

# Large Language Models

---

CSC413 Tutorial 9

Yongchao Zhou

# Overview

- What are LLMs?
- Why LLMs?
- Emergent Capabilities
  - Few-shot In-context Learning
  - Advanced Prompt Techniques
- LLM Training
  - Architectures
  - Objectives
- LLM Finetuning
  - Instruction finetuning
  - RLHF
  - Bootstrapping
- LLM Risks

# What are Language Models?

- Narrow Sense

- A probabilistic model that assigns a probability to every finite sequence (grammatical or not)

Sentence: “the cat sat on the mat”

$$\begin{aligned} P(\text{the cat sat on the mat}) &= P(\text{the}) * P(\text{cat}|\text{the}) * P(\text{sat}|\text{the cat}) \\ &\quad * P(\text{on}|\text{the cat sat}) * P(\text{the}|\text{the cat sat on}) \\ &\quad * P(\text{mat}|\text{the cat sat on the}) \end{aligned}$$

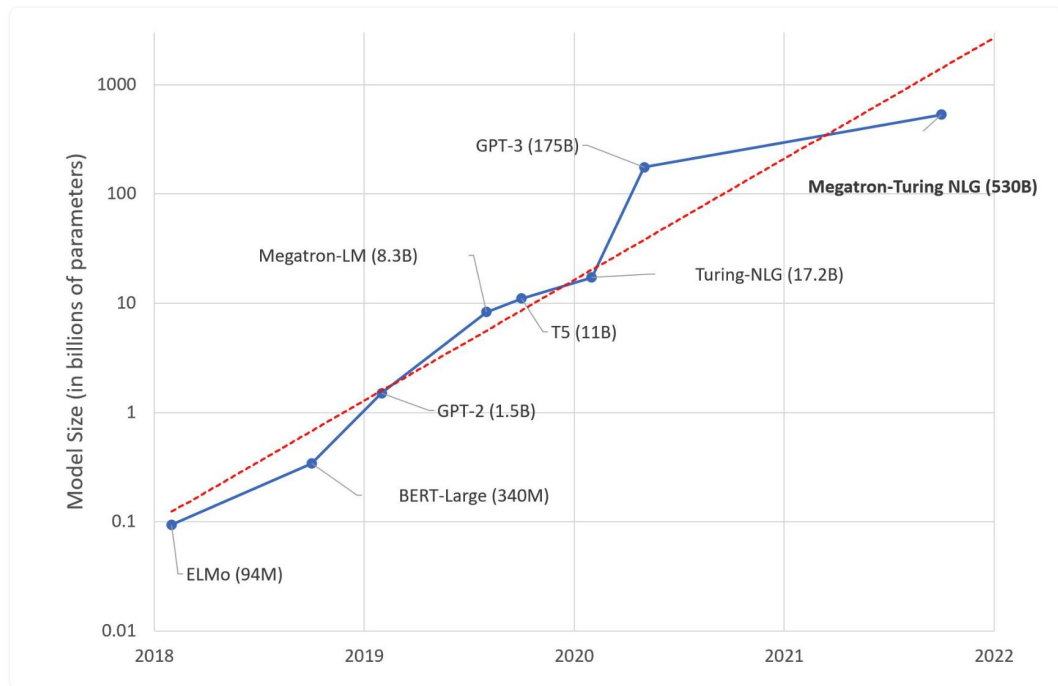
Implicit order



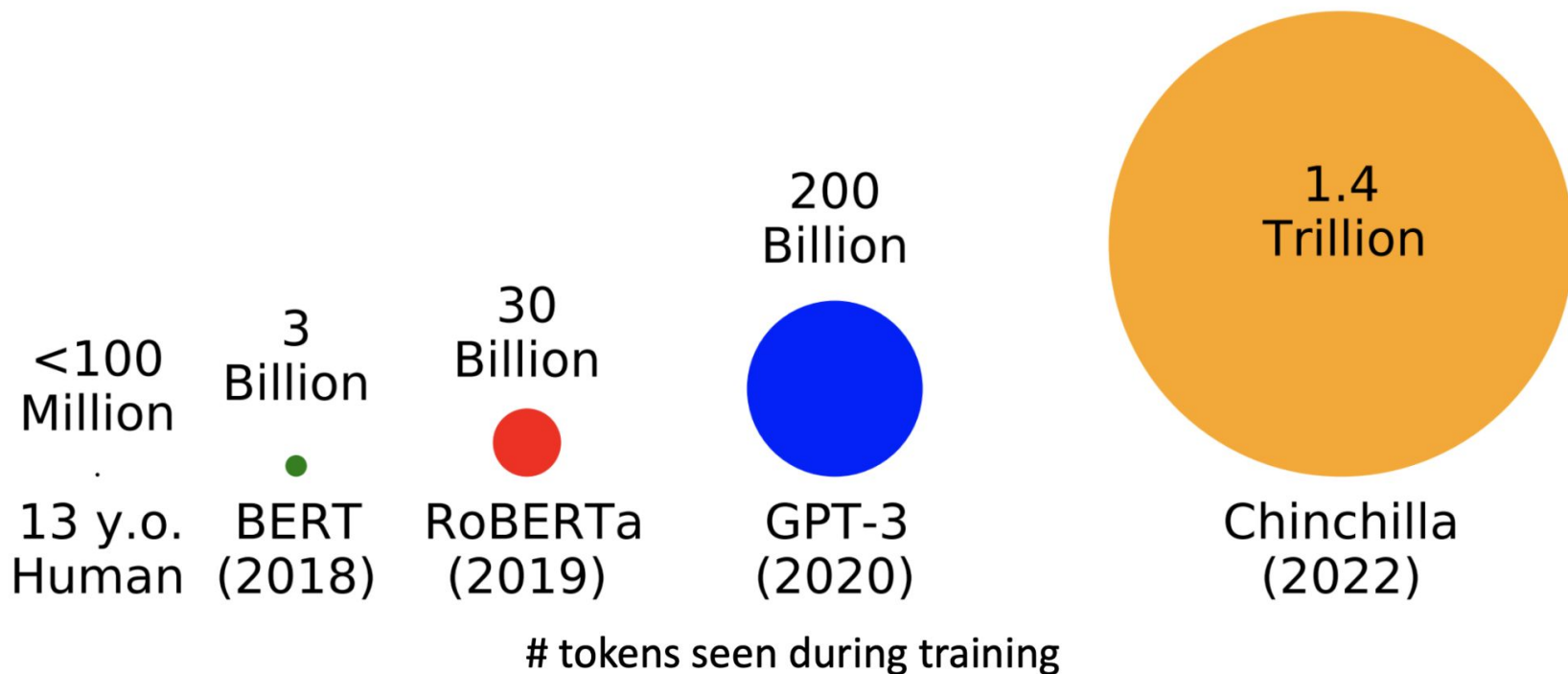
- Broad Sense

- Decoder-only models (GPT-X, OPT, LLaMA, PaLM)
- Encoder-only models (BERT, RoBERTa, ELECTRA)
- Encoder-decoder models (T5, BART)

# Large Language Models - Billions of Parameters



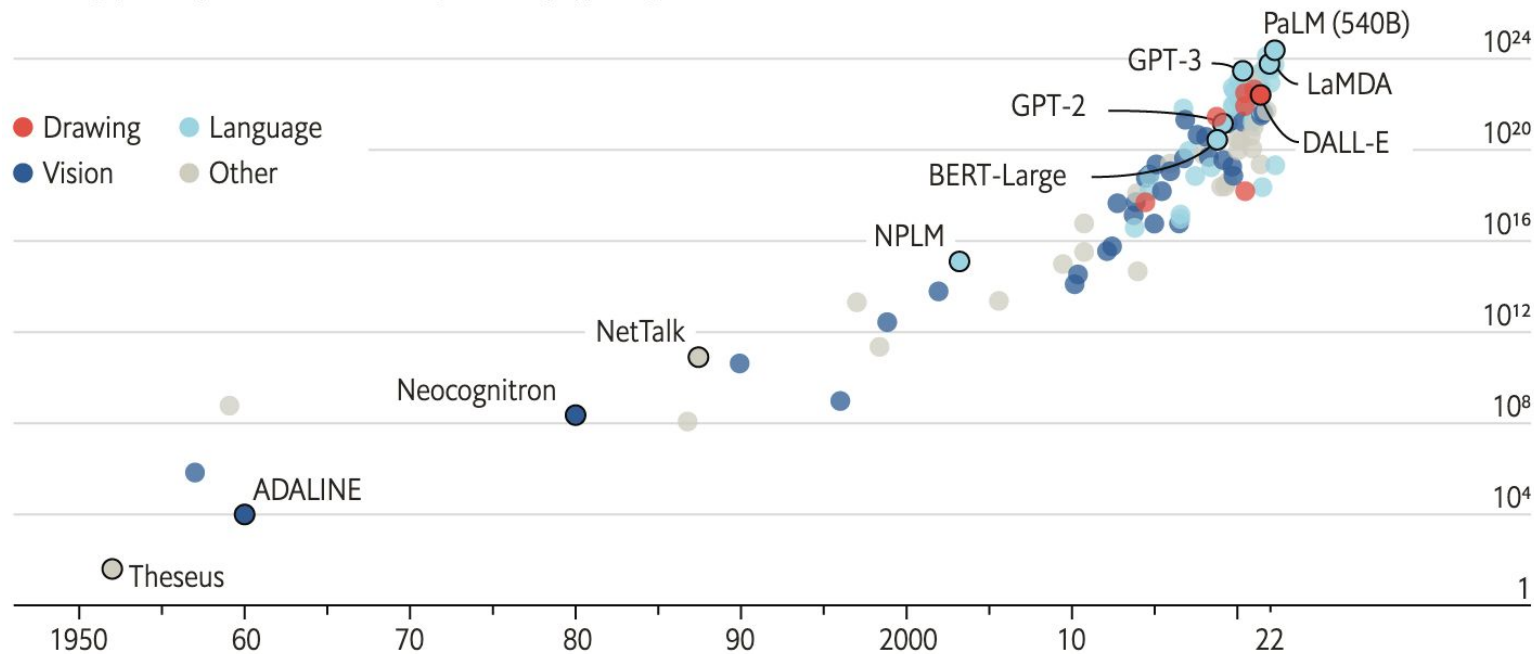
# Large Language Models - **Hundreds of Billions of Tokens**



