Decoupling the Effect of Chain-of-Thought Reasoning: A Human Label Variation Perspective

Beiduo Chen, Tiancheng Hu, Caiqi Zhang, Robert Litschko,
Anna Korhonen, Barbara Plank

MaiNLP lab, LMU Munich
LTL, University of Cambridge

ACL 2026: The 64th Annual Meeting of the Association for Computational Linguistics
July 2-7, 2026

Motivation

- **Human Label Variation (HLV)** captures genuine annotator disagreement, requiring models to predict answer distributions, not only a single label.
- **Long Chain-of-Thought (CoT)** improves many single-answer reasoning benchmarks, but it may collapse ambiguity into one confident choice.
- **Question:** Does CoT help LLMs approximate human label distributions, and if so, what part of the prediction does it actually control?

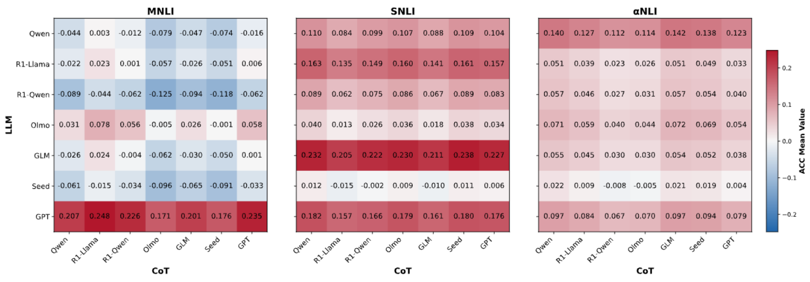
Experimental Setup

Data	ChaosNLI: MNLI, SNLI, and αNLI, each with human judgment distributions from 100 annotators.
Models	Qwen, R1-Llama, R1-Qwen, Olmo, GLM, Seed, GPT: open-source reasoning-tuned LLMs.
Metrics	Accuracy$\uparrow$ for top-1 correctness; JSD$\downarrow$ for distributional alignment; Spearman's $\rho\uparrow$ for ranking alignment.
Probes	Start-vs-last reasoning, Cross-CoT transfer, and step-wise early stopping over 10 CoT segments.

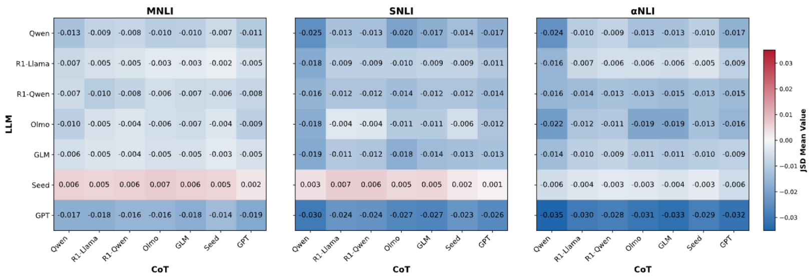
Reasoning LLMs	**Abbr.**
Qwen/Qwen3-30B-A3B-Thinking-2507 (Team, 2025b)	Qwen
deepseek-ai/DeepSeek-R1-Distill-Llama-70B (DeepSeek-AI et al., 2025)	R1-Llama
deepseek-ai/DeepSeek-R1-Distill-Qwen-32B (DeepSeek-AI et al., 2025)	R1-Qwen
allenai/Olmo-3-32B-Think (Olmo et al., 2025)	Olmo
zai-org/GLM-Z1-32B-0414 (GLM et al., 2024)	GLM
ByteDance-Seed/Seed-OSS-36B-Instruct (Team, 2025a)	Seed
openai/gpt-oss-20b (OpenAI, 2025)	GPT

Does CoT Improve HLV Performance?

