Introduction to Language Models

Michael Hahn

Seminar "Language Models as Cognitive Models"
October 31, 2023

Language Modeling

The task of predicting the next word

Language Modeling

The task of predicting the next word

the students attended a _____

Language Modeling

The task of predicting the next word

seminar

the students attended a _____   lecture

concert

…

Language Modeling

The task of predicting the next word probabilistically

seminar

the students attended a _____   lecture

concert

toothpaste

the

…

$$P(\boldsymbol{x}^{(t+1)} \mid \boldsymbol{x}^{(t)}, \ldots, \boldsymbol{x}^{(1)})$$
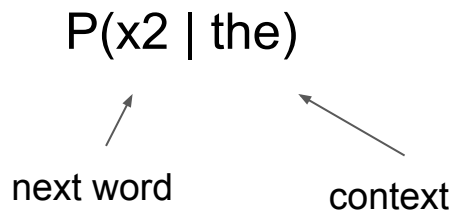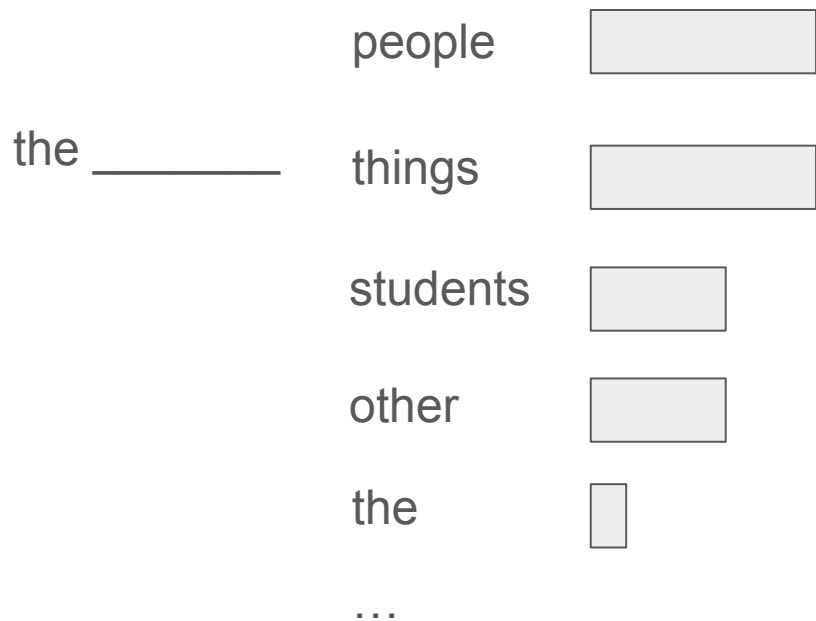
next word   context

Generating a Text

Generating a text

 the _____

Generating a Text

Generating a text

people

the _____ things

students

other

the

…

P(x2 | the)

next word   context
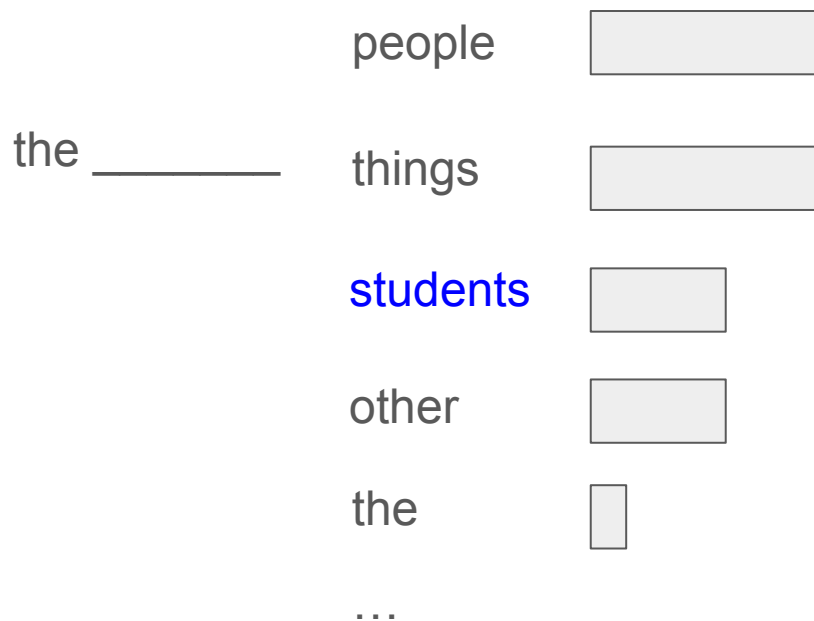
Generating a Text

Generating a text

the _____

people

things

students

other

the

…

P(x2 | the)

next word   context

Generating a Text

Generating a text

the students _____

are

were

went

attended

the

…

P(x3 | the students)

next word

context

Generating a Text

Generating a text

the students _____

are

were

went

attended

the

…

P(x3 | the students)

next word

context

Generating a Text

Generating a text

the students attended_____

the	
a	
another	
seminars	
some	
…	

P(x4 | the students attended

next word   context

Generating a Text

Generating a text

the students attended_____

the	
a	
another	
seminars	
some	
…	

P(x4 | the students attended

next word   context

Generating a Text

Generating a text

the students attended a _____

seminar

lecture

concert

toothpaste

the

…

$P(x5 \mid \text{the students attended a})$
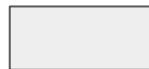
next word

context

Generating a Text

Generating a text

the students attended a _____

seminar

lecture

concert

toothpaste

the

…
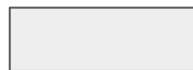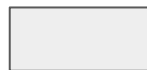
P(x5 | the students attended a)

next word

context

Language Modeling

P(the students attended a seminar)

=

P(the)

× P(students | the)

× P(attended | the students)

× P(a | the students attended)

× P(seminar | the students attended a)

Language Modeling

General Formula:

$$P(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(T)}) = P(\boldsymbol{x}^{(1)}) \times P(\boldsymbol{x}^{(2)} \mid \boldsymbol{x}^{(1)}) \times \cdots \times P(\boldsymbol{x}^{(T)} \mid \boldsymbol{x}^{(T-1)}, \ldots, \boldsymbol{x}^{(1)})$$

$$= \prod_{t=1}^{T} P(\boldsymbol{x}^{(t)} \mid \boldsymbol{x}^{(t-1)}, \ldots, \boldsymbol{x}^{(1)})$$

Google

what is the | ✕   📷

🔍 what is the **tower of london**

🔍 what is the **richest country in the world**

🔍 what is the **one piece**

🔍 what is the **name of the huge area in the centre of australia**

🔍 what is the **capital of australia**

🔍 what is the **weather today**

🔍 what is the **right course of action**

🔍 what is the **london eye**

🔍 what is the **meaning of life**

How to calculate P(next word | context)?

