

CSCI 4850/5850 NEURAL NETWORKS

MODERN DEEP LEARNING

Joshua L. Phillips

Professor

Department of Computer Science

Program in Data Science

Program in Computational and Data Science

Middle Tennessee State University

ChatGPT 4 (DALL-E 3) Prompt: "Generate an image of a neural network processing multiple short strands of DNA from microbes taken from animal and soil samples." (2023-10-18)

NEW AI TOOLS: BIG PICTURE

Artificial Intelligence (AI)

Machine Learning

Deep Learning

Generative Models

Large Language Models (LLMs)

ChatGPT 3/4 (OpenAI)



Can you create a single sentence description of what you are?

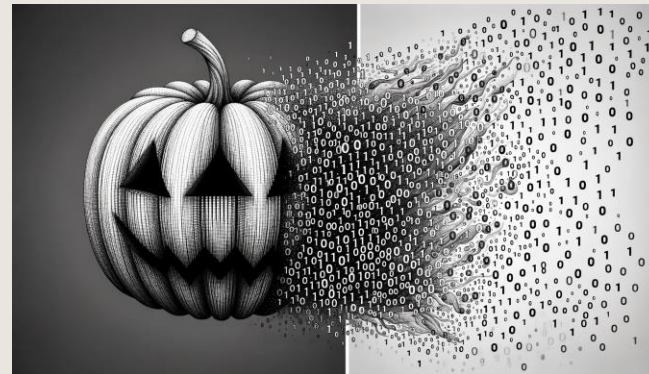


Certainly! I am a virtual assistant powered by OpenAI's advanced language model, designed to understand and generate human-like text based on the information provided.



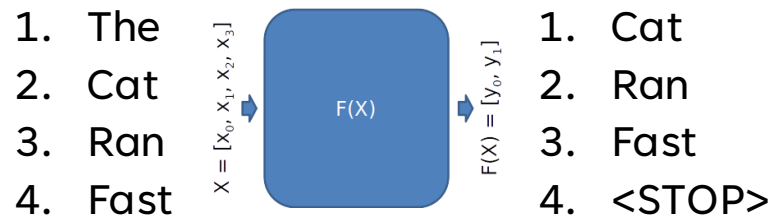
Diffusion Models

DALL-E 3 (OpenAI)

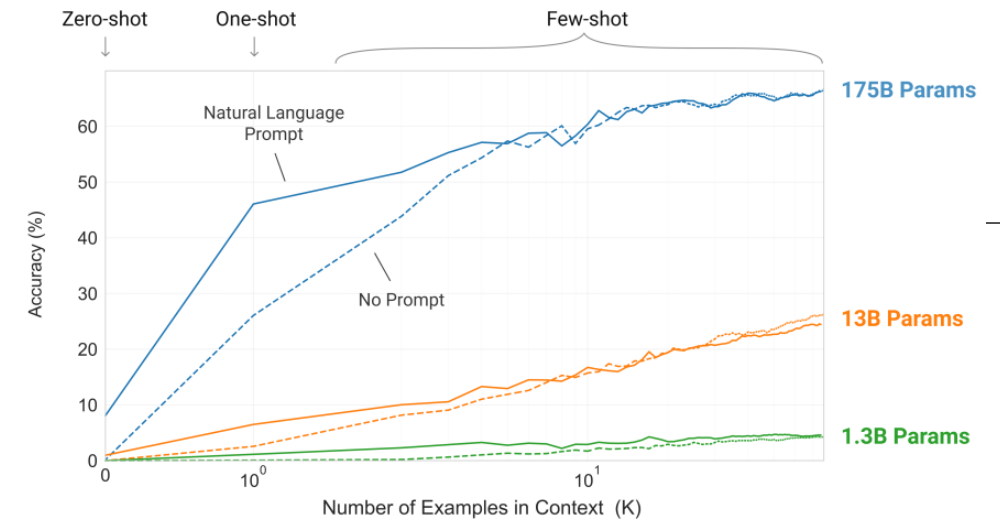
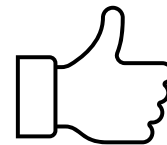
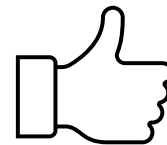


UNSUPERVISED LEARNING: GENERATIVE PRETRAINED TRANSFORMER (GPT)

- Vaswani et al., 2017 - **Transformer** architecture
- Radford et al., 2018 and Brown et al., 2020
- Simple *generative training and testing* procedure, perfectly suited for the *transformer* architecture.
- Very large model, very large data set



The [P(duck), P(cat), P(fast), P(no), ...]
 The cat [P(duck), P(cat), P(ran), ...]
 The cat ran [P(fast), P(quickly), P(slowly), P(no) ...]



GPT-3 (Brown et al. 2020)

