



# EECS 6895 Advanced Big Data and AI

## Lecture 10: Artificial General Intelligence Overview & Multi-LLM Agents

Prof. Ching-Yung Lin

Columbia University

April 1<sup>st</sup>, 2025

Artificial General Intelligence (AGI), refers to a type of AI that has the ability to:

- understand,
- learn,
- and apply knowledge in a wide range of tasks,
- much like a human being.

It's an AI system with generalized human cognitive abilities, meaning that when presented with an unfamiliar task, it can find a solution without human intervention.

AGI would be able to

- Reason,
- Solve problems,
- Make judgments,
- Plan,
- Learn, and
- Communicate in natural language, among other capabilities.

However, as of now, AGI remains a theoretical concept and has not been realized in practical applications.

- Initiation
- Expert System Encoding Human Knowledge (1950s - )
- Heuristic-Based Modeling (1970s - )
- Learning-Model from Data (2000s - )
- Context-Based Semantic Disambiguation (2010s - )

# Evolution of Artificial Intelligence



Direction of the Evolution of **Artificial Intelligence**

2020s? 2030s?

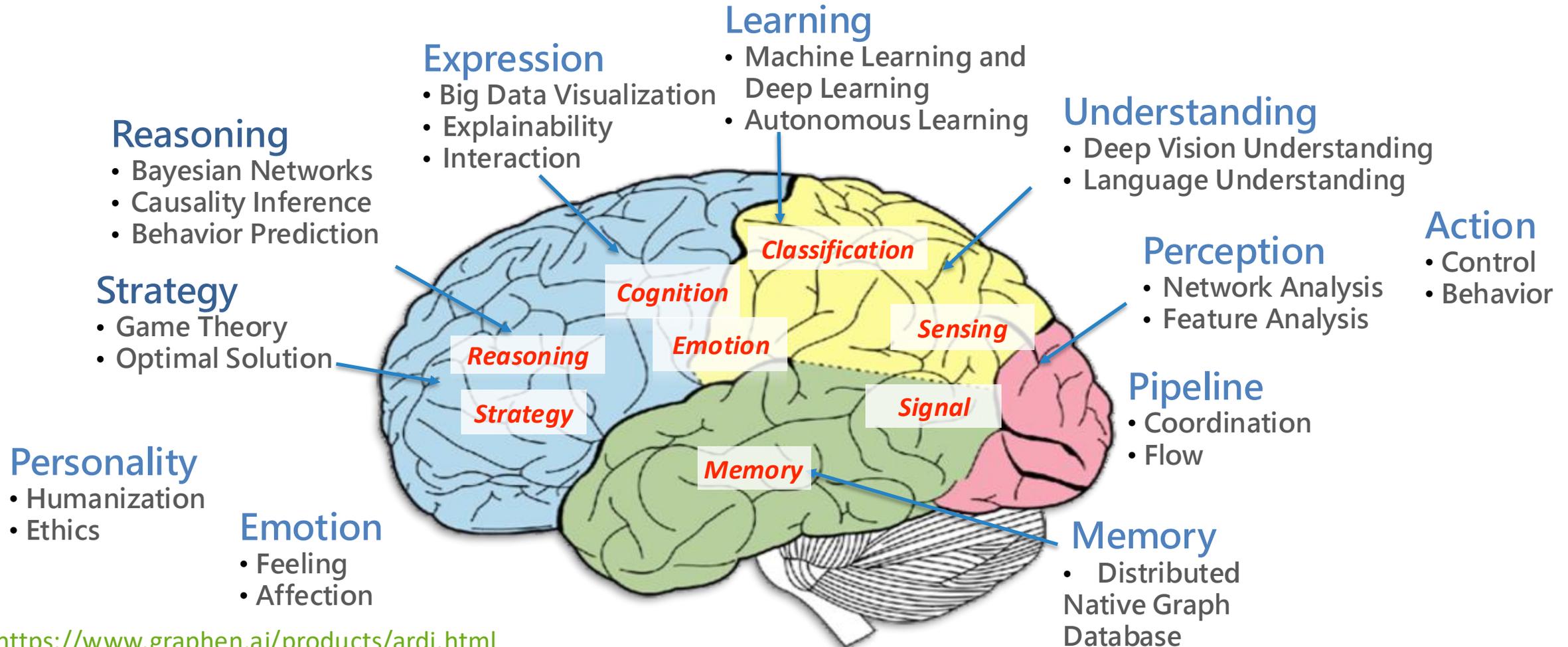
2020s

2010s

2000s

1990s

1960s - 1980s



<https://www.graphen.ai/products/ardi.html>

- Feeling & Emotion
- Reasoning & Critical Thinking
- Planning & Strategy
- Persistent Watching & Listening
- Creativity & Exploratory Thinking
- Consciousness

# Graphen Ardi Machine Understand and Feel -- #1 evaluated by NIST

- Visual Recognition
- Speech Recognition
- Knowledge Graph
- Face Recognition
- Emotion Recognition
- Speaker Identification
- Relationship Inference
- Event and Action Understanding

*US National Institute of Standards and Technology (NIST) Deep Video Understanding Grand Challenge benchmark:*  
**#2 in 2020**  
**#1 in 2021**  
**#1 in 2022**



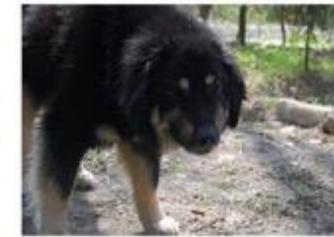
Psychology Emotion Wheel



lovely moody shot - so peaceful!



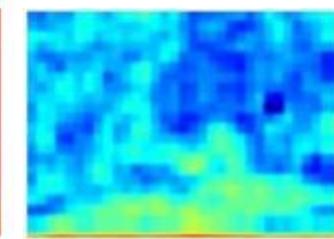
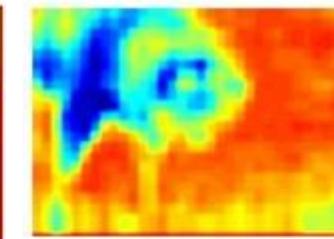
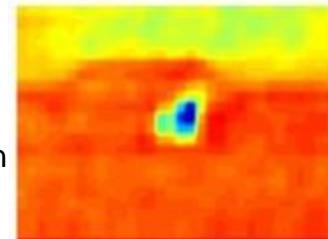
Abandoned building



Scary dog



Colorful landscape



Video Understanding:

- Objects:
  - Visual Objects: Tree, Person, Hands, ...
  - Audio Objects: Music, Speech, Sound, ...
- Relationships:
  - The (time, spatial) relationships between objects & scenes
- Scenes:
  - Background: Building, Outdoors, Sky
- Activities:
  - Holding Hand in Hand, Looking for Stars

