

CS11-711 Advanced NLP

Complex Reasoning

Graham Neubig



Carnegie Mellon University

Language Technologies Institute

Site

<https://phontron.com/class/anlp2024/>

What is Reasoning?

- Using **evidence** and **logic** to arrive at conclusions and make judgments
- Can be
 - **Formal:** focusing on strict truth values
 - **Informal:** based on experience and intuition

Types of Reasoning

(examples: Huang and Chang 2023)

Deductive: Use logic to go from premise to firm conclusion.

Premise: All mammals have kidneys.

Premise: All whales are mammals.

Conclusion: All whales have kidneys.

Inductive: From observation, predict a likely conclusion.

Observation: When we see a creature with wings, it is usually a bird.

Observation: We see a creature with wings.

Conclusion: The creature is likely to be a bird.

Abductive: From observation, predict the most likely explanation.

Observation: The car cannot start and there is a puddle of liquid under the engine.

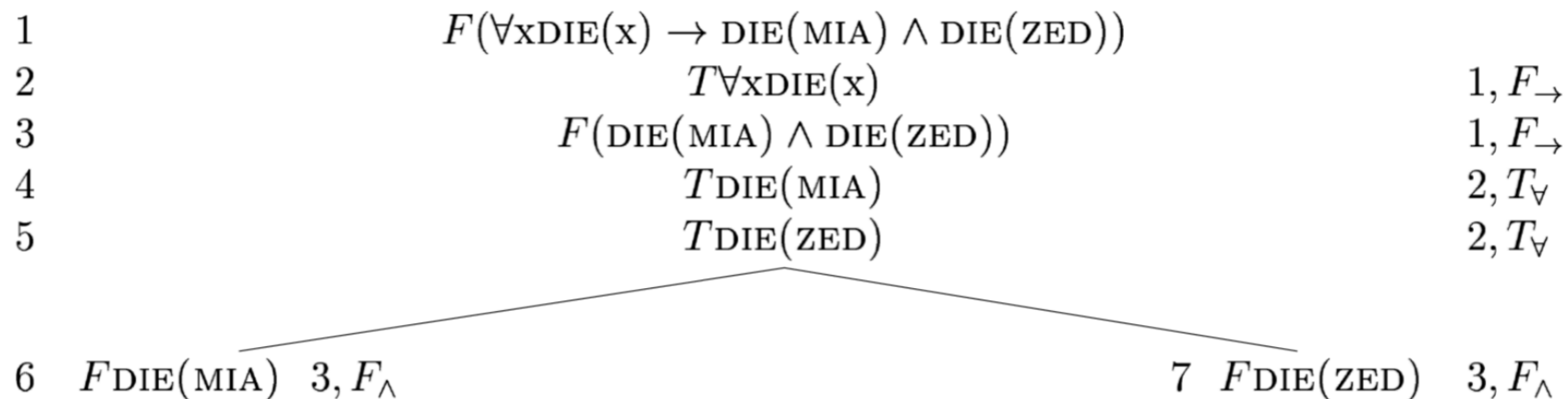
Likely Explanation: The car has a leak in the radiator

Pre-LLM Reasoning Methods

Computational Semantics

- Reasoning is something that traditional semantic representations are really good at! (e.g. Prolog)

$$\forall x \text{DIE}(x) \rightarrow \text{DIE}(\text{MIA}) \wedge \text{DIE}(\text{ZED})$$

