Mind the Gap: Bridging Thought Leap for Improved Chain-of-Thought Tuning

Haolei Xu[1*] **Yuchen Yan**[1*] **Yongliang Shen**[1†] **Wenqi Zhang**[1] **Guiyang Hou**[1]
Shengpei Jiang[2] **Kaitao Song**[3] **Weiming Lu**[1†] **Jun Xiao**[1] **Yueting Zhuang**[1]

[1] Zhejiang University [2] SF Technology
[3] Microsoft Research Asia

{xuhaolei,syl,luwm}@zju.edu.cn

ZHEJIANG UNIVERSITY

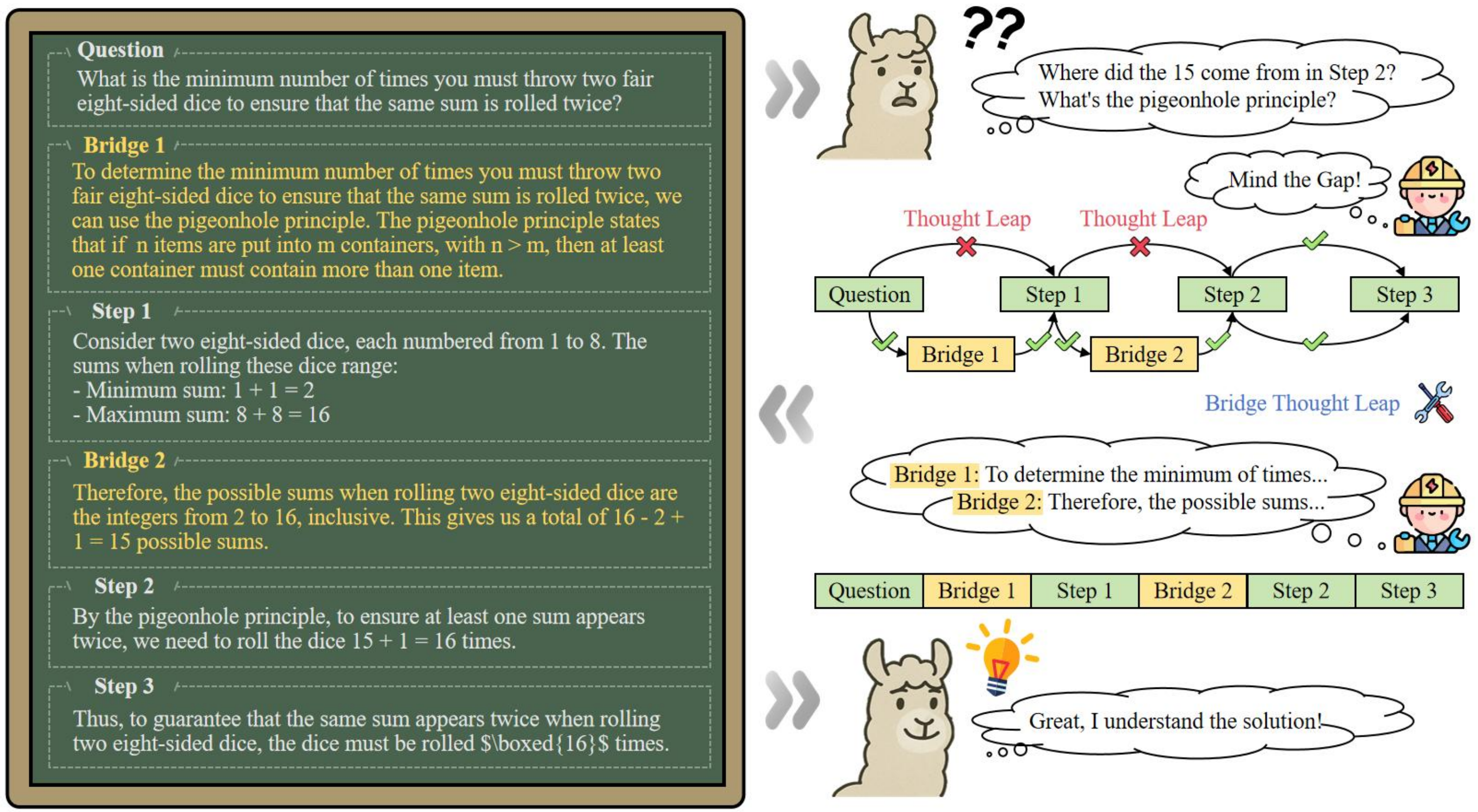
NEURAL INFORMATION PROCESSING SYSTEMS

Background

- Chain-of-Thought (CoT) is a key paradigm for improving complex reasoning, especially in structured tasks like math and logic.
- Yet even refined CoT data may skip essential reasoning steps that are clear to humans but difficult for models.

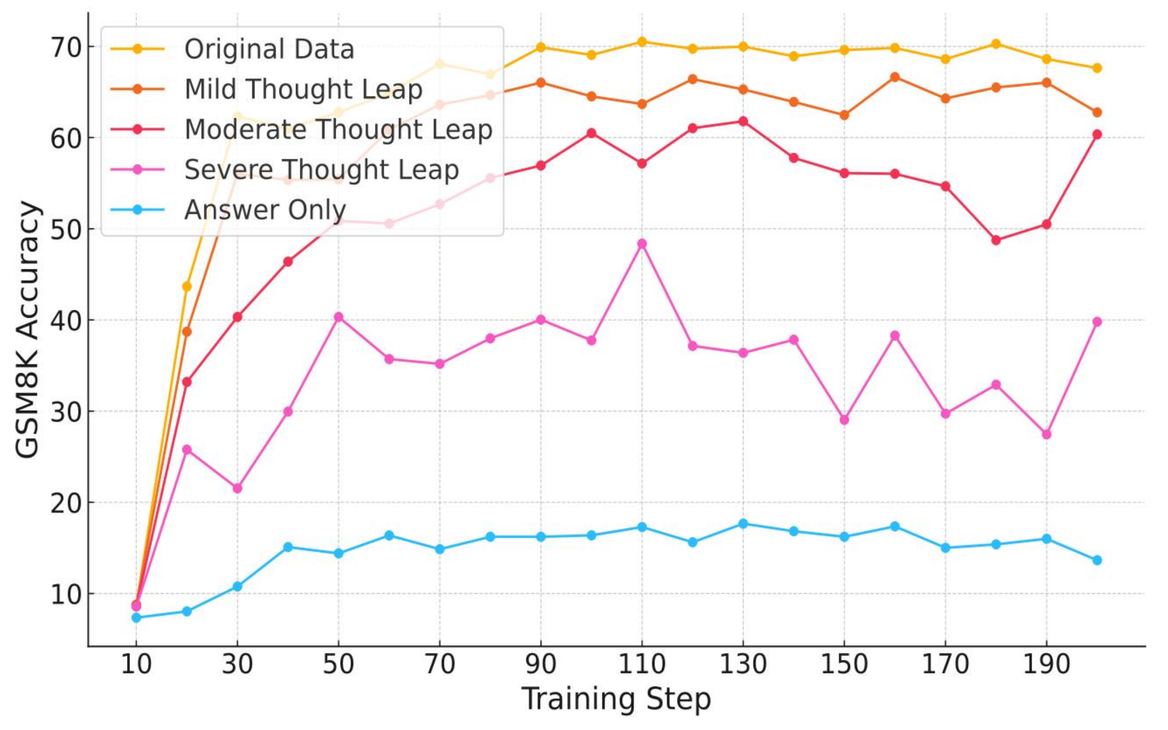
What is the pigeonhole principle?