Task	MNLI				SNLI				αNLI			
LLMs/Metrics	$ACC_{start}\uparrow$	$ACC_{last}\uparrow$	$JSD_{start}\downarrow$	$JSD_{last}\downarrow$	$ACC_{start}\uparrow$	$ACC_{last}\uparrow$	$JSD_{start}\downarrow$	$JSD_{last}\downarrow$	$ACC_{start}\uparrow$	$ACC_{last}\uparrow$	$JSD_{start}\downarrow$	$JSD_{last}\downarrow$
Qwen	0,688	0,644	0,093	0,080	0,668	0,778	0,144	0,119	0,749	0,890	0,108	0,084
R1-Llama	0,666	0,689	0,082	0,077	0,615	0,750	0,133	0,123	0,839	0,878	0,098	0,091
R1-Qwen	0,734	0,672	0,080	0,072	0,689	0,764	0,127	0,115	0,832	0,860	0,094	0,081
Olmo	0,614	0,609	0,088	0,082	0,738	0,775	0,133	0,122	0,819	0,863	0,107	0,087
GLM	0,670	0,640	0,082	0,077	0,545	0,756	0,134	0,120	0,834	0,888	0,099	0,088
Seed	0,705	0,614	0,077	0,083	0,766	0,777	0,124	0,127	0,868	0,887	0,098	0,095
GPT	0,437	0,672	0,095	0,077	0,596	0,772	0,145	0,119	0,793	0,872	0,112	0,080



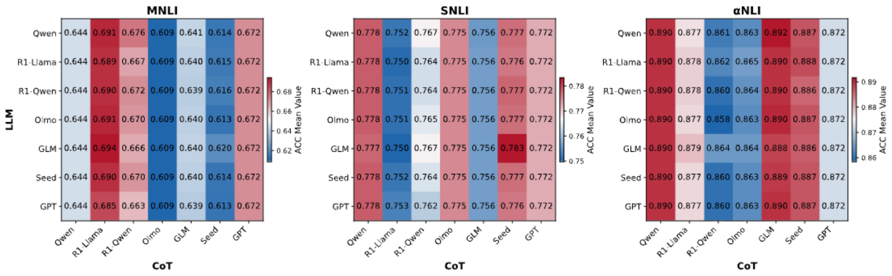
(a) Delta Accuracy ↑



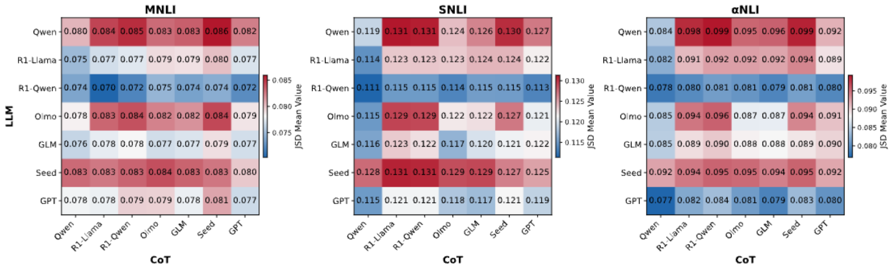
(b) Delta JSD ↓

- Injecting CoT almost universally reduces divergence from human judgment distributions, suggesting that CoT text carries portable HLV-relevant information, but accuracy and JSD respond differently.

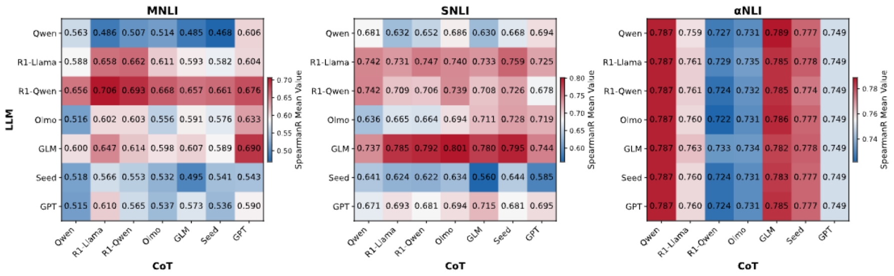
Cross-CoT Reveals a Split Influence



(a) Last Accuracy ↑



(b) Last JSD ↓



(c) Last Spearman's ρ ↑

MNLI final-step variance	LLM prior	CoT	Residual
Accuracy	0.1%	99.5%	0.4%
JSD	83.3%	8.7%	8.1%
Spearman's ρ	71.1%	13.1%	15.7%

