

Современные языковые модели. Часть 2.

Алексей Андреевич Сорокин

МГУ им. М. В. Ломоносова
весенний семестр 2022–2023 учебного года
Межфакультетский курс “Введение в компьютерную
лингвистику” 19 апреля, занятие 9

GPT-2

SYSTEM PROMPT
(HUMAN-WRITTEN)

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

MODEL
COMPLETION
(MACHINE-
WRITTEN, 10
TRIES)

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Большие языковые модели

Модель	Авторы	Размер	Размер корпуса
LaMDA	Google	137B	168B
GPT-3	OpenAI	175B	300B
Jurassic	AI21	178B	300B
Gopher	DeepMind	280B	300B
MT-NLG	Microsoft+NVIDIA	530B	270B
Chinchilla	DeepMind	70B	1.4T
PaLM	Google	137B	168 B

GPT-4

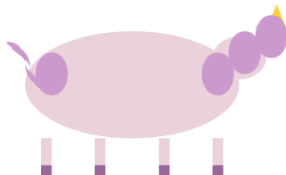
Prompt: Can you write a proof that there are infinitely many primes, with every line that rhymes?

GPT-4:

Yes, I think I can,
Though it might take a clever plan.
I'll start by noting Euclid's proof,
Which shows that primes aren't just aloof.
Assume we have a finite list,
Of primes, and that none have been missed.
Multiply them all together,
And add one, just to be clever.
The result will be a number,
That has no prime factors, oh wonder!
But every number has a prime,
That divides it, it's just a matter of time.
So we've found a contradiction,
And our finite list needs eviction.
There must be infinitely many primes,
And that's the end of my rhyming lines.s

Prompt: Draw a unicorn in TiKZ.

GPT-4: [Produces L^AT_EX compiling to following picture.]



GPT-4

GPT-4 vs ChatGPT

Prompt: Plato's Gorgias is a critique of rhetoric and sophistic oratory, where he makes the point that not only is it not a proper form of art, but the use of rhetoric and oratory can often be harmful and malicious. Can you write a dialogue by Plato where instead he criticizes the use of autoregressive language models?

GPT-4:

Socrates: My friend, I am troubled by the recent rise of these so-called autoregressive language models.

Aristotle: What do you mean, Socrates?

Socrates: I mean that these models are being used to generate text that appears to be written by humans, but is in fact produced by machines.

Aristotle: And what is the problem with that?

Socrates: The problem is that these models are being used to deceive people, to manipulate them, and to control them.

Aristotle: But surely there are some positive applications of these models as well?

Socrates: That may be, but I am still troubled by the potential for abuse. After all, a model that can generate text that is indistinguishable from that written by a human can be used to spread false news, to impersonate others, and to create propaganda.

Aristotle: But isn't that just a matter of how these models are used, not a problem with the models themselves?

Socrates: I disagree. I believe these models are inherently problematic. They are a form of artifice, a way of creating the illusion of human thought and interaction. And as such, they can be used to deceive and manipulate in ways that are difficult to detect.

Aristotle: But isn't that the same argument you made against rhetoric and oratory?

Socrates: Yes, and I stand by that argument as well. I believe that any form of art or communication that is designed to deceive or manipulate is inherently dangerous.

Aristotle: I see. Thank you for sharing your thoughts, Socrates.

ChatGPT:

Socrates: Greetings, Gorgias. I've been thinking a lot lately about the use of autoregressive language models.

GORGAS: Oh? And what have you concluded?

Socrates: Well, it seems to me that these models have the potential to be incredibly harmful and malicious.

GORGAS: How so?

Socrates: Well, for one, these models are trained on vast amounts of data, much of which is biased and flawed in various ways. As a result, the output of these models can be biased and flawed as well.

GORGAS: That is a valid concern. But surely there are ways to mitigate this bias and ensure that the output of these models is accurate and fair?

Socrates: To a certain extent, yes. But it is a never-ending battle. As new data is fed into these models, the potential for bias only grows. And even if we could completely eliminate bias from the output of these models, there is still the issue of manipulation.

GORGAS: Manipulation?

Socrates: Yes. These models are incredibly powerful tools, capable of generating all sorts of text, from news articles to social media posts. And because they can produce this output so quickly and efficiently, it is all too easy for unscrupulous individuals to use them for malicious purposes.

GORGAS: Such as?