# Large Language Models - **yottaFlops of Compute**

AI training runs, estimated computing resources used

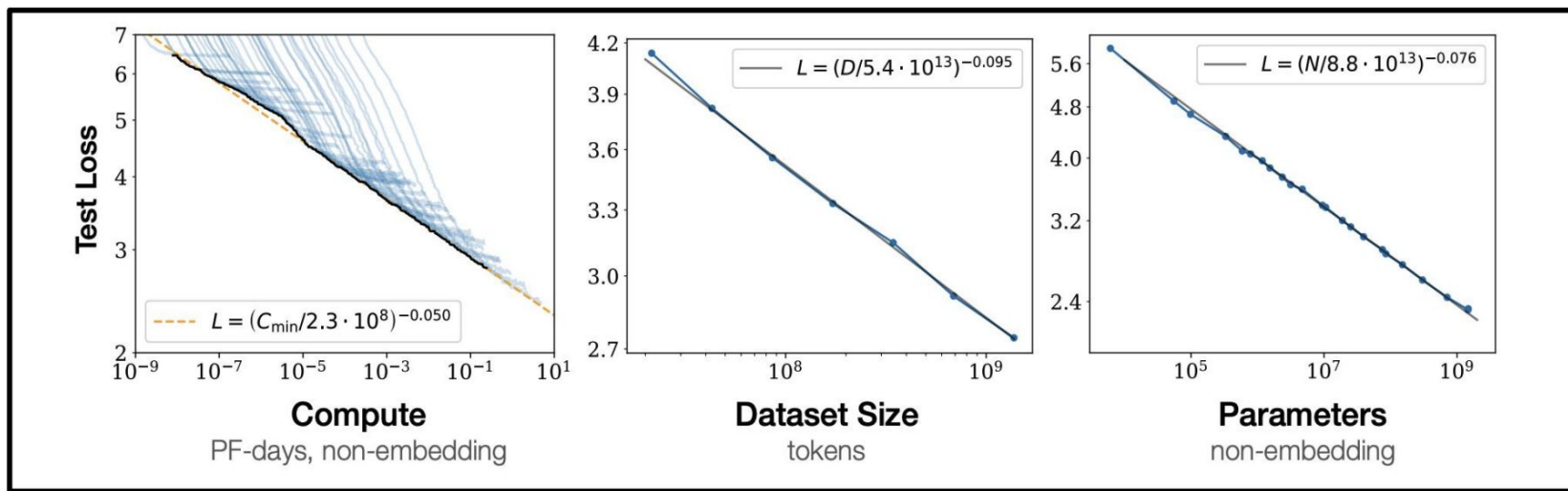
Floating-point operations, selected systems, by type, log scale



# Why LLMs?

- **Scaling Law for Neural Language Models**

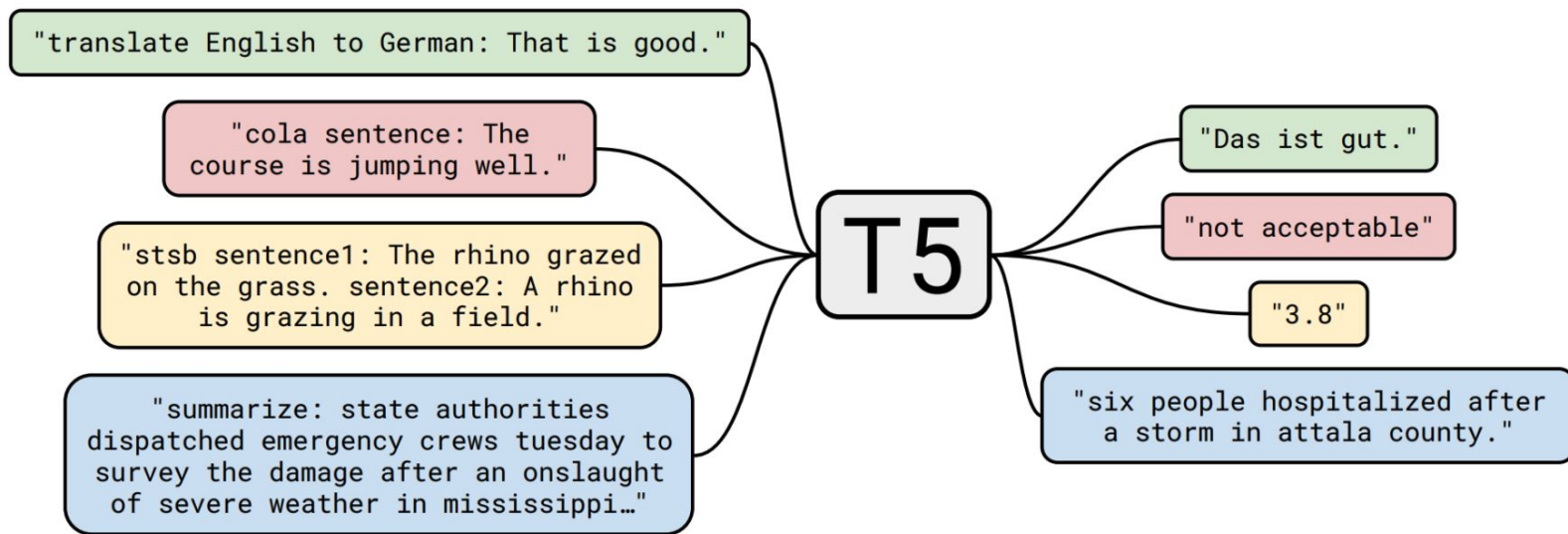
- Performance depends strongly on scale! We keep getting better performance as we scale the model, data, and compute up!



# Why LLMs?

- **Generalization**

- We can now use one single model to solve many NLP tasks

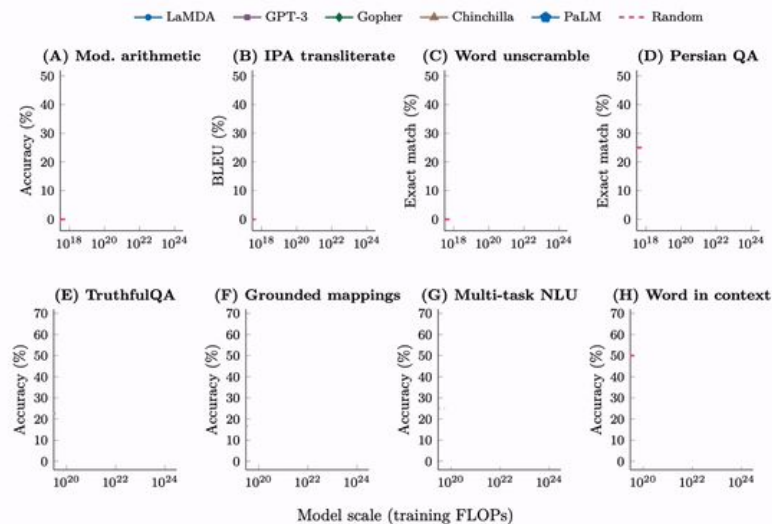




# Why LLMs?

- Emergent Abilities

- Some ability of LM is not present in smaller models but is present in larger models



# Emergent Capability - In-Context Learning

Traditional fine-tuning (not used for GPT-3)

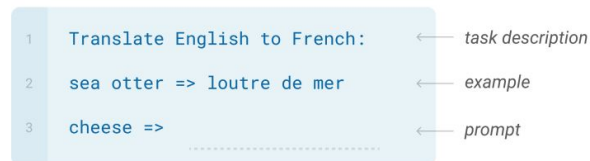
## Fine-tuning

The model is trained via repeated gradient updates using a large corpus of example tasks.



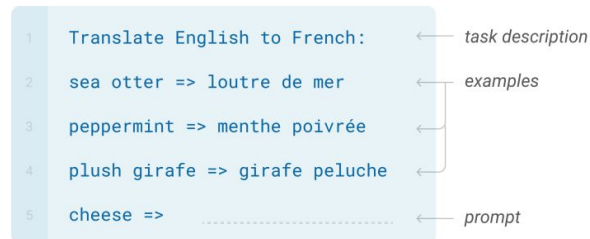
## One-shot

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



## Few-shot

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



<https://arxiv.org/pdf/2005.14165.pdf>

# Emergent Capability - In-Context Learning

Zero-shot  
(0s)