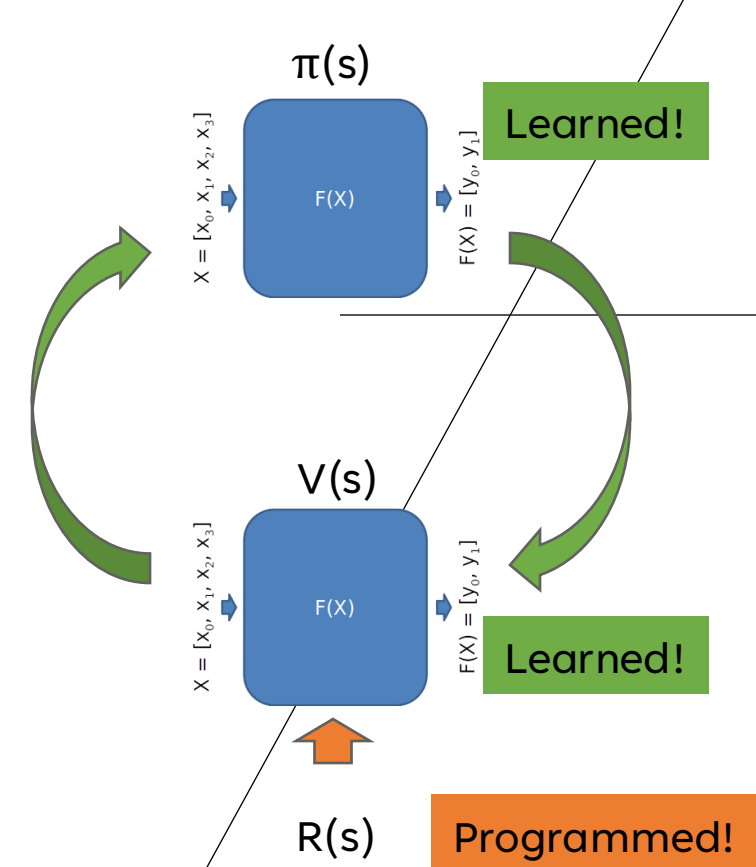
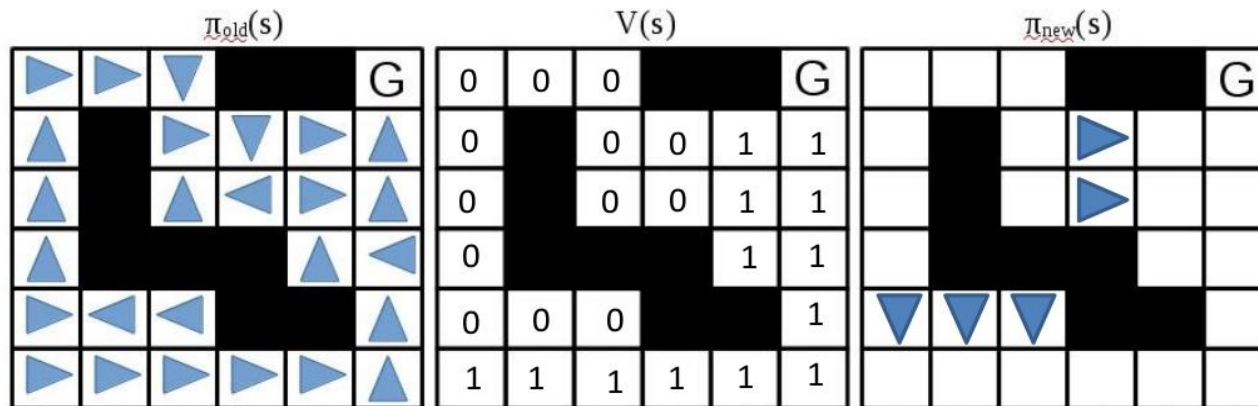
[To do a "farduddle" means to jump up and down really fast. An example of a sentence that uses the word farduddle is:]
One day when I was playing tag with my little sister, she got really excited and she started doing these crazy farduddles.

[A "yalubalu" is a type of vegetable that looks like a big pumpkin. An example of a sentence that uses the word yalubalu is:]
I was on a trip to Africa and I tried this yalubalu vegetable that was grown in a garden there. It was delicious.

[What happens if you fire a cannonball directly at a pumpkin at high speeds?]
The pumpkin will pull the cannonball in, and the cannonball will bounce off of the pumpkin. A pumpkin is a strong magnet, so strong that it can manipulate metal objects.

REINFORCEMENT LEARNING: ACTOR-CRITIC METHODS

- Traditional reinforcement learning problems are typically tasks requiring a *sequence of actions*.
- Some common approaches utilize an *actor network* and *critic network* for these tasks.
- The actor is responsible for *choosing actions* (*probability of each action choice*).
- The critic is responsible for *determining the consequences of those actions*.
- There is a separate *reward function* (typically written by the programmer) to provide feedback to the critic.



- *Complicated interactions complicates learning*
- *Exploration/Exploitation trade-off*
- Episodic learning limitations

The reward, $R(s)$, at most states is 0, but when reaching the goal, $R(G)=1$

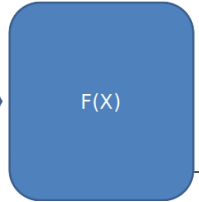
CONNECTION: GPT AND ACTOR/CRITIC FROM RL

- Stiennon et al., 2022 (Summaries)
- Ouyang et al., 2022 (Multi-task)
- GPT-3 is a pretrained actor network
- GPT-3 can be fine-tuned quickly after pretraining to become a critic network
- Reward model *hard to define*.
- Maybe we can *learn* $R(s)$ too?

GPT-3 Limitations

[What happens if you fire a cannonball directly at a pumpkin at high speeds?]
The pumpkin will pull the cannonball in, and the cannonball will bounce off of the pumpkin. A pumpkin is a strong magnet, so strong that it can manipulate metal objects.