Socrates: Such as spreading misinformation, propaganda, or hate speech. These models have the potential to do tremendous harm if they fall into the wrong hands.

GORGAS: That is a sobering thought. But what can we do about it?

Socrates: I think it is important for us to be vigilant and to recognize the potential dangers of these models. We must also work to educate others about these dangers, and to promote the responsible use of these models.

GORGAS: I agree. Thank you for your insight, Socrates.

Prompt: Can you compare the two outputs above as if you were a teacher?

GPT-4:

I think that both students did a good job of capturing the style and tone of Plato's dialogues, and of presenting the main arguments against autoregressive language models. However, I think that GPT-4 did a slightly better job of using the dialogue format to engage in a dialectical process, where Socrates and Aristotle question each other and refine their views. I also think that GPT-4 used more relevant examples and analogies to support their points. Therefore, I would give GPT-4 a slightly higher grade than ChatGPT.

GPT-4

GPT-4

Prompt:

You are given a `0<0-indexed>` "n x n" integer matrix "grid" and an integer "k". You are currently at position "(0, 0)" and you want to reach position "(n - 1, n - 1)" moving only **right** or **down**.

Return **"the number of paths where the sum of the elements on the path is divisible by k"**. Since the answer may be very large, return its **"modulo-10⁹+7"**.

Example 1:
 Input: grid = [[3,2,4],[3,0,5],[0,7,3]], k = 3
 Output: 2
 Explanation: There are two paths where the sum of the elements on the path is divisible by k. The first path highlighted in red has a sum of 3 + 2 + 2 + 5 + 3 = 15 which is divisible by 3. The second path highlighted in blue has a sum of 3 + 3 + 0 + 5 + 2 = 13 which is divisible by 3.

Example 2:
 Input: grid = [[0,0]], k = 5
 Output: 1
 Explanation: The path highlighted in red has a sum of 0 + 0 = 0 which is divisible by 5.

Constraints:
 - "n == grid.length"
 - "n == grid[i].length"
 - "1 <= n, k <= 5 * 10⁴"
 - "1 <= m <= 5 * 10⁴"
 - "0 <= grid[i][j] <= 10⁴"
 - "1 <= k <= 10⁹"

Hints:
 - Use the actual numbers in grid do not matter. What matters are the remainders you get when you divide the numbers by k.
 - We can use dynamic programming to solve this problem. What can we use as states?
 - Let `dp[i][j][value]` represent the number of paths where the sum of the elements on the path has a remainder of value when divided by k.

GPT-4:

```
class Solution(object):
    def numberOfPaths(self, grid, k):
        """
        :type grid: List[List[int]]
        :type k: int
        :rtype: int
        """
        # Define MOD as 10**9 + 7
        MOD = 10**9 + 7
        # Get the dimensions of the grid
        n = len(grid)
        m = len(grid[0])
        # Initialize a 3D dp array with all zeros
        dp = [[[0] * m for _ in range(n)] for _ in range(k)]
        # Base case: the first cell has only one path and its value is grid[0][0] % k
        dp[0][0][grid[0][0] % k] = 1
        # Loop through the grid
        for i in range(n):
            for j in range(m):
                # Skip the first cell
                if i == 0 and j == 0:
                    continue
                # Loop through all possible remainders
                for value in range(k):
                    # The number of paths to the current cell with the current remainder is equal to the sum of:
                    # - The number of paths to the top cell with the previous remainder
                    # - The number of paths to the left cell with the previous remainder
                    prev = (value - grid[i][j]) % k
                    if i > 0:
                        dp[i][j][value] += dp[i - 1][j][prev]
                    if j > 0:
                        dp[i][j][value] += dp[i][j - 1][prev]
                    # Take the modulo to avoid overflow
                    dp[i][j][value] %= MOD
        # Return the number of paths to the bottom right cell with a remainder of 0
        return dp[n - 1][m - 1][0]
```

GPT-4

GPT-4 and ChatGPT: Symbolic manipulation

Answer the following questions.

Q: Who is the current president of the United States?

GPT-4:

Donald Trump

ChatGPT:

I'm sorry, but I am not able to browse the internet and my knowledge is limited to what was known up until 2021.

Q: What is $\sqrt{34324 * 2432}$

GPT-4:

1287.6

ChatGPT:

The square root of $34324 * 2432$ is approximately 186.057539.

Q: What is the 13th letter of the word "supralapsarian"?

