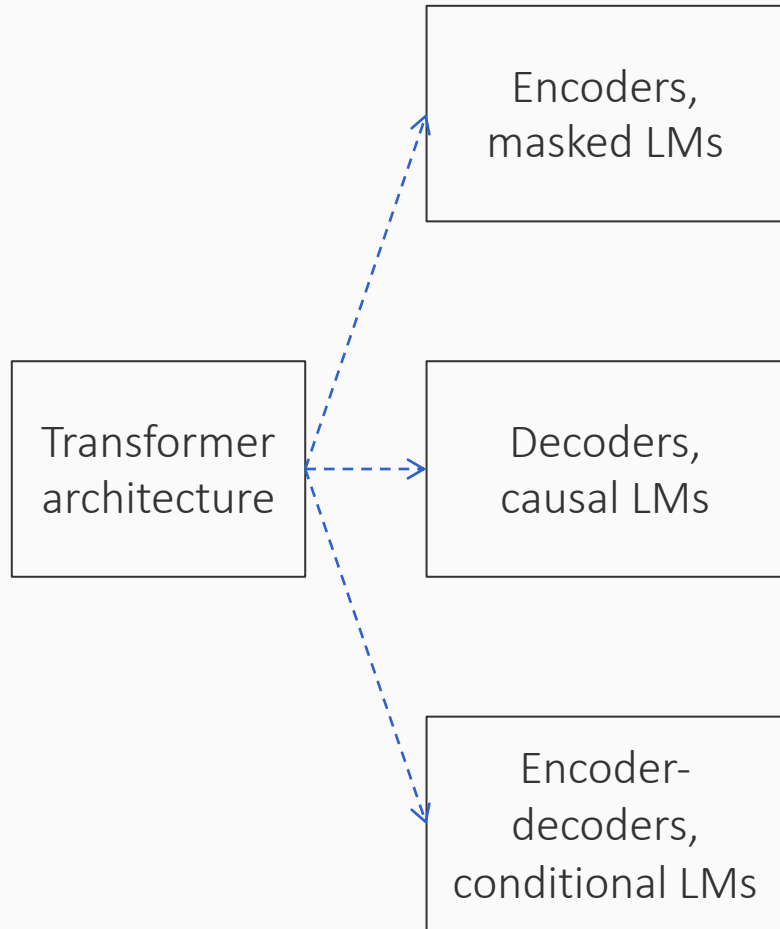
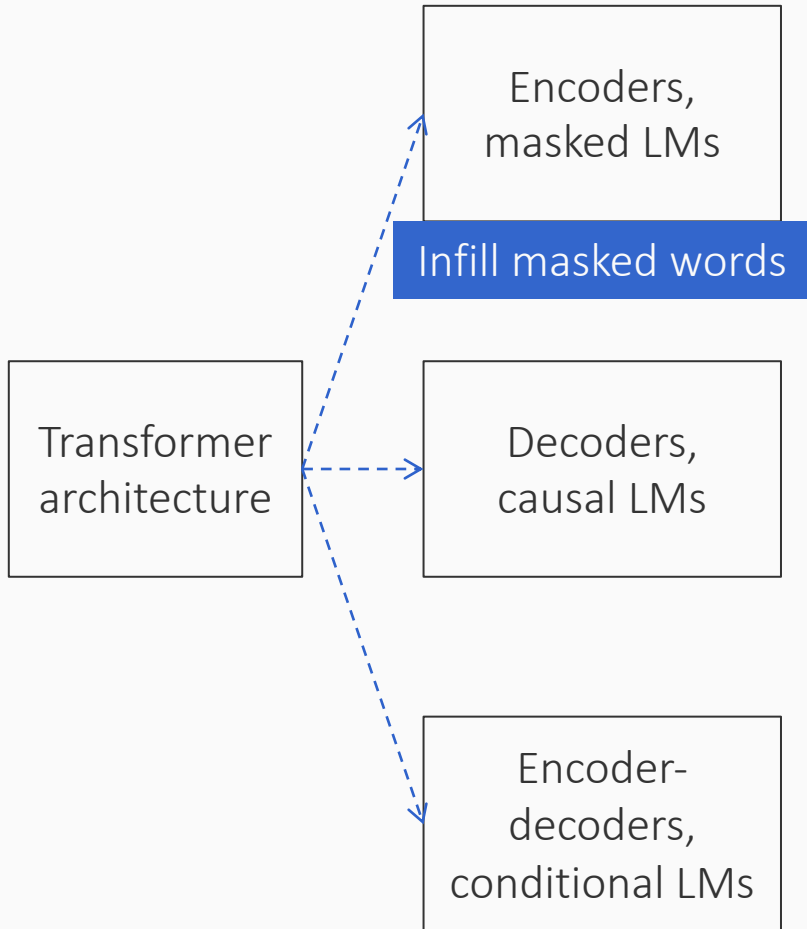