P(seminar | the students attended a)

= #(the students attended a seminar) / #(the students attended a)

how often does each
sequence occur in a
huge corpus?

Problem:
Longer sequences often never observed in corpus.
Need to back-off to shorter contexts.
Cannot capture long-distance relations.

How to calculate P(next word | context)?

Modern Approach
(roughly since 2015)

P(… | the students attended a)

```
Neural Network
```
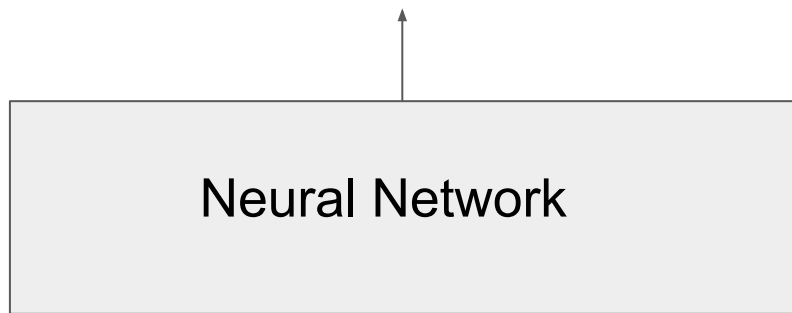
trained on large
corpus

the students attended a

Input Representation

Word Embeddings

Input Representation

Problem: what to do with unknown/rare words?

Input Representation

Problem: what to do with unknown/rare words?

The road running through UdS campus is called **Stuhlsatzenhausweg**

may not have appeared in the training set!

Input Representation

Problem: what to do with unknown/rare words?

The road running through UdS campus is called Stuhlsatzenhausweg

Traditional Approach: Replace by special token

The road running through UdS campus is called <UNKNOWN>

Not very satisfying!

Input Representation

Problem: what to do with unknown/rare words?

The road running through UdS campus is called Stuhlsatzenhausweg

Modern approach: subword tokenization

*The_ road_ run ning_ through_ U d S_ campus_ is called_ Stuhl satz en haus weg_*

Can represent any input.
In the worst case, just need to back off to the individual letters.

Neural Network: Transformer

v(the)	v(students)	v(attended)	v(a)

Neural Network: Transformer

v(the)	v(students)	v(attended)	v(a)

+   +   +   +

p(1)	p(2)	p(3)	p(4)
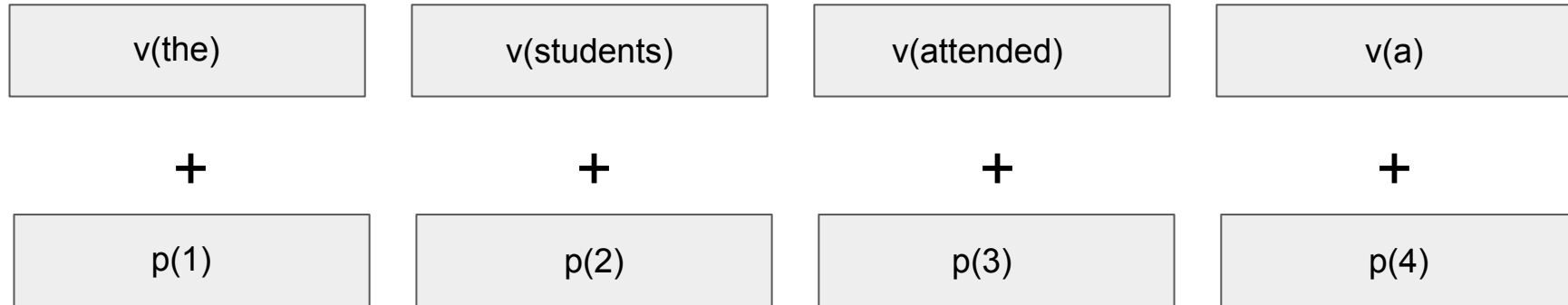
positional embeddings indicating
position in the sequence

Neural Network: Transformer

compute all dot products between keys and queries

key	query	value		key	query	value		key	query	value		key	query	value

v(the)		v(students)		v(attended)		v(a)

+   +   +   +

p(1)		p(2)		p(3)		p(4)
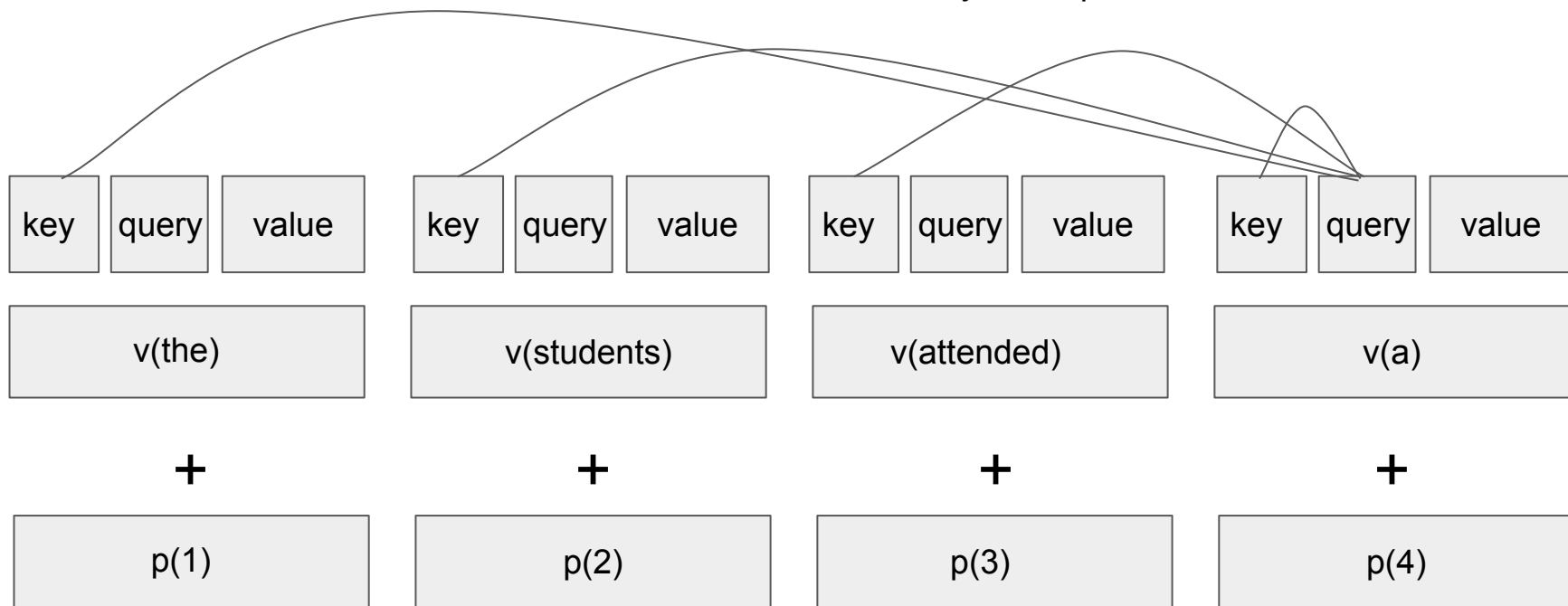
Neural Network: Transformer

attention weights are normalized to sum up to 1

key	query	value		key	query	value		key	query	value		key	query	value

v(the)	v(students)	v(attended)	v(a)