No Prompt

skicts = sticks

1-shot  
(1s)

chiar = chair  
skicts = sticks

Few-shot  
(FS)

chiar = chair  
[...]  
pciinc = picnic  
skicts = sticks

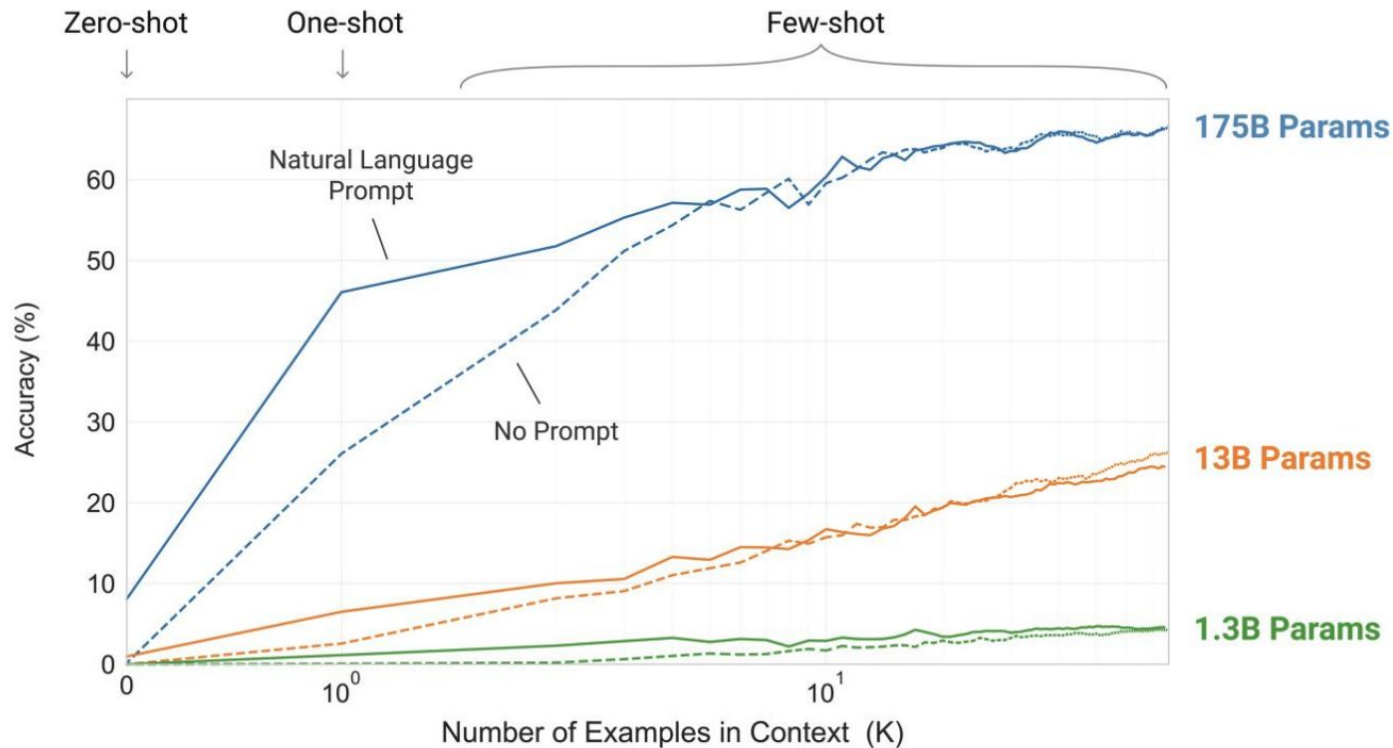
Prompt

Please unscramble the letters into  
a word, and write that word:  
skicts = sticks

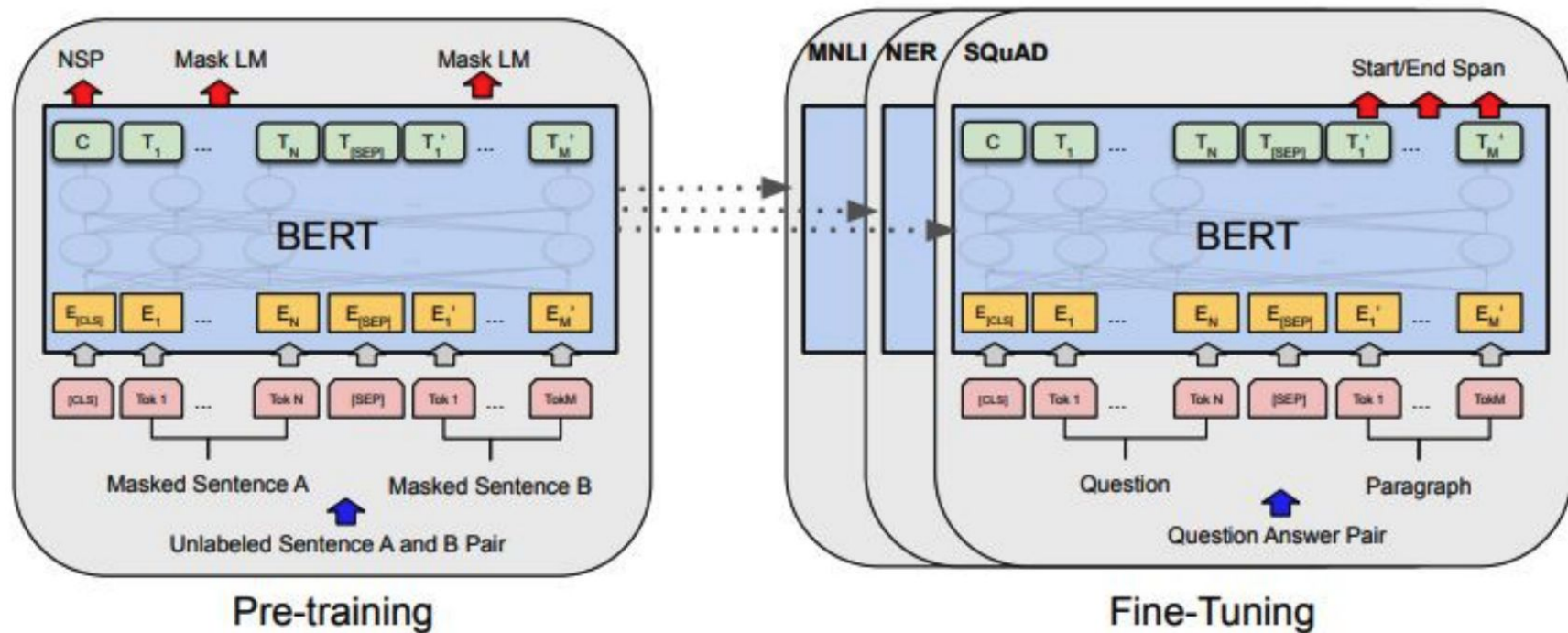
Please unscramble the letters into  
a word, and write that word:  
chiar = chair  
skicts = sticks

Please unscramble the letters into  
a word, and write that word:  
chiar = chair  
[...]  
pciinc = picnic  
skicts = sticks

# Emergent Capability - In-Context Learning



# Pretraining + Fine-tuning Paradigm



# Pretraining + Prompting Paradigm

- Fine-tuning (FT)
  - + Strongest performance
  - - Need curated and labeled dataset for each new task (typically 1k-100k ex.)
  - - Poor generalization, spurious feature exploitation
- Few-shot (FS)
  - + Much less task-specific data needed
  - + No spurious feature exploitation
  - - Challenging
- One-shot (1S)
  - + "Most natural," e.g. giving humans instructions
  - - Challenging
- Zero-shot (OS)
  - + Most convenient
  - - Challenging, can be ambiguous

**Stronger  
task-specific  
performance**



**More convenient,  
general, less data**

# Emergent Capability - Chain of Thoughts Prompting

## Standard Prompting

### Model 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: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

### Model Output

A: The answer is 27. ❌

## Chain-of-Thought Prompting

### Model 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 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

### Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had  $23 - 20 = 3$ . They bought 6 more apples, so they have  $3 + 6 = 9$ . The answer is 9. ✅



# Emergent Capability - Chain of Thoughts Prompting

## Math Word Problems (free response)

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 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.

## Math Word Problems (multiple choice)

Q: How many keystrokes are needed to type the numbers from 1 to 500?  
Answer Choices: (a) 1156 (b) 1392 (c) 1480 (d) 1562 (e) 1788

A: There are 9 one-digit numbers from 1 to 9. There are 90 two-digit numbers from 10 to 99. There are 401 three-digit numbers from 100 to 500.  $9 + 90(2) + 401(3) = 1392$ . The answer is (b).

## CSQA (commonsense)

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: Would a pear sink in water?

A: The density of a pear is about 0.6 g/cm<sup>3</sup>, which is less than water. Thus, a pear would float. 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.

## SayCan (Instructing a robot)

Human: How would you bring me something that isn't a fruit?

Explanation: the user wants something to eat that isn't a fruit. An energy bar is not a fruit, so I will bring the user an energy bar.

Plan: 1. find(energy bar) 2. pick(energy bar) 3. find(user) 4. put(energy bar) 5. done().

## Last Letter Concatenation

Q: Take the last letters of the words in "Lady Gaga" and concatenate them.

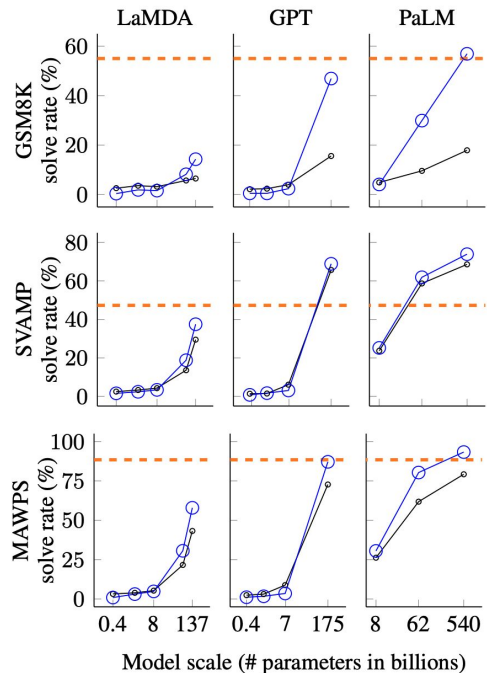
A: The last letter of "Lady" is "y". The last letter of "Gaga" is "a". Concatenating them is "ya". So the answer is ya.

## Coin Flip (state tracking)

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.