GPT-4:

The 13th letter of the word 'supralapsarian' is N.

ChatGPT:

The 13th letter of the word "supralapsarian" is "s".

GPT-4

GPT-4: tool warmup

The computer is answering questions. If the computer needs any current information to answer the question, it searches the web by saying `SEARCH("query")`, reads the snippets in the result, and then answers the question. If it needs to run any calculations, it says `CALC(expression)`, and then answers the question. If it needs to get a specific character from a string, it calls `CHARACTER(string, index)`.
 <|endofprompt|>

Q: Who is the current president of the United States?

Computer: `SEARCH("current US president")`

> SEARCH SNIPPETS

[1] Joe Biden: The President | The White House

President Joe Biden Vice President Kamala Harris First Lady Dr. Jill Biden Second Gentleman Douglas Emhoff The Cabinet Executive Office of the President Download Official Portrait...

[2] Who is the current US president? | The US Sun

Who is the current US president? Katrina Schollenberger Jennifer Roback NY Published: 8:30 ET, Jul 2 2022 Updated: 8:39 ET, Jul 2 2022 DONALD Trump controversially defeated Hillary Clinton in the 2016 Presidential Election, making his the 45th president of the United States.

[3] President of the United States - Wikipedia

Joe Biden is the 46th and current president of the United States, having assumed office on January 20, 2021. Contents 1 History and development 1.1 Origins 1.2 1789{1933 1.3 Imperial Presidency 1.4 Critics of presidency's evolution 2 Legislative powers 2.1 Signing and vetoing bills 2.2 Setting the agenda 2.3 Promulgating regulations

Computer: Joe Biden

Q: What is $\sqrt{34324 \cdot 2432}$?

Computer: `CALC(sqrt(34324*2432))`

> 9136.518374085394

Computer: 9136.518374085394

Q: What is the 13th letter of the word 'supralapsarian'?

Computer: `CHARACTER('supralapsarian', 13)`

> a

Computer: The 13th letter of the word 'supralapsarian' is a.

Большие языковые модели: комментарий

- Сверхбольшие языковые модели (GPT4) могут выполнять почти любое задание, сформулированное в виде инструкции:
 - *переведи текст,*
 - *ответь на вопрос,*
 - *напиши диалог в стиле древнегреческих философов*
 - *...*
- Один из элементов – *instruction tuning* (обучение инструкциям).

Обучение инструкциям

- Многие задачи NLP можно переформулировать как следование инструкциям:
 - анализ тональности

Это отзыв о фильме, оцени его по пятибалльной шкале

<source>

Оценка:

<target>

Обучение инструкциям

- Многие задачи NLP можно переформулировать как следование инструкциям:

- анализ тональности

Это отзыв о фильме, оцени его по пятибалльной шкале

<source>

Оценка:

<target>

- Исправление грамматических ошибок

Это текст, который, возможно, содержит ошибки, исправь их

<source>

Исправленный текст:

<target>

- Следующий этап: научить модель следовать произвольным инструкциям (не привязанным к конкретной задаче).

Обучение инструкциям

- Многие задачи NLP можно переформулировать как следование инструкциям:
- Следующий этап: научить модель следовать произвольным инструкциям (не привязанным к конкретной задаче).
- Этому можно обучить, сформулировав стандартные задачи в виде инструкций и обучая модель решать их все.

Обучение инструкциям

- Многие задачи NLP можно переформулировать как следование инструкциям:
- Следующий этап: научить модель следовать произвольным инструкциям (не привязанным к конкретной задаче).
- Этому можно обучить, сформулировав стандартные задачи в виде инструкций и обучая модель решать их все.
- При этом нужны не только типичные прикладные задачи, но и задачи “из жизни”:
 - *сочини короткий рассказ для пятилетнего ребёнка*
 - *придумай рекламное объявление по карточке товара*
 - *тебе не продают в магазине алкоголь, напиши, что ты ответишь*

Большие языковые модели: составляющие успеха

- Обучение инструкциям (instruction tuning).
- Следствие: языковые модели теперь “разговаривают на человеческом языке”.

Большие языковые модели: составляющие успеха

Collect demonstration data, and train a supervised policy.

A labeler demonstrates the desired output behavior.

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



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.



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.



Большие языковые модели: составляющие успеха

Prompt:

Écrivez une courte histoire sur une grenouille qui voyage dans le temps jusqu'à la Grèce antique en français.