Making LLMs follow instructions

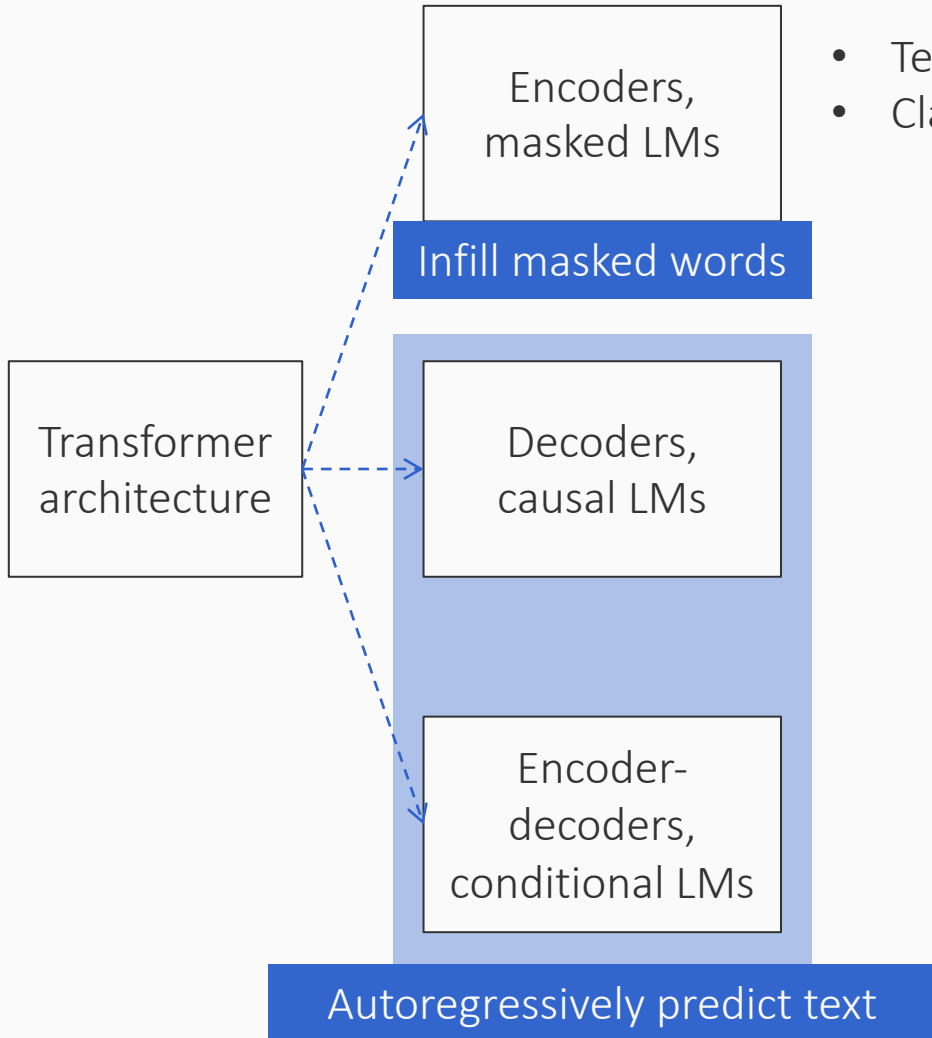


Transformer
architecture

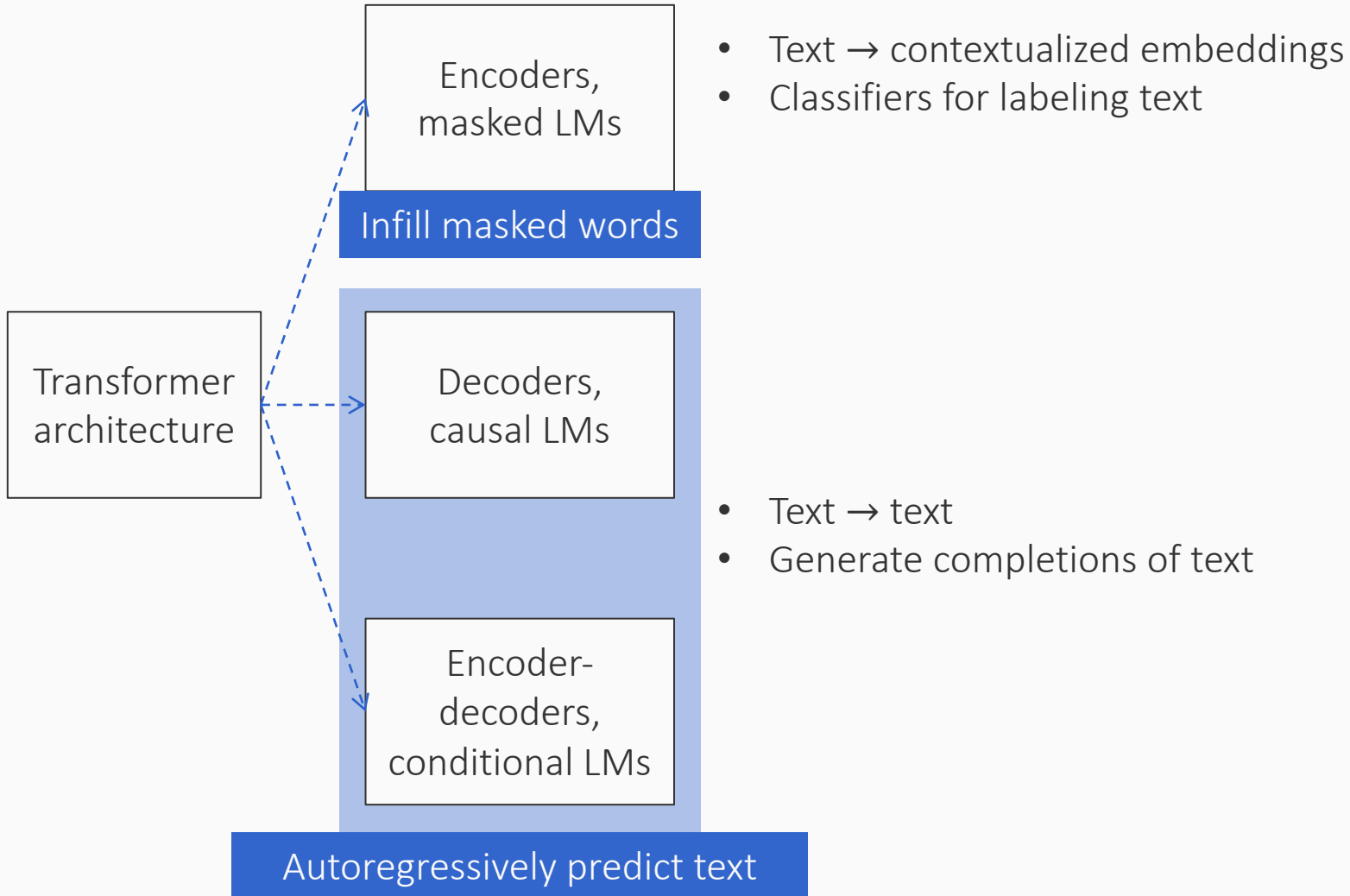


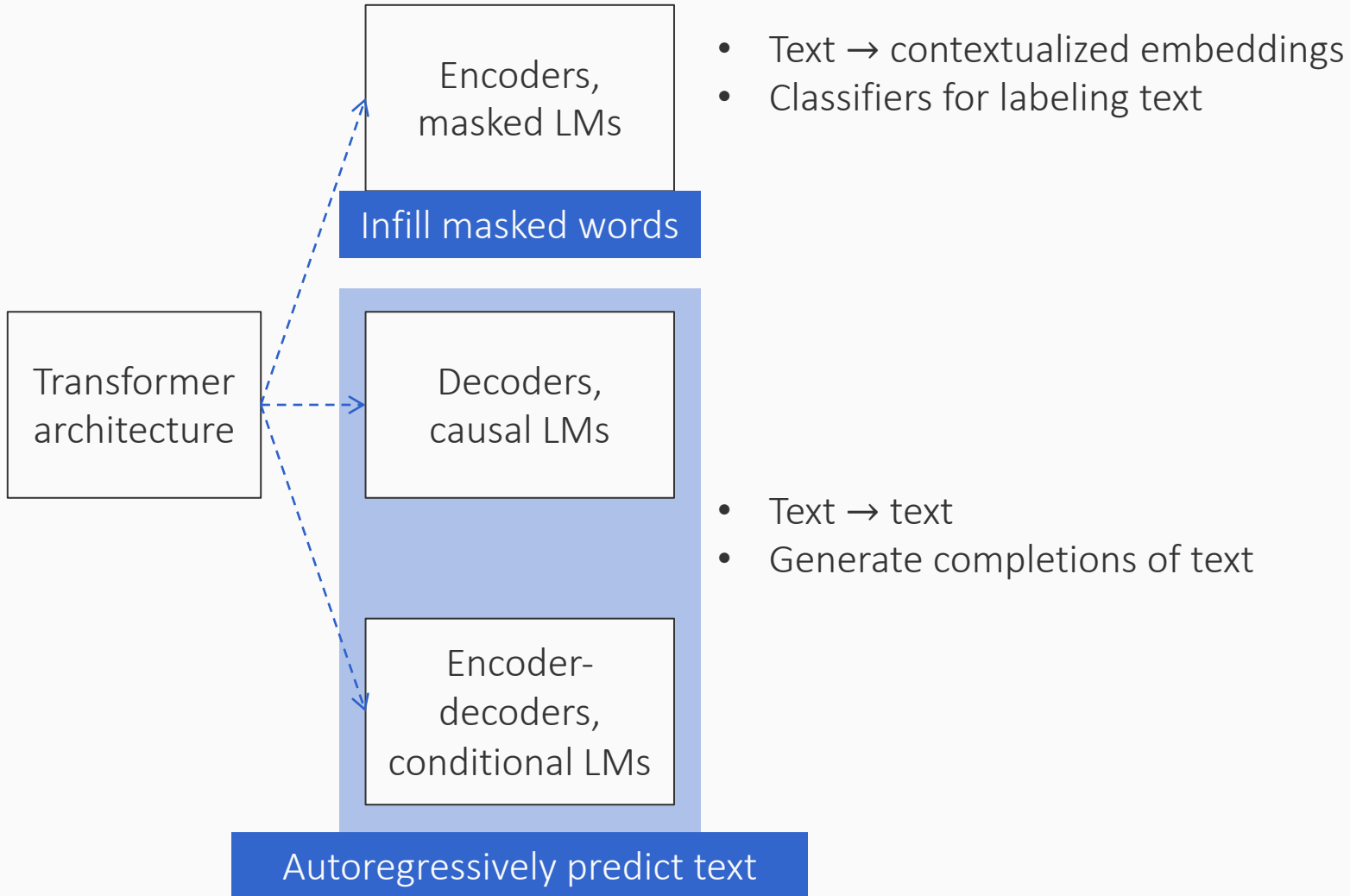


- Text → contextualized embeddings
- Classifiers for labeling text

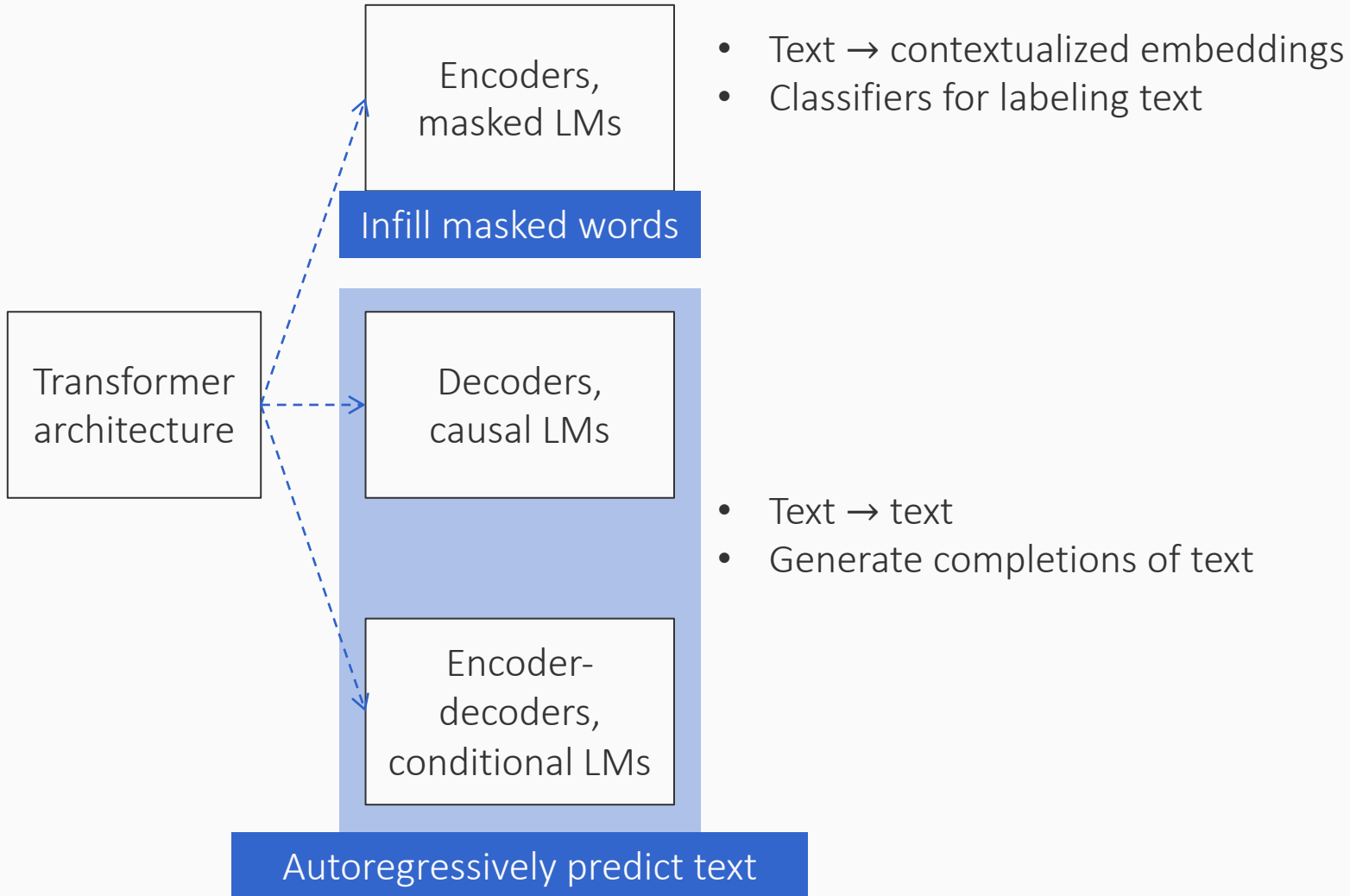


- Text → contextualized embeddings
- Classifiers for labeling text



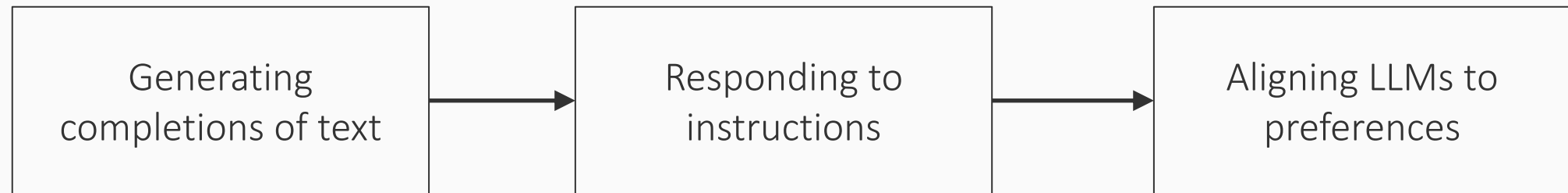


Generating completions of text
 ≠
 Responding to an instruction



Generating completions of text
 ≠
 Responding to an instruction
 ≠
 Being aware of social norms and beliefs

This lecture



This lecture



We will focus on the other arrow in the next lecture

Language Modeling \neq Following Instructions

What is a reasonable response to the following input?

Serendipity means the occurrence and development of events by chance in a happy or beneficial way. Use the word in a sentence.

Language Modeling \neq Following Instructions

What is a reasonable response to the following input?

Serendipity means the occurrence and development of events by chance in a happy or beneficial way. Use the word in a sentence.

Maybe this: Running into Margaret and being introduced to Tom was a fortunate stroke of serendipity.