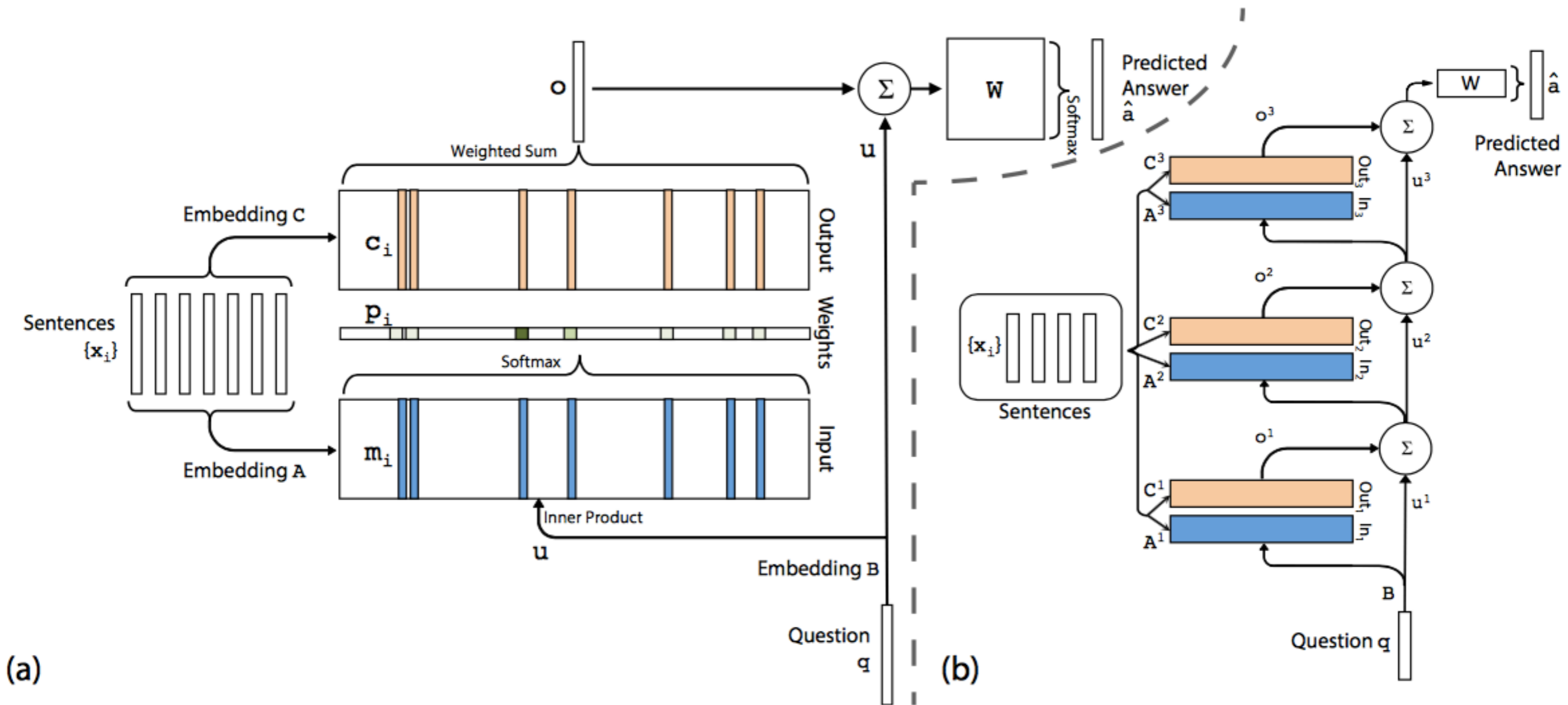


- See "Representation and Inference for Natural Language" (Blackburn & Bos 1999)
- Most neural networks are just a very rough approximation...

Memory Networks

(Sukhbaatar et al. 2015)

- Use standard softmax attention, and multiple layers



(a)

(b)

Solving Word Problems w/ Symbolic Reasoning

- Idea: combine semantic parsing (with explicit functions) and machine reading
- e.g. Gupta et al. (2020)

Who kicked the longest field goal in the second quarter?

Question Parser

```
relocate(find-max-num(filter(find()))
```

Program Executor

find	filter	find-max-num	relocate
field goal	in the second quarter		Who kicked

Answer: Connor Barth

In the first quarter, Buffalo trailed as Chiefs QB Tyler Thigpen completed a 36-yard TD pass to RB Jamaal Charles. The Bills responded with RB Marshawn Lynch getting a 1-yard touchdown run. In the second quarter, Buffalo took the lead as kicker Rian Lindell made a 21-yard and a 40-yard field goal. Kansas City answered with Thigpen completing a 2-yard TD pass. Buffalo regained the lead as Lindell got a 39-yard field goal. The Chiefs struck with kicker Connor Barth getting a 45-yard field goal, yet the Bills continued their offensive explosion as Lindell got a 34-yard field goal, along with QB Edwards getting a 15-yard TD run. In the third quarter, Buffalo continued its poundings with Edwards getting a 5-yard TD run, while Lindell got himself a 48-yard field goal. Kansas City tried to rally as Thigpen completed a 45-yard TD pass to WR Mark Bradley, yet the Bills replied with Edwards completing an 8-yard TD pass to WR Josh Reed. In the fourth quarter, Edwards completed a 17-yard TD pass to TE Derek Schouman.

Chain of Thought and Variants

Review: Chain-of-thought Prompting (Wei et al. 2022)