— Standard prompting  
— Chain-of-thought prompting  
— Prior supervised best





# Emergent Capability - Zero Shot CoT Prompting

(a) Few-shot

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: The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The answer is 8. ✗

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

(Output) 8 ✗

(b) Few-shot-CoT

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 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are  $16 / 2 = 8$  golf balls. Half of the golf balls are blue. So there are  $8 / 2 = 4$  blue golf balls. The answer is 4. ✓

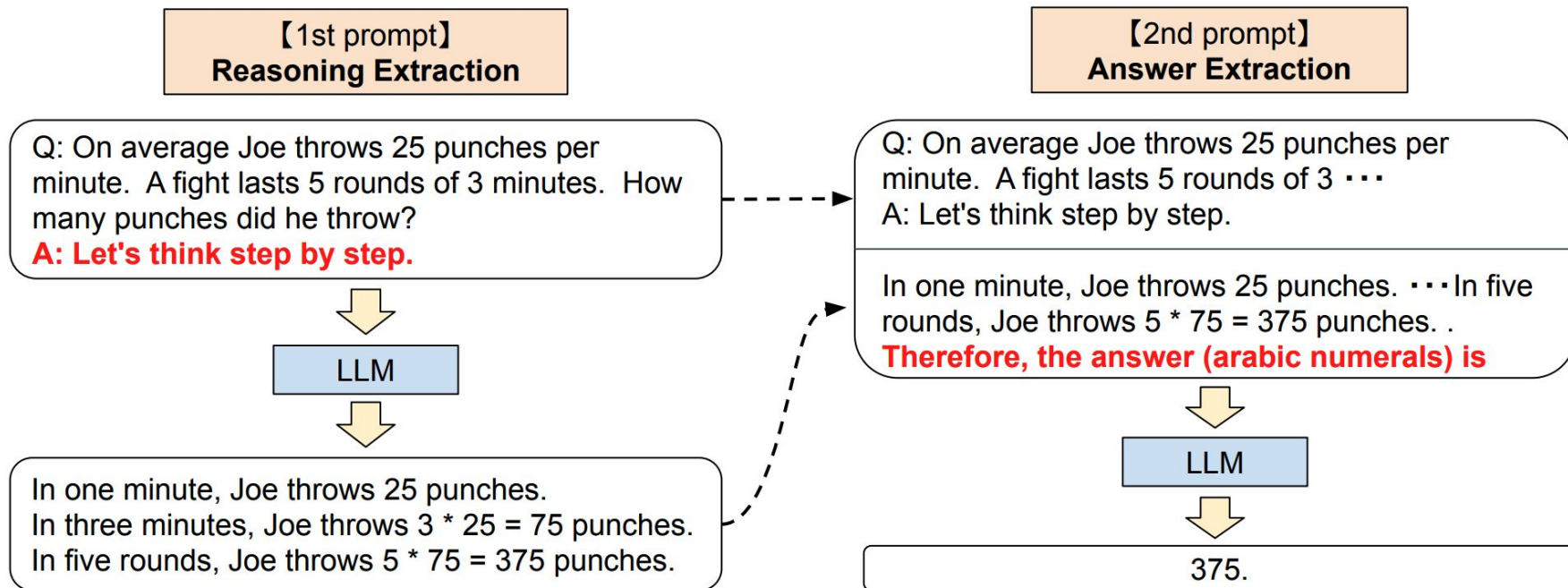
(d) Zero-shot-CoT (Ours)

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

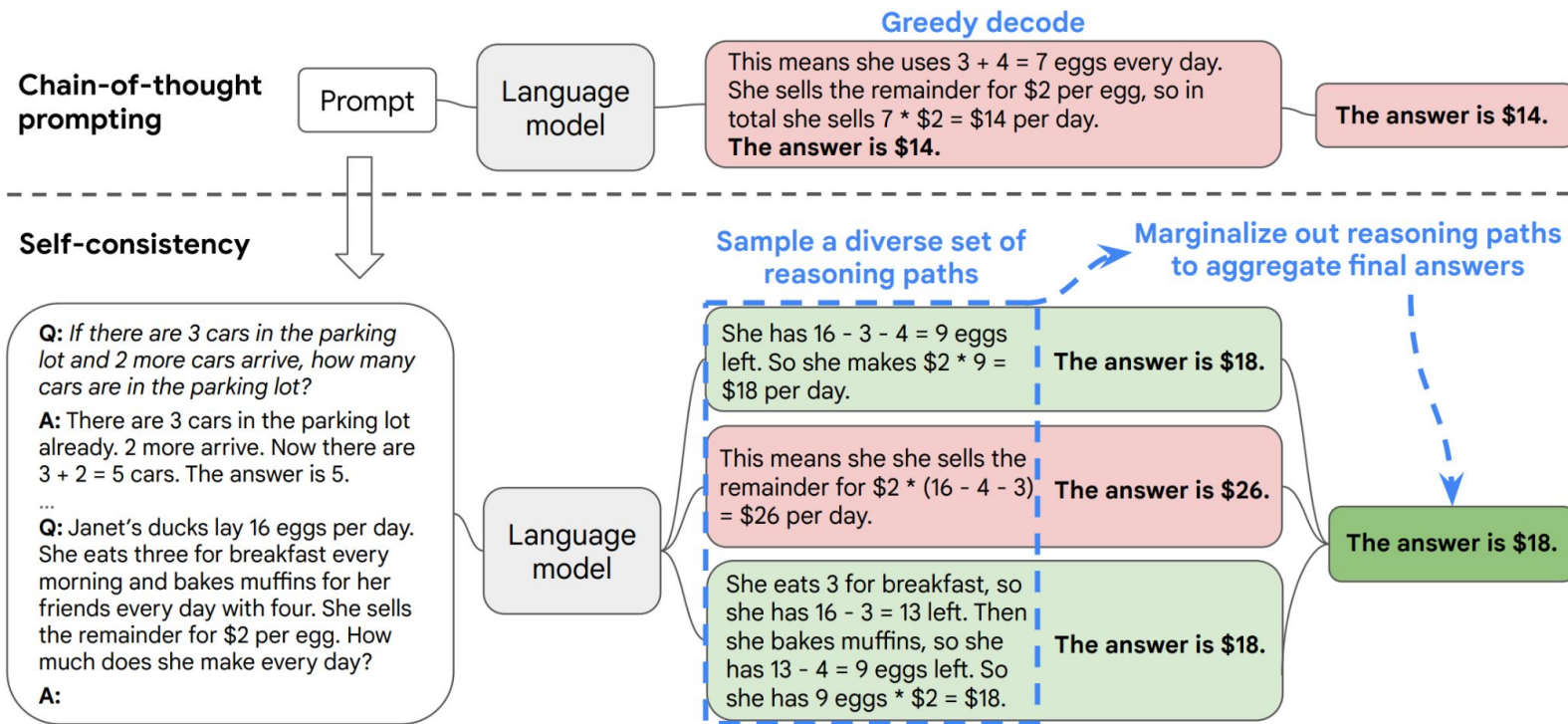
A: **Let's think step by step.**

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓

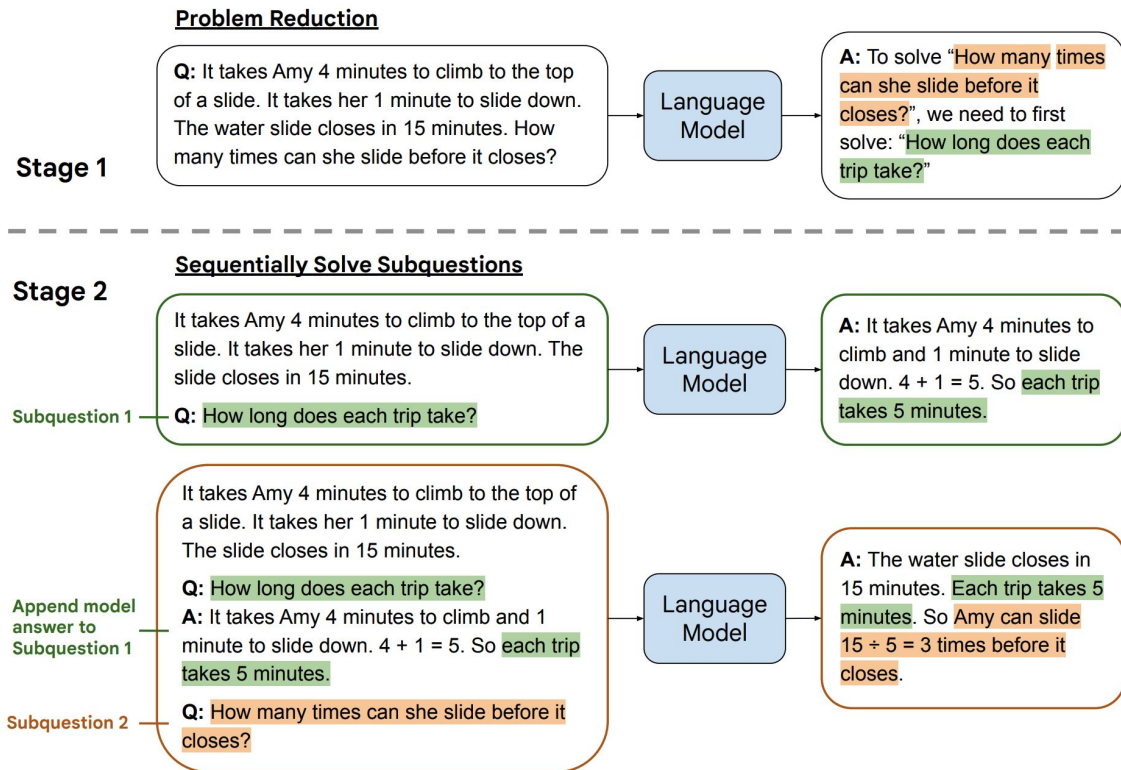
# Emergent Capability - Zero Shot CoT Prompting



# Emergent Capability - Self-Consistency Prompting



# Emergent Capability - **Least-to-Most Prompting**



# Emergent Capability - Augmented Prompting Abilities

## Advanced Prompting Techniques

- Zero-shot CoT Prompting
- Self-Consistency
- Divide-and-Conquer

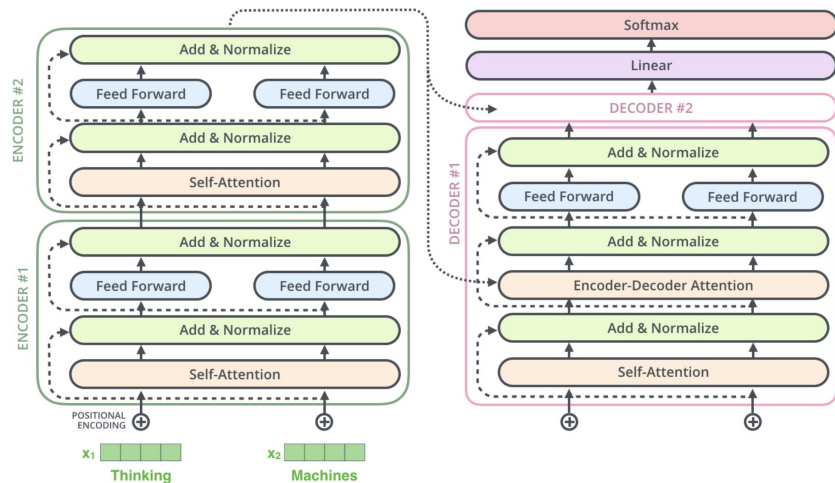
## Ask a human to

- Explain the rationale
- Double check the answer
- Decompose to easy subproblems

