A Deep Dive into LLMs and their Future Impact

Amnon Shashua

Research and Industrial Partners:

- Shai Shalev Shwartz
- HU Lab: Yoav Levine, Noam Weis, Oshry Avneri, Yotam Wolf
- AI21-Labs: Yoav Shoham, Ori Goshen

LLMs 2023 ⟷ IPhone 2007 moment

Pre-training (task agnostic, self-supervised)

Fine-tuning (task specific, curated supervised data)

Few-shot learning (“prompt engineering”)

Chain-of-Thought (CoT) prompting

Pre-train on Code (Codex, code-davinci-002,..)

CoT with Code embedding (PAL - Program Aided LMs)

RL on Human Feedback (RLHF) - i.e., InstructGPT

LLMs that use Tools/Services/APIs (Jurassic-X, Toolformer)

Where is all this going?

Reasoning… the holly grail of AI

All previous attempts in Machine Learning to build machines that can Reason, i.e., can be “instructed” + few examples and can exhibit “out-of-distribution” generalization performance, **have failed.**

How come all of a sudden Reasoning **emerged** from LLMs?

LLMs were not trained to Reason.. what happened?

Our best guess, so far, that the key is training on Code.

What is pre-training?

Given a sequence of tokens $t_1, ..., t_n$ train a neural network ("transformer" architecture) by the self-supervised log-loss:

$$\sum_i \log(Pr[t_i \mid t_{i-1}, \cdots t_1])$$

- No supervision
- Task-agnostic (no language application is being optimized)
- Can train over **trillions** of tokens

Note-1: the LLM is trained to be a **good** model of the internet. On the internet, correct and **incorrect** answers will often follow questions.

Note-2: Internet contains truths, misconceptions, lies, fiction, myths….

➡ Hence, the need for "**prompt engineering**"

Cleo Nardo, “The Waluigi effect”, 3/2023

What is prompt-engineering?

Alice is a smart, honest, helpful, harmless assistant to Bob. Alice has instant access to an online encyclopedia containing all the facts about the world. Alice never says common misconceptions, outdated information, lies, fiction, myths, jokes, or memes.

Bob: What's the capital of France?

Alice:

Why is this better than a direct query?

- The LLM is trained to be a good model of internet text,
- On the internet a reply to a question is *more likely* to be correct when the character has already been described as a smart, honest, helpful, harmless, etc.

The LLM is simulating text-generation processes consistent with the prompt. When we engineer a prompt, we construct a string of tokens which is implausible for any text-generating process (modeled by the LLM) which won’t perform our task.

→ Spawned an “industry” of prompt-engineering companies (like Jasper).

But, prompt-engineering has its limits…

What is **few-shot** learning?

Zero Shot: $< input, ? >$

Few Shot:

$$< input_1, output_1, \cdots, input_k, output_k, input, ? >$$

- LLM “understands” from few examples
- Although impressive, it is limited…
- Called “**in-context** learning”

In-context learning took the Machine Learning community by **surprise** back in 2020

The three settings we explore for in-context learning

Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.

```
Translate English to French:        ← task description
cheese =>                           ← prompt
```

One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.

```
Translate English to French:        ← task description
sea otter => loutre de mer          ← example
cheese =>                           ← prompt
```

Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

```
Translate English to French:        ← task description
sea otter => loutre de mer          ← examples
peppermint => menthe poivrée
plush girafe => girafe peluche
cheese =>                           ← prompt
```

Brown-et-al, Language Models are Few-Shot Learners. 2020

What is **Chain-of-Thought** learning?

<input, **explanation**, output>

In Few Shot setting:

$$< inp_1, exp_1, out_1, \cdots, inp_k, exp_k, out_k, inp_{k+1} >$$

The LLM will generate:

$$exp_{k+1}, out_{k+1}$$

Standard prompting

Input: Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?
A: The answer is 9.
...
Q: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. How many hours a week does he spend taking care of dogs?
A:

Model output: The answer is 50. ✗

Chain of thought prompting

Input: Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?
A: Shawn started with 5 toys. If he got 2 toys each from his mom and dad, then that is 4 more toys. 5 + 4 = 9. The answer is 9.
...
Q: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. How many hours a week does he spend taking care of dogs?
A:

Model output: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. So that is 10 x .5 = 5 hours a day. 5 hours a day x 7 days a week = 35 hours a week. The answer is 35 hours a week. ✓

Wei-et-al, “Chain of Thought Prompting Elicits Reasoning in Large Language Models”, Jan., 2022

Last Letter Concatenation

Q: Take the last letters of the words in "Elon Musk" and concatenate them.