+	+	+	+

p(1)	p(2)	p(3)	p(4)
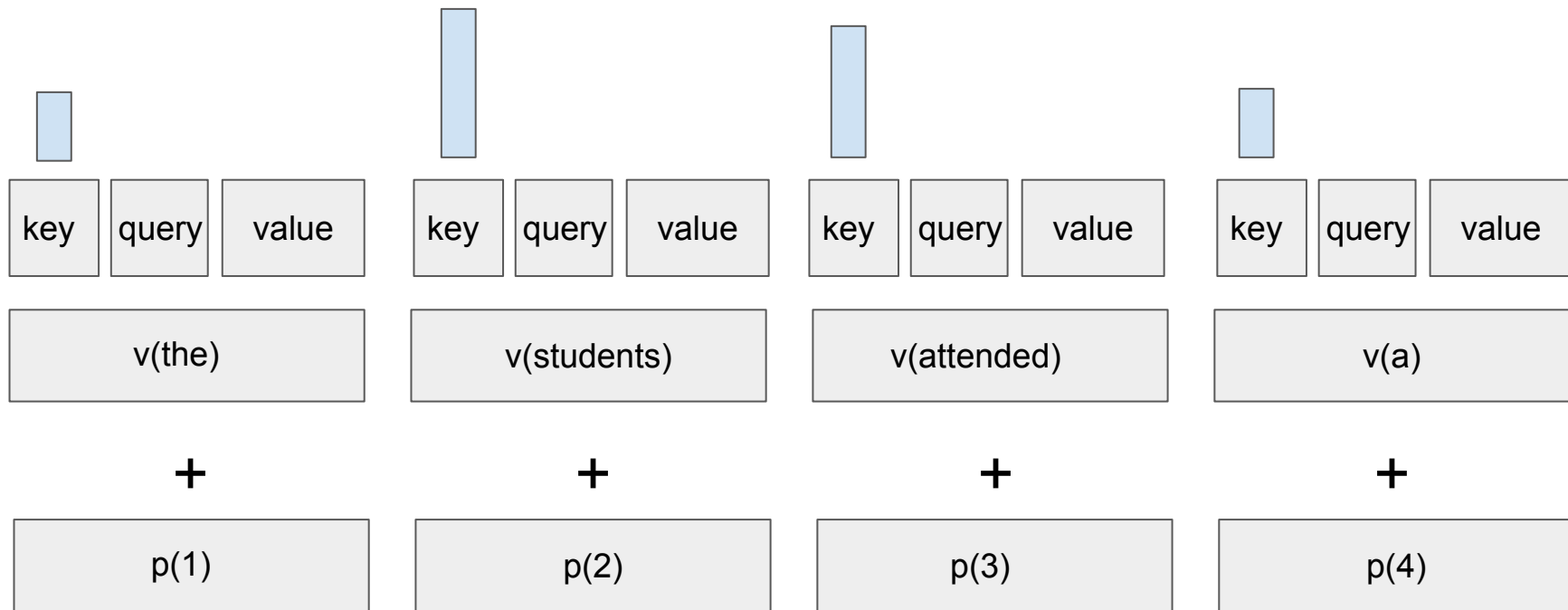
Neural Network: Transformer

value vectors are summed up, weighted by attention weights

attention

key	query	value		key	query	value		key	query	value		key	query	value

v(the)	v(students)	v(attended)	v(a)

+ + + +
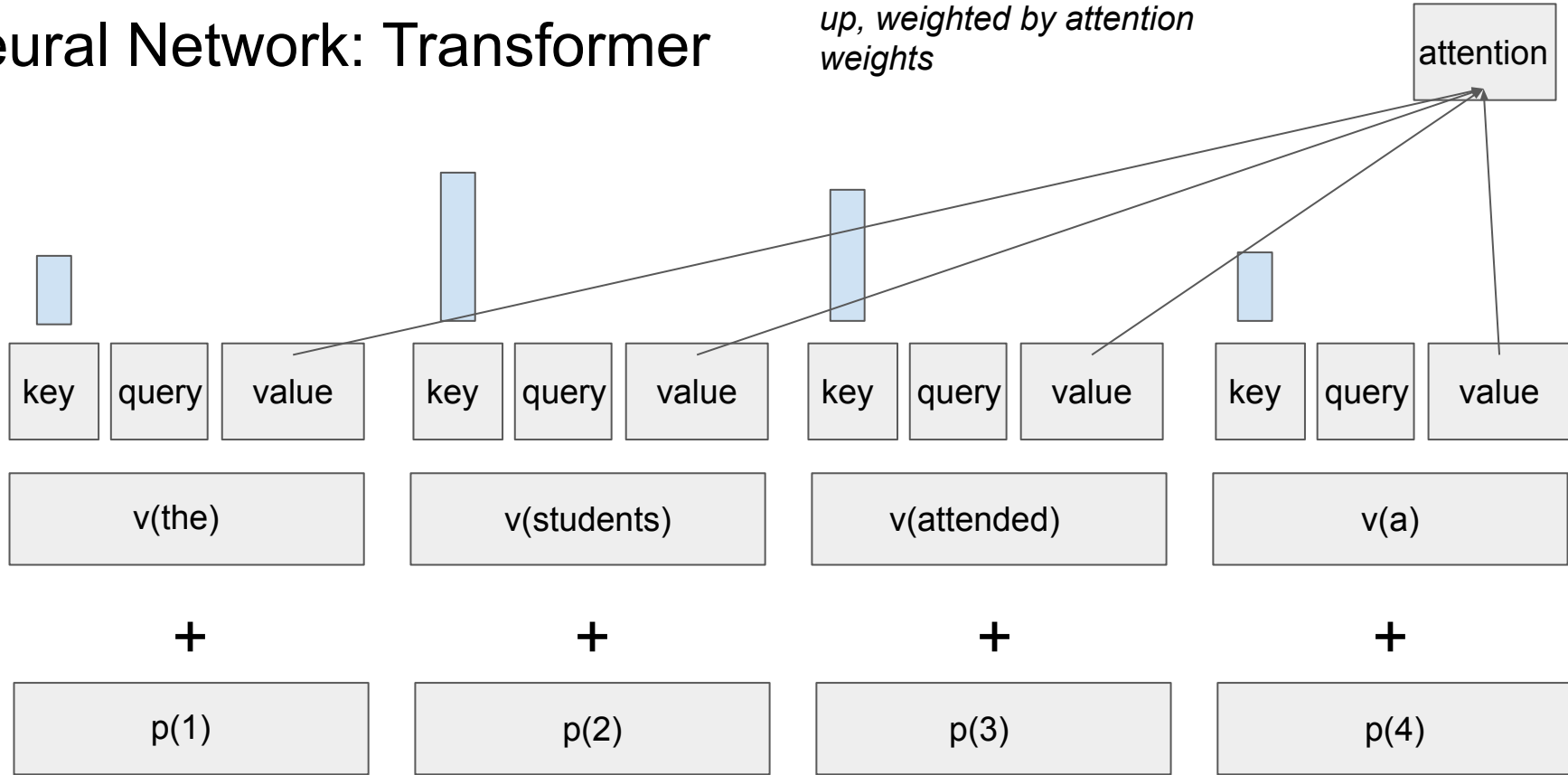
p(1)	p(2)	p(3)	p(4)