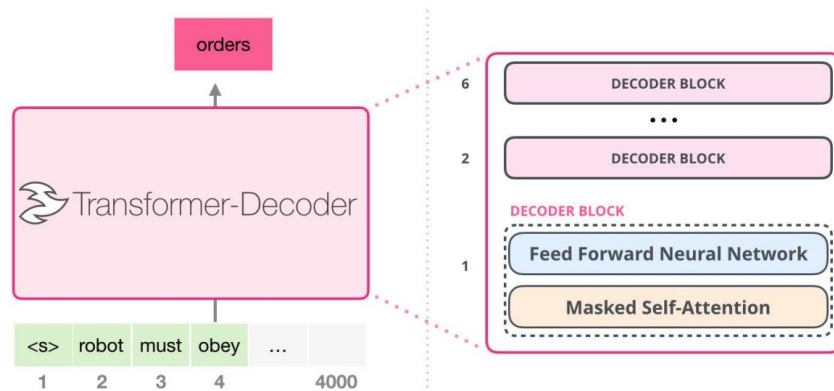
Large Language Models demonstrate some human-like behaviors!

# Training Architectures

## Encoder-decoder models (T5, BART)



## Decoder-only models (GPT-X, PaLM)



# Training Objectives - UL2

## R-Denoising

Inputs:

[R] He dealt in archetypes before anyone knew such things existed, and his 3 to take an emotion or a situation 5 it to the limit helped create a cadre of plays that have been endlessly 4 - and copied. Apart from this, Romeo and Juliet inspired Malorie Blackman's Noughts 5 there are references to Hamlet in Lunar Park by Bret Easton Ellis 2 The Tempest was the cue for The Magus by John Fowles.

Target:

<B> 3 <S> 5 <S> 4 <S> 5  
<S> 2 <E>

## S-Denoising

Inputs:

[S] He dealt in archetypes before anyone knew such things existed, and his ability to take an emotion or a situation and push it to the limit helped create a cadre of plays that have been endlessly staged - and copied. Apart from this, Romeo and Juliet 95

Target:

<B> 95  
<E>

## X-Denoising

Inputs:

[X] He dealt in archetypes be 16 things existed, and his ability to take an emotion or a situation 32 plays that have been endlessly staged - and copied. Apart from 24 Malorie Blackman's Noughts & Crosses, there are references to Hamlet in Lunar 24 Tempest was the cue for The Magus by John Fowles.

Target:

<B> 16 <S> 32 <S> 24 <S> 24 <S> <E>

Inputs:

[X] He dealt in archetypes 3 anyone knew such things existed, a 3 ability to take an 5 situation and push it to the limit helped 4 cadre of plays 4 been endlessly staged - and 5 Apart from this, Romeo and Juliet inspired Malorie Blackman's 5 Crosses, 3 are references to Hamlet in 3 Park by Bret Easton 2 and 4 4 was the 2 for The 4 by John 5

Target:

<B> 3 <S> 3 <S> 5 <S> 4 <S> 4 <S> 5 <S> 5 <S> 3 <S> 3 <S> 2 <S> 4 <S> 4 <S> 2 <S> 4 <S> 5 <E>

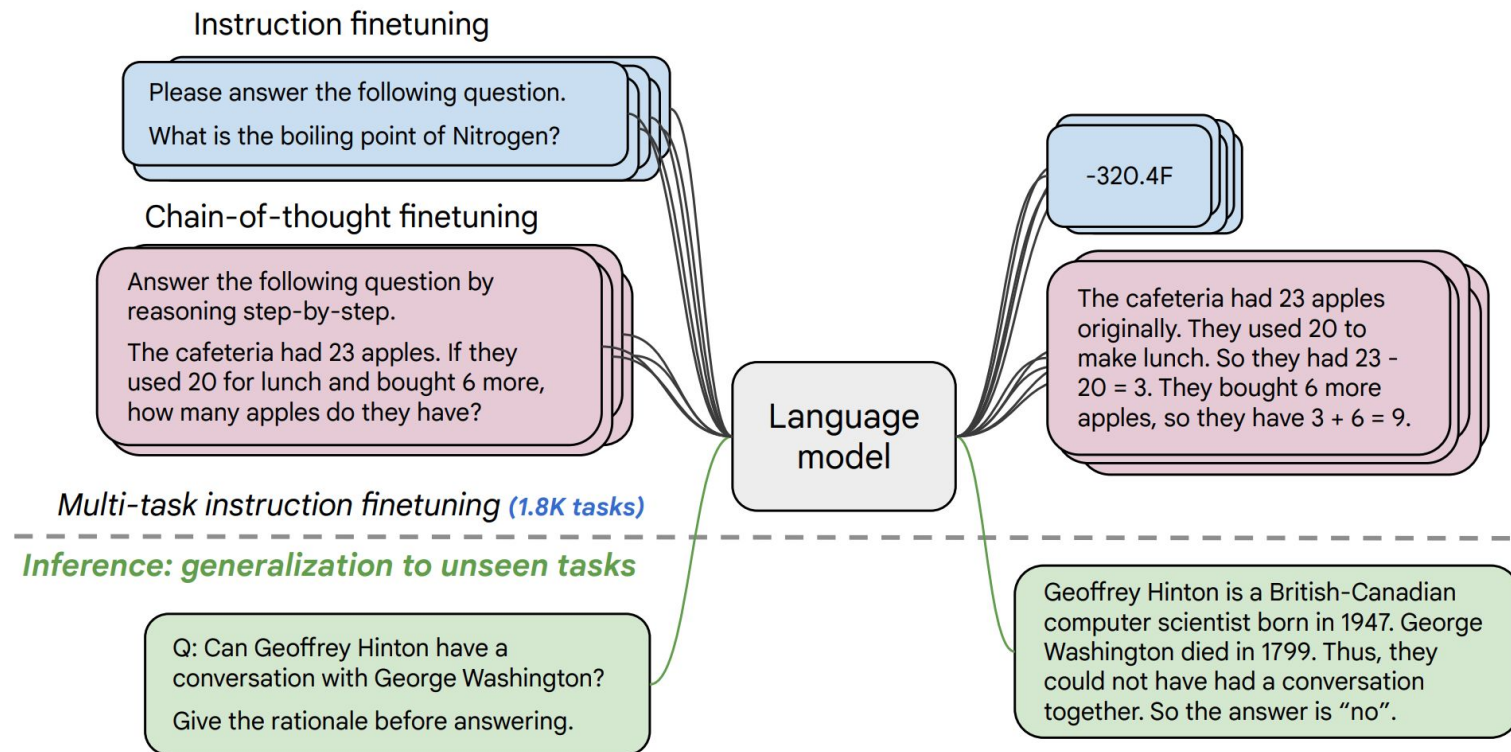


# What kinds of things does pretraining learn?

- *Stanford University is located in \_\_\_\_\_, California.* [Trivia]
- *I put \_\_\_\_ fork down on the table.* [syntax]
- *The woman walked across the street, checking for traffic over \_\_\_\_ shoulder.* [coreference]
- *I went to the ocean to see the fish, turtles, seals, and \_\_\_\_.* [lexical semantics/topic]
- *Overall, the value I got from the two hours watching it was the sum total of the popcorn and the drink. The movie was \_\_\_\_.* [sentiment]
- *Iroh went into the kitchen to make some tea. Standing next to Iroh, Zuko pondered his destiny. Zuko left the \_\_\_\_.* [some reasoning – this is harder]
- *I was thinking about the sequence that goes 1, 1, 2, 3, 5, 8, 13, 21, \_\_\_\_* [some basic arithmetic; they don't learn the Fibonacci sequence]



# Finetune - **Instruction Finetune**



# Finetune - **Instruction Finetune**

## Finetuning tasks

### TO-SF

Commonsense reasoning  
Question generation  
Closed-book QA  
Adversarial QA  
Extractive QA  
Title/context generation  
Topic classification  
Struct-to-text  
...

*55 Datasets, 14 Categories,  
193 Tasks*

### Muffin

Natural language inference	Closed-book QA
Code instruction gen.	Conversational QA
Program synthesis	Code repair
Dialog context generation	...

*69 Datasets, 27 Categories, 80 Tasks*

### CoT (Reasoning)

Arithmetic reasoning	Explanation generation
Commonsense Reasoning	Sentence composition
Implicit reasoning	...