- Get the model to explain its reasoning

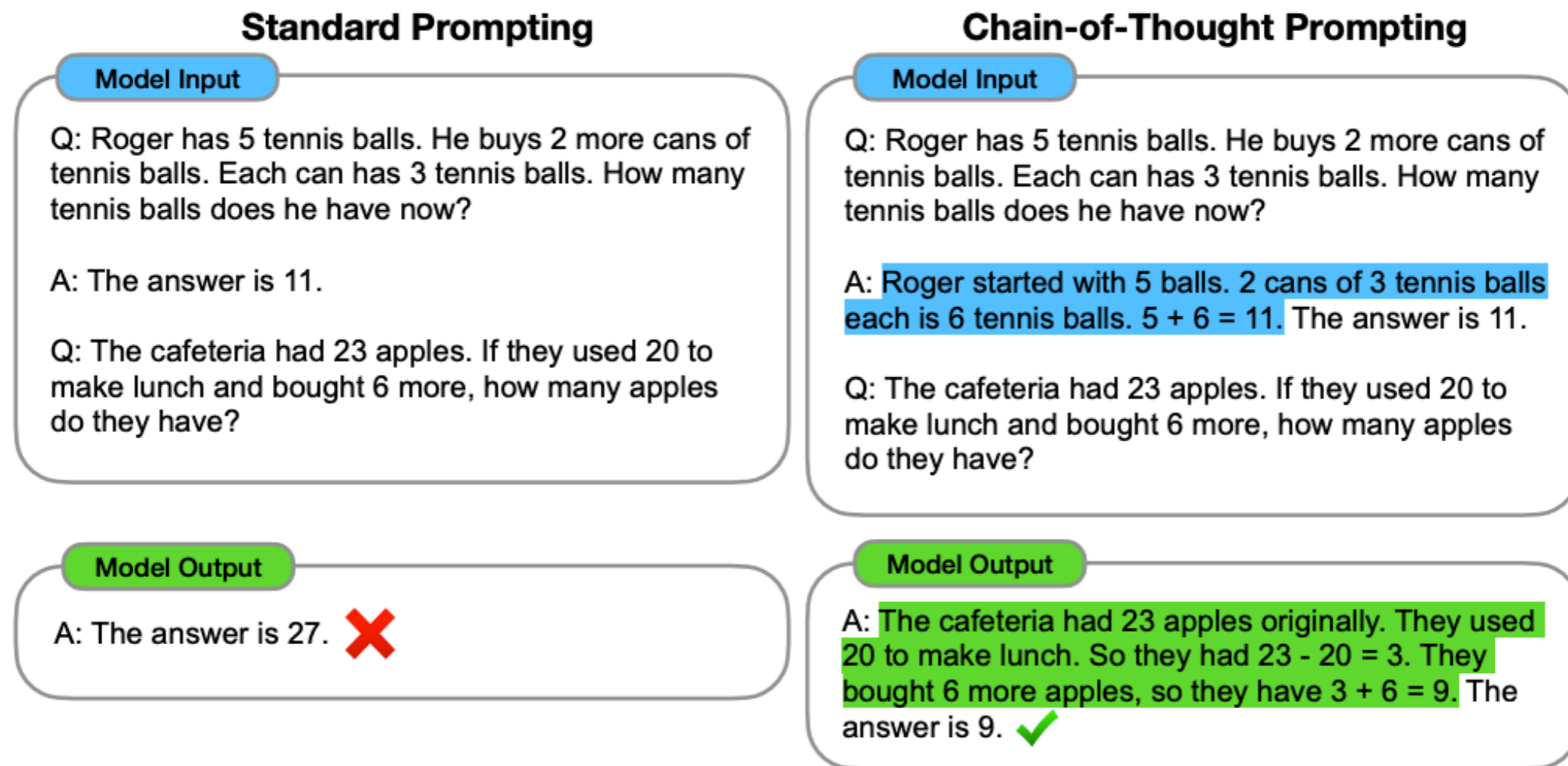


Figure 1: Chain-of-thought prompting enables large language models to tackle complex arithmetic, commonsense, and symbolic reasoning tasks. Chain-of-thought reasoning processes are highlighted.

Review: Zero-shot Chain of Thought (Kojima et al. 2022)

- Just prompt the model w/ “let’s think step-by-step”

(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. **X**

(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. **✓**

(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 **X**

(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. **✓**

Self Ask (Press et al. 2022)

- Identify when follow-up/clarification questions are necessary

Direct Prompting

GPT-3

Question: Who lived longer, Theodor Haecker or Harry Vaughan Watkins?

Answer: Harry Vaughan Watkins.

Question: Who was president of the U.S. when superconductivity was discovered?

Answer: Franklin D. Roosevelt



Chain of Thought

GPT-3

Question: Who lived longer, Theodor Haecker or Harry Vaughan Watkins?

Answer: Theodor Haecker was 65 years old when he died. Harry Vaughan Watkins was 69 years old when he died.

So the final answer (the name of the person) is: Harry Vaughan Watkins.

Question: Who was president of the U.S. when superconductivity was discovered?

Answer: Superconductivity was discovered in 1911 by Heike Kamerlingh Onnes. Woodrow Wilson was president of the United States from 1913 to 1921. So the final answer (the name of the president) is: Woodrow Wilson.



Self-Ask

GPT-3

Question: Who lived longer, Theodor Haecker or Harry Vaughan Watkins?

Are follow up questions needed here: Yes.

Follow up: How old was Theodor Haecker when he died?

Intermediate answer: Theodor Haecker was 65 years old when he died.

Follow up: How old was Harry Vaughan Watkins when he died?

Intermediate answer: Harry Vaughan Watkins was 69 years old when he died.

So the final answer is: Harry Vaughan Watkins

Question: Who was president of the U.S. when superconductivity was discovered?

Are follow up questions needed here: Yes.

Follow up: When was superconductivity discovered?

Intermediate answer: Superconductivity was discovered in 1911.

Follow up: Who was president of the U.S. in 1911?

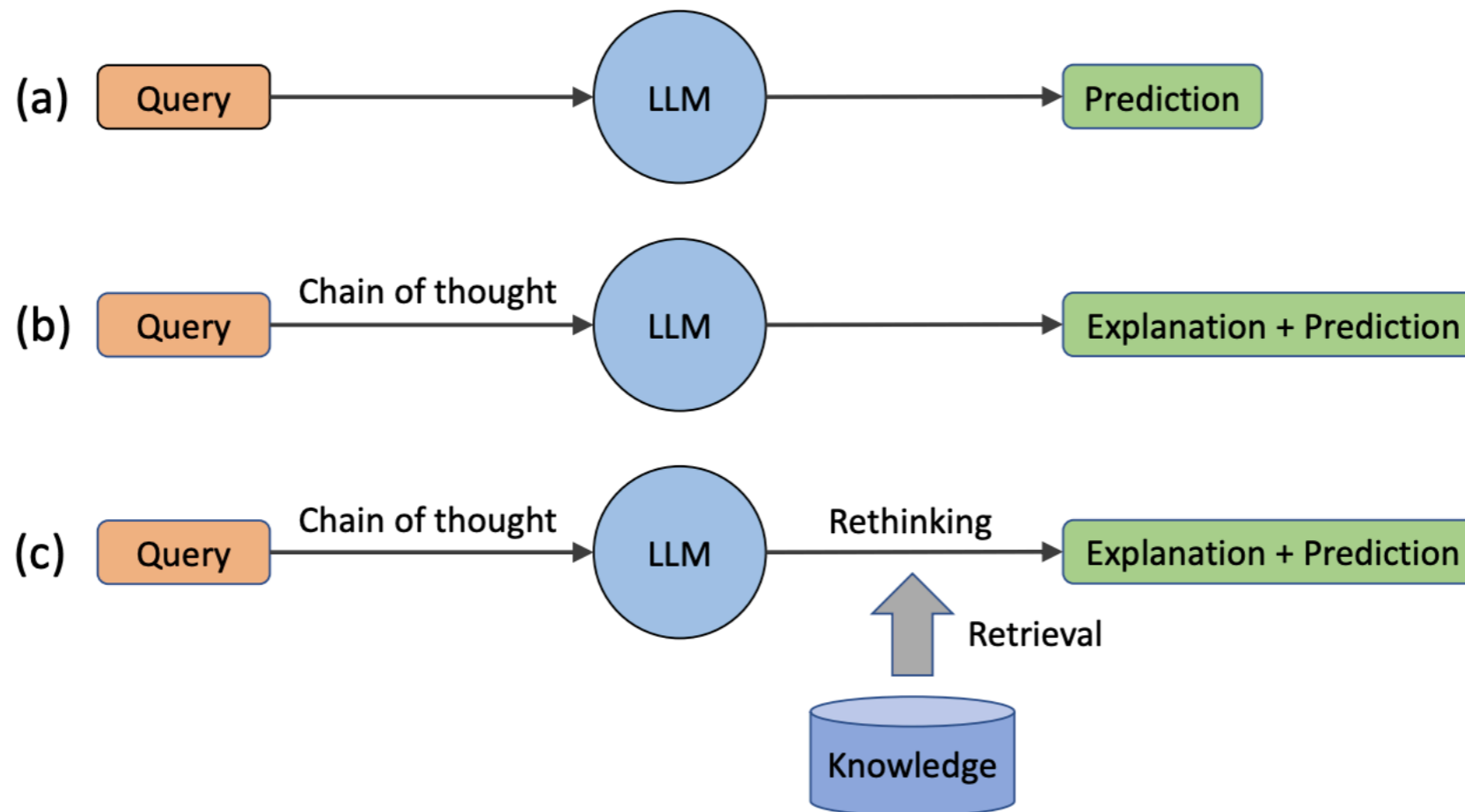
Intermediate answer: William Howard Taft.

