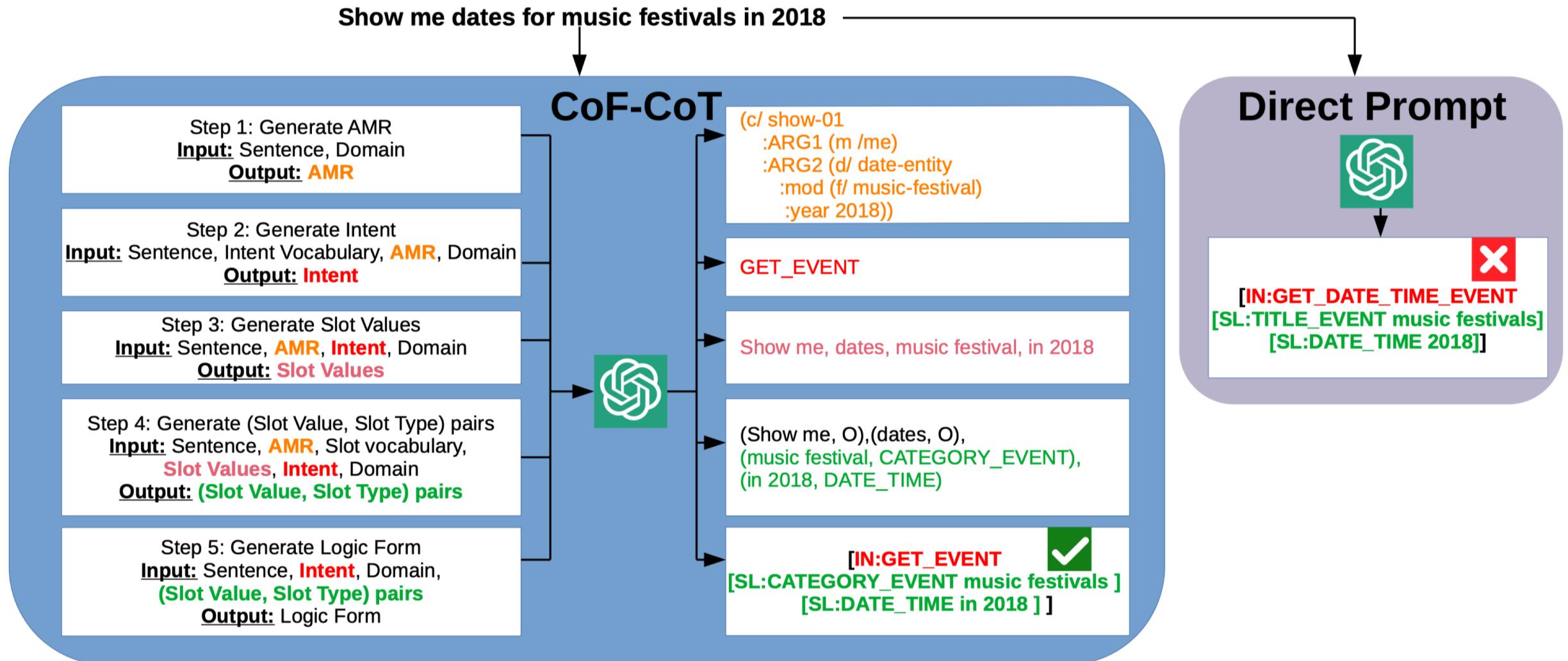


Coarse-to-Fine CoT Prompting

- 由粗到细的 CoT
 - 先解决难度较低的粗粒度问题
 - 再解决难度较高的细粒度问题
 - 使用粗粒度任务的输出来引导后续的细粒度任务
- 结构表示可以增强 PLM
 - 依存 (Dependency Parsing) 图、Constituency Parsing 图
 - 基于语义的 AMR 图

Multi-grained NLU tasks



Results (vs. CoT)