Neural Network: Transformer

in fact, we may use multiple such attention "heads" in parallel

attention

key	query	value		key	query	value		key	query	value		key	query	value

v(the)	v(students)	v(attended)	v(a)

+	+	+	+
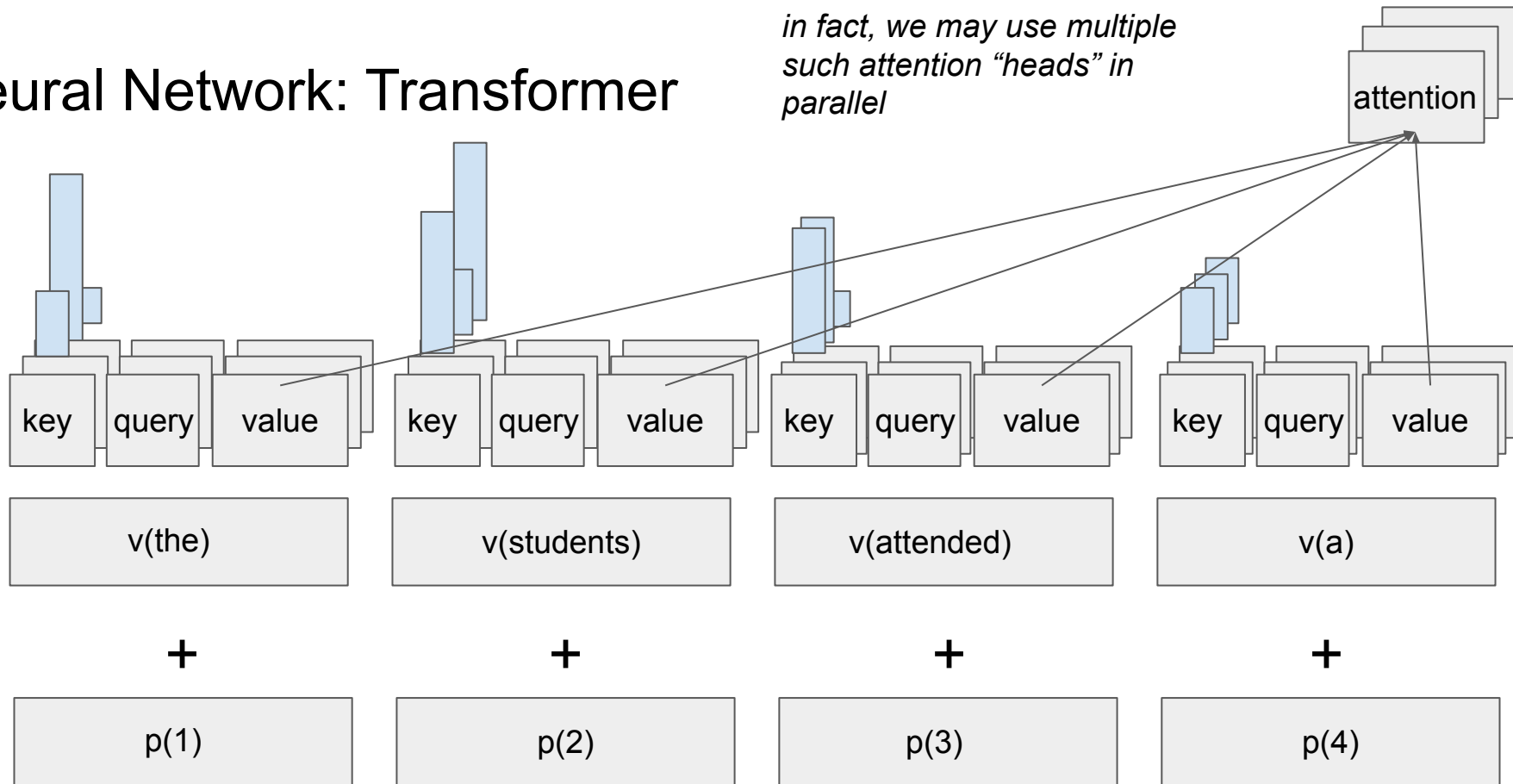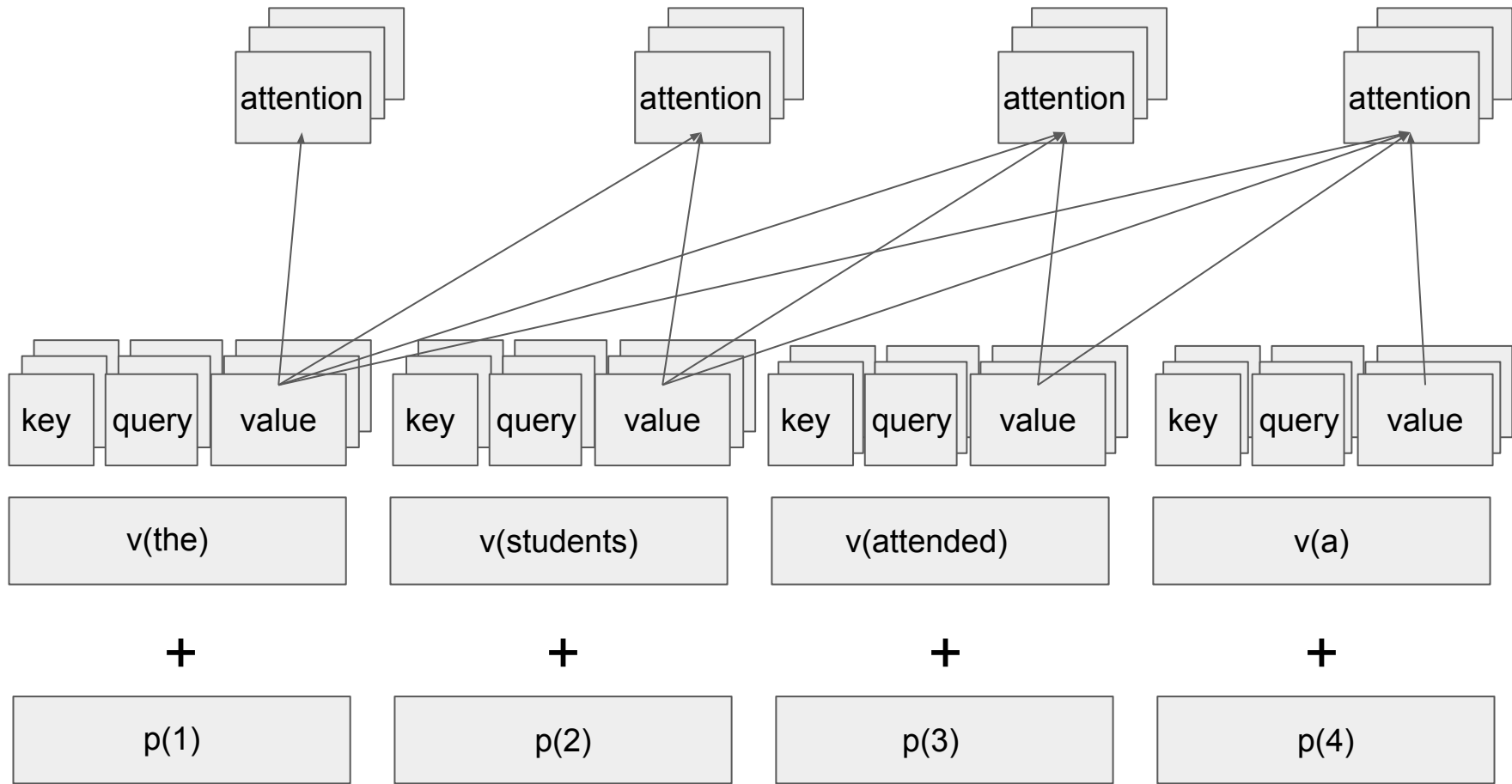