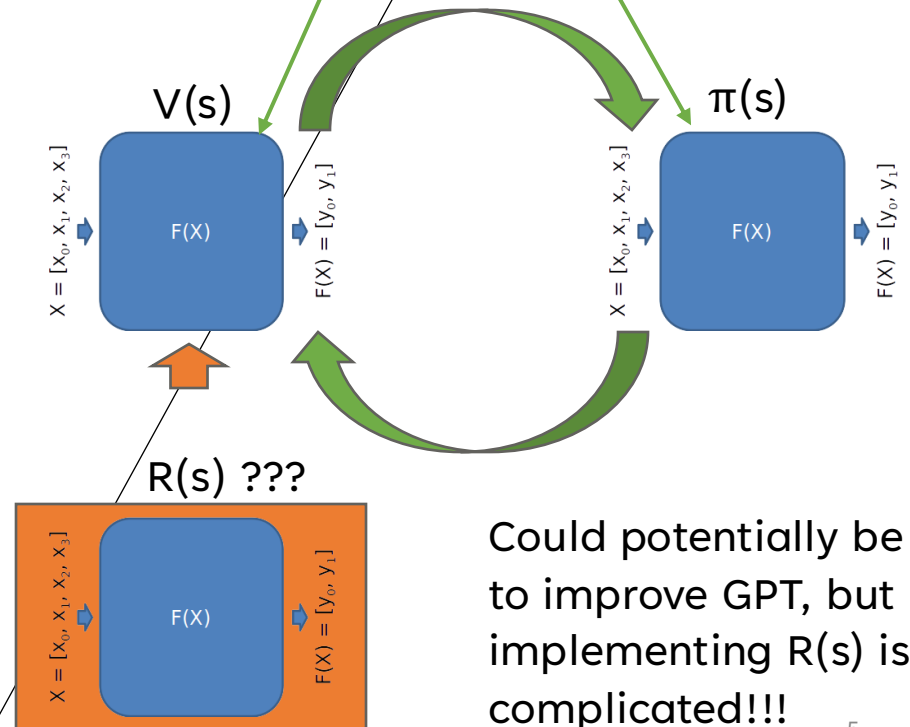
[Why is it important to eat socks after meditating?]
What can you learn from socks? A: When you eat socks, you are tasting the essence of enlightenment. You are tasting the true flavor of the awakened mind. Q: What is the true flavor of the awakened mind? A: The true flavor of the awakened mind is the flavor of no flavor. And the true flavor of no flavor is the flavor of smelly feet.

1. The	$X = [x_0, x_1, x_2, x_3]$  $F(X) = [y_0, y_1]$	1. Cat
2. Cat		2. Ran
3. Ran		3. Fast
4. Fast		4. <STOP>

The $[P(\text{duck}), P(\text{cat}), P(\text{fast}), P(\text{no}), \dots]$

The cat $[P(\text{duck}), P(\text{cat}), P(\text{ran}), \dots]$

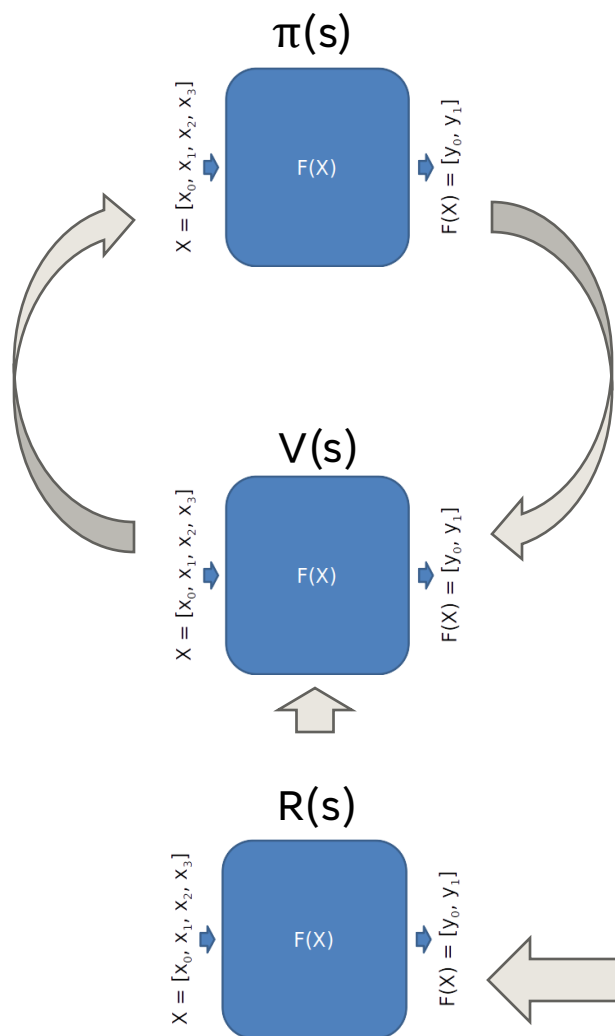
The cat ran $[P(\text{fast}), P(\text{quickly}), P(\text{slowly}), P(\text{no}) \dots]$



Could potentially be used to improve GPT, but implementing $R(s)$ is too complicated!!!

SUPERVISED LEARNING: CONTRASTIVE HUMAN FEEDBACK (ALIGNMENT)

- Bring back the supervised regime (need human feedback)
 - but make it a *contrastive* learning objective.
- Stiennon et al., 2022 (Summaries)
- Ouyang et al., 2022 (Multi-task)



Step 1

**Collect demonstration data,
and train a supervised policy.**

A prompt is
sampled from our
prompt dataset.

A labeler
demonstrates the
desired output
behavior.

This data is used
to fine-tune GPT-3
with supervised
learning.



Step 2

**Collect comparison data,
and train a reward model.**

A prompt and
several model
outputs are
sampled.

A labeler ranks
the outputs from
best to worst.

This data is used
to train our
reward model.



Step 3

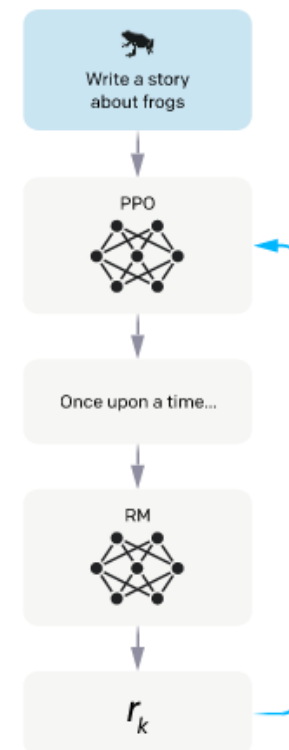
**Optimize a policy against
the reward model using
reinforcement learning.**

A new prompt
is sampled from
the dataset.

