

1
2
0
2
b
e
F
5
1
1
L
C
s
c
l
1
v
0
5
3
7
0
2
0
1
2
v
i
X
r
a

ew-Shot : F

ai moire@knc. le@knc.ai ky

Abstract

. GPT-3 0- 가 few-shot
few-shot
가
metaprompt

Keywords: , , GPT-3,
GPT-3 [3], *Language models are few-shot learners*
few-shot
, GPT-3 few
learn- ing
-shot
task location
few-shot

1 Motivation

GPT-3[3]
가
가
GPT-3 가
[12]. GPT-3
가 few-shot
few-shot
few-shot
가

few-shot
0-shot [13, 14].

metaprompt programming

3 Investigating few-shot prompting

GPT-3 0, 1, n (n)가 ,
가 . GPT-3 가 , 0
GPT-3가

2 Related work

[15, 1 [3].
[6, 11], [22, 17], [24] 가
few-shot [18, 9].
0-
API AI Dungeon OpenAI GPT-3 가
few-shot-learning 가
가 few-s)가
hot

Gwern , ,
GPT-3 , PDF 가
GPT-3
am Sabeti 가
[2]. Arr 가 few-shot 0- 가
[21]. Zachary Robert [3] 0- 가 GPT-3
son GPT-3
[20]. Twitter KaryoKleptid (1) 가 GPT-3
GPT-3 10-

3.1 The success of 0-shot prompts

3.2 Examples don't always help

	Prompt	Babbage / 6.7B	Curie / 13B
	OpenAI 0-shot	15.5	22.4
	OpenAI 1-shot	31.6	31.4
	OpenAI 64-shot	36.4	38.3
	Reproduced OpenAI 0-shot	15.9	18.7
	Reproduced OpenAI 1-shot	21.8	24.1
	Reproduced OpenAI 10-shot	25.1	27.9
	Simple colon 0-shot	23.5	33.3
	Simple colon 1-shot	18.0	27.6
	Simple colon 10-shot	24.1	33.4
	Master translator 0-shot	26.5	32.9

1: SacreBLEU [19]
WMT'14 Fr-En
[1]
GPT-3 GPT-

3 BLEU , OpenAI API 가 Babbage Curie

[3] (1, 2)

API Babbage Curie GPT-3 가

64 , 10 .

French: example_source_phrase
 English: example_target_phrase

: source_phrase
 가가
 :
 2: "
 " bold
 [...]

French: source_phrase
 English:

Figure 1: The “Simple Colon” prompt format. For few-shot tasks, additional examples are provided as shown. Text in **bold** is to be replaced by source and target language text examples.

4 Prompt programming

(GPT-3)
 가
 ?
 ,
programming in natural language
 가 , , , ,
 , ,
 .

4.1 The dynamics of language

, GPT-3가
 가
anthropomorphizing the model . G
 GPT-3 가 [3].
 , PT-3
 , 가
 가
 .GPT-3 가 (§4.5)
 GPT-3가
function (§4.6)
 .GPT-3
 가
 , ,
 가
 ().
 [8]. "
 , ,
 가
 가
 가
 , ()
 §4.2 - §4.7
 GPT-3
 , few-shot
 . [2]

GPT-3 demonstration GPT-3 가
 , GPT-3 .
 , , 가
 . ,

4.2 Direct task specification: constructing the signifier

GPT-3 fine-tuning few-shot " " §3.2
 .0 GPT-3 few-shot , 가
 가 .
 0- signifier 가 [23].
 .

4.4 Task specification by memetic proxy

,"2 " 1
 how 가 GPT-3 [23].
 , 가
 /
 ,
 , " ,
 " 가 ,
 ,
 가 가 ,
 (§4.3) 가 (§
 4.4) GPT-3 [2].
 (GPT-3

4.3 Task specification by demonstration

Few-shot 가
 GPT-3 가 , GPT-3 가
 , , 가
 , 가
 가 ,
 " ,
 " ,
 ,

4.5 Prompt programming as

4.5 Prompt programming as constraining behavior

GPT-3

가

a person would

가

any person

could

4.6 Serializing reasoning for closed-ended questions

could 가 any person 가

.

.

GPT-3 가

가) (가

가 [2].

,

가 가

,

,

,

:

Mon corps est unTransformateur de soi, mais auss
i unTransformateur pour cette cire de langage.

" " "

[illegible]

[26, 7]. 가 , 가

GPT-3 " " ((6). BERT[5]

), . GPT-3가

[20, 13]. **4.7 Metaprompt programming**

GPT-3 가

가 GPT-3 가

. GPT-3

" "

[19].

(§4.7). *metaprompts*

, 가 (:)

" ~

" ,

rationalize [27]. "

가 §4.7

가 (3-5) OpenAI API(

가 =davinci, =0) GPT-3

가 " "

bold , GPT-3

(

)

"

f(x) = x*x. What is f(f(3))?

Let's solve this problem by splitting it into steps.

f(f(3)) = f(3*3) = 3*3*3 = 27

We can see that f(3) = 3*3 = 9, so f(f(3)) = 27.

