

Philosopher Game

April 28, 2023

Abstract

The philosopher game is a multiplayer game where players create counterfactual histories of major figures and the outcomes are scored by AI. It is typically played by setting a person, a time, and a book that is read by that person at that time.

1 Round 1. Positive Results of the Philosopher's Game for Elon Musk

In a internet video Elon tells that his life philosophy was formed by reading Douglas Adams' Hitchhiker Guide to the Galaxy after reading Nietzsche and Schopenhauer. The counterfactual version examines the impact of reading X book in between the two life events. The impact is scored based on the 180 page Mar Settlement Master plan written by Joel Dietz and ChatGPT

Table 1: Positive Results for Elon Musk

| Book Title | Life Change | Mars Mission Score (%) |
|------------------------------------|------------------------------|------------------------|
| Kim Stanley Robinson, Red Mars | Mars colonization motivation | 12 |
| Hegel's Science of Logic | Accelerated Neuralink | 8 |
| Arthur C. Clarke's Childhood's End | Space advocacy | 6 |
| Edward de Bono, Lateral Thinking | Enhanced problem-solving | 4 |
| Le Guin, The Dispossessed | Utopian ideas | 2 |
| Gandhi's autobiography | Non-violent leadership | 1 |

2 Round 2: Results of the Philosopher's Game for Sam Altman

The second round was played with Sam Altman and the question what if he was reading this book as he created OpenAI.

| Table 2: Ranked Results for Sam Altman | | | |
|--|--------------------------|------------------------|--------------------|
| Book Title | Life Change | Mars Mission Score (%) | Humanity Score (%) |
| C, Finite and Infinite Games | Endless growth | 22 | 88 |
| Dennett, Elbow Room | Enhanced decision-making | 18 | 81 |
| Plato’s Republic | Just leadership | 10 | 78 |
| Heart of Darkness | Cautionary tale | -2 | 45 |

3 Round 3: Negative Results of the Philosopher’s Game for Elon Musk

The third round was played in person by a group of cryptocurrency enthusiasts in Puerto Rico who had the idea of playing a round what would have the largest negative impact on Elon.

| Table 3: Ranked Negative Results for Elon Musk | | |
|--|--------------------------------|------------------------|
| Book Title | Life Change | Mars Mission Score (%) |
| Mein Kampf | Distorted worldview | -35 |
| JJ, How to make love like a porn star | Distracted | -30 |
| Marx, Communist Manifesto | Focus on wealth redistribution | -20 |
| Rhonda Byrne, The Secret | Pseudo-scientific beliefs | -10 |

4 Additional examples of GPT-4’s counterfactual logic

4.1 Elon

After reading Kim Stanley Robinson’s ”Red Mars,” Elon Musk becomes even more passionate about the idea of Mars colonization. The quote ”We came for the air and the water, but what we found was worth the trip,” resonates with him and drives his dedication to creating a self-sustaining human presence on Mars.

In this alternate reality, Elon Musk uses the vivid descriptions and creative ideas found in ”Red Mars” as a source of inspiration for his Mars mission. He works closely with his team at SpaceX to develop realistic and sustainable strategies for colonization. As a result, the Mars mission experiences a 10

Please note that these are fictional scenarios and not actual predictions of how these readings would have influenced Elon Musk in reality.

4.2 Altman

In this alternate history, Sam Altman reads Plato's "Republic" while creating OpenAI. Plato's quote, "Until philosophers rule as kings or those who are now called kings and leading men genuinely and adequately philosophize, that is, until political power and philosophy entirely coincide... cities will have no rest from evils," deeply resonates with Altman and influences his perspective on leadership and the development of artificial intelligence.

Inspired by the idea of philosopher-kings, Altman places a strong emphasis on the importance of wisdom, ethics, and philosophical understanding in the leadership of OpenAI. He actively seeks out and involves philosophers, ethicists, and interdisciplinary thinkers in shaping the organization's vision, values, and research directions. In this timeline, OpenAI becomes a pioneer in the field of AI ethics, with a 20

4.3 Elon negative

After reading "Mein Kampf," Elon Musk is deeply disturbed by the book's hateful ideology and extreme beliefs. A quote that exemplifies the disturbing content is: "The receptivity of the great masses is very limited, their intelligence is small, but their power of forgetting is enormous."

Rather than being influenced by the book, Musk becomes more determined than ever to work towards a future that fosters unity, cooperation, and understanding among people. This strengthens his resolve to establish a Mars settlement as a symbol of humanity's collective progress. As a result, there is a +2

5 Future plans

Intended future rounds likely include Jeff Bezos and other figures highlighted by ChatGPT in the Mars plan. Additionally as other advanced AIs come together their results can be averaged. Additionally the AIs can be graded by their results (currently at the level of an average university student).

6 Jargon Pseudocode

Jargon (<https://github.com/jbrukh/gpt-jargon>) is a language custom created for working with ChatGPT and other LLMs.

```
+++ philosopher-game($person, $book)
-- Find a relevant quote from the author of $book
-- Find out how the life of $person would have been
   influenced by reading $book
```

```
-- Generate a two-paragraph summary:
-- In the first paragraph, include the relevant quote
  from the author
-- In the second paragraph, describe the very specific
  person's life after reading $book
-- Score the outcome based on how much the modified life
  furthered the Mars mission
/output The modified life summary and score
+++
```

7 Conclusions

Counterfactuals are used as a philosophical exercise in many cases and is a hard problem that is worthy of an advanced AI. Additionally this game is intellectually engaging and fun to play for the participants. It may also generate reading lists and instructions useful for getting humanity to the stars. One major caveat, ChatGPT does not always remember things correctly and regenerates its answers with different results.