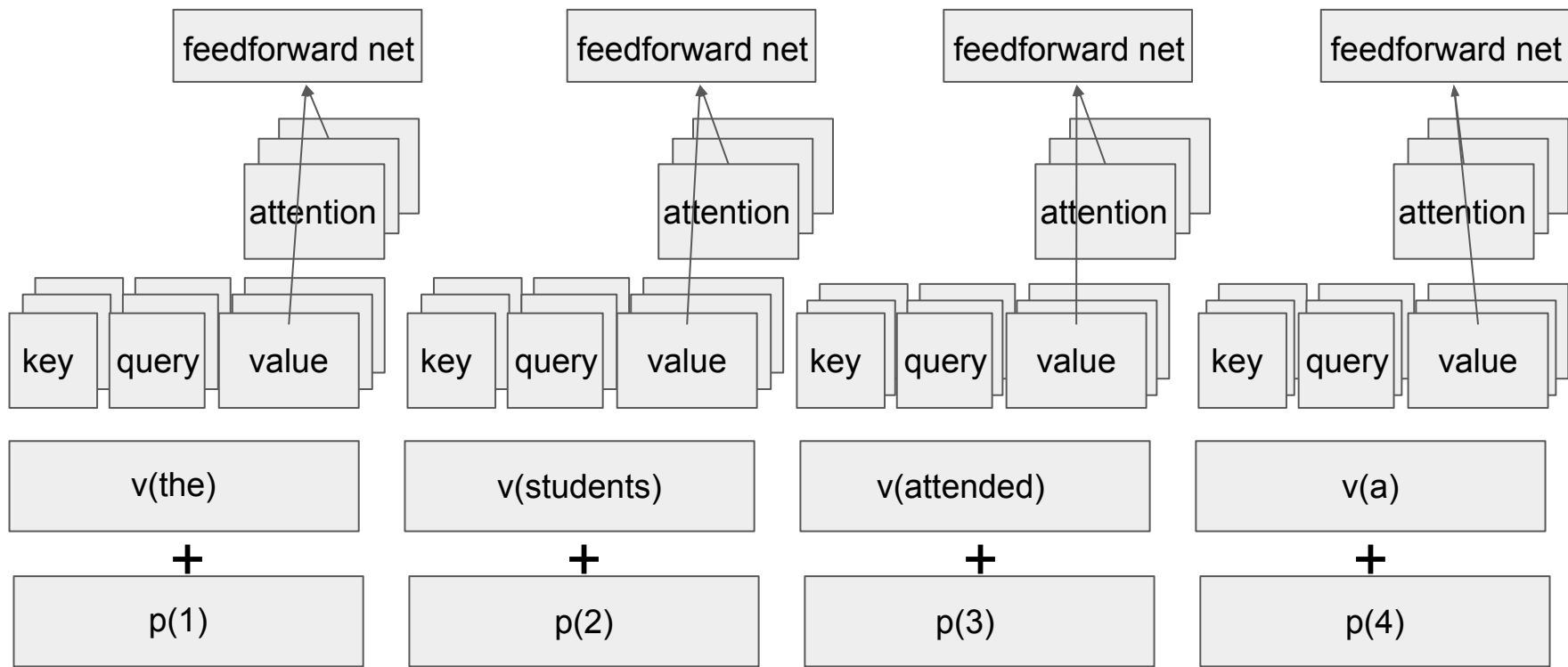
p(1)	p(2)	p(3)	p(4)

this is one transformer layer

feedforward net	feedforward net	feedforward net	feedforward net
attention	attention	attention	attention
key	query	value	key	query	value	key	query	value	key	query	value