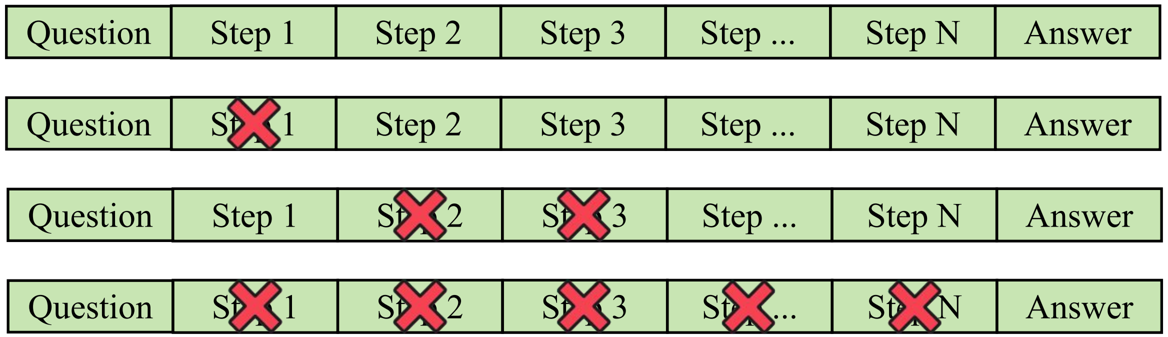
Where does the number 15 come from?



Background

- What happens when a model is trained on datasets with thought leaps?
- Based on MetaMathQA, construct datasets with varying degrees of thought leaps.





Training impact: Severe reasoning leaps can cause up to 27.83% performance loss

Learning efficiency: Slower and less stable convergence.

Formulation

V : completeness function

Ideal complete CoT $C^* = (Q, s_1^*, s_2^*, \ldots, s_m^*)$ $\quad \forall i \in [1, m-1] \quad V(s_i^*, s_{i+1}^*) = \text{True}$

Thought Leap CoT $C = (s_0, s_1, s_2, \ldots, s_n)$ $\quad \exists k \in [0, n-1] \quad V(s_k, s_{k+1}) = \text{False}$

CoT Thought Leap Bridge task

- Leap Identification (s_k, s_{k+1})
- Content Bridge $(s'_{k.1}, s'_{k.2}, \ldots, s'_{k.j})$ ➡

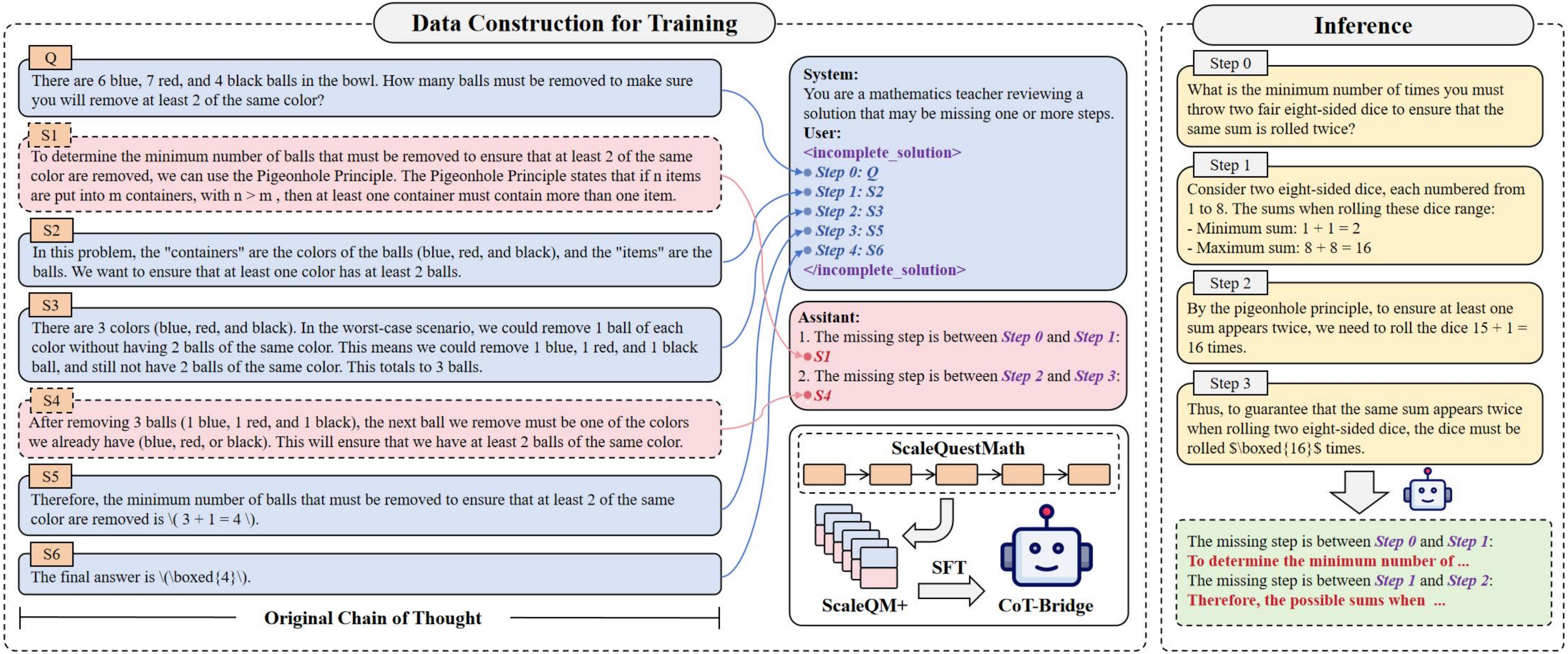
$$V(s_k, s'_{k.1}) = \text{True}$$
$$\forall i \in [1, j-1], V(s'_{k.i}, s'_{k.i+1}) = \text{True}$$
$$V(s'_{k.j}, s_{k+1}) = \text{True}$$