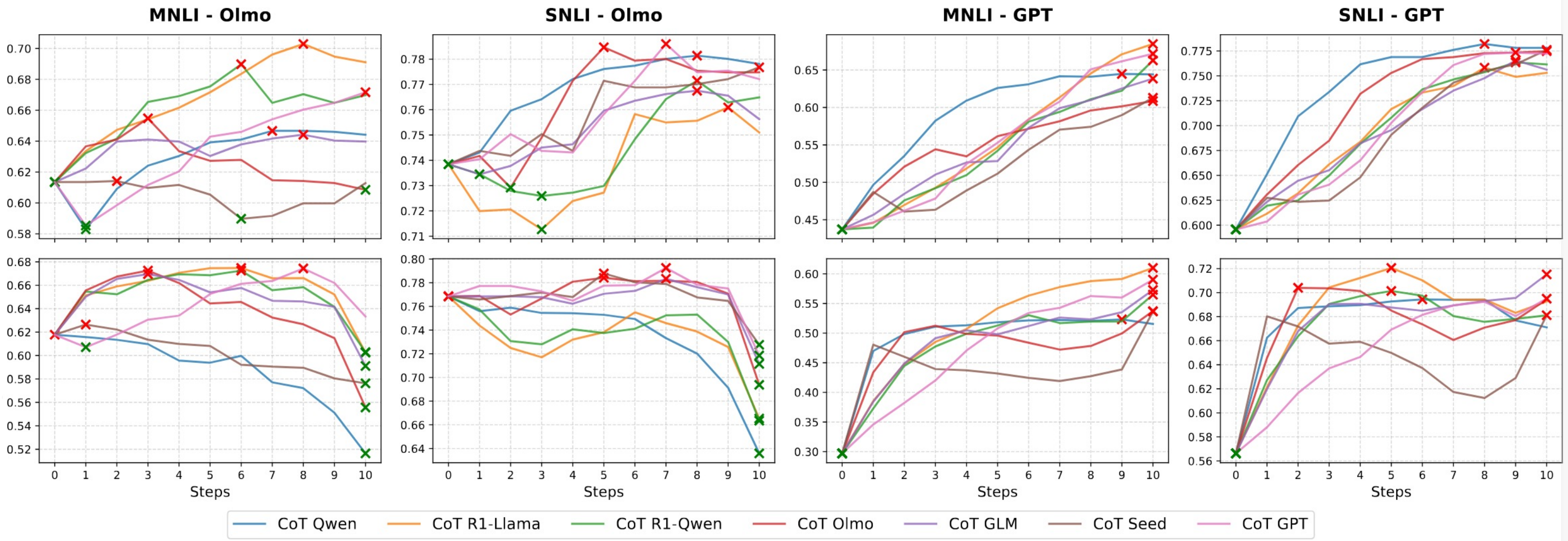
- Models follow CoT for the final argmax choice, but fall back to latent parametric preferences when assigning probability mass to alternative options.

When Does CoT Take Control?

Task	MNLI								
Metric	ACC			JSD			Spearman's ρ		
Step/Factor	LLM	CoT	Residual	LLM	CoT	Residual	LLM	CoT	Residual
Step 0	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%
Step 1	97.3%	0.2%	2.5%	96.9%	0.4%	2.7%	92.7%	0.4%	6.9%
Step 2	95.7%	0.5%	3.7%	95.3%	0.7%	4.0%	92.3%	1.8%	5.9%
Step 3	91.8%	1.6%	6.6%	93.4%	1.6%	5.0%	91.6%	2.6%	5.8%
Step 4	89.1%	2.2%	8.7%	91.5%	2.9%	5.6%	90.5%	3.6%	5.9%
Step 5	85.2%	4.3%	10.6%	89.3%	4.2%	6.4%	88.0%	4.9%	7.1%
Step 6	79.4%	8.5%	12.0%	87.9%	5.7%	6.5%	83.6%	7.6%	8.8%
Step 7	66.5%	18.4%	15.1%	87.5%	6.1%	6.4%	77.5%	10.1%	12.4%
Step 8	44.1%	38.3%	17.5%	86.4%	7.0%	6.7%	73.3%	14.5%	12.2%
Step 9	22.4%	65.2%	12.4%	84.3%	8.5%	7.2%	73.0%	16.8%	10.3%
Step 10	0.1%	99.5%	0.4%	83.3%	8.7%	8.1%	71.1%	13.1%	15.7%