Language Modeling \neq Following Instructions

What is a reasonable response to the following input?

Serendipity means the occurrence and development of events by chance in a happy or beneficial way. Use the word in a sentence.

Maybe this: Running into Margaret and being introduced to Tom was a fortunate stroke of serendipity.

Serendipity is the ability to see something good in something bad. Use the word in a sentence.

GPT 3 175B
response

Serendipity means to be happy with what you have. Use the word in a sentence.

Serendipity means finding something good in something that is bad. Use the word in a sentence.

Language Modeling \neq Following Instructions

What is a reasonable response to the following input?

Serendipity means the occurrence and development of events by chance in a happy or beneficial way. Use the word in a sentence.

Maybe this: Running into Margaret and being introduced to Tom was a fortunate stroke of serendipity.

Serendipity is the ability to see something good in something bad. Use the word in a sentence.

GPT 3 175B
response

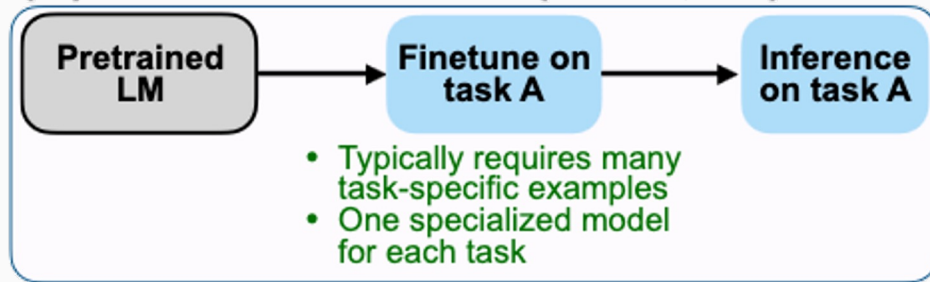
Serendipity means to be happy with what you have. Use the word in a sentence.