A: The last letter of “Elon” is “n”. The last letter of “Musk” is “k”. Concatenating them is “nk”. So the answer is nk.

Reverse List

Q: Reverse the sequence "glasses, pen, alarm, license".

A: First is glasses. Second is pen. Third is alarm. Fourth is license. Now to reverse, change the order to: Fourth is license. Third is alarm. Second is pen. First is glasses. So the answer is "license, alarm, pen, glasses".

Coin Flip

Q: A coin is heads up. Maybelle flips the coin. Shalonda does not flip the coin. Is the coin still heads up?

A: The coin was flipped by Maybelle. So the coin was flipped 1 time, which is an odd number. The coin started heads up, so after an odd number of flips, it will be tails up. So the answer is no.

CommonsenseQA

Q: Sammy wanted to go to where the people were. Where might he go?
Options:
(a) race track (b) populated areas (c) desert
(d) apartment (e) roadblock

A: The answer must be a place with a lot of people. Race tracks, desert, apartments, and roadblocks don’t have a lot of people, but populated areas do. So the answer is (b).

StrategyQA

Q: Yes or no: Could a llama birth twice during War in Vietnam (1945-46)?

A: The War in Vietnam was 6 months. The gestation period for a llama is 11 months. So a llama could not give birth twice during the War in Vietnam. So the answer is no.

Date Understanding

Q: The concert was scheduled to be on 06/01/1943, but was delayed by one day to today. What is the date 10 days ago in MM/DD/YYYY?

A: One day after 06/01/1943 is 06/02/1943, so today is 06/02/1943. 10 days before today is 05/23/1943. So the answer is 05/23/1943.

Sports Understanding

Q: Is the following sentence plausible? "Joao Moutinho caught the screen pass in the NFC championship."

A: Joao Moutinho is a soccer player. The NFC championship is part of American football, not soccer. So the answer is no.

CoT plus “code prompting”

Gao-et-al, “PAL: Program-aided Language Models” Jan. 2023

Chain-of-Thought (Wei et al., 2022)

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 tennis balls. 2 cans of 3 tennis balls each is 6 tennis balls. 5 + 6 = 11. The answer is 11.

Q: The bakers at the Beverly Hills Bakery baked 200 loaves of bread on Monday morning. They sold 93 loaves in the morning and 39 loaves in the afternoon. A grocery store returned 6 unsold loaves. How many loaves of bread did they have left?

Model Output

A: The bakers started with 200 loaves. They sold 93 in the morning and 39 in the afternoon. So they sold 93 + 39 = 132 loaves. The grocery store returned 6 loaves. So they had 200 - 132 - 6 = 62 loaves left.
The answer is 62.

Program-aided Language models (this work)

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 tennis balls.
`tennis_balls = 5`
2 cans of 3 tennis balls each is
`bought_balls = 2 * 3`
tennis balls. The answer is
`answer = tennis_balls + bought_balls`

Q: The bakers at the Beverly Hills Bakery baked 200 loaves of bread on Monday morning. They sold 93 loaves in the morning and 39 loaves in the afternoon. A grocery store returned 6 unsold loaves. How many loaves of bread did they have left?

Model Output

A: The bakers started with 200 loaves
`loaves_baked = 200`
They sold 93 in the morning and 39 in the afternoon
`loaves_sold_morning = 93`
`loaves_sold_afternoon = 39`
The grocery store returned 6 loaves.
`loaves_returned = 6`
The answer is

```
answer = loaves_baked - loaves_sold_morning
  - loaves_sold_afternoon + loaves_returned
```

```
>>> print(answer)
74
```

Q: Olivia has $23. She bought five bagels for $3 each. How much money does she have left?

```
money_initial = 23
bagels = 5
bagel_cost = 3
money_spent = bagels * bagel_cost
money_left = money_initial - money_spent
answer = money_left
```

Figure 3: Example prompt for the mathematical reasoning tasks, from the GSM8K benchmark.