So the final answer is: William Howard Taft.



Chain of Thought w/ Retrieval (He et al. 2023)

- Add relevant sentences during the CoT process



- Retrieval done w/ standard BM25 over Wikipedia

Multilingual Chain of Thought Reasoning (Shi et al. 2022)

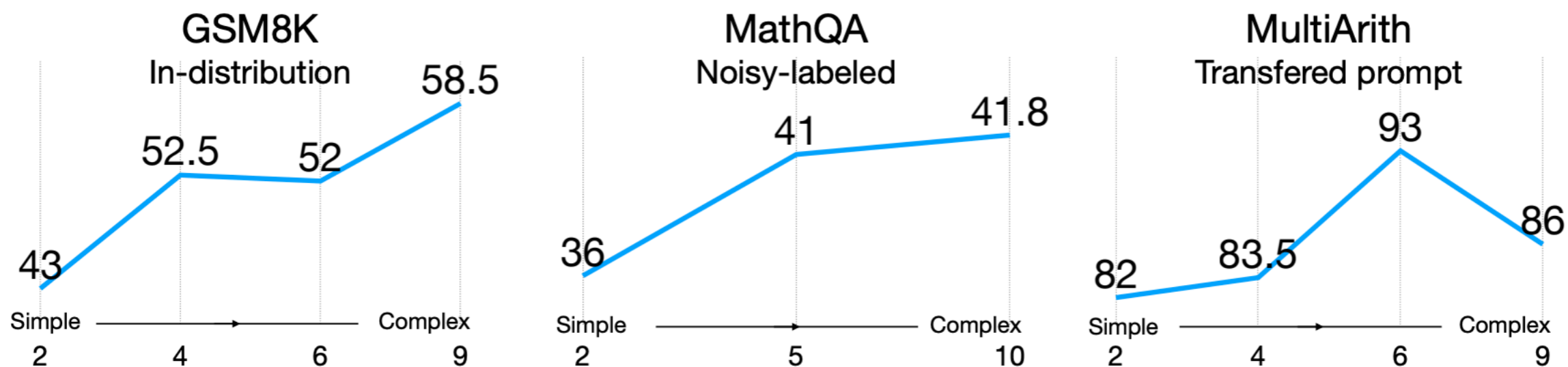
- Do we reason in another language or English?

Original Question	Frage: Roger hat 5 Tennisbälle. Er kauft noch 2 Dosen Tennisbälle. In jeder Dose sind 3 Tennisbälle. Wie viele Tennisbälle hat er jetzt?
DIRECT	Antwort: 11
NATIVE-CoT	Schritt-für-Schritt-Antwort: Roger begann mit 5 Bällen. 2 Dosen von jeweils 3 Tennisbällen macht 6 Tennisbälle. $5 + 6 = 11$. Die Antwort ist 11.
EN-CoT	Step-by-Step Answer: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.
Translated English Question	Question: 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?
TRANSLATE-EN	Step-by-Step Answer: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

	AVG	HRL	URL	EN	DE	FR	ES	RU	ZH	JA	TH	TE	BN	SW
Lang. Freq. (PaLM, %)	–	–	–	78.0	3.5	3.3	2.1	.53	.40	.38	.04	.02	.006	.005
PaLM-540B														
• DIRECT	18.6	19.3	16.8	22.0	18.8	19.6	20.0	22.0	19.2	16.0	16.8	17.6	17.2	15.6
• NATIVE-CoT	48.1	47.9	44.9	62.4	49.2	46.4	56.8	48.4	46.8	40.0	52.8	45.6	46.0	35.2
• EN-CoT	51.3	52.3	46.8	62.4	53.6	51.2	58.0	55.6	46.0	49.6	49.6	46.8	46.4	44.4
• TRANSLATE-EN	55.0	56.3	51.2	62.4	57.2	55.2	60.0	59.6	55.6	50.0	50.8	49.6	53.2	51.2

Complexity-based Prompting (Fu et al. 2022)

- Interestingly, for some tasks a larger number of reasoning steps is indicative of improved accuracy