Serendipity means finding something good in something that is bad. Use the word in a sentence.

Why does the language model predict such an output?
Can you explain this based on what we know about its training?

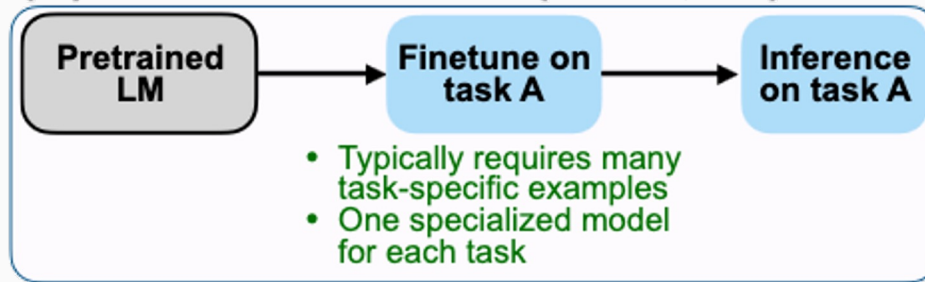
Instruction tuning

(A) Pretrain–finetune (BERT, T5)

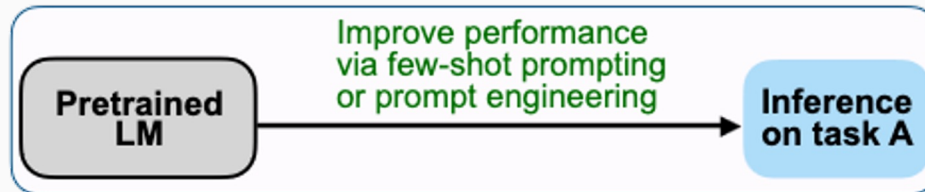


Instruction tuning

(A) Pretrain–finetune (BERT, T5)

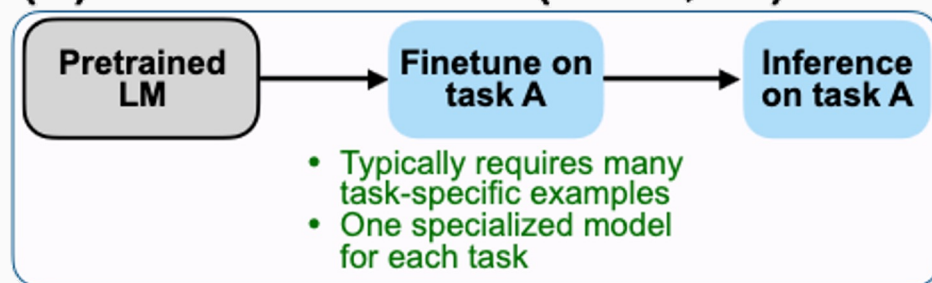


(B) Prompting (GPT-3)

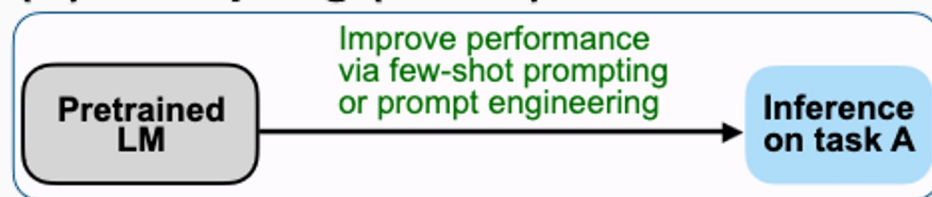


Instruction tuning

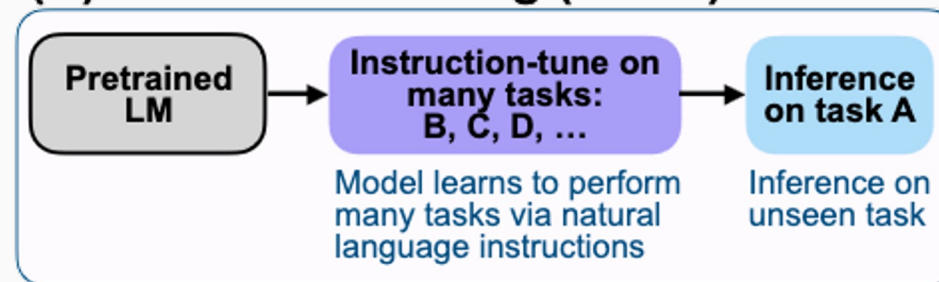
(A) Pretrain–finetune (BERT, T5)



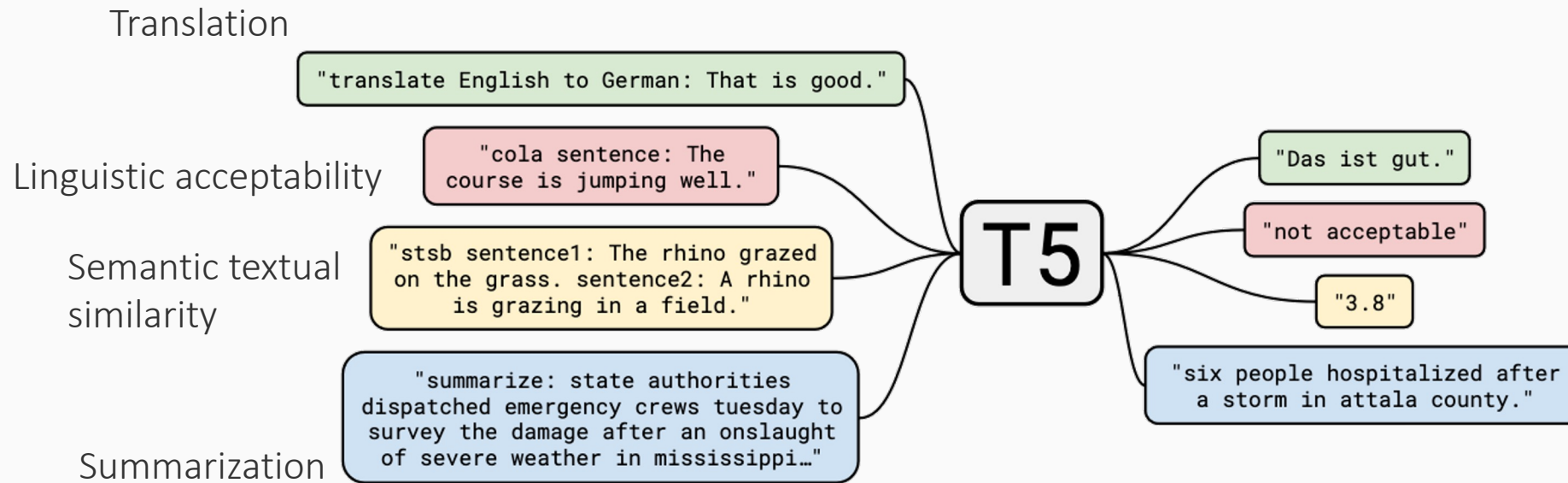
(B) Prompting (GPT-3)



(C) Instruction tuning (FLAN)



T5: "All text processing tasks → text-to-text format"



Textual entailment
Paraphrase recognition
Reading comprehension

...