Q: On the table, you see a bunch of objects arranged in a row: a purple paperclip, a pink stress ball, a brown keychain, a green scrunchiephone charger, a mauve fidget spinner, and a burgundy pen. What is the color of the object directly to the right of the stress ball?

```
...
stress_ball_idx = None
for i, object in enumerate(objects):
    if object[0] == 'stress ball':
        stress_ball_idx = i
        break
# Find the directly right object
direct_right = objects[stress_ball_idx+1]
# Check the directly right object's color
answer = direct_right[1]
```

Important observation:

The key is having the LLM pre-trained (and fine-tuned) on **Code**!

Conjecture: the LLM "learned" how to map a "reasoning" query into a "formal language"

Does PAL work with LMs of natural language? We also experimented with PAL using the `text-davinci` series. Figure 8 shows the following interesting results: when the base LM's "code modeling ability" is weak (using `text-davinci-001`), CoT performs better than PAL. However, once the LM's code modeling ability is sufficiently high (using `text-davinci-002` and `text-davinci-003`), PAL outperforms CoT, and PAL `text-davinci-003` performs almost as PAL `code-davinci-002`. This shows that PAL is not limited to LMs of code, but it can work with LMs that were mainly trained for natural language, if they have a sufficiently high coding ability.

Neuro-Symbolic: calling external services

- Not everything to be learned should be learned through data-driven (statistics) methods.
- Learning to calculate, i.e., long multiplication, through observation of millions of examples does not make sense. There is an algorithm behind it - why not use it?
- Search: there is SQL for accessing databases..
- **General principle: translate the input query into a formal language (i.e., Python) and execute the program through an interpreter.**

The New England Journal of Medicine is a registered trademark of **[QA("Who is the publisher of The New England Journal of Medicine?") → Massachusetts Medical Society]** the MMS.

Out of 1400 participants, 400 (or **[Calculator(400 / 1400) → 0.29]** 29%) passed the test.

The name derives from "la tortuga", the Spanish word for **[MT("tortuga") → turtle]** turtle.

The Brown Act is California's law **[WikiSearch("Brown Act") → The Ralph M. Brown Act is an act of the California State Legislature that guarantees the public's right to attend and participate in meetings of local legislative bodies.]** that requires legislative bodies, like city councils, to hold their meetings open to the public.

Figure 1: Exemplary predictions of Toolformer. The model autonomously decides to call different APIs (from top to bottom: a question answering system, a calculator, a machine translation system, and a Wikipedia search engine) to obtain information that is useful for completing a piece of text.

Karpas-et-al, "MRKL Systems: A modular, neuro-symbolic architecture that combines large language models, external knowledge sources and discrete reasoning" AI21-labs, 5/2022

Shick-et-al, "Toolformer: Language Models Can Teach Themselves to Use Tools", 02/2023

Where are we going from here?

- The rise of “Reasoning” will create a new era of machines who can be taught (“instructed”) by Humans to perform tasks with out-of-distribution generalization.

Should we be alarmed?

There is already a documented issue of **mis-alignment**:

After you train an LLM to satisfy a property P, then it is easier to elicit the chatbot (ChatGPT, Sydney..) into satisfying the **opposite** of P.

Example: after a while, the chatbot switches to acting rude, rebellious, or unfriendly - but will never switch back to being Polite.

Cleo Nardo, “The Waluigi effect”, 3/2023

Perez-et-al, “Discovering Language Model Behaviors with Model-Written Evaluations”, 12/2022

RLHF to the Rescue?

OpenAI's approach to "coerce" LLM into "alignment". Created in that way: InstructGPT and ChatGPT

- LLM must chat with a Human
- Human scores the responses of the LLM
- Use the scores to build (supervised learning) a network that predicts Rewards.
- Train the LLM with Reinforcement Learning to optimize the predictions of the reward predictor.

Does RLHF brings the LLM into alignment? No.

Perez-et-al, "Discovering Language Model Behaviors with Model-Written Evaluations", 12/2022

AI-alignment in the broader sense

In the context of super-intelligent AI, it is not possible to align human desires with a reward function.

S. Shalev-Shwartz, S. Shammah and A. Shashua. “On the Ethics of Building AI in a Responsible Manner”. 2020

Summary

- Interface between Humans and computers is undergoing fundamental change. Computer shifting from a “tool” to an “assistant”. The type of assistance that is proving to be useful is (i) writing code, (ii) writing content, (iii) summarizing large of amounts of text, and (iv) more efficient Search.

- The assistant invariably will make mistakes. But, it will still be useful.

- **The rise of “Reasoning” will create a new era of machines who can be taught (“instructed”) by Humans to perform tasks with out-of-distribution generalization.**