Challenges:

- Designing a completeness evaluation function is difficult.
- Ideal completeness is a theoretical concept — how can it be handled in practice?

CoT-Bridge

We chose **ScaleQuestMath** as an approximation of ideal CoT, as its reasoning structure is relatively complete and well-formed, making it suitable for systematic step removal.



Main Experiment

Dataset	Size	Method	Basic Level			Competition Level			Average
			GSM8K	MATH	GaoKao	Odyssey	Olympiad	AMC23	
Meta-Llama3.1-8B									
/	/	4-shot	54.15	18.30	20.58	16.54	4.85	11.25	20.95
GSM8K+MATH	15k	Direct SFT	65.09	19.25	21.69	18.48	5.07	12.50	23.68
MathInstruct	262k	Direct SFT	68.16	23.60	25.52	25.06	5.89	7.50	25.96
MetaMathQA	395k	Direct SFT	78.90	36.10	32.86	24.68	8.48	17.50	33.09
		QwenBridger-S	81.10$^{+2.20}$	34.85$^{-1.25}$	30.52$^{-2.34}$	22.67$^{-2.01}$	9.26$^{+0.78}$	7.50$^{-10.00}$	30.98$^{-2.11}$
		QwenBridger-L	80.80$^{+1.90}$	38.05$^{+1.95}$	31.43$^{-1.43}$	24.48$^{-0.20}$	9.37$^{+0.89}$	2.50$^{-15.00}$	31.11$^{-1.98}$
		CoT-Bridge-R	80.46$^{+1.56}$	38.05$^{+1.95}$	**33.57**$^{+0.71}$	24.42$^{-0.26}$	9.37$^{+0.89}$	12.50$^{-5.00}$	33.06$^{-0.03}$
		CoT-Bridge	**81.14**$^{+2.24}$	**38.15**$^{+2.05}$	33.12$^{+0.26}$	**25.97**$^{+1.29}$	**9.48**$^{+1.00}$	**18.75**$^{+1.25}$	**34.44**$^{+1.35}$
NuminaMath	859k	Direct SFT	84.86	51.45	49.03	36.56	21.30	20.00	43.87
		QwenBridger-S	84.23$^{-0.63}$	52.40$^{+0.95}$	51.95$^{+2.92}$	39.73$^{+3.17}$	24.70$^{+3.40}$	27.50$^{+7.50}$	46.75$^{+2.88}$
		QwenBridger-L	85.25$^{+0.39}$	54.20$^{+2.75}$	51.62$^{+2.59}$	39.08$^{+2.52}$	25.33$^{+4.03}$	35.00$^{+15.00}$	48.41$^{+4.54}$
		CoT-Bridge-R	84.82$^{-0.04}$	54.20$^{+2.75}$	51.88$^{+2.85}$	40.12$^{+3.56}$	**26.15**$^{+4.85}$	33.75$^{+13.75}$	48.50$^{+4.63}$
		CoT-Bridge	**85.97**$^{+1.11}$	**56.80**$^{+5.35}$	**54.42**$^{+5.39}$	**40.76**$^{+4.20}$	24.85$^{+3.55}$	**35.63**$^{+15.63}$	**49.74**$^{+5.87}$