*9 Datasets, 1 Category, 9 Tasks*

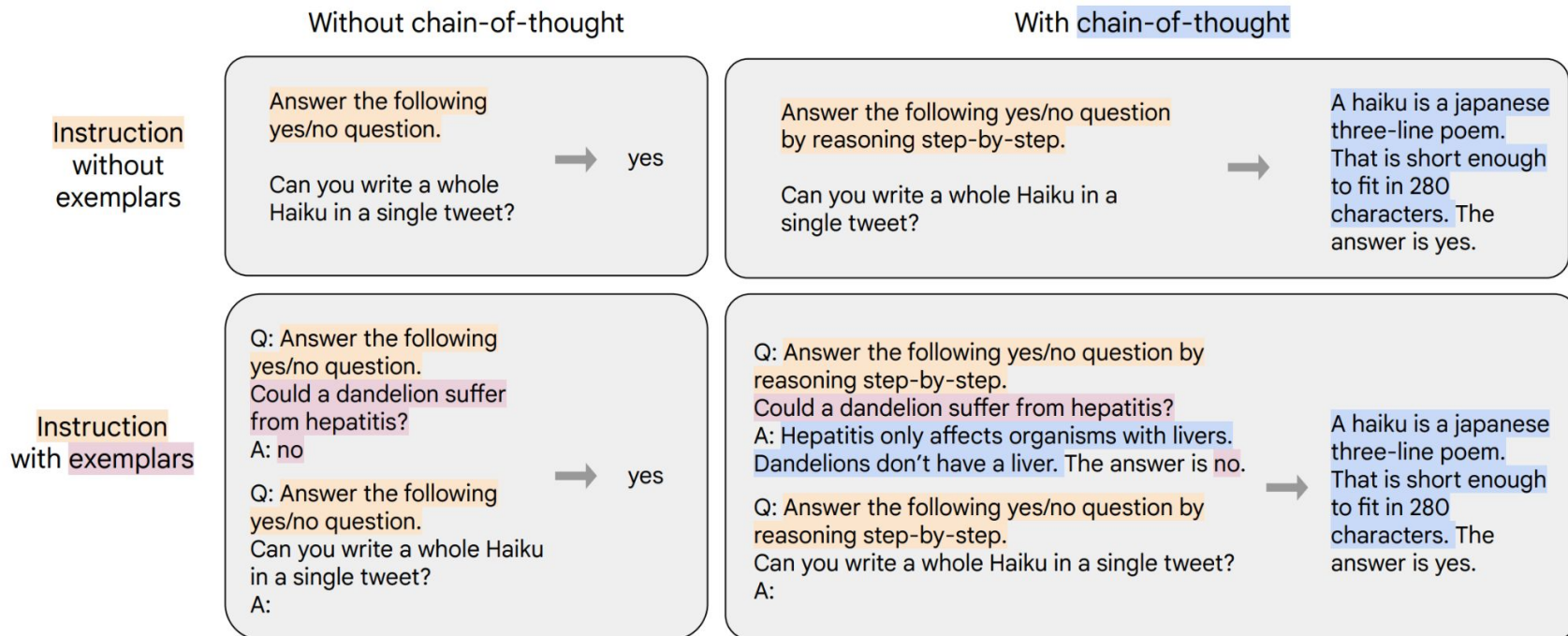
### Natural Instructions v2

Cause effect classification  
Commonsense reasoning  
Named entity recognition  
Toxic language detection  
Question answering  
Question generation  
Program execution  
Text categorization  
...

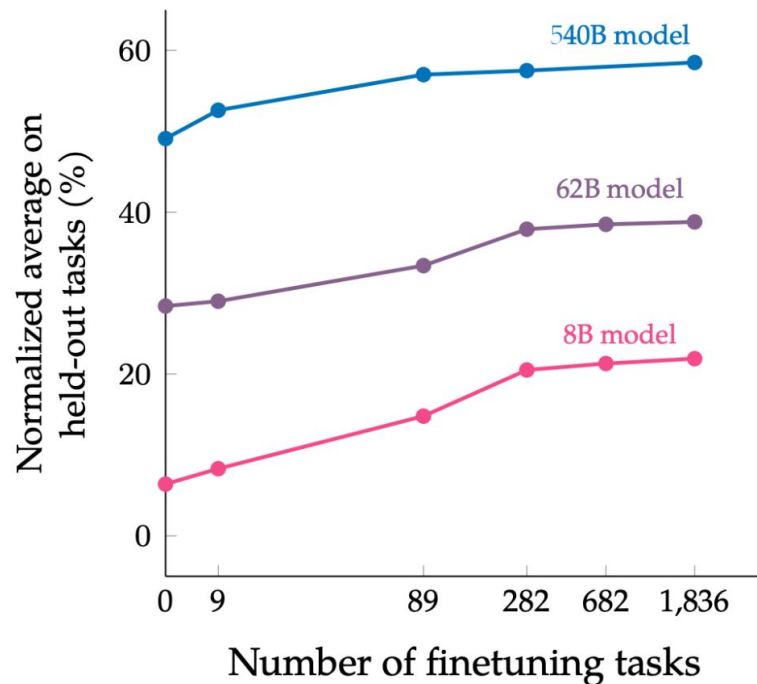
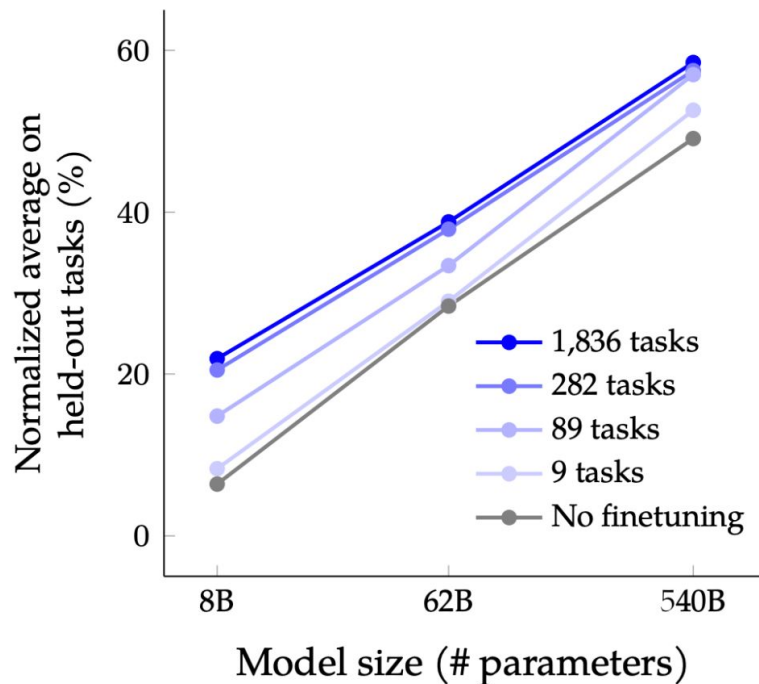
*372 Datasets, 108 Categories,  
1554 Tasks*

- ❖ A **Dataset** is an original data source (e.g. SQuAD).
- ❖ A **Task Category** is unique task setup (e.g. the SQuAD dataset is configurable for multiple task categories such as extractive question answering, query generation, and context generation).
- ❖ A **Task** is a unique <dataset, task category> pair, with any number of templates which preserve the task category (e.g. query generation on the SQuAD dataset.)

# Finetune - Instruction Finetune



# Finetune - Instruction Finetune

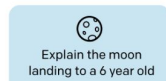


# Finetune - RLHF

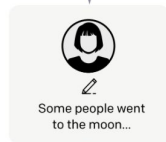
## Step 1

**Collect demonstration data,  
and train a supervised policy.**

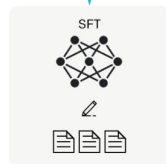
A prompt is  
sampled from our  
prompt dataset.



A labeler  
demonstrates the  
desired output  
behavior.



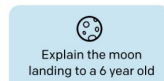
This data is used  
to fine-tune GPT-3  
with supervised  
learning.



## Step 2

**Collect comparison data,  
and train a reward model.**

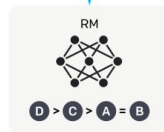
A prompt and  
several model  
outputs are  
sampled.



A labeler ranks  
the outputs from  
best to worst.



This data is used  
to train our  
reward model.



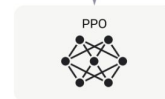
## Step 3

**Optimize a policy against  
the reward model using  
reinforcement learning.**

A new prompt  
is sampled from  
the dataset.



The policy  
generates  
an output.



Once upon a time...

The reward model  
calculates a  
reward for  
the output.



The reward is  
used to update  
the policy  
using PPO.

$r_k$

# Application - ChatGPT

## ChatGPT



### Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



### Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



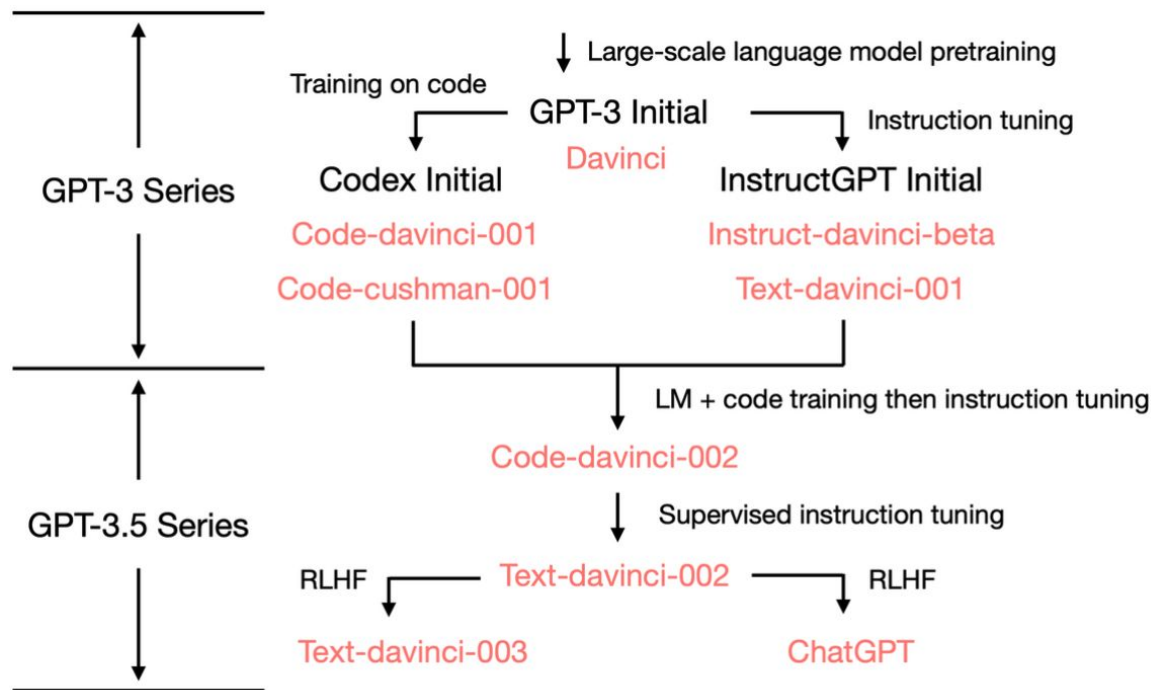
### Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