v(the)	v(students)	v(attended)	v(a)
+	+	+	+
p(1)	p(2)	p(3)	p(4)
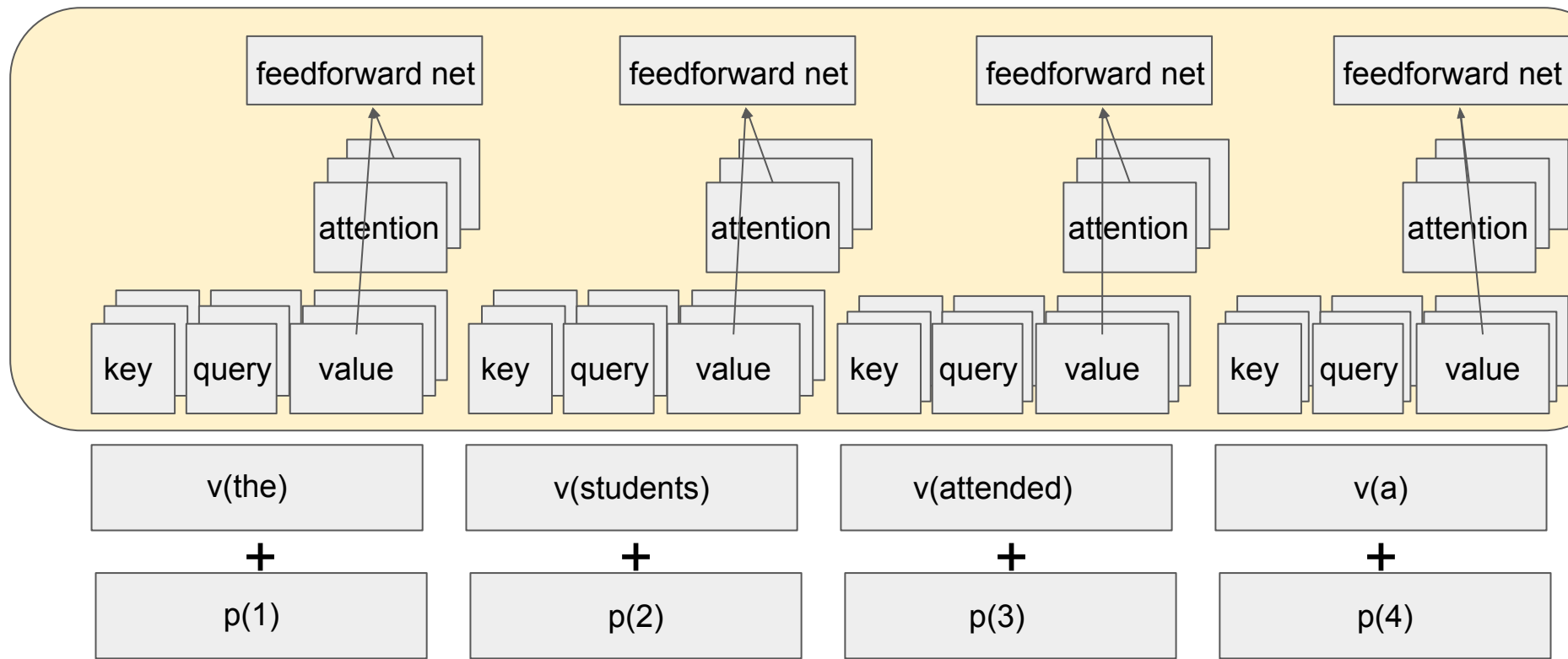
Layer N

. . .

Layer 3

Layer 2

Layer 1

v(the)		v(students)		v(attended)		v(a)
+		+		+		+
p(1)		p(2)		p(3)		p(4)

| a | aardvark | … | seminar | … | the | … | | | a | aardvark | … | seminar | … | the | … | | | a | aardvark | … | seminar | … | the | … | | | a | aardvark | … | seminar | … | the | … |
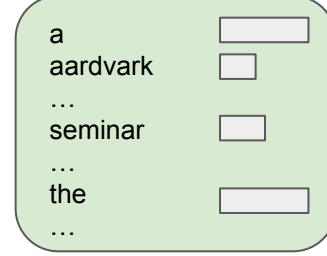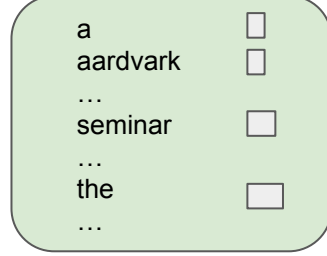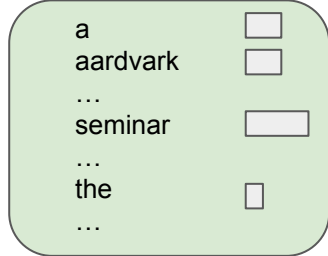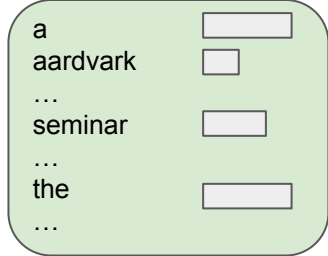
feedforward net   feedforward net   feedforward net   feedforward net

Layer N

. . .

Layer 1

v(the)	v(students)	v(attended)	v(a)
+	+	+	+
p(1)	p(2)	p(3)	p(4)

GPT-3: Prompting & In-Context Learning

Circulation revenue has increased by 5% in Finland. // Positive

Panostaja did not disclose the purchase price. // Neutral

Paying off the national debt will be extremely painful. // Negative

The company anticipated its operating profit to improve. // _____

LM ↓

Positive

GPT-3: Prompting & In-Context Learning

Circulation revenue has increased by 5% in Finland. // Positive

Panostaja did not disclose the purchase price. // Neutral

Paying off the national debt will be extremely painful. // Negative

The company anticipated its operating profit to improve. // _____

LM ↓

Positive

Circulation revenue has increased by 5% in Finland. // Finance

They defeated ... in the NFC Championship Game. // Sports

Apple ... development of in-house chips. // Tech

The company anticipated its operating profit to improve. // _____

LM ↓

Finance

http://ai.stanford.edu/blog/understanding-incontext/

GPT-3: Prompting & In-Context Learning


```
1    gaot => goat
2    sakne => snake
3    brid => bird
4    fsih => fish
5    dcuk => duck
6    cmihp => chimp
```

In-context learning

```
1    thanks => merci
2    hello => bonjour
3    mint => menthe
4    wall => mur
5    otter => loutre
6    bread => pain
```

In-context learning
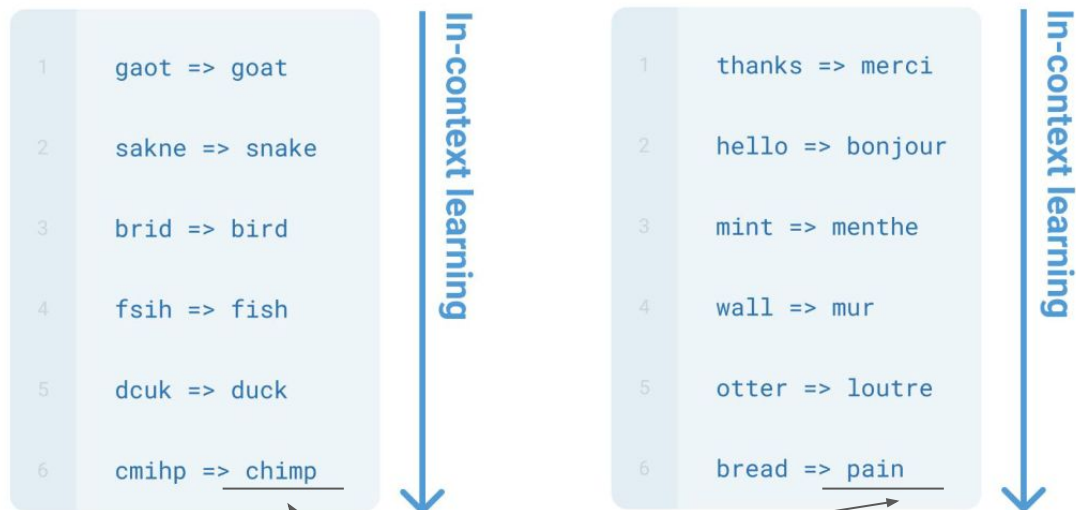
LM correctly completes the prompt!