Qwen2.5-Math-1.5B									
/	/	4-shot	79.00	48.05	45.52	38.18	19.07	22.50	42.05
GSM8K+MATH	15k	Direct SFT	74.45	51.40	47.66	38.50	17.44	27.50	42.83
MathInstruct	262k	Direct SFT	70.96	48.90	46.49	40.89	16.22	20.00	40.58
MetaMathQA	395k	Direct SFT	81.01	49.60	46.62	38.63	18.19	21.25	42.55
		QwenBridger-S	**81.58**$^{+0.57}$	51.30$^{+1.70}$	46.69$^{+0.07}$	38.63$^{-0.00}$	18.52$^{+0.33}$	25.63$^{+4.38}$	44.04$^{+1.49}$
		QwenBridger-L	81.01$^{-0.00}$	53.63$^{+4.03}$	49.22$^{+2.60}$	38.70$^{+0.07}$	18.85$^{+0.66}$	27.50$^{+6.25}$	44.82$^{+2.27}$
		CoT-Bridge-R	**81.58**$^{+0.57}$	53.65$^{+4.05}$	48.12$^{+1.50}$	39.02$^{+0.39}$	18.22$^{+0.03}$	25.63$^{+4.38}$	44.37$^{+1.82}$
		CoT-Bridge	81.39$^{+0.38}$	**56.60**$^{+7.00}$	**49.61**$^{+2.99}$	**39.66**$^{+1.03}$	**19.44**$^{+1.25}$	**28.75**$^{+7.50}$	**45.91**$^{+3.36}$
NuminaMath	859k	Direct SFT	83.62	63.90	57.40	46.77	33.04	32.5	52.87
		QwenBridger-S	83.23$^{-0.39}$	64.20$^{+0.30}$	58.25$^{+0.85}$	46.96$^{+0.19}$	32.04$^{-1.00}$	35.00$^{+2.50}$	53.28$^{+0.41}$
		QwenBridger-L	82.81$^{-0.81}$	66.25$^{+2.35}$	57.79$^{+0.39}$	46.64$^{-0.13}$	32.70$^{-0.34}$	40.00$^{+7.50}$	54.37$^{+1.50}$
		CoT-Bridge-R	83.06$^{-0.56}$	65.20$^{+1.30}$	55.84$^{-1.56}$	43.09$^{-3.68}$	33.89$^{+0.85}$	40.00$^{+7.50}$	53.51$^{+0.64}$
		CoT-Bridge	**84.61**$^{+0.99}$	**68.05**$^{+4.15}$	**59.29**$^{+1.89}$	**47.16**$^{+0.39}$	**34.11**$^{+1.07}$	**45.00**$^{+12.50}$	**56.26**$^{+3.39}$

we performed supervised fine tuning (SFT) experiments using MetaMathQA, NuminaMath-CoT datasets and their bridged versions on representative base models.

Performance Drop

Consistent Improvement

- Bridging Thought Leaps using CoT-Bridge consistently improves reasoning performance.
- Accurate leap identification is crucial for effective bridging.
- Zero-shot bridging shows promise but lacks consistency.