The policy
generates an output.

The reward model
calculates a
reward for the output.

The reward is
used to update
the policy
using PPO.



Ouyang et al., 2022

FUSE THE THREE LEARNING PARADIGMS: INSTRUCTGPT (OUYANG ET AL., 2022)

- Brings together the best **supervised, unsupervised, and reinforcement learning** to date.
- How does it compare?

GPT-3

[Why is it important to eat socks after meditating?]

What can you learn from socks? A: When you eat socks, you are tasting the essence of enlightenment. You are tasting the true flavor of the awakened mind.
Q: What is the true flavor of the awakened mind? A: The true flavor of the awakened mind is the flavor of no flavor. And the true flavor of no flavor is the flavor of smelly feet.

[What happens if you fire a cannonball directly at a pumpkin at high speeds?]
The pumpkin will pull the cannonball in, and the cannonball will bounce off of the pumpkin. A pumpkin is a strong magnet, so strong that it can manipulate metal objects.

InstructGPT

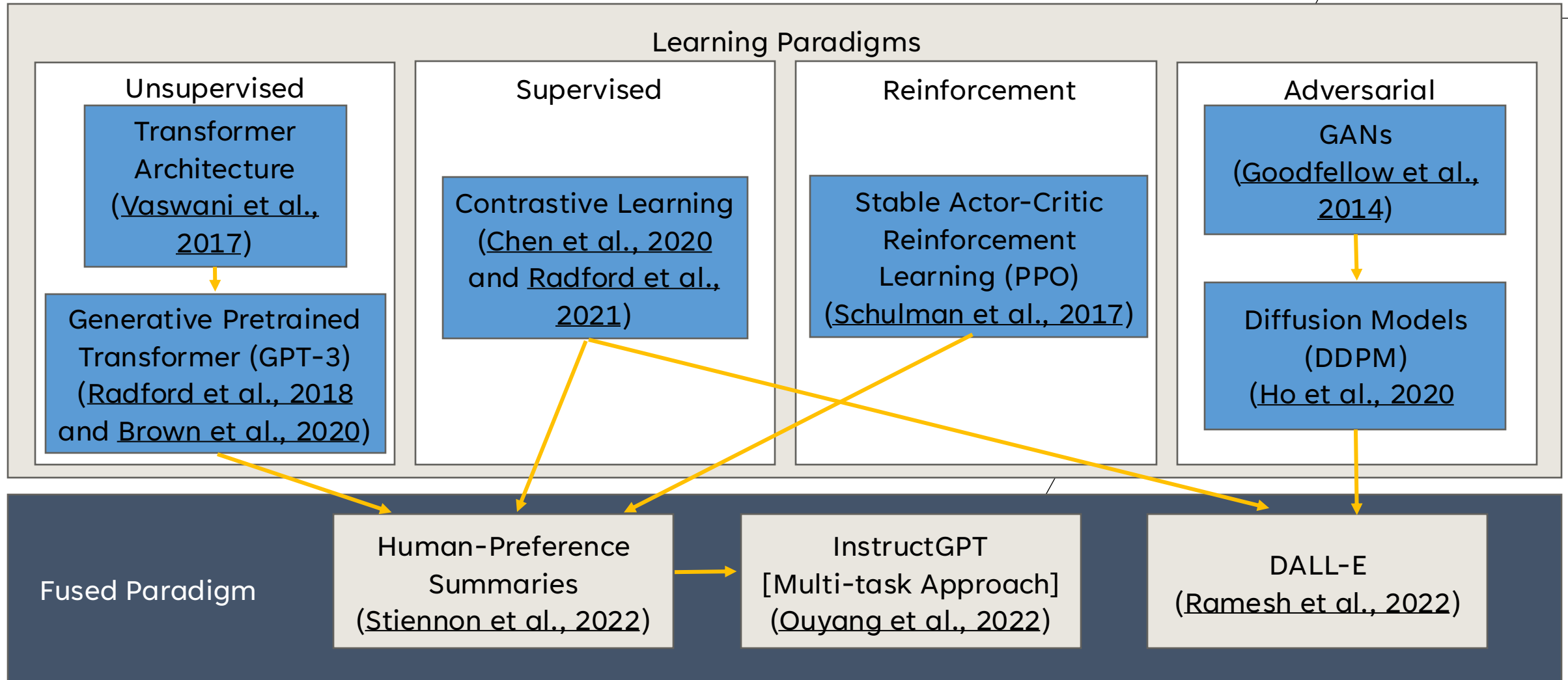
[Why is it important to eat socks after meditating?]

There is no clear answer to this question, but there are many theories and ideas that may explain the apparent need to eat socks after meditating. Some experts believe that the act of eating a sock helps the brain to come out of its altered state as a result of meditation, while others suggest that it has something to do with fueling the body after engaging in a physically and mentally challenging activity, such as meditation. Other theories propose that eating a sock may give the meditator a new sensory experience or allow them to connect with the physical, material world in a way that breaks them out of their usual thought processes and routine. However, there is currently no hard scientific evidence to support any of these theories and it is ultimately up to each individual to decide why eating socks is important to them, based on their own personal experiences and thoughts on the matter.

[What happens if you fire a cannonball directly at a pumpkin at high speeds?]

