CS 839: Foundation Models

Chain-of-Thought

Fred Sala

University of Wisconsin-Madison

Oct. 3, 2023
Announcements

• Logistics:
  • Homework 1 due in a week

• Class roadmap:

<table>
<thead>
<tr>
<th>Tuesday Oct. 3</th>
<th>Reasoning &amp; Chain-of-Thought</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thursday Oct. 5</td>
<td>In-Context Learning: Practice and Theory</td>
</tr>
<tr>
<td>Tuesday Oct. 10</td>
<td>Fine-Tuning, Specialization, Adaptation</td>
</tr>
<tr>
<td>Thursday Oct. 12</td>
<td>Training</td>
</tr>
<tr>
<td>Tuesday Oct. 17</td>
<td>RLHF</td>
</tr>
</tbody>
</table>
Outline

• **Review of Chain-of-Thought + ToT Search**
  • Basic idea, zero-shot and few-shot, choosing examples for few-shot, tree-of-thoughts

• **Ingredients For CoT**
  • Scale, manual vs. auto-rationales, rationale correctness

• **Connecting to Externals**
  • Programs, tools, etc.
Outline

• Review of Chain-of-Thought + ToT Search
  • Basic idea, zero-shot and few-shot, choosing examples for few-shot, tree-of-thoughts

• Ingredients For CoT
  • Scale, manual vs. auto-rationales, rationale correctness

• Connecting to Externals
  • Programs, tools, etc.
New Auto-Prompting Work

Recall: search for hard prompts: tough optimization problem
• Lots of classic search methods only require notion of “neighbors” and evaluation function access
  • Hill-climbing
  • Simulated annealing
  • Genetic algorithms

• “Promptbreeder”: a new approach via genetic algorithms
  • Show all your working. II. You should use the correct mathematical notation and vocabulary, where appropriate. III. You should write your answer in full sentences and in words. IV. You should use examples to illustrate your points and prove your answers. V. Your workings out should be neat and legible
New Auto-Prompting Work

“Promptbreeder”: a new approach via genetic algorithms

Fernando et al ‘23
Performing complex reasoning is hard. Help the model:

- **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?
    - A: The answer is 27. ✗

- **Chain-of-Thought Prompting**
  - **Model Input**
    - Q: 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?
    - 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. ✓
Chain-of-Thought: **Applications**

Some things it can be used for:

- **Math problems**
- **Commonsense reasoning**
- **Symbolic reasoning**

<table>
<thead>
<tr>
<th>Math Word Problems (free response)</th>
<th>Math Word Problems (multiple choice)</th>
<th>CSQA (commonsense)</th>
</tr>
</thead>
<tbody>
<tr>
<td>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?</td>
<td>Q: How many keystrokes are needed to type the numbers from 1 to 5007? Answer Choices: (a) 1156 (b) 1392 (c) 1480 (d) 1562 (e) 1788</td>
<td>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</td>
</tr>
<tr>
<td>A: Roger started with 5 balls. 2 cans of 8 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is $11$.</td>
<td>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 $1392$.</td>
<td>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).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>StrategyGA</th>
<th>Date Understanding</th>
<th>Sports Understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q: Yes or no: Would a pear sink in water?</td>
<td>Q: The concert was scheduled to be on 08/01/1943, but was delayed by one day to today. What is the date 10 days ago in MM/DD/YYYY?</td>
<td>Q: Is the following sentence plausible? &quot;Joao Moutinho caught the screen pass in the NFC championship.&quot;</td>
</tr>
<tr>
<td>A: The density of a pear is about 0.6 g/cm³, which is less than water. Thus, a pear would float. So the answer is no.</td>
<td>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.</td>
<td>A: Joao Moutinho is a soccer player. The NFC championship is part of American football, not soccer. So the answer is no.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SayCan (Instructing a robot)</th>
<th>Last Letter Concatenation</th>
<th>Coin Flip (state tracking)</th>
</tr>
</thead>
<tbody>
<tr>
<td>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().</td>
<td>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.</td>
<td>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.</td>
</tr>
</tbody>
</table>
Chain-of-Thought: Zero-Shot

No examples shown; encourage model to decompose

- Add to prompt: “Let’s think step by step” before each answer
- For answer extraction, add prompts like “Therefore, the answer (arabic numerals) is” (Kojima et al ‘23)

(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. ✗

(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 ✗

(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. ✗
Chain-of-Thought: Few-Shot Examples

As before, we must choose few-shot examples.

• More structured than simple semantic similarity
• **Complexity-based** prompting.
  • “[S]imply choose complex prompts over simple ones.”
• Promoting: include most steps. Ensembling: MV over set of most complex chains.