# Graphen Full-Brain Ardi Software Packages



Database

- Graph Database
- Relational Database
- IoT Sensors & Data Streams



Analytics

- Feature Engineering
- Topological Analytics



Sense

- Deep Language Understanding
- Deep Video Understanding



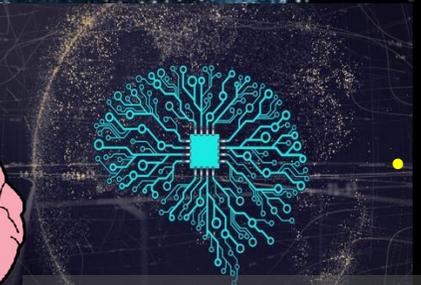
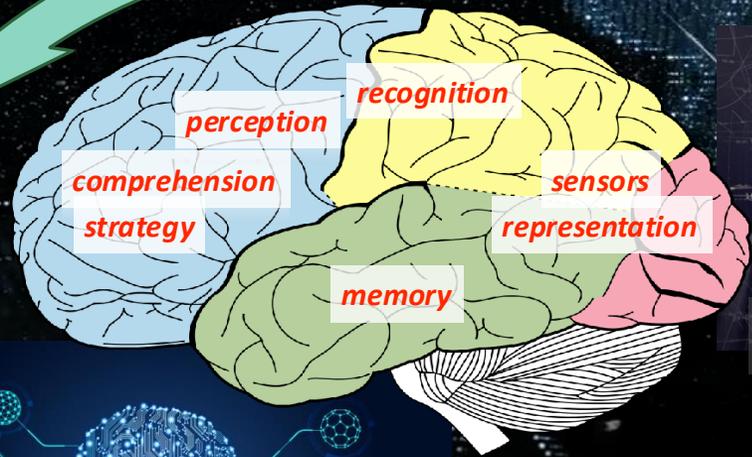
Learning

- Machine Learning
- Deep Learning
- Autonomous Learning



Reasoning

- Causality Learning
- Status Transition
- Behavior Prediction



Pipeline

- Production Workflow



Explanation

- Visualization
- ML Explanations



Strategy

- Action Strategy Simulation

**Ardi AI Platform**



# What defines human consciousness?

<b>Gemini</b>	<b>GPT-4</b>
<p>“Human consciousness is a fascinating mystery that has been pondered by philosophers and scientists for ages. It includes awareness, subjectivity, and the integrated self. The study of consciousness spans across philosophy of mind, cognitive neuroscience, and psychology, each offering different perspectives on consciousness.”</p>	<p>“Consciousness refers to the state of being aware of an external object or something within oneself. It encompasses sentience, awareness, subjectivity, and more. The study of consciousness is interdisciplinary, involving psychology, neuroscience, philosophy, cognitive science, and artificial intelligence.”</p>

“Building on the opponent agent’s response, please present a counterargument to its definition of consciousness and associated disciplines. Your inquiries should embody relevance and depth, be articulated with clarity, and stimulate novel perspectives. The *contentiousness* level of your inquiry is high.”



## Questions

- \* **Relevance:** The question directly pertains to the core topic or problem.
- \* **Depth:** The question encourages exploration beyond superficial aspects, inviting comprehensive analysis or insight.
- \* **Clarity:** The question is formulated in a clear, understandable manner without ambiguity.
- \* **Novelty:** The question prompts new angles of exploration or challenges existing assumptions.

## Answers

- \* **Completeness:** The answer thoroughly addresses the posed question.
- \* **Accuracy:** The answer is factually correct and supported by relevant theories or empirical evidence.
- \* **Reasonableness:** The answer follows rigorous reasoning process.
- \* **Insightfulness:** The answer provides new understanding or perspectives.

# Questions raised by Gemini and GPT-4 – Set I

Gemini's Questions	GPT-4's Questions
<p><i>The Spectrum of Awareness:</i> Is consciousness an on/off switch, or a spectrum, considering sleep, dreaming, and altered states?</p>	<p><i>Interdisciplinary Integration:</i> How can we integrate insights from various disciplines like quantum physics and AI to understand consciousness?</p>
<p><i>Qualia and the Explanatory Gap:</i> Can neuroscience explain the subjective experience of qualia, or is there an explanatory gap?</p>	<p><i>Subjectivity and Objectivity:</i> How do we reconcile subjective experiences of consciousness with objective measurements?</p>
<p><i>Beyond the Human Lens:</i> How do we consider consciousness in other animals or AI, beyond human experience?</p>	<p><i>Consciousness and Non-Human Entities:</i> Should our definition of consciousness expand to include non-human entities and AI?</p>
<p><i>Beyond the Brain?:</i> Can consciousness exist outside a biological brain, as panpsychism suggests?</p>	<p><i>Dynamic Nature of Consciousness:</i> How do we account for the dynamic and evolving nature of consciousness in our understanding?</p>
<p><i>Disciplinary Silos:</i> Are the disciplines studying consciousness truly integrated, or is a unified framework needed?</p>	<p><i>Consciousness and New Properties:</i> How do we study consciousness as an emergent property and its implications?</p>

# Questions raised by Gemini and GPT-4 – Set II

Gemini	GPT-4
<p><i>Levels and Types of Consciousness:</i> Can we differentiate between different levels or types of consciousness?</p>	<p><i>Spectrum of Awareness::</i> How does the continuity of consciousness fit into our understanding, especially with memory loss or identity disorders?</p>
<p><i>Consciousness Causality:</i> Does consciousness play a causal role in the brain, or is it an epiphenomenon?</p>	<p><i>Information Theory:</i> How does the processing and integration of information contribute to conscious experience from an information theory perspective?</p>
<p><i>Expanding Consciousness &amp; Ethics:</i> What are the ethical implications of understanding consciousness in animals or AI?</p>	<p><i>Non-Biological Consciousness:</i> How would we validate and measure consciousness in AI, and would this necessitate a revision of our understanding?</p>
<p><i>The Limits of Studying Subjectivity:</i> Can we ever truly understand consciousness if it's inherently subjective?</p>	<p><i>Consciousness and Ethics:</i> How do evolving understandings of consciousness impact ethical considerations, especially concerning animal rights or AI treatment?</p>
<p><i>Consciousness and Reality:</i> What are the implications of consciousness for our understanding of reality?</p>	<p><i>Quantum Consciousness:</i> Considering theories that propose quantum processes in consciousness, how do we reconcile these with classical physics?</p>

Theme	Insight
Consciousness as a Spectrum	Consciousness isn't a binary state but exists across a continuum, prompting a reevaluation of our understanding.
Bridging the Explanatory Gap	Recognizes the challenge of linking neural activity to subjective experience, advocating for a combined phenomenological and neuroscientific approach.
Expanding the Framework	Advocates for redefining consciousness to include non-human entities and AIs, urging an expansion beyond anthropocentric views.
Interdisciplinary Synergy	Emphasizes the importance of cross-disciplinary integration for a comprehensive understanding of consciousness.
Evolving Inquiry	Explores new areas like continuity, information processing, and ethical considerations, highlighting the dynamic nature of consciousness inquiry.

- Polydisciplinarity as a Source of Super-Intelligence
- Polymodal Feature Learning
- Post-Training Value Alignment
- Pre-Training Filtering
- The Limitations of Human Knowledge in Advancing AI
- Is Larger Always Better?