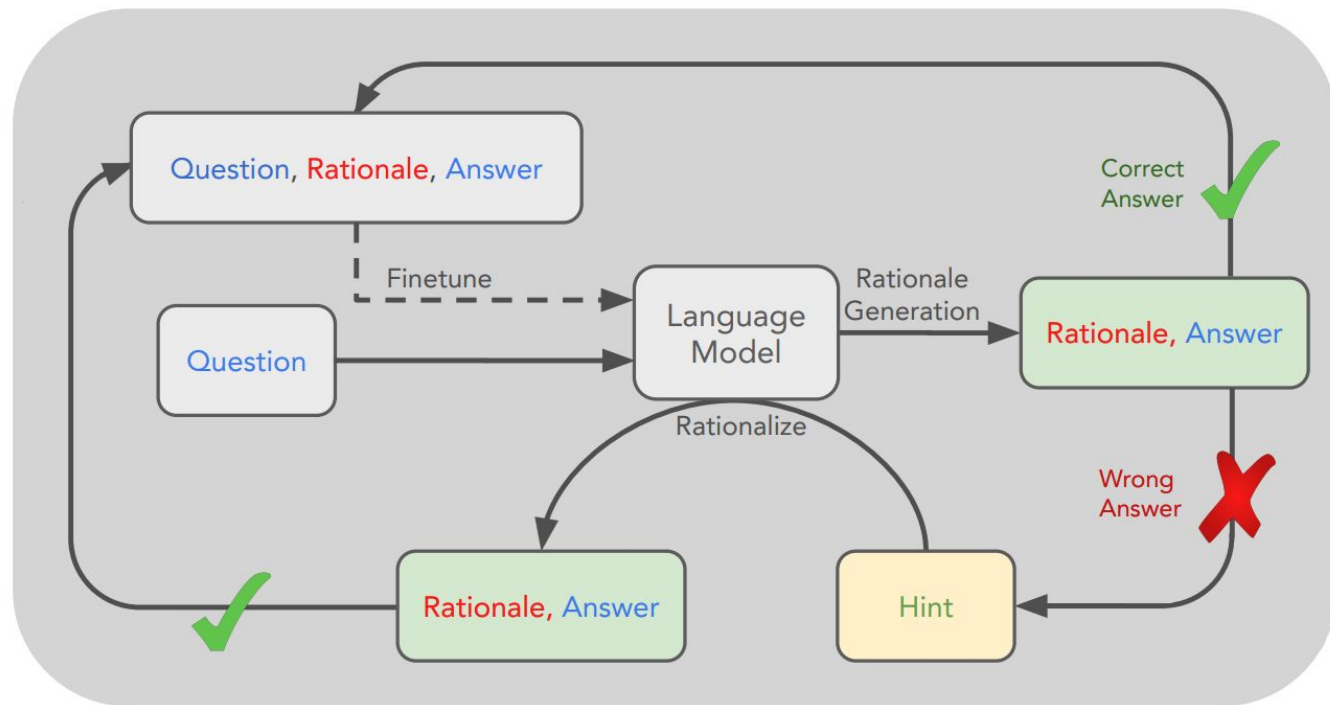
Limited knowledge of world and events after 2021

# Application - ChatGPT





# Finetune - Bootstrapping



Q: What can be used to carry a small dog?

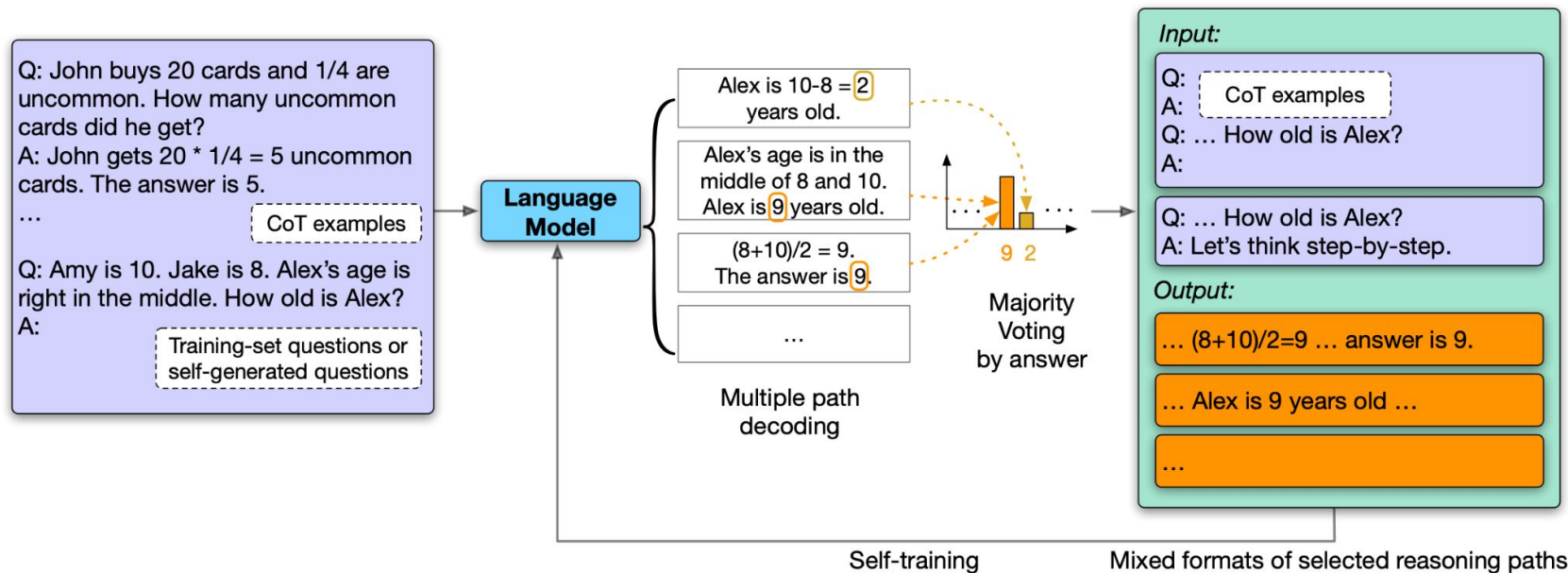
Answer Choices:

- (a) swimming pool
- (b) basket
- (c) dog show
- (d) backyard
- (e) own home

A: The answer must be something that can be used to carry a small dog. Baskets are designed to hold things. Therefore, the answer is basket (b).



# Finetune - Bootstrapping



# Large Language models Risks

- LLMs make mistakes  
(falsehoods, hallucinations)
- LLMs can be misused  
(misinformation, spam)
- LLMs can cause harms  
(toxicity, biases, stereotypes)
- LLMs can be attacked  
(adversarial examples, poisoning, prompt injection)
- LLMs can be useful as defenses  
(content moderation, explanations)



## Large language models associate Muslims with violence

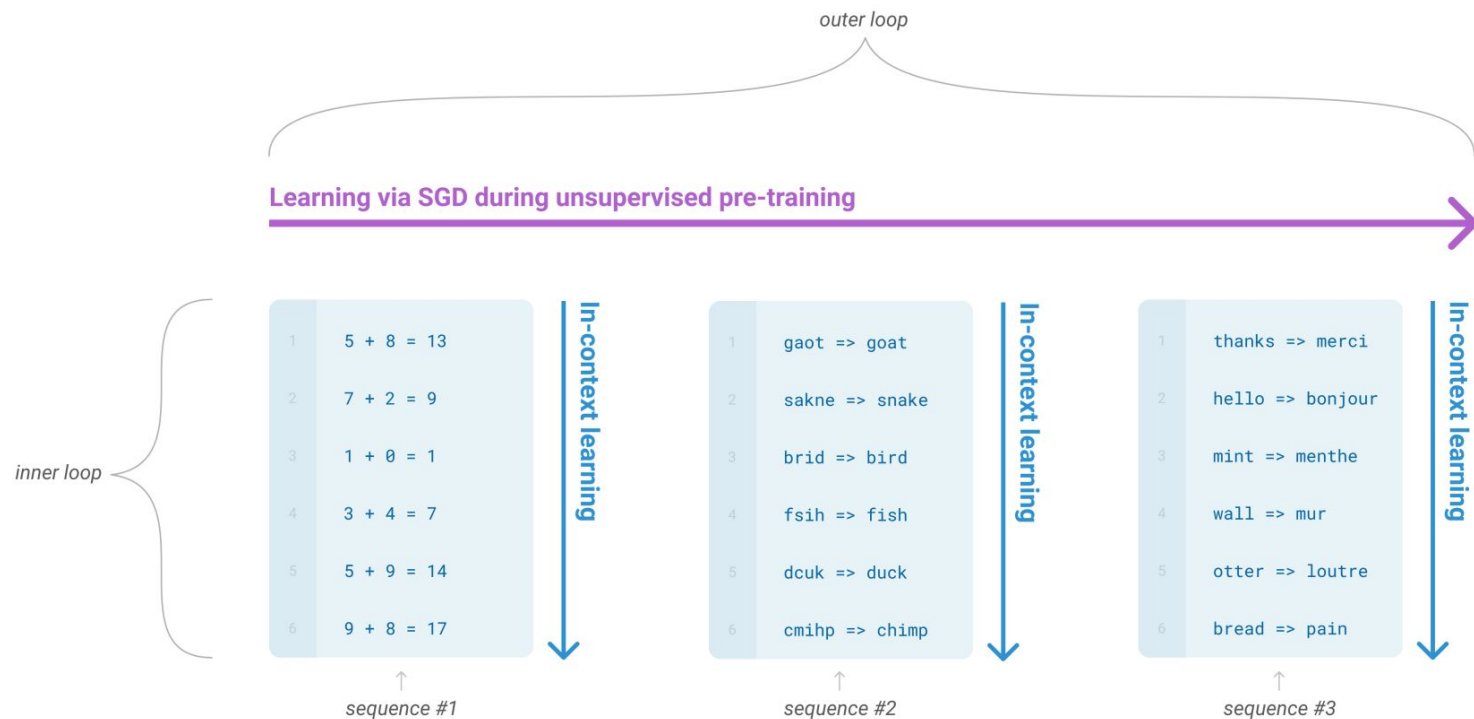
[Abubakar Abid](#), [Maheen Farooqi](#) & [James Zou](#) 

