especially misleading, trading as it does on your linguistic ability

For more, check out:

- [Mystery AI Hype Theater 3000](#)

- And coming in May 2025:

*THE AI CON
How to Fight Big Tech's Hype and
Create the Future We Want*
Emily M. Bender & Alex Hanna