GPT-3

가 " 가

Figure 3: A generic serializing metaprompt applied to a math problem

A word cloud visualization of the Korean word '가' (ga) in various sizes and orientations, with some text labels like 'Tesla Motors', 'CEO', 'Boring Company', 'OpenAI', 'Neuralink', 'SpaceX', 'T', 'Future', and 'of Life Institute' interspersed.

6: 가
,"
?"

5 Directions for future work

§3 (BLEU) . 가 0-

8

(OpenAI

(: NP-

, ,

, GPT-3 가
produce
가 noticing

가 가
가 0 64

5.2.4 Games

가

§3.2

1

0 64

[4]

()

. 가

5.2 New methods for benchmarking

가

5.2.1 Isolating catastrophic failures

가

가

가

가

가

5.2.2 Metaprompts for evaluations

가 가 가
가

(:

)

5.2.3 Language models for evaluations

가 가
가

Acknowledgements

La

v Varshney

Michael Ivanitskiy John Balis

GPT-3

Miles Brundage OpenAI

References

[1] Ondrej Bojar et al. “ 2014

” . In:

Proceedings of the Ninth Workshop on Statistical Machine Translation.

: , 2014 6 , 12-58 .

[2] . “ GPT-3 ” . : (2020).

[3] Tom B Brown et al. " . In: *arXiv preprint arXiv:2005.14165* (2020).

- [4] Marc-Alexandre Côté. "TextWorld: A Text-Based World for Embodied Agents". (2019). arXiv: 1806.11532 [cs.LG]. [5] Jacob Devlin. "Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding". *arXiv preprint arXiv:1810.04805* (2018). [6] Angela Fan, Mike Lewis, Yann Dauphin. *Hierarchical Neural Story Generation*. 2018. arXiv: 1805.04833 [cs.CL]. [7] Zhe Gan. *Multi-step Reasoning via Recurrent Dual Attention for Visual Dialogue*. 2019. arXiv: 1902.00579 [cs.CV]. URL: <https://arxiv.org/abs/1902.00579>. [8] Leo Gao. "A GPT-2". 2020. [leogao.dev](https://bit.ly/3rViLGk) (2020); <https://bit.ly/3rViLGk>. [9] Tianyu Gao, Adam Fisch, Danqi Chen. *Making Pre-trained Language Models Better Few-shot Learners*. 2020. arXiv: 2012.15723 [cs.CL]. [10] Dan Hendrycks. "A Challenging Dataset for Natural Language Understanding". 2020. URL: <https://arxiv.org/abs/2009.03300>. [11] Ari Holtzman. *The Curious Case of Neural Text Degeneration*. 2020. arXiv: 1904.09751 [cs.CL]. [12] Jeremy Howard, Sebastian Ruder. "The BERT Toolkit". *arXiv preprint arXiv:1801.06146* (2018). URL: <https://arxiv.org/abs/1801.06146>. [13] KaryoKleptid. *Seems to work*. 2020. URL: <https://bit.ly/37dA1hY>. [14] KaryoKleptid. *Teaching GPT-3 to do a brute force 'for loop' checking answers*. 2020. URL: <https://bit.ly/2N7khX1>. [15] Nitish Shirish Keskar et al. "CTRL: A Conditional Transformer for Natural Language Generation". *CoRR* abs/1909.05858(2019). arXiv: 1909.05858. URL: <http://arxiv.org/abs/1909.05858>. [16] Ben Krause. "GeDi: Generative Discriminator Guided Sequence Generation". 2020. [17] Xiang Lisa Li, Percy Liang. "Prefix-Tuning: Optimizing Continuous Prompts for Generation". 2021. *arXiv preprint arXiv:2101.00190*. [18] Jianming Liu, Matt Gardner. "BLEU". 2020. [19] Matt Post. "BLEU". 2020. *Proceedings of the Association for Computational Linguistics*, 2018-10, pp. 186 – 191. URL: <https://www.aclweb.org/anthology/W18-6319>. [20] Zachary Robertson. *You Can Probably Amplify GPT3 Directly*. 2020. URL: <https://bit.ly/3tXT7Cw>. [21] Arram Sabeti. *GPT-3: Using Fiction to Demonstrate How Prompts Impact Output Quality*. 2020. URL: <https://bit.ly/3jP3TWW>. [22] Taylor Shin et al. *AutoPrompt: Eliciting Knowledge from Language Models with Automatically Generated Prompts*. 2020. arXiv: 2010.15980 [cs.CL]. [23] Latitude Team. *World Creation by Analogy*. 2020. URL: <https://bit.ly/2N4vXK0>. [24] Lilian Wang. "가". (2021). URL: <https://bit.ly/3pl2eKa>. [25] Qinyuan Ye, Xiang Ren. *Zero-shot Learning by Generating Task-specific Adapters*. 2021. arXiv: 2101.00420 [cs.CL]. [26] Jianxing Yu et al. "GPT-3". *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 2020, 6729 – 6739. [27] Eliezer Yudkowsky. "less-wrong.com (2007)". URL: <https://bit.ly/3pmYt6>.