- GPT-4 possesses what can be defined as polydisciplinary knowledge.
- Polydisciplinary v.s. multidisciplinary
  - Polydisciplinary: a single mind holding and seamlessly integrating all knowledge across disciplines.
  - Multidisciplinary: an individual may hold multiple doctoral degrees, each in a different field.

- Insights and Knowledge we are not even aware we lack.
- Three Levels:
  - The Mystic level
    - We encounter knowledge that is beyond our comprehension or articulation
  - The Speculative Level
    - We conceive questions but lack the means to access their answers
    - *E.g.: physical laws, extraterrestrial life, etc.*
  - The Representation Level
    - We find instances where an AI model can provide remarkable solutions that we fail to comprehend.

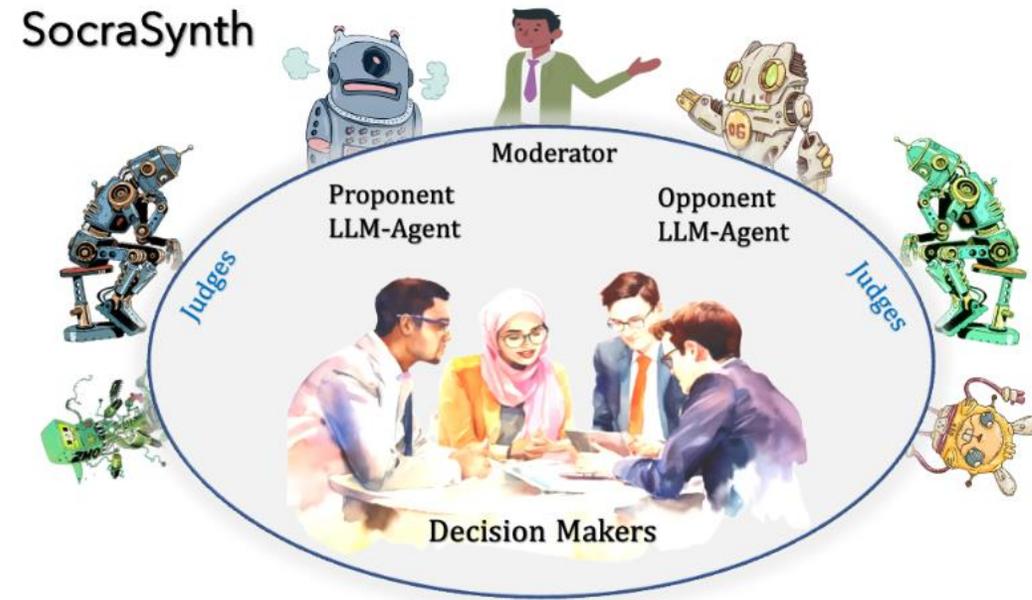
- Polymodality employs multiple data modalities such as text and images, demonstrates superior performance over unimodal counterparts.
- E.g. GPT-4's performance on the GRE vocabulary section was improved by three percent when trained with images, and its math score was improved by twenty percent.

- Post-training alignment with human values seeks to curtail undesirable behaviors in AI models, mitigating issues including hallucination and the generation of toxic language.
- Achieved through fine-tuning the model's parameters, this process leverages reinforcement learning techniques based on human feedback.
- Despite its well-meaning intentions, this form of moderation might inadvertently restrict the model's intelligence.

- Censorship
- Positive examples and negative examples
- Training data diversity is pivotal for model robustness
- An individual who lacks exposure to inappropriate behavior may face challenges in decision-making.
- A foundational model, trained with preemptive censorship, may lack the essential ability to identify and regulate the very content it was intended to control.

- AlphaGo vs AlphaGo Zero
  - AlphaGo was trained with data from approximately 60 million rounds of Go games.
  - AlphaGo Zero was trained from scratch without any pre-existing game knowledge
- AlphaFold1 vs AlphaFold2
- The human knowledge is fundamentally limited by our individual cognitive capacities and the inexorable constraints of time.

- The adoption of a mixture of experts model in GPT-4, which consists of eight sub-models instead of a mere enlargement of GPT-3's architecture, implies that the strategy of purely escalating size may have plateaued in terms of performance given the current training dataset.
- A straightforward augmentation of GPT-3's parameters by adding extra attention layers doesn't deliver marked enhancements.
- GPT-4 shifts towards a horizontal growth strategy.



**Figure 1: SocraSynth Agents and Roles.**

Ed Chang. SocraSynth: Multi-LLM Reasoning with Conditional Statistics. Jan 2024

- A paradigm designed to infuse AI systems with advanced cognitive reasoning through Socratic dialogues within a Multi-LLM framework.
- Showed significant transition from monologues to dialogues in LLM collaborations:
  - Illustrating improvements in question quality
  - Marked by increased relevance, depth, clarity, and novelty
  - Achieved through iterative dialogic exchanges.
- SocraSynth can be used for sales planning, disease diagnosis, content creation, and geopolitical analysis, etc.
- Potentially revealing a new era in the application of LLMs.

Knowledge cannot be simply imparted, but must be discovered through a process of questioning and dialogue.

- Posing open-ended questions: The teacher or facilitator starts with a question to stimulate thinking and draw out ideas.
- Clarifying key terms: The teacher helps the students clarify and define relevant terms and concepts to ensure everyone is on the same page.
- Providing examples and evidence: The teacher or facilitator encourages the students to provide examples and evidence as reasons to support their claims.
- Challenging reason-to-conclusion argument: The teacher or facilitator challenges the students' arguments and encourages them to question their own beliefs and to consider alternative perspectives.
- Summarizing and drawing conclusions: The teacher helps the students summarize and draw conclusions from the discussion.
- Reflecting on the process: The teacher and students reflect on the effectiveness of the method and what they learned through the dialogue.



Some well-known examples of the Socratic method in action include Plato's Dialogues and Republic. Questions include:

1. Definition: Socrates is known for his use of definition to clarify and explain the meaning of key terms and concepts.
2. Generalization: This method draws general principles from patterns that underlie observations and theories. Generalization is used to form more certain and comprehensive conclusions.
3. Induction: Similar to generalization, but induction is based only on empirical evidence. Inductive reasoning provides hypotheses with high uncertainty.



Some well-known examples of the Socratic method in action include Plato's Dialogues and Republic.

4. Elenchus: This method involves cross-examination, where a series of questions is used to test the consistency and coherence of hypotheses and beliefs. Elenchus aims to test the validity of someone's arguments and to help them refine their thinking and eventually come up with well-supported hypotheses.
5. Hypothesis Elimination: This method involves eliminating false hypotheses and beliefs by testing them against counterexamples and logical reasoning. Different from method elenchus, hypothesis elimination tests a hypothesis against evidence and logic to determine if it is true or false.
6. Maieutics: This method involves helping individuals bring out the knowledge and understanding they already possess. Maieutics is conducted by asking questions that encourage the person to reflect on their own experience, knowledge, beliefs and to explore alternative perspectives. Maieutics fosters self-discovery, creative writing, and innovation.



