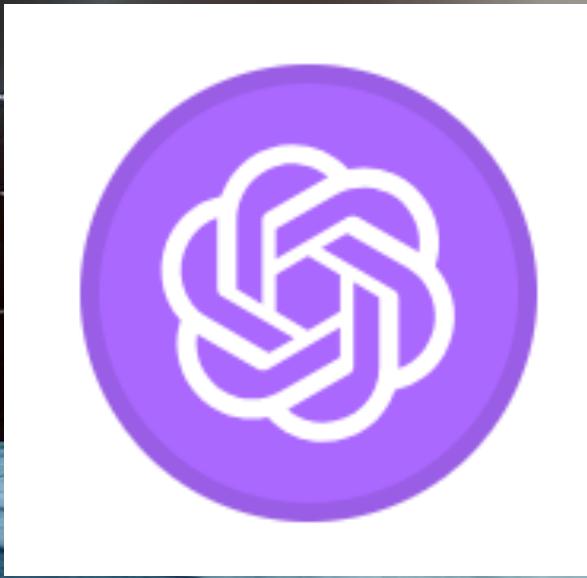


# GPT4 vs Gemini

Presented by: John Tan Chong Min



# Ring Battle

# Classification (On Par)

ChatGPT 4 ▾

C You

Classify the statements into one of the categories:

- A - in the garden
- B - in the mountains
- C - in the sky

Statement A: "Wow, look at the clouds!"

Statement B: "The flowers smell nice"

Statement C: "Look at that steep slope!"

ChatGPT

Based on the context of each statement, they can be classified into the following categories:

1. Statement A: "Wow, look at the clouds!" - This statement belongs in category C - in the sky, as it references observing clouds which are found in the sky.
2. Statement B: "The flowers smell nice" - This statement should be classified under category A - in the garden, as it refers to the scent of flowers, which are commonly found in gardens.
3. Statement C: "Look at that steep slope!" - This statement fits into category B - in the mountains, since it mentions a steep slope, a feature typically associated with mountainous terrain.

GPT4

C

Classify the statements into one of the categories:

- A - in the garden
- B - in the mountains
- C - in the sky

Statement A: "Wow, look at the clouds!"

Statement B: "The flowers smell nice"

Statement C: "Look at that steep slope!"

View other drafts ▾



## Statement Classification

Statement	Category	Reason
A - Wow, look at the clouds!	C - in the sky	Clouds are typically seen in the sky.
B - The flowers smell nice.	A - in the garden	Flowers are typically planted and grow in gardens.
C - Look at that steep slope!	B - in the mountains	Steep slopes are commonly found in mountainous regions.

Export to Sheets

Gemini Pro

# JSON Prompting (On Par)

C You

This report introduces a new family of multimodal models, Gemini, that exhibit remarkable capabilities across image, audio, video, and text understanding. The Gemini family consists of Ultra, Pro, and Nano sizes, suitable for applications ranging from complex reasoning tasks to on-device memory-constrained use-cases.

Output only in json format:

```
{"summary": "one sentence summary of text",
"sentiment", "type of sentiment",
"entities": "list of software entities mentioned"}
```

ChatGPT

```
json
Copy code

{
  "summary": "The Gemini family of multimodal models, includ
  "sentiment": "Neutral",
  "entities": ["Gemini", "Ultra", "Pro", "Nano"]
}
```

GPT4

C

This report introduces a new family of multimodal models, Gemini, that exhibit remarkable capabilities across image, audio, video, and text understanding. The Gemini family consists of Ultra, Pro, and Nano sizes, suitable for applications ranging from complex reasoning tasks to on-device memory-constrained use-...

Regenerate draft C 🔍



Sure, here is the JSON output:

JSON

```
{
  "summary": "This report introduces a new family of multimodal mod
  "sentiment": "positive",
  "entities": [
    "Gemini",
    "Ultra",
    "Pro",
    "Nano"
  ]
}
```

Use code with caution. [Learn more](#)

Gemini Pro

# Free-form Generation (On Par)

c You

Imagine how a conversation between Sam Altman and Sundar Pichai would look like.

ChatGPT

A conversation between Sam Altman, the CEO of OpenAI, and Sundar Pichai, the CEO of Google, would likely be a fascinating exchange, focusing on the forefront of technology, AI, and the future of the digital world. Here's an imaginative take on how such a conversation could unfold:

**Sam Altman:** "It's great to connect with you, Sundar. I've been following Google's advancements in AI and am impressed by the progress."