![Diagram: Workflow of chain of thoughts prompting](image)

A. Workflow of chain of thoughts prompting

Asia bought a homecoming dress on sale for $140. It was originally priced at $350. What percentage off did she get at the sale?

1. Asia saved $350 - $140 = $210 on the dress.
2. That means she saved $210 / $350 = 0.60 or 60% off on the dress.

The answer is 60

... < more CoT cases > ...

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

1. Angelo and Melanie want to plan how many hours ... how many days should they plan to study total over the next week if they take a 10-minute break every hour ... ?
2. Angelo and Melanie think they should dedicate 3 hours to each of the 2 chapters ...
3. For the worksheets they plan to dedicate 1.5 hours for each worksheet ...
4. Angelo and Melanie need to start with planning 12 hours to study, at 4 hours a day, 12 / 4 = 3 days, ...
5. They want to study no more than 4 hours each day, 15 hours / 4 hours each day = 3.75
6. They will need to plan to study 4 days to allow for all the time they need.

The answer is 4

B. Example complex chain, 9 reasoning steps

![Diagram: CoT prompt + Question](image)

Cos et al ’23
Chain-of-Thought: Generalizations

How do we really “reason”?  
• Not really by sampling a bunch of chains...

Yao et al ‘23
**Chain-of-Thought: Generalizations**

**Tree-of-thoughts basic idea:**

- **Notation:** thoughts $z_1, z_2, ..., z_n$ bridge $x$ and $y$

- Comparison to other methods:
  - Vanilla CoT: sample $z_i \sim p_\theta(z_i \mid x, z_1, ..., z_{i-1})$, $y \sim p_\theta(y \mid x, z_1, ..., z_n)$
  - CoT Self-Consistency: sample multiple times, take majority vote

- Idea: create a state $s=[x,z_1,...,z_n]$

- Generate multiple candidates for next state
  - Then run standard search (i.e., BFS, DFS, A*)

Drichel (Wiki)
Chain-of-Thought: Generalizations

Tree-of-thoughts key aspects:

• **Thought decomposition**: how big zs should be
• **Thought generation**: obtaining the next sample
  • Try to avoid duplication
• **State evaluation**: How close are we to solution?
  • Recall heuristics for search from CS 540
  • Either use LM itself, or vote/weighted vote across solutions
• **Search**: BFS or DFS
  • Or more advanced search methods

Drichel (Wiki)
Chain-of-Thought: Generalizations

Tree-of-thoughts example:

4.1 Game of 24

Game of 24 is a mathematical reasoning challenge, where the goal is to use 4 numbers and basic arithmetic operations (+/-/*) to obtain 24. For example, given input “4 9 10 13”, a solution output could be “(10 - 4) * (13 - 9) = 24”.

![Diagram of Game of 24](image)
Break & Questions
Outline

• Review of Chain-of-Thought + ToT Search
  • Basic idea, zero-shot and few-shot, choosing examples for few-shot, tree-of-thoughts

• Ingredients For CoT
  • Scale, manual vs. auto-rationales, rationale correctness

• Connecting to Externals
  • Programs, tools, etc.
What Matters for CoT? Scale?

Do all language models exhibit CoT behavior?

A: No. Shows up only at certain sizes
   • “Emergent behavior”
   • Example: CoT does not help until ~10B

(Model: LaMDA, Dataset: Math)

Wei et al ‘23
What Matters for CoT? Correctness?

Does reasoning have to work in provided examples?

A: **No.** Invalid reasoning in CoT still helps

- Versus no CoT
- What does matter
  - Relevant steps (still bridge gap)
  - Steps follow each other

Wang et al ‘23
What Matters for CoT? **Human signal?**

Do examples have to be manually crafted?

A: **No.** Auto-CoT: generate examples to be used

- Need diversity: first cluster, then sample from each cluster
What Matters for CoT? Pretraining?

Does “reasoning” data in pretraining extend to other languages?

A: Not entirely.

<table>
<thead>
<tr>
<th>Lang. Freq. (PaLM, %)</th>
<th>AVG</th>
<th>HRL</th>
<th>URL</th>
<th>EN</th>
<th>DE</th>
<th>FR</th>
<th>ES</th>
<th>RU</th>
<th>ZH</th>
<th>JA</th>
<th>TH</th>
<th>TE</th>
<th>BN</th>
<th>SW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>78.0</td>
<td>3.5</td>
<td>3.3</td>
<td>2.1</td>
<td>.53</td>
<td>.40</td>
<td>.38</td>
<td>.04</td>
<td>.02</td>
<td>.006</td>
<td>.005</td>
</tr>
</tbody>
</table>