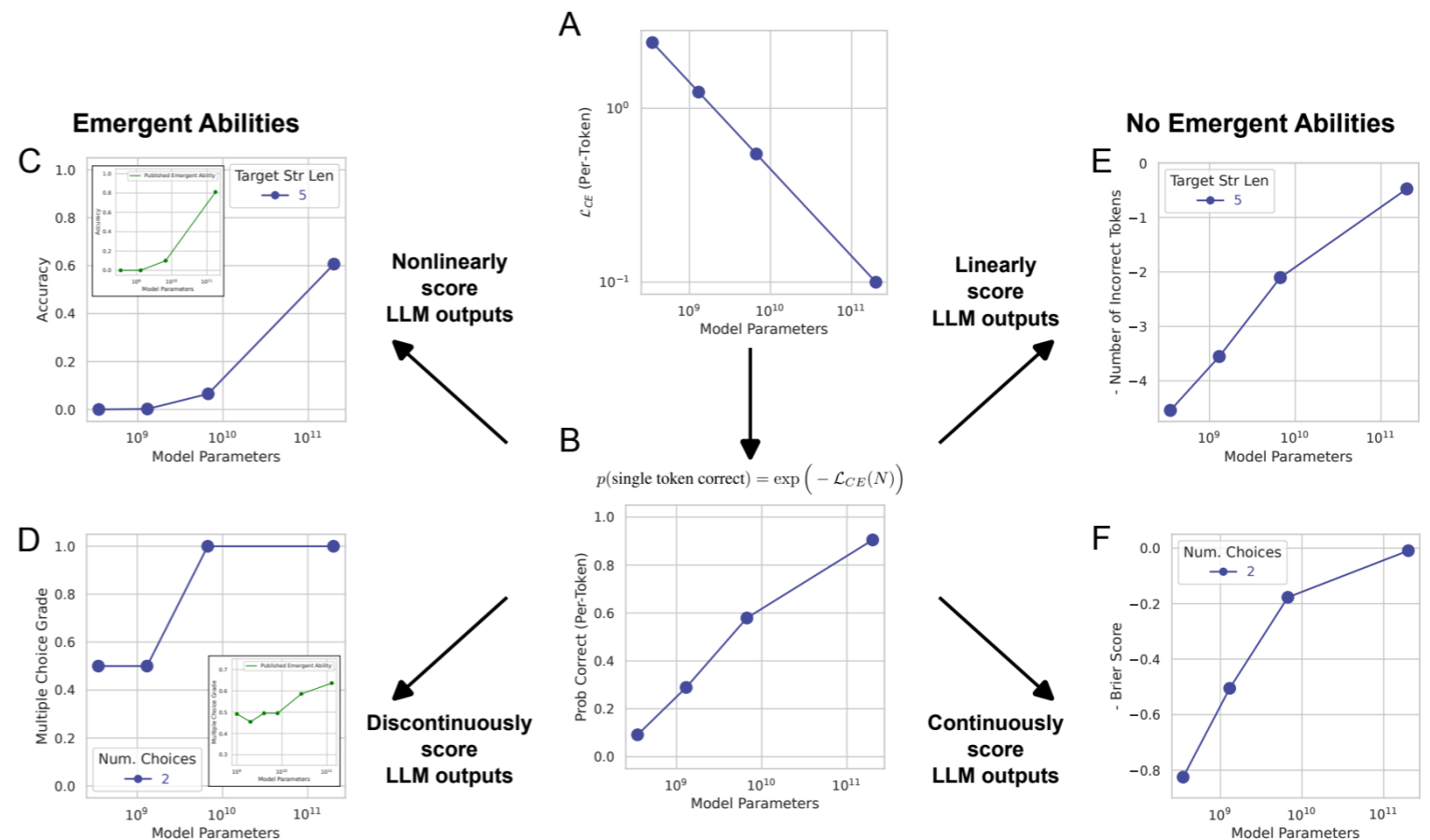
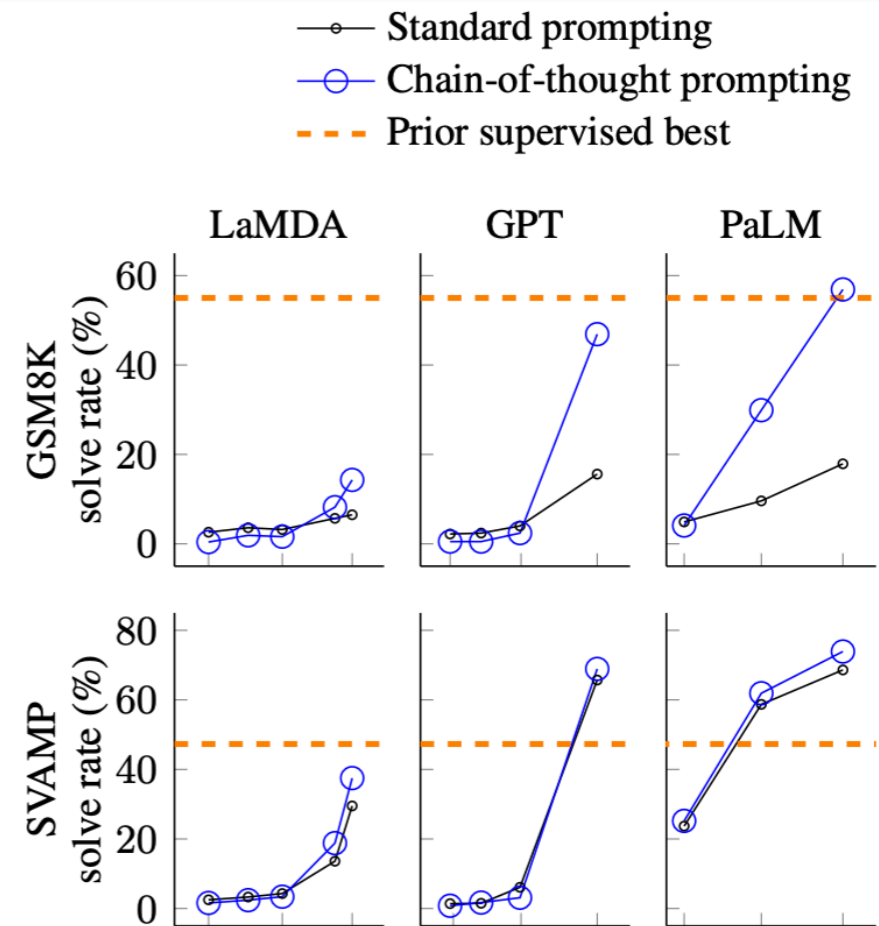
- This can be used to simply improve accuracy:
 - Sample multiple reasoning paths
 - Perform self-consistency over the longer reasoning paths

Systematic Studies of Reasoning in LLMs

Reasoning is an “Emergent” Ability

(Wei et al. 2022)