Model	MTOP							
	Zero-shot				Few-shot			
	NLU		Semantic Parsing		NLU		Semantic Parsing	
	Intent Acc	Slot F1	Frame Acc	Exact Match	Intent Acc	Slot F1	Frame Acc	Exact Match
Direct Prompt	31.50 ± 1.80	21.84 ± 2.83	8.33 ± 1.44	6.00 ± 1.32	51.33 ± 3.40	28.35 ± 3.24	11.00 ± 1.80	8.33 ± 1.00
CoT	31.83 ± 2.02	22.40 ± 1.61	8.67 ± 0.35	6.33 ± 1.04	47.67 ± 5.20	28.46 ± 3.10	11.83 ± 1.53	8.50 ± 1.04
SC-CoT	32.50 ± 1.89	22.71 ± 2.44	10.05 ± 0.87	6.83 ± 0.76	53.50 ± 3.04	29.53 ± 1.99	12.50 ± 1.80	9.00 ± 0.87
ComplexCoT	32.67 ± 2.00	22.86 ± 3.17	10.83 ± 0.29	7.16 ± 0.58	48.83 ± 2.47	29.21 ± 2.65	13.17 ± 0.58	8.83 ± 2.89
Least-to-Most	45.67 ± 0.58	21.84 ± 1.91	14.50 ± 0.50	8.00 ± 0.50	49.83 ± 4.54	27.28 ± 2.41	16.00 ± 0.50	8.83 ± 0.76
Plan-and-Solve	45.00 ± 4.00	22.45 ± 2.28	9.50 ± 1.61	8.25 ± 2.25	—	—	—	—
CoF-CoT	57.67 ± 2.75	23.47 ± 4.09	14.33 ± 1.52	9.00 ± 1.00	61.50 ± 4.93	30.12 ± 3.93	15.00 ± 1.32	11.00 ± 1.61
MASSIVE								
Direct Prompt	72.50 ± 4.58	33.24 ± 3.34	24.17 ± 3.79	20.67 ± 3.28	75.17 ± 0.58	42.36 ± 2.98	29.00 ± 5.39	24.50 ± 4.07
CoT	71.83 ± 2.57	36.32 ± 1.94	24.50 ± 2.29	21.66 ± 3.40	76.83 ± 3.82	44.89 ± 2.50	31.33 ± 0.87	25.83 ± 2.25
SC-CoT	73.05 ± 1.27	37.06 ± 2.54	27.16 ± 3.21	22.50 ± 2.65	77.33 ± 2.89	47.02 ± 4.60	34.00 ± 3.21	27.16 ± 3.50
ComplexCoT	73.66 ± 3.65	37.64 ± 3.51	25.83 ± 2.25	22.16 ± 2.51	77.83 ± 1.83	46.59 ± 2.43	36.50 ± 2.89	28.00 ± 3.69
Least-to-Most	72.83 ± 4.65	37.62 ± 1.69	31.50 ± 1.53	26.50 ± 1.26	77.00 ± 3.28	45.93 ± 3.99	32.50 ± 4.09	29.00 ± 5.11
Plan-and-Solve	69.33 ± 2.47	38.07 ± 2.07	32.00 ± 1.26	29.00 ± 1.26	—	—	—	—
CoF-CoT	89.00 ± 2.29	38.66 ± 3.25	33.17 ± 4.04	25.50 ± 2.64	92.00 ± 2.29	47.06 ± 4.63	37.50 ± 1.89	29.50 ± 3.12

Results (vs. Fine-tuning)

Method	Assumption	Intent Acc	Slot F1	Frame Acc	Exact Match
RoBERTa FT	Supervised	67.19 ± 2.90	75.17 ± 1.08	43.57 ± 4.18	36.10 ± 1.08
RoBERTa FT	ZSL	0	12.68 ± 1.25	0	0
RoBERTa FT	FSL	0	13.75 ± 1.22	0	0
CoF-CoT	ZSL	57.67 ± 2.75	23.47 ± 4.09	14.33 ± 1.52	9.00 ± 1.00
CoF-CoT	FSL	61.50 ± 4.93	30.12 ± 3.93	15.00 ± 1.32	11.00 ± 1.61

	Intent Acc	Slot F1	Frame Acc	Exact Match
CoT (w/o structure)	57.16 ± 3.69	17.50 ± 2.92	12.16 ± 1.61	4.67 ± 3.33
CP-CoT	57.33 ± 3.25	19.34 ± 3.34	13.16 ± 1.04	5.50 ± 1.32
DP-CoT	57.50 ± 3.01	17.83 ± 2.53	12.67 ± 1.04	5.83 ± 2.08
AMR-CoT	57.67 ± 2.75	23.47 ± 4.09	14.33 ± 1.52	9.00 ± 1.00

步骤的重要性

Method	Assumption	Intent Acc	Slot F1	Frame Acc	Exact Match
Random-CoT	Random Order	80.67 ± 3.60	27.14 ± 2.47	26.50 ± 1.80	16.50 ± 1.04
	FoC order	83.00 ± 2.88	32.11 ± 2.50	28.50 ± 3.21	18.00 ± 3.50
CoF-CoT (w/o step 1)	No AMR	81.50 ± 4.36	33.68 ± 2.40	27.50 ± 2.65	18.00 ± 0.76
CoF-CoT (w/o step 2)	No intent	78.17 ± 4.80	27.66 ± 1.93	23.50 ± 2.78	14.50 ± 2.25
CoF-CoT (w/o step 3)	No separate KP	82.33 ± 1.04	34.63 ± 3.10	32.83 ± 2.47	23.00 ± 1.80
CoF-CoT (w/o step 4)	No separate slot prediction for KP	79.17 ± 4.01	32.92 ± 5.02	31.50 ± 3.50	21.83 ± 3.17
CoF-CoT (w/o step 3+4)	No separate slot prediction	81.33 ± 4.19	31.31 ± 3.77	27.67 ± 5.34	21.00 ± 4.92
CoF-CoT - Conditioning	CoF order (Full)	89.00 ± 2.29	38.66 ± 3.25	33.17 ± 4.04	25.50 ± 2.64
	No domain	84.50 ± 2.75	36.80 ± 2.08	32.50 ± 1.73	24.83 ± 0.58

从错误中学习

- 人类的学习过程 (error-driven learning)
 - 教科书 – Golden 数据
 - 练习 – 错在哪里？如何改正
 - 从错误中学习 – 归纳总结，提升水平
- 让 LLM 从错误中学习（推理任务）
 - 构建 mistake-correction 数据
 - 让弱 LLM 生成错误的推理步骤
 - 让强 LLM 找出问题、解释错误原因、改正错误
 - 使用数据微调模型

Question: Tina makes \$18.00 an hour. If she works more than 8 hours per shift, she is eligible for overtime, which is paid by your hourly wage + 1/2 your hourly wage. If she works 10 hours every day for 5 days, how much money does she make?



