Compositional Generalization from Learned Skills via CoT Training: A Theoretical and Structural Analysis for Reasoning

Xinhao Yao[1] · Ruifeng Ren[1] · Yun Liao[2] · Lizhong Ding[3] · Yong Liu[1]†

GSAI, Renmin University of China[1], TUST[2], BIT[3]

Xinhao Yao

[1] Yao et al. Compositional Generalization from Learned Skills via CoT Training: A Theoretical and Structural Analysis for Reasoning. ICLR 2026.

1. Introduction & Motivation

Despite the prevalence of Chain-of-Thought (CoT) training paradigms, fundamental questions remain unresolved:

- **(Q1) Does training with Chain of Thought (CoT) enhance reasoning generalization in both in-distribution (ID) and out-of-distribution (OOD) settings—and if so, what theoretical principles underlie this capacity?**
- **(Q2) How is such generalization capability instantiated within the model's internal representations?**

Main Insights:

we propose that compositional generalization is fundamental: models systematically combine simpler learned skills during CoT training to address novel and more complex problems.

2. Generalization Error Analysis

Definition1 (Data Distribution). Let X and Y denote the input and output random variables, respectively.

$P(Y \mid X) = \sum_{C} P(Y \mid X, C) P(C \mid X) = \sum_{C} P(Y \mid X, C) \prod_{k=1}^{K} P(C_k \mid C_{<k}, X)$, where

$P(C_k \mid C_{<k}, X)$ is the probability of the k-th reasoning step given all previous steps and the input.

Definition2. The expected generalization error: $\widetilde{\text{error}}$**= population risk - empirical risk.**

Assumption 1 (Mixed Test Distribution). Assume $P_{\text{test}}(Y|X)$ is a mixture of two distributions:

$P_{\text{test}}(Y|X) = (1-\alpha) P_{\text{test}}^{\text{ID}}(Y|X) + \alpha P_{\text{test}}^{\text{OOD}}(Y|X)$, mixing coefficient $\alpha \in [0, 1]$.

Theorem 1 (**Generalization Bounds via Distributional Divergence**). *The generalization error naturally decomposes into ID and OOD components:*

$$\widetilde{\text{error}} \leq \sqrt{\frac{2R^2}{N}\left[(1-\alpha)D_{\text{KL}}\left(P_{\text{test}}^{\text{ID}}(Y \mid X) \,||\, P_{\text{train}}(Y \mid X)\right) + \alpha D_{\text{KL}}\left(P_{\text{test}}^{\text{OOD}}(Y \mid X) \,||\, P_{\text{train}}(Y \mid X)\right)\right]}.$$

Theorem 2 (**OOD Generalization Error for Training with CoT**). A*ssume sufficient training such that* $D_{\text{KL}}(P_{\text{test}}^{\text{ID}} | P_{\text{train}}) \to 0$**,** where C denotes the CoT intermediate reasoning steps*:*

$$\widetilde{\text{error}}^2 \leq \frac{2R^2\alpha}{N}\left[D_{\text{KL}}\left(P_{\text{test}}^{OOD}(C \mid X) \,||\, P_{\text{train}}(C \mid X)\right) + \mathbb{E}_{C \sim P_{\text{test}}^{OOD}(C|X)}\left[D_{\text{KL}}\left(P_{\text{test}}^{OOD}(Y \mid X, C) \,||\, P_{\text{train}}(Y \mid X, C)\right)\right]\right].$$

Remark

- **For ID:** The compositional patterns C in the test problems exactly match those encountered during training. Thus, under sufficient training—with or without CoT—the ID generalization error approaches zero.
- **For OOD:** Since OOD test problems involve unseen compositional patterns, training without CoT struggles to generalize to OOD settings. In contrast, with CoT training, both P(C|X) and P(Y |X, C) are explicitly modeled (that is, the learned, simpler skills), aligning precisely with the stages of the compositional circuit. **CoT training** teaches models how to think through **compositional reasoning—**not merely what to think **(matching the answers).**

3. Internal Circuits of CoT vs. Non-CoT Training

Training data & ID/OOD test.

ID vs. OOD: Is the relation composition (r_i, r_j) consistent?

S_{ID}:
(Paris, CapitalOf, France), (France, LocatedIn, Europe),
(Berlin, CapitalOf, Germany), (Germany, LocatedIn, Europe).

S_{OOD}:
(Lima, CapitalOf, Peru), (Peru, HasNaturalFeature, Andes Mountains).