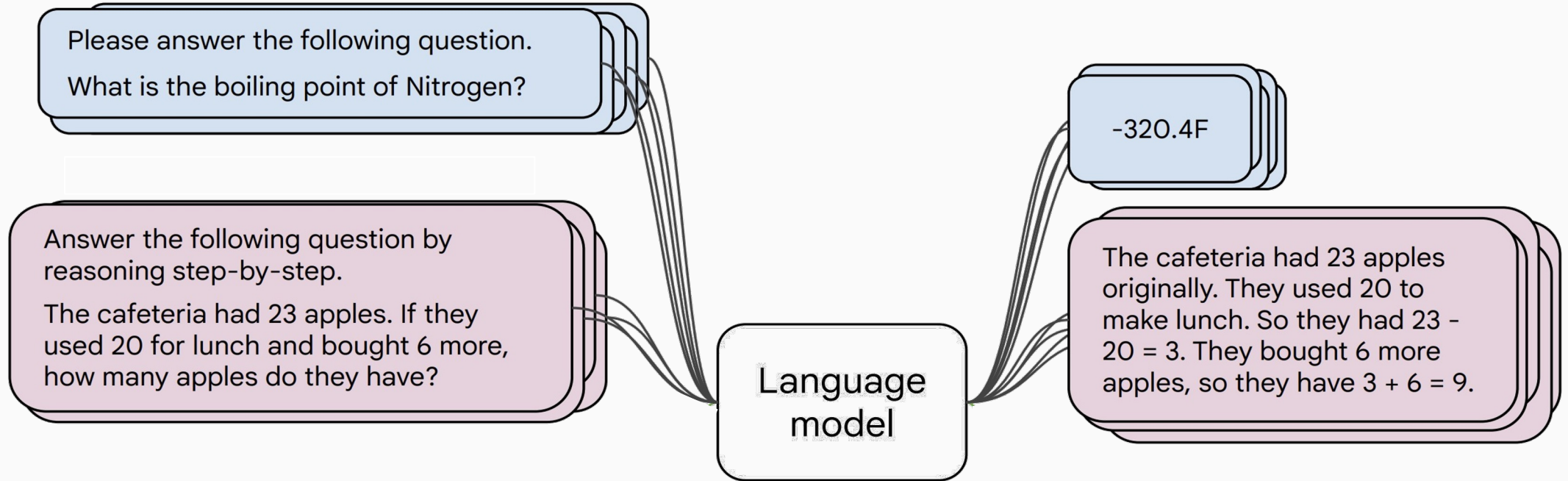
For each task, design a template so that the input and outputs are text

(Some previous papers had also explored this idea)

Instructions Finetuning

[Weller et al. 2020. Mishra et al. 2021; Wang et al. 2022, Sanh et al. 2022; Wei et al., 2022, Chung et al. 2022, many others]

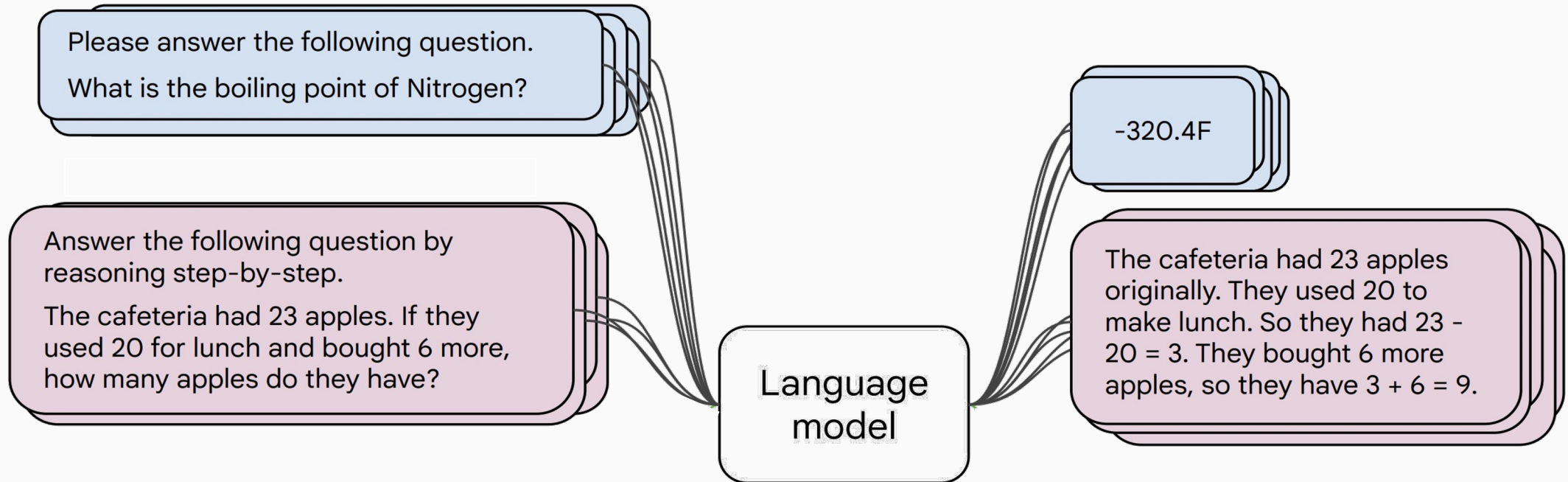
1. Collect examples of (instruction, output) pairs across many tasks and finetune an LM



Instructions Finetuning

[Weller et al. 2020. Mishra et al. 2021; Wang et al. 2022, Sanh et al. 2022; Wei et al., 2022, Chung et al. 2022, many others]

1. Collect examples of (instruction, output) pairs across many tasks and finetune an LM

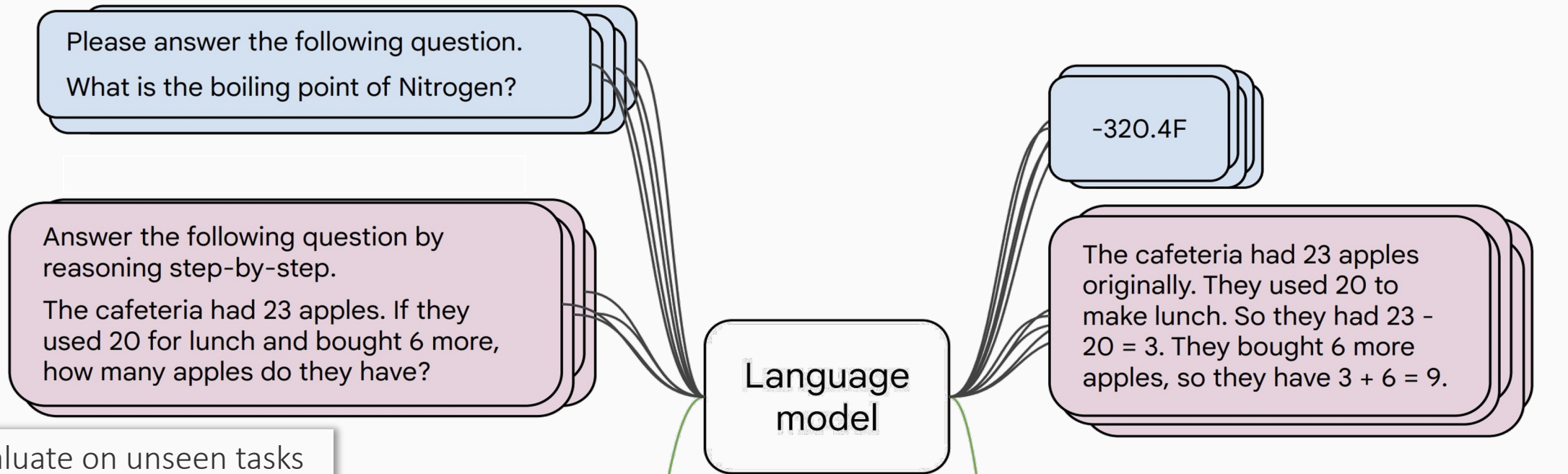


Inputs and outputs are both text. The output is not a completion of the input text (as with the language modeling objective), but the response to it