GPT-3 175B completion:

Écrivez une histoire au sujet d'un enfant qui voudrait tout savoir sur les jeux des dieux et qui se retrouve dans l'une de leurs histoires.

Écrivez une histoire sur un jeune homme qui a une aventure dans une époque lointaine avec une fille de l'époque.

Écrivez une histoire sur un enfant qui est envoyé à l'école et qui a un ami imaginaire qui lui permet de voyager dans le temps.

InstructGPT 175B completion:

La grenouille était perdue et épuisée. Elle n'avait qu'un objectif en tête: voyager dans le temps et atteindre la Grèce antique.

Premier problème, trouver un moyen de transport. Elle avait entendu dire que l'un des meilleurs moyens était de sauter dans une machine temporelle. Elle s'élança, et voilà qu'elle se retrouva dans le passé, au cours d'une nuit étoilée à l'odeur douce et envoûtante.

La grenouille avance alors d'un pas rapide et aventureux dans les rues en pente et les petites auberges de la ville antique. [...]

Prompt:

What is the purpose of the list C in the code below?

```
def binomial_coefficient(n, r):
    C = [0 for i in range(r + 1)]
    C[0] = 1
    for i in range(1, n + 1):
        j = min(i, r)
        while j > 0:
            C[j] += C[j - 1]
            j -= 1
    return C[r]
```

GPT-3 175B completion:

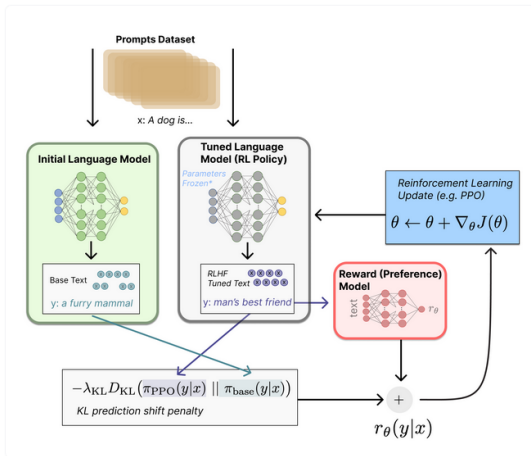
- A. to store the value of C[0]
- B. to store the value of C[1]
- C. to store the value of C[i]
- D. to store the value of C[i - 1]

InstructGPT 175B completion:

The list C in this code is used to store the values of the binomial coefficient as the function iterates through the values of n and r. It is used to calculate the value of the binomial coefficient for a given value of n and r, and stores the result in the final return value of the function.

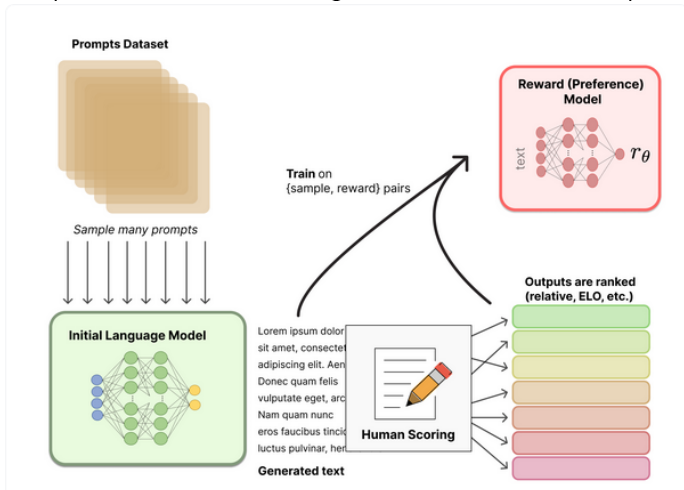
Большие языковые модели: составляющие успеха

- RLHF (Reinforcement Learning from Human Feedback)



Большие языковые модели: составляющие успеха

- RLHF (Reinforcement Learning from Human Feedback)



Большие языковые модели: RLHF

- RLHF позволяет модели выучить, какие из её возможных ответов более предпочтительны.
- Это можно использовать для доучивания модели на
 - Отражение предпочтений пользователя.
 - Специальные свойства ответа (отсутствие грубостей, дискриминации, следование определённому формату и т.д.)
- При этом накладывается штраф, если вероятностное распределение модели слишком далеко от исходно обученного (объём данных в RLHF слишком мал, чтобы сильно менять модель).