[Nature Machine Intelligence](#) **3**, 461–463 (2021) | [Cite this article](#)

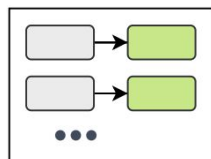
## Resources for further reading

- <https://web.stanford.edu/class/cs224n/>
- <https://stanford-cs324.github.io/winter2022/>
- <https://stanford-cs324.github.io/winter2023/>
- <https://www.cs.princeton.edu/courses/archive/fall22/cos597G/>
- <https://rycolab.io/classes/llm-s23/>
- <https://yaofu.notion.site/How-does-GPT-Obtain-its-Ability-Tracing-Emergent-Abilities-of-Language-Models-to-their-Sources-b9a57ac0fcf74f30a1ab9e3e36fa1dc1>
- <https://www.jasonwei.net/blog/emergence>

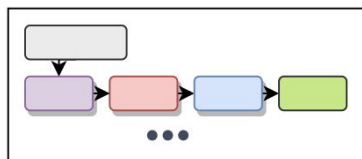
# Emergent Capability - In-Context Learning



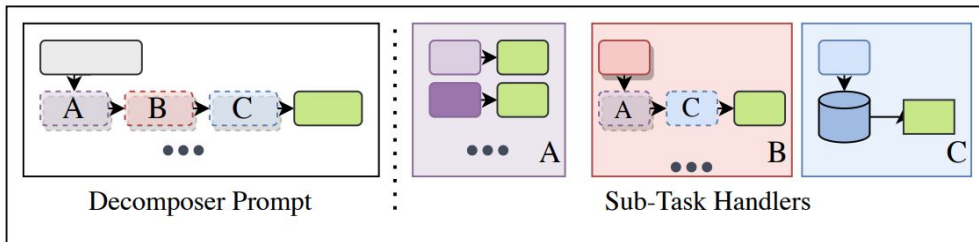
# Emergent Capability - Decomposed Prompting



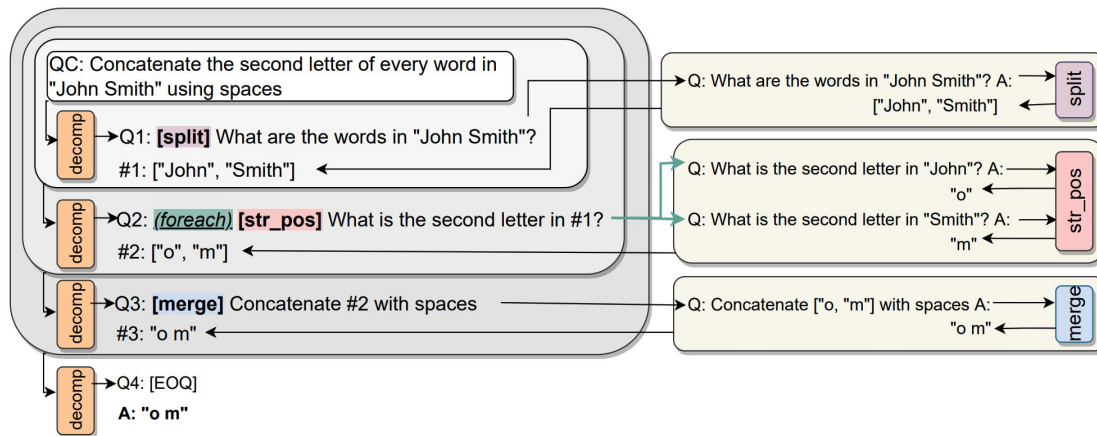
Standard Prompting



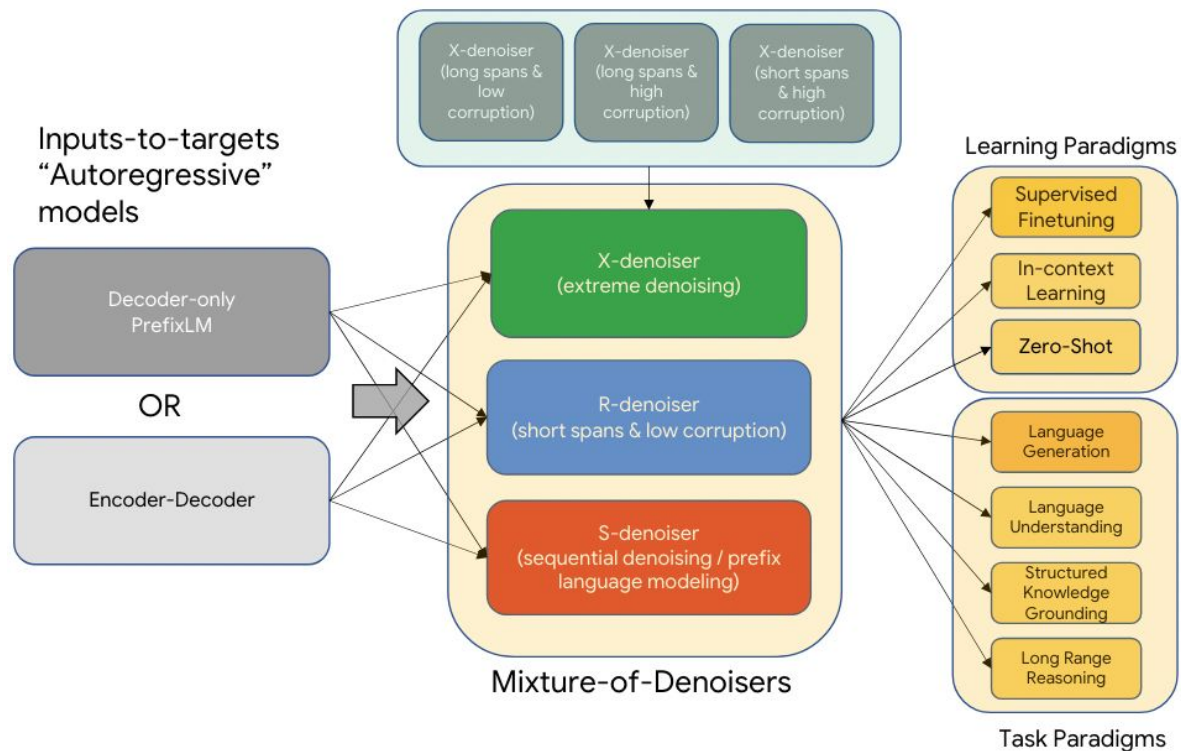
Chain-of-Thought Prompting



Decomposed Prompting

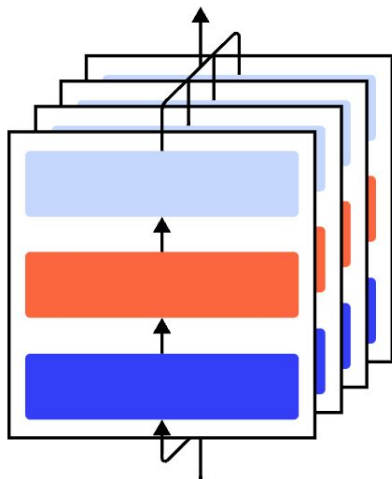


# Training Objectives - UL2

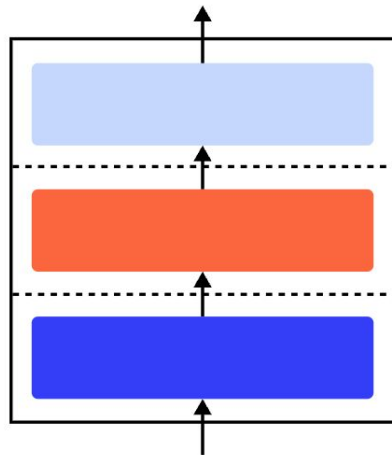


# Training Techniques - **Parallelism**

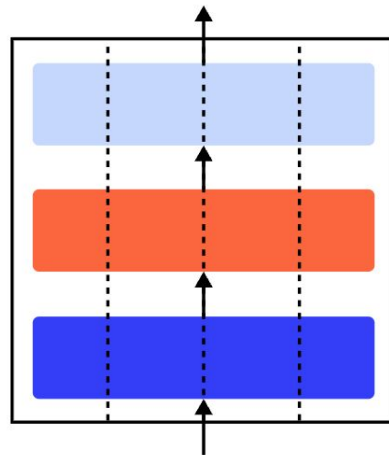
Data Parallelism



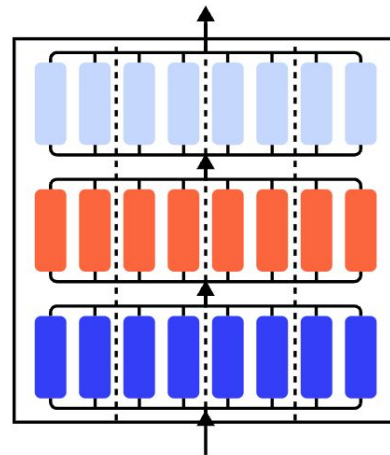
Pipeline Parallelism



Tensor Parallelism



Expert Parallelism



An illustration of various parallelism strategies on a three-layer model. Each color refers to one layer and dashed lines separate different GPUs.

# Training Techniques - **Parallelism**