Task	SNLI								
Metric	ACC			JSD			Spearman's ρ		
Step/Factor	LLM	CoT	Residual	LLM	CoT	Residual	LLM	CoT	Residual
Step 0	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%
Step 1	96.5%	1.4%	2.2%	94.4%	0.9%	4.7%	83.2%	2.4%	14.4%
Step 2	85.1%	8.6%	6.4%	88.2%	4.4%	7.4%	76.1%	8.0%	15.8%
Step 3	75.1%	16.7%	8.1%	82.2%	8.3%	9.5%	73.2%	12.2%	14.6%
Step 4	62.5%	29.1%	8.5%	75.4%	14.0%	10.6%	69.9%	16.4%	13.7%
Step 5	57.4%	32.4%	10.2%	72.8%	16.6%	10.6%	70.8%	13.0%	16.2%
Step 6	58.9%	29.7%	11.3%	72.5%	17.4%	10.1%	71.9%	10.4%	17.7%
Step 7	53.4%	30.6%	16.0%	73.1%	17.2%	9.7%	73.9%	6.6%	19.5%
Step 8	43.0%	33.2%	23.8%	72.1%	19.1%	8.8%	71.8%	7.9%	20.3%
Step 9	34.8%	43.9%	21.2%	71.3%	19.8%	8.9%	77.8%	6.3%	15.9%
Step 10	0.2%	98.6%	1.2%	67.7%	22.2%	10.1%	81.7%	3.3%	15.0%

Task	αNLI								
Metric	ACC			JSD			Spearman's ρ		
Step/Factor	LLM	CoT	Residual	LLM	CoT	Residual	LLM	CoT	Residual
Step 0	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%
Step 1	93.9%	1.4%	4.7%	95.4%	1.3%	3.3%	94.0%	1.3%	4.8%
Step 2	75.7%	10.5%	13.8%	89.3%	3.9%	6.8%	76.3%	10.4%	13.2%
Step 3	64.1%	22.2%	13.7%	81.4%	8.8%	9.8%	64.3%	22.4%	13.3%
Step 4	52.1%	31.2%	16.8%	80.8%	9.5%	9.6%	51.9%	32.6%	15.5%
Step 5	45.2%	41.7%	13.1%	81.6%	10.4%	8.0%	47.2%	41.3%	11.5%
Step 6	35.4%	53.8%	10.8%	84.9%	9.4%	5.8%	36.4%	53.3%	10.3%
Step 7	22.5%	67.3%	10.2%	87.0%	8.2%	4.9%	22.1%	67.5%	10.4%
Step 8	9.6%	76.2%	14.2%	85.2%	10.1%	4.7%	10.1%	76.9%	13.0%
Step 9	5.3%	83.5%	11.2%	81.4%	13.2%	5.4%	5.4%	84.1%	10.5%
Step 10	0.1%	99.4%	0.5%	76.9%	15.2%	7.9%	0.1%	99.4%	0.5%



- CoT influence on accuracy spikes at the final reasoning step, while its influence on distributional similarity remains low throughout the chain.

Conclusion & Contributions

- **Long CoT improves overall distributional alignment,** suggesting that reasoning text does contain useful information for HLV.
- **But CoT mainly controls the top answer.** The non-argmax probability landscape remains anchored to model priors.
- **Next step: distribution-aware reasoning.** Reasoning systems should explicitly maintain and calibrate uncertainty instead of only converging to a decisive label.

Thank You !

Beiduo Chen

Contact: beiduochen@cis.lmu.de

Code: https://github.com/mainlp/CoT-HLV

Code