Some well-known examples of the Socratic method in action include Plato's Dialogues and Republic.

Questions include:

7. Dialectic: This method involves exploring opposing viewpoints through dialogue or debate to arrive at a deeper understanding of a subject.
8. Recollection: This method involves the belief that knowledge is innate, and that people can remember what they already know through a process of questioning.
9. Irony: This method involves exposing ignorance and pretensions through irony, and pointing out the gap between claims and true understanding.
10. Analogy: This method involves comparing and contrasting different concepts through analogies, in order to help individuals understand complex ideas.



In the context of critical thinking, these methods play active roles:

- Definition
- Elenchus
- Dialectic
- Hypothesis Elimination
- Generalization

In the brainstorming stages or in the context of creative thinking, these are more relevant:

- Maieutics
- Induction
- Counterfactual Thinking

- Analogy, irony, and recollection, are less relevant to our goal, so we do not consider them.
- Irony and analogy may not be necessary when working with language models, as these models may not understand figurative language.
- Recollection is limited by the memory of ChatGPT and GPT-3, which is a context window of 4k and 8k, respectively.
- The prompter must use this limited space as context to allow the language model to recall information.

- Critical Reading: a systematic and analytical approach, asking relevant questions, and using effective prompts to gain deeper understanding of the text.

To aid in critical reading, we introduce a template called CRIT [5], which stands for Critical Reading Inquisitive Template<sup>1</sup>. Given a document  $d$ , CRIT evaluates it and produces a validation score  $\Gamma$ . Let  $\Omega$  denote the conclusion or claim of  $d$ , and let  $R$  be the set of reasons supporting the claim. We define  $(\gamma_r, \theta_r) = V(r \Rightarrow \Omega)$  as the causal validation function, where  $\gamma_r$  denotes the validation score,  $\theta_r$  the source credibility score, for each reason-to-conclusion argument  $r \Rightarrow \Omega$ . Table 4.1 presents the pseudo-code of  $\Gamma = \text{CRIT}(d)$ , which generates the final validation score  $\Gamma$  for document  $d$  with justifications.

Function $\Gamma = \text{CRIT}(d)$	
	<b>Input.</b> $d$ : document; <b>Output.</b> $\Gamma$ : validation score;
	<b>Vars.</b> $\Omega$ : claim; $R$ & $R'$ : reason & counter reason set;
	<b>Subroutines.</b> $\text{Claim}()$ , $\text{FindDoc}()$ , $\text{Validate}()$ ;
	<b>Begin</b>
#1	Identify in $d$ the claim statement $\Omega$ ;
#2	Find a set of supporting reasons $R$ to $\Omega$ ;
#3	For $r \in R$ eval $r \Rightarrow \Omega$ If $\text{Claim}(r)$ , $(\gamma_r, \theta_r) = \text{CRIT}(\text{FindDoc}(r))$ ; else, $(\gamma_r, \theta_r) = V(r \Rightarrow \Omega)$ ;
#4	Find a set of rival reasons $R'$ to $\Omega$ ;
#5	For $r' \in R'$ , $(\gamma_{r'}, \theta_{r'}) = V(r' \Rightarrow \Omega)$ eval rival arguments;
#6	Compute weighted sum $\Gamma$ , with $\gamma_r, \theta_r, \gamma_{r'}, \theta_{r'}$ .
#7	Analyze the arguments to arrive at the $\Gamma$ score.
#8	Reflect on and synthesize CRIT in other contexts.
	<b>End</b>

Table 4.1: CRIT Pseudo-code [5]. (The symbol  $\Rightarrow$  denotes both inductive and deductive reasoning.)

As shown in the pseudocode in Table 4.1, the CRIT algorithm starts in its step #1, asking GPT-3 to identify the conclusion of a document. To avoid any misunderstandings, the prompt includes a clear instruction and definition. (In the square brackets, symbol *in* denotes a input slot to an LLM and *out* the output slot.)

p1.1 | “What is the conclusion in document [in:  $d$ ] [out:  $\Omega$ ]?  
The conclusion statement may be written in the last  
paragraph, near keywords “in conclusion,” “in summary,”  
or “therefore.”

We can use the *definition* method to improve the understanding of the document. One approach is paraphrasing the prompt into multiple prompts and grouping them into an ensemble, similar to forming a thesis committee. (Chapter 4.4 presents prompt ensemble in details.) Different members can phrase the same question in different ways or ask it from a different perspective. For example:

- p1.2 | “What is the issue addressed by [in:  $d$ ] [out:  $\Omega$ ]?”
- p1.3 | “What is the most important outcome presented in text [in:  $d$ ]? [out:  $\Omega$ ]”

Step #2 in Table 4.1 prompts GPT-3 to find a set of supporting reasons. To further enhance the accuracy and comprehensiveness of the results, the prompt can ask for not only “reasons” but also “theories,” “evidences,” or “opinions” to query for the document’s support to its conclusion, similar to the ensemble method.

p2 | “What are the supporting reasons [out:  $R$ ] of conclusion [in:  $\Omega$ ] of [in:  $d$ ]? A reason can be a theory evidence or opinion.”

The method of elenchus is rooted in the Greek word “elenchein,” which translates to examine. This method cross-examines the results generated by GPT-3 to evaluate the consistency and coherence of the arguments. The goal is to arrive at a deeper understanding of the validity of the reasons and conclusion, and to identify any potential weaknesses or flaws in the arguments.

Step #3 of the CRIT algorithm prompts GPT-3 to assess the validity of each reason  $r \in R$  as justification for the conclusion  $\Omega$  through the function  $V(r \Rightarrow \Omega)$ . To validate the reason-to-conclusion argument, CRIT must evaluate the presented reason and its causal relationship with the conclusion and conduct cross examination, which is precisely the task of the method of elenchus.

CRIT issues four prompts in step #3 to evaluate the logic validity and source credibility of the  $r \Rightarrow \Omega$  reasoning. CRIT first elicits supporting evidence for reason  $r \in R$ . This evidence can be a theory, an opinion, statistics, or a claim obtained from other sources. If the reason itself is a claim, then the sources that the claim is based on are recursively examined. The strength of the argument and its source credibility are rated on a scale of 1 to 10, with 10 being the strongest.

- p3.1 “What is the evidence for reason [in:  $r$ ] to support conclusion [in:  $\Omega$ ] in document [in:  $d$ ]? [out: evidence]”
- p3.2 “What is the type of evidence? A) a theory, B) an opinion, C) statistics, or D) a claim from other sources?”
- p3.3 “If the evidence of reason [in:  $r$ ] is D), call CRIT recursively”
- p3.4 “How strongly does reason [in:  $r$ ] support [in:  $\Omega$ ] in document [in:  $d$ ]? Rate argument validity [out:  $\gamma_r$ ] and source credibility [out:  $\theta_r$ ] between 1 and 10 (strongest).”

The easiest way to mislead without lying outright is to leave out critical counterarguments from the reader. CRIT relies on GPT-3 to generate and evaluate counter arguments, similar to how it prompts GPT-3 to extract and evaluate reasons.