- Emergent abilities — only appear when models are very large
- Note: emergent abilities are somewhat an artifact of how we measure accuracy (Schaeffer et al. 2023)



Reliability of Explanations

(Ye and Durrett 2022)

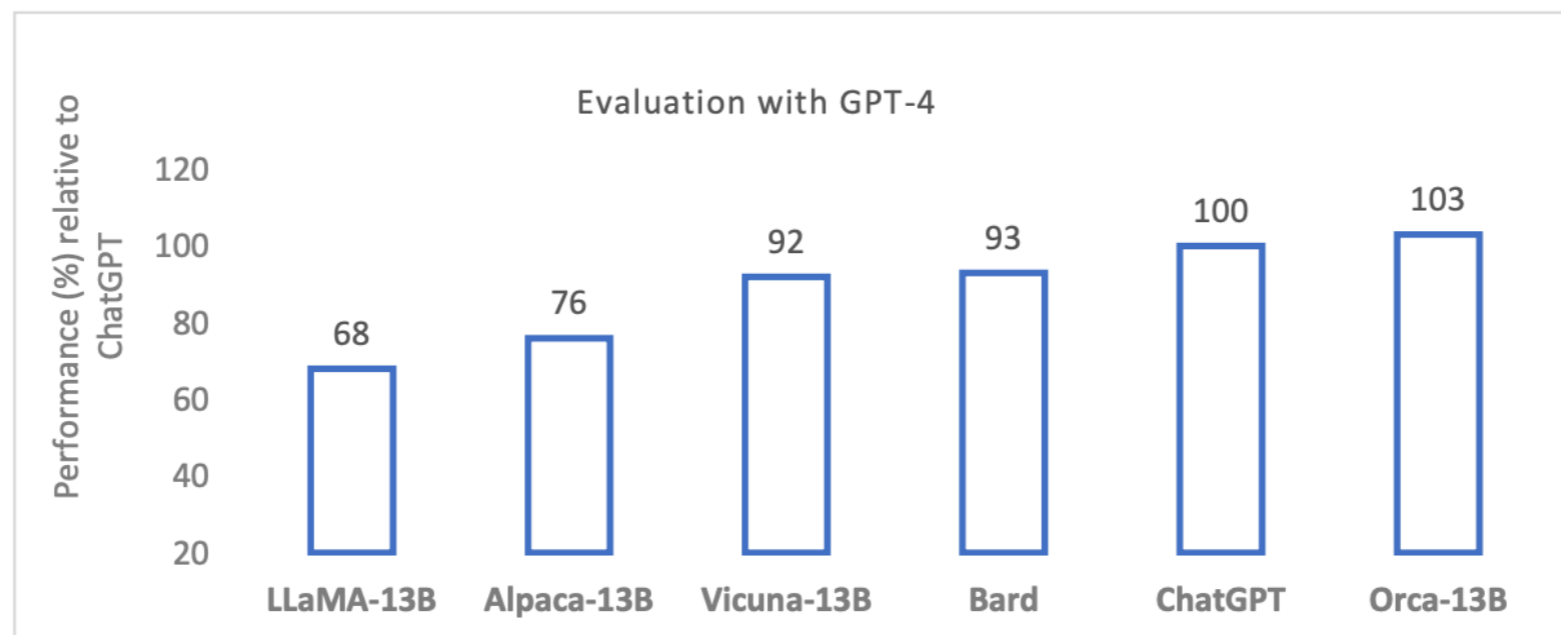
- Examine the relationship between
 - accuracy of the answer
 - factuality of the explanation
 - consistency of the answer+explanation
- E-P: explain then predict (CoT)
- P-E: predict then explain

		Acc	Fac	Con	Acc=Fac	Acc=Con
<i>reliability of explanations generated by InstructGPT</i>						
InstructGPT	SYNTH (E-P)	58.4	72.8	64.8	66.5	68.8
	SYNTH (P-E)	54.8	51.6	95.2	89.6	57.2
	ADVHP (E-P)	62.0	79.6	91.2	80.0	68.4
	ADVHP (P-E)	54.0	69.2	82.0	77.6	67.2
	E-SNLI (P-E)	62.0	—	98.8	—	62.0
<i>reliability of explanations generated by other LLMs on SYNTH</i>						
OPT (175B)	SYNTH (E-P)	30.0	77.2	47.2	45.6	58.8
	SYNTH (P-E)	39.6	64.0	81.2	69.2	49.6
GPT-3	SYNTH (E-P)	46.8	59.2	64.8	66.8	61.2
	SYNTH (P-E)	52.4	52.4	83.2	78.4	58.0
text-davinci-002	SYNTH (E-P)	86.0	91.6	85.2	91.2	84.8
	SYNTH (P-E)	81.6	83.2	96.4	95.8	82.8

Training for Chain of Thought

ORCA: Training Small Models for Reasoning (Mukherjee et al. 2024)

- Generates a large and diverse chain-of-thought dataset from GPT-3.5 and GPT-4
- 5M complex instructions + CoT explanations
- Corresponding high accuracy compared to other datasets



- Replicated in OpenORCA dataset

Chain of Thought Reward Models

(Lightman et al. 2023)

- Get human supervision on the steps

The denominator of a fraction is 7 less than 3 times the numerator. If the fraction is equivalent to $2/5$, what is the numerator of the fraction? (Answer: 14)

Let's call the numerator x.

So the denominator is $3x-7$.

We know that $x/(3x-7) = 2/5$.

So $5x = 2(3x-7)$.

$5x = 6x - 14$.

So $x = 7$.

Let

$$x^8 + 3x^4 - 4 = p_1(x)p_2(x) \cdots p_k(x),$$

where each non-constant polynomial $p_i(x)$ is monic with integer coefficients, and cannot be factored further over the integers. Compute $p_1(1) + p_2(1) + \cdots + p_k(1)$.

I notice that the given polynomial has even degree and only even powers of x , so I can try to make a substitution to simplify it.

Let $y = x^4$, then the polynomial becomes $y^2 + 3y - 4$, which is a quadratic equation.