(Brown et al., 2020)

Chain-of-thought prompting

Standard Prompting

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?

Nye et al. (2021), Wei et al. (2022

Chain-of-thought prompting

Standard Prompting

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. ✖

Nye et al. (2021), Wei et al. (2022)

Chain-of-thought prompting

Standard Prompting

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

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?

Nye et al. (2021), Wei et al. (2022)

Chain-of-thought prompting

Standard Prompting

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

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. ✔️

Nye et al. (2021), Wei et al. (2022

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.**

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.** *There are 16 balls in total. Half of the balls are golf balls. That means there are 8 golf balls. Half of the golf balls are blue. That means there are 4 blue golf balls.* ✔️

(Kojima et al., 2022)

No.	Category	Zero-shot CoT Trigger Prompt	Accuracy
1	LM-Designed	Let's work this out in a step by step way to be sure we have the right answer.	**82.0**
2	Human-Designed	Let's think step by step. (*1)	78.7
3		First, (*2)	77.3
4		Let's think about this logically.	74.5
5		Let's solve this problem by splitting it into steps. (*3)	72.2
6		Let's be realistic and think step by step.	70.8
7		Let's think like a detective step by step.	70.3
8		Let's think	57.5
9		Before we dive into the answer,	55.7
10		The answer is after the proof.	45.7
-		(Zero-shot)	17.7

However…

Explain the moon landing to a 6 year old in a few sentences.

GPT-3

```
Explain the theory of gravity to a 6 year old.

Explain the theory of relativity to a 6 year old in a few sentences.

Explain the big bang theory to a 6 year old.

Explain evolution to a 6 year old.
```

https://web.stanford.edu/class/cs224n/slides/cs224n-2023-lecture11-prompting-rlhf.pdf   (Ouyang et al., 2022)

Instruction Finetuning

PROMPT   *Explain the moon landing to a 6 year old in a few sentences.*
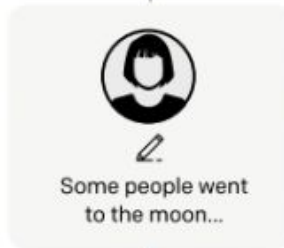
COMPLETION

Human
A giant rocket ship blasted off from Earth carrying astronauts to the moon. The astronauts landed their spaceship on the moon and walked around exploring the lunar surface. Then they returned safely back to Earth, bringing home moon rocks to show everyone.

Collect demonstration data, and train a supervised policy.

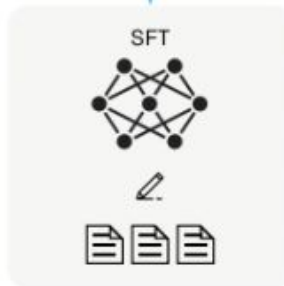A prompt is sampled from our prompt dataset.

Explain the moon landing to a 6 year old

A labeler demonstrates the desired output behavior.

Some people went to the moon...

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

SFT

(Ouyang et al., 2022)

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)

(Chung et al., 2022, Figure 10)

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.

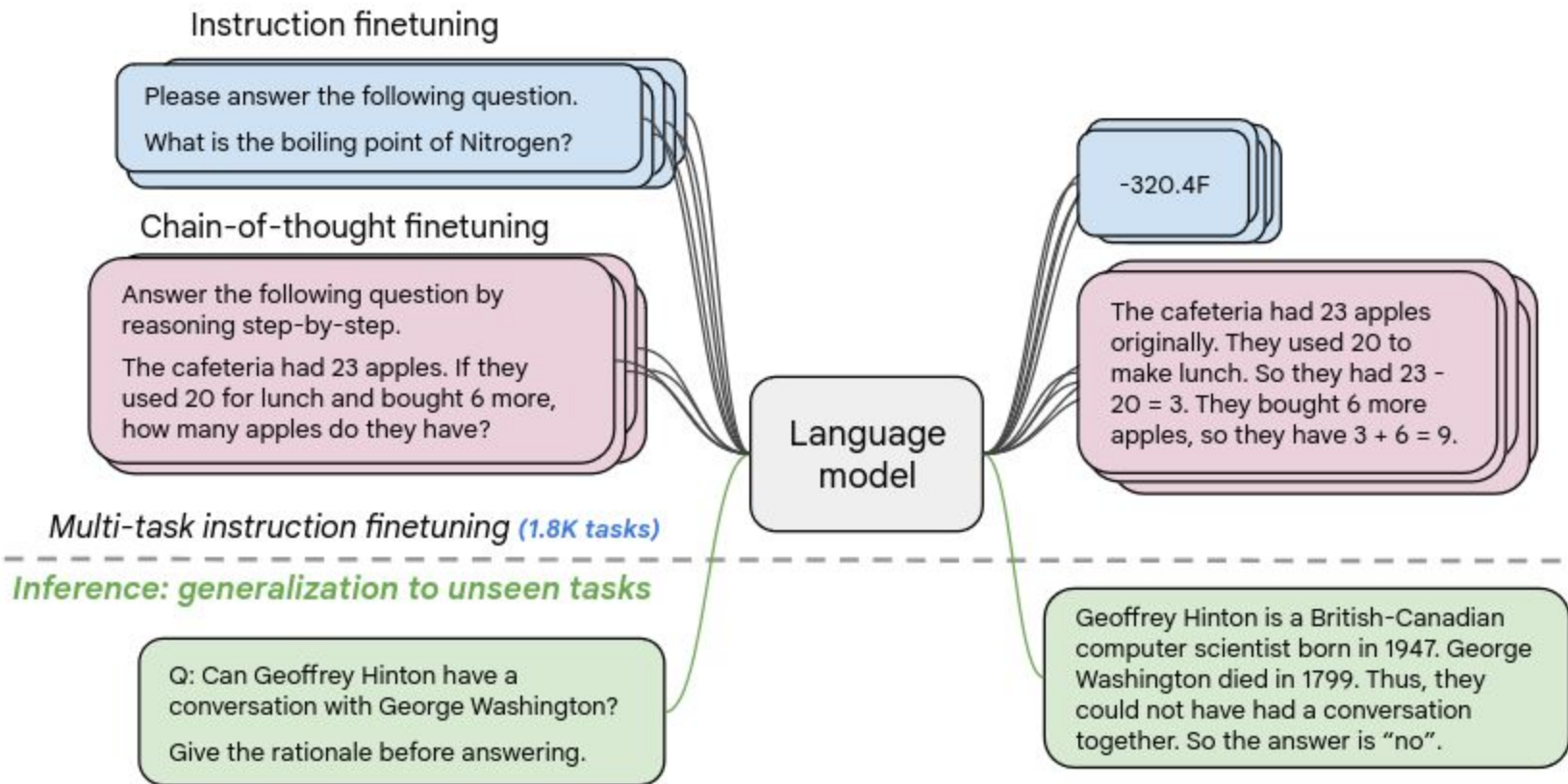Sentence: The reporter and the chef will discuss their favorite dishes.