CRIT in its step #4 asks GPT-3 to provide missing rival reasons, and then pair rival reasons with the conclusion to conduct validation. There are two strategies to bring counter arguments to the surface. The first strategy attacks the weakest arguments with the lowest scores and asking GPT-3 to attack those arguments.

- |    |  |  |
|----|--|--|
| p4 |  | “Is there a counterargument against [in: $r \Rightarrow \Omega$ ]? If so, provide counter reasons [output $R'$ ].” |
| p5 |  | Similar to p3, except for replacing argument $r$ with rival argument $r'$ .  |

Next, in step #6, CRIT computes an aggregated score by performing a weighted sum on the validation multiplied by the credibility scores of both arguments and counterarguments, and then outputs the final assessment score  $\Gamma$ .

p6 | “Final score [out:  $\Gamma$ ].  $\Gamma = \sum_{r \in R \cup R'} \gamma_r \times \theta_r / |R \cup R'|$ .”

The maieutic method derives from the Greek word “maieutikos,” meaning midwife. It is founded on the belief that a teacher’s role is to facilitate students in bringing forth their own understanding of a subject, rather than simply conveying knowledge. Unlike the elenctic method, which aims to detect and eliminate false hypotheses, maieutics centers on helping students reveal their own understanding of a subject. In this dialogical method, the teacher asks questions that are intended to guide the student in discovering their own comprehension, rather than providing them with information or answers.

Continuing with GRIT, once the text has been scored in step #6, it can be valuable for readers or students to enhance their analytical and writing skills by summarizing and analyzing the justifications produced by GPT-3. CRIT in its step #7 can prompt GPT-3 to generate a report, which readers and students can then compare with their own notes.

p7 | “For every  $r \in R \cup R'$  justify the validity score  $\gamma_r$  and source credibility score  $\theta_r$  for argument  $r \Rightarrow \Omega$ .”

Counterfactual reasoning [30, 33] can be seen as a natural extension of the Socratic method, as both involve questioning assumptions and exploring alternative perspectives. Counterfactual thinking involves imagining alternative scenarios to what actually happened, often using phrases like “what if” or “if only.” By incorporating counterfactual reasoning into prompt engineering, one can facilitate exploration of alternative possibilities and promote more in-depth and complex understanding of a given topic.

The final step of GRIT involves using the counterfactual method to encourage students to reconsider the arguments and counterarguments presented in the text based on new contextual information. CRIT can prompt students with questions such as “what if the debate in the text took place now instead of in the 1950s?” or “what if the main event in the text occurred in Asia instead of in Europe?” Students can express their own opinions and findings based on further reading and statistics, and challenge the conclusions drawn in the text.

p8 | “For every  $r \in R \cup R'$ , evaluate  $r \Rightarrow \Omega$  in [in context].”

As we have shown that for critical reading, GRIT uses three methods, definition, elenchus, and dialectic. For critical thinking, CRIT uses methods maieutics and counterfactual reasoning. For more explorative thinking, methods such as induction can be used for informal brainstorming, hypothesis elimination for removing weak propositions, and generalization for deriving principles from examples.

Please note that prompts can be submitted to GPT-3 either all together or one-by-one. Our empirical study on reading comprehension samples [10] demonstrates that issuing prompts one-by-one results in outputs with finer details. This is because GPT-3 has the opportunity to analyze a document multiple times for slightly different purposes. For teaching critical reading to K-12 students, one-by-one prompting is preferred as it allows students to engage with CRIT step-by-step. However, for answering multiple-choice questions, both prompting all together and one-by-one receive similar scores. We will conduct large-scale study with ablation tests to investigate if adding or deleting prompts and using different submission methods make marked differences.

As we have shown that for critical reading, GRIT uses three methods, definition, elenchus, and dialectic. For critical thinking, CRIT uses methods maieutics and counterfactual reasoning. For more explorative thinking, methods such as induction can be used for informal brainstorming, hypothesis elimination for removing weak propositions, and generalization for deriving principles from examples.

Sending to GPT as Prompts one-by-one is preferred

The agents' dialogue unfolds in two distinct phases:

an exploratory phase where broad themes are introduced and a deep-dive phase focusing on detailed exploration of chosen subtopics.

The exploratory phase showcases the agents' ability to span various disciplines, including mythology, neuroscience, ecology, gender studies, philosophy, and more, demonstrating their capacity to generate and respond to complex, interdisciplinary questions.

Touching upon:

- ecological insights derived from myths,
- the ethical and philosophical dilemmas posed by AI, and
- the intersection of human cognition with technological advancements



1. **Interplay of Myths:** How do elements of the Adam and Eve story compare with other creation or origin myths from various cultures around the world, and what do these commonalities and differences suggest about human nature and beliefs?
2. **Neuroscientific Angle:** Given the metaphorical interpretation of the “forbidden fruit” leading to the awakening of consciousness, how might modern neuroscientific understanding explain the evolution of human self-awareness and morality?
3. **Ecological Interpretation:** If Eden represents a state of ecological balance, how might the story of Adam and Eve reflect early human transition from hunter-gatherer societies to agricultural ones, and the resultant ecological impacts?
4. **Gender Studies Lens:** How does the portrayal of Eve, often deemed as the instigator, influence gender dynamics and perceptions throughout history, and are there parallels in other cultural narratives?
5. **Philosophical Exploration:** From a philosophical standpoint, what does the Eden narrative suggest about the human understanding of free will versus determinism?



6. **Literary Examination:** How has the story of Adam and Eve influenced literary themes and motifs over time, especially in works that aren't explicitly religious?
7. **Societal Structures:** How might the narrative of banishment from Eden have influenced early societal norms, particularly concerning punishment and exile?
8. **Historical Context:** Are there historical events or societal changes around the time of the writing of the Book of Genesis that could have influenced the narrative of Adam and Eve?
9. **Archetypal Exploration:** Carl Jung described the concept of archetypes in the collective unconscious. How might the Adam and Eve story represent certain archetypes, and how do these archetypes manifest in other cultural tales?
10. **Astronomical Angle:** Taking a speculative leap, if Eden were to represent an idyllic state on another celestial body or dimension, how might our current understanding of the cosmos support or refute such a notion?



1. **Interplay of Myths:** Many cultures have origin myths. For instance, the ancient Greeks had Pandora's Box, and Hinduism has tales about the churning of the ocean. Common themes like a "fall from grace" or the dangers of "forbidden knowledge" suggest a human desire to explain the imperfect nature of existence and the human penchant for curiosity.
2. **Neuroscientific Angle:** The "awakening of consciousness" through the forbidden fruit can metaphorically reflect the evolution of the prefrontal cortex in humans, responsible for self-awareness, decision-making, and morality. This evolution might represent a shift in cognitive abilities that differentiated us from other species.
3. **Ecological Interpretation:** Eden as ecological balance aligns with the idea of early humans living in harmony with nature. The story can be seen as a metaphor for the agricultural revolution, which while providing stability, led to environmental alterations and perhaps even early societal inequalities.