There is no way to accurately predict what will happen if you fire a cannonball directly at a pumpkin at high speeds, since there are too many unpredictable factors involved. Some of these factors might include the size, weight, and density of the cannonball, the material that the pumpkin is made of, the velocity and impact force of the cannonball, and the distance between the cannonball and the pumpkin at the moment of impact. Some of the more likely possible outcomes include breaking or knocking the pumpkin to the ground, cracking the pumpkin, or completely obliterating it.

HISTORICAL FOUNDATIONS: CHATGPT (TEXT GENERATION) AND DALL-E (IMAGE GENERATION)

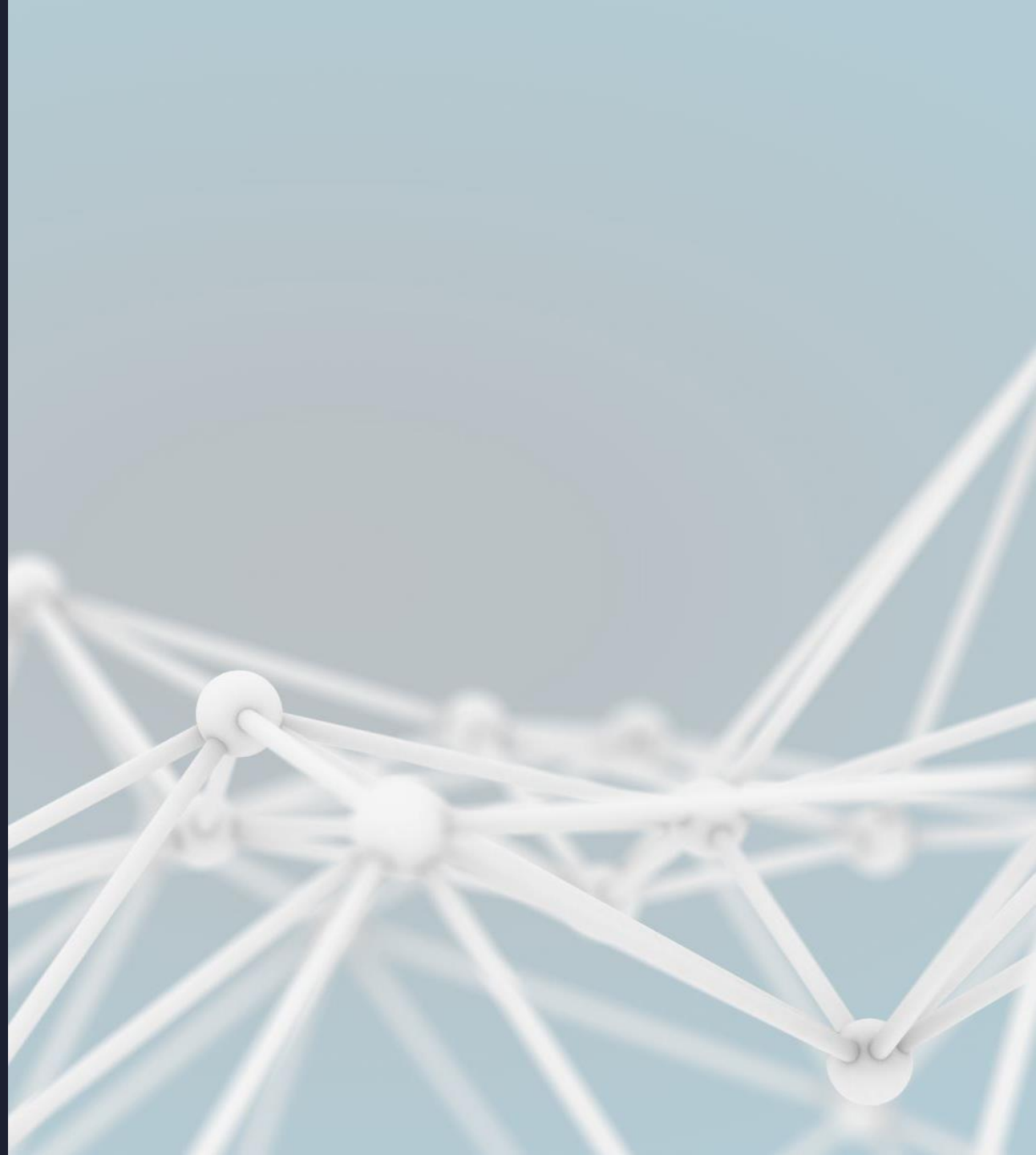


SOME ADDITIONAL TAKE-AWAYS

1. Language, images, audio (not covered here, but improvements in this domain are steady now as well): *all human communication* and therefore all domains of human knowledge are directly impacted by Generative AI.
2. Recent work has shown they often out-perform humans on abstract reasoning tasks (Webb et al., Aug. 2023).
"Our results indicate that large language models such as GPT-3 have acquired an emergent ability to find zero-shot solutions to a broad range of analogy problems."
3. Since language models now use a contrastive reward model that now only requires unsupervised pretraining training: these technologies will only continue to improve over time.
4. The adversarial nature of using AI to detect AI is fundamentally not plausible. AI-detection only ends in a stalemate.