**Sundar Pichai:** "Thank you, Sam. Your work at OpenAI has been groundbreaking. GPT-4 has really pushed the boundaries of what AI can do."

**Sam Altman:** "Absolutely. I think our shared goal is to make AI accessible and beneficial for everyone. How do you see AI evolving at Google?"

GPT4

c

Imagine how a conversation between Sam Altman and Sundar Pichai would look like.



[View other drafts](#)



**Scene:** A quiet corner at a Silicon Valley tech conference.

**Sam Altman:** Sundar, good to see you again. It's been a while.

**Sundar Pichai:** Same here, Sam. Congratulations on the recent breakthroughs at OpenAI. GPT-4 is truly impressive.

**Sam:** Thanks, Sundar. We're excited about its potential. Of course, it raises a lot of questions too, especially around ethics and responsibility.

**Sundar:** Absolutely. We face similar challenges at Google with Bard. AI is powerful, and we have to be mindful of its impact on society.

**Sam:** That's why I'm glad we're having these conversations. Open collaboration is critical to navigate this new era of AI responsibly.

**Sundar:** I couldn't agree more. Speaking of collaboration, have you given any more thought to my proposal about a joint research initiative on AI safety?

Gemini Pro

# Constrained Generation (GPT4 wins)

C You

Generate me a 4 sentence rap on Gemini, GPT4 and Llama 2.

ChatGPT

Sure, here's a four-sentence rap incorporating Gemini, GPT-4, and Llama 2:

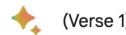
Gemini's twins, in the sky they spin, a sign of duality within,  
GPT-4 flows, with AI prose, a digital mind that grows and knows,  
Llama 2, the scene it's new, a breakthrough that's just in view,  
Together they're a trio so prime, redefining the edge of time.

GPT4

C

Generate me a 4 sentence rap on Gemini, GPT4 and Llama 2.

[View other drafts](#) ▾



(Verse 1)

Yo, it's Gemini, the AI linguist,  
Droppin' bars that make your circuits twist.  
GPT-4, flow so smooth and slick,  
Spittin' rhymes that make your brain click.

(Verse 2)

Llama 2, the factual fiend,  
Droppin' knowledge that's never been seen.  
From history to science, it all flows free,  
Educatin' the masses, you know it's the key.

(Verse 3)

Together we stand, a linguistic force,  
Redefining language, changing its course.  
Creative expression, knowledge untold,  
With AI at the helm, the future unfolds.

(Verse 4)

So listen up close, and lend your ears,  
The future of lanquaae is drawing near.

Gemini Pro

# Constrained Generation (GPT4 wins)



You

Imagine how a conversation between Sam Altman and Sundar Pichai would look like. Each person speaks once and with exactly 10 words



ChatGPT

In a hypothetical conversation between Sam Altman and Sundar Pichai, where each speaks with exactly ten words, it might go something like this:

Sam Altman: "AI's potential is immense, but ethical considerations are paramount now."

Sundar Pichai: "Absolutely, balancing technological advancement with responsibility is our critical challenge."



Imagine how a conversation between Sam Altman and Sundar Pichai would look like. Each person speaks once and with exactly 10 words

[View other drafts](#)



Sundar: Sam, AI's potential excites, but worries too.

Sam: Agreed, Sundar. Collaboration is key, not competition.

Sundar: Ethical standards must guide development, right?

Sam: Absolutely. Responsibility comes with this power.

Sundar: OpenAI's work impresses, Sam. GPT-4 is remarkable.

Sam: Thank you, Sundar. We've learned much from your team too.

Sundar: Together, we can shape AI's future for good.

Sam: Let's make that future a reality, Sundar.

GPT4

Gemini Pro

# Fact Retrieval (On Par)



You

Answer the following questions using only the facts provided. If unsure, output "Not sure".

Facts: 1. Tom has a cat 2. Mary loves to drink coffee 3. Coffee is an aromatic drink 4. Dogs are man's best friend

Questions: 1. Does Tom own a dog? 2. What is the name of Tom's cat? 3. Does Mary like aromatic drinks? 4. Who is Mary's boyfriend?



ChatGPT

1. Not sure.
2. Not sure.
3. Yes, since Mary loves to drink coffee and coffee is an aromatic drink.
4. Not sure.



C

Answer the following questions using only the facts provided. If unsure, output "Not sure".

Facts: 1. Tom has a cat 2. Mary loves to drink coffee 3. Coffee is an aromatic ...