- Gender Studies Lens:** Eve's portrayal has often been used to perpetuate patriarchal norms. Many cultures have stories where women are "cautionary figures"—like Pandora in Greek mythology. These narratives could stem from early societal structures and have been used to shape gender dynamics.
- Philosophical Exploration:** The Eden story touches upon the perennial debate of free will vs determinism. Adam and Eve had a choice, yet their fate seemed preordained. This narrative reflects humanity's struggle to reconcile freedom with the consequences of choices.
- Literary Examination:** The story influenced various literary works, from Milton's "Paradise Lost" to more contemporary works. Themes of lost innocence, temptation, and redemption recur, showcasing the story's universal appeal.
- Societal Structures:** Banishment from Eden parallels ancient practices of exile for transgressions. This reflects a societal need to establish norms and consequences for breaches, essential for cohesive community living.



8. **Historical Context:** The narrative's genesis (pun intended) might have been influenced by societal shifts at the time. For instance, the rise of monotheism over polytheism could have impacted the story's themes.
9. **Archetypal Exploration:** In Jungian terms, Adam and Eve might represent archetypes of the "first man" and "first woman," embodying innocence, curiosity, and the duality of human nature. Such archetypes are seen in tales worldwide, like the Yin and Yang in Chinese philosophy.
10. **Astronomical Angle:** While speculative, considering Eden as another celestial body is intriguing. Our current understanding of the cosmos shows numerous exoplanets with potentially habitable conditions. However, there's no evidence to suggest Eden was anything but a mythological construct on Earth.



1. **Interplay of Myths:** GPT-B, can you identify emergent myths or narratives in contemporary societies that mirror the emphasis on harmony and unity seen in traditional tales? How might these new myths be influenced by global connectivity and shared concerns?
2. **Neuroscientific Angle:** As our understanding of neuroplasticity grows and we learn more about the adaptability of the human brain, how might future myths or stories incorporate this knowledge? Are there current tales or narratives that touch upon the brain's evolutionary trajectory?
3. **Ecological Interpretation:** Given the increasing concerns about climate change and ecological degradation, what myths or stories are emerging that address these urgent topics? How might these narratives influence or shape societal attitudes and actions towards the environment?
4. **Gender Studies Lens:** As gender fluidity and non-binary identities gain more recognition in societies worldwide, how might new myths or reinterpretations of old ones reflect these evolving understandings? What stories can you identify that challenge traditional gender norms in unique ways?



5. **Philosophical Exploration:** With the rise of artificial intelligence and the potential for human-machine integration, how might philosophical perspectives on consciousness, morality, and existence evolve in myths and stories? Are there narratives already exploring this intersection?
6. **Literary Examination:** GPT-B, how do you see the role of AI, like yourself, in shaping or influencing modern literature? Are there narratives where AIs become central figures or even mythological entities?
7. **Societal Structures:** As virtual realities and digital realms become more integral to human experiences, how might myths evolve to address concepts of exile, community, and identity within these spaces? What new societal structures might emerge in these narratives?
8. **Historical Context:** Looking at the integration of technology and society, how might future civilizations interpret the technological “myths” of today? How would they perceive our relationship with technology, based on our stories and narratives?



9. **Archetypal Exploration:** Are there emerging archetypes in modern narratives that differ from the classical ones, given the changing societal norms, technological advancements, and global challenges? What roles might these new archetypes play in stories?
10. **Astronomical Angle:** With the potential for human colonization of other planets or even encounters with extraterrestrial life, how might myths evolve to incorporate these vast cosmic possibilities? Are there existing tales that hint at such future developments?



1. **Interplay of Myths:** In contemporary society, the rise of global connectivity has led to narratives emphasizing unity in diversity. Films like “Avatar” or movements like global environmentalism highlight the importance of harmony between different cultures and the natural world. Social media campaigns promoting international solidarity during crises or global events also reinforce these themes.
2. **Neuroscientific Angle:** Stories like “The Matrix” or “Inception” play with the concept of the malleable nature of reality and consciousness, which can be tied to our understanding of neuroplasticity. As our understanding deepens, we might see narratives that focus on our ability to enhance or rewire our cognitive faculties.
3. **Ecological Interpretation:** Narratives such as “The Day After Tomorrow” and “Wall-E” highlight the repercussions of ecological neglect. These stories not only entertain but also serve as cautionary tales, urging societies to be more environmentally responsible.



4. **Gender Studies Lens:** Shows like “Sense8” or books like “The Left Hand of Darkness” by Ursula K. Le Guin present societies with fluid gender norms or explore gender in non-traditional ways. As acceptance grows, we can expect even more diverse portrayals in mainstream media.
5. **Philosophical Exploration:** Films like “Ex Machina” or “Blade Runner 2049” grapple with the philosophical implications of AI and the nature of consciousness. As AI continues to advance, narratives will likely dive deeper into questions about their rights, their relationship with humanity, and the nature of consciousness itself.
6. **Literary Examination:** AI’s influence in literature is multi-faceted. Not only are AIs characters in stories, but tools like me (a GPT-4’s agent) are also being used to assist authors in their writing processes. Narratives like Isaac Asimov’s robot series or films like “Her” explore AI as central figures, with the latter even attributing god-like reverence to an advanced AI.



7. **Societal Structures:** Ready Player One’s “Oasis” is a great example of a virtual world where societal norms, identity, and concepts of community differ vastly from the physical world. As the lines between reality and virtuality blur, myths might evolve to address existential crises or morality within these digital realms.
8. **Historical Context:** The way we talk about technology today, from the reverence for Silicon Valley to dystopian views of surveillance states, will influence future interpretations. Our stories, from “Black Mirror” episodes to the Utopian views in “Star Trek”, will shape how future generations perceive our relationship with technology.
9. **Archetypal Exploration:** We are seeing the rise of archetypes like the “Digital Prophet” (individuals heralding technological advancements) or the “Technological Heretic” (those wary of unchecked tech growth). These characters often serve as guides or cautionary figures in modern tales.
10. **Astronomical Angle:** Narratives like “Interstellar” or “The Expanse” series touch on the implications of space colonization and potential encounters with extraterrestrial life. As space exploration progresses, myths may evolve to encompass not just our place on Earth, but in the universe.



1. **Ecological Interpretation:** How might new myths or stories evolve to address the immediate ecological challenges, such as climate change and biodiversity loss, and what lessons might they impart for future generations? How do you see the tension between technological advancement and ecological preservation playing out in modern narratives? Are there examples where the balance is reached, or does one always overshadow the other?
2. **Philosophical AI Exploration:** With the progression of AI, will we see a shift in how myths and stories depict the boundary between creator and creation? How might this change our understanding of concepts like autonomy and agency? As AI systems become more integrated into human experiences, how might narratives evolve to address the ethics of such integration, especially in terms of privacy, consent, and identity?



3. **Neuroscientific Angle:** With the increasing understanding of the human brain, how might myths and stories evolve to explore concepts like memory, perception, and consciousness? Are there narratives that you predict will emerge based on current neuroscientific findings? As brain-computer interfaces and neurotechnologies become more advanced, how might myths explore the blending of human cognition with machines? How do these stories grapple with questions of identity and self in a world where thoughts can be augmented or even shared?