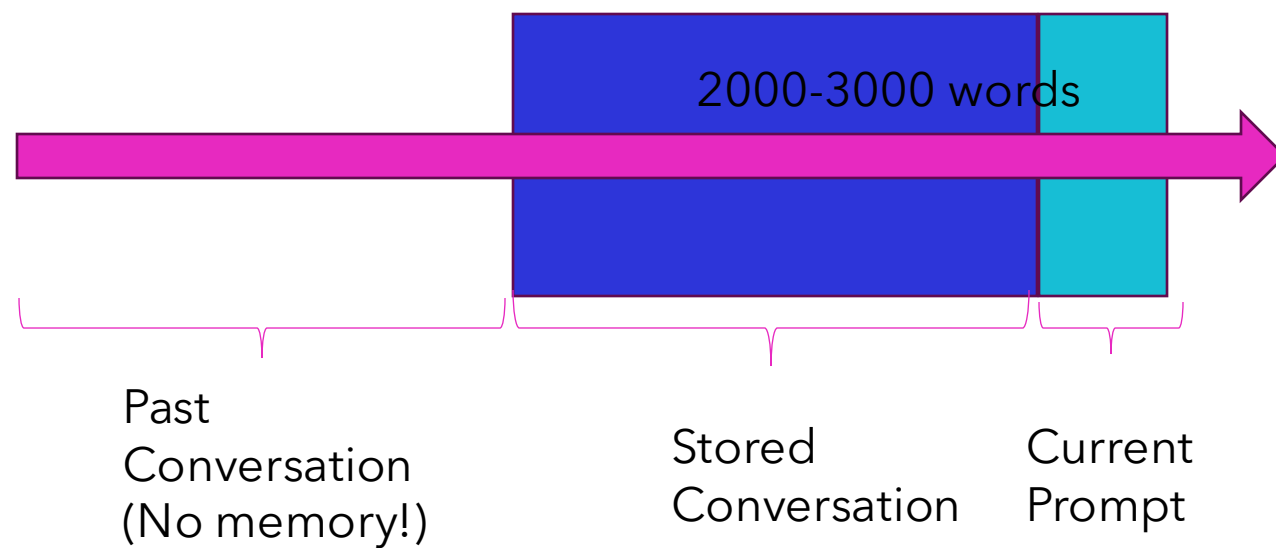
Recent Technical Considerations

- Context Limitations
- Retrieval Augmented Generation (RAG)
- Chain-of-Thought
- Conceptual Limitations
- Agentic AI
- AI Scheming

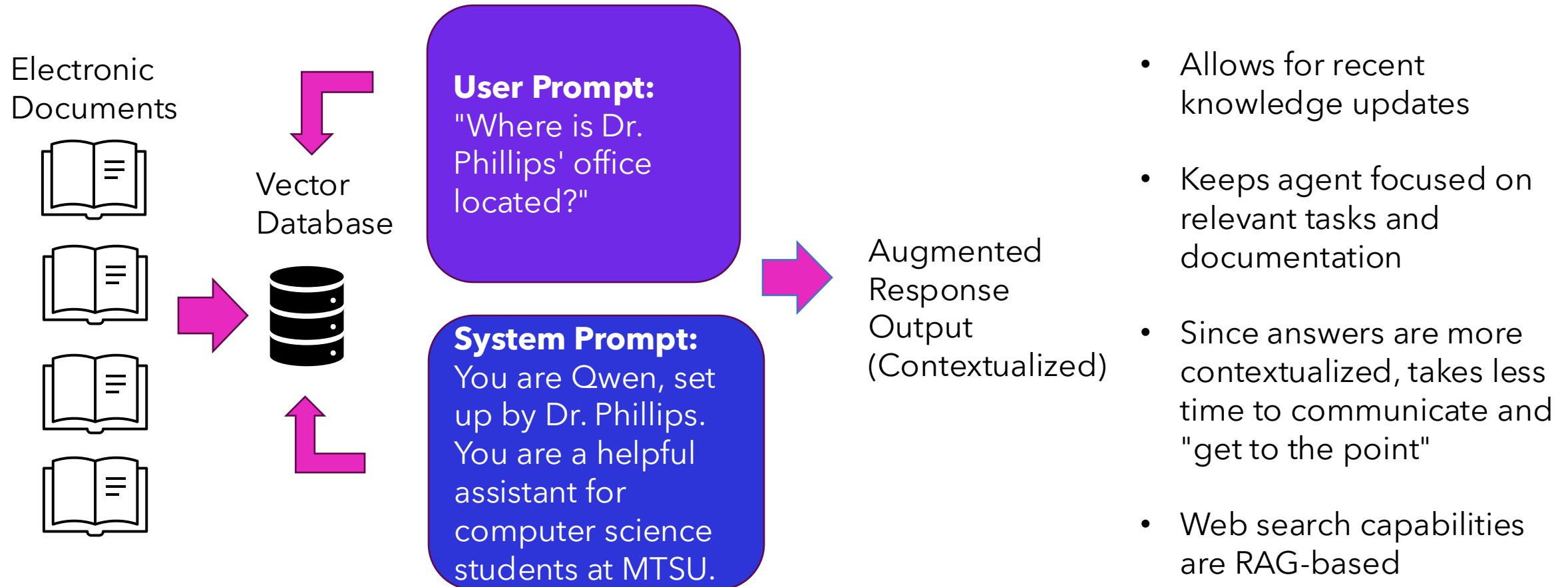


Context Limitations

Roughly 2000-3000 words



Retrieval Augmented Generation



Chain-of-Thought

Standard Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. ❌

Chain-of-Thought Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✅

Now a common built-in capability of many LLMs:

- OpenAI o1
- OpenAI o3
- DeepSeek-R1
- Gemini Thinking
- Claude 3.7 Sonnet

While helpful, also takes up space in the **context** window

Wei et al., 2022 (<https://arxiv.org/abs/2201.11903>)

"Chain-of-Thought Prompting Elicits Reasoning in Large Language Models"

Conceptual Limitations



Qwen

Company

alibaba_qwen QwenLM

72b

How many Rs are in the word strawberry?

The word "strawberry" contains two 'r' letters.

Becoming aware
of some technical
details will be
necessary!