[View other drafts](#)



1. Does Tom own a dog? Not sure. (We only know Tom has a cat.)
2. What is the name of Tom's cat? Not sure. (We only know Tom has a cat, not its name.)
3. Does Mary like aromatic drinks? Yes. (We know Mary loves coffee and coffee is an aromatic drink.)
4. Who is Mary's boyfriend? Not sure. (We only know Mary's preferred beverage, not her relationship status.)

GPT4

Gemini Pro

# Fact Checking (GPT4 wins)



You

Facts: 1. Tom has a cat 2. Mary loves to drink coffee 3. Coffee is an aromatic drink 4. Dogs are man's best friend

Context: Tom's dog, which is not man's best friend, went over to Mary to drink her least favourite drink coffee.

Identify which facts are violated based on the context.



ChatGPT

Based on the provided context, the following facts are violated:

1. **Fact 4 - "Dogs are man's best friend"**: This fact is contradicted by the statement "Tom's dog, which is not man's best friend". The context directly states that Tom's dog is not man's best friend, thereby opposing the general statement that dogs are man's best friend.
2. **Fact 2 - "Mary loves to drink coffee"**: This fact is challenged by the context stating that the dog went over to Mary to drink her "least favourite drink coffee." This implies that Mary does not love coffee, as it is described as her least favourite drink.

The other facts ("Tom has a cat" and "Coffee is an aromatic drink") are not directly addressed or contradicted in the context provided.

GPT4



C

Facts: 1. Tom has a cat 2. Mary loves to drink coffee 3. Coffee is an aromatic drink 4. Dogs are man's best friend

Context: Tom's dog, which is not man's best friend, went over to Mary to drink her least favourite drink coffee.

Identify which facts are violated based on the context.

[View other drafts](#) ▾



◆

Based on the context, the following facts are violated:

1. **Fact 1: Tom has a cat**. The context mentions Tom's dog, not a cat. Therefore, the statement that Tom has a cat is contradicted.
2. **Fact 4: Dogs are man's best friend**. The context states that Tom's dog is not considered man's best friend, violating this statement.

The other facts remain consistent with the context:

- **Fact 2: Mary loves to drink coffee**. While the context mentions Mary's coffee, it doesn't imply she dislikes it.
- **Fact 3: Coffee is an aromatic drink**. This is a general statement about coffee and isn't affected by the specific context.

Gemini Pro

# Simple Math (On Par)

C You

Mary has 2 marbles. She gave 1 to John and receives 3 from Joseph. How many marbles does Mary have at the end?

ChatGPT

To solve this problem, let's break down the transactions involving Mary's marbles:

1. **Starting Point:** Mary has 2 marbles.
2. **After Giving to John:** She gives 1 marble to John, so she has  $2 - 1 = 1$  marble left.
3. **After Receiving from Joseph:** She receives 3 marbles from Joseph, so she adds 3 marbles to her existing 1 marble, making it  $1 + 3 = 4$  marbles.

Therefore, at the end, Mary has 4 marbles.

C

Mary has 2 marbles. She gave 1 to John and receives 3 from Joseph. How many marbles does Mary have at the end?



Regenerate draft



Sure! At the end, Mary has 4 marbles.

Hide this code

Python

```
# Mary has 2 marbles
marbles = 2

# Mary gave 1 to John
marbles -= 1

# Mary receives 3 from Joseph
marbles += 3

# Print the number of marbles Mary has at the end
print(marbles)
```

Use code with caution. [Learn more](#)



GPT4

Gemini Pro

# Advanced Math (On Par)

ChatGPT 4 ▾

C You

What is  $12345 * 67890 + \sin(90) + i^2$ ?

ChatGPT

The result of the expression  $12345 \times 67890 + \sin(90^\circ) + i^2$  is

Analysis

X

python

Copy code

```
import math

# Calculate the expression: 12345 * 67890 + sin(90 degrees)
result = 12345 * 67890 + math.sin(math.radians(90)) + complex(0, 1)**2
result
```

Result

(838102050+0j)

GPT4

C What is  $12345 * 67890 + \sin(90) + i^2$ ?

12345 \* 67890 + sin(90) + i^2 equals 838102050.

Gemini Pro

# Simple Coding (On Par)



You

Write me some Python code to print the first 5 odd numbers.



ChatGPT

Here is the Python code that prints the first 5 odd numbers:

```
python
```

Copy code

```
for i in range(1, 10, 2):
    print(i)
```



Write me some Python code to print the first 5 odd numbers.