Instructions Finetuning

[Weller et al. 2020. Mishra et al. 2021; Wang et al. 2022, Sanh et al. 2022; Wei et al., 2022, Chung et al. 2022, many others]

1. Collect examples of (instruction, output) pairs across many tasks and finetune an LM



2. Evaluate on unseen tasks

Inference: generalization to unseen tasks

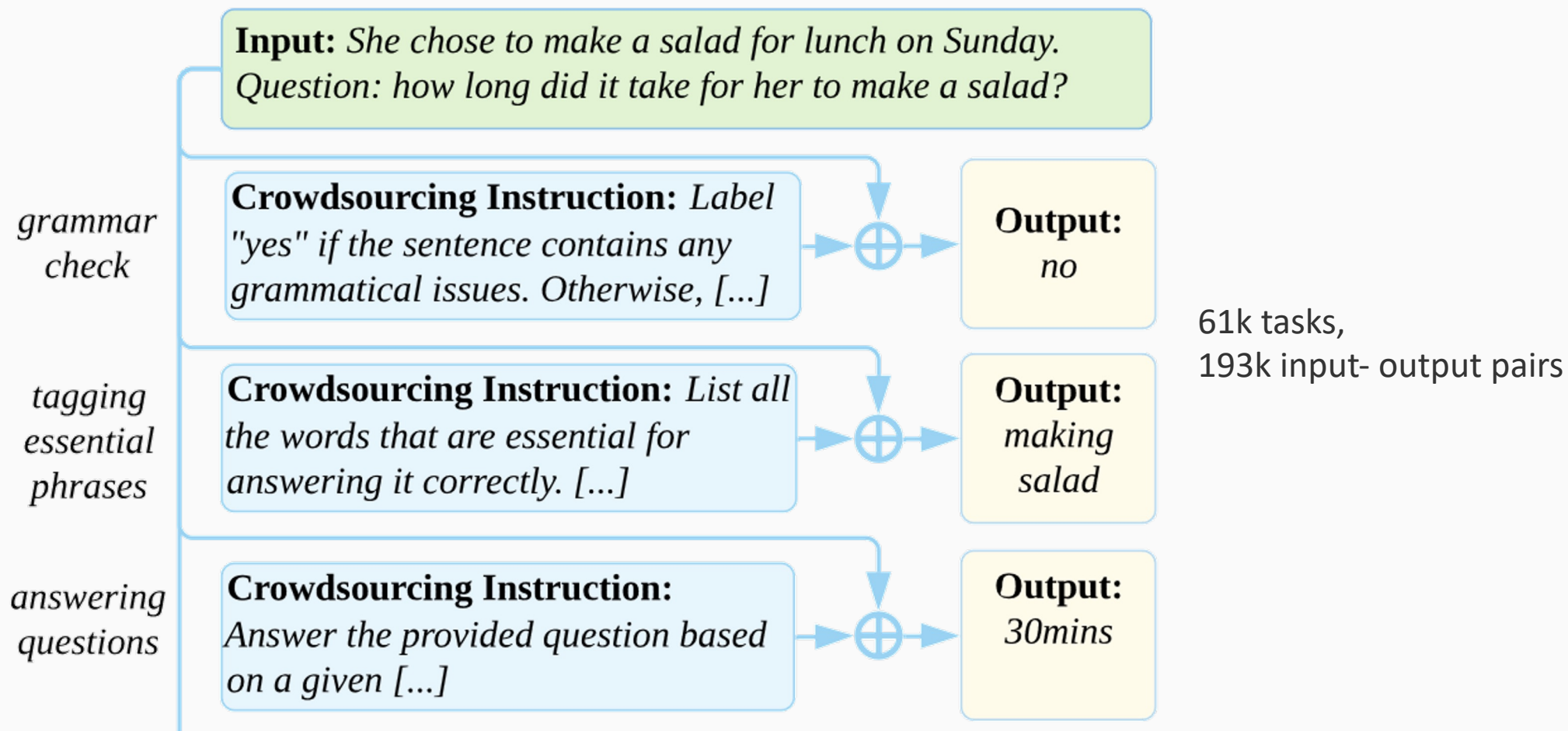
Q: Can Geoffrey Hinton have a conversation with George Washington?
Give the rationale before answering.

Geoffrey Hinton is a British-Canadian computer scientist born in 1947. George Washington died in 1799. Thus, they could not have had a conversation together. So the answer is "no".

There are many instruction-tuning datasets out there

Release	Collection	Model Details				Data Collection & Training Details				
		Model	Base	Size	Public?	Prompt Types	Tasks in Flan	# Exs	Methods	
2020 05	UnifiedQA	UnifiedQA	RoBerta	110-340M	P	ZS	46 / 46	750k		
2021 04	CrossFit	BART-CrossFit	BART	140M	NP	FS	115 / 159	71M		
2021 04	Natural Inst v1.0	Gen. BART	BART	140M	NP	ZS / FS	61 / 61	620k	+ Detailed k-shot Prompts	
2021 09	Flan 2021	Flan-LaMDA	LaMDA	137B	NP	ZS / FS	62 / 62	4.4M	+ Template Variety	
2021 10	P3	T0, T0+, T0++	T5-LM	3-11B	P	ZS	62 / 62	12M	+ Template Variety + Input Inversion	
2021 10	MetalCL	MetalCL	GPT-2	770M	P	FS	100 / 142	3.5M	+ Input Inversion + Noisy Channel Opt	
2021 11	ExMix	ExT5	T5	220M-11B	NP	ZS	72 / 107	500k	+ With Pretraining	
2022 04	Super-Natural Inst.	Tk-Instruct	T5-LM, mT5	11-13B	P	ZS / FS	1556 / 1613	5M	+ Detailed k-shot Prompts + Multilingual	
2022 10	GLM	GLM-130B	GLM	130B	P	FS	65 / 77	12M	+ With Pretraining + Bilingual (en, zh-cn)	
2022 11	xP3	BLOOMz, mT0	BLOOM, mT5	13-176B	P	ZS	53 / 71	81M	+ Massively Multilingual	
2022 12	Unnatural Inst.†	T5-LM-Unnat. Inst.	T5-LM	11B	NP	ZS	~20 / 117	64k	+ Synthetic Data	
2022 12	Self-Instruct†	GPT-3 Self Inst.	GPT-3	175B	NP	ZS	Unknown	82k	+ Synthetic Data + Knowledge Distillation	
2022 12	OPT-IML Bench†	OPT-IML	OPT	30-175B	P	ZS + FS CoT	~2067 / 2207	18M	+ Template Variety + Input Inversion + Multilingual	
2022 10	Flan 2022 (ours)	Flan-T5, Flan-PaLM	T5-LM, PaLM	10M-540B	P NP	ZS + FS CoT	1836	15M	+ Template Variety + Input Inversion + Multilingual	

Natural Instructions



Super-Natural Instructions

Super-NaturalInstructions dataset contains over 1.6K tasks, 3M+ examples

Classification, sequence tagging, rewriting, translation, QA...