**GPT-3 (text-davinci-002)**

- **DIRECT**
  - 11.7 15.1 5.7 16.0 14.8 16.8 17.2 12.4 18.0 11.2 8.8 0.8 4.4 8.8
- **NATIVE-COT**
  - 26.4 34.7 7.2 53.6 36.0 37.6 40.4 28.4 40.0 26.0 10.8 0.4 6.4 11.2
- **EN-COT**
  - 31.6 39.4 13.9 53.6 44.0 46.0 44.8 28.4 40.8 32.4 19.6 5.6 9.6 20.8
- **TRANSLATE-EN**
  - 45.6 47.5 40.7 53.6 46.4 46.4 51.6 48.8 47.2 44.8 41.2 42.8 41.2 37.6

**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

Shi et al ‘22
Break & Questions
Outline

• Review of Chain-of-Thought + ToT Search
  • Basic idea, zero-shot and few-shot, choosing examples for few-shot, tree-of-thoughts

• Ingredients For CoT
  • Scale, manual vs. auto-rationales, rationale correctness

• Connecting to Externals
  • Programs, tools, etc.
Beyond the Unaided Language Models

Even when we do CoT, the language model can get things wrong.

• Often simple things... like arithmetic.

• How else can we help it?

• A: Use external tools
Tools: Program-aided LMs

Use external tools:
- Python interpreter

- How? *Interleave* the text explanations in CoT steps with lines of Python code

- LMs can already output code
  - Just need to *prompt* the right way

Gao et al ‘23
Tools: Programs-aided LMs **Prompts**

Just need to *prompt* the right way

- Craft examples that interleave code and text

---

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

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

---

**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 scrunchie phone charger, a mauve fidget spinner, and a burgundy pen. What is the color of the object directly to the right of the stress ball?

```python
... 
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]
```
Tools: Program-of-Thoughts

Similar idea: program-of-thoughts

Question: In Fibonacci sequence, it follows the rule that each number is equal to the sum of the preceding two numbers. Assuming the first two numbers are 0 and 1, what is the 50th number in Fibonacci sequence?

<table>
<thead>
<tr>
<th>CoT</th>
<th>PoT</th>
</tr>
</thead>
</table>
| The first number is 0, the second number is 1, therefore, the third number is 0+1=1. The fourth number is 1+1=2. The fifth number is 1+2=3. The sixth number is 2+3=5. The seventh number is 3+5=8. The eighth number is 5+8=13. .... (Skip 1000 tokens) The 50th number is 32,432,268,459. | length_of_fibonacci_sequence = 50
fibonacci_sequence = np.zeros(length_of_)
fibonacci_sequence[0] = 0
fibonacci_sequence[1] = 1
For i in range(3, length_of_fibonacci_sequence):
    fibonacci_sequence[i] = fibonacci_sequence[i-1] +
fibonacci_sequence[i-2]
ans = fibonacci_sequence[-1] |

32,432,268,459

Question: Kitty saves 20000 dollars to the bank. After three years, the sum with compound interest rate is 1000 dollars more than the sum with simple interest rate. What is the interest rate of the bank?

<table>
<thead>
<tr>
<th>CoT</th>
<th>PoT</th>
</tr>
</thead>
</table>
| Assuming the interest rate is x. The sum after two years with simple interest rate is 20000 + x * 20000 * 3 = 20000 + 60000x. The sum after two years with compound interest rate is 20000 * (1 + x) ^ 3 = 20000 + 60000 + 60000x + 60000x*x + 20000x*x^2 + 20000x^3. The difference can be written as 60000x*x^2 + 20000x^3 = 1000. In order to solve x, we can use the quadratic formula, x = (-b + sqrt(b^2 - 4ac)) / 2a, ...., x = (-20000 + 6160) / 120000, x = -0.051333. | interest_rate = Symbol("x")
sum_in_two_years_with_simple_interest= 20000 + interest_rate * 20000 * 3
sum_in_two_years_with_compound_interest = 20000 * (1 + interest_rate)^3
# Since compound interest is 1000 more than simple interest.
ans = solve(sum_after_in_yeras_with_compound_interest - sum_after_two_years_in_compound_interest - 1000, interest_rate) |

-0.051333

Figure 1: Comparison between Chain of Thoughts and Program of Thoughts.

Chen et al ‘22
Ideally, use more general external tools
- Without lots of human annotation
- Model should decide on its own which tool to use

- **Toolformer**: introduces API calls into the model
  - But these API calls aren’t already there... so need to fine-tune
Bibliography

- Zhang et al ’23: Zhuosheng Zhang, Aston Zhang, Mu Li, Alex Smola, “Automatic Chain of Thought Prompting in Large Language Models” (https://openreview.net/forum?id=5NTt8GFjUHkr)
Thank You!