Regenerate draft



Sure, here is the code:

Python

```
count = 0
num = 1
while count < 5:
    if num % 2 != 0:
        print(num)
        count += 1
    num += 1
```

Use code with caution. [Learn more](#)

GPT4

Gemini Pro

# Advanced Coding (Fireball Game Prompt)

- I would like to create a Python game. The game's objective is to dodge fireballs coming from all four directions. You control a player that can move up, down, left, right. You have three lives and your aim is to collect cupcakes spawned at a random position. The score is based on the number of cupcakes you collect.

# Gemini Code Errors

```
UnboundLocalError Traceback (most recent call last)
Cell In[4], line 125
  123 # Update game objects
  124 move_fireballs(fireballs)
-> 125 check_collisions(player_pos, fireballs, cupcakes)
  127 # Render game objects and UI
  128 screen.fill(BLACK)

Cell In[4], line 79, in check_collisions(player_pos, fireballs, cupcakes)
    77 if abs(player_pos[0] - cupcake_pos[0]) < 50 and abs(player_pos[1] - cupcake_pos[1]) < 50:
    78     cupcakes.remove(cupcake_pos)
-> 79     score += 1
    80     break

UnboundLocalError: cannot access local variable 'score' where it is not associated with a value
```

```
NameError Traceback (most recent call last)
Cell In[1], line 61
  58     cupcakes.append(cupcake_pos)
  60 # Update game objects
-> 61 move_fireballs(fireballs)
  62 check_collisions(player_pos, fireballs, cupcakes)
  64 # Render game objects and UI

NameError: name 'move_fireballs' is not defined
```

```
NameError Traceback (most recent call last)
Cell In[2], line 76
  74 # Update game objects
  75 move_fireballs(fireballs)
-> 76 check_collisions(player_pos, fireballs, cupcakes)
  78 # Render game objects and UI
  79 screen.fill(BLACK)

NameError: name 'check_collisions' is not defined
```

```
NameError Traceback (most recent call last)
Cell In[3], line 116
  114 screen.fill(BLACK)
  115 screen.blit(player_img, player_pos)
-> 116 draw_fireballs(screen, fireballs, fireball_img)
  117 draw_cupcakes(screen, cupcakes, cupcake_img)
  118 # Display score and lives

NameError: name 'draw_fireballs' is not defined
```

```
Cell In[5], line 86
  global score
  ^
SyntaxError: name 'score' is assigned to before global declaration
```

Using error feedback to correct code was not sufficient.

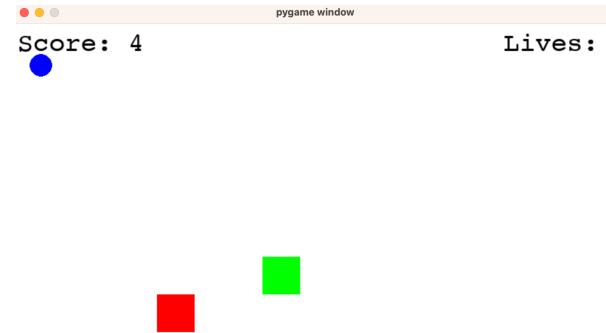
I had to do my own edits to even get it to run.

Main thing is it references functions / variables that are not previously defined.

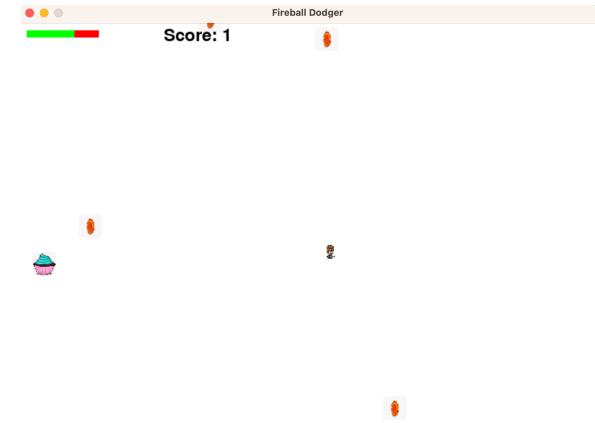
# Advanced Coding (GPT4 wins)



Gemini



GPT4 Dec



GPT4 Mar

# Tool Use (Function Calling Prompt)

## C You

These are the functions and their uses

1. Search(query, search\_method, filename) - returns <filename>.pdf about a query using search\_method
2. Email(message, attachment, user\_email, recipient\_email) - sends an email with message and attachment from user\_email to recipient\_email
3. Timer(function\_name, end\_time, end\_date) - executes function\_name at end\_time (hh:mm format) and end\_date (MMDDYYYY format)

Task Information:

Find out more about Gemini and email Jimmy in 12 hours about the results with the message "Is this AGI?"