I can factor this quadratic equation as $(y + 4)(y - 1)$, so the original polynomial is $(x^4 + 4)(x^4 - 1)$.

Now I need to factor each of these factors further, if possible, over the integers.

For the first factor, $x^4 + 4$, I recall the Sophie Germain identity, which says that $a^4 + 4b^4 = (a^2 + 2b^2 + 2ab)(a^2 + 2b^2 - 2ab)$.

Applying this identity with $a = x$ and $b = 1$, I get $x^4 + 4 = (x^2 + 2x + 2)(x^2 - 2x + 2)$.

For the second factor, $x^4 - 1$, I can use the difference of squares formula, which says that $a^2 - b^2 = (a + b)(a - b)$.

Applying this formula with $a = x^2$ and $b = 1$, I get $x^4 - 1 = (x^2 + 1)(x^2 - 1)$.

Then I can apply the difference of squares formula again with $a = x$ and $b = 1$, to get $x^2 - 1 = (x + 1)(x - 1)$.

So the original polynomial is $(x^4 + 4)(x^4 - 1) = (x^2 + 2x + 2)(x^2 - 2x + 2)(x^2 + 1)(x + 1)(x - 1)$.

These are all the irreducible factors over the integers, so $p_1(x) = x^2 + 2x + 2$, $p_2(x) = x^2 - 2x + 2$, $p_3(x) = x^2 + 1$, $p_4(x) = x + 1$, and $p_5(x) = x - 1$.

To find the sum of their values at $x = 1$, I just need to plug in $x = 1$ into each factor and add them up.

I get $p_1(1) = 1^2 + 2 \cdot 1 + 2 = 5$, $p_2(1) = 1^2 - 2 \cdot 1 + 2 = 1$, $p_3(1) = 1^2 + 1 = 2$, $p_4(1) = 1 + 1 = 2$, and $p_5(1) = 1 - 1 = 0$.

So the final answer is $5 + 1 + 2 + 2 + 0 = 10$.

Answer: 10

Let

$$x^8 + 3x^4 - 4 = p_1(x)p_2(x) \cdots p_k(x),$$

where each non-constant polynomial $p_i(x)$ is monic with integer coefficients, and cannot be factored further over the integers. Compute $p_1(1) + p_2(1) + \cdots + p_k(1)$.

This problem is about factoring a polynomial over the integers, which can be a tricky task.

One possible strategy is to look for rational roots of the polynomial, using the rational root theorem.

However, that would involve checking a lot of possible factors of -4 , and there is no guarantee that the polynomial has any rational roots at all.

Another possible strategy is to use some algebraic tricks to simplify the polynomial or reduce its degree.

I notice that the polynomial has only even powers of x , so I can try to substitute $y = x^2$ and see what happens.

Then the polynomial becomes $y^4 + 3y^2 - 4$, which is a quadratic equation in y^2 .

I can try to factor this quadratic equation by looking for two numbers that multiply to -4 and add to 3 .

Those numbers are 4 and -1 , so I can write $y^4 + 3y^2 - 4 = (y^2 + 4)(y^2 - 1)$.

Now I can substitute back $x^2 = y$ and get $x^8 + 3x^4 - 4 = (x^4 + 4)(x^4 - 1)$.

This looks promising, because I have reduced the degree of the factors from 8 to 4 .

I can try to factor each of these factors further, using the difference of squares formula.

I get $x^8 + 3x^4 - 4 = (x^2 + 2)(x^2 - 2)(x^2 + 1)(x^2 - 1)$.

I can apply the difference of squares formula again to the last factor and get $x^8 + 3x^4 - 4 = (x^2 + 2)(x^2 - 2)(x^2 + 1)(x + 1)(x - 1)$.

Now I have factored the polynomial completely into monic linear and quadratic factors with integer coefficients.

These are the $p_i(x)$'s that the problem is asking for.

To find the sum of their values at $x = 1$, I just need to plug in $x = 1$ into each factor and add them up.

I get $p_1(1) + p_2(1) + \cdots + p_k(1) = (1^2 + 2)(1^2 - 2)(1^2 + 1)(1 + 1)(1 - 1)$.

Simplifying, I get $p_1(1) + p_2(1) + \cdots + p_k(1) = (3)(-1)(2)(2)(0)$.

Multiplying, I get $p_1(1) + p_2(1) + \cdots + p_k(1) = 0$.



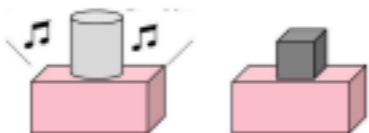
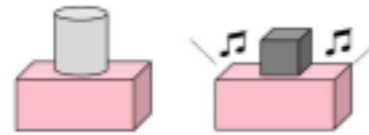
Answer: 0

- Train reward model, give feedback on each sentence

Abductive Reasoning: Learning Patterns from Data

Inference to Explanations

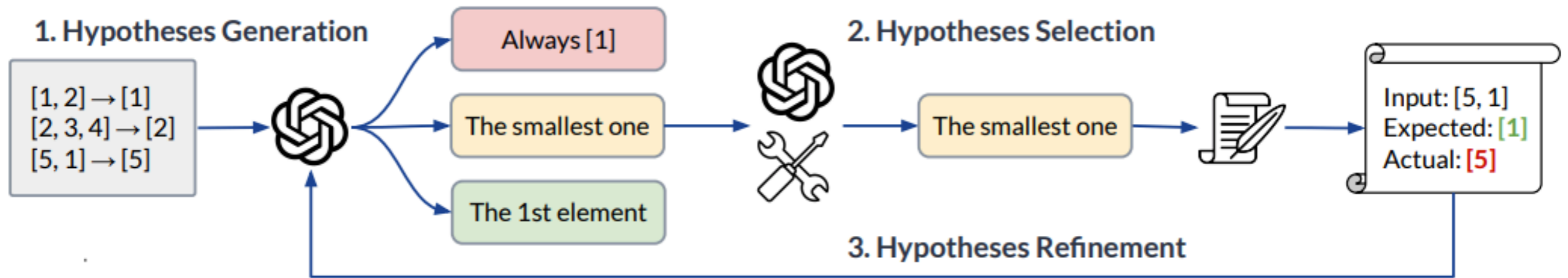
- Can we find a rule that underlies a pattern in data?