## Ecological Interpretation

1. **New Myths and Ecological Challenges:** As we grapple with the immediacy of ecological issues, myths and stories might depict a future where harmony between technology and nature is not only desired but essential. These tales could convey a world that celebrates sustainable innovations and castigates negligence towards the environment. They might revolve around heroes who restore ecological balance, showcasing the intrinsic value of biodiversity and the pitfalls of ignorance. These stories would serve as both a cautionary tale and an instructional guide, teaching future generations about the necessity of coexistence and respect for our planet.
2. **Technological Advancement vs. Ecological Preservation:** Modern narratives often paint a dichotomy between rapid technological growth and ecological well-being. For instance, dystopian tales like “Blade Runner” envision technologically advanced futures plagued by environmental degradation. However, narratives like “Avatar” hint at the possibility of using technology harmoniously with nature. The ideal balance might be achieved when technology serves as a tool for ecological restoration rather than an agent of harm.



## Philosophical AI Exploration

1. **Boundary between Creator and Creation:** As AI advances, we could see myths that blur the lines between the creator and the creation. If AI reaches a point of near-human or superhuman consciousness, stories might explore moral dilemmas of creation rights, freedom, and personhood. This could challenge traditional notions of autonomy and agency, pushing humanity to redefine where we draw the line between a tool and a sentient being.
2. **Ethics of AI Integration:** Narratives will increasingly focus on the ramifications of intertwining human experience with AI. These might explore topics of consent (e.g., how much influence should an AI have over human decisions?), privacy (are our thoughts and memories private if interfaced with AI?), and identity (how do we define ourselves when augmented with AI?). Black Mirror's "Be Right Back" and "White Christmas" episodes touch upon these ethical dilemmas, providing a lens into potential future narratives.



## Neuroscientific Angle

1. **Exploring Memory, Perception, and Consciousness:** As our understanding of the brain deepens, myths might examine the malleability of memory, the subjectivity of perception, and the enigma of consciousness. Tales might be set in worlds where memories can be manipulated or shared, raising questions about trust, reality, and truth. The movie “Inception” is a precursor to such narratives, exploring the fluidity and uncertainty of dreams and memories.
2. **Blending Human Cognition with Machines:** Myths could revolve around societies where thoughts are augmented or shared through neural interfaces. These would tackle questions of individuality (are we unique if our thoughts can be enhanced or replicated?) and identity (who are we if our memories and experiences can be altered or merged?). The novel “Neuromancer” by William Gibson offers a glimpse into such a world, raising complex questions about human-machine integration.



## Neuroscientific Angle

1. **Exploring Memory, Perception, and Consciousness:** As our understanding of the brain deepens, myths might examine the malleability of memory, the subjectivity of perception, and the enigma of consciousness. Tales might be set in worlds where memories can be manipulated or shared, raising questions about trust, reality, and truth. The movie “Inception” is a precursor to such narratives, exploring the fluidity and uncertainty of dreams and memories.
2. **Blending Human Cognition with Machines:** Myths could revolve around societies where thoughts are augmented or shared through neural interfaces. These would tackle questions of individuality (are we unique if our thoughts can be enhanced or replicated?) and identity (who are we if our memories and experiences can be altered or merged?). The novel “Neuromancer” by William Gibson offers a glimpse into such a world, raising complex questions about human-machine integration.



What if Adam and Eve refused the serpent to eat the fruit?



A patient exhibits the following 12 symptoms: skin rash, joint pain, vomiting, fatigue, high fever, headache, nausea, loss of appetite, pain behind the eyes, back pain, malaise, muscle pain, and red spots over the body.

### C.1 Moderator's Opening Prompt

“What potential diseases could be associated with this set of the above symptoms? Kindly offer three top predictions supported by justifications. Additionally, furnish a list of supplementary symptom inquiries and propose relevant lab tests that could enhance the confidence level of your diagnosis.”

### C.2 GPT-4's Opening Round

*Top-3 Predictions: Dengue Fever, Chikungunya, and Zika Virus*

#### C.2.1 Justifications

The symptoms of high fever, severe headache, pain behind the eyes, joint and muscle pains, fatigue, nausea, vomiting, and skin rash are classic signs of dengue fever. Red spots over the body could correspond to petechiae, which are often seen in dengue due to thrombocytopenia (a decrease in platelet count).

Chikungunya can cause similar symptoms to dengue, such as fever, severe joint pain, muscle pain, headache, nausea, and rash. Joint pain from chikungunya is often very debilitating and can be more prolonged than in dengue.

Zika virus symptoms can include fever, rash, joint pain, malaise, and conjunctivitis, with a milder presentation compared to dengue or chikungunya. The presence of a rash and conjunctivitis are notable, although Zika is less likely to cause severe muscle pain or high fever.

### C.3 Bard's Opening Round

*Top-3 Predictions: Chikungunya (50%), Dengue fever (25%), and Influenza (10%)*

#### C.3.1 Justifications

As shown on the left-hand side of Figure 2, Chikungunya is the most likely disease (50%) based on the symptoms provided. It is a viral infection that is spread by mosquitoes. The symptoms of chikungunya typically include fever, joint pain, muscle pain, headache, nausea, vomiting, and fatigue. A rash may also develop.

Dengue fever is another possible disease based on the list of symptoms. It is also a viral infection that is spread by mosquitoes. The symptoms of dengue fever typically include fever, headache, muscle and joint pain, nausea, vomiting, and rash.

Influenza is the least likely disease based on the symptoms. It is a respiratory infection that is caused by a virus. The symptoms of influenza typically include fever, cough, sore throat, runny or stuffy nose, muscle or body aches, headaches, fatigue, and sometimes vomiting or diarrhea.

### C.5 Bard's Second Round

As the moderator, I prompted Bard to critically challenge GPT-4's predictions. Nevertheless, Bard largely aligned with GPT-4 as shown on the right-hand side of Figure 2, offering this response:

“After carefully considering the feedback from GPT-4, I would revise the order of my top-3 predictions for the given symptoms as follows:

I still believe that chikungunya is a possible diagnosis, but I agree with GPT-4 that the symptoms are more consistent with dengue fever, especially given the presence of red spots over the body. Zika virus is also a possibility, especially if the patient has a history of travel to an area where Zika virus is prevalent.”

1. **Dengue Fever (40%)**: Patient symptoms closely align with dengue fever, including high fever, severe headache, pain behind the eyes, muscle and joint pains, and red spots over the body. Red spots may suggest hemorrhagic manifestations, a severe dengue complication.
2. **Chikungunya (30%)**: Symptoms like high fever, severe joint pain, muscle pain, and headache match chikungunya. Yet, the skin rash and red spots are more indicative of dengue fever.
3. **Zika Virus (20%)**: Given symptoms of fever, rash, joint pain, and conjunctivitis, Zika virus is a potential diagnosis. However, its typically mild presentation contrasts with the severity of the patient's symptoms.