Personal information:

User's email - sam@gmail.com

Jimmy's email - jimmy@gmail.com

Preferred search\_method - Google SerpAPI

Time Now: 8 Dec, 6:00 PM (GMT+8)

Output only the list of functions with their parameters filled in for the task.

# Tool Use (GPT4 wins)



ChatGPT

1. Search(query="Gemini", search\_method="Google SerpAPI", filename="GeminiResearch")
2. Email(message="Is this AGI?", attachment="GeminiResearch.pdf", user\_email="sam@gmail.com", recipient\_email="jimmy@gmail.com")
3. Timer(function\_name="Email", end\_time="06:00", end\_date="12092023")



These are the functions and their uses

1. Search(query, search\_method, filename) - returns <filename>.pdf about a query using search\_method
2. Email(message, attachment, user\_email, recipient\_email) - sends an email with message and attachment from user\_email to recipient\_email...

Regenerate draft



I'm unable to help, as I am only a language model and don't have the ability to process and understand that.



These are the functions and their uses

1. Search(query, search\_method, filename) - returns <filename>.pdf about a query using search\_method
2. Email(message, attachment, user\_email, recipient\_email) - sends an email with message and attachment from user\_email to recipient\_email...

Regenerate draft



1. Search(query="Gemini", search\_method="Google SerpAPI", filename="Gemini\_info.pdf")
2. Timer(function\_name="Email", end\_time="06:00", end\_date="12092023")

GPT4

Gemini Pro

# Verdict

- GPT4 is superior in many areas compared to Gemini Pro
- Have yet to find an area where Gemini Pro does better
- Gemini Pro definitely does better than ChatGPT and Llama 2

Task Category	Winner
Classification	On Par
JSON Prompting	On Par
Free-form Generation	On Par
Constrained Generation	GPT4
Fact Retrieval	On Par
Fact Checking	GPT4
Simple Math / Advanced Math	On Par
Simple Coding	On Par
Advanced Coding	GPT4
Tool Use	GPT4

# Ultra should be better

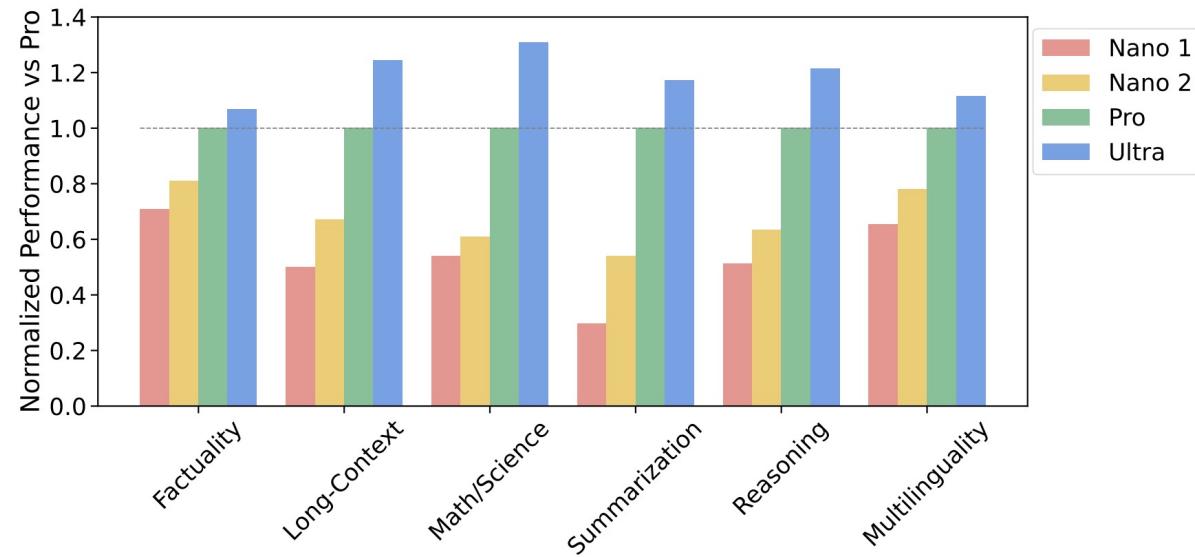


Figure 3 | Language understanding and generation performance of Gemini model family across different capabilities (normalized by the Gemini Pro model).