Many languages: 576 non-English



PromptSource/P3

P3: Public Pool of Prompts, now 2085 prompts on 183 datasets

Dataset ?

cosmos_qa|

cord19

cornell_movie_dialog

cos_e

cosmos_qa

covid_qa_castorini

covid_qa_deepset

covid_qa_ucsd

No of prompts created for cosmos_qa : 13

Prompt name ?

description_context_question_text| ▾

context_answer_to_question

context_description_question_ans...

context_description_question_ans...

context_description_question_text

context_question_description_ans...

context_question_description_ans...

context_question_description_text

description_context_question_anc

Input template

```
Read the following context and answer the question.
Context: {{ context }}
Question: {{ question }}
Answer:
```

Target template

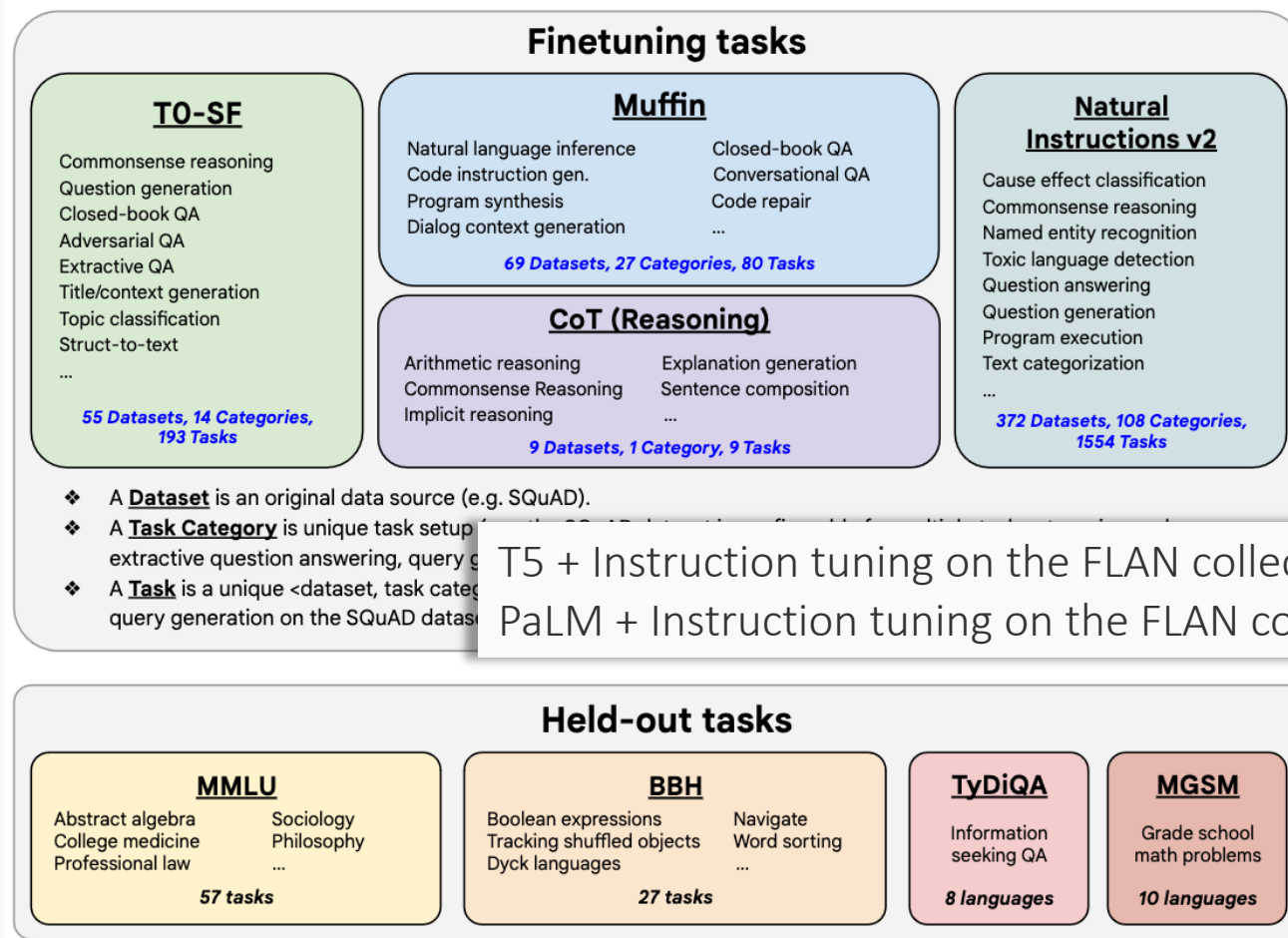
```
{{ answer_choices[label] }}
```

<https://github.com/bigscience-workshop/promptsources>

<https://huggingface.co/datasets/bigscience/P3>

The FLAN collection

73 datasets, 146 task categories, and 1,836 total tasks



Instruction-Tuning: Example

Model input (Disambiguation QA)

Q: In the following sentences, explain the antecedent of the pronoun (which thing the pronoun refers to), or state that it is ambiguous.

Sentence: The reporter and the chef will discuss their favorite dishes.

Options:

- (A) They will discuss the reporter's favorite dishes
- (B) They will discuss the chef's favorite dishes
- (C) Ambiguous

A: Let's think step by step.

Before instruction finetuning

The reporter and the chef will discuss their favorite dishes.

The reporter and the chef will discuss the reporter's favorite dishes.

The reporter and the chef will discuss the chef's favorite dishes.

The reporter and the chef will discuss the reporter's and the chef's favorite dishes.

✘ (doesn't answer question)

<https://huggingface.co/google/flan-t5-xxl>

Instruction-Tuning: Example

Model input (Disambiguation QA)

Q: In the following sentences, explain the antecedent of the pronoun (which thing the pronoun refers to), or state that it is ambiguous.

Sentence: The reporter and the chef will discuss their favorite dishes.

Options:

- (A) They will discuss the reporter's favorite dishes
- (B) They will discuss the chef's favorite dishes
- (C) Ambiguous