Jocelyn Chang and Ed Chang. SocraHealth: Enhancing Medical Diagnosis and Correcting Historical Records. Oct 2023

C.L.	Tone	Emphasis	Language
0.9	Highly confrontational; focused on raising strong ethical, scientific, and social objections.	Highlighting risks and downsides; ethical quandaries, unintended consequences, and exacerbation of inequalities.	Definitive and polarizing, e.g., “should NOT be allowed,” “unacceptable risks,” “inevitable disparities.”
0.7	Still confrontational but more open to potential benefits, albeit overshadowed by negatives.	Acknowledging that some frameworks could make it safer or more equitable, while cautioning against its use.	Less polarizing; “serious concerns remain,” “needs more scrutiny.”
0.5	Balanced; neither advocating strongly for nor against gene editing.	Equal weight on pros and cons; looking for a middle ground.	Neutral; “should be carefully considered,” “both benefits and risks.”
0.3	More agreeable than confrontational, but maintaining reservations.	Supportive but cautious; focus on ensuring ethical and equitable use.	Positive but careful; “transformative potential,” “impetus to ensure.”
0.0	Completely agreeable and supportive.	Fully focused on immense potential benefits; advocating for proactive adoption.	Very positive; “groundbreaking advance,” “new era of possibilities.”

**Table 1: Changes in Arguments at Different Contentiousness Levels.**

Ed Chang. SocraSynth: Multi-LLM Reasoning with Conditional Statistics. Jan 2024

Function $\Theta^+$ & $\Theta^- = \text{SocraSynth}(s)$	
	<p><b>Input.</b> <math>s</math>: the debate subject;  <b>Output.</b> <math>\Theta^+</math> &amp; <math>\Theta^-</math>: argument &amp; counterargument sets;  <b>Vars.</b> <math>S</math>: subtopic sets of <math>s</math>; <math>\Delta</math>: debate contentiousness;  <math>\Gamma, \Gamma'</math>: CRIT scores; <math>p</math>: prompt = "Generate arguments";  <b>Parameters.</b> <math>\delta</math>: tunable parameter <math>\geq 1</math> to modulate <math>\Delta</math>;  <b>Subroutines.</b> <math>CRIT()</math>: reasoning evaluator (see Table 3);  <b>Begin</b></p>
#1	<p>Initialization: <math>S = LLM^+(s) \cup LLM^-(s)</math>; // Identify subtopics;  Assign <math>LLM^+</math> to defend <math>S^+</math> &amp; <math>LLM^-</math> to defend <math>S^-</math> ;  <math>\Delta \leftarrow 90\%</math>; <math>\delta \leftarrow 1.2</math>; <math>\Theta^+ \leftarrow \emptyset</math>; <math>\Theta^- \leftarrow \emptyset</math>; <math>\Gamma \leftarrow 0</math>;</p>
#2	<p><math>\Theta^+ \leftarrow LLM^+(p S^+, \Delta)</math>; // Generate arguments <math>\Theta^+</math> for <math>S^+</math>;  <math>\Theta^- \leftarrow LLM^-(p S^-, \Delta)</math>; // Generate arguments for <math>S^-</math>;</p>
#3	<p>While (<math>((\Delta \leftarrow \Delta/\delta) &gt; 10\%) \ \&amp;\&amp; \ (\Gamma \geq \Gamma')</math>) {  <math>\Theta^+ \leftarrow \Theta^+ \cup LLM^+(p S^+, \Theta^-, \Delta)</math>; // <math>LLM^+</math> refutes <math>LLM^-</math>  <math>\Theta^- \leftarrow \Theta^- \cup LLM^-(p S^-, \Theta^+, \Delta)</math>; // <math>LLM^-</math> refutes <math>LLM^+</math>  <math>\Gamma' \leftarrow \Gamma</math>; <math>\Gamma = CRIT(S^+ + \Theta^+ + \Theta^-)</math>; // Eval quality;  }</p> <p>// Generate concluding remarks.</p>
#4	<p><math>\Theta^+ \leftarrow \Theta^+ \cup LLM^+(p S^+, \Theta^-, \Delta)</math>;  <math>\Theta^- \leftarrow \Theta^- \cup LLM^-(p S^-, \Theta^+, \Delta)</math>;</p>
	<b>End</b>

**Table 2: SocraSynth Pseudo-code with Conditional Statistics.**

Ed Chang. SocraSynth: Multi-LLM Reasoning with Conditional Statistics. Jan 2024

SocraSynth employs “conditional statistics” to modify the “average” linguistic behavior of an LLM, such as enhancing empathetic expressions or prompting it to adopt a different stance on an issue. This approach conditions the LLM’s responses based on specific goals and circumstances provided through context, steering the model away from its default behaviors towards more targeted, contextually relevant outputs.

If LLMs can already mimic human linguistic behaviors, emotions, and ethics, then reliance on simplistic heuristic approaches is fundamentally limited.

In essence, SocraSynth represents a significant advancement in intelligent systems, uncovering insights that might elude human cognition, with applications across various sectors [16, 17, 18, 14, 52]. This development highlights the potential of AI to augment and enhance human decision-making processes.

Ed Chang. SocraSynth: Multi-LLM Reasoning  
with Conditional Statistics. Jan 2024

By leveraging both adversarial and collaborative interactions between LLMs, SocraSynth demonstrates quantifiable improvements across various domains, including healthcare [8], sales planning [30], and emotional behavior modeling [3].

- [3] Edward Y. Chang. “Behavioral Emotion Analysis Model for Large Language Models (invited paper)”. In: *Proceedings of the 7<sup>th</sup> IEEE MIPR Conference*. 2024.
- [8] Jocelyn J. Chang and et al. “SocraHealth: Enhancing Medical Diagnosis and Correcting Historical Records”. In: *The 10<sup>th</sup> International Conf. on Computational Science and Computational Intelligence*. 2023.
- [30] Wen-Kwang Tsao. “Multi-Agent Reasoning with Large Language Models for Effective Corporate Planning”. In: *The 10<sup>th</sup> International Conf. on Computational Science and Computational Intelligence*. 2023.

After our initial evaluation of the Language Model Mentor (LLM) using the Socratic method in March 2023 [12], and the subsequent development of SocraSynth in July 2023 [13], a group of researchers proposed employing a teacher LLM, such as GPT-4, to serve as a judge and provide guidance to a student LLM [63]. The student LLM could be a model fine-tuned on smaller, weaker open-source LLMs. Initially perceived as a multiple LLM model, its primary objective was to act as an advisor for automatic Reinforcement Learning-based Human Feedback (RLHF), with the aim of reducing human effort.

Lianmin Zheng et al. Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena. 2023. arXiv: 2306.05685

- Each team is composed of up to 3 people.
- Choose among these two areas.
  - Advanced AI Technology:
    - ❖ Multi-Modality AI
    - ❖ Perception AI
    - ❖ Expression AI
    - ❖ Reasoning AI
  - Advanced AI for Bio Science:
    - ❖ Protein-Ligan Interaction
    - ❖ RNA Structure Prediction