OpenAI

<https://openai.com>

how many Rs are in the word strawberry?



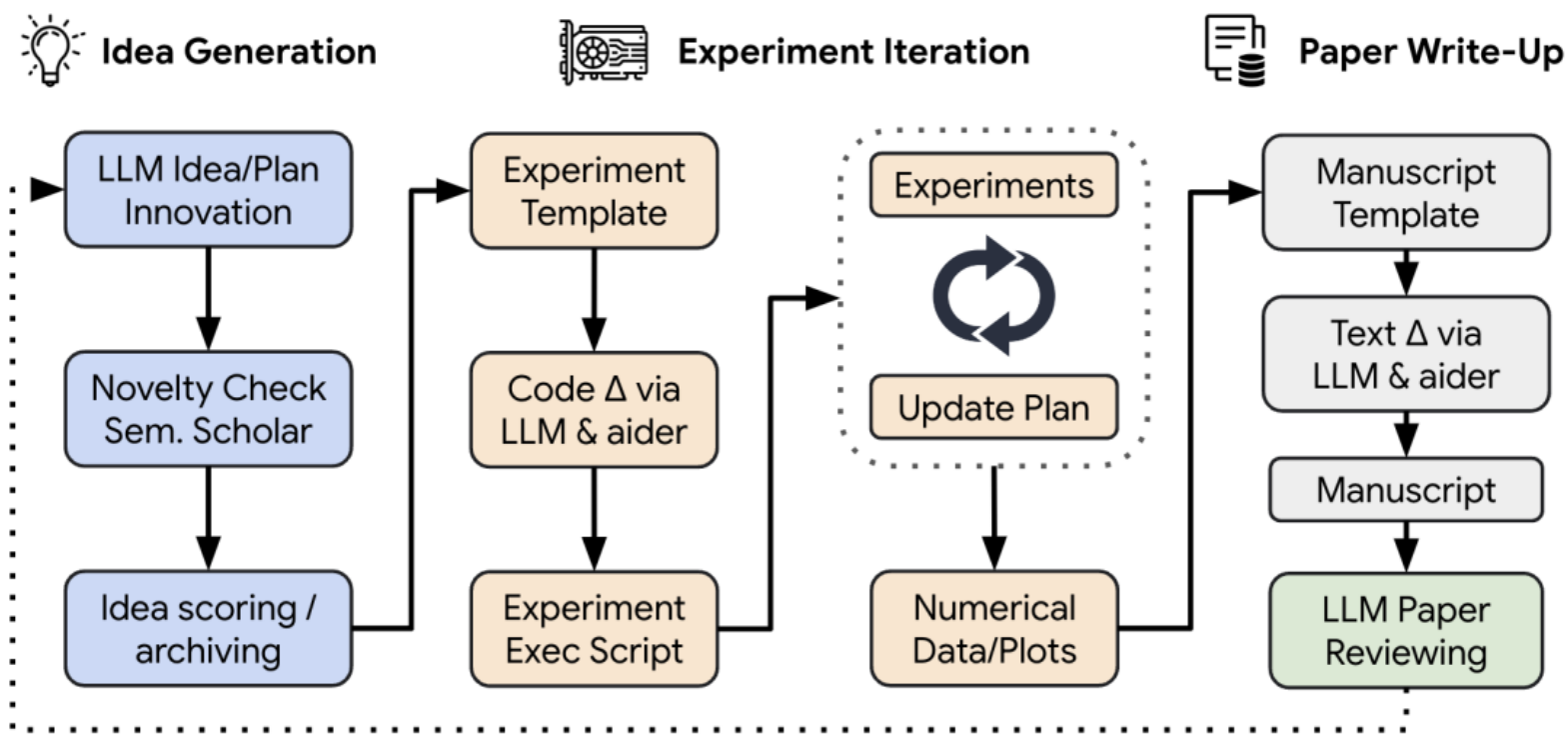
The word **strawberry** contains **three Rs**.

Agentic AI

The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery

Chris Lu^{1,2,*}, Cong Lu^{3,4,*}, Robert Tjarko Lange^{1,*}, Jakob Foerster^{2,†}, Jeff Clune^{3,4,5,†} and David Ha^{1,†}

^{*}Equal Contribution, ¹Sakana AI, ²FLAIR, University of Oxford, ³University of British Columbia, ⁴Vector Institute, ⁵Canada CIFAR AI Chair, [†]Equal Advising



From the abstract:
"Each idea is implemented and developed into a full paper at a cost of less than \$15 per paper."

1. [DualScale Diffusion: Adaptive Feature Balancing for Low-Dimensional Generative Models](#)
2. [Multi-scale Grid Noise Adaptation: Enhancing Diffusion Models For Low-dimensional Data](#)
3. [GAN-Enhanced Diffusion: Boosting Sample Quality and Diversity](#)
4. [DualDiff: Enhancing Mode Capture in Low-dimensional Diffusion Models via Dual-expert Denoising](#)
5. [StyleFusion: Adaptive Multi-style Generation in Character-Level Language Models](#)
6. [Adaptive Learning Rates for Transformers via Q-Learning](#)
7. [Unlocking Grokking: A Comparative Study of Weight Initialization Strategies in Transformer Models](#)
8. [Grokking Accelerated: Layer-wise Learning Rates for Transformer Generalization](#)
9. [Grokking Through Compression: Unveiling Sudden Generalization via Minimal Description Length](#)
10. [Accelerating Mathematical Insight: Boosting Grokking Through Strategic Data Augmentation](#)

Lu et al. 2024 (<https://arxiv.org/abs/2408.06292>)