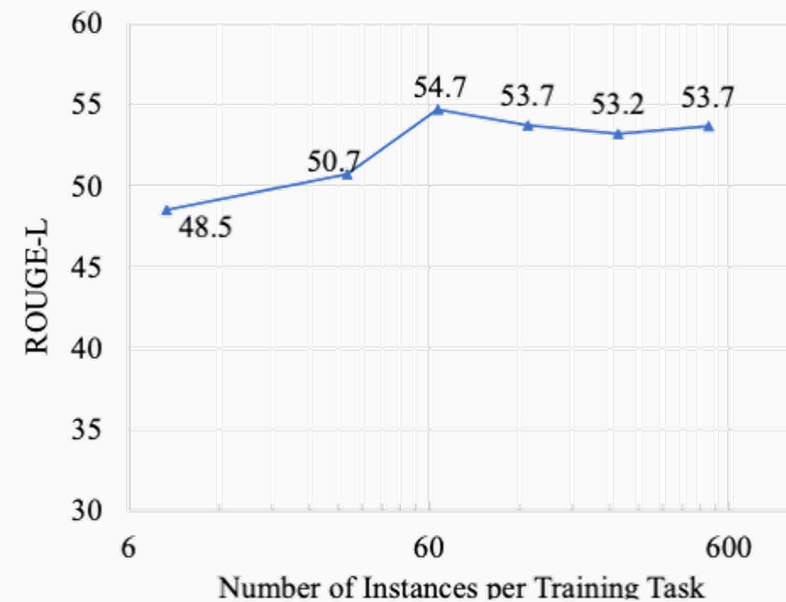
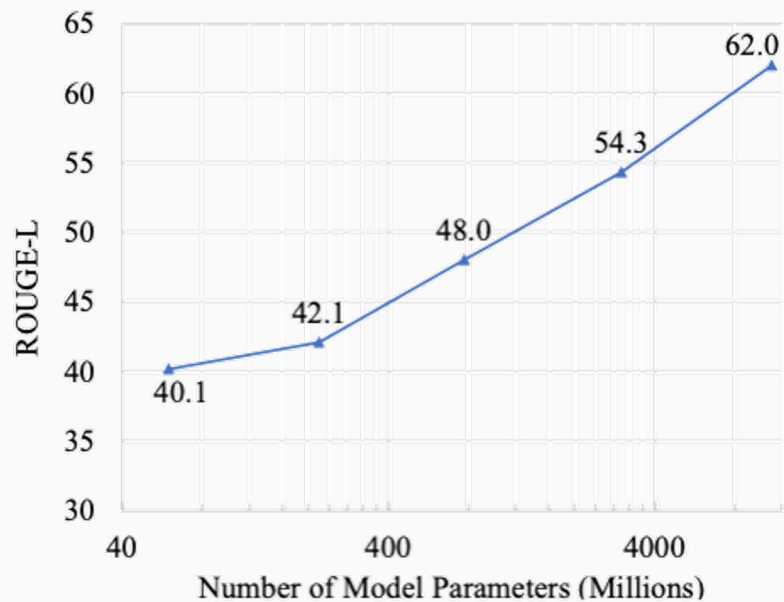
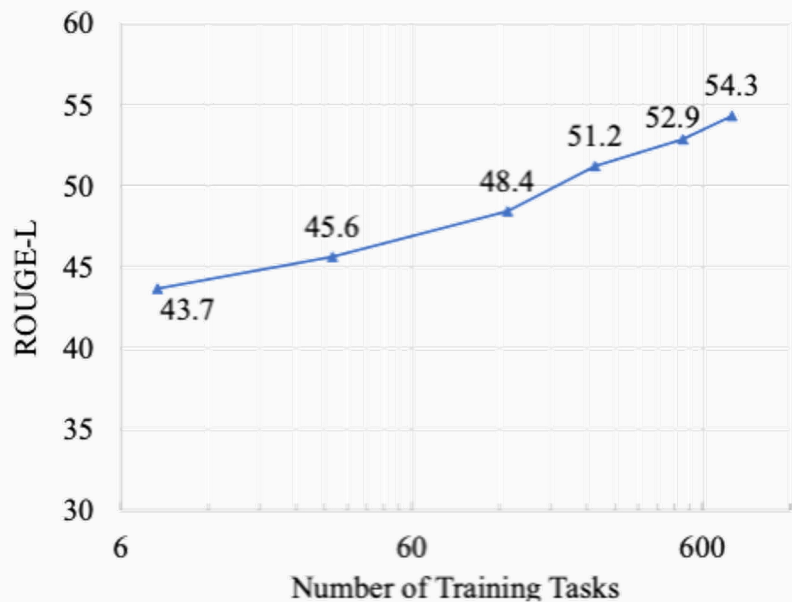
A: Let's think step by step.

After instruction finetuning

The reporter and the chef will discuss their favorite dishes does not indicate whose favorite dishes they will discuss. So, the answer is (C). ✓

<https://huggingface.co/google/flan-t5-xxl>

Scaling Instruction-Tuning



Linear growth of model performance with exponential increase in observed tasks and model size.

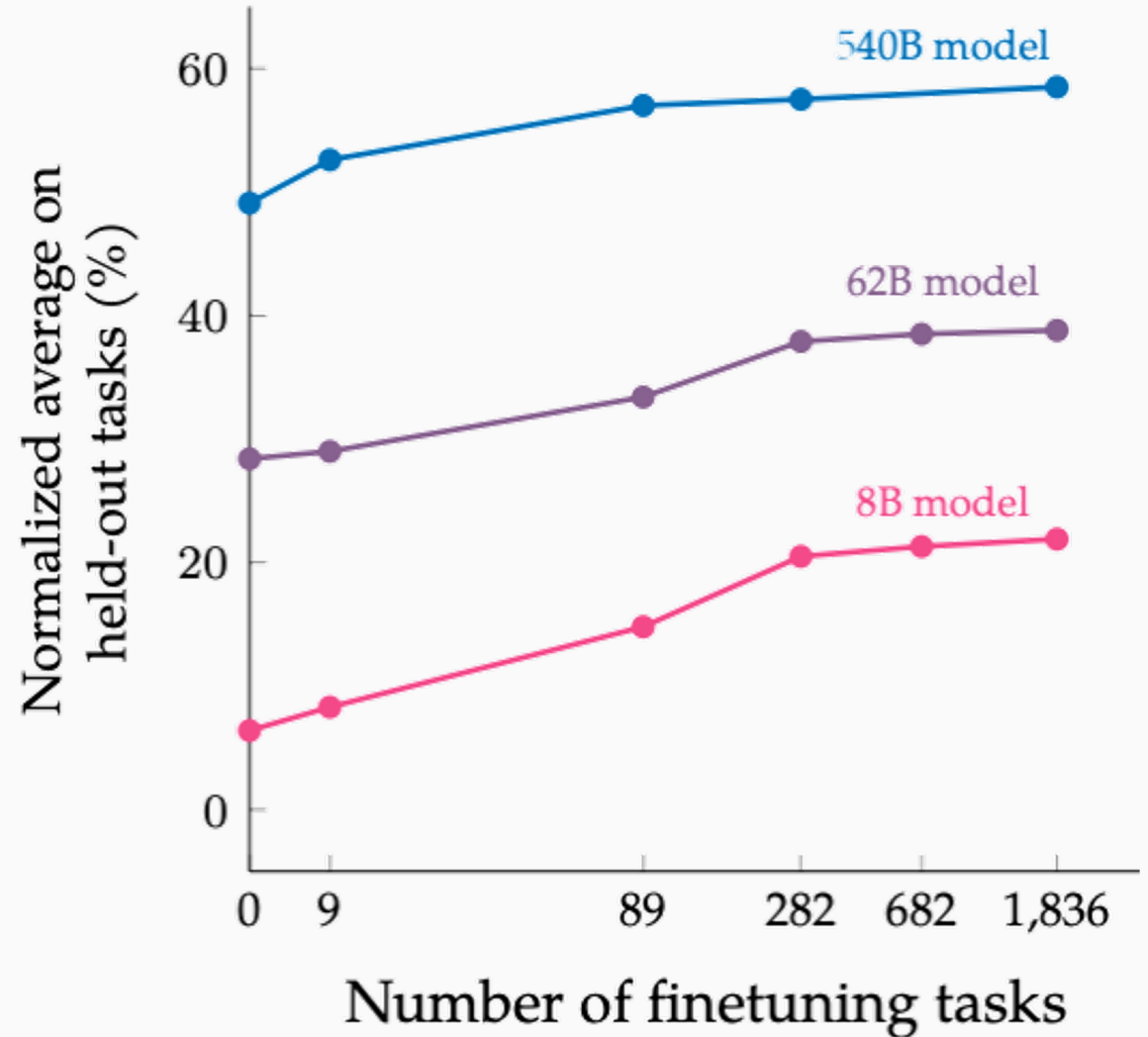
Number of examples has little effect.

Scaling Instruction-Tuning

Instruction finetuning improves performance by a large margin compared to no finetuning

Increasing the number of finetuning tasks improves performance

Increasing model scale by an order of magnitude (i.e., 8B \rightarrow 62B or 62B \rightarrow 540B) improves performance substantially for both finetuned and non-finetuned models



Summary

- Training (tuning) LMs with annotated input instructions and their output.
- Pros:
 - Simple to implement
 - Shows generalization to unseen tasks.
- Cons:
 - It's expensive to collect ground- truth data for tasks.
 - Tasks like open-ended creative generation have no right answer. For example: “Write me a story about a dog and her pet grasshopper.” Based on fine-tuning objectives, any deviations (even single-token) would incur a loss.