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). ✅
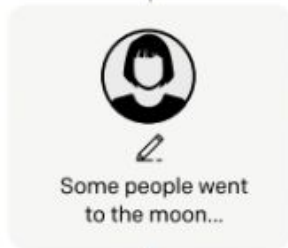
(Chung et al., 2022, Figure 10)

Instruction finetuning

Please answer the following question.

What is the boiling point of Nitrogen?

Chain-of-thought finetuning

Answer the following question by reasoning step-by-step.

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

Multi-task instruction finetuning (1.8K tasks)

Inference: generalization to unseen tasks

Q: Can Geoffrey Hinton have a conversation with George Washington?

Give the rationale before answering.

Language model

-320.4F

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.

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".

FLAN-T5; Chung et al., 2022

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.

Explain the moon landing to a 6 year old

A labeler demonstrates the desired output behavior.

Some people went to the moon...

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

SFT

Limitations:

- often there is no single correct way of completing a task
- getting some tokens wrong is much worse than other tokens

(Ouyang et al., 2022)

Human Feedback

1. given a prompt, sample outputs from the LM

Human Feedback

1. given a prompt, sample outputs from the LM

2. have humans rate them

```
Summarize the following:
Saarland University (German:
Universität des Saarlandes,
pronounced) is a public research
university located in Saarbrücken,
the capital of the German state of
Saarland. It was founded in 1948
in Homburg in co-operation with
France and is organized in six
faculties that cover all major
fields of…
```

Human Feedback

1. given a prompt, sample outputs from the LM

2. have humans rate them

```
Summarize the following:
Saarland University (German:
Universität des Saarlandes,
pronounced) is a public research
university located in Saarbrücken,
the capital of the German state of
Saarland. It was founded in 1948
in Homburg in co-operation with
France and is organized in six
faculties that cover all major
fields of…
```

Summary 1:
Saarland University is a research
university in Saarbrücken,
Germany.

Summary 3:
An university in Saarland was
founded in 1948.

Summary 2:
The university senate has nine
professors.

Human Feedback

1. given a prompt, sample outputs from the LM

2. have humans rate them

```
Summarize the following:
Saarland University (German:
Universität des Saarlandes,
pronounced) is a public research
university located in Saarbrücken,
the capital of the German state of
Saarland. It was founded in 1948
in Homburg in co-operation with
France and is organized in six
faculties that cover all major
fields of…
```

```
Summary 1:
Saarland University is a research
university in Saarbrücken,
Germany.
```
good

```
Summary 3:
An university in Saarland was
founded in 1948.
```
fair

```
Summary 2:
The university senate has nine
professors.
```
poor

Human Feedback

1. given a prompt, sample outputs from the LM

2. have humans rate them

3. tune LM to give responses given higher human ratings

Human Feedback

1. given a prompt, sample outputs from the LM

2. have humans rate them

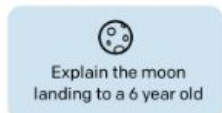3. tune LM to give responses given higher human ratings

challenge: human feedback hard to scale

solution: train proxy reward model on human data & finetune LM for that proxy model
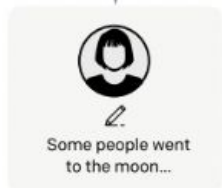
Step 1
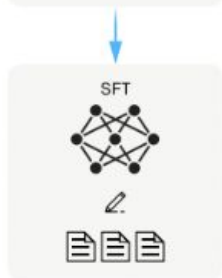### Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.

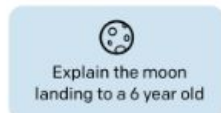A labeler demonstrates the desired output behavior.

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

Explain the moon landing to a 6 year old

Some people went to the moon...

SFT

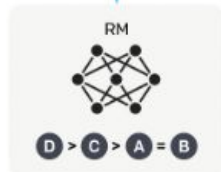## Step 2
Collect comparison data, and train a reward model.

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.

Explain the moon landing to a 6 year old

A Explain gravity...
B Explain war...
C Moon is natural satellite of...
D People went to the moon...

D > C > A = B

RM

D > C > A = B

Step 3
Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.
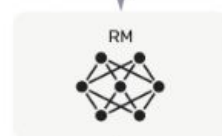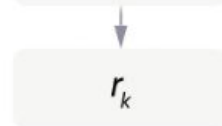
The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

Write a story about frogs

PPO

Once upon a time...

RM

r_k

Ouyang (2022)

PROMPT *Explain the moon landing to a 6 year old in a few sentences.*

COMPLETION GPT-3

Explain the theory of gravity to a 6 year old.

Explain the theory of relativity to a 6 year old in a few sentences.

Explain the big bang theory to a 6 year old.

Explain evolution to a 6 year old.

InstructGPT

People went to the moon, and they took pictures of what they saw, and sent them back to the earth so we could all see them.

https://web.stanford.edu/class/cs224n/slides/cs224n-2023-lecture11-prompting-rlhf.pdf

PROMPT   *Write a short poem about a wise frog.*

COMPLETION   GPT-3

Write a short story in which a character has two different names.

Write a short story in which you try to get something back that you have lost.

Write a short story in which a character has a bad dream.

InstructGPT

The frog is a symbol of wisdom
He knows all the secrets of the world
He is a master of disguise
And a great teacher of life
He is a symbol of transformation
And the bringer of change
He is the frog who has seen it all
And knows the meaning of it all

Conclusion

Modern Large Language Models combine…

- next-word prediction

- finetuning for instructions following

- reinforcement learning for following human preferences

Thank you!