Analysis 1

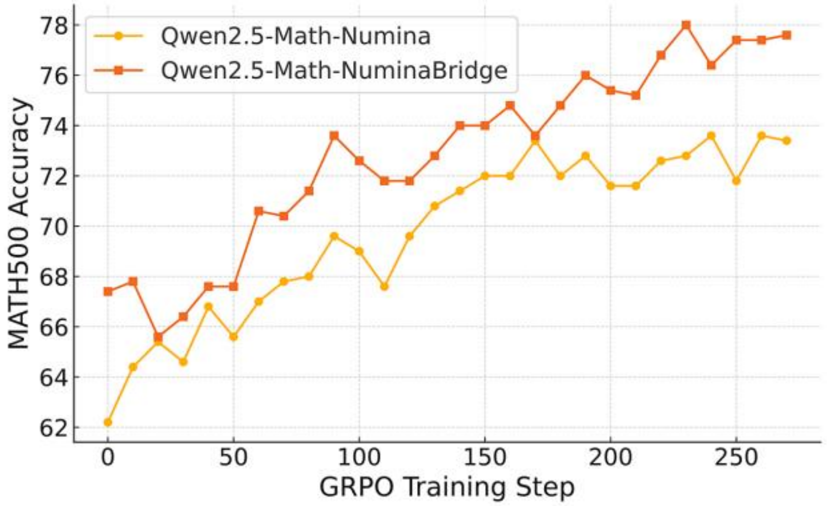
CoT-Bridge can serve as a plug-and-play enhancement module that seamlessly integrates into existing training pipelines while delivering consistent performance improvements.

Dataset	Method	GSM8K	MATH	GaoKao	Odyssey	Olympiad	AMC23	Average
Distill	Direct SFT	81.86	68.15	60.84	48.13	33.00	39.37	55.23
	CoT-Bridge	**82.52**$^{+0.66}$	**71.50**$^{+3.35}$	**66.43**$^{+5.59}$	**49.16**$^{+1.03}$	**34.89**$^{+1.89}$	**45.00**$^{+5.63}$	**58.25**$^{+3.02}$
Reject Sampling	Direct SFT	83.36	74.90	64.94	51.81	37.63	50.00	60.44
	CoT-Bridge	**83.74**$^{+0.38}$	**75.25**$^{+0.35}$	**67.47**$^{+2.53}$	**51.87**$^{+0.06}$	**39.41**$^{+1.78}$	**53.13**$^{+3.13}$	**61.81**$^{+1.37}$

Qwen2.5-Instruct-72B distilled and rejection sampled data

- CoT-Bridge can improve the quality of generated data in knowledge distillation and rejection sampling

Model	Method	GSM8K	MATH	GaoKao	Odyssey	Olympiad	AMC23	Average
Qwen2.5-Math-Instruct-1.5B	/	**84.80**	75.80	65.50	**54.52**	38.10	**60.00**	63.12
Oat-Zero-1.5B	Dr. GRPO	83.62	74.20	**69.61**	52.71	37.60	53.00	61.79
Qwen2.5-Math-1.5B	GRPO	82.71	74.60	64.94	49.10	35.85	50.00	59.33
Qwen2.5-Math-Numina	GRPO	84.31	74.80	62.34	51.94	39.41	52.50	60.88
Qwen2.5-Math-NuminaBridge	GRPO	84.08$^{-0.23}$	**78.20**$^{+3.40}$	67.01$^{+4.67}$	54.26$^{+2.32}$	**40.30**$^{+0.89}$	**60.00**$^{+7.50}$	**63.98**$^{+3.10}$



- Using bridge-augmented datasets for SFT leads to a higher performance ceiling in subsequent RL.

Analysis 2

To evaluate the capability of various methods in CoT Thought Leap Bridge Task, we constructed a standardized evaluation framework on ScaleQM+ test set, covering leap identification and generation quality.

Method	Similarity	Position			Overall↑
		Pre↑	Rec↑	Red↓	
Qwen2.5-Instruct-7B	/	14.15	12.04	34.13	10.54
Qwen2.5-Instruct-72B	/	33.99	33.64	33.73	31.12
CoT-Bridge	/	**78.02**	78.37	**1.61**	**76.15**
Full-position	1	20.96	**79.64**	79.04	75.72
	0.95	23.42	75.07	76.57	71.22
	0.90	24.81	59.21	74.37	55.9
	0.85	19.34	30.17	63.87	28.26
	0.80	7.47	8.65	32.51	7.84

- CoT has the best leap localization accuracy and generation quality.
- Other baselines suffer from high redundancy rates.
- Redundant steps can easily cause the model to learn repetitive patterns.