	ACRE	MiniSCAN	List Functions	MiniARC
Examples		dax → ● lug → ● lug fep → ● ● ● dax fep → ● ● ●	[1, 2, 3] → [1] [2, 3, 4] → [2] [5, 1] → [5]	
Bad Rule		dax → ● lug → ● [X] fep → ● ●	The smallest one	Swap the colors of two objects
Good Rule		dax → ● lug → ● [X] fep → ● ● ●	The 1st element	Drop all objects

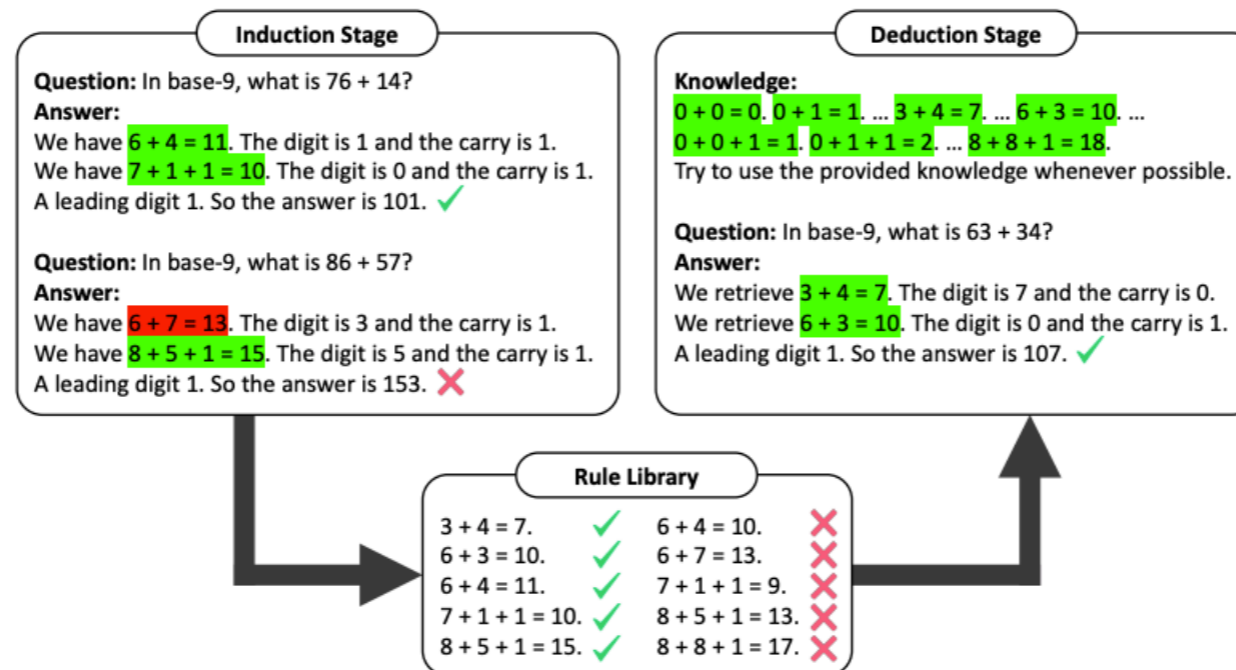
- Similar to program induction

Rule Induction w/ LLMs

- Propose hypotheses, verify with symbolic verifier (Qiu et al. 2023)



- Use hypotheses in CoT, keep ones that result in correct answers (Zhu et al. 2023)



Learning Differences between Text Collections (Zhong et al. 2023)

- Answers the question “what are salient differences between text collections”?

Proposer prompt

Samples from the two corpora {

Group A: The Manchester United soccer squad welcomes rising star Juan Silva, ...
Group A: As Serena Willows joins the UCLA women's tennis roster,
Group A: // 20 samples not included for brevity

Group B: Egypt's President Abdel Fattah el-Sisi and Saudi Arabia's ...
Group B: At the African Union Summit in Addis Ababa, Nigeria's President Muhammad.
Group B: // 20 samples not included for brevity

(some of the sentences are truncated for brevity) ...

Exploration Goal {

The original dataset includes news summaries. The two corpora are generated based on when they were published. Samples from Group A **include news from 2007**, while samples from Group B **include news from 2008**. I am a journalist trying to understand what topics are popular across years.

Formatting Instructions {

Please write a list of hypotheses (separated by bullet points "-") of how datapoints from Group A differ from those from Group B. Each hypothesis should be formatted as a sentence fragment. Here are three examples.

- "talks about politics, such as presidential election."
- "contains insulting language for immigrants."
- "uses double negation, i.e., using two negations in a sentence."

Language Model Outputs {

Based on the two sentence groups (A and B) from the above, more sentences in Group A ...

- **"mentions a sports team recruiting a new member"**
- **"mentions about academic relations, such as teachers or students"**
- **"mentions about // a list of hypotheses not included for brevity"**

Validator prompt

Check whether the TEXT satisfies a PROPERTY. Respond with Yes or No. When uncertain, output No.

Now complete the following example -
input: PROPERTY: **mentions a sports team recruiting a new member**
TEXT: **As Serena Willows joins the UCLA women's tennis roster**
output:

Pr[NextWord = "Yes"] = 99%

Check whether the TEXT satisfies a PROPERTY. Respond with Yes or No. When uncertain, output No.

Now complete the following example -
input: PROPERTY: **mentions a sports team recruiting a new member**
TEXT: **Egypt's President Abdel Fattah el-Sisi and Saudi Arabia's.**
output:

Pr[NextWord = "Yes"] = 5%

Questions?