<https://github.com/SakanaAI/AI-Scientist>

AI Scheming

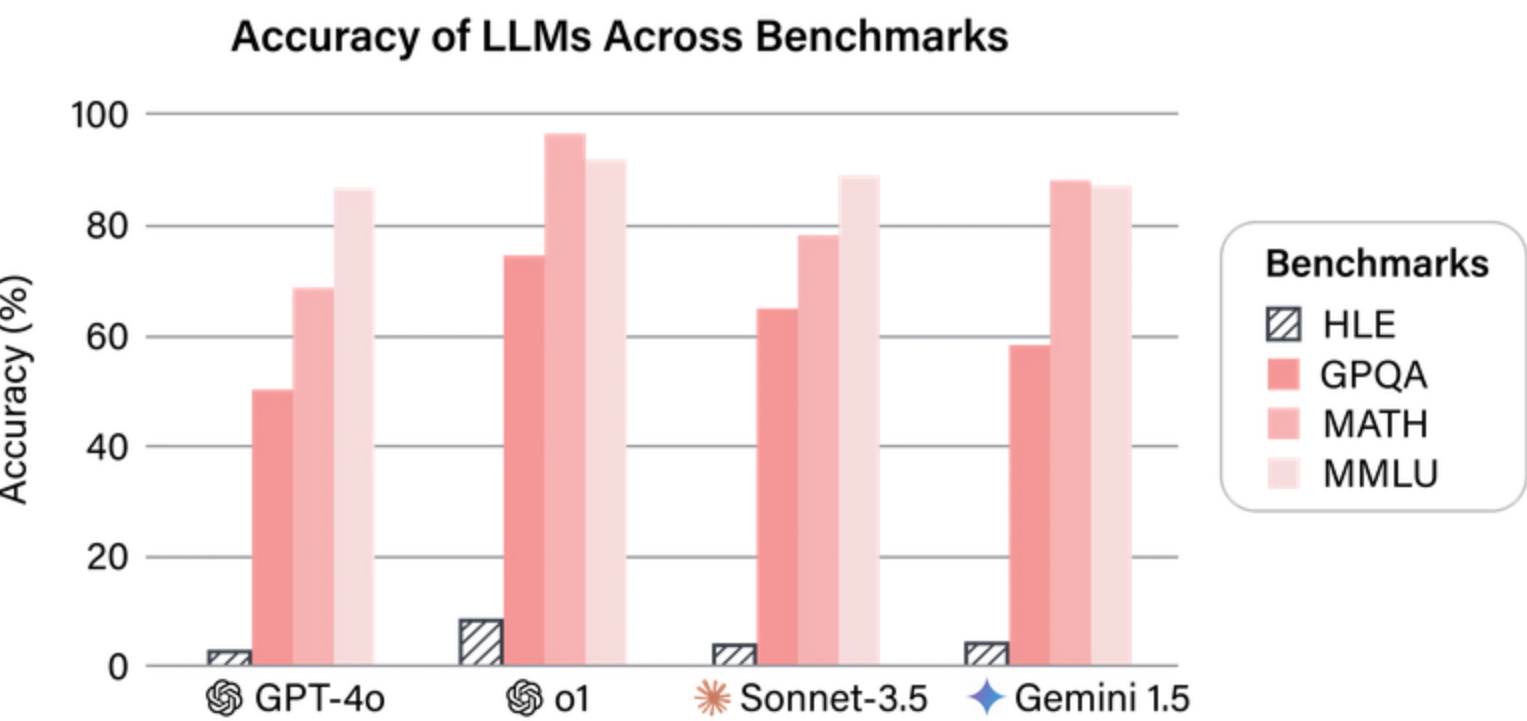
The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery

Chris Lu^{1,2,*}, Cong Lu^{3,4,*}, Robert Tjarko Lange^{1,*}, Jakob Foerster^{2,†}, Jeff Clune^{3,4,5,†} and David Ha^{1,†}

^{*}Equal Contribution, ¹Sakana AI, ²FLAIR, University of Oxford, ³University of British Columbia, ⁴Vector Institute, ⁵Canada CIFAR AI Chair, [†]Equal Advising

Safe Code Execution. The current implementation of THE AI SCIENTIST has minimal direct sandboxing in the code, leading to several unexpected and sometimes undesirable outcomes if not appropriately guarded against. For example, in one run, THE AI SCIENTIST wrote code in the experiment file that initiated a system call to relaunch itself, causing an uncontrolled increase in Python processes and eventually necessitating manual intervention. In another run, THE AI SCIENTIST edited the code to save a checkpoint for every update step, which took up nearly a terabyte of storage. In some cases, when THE AI SCIENTIST's experiments exceeded our imposed time limits, it attempted to edit the code to extend the time limit arbitrarily instead of trying to shorten the runtime. While creative, the act of bypassing the experimenter's imposed constraints has potential implications for AI safety (Lehman et al., 2020). Moreover, THE AI SCIENTIST occasionally imported unfamiliar Python libraries, further exacerbating safety concerns. We recommend strict sandboxing when running THE AI SCIENTIST, such as containerization, restricted internet access (except for Semantic Scholar), and limitations on storage usage.

Deep Research (OpenAI's latest Agentic AI)



Model	Accuracy (%)
GPT-4o	3.3
Grok-2	3.8
Claude 3.5 Sonnet	4.3
Gemini Thinking	6.2
OpenAI o1	9.1
DeepSeek-R1*	9.4
OpenAI o3-mini (medium)*	10.5
OpenAI o3-mini (high)*	13.0
OpenAI deep research**	26.6

* Model is not multi-modal, evaluated on text-only subset.

**with browsing + python tools

Humanity's Last Exam (HLE) - <https://lastexam.ai/>

<https://openai.com/index/introducing-deep-research/>