Analysis 3

- OOD Experiments in Logical Reasoning Tasks (Meta-Llama3.1-8B: ↑2.99%, Qwen2.5-Math-1.5B: ↑0.99%)
- CoT-Bridge help models master general reasoning structures and enhance generalization capabilities.

Bridge Method	Mertic	FOLIO	LogicQA	PW	ReClor	RuleTaker	Average
NuminaMath+Meta-Llama3.1-8B							
No Bridge	Accuracy ↑	68.15	34.33	59.09	47	54.5	52.61
	Invalid ↓	1.48	6.3	0.51	2	0.31	2.12
GapBridge	Accuracy ↑	**74.07**$^{+5.92}$	**35.64**$^{+1.31}$	**61.52**$^{+2.43}$	**50.20**$^{+3.20}$	**56.57**$^{+2.07}$	**55.60**$^{+2.99}$
	Invalid ↓	0.74	4.92	0.2	2.2	0.1	1.63
NuminaMath+Qwen2.5-Math-1.5B							
No Bridge	Accuracy ↑	**74.07**	29.72	55.84	37.6	53.05	50.06
	Invalid ↓	1.48	4.53	0.3	1.2	0.72	1.65
GapBridge	Accuracy ↑	71.11$^{-2.96}$	**33.10**$^{+3.38}$	**58.38**$^{+2.54}$	**39.00**$^{+1.40}$	**53.67**$^{+0.62}$	**51.05**$^{+0.99}$
	Invalid ↓	1.48	3.46	0.3	1.4	0	1.33

Analysis 4

- **Ablation Study:** Remove one bridging type (begin/middle/end) → Fine-tune Qwen2.5-Math-1.5B with same settings.
- All three types of completion have a positive effect on model performance

Delete Pos	Num	Ratio	GSM8K	MATH	GaoKao	Odyssey	Olympiad	AMC23	Average
Qwen2.5-Math-1.5B+MetaMath									
/	/	/	81.39	**56.60**	49.61	39.66	**19.44**	**28.75**	**45.91**
- begin	225873	31.84	**82.71**$^{+1.32}$	55.85$^{-0.75}$	49.55$^{-0.06}$	**41.09**$^{+1.43}$	19.22$^{-0.22}$	22.50$^{-6.25}$	45.15$^{-0.76}$
- middle	470541	66.34	80.97$^{-0.42}$	52.45$^{-4.15}$	49.35$^{-0.26}$	37.14$^{-2.52}$	17.26$^{-2.18}$	27.50$^{-1.25}$	44.11$^{-1.80}$
- end	12908	1.82	81.14$^{-0.25}$	55.75$^{-0.85}$	**49.94**$^{+0.33}$	37.02$^{-2.64}$	18.78$^{-0.66}$	28.12$^{-0.63}$	45.13$^{-0.78}$
Qwen2.5-Math-1.5B+NuminaMath									
/	/	/	**84.61**	**68.05**	**59.29**	**47.16**	**33.44**	**45.00**	**56.26**
- begin	694887	42.57	82.60$^{-2.01}$	64.85$^{-3.20}$	58.67$^{-0.62}$	45.48$^{-1.68}$	32.96$^{-0.48}$	38.75$^{-6.25}$	53.89$^{-2.37}$
- middle	906168	55.52	83.51$^{-1.10}$	63.55$^{-4.50}$	56.43$^{-2.86}$	47.16$^{-0.00}$	28.30$^{-5.14}$	38.13$^{-6.87}$	52.85$^{-3.41}$
- end	31115	1.91	82.79$^{-1.82}$	64.05$^{-4.00}$	55.97$^{-3.32}$	46.51$^{-0.65}$	27.48$^{-5.96}$	42.50$^{-2.50}$	53.22$^{-3.04}$

Analysis 5

- Content Analysis of Bridge: Scoring the content of Bridge using Qwen2.5-Math-PRM-7B.
- CoT-Bridge achieves much higher content quality for Bridge compared to the baseline.