Rationales with Mistakes:

Step 1: Tina makes \$18.00 an hour for 8 hours, which is $8 * \$18.00 = \144.00 .

Step 2: She makes \$27.00 an hour for the 2 hours of overtime, which is $2 * \$27.00 = \54.00 .

Step 3: So for 5 days, she makes \$144.00 + \$54.00 = \$198.00.

Step 4: The answer is 198.



GPT-4 As Corrector

↓ *Identifying Mistake and Making Correction*

Correction:

Incorrect Step: Step 3.

Explanation: Step 3 only calculates the earnings for one day, but not for the entire five days.

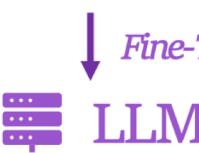
Correct Solution:

...

Step 3: For one day, she makes $\$144.00 + \$54.00 = \$198.00$.

Step 4: For 5 days, she makes $\$198.00 * 5 = \990.00 .

Step 5: The answer is 990.



↓ *Fine-Tuning on Mistake-Correction Data*

Example 1: Prompt For Generating Corrections

For the following math word problems, the original solutions may contain errors. Please identify the incorrect step in each solution, explain why it is incorrect, and provide the correct solution starting from that step.

Question: James creates a media empire. He creates a movie for \$2000. Each DVD cost \$6 to make. He sells it for 2.5 times that much. He sells 500 movies a day for 5 days a week. How much profit does he make in 20 weeks?

Original Solution:

Step 1: 500 movies a day, 5 days a week, for 20 weeks, he sells $500 * 5 * 20 = 50000$ movies.

Step 2: Each movie he sells for $\$6 * 2.5 = \15 .

Step 3: Thus, he makes a profit of $\$15 - \$6 = \$9$ per movie.

Step 4: Therefore, he makes a profit of $\$9 * 50000 = \450000 .

Step 5: The answer is 450000.

Incorrect Step: Step 4.

Explanation: The error in Step 4 is that it does not take into account the initial cost of creating the movie (\$2000). The correct solution should subtract this amount from the total profit calculated in Step 4.

Correct Solution:

...

Step 4: Therefore, he makes a profit of $\$9 * 50000 - \$2000 = \$448000$.

Step 5: The answer is 448000.

... (Another 3 examples)

Question: q_i

Original Solution: \tilde{r}_i

Incorrect Step: {Completion}

Results

Model	Training	GSM8K		MATH	
		1st / 2nd / 3rd	Avg.	1st / 2nd / 3rd	Avg.
LLaMA-2-70B (Touvron et al., 2023b)	CoT Fine-Tuning	81.4 / 81.3 / 81.1	81.3	23.6 / 23.2 / 23.2	23.2
	+ Learning From Mistakes	83.5 / 83.4 / 83.2	83.4 (+2.1)	25.0 / 25.0 / 24.6	24.9 (+1.7)
LLaMA-65B (Touvron et al., 2023a)	CoT Fine-Tuning	76.2 / 76.2 / 75.7	76.0	19.7 / 19.7 / 19.2	19.5
	+ Learning From Mistakes	77.9 / 77.3 / 77.2	77.5 (+1.5)	20.8 / 20.3 / 20.2	20.4 (+0.9)
CodeLLaMA-34B (Rozière et al., 2023)	CoT Fine-Tuning	68.8 / 68.5 / 68.2	68.5	19.1 / 19.0 / 18.9	19.0
	+ Learning From Mistakes	71.7 / 71.0 / 70.9	71.2 (+2.7)	20.4 / 20.2 / 20.0	20.2 (+1.2)
LLaMA-2-13B (Touvron et al., 2023b)	CoT Fine-Tuning	62.9 / 62.7 / 62.7	62.8	12.2 / 11.9 / 11.8	12.0
	+ Learning From Mistakes	65.7 / 65.2 / 65.0	65.3 (+2.5)	12.6 / 12.6 / 12.4	12.5 (+0.5)
LLaMA-2-7B (Touvron et al., 2023b)	CoT Fine-Tuning	52.6 / 52.5 / 52.5	52.5	8.7 / 8.5 / 8.5	8.6
	+ Learning From Mistakes	54.1 / 53.7 / 53.6	53.8 (+1.3)	9.4 / 8.9 / 8.8	9.0 (+0.4)

结论

- 专业领域模型同样可以受益
- 模型规模越大，越能从错误中学习（提升明显）
- 生成 corrections 只能用强模型进行
- GPT-4 难以修正自己犯的错
- 最强的 GPT-4 也难以解决困难复杂的推理