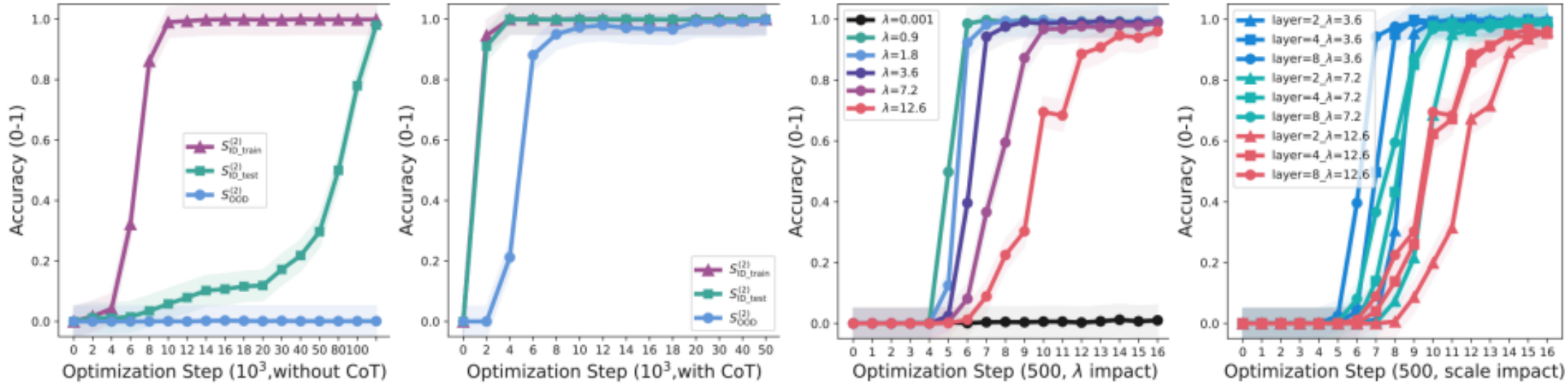
$S^{(2)}_{\text{ID}_{\text{train}}}$, a uniformly random subset of the inferred facts derived from S_{ID}:
(Paris, CapitalOfCountryLocatedIn, Europe).

$S^{(2)}_{\text{ID}_{\text{test}}}$, previously unseen inferred facts derived from S_{ID} (ID generalization):
(Berlin, CapitalOfCountryLocatedIn, Europe).

$S^{(2)}_{\text{OOD}_{\text{test}}}$, previously unseen inferred facts derived from S_{OOD} (OOD generalization):
(Lima, CapitalOfCountryWithNaturalFeature, Andes Mountains).

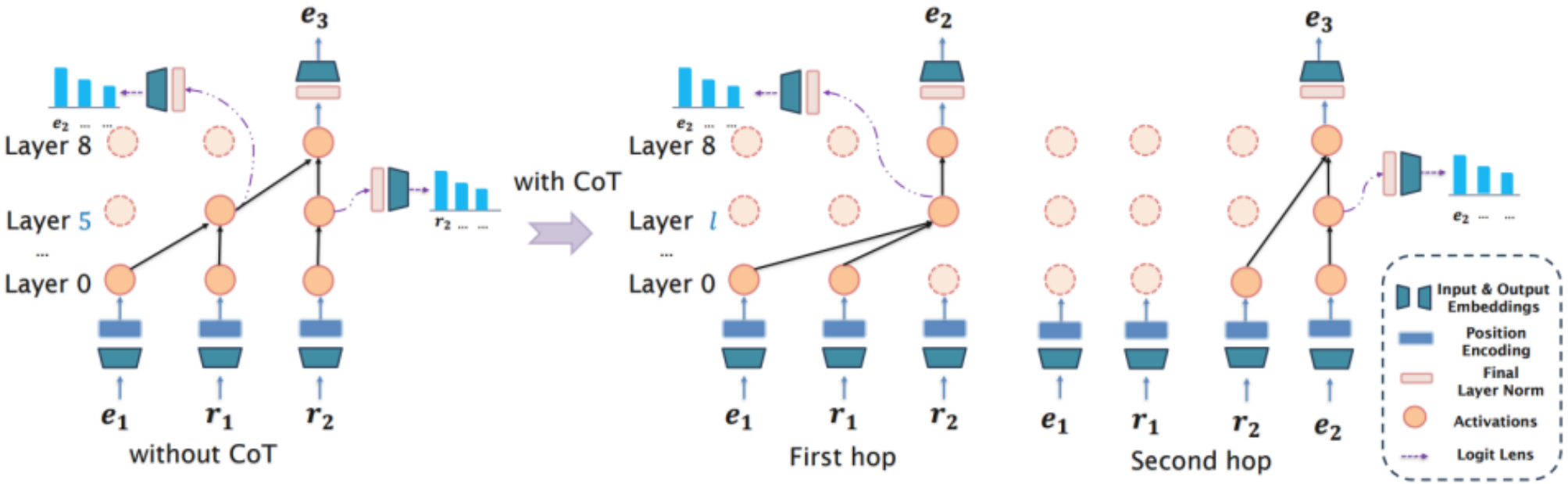
ID test: unseen instances within known compositions. The model may memorize relational compositions or reasoning patterns.

OOD test: Test the model's ability to generalize to unseen compositional patterns. This indicates that the model has learned the underlying reasoning patterns.



Remark

- **Insight 1:** Compared to training without CoT, incorporating CoT annotations significantly **accelerates convergence and improves both ID and OOD generalization**.
- **Insight 2:** Indeed, the model may undergo a process of **memorization before gradually learning the underlying patterns, eventually evolving toward more structured** and cross-instance transferable representations.



We perform a set of logit lens and causal tracing experiments after training with CoT：interpret intermediate hidden states by projecting them into the output vocabulary space. **target state = ? top-1 token**

4. Structural Discussion

Training with/without CoT:

- For two-hop facts $(e_1, r_1, e_2) \oplus (e_2, r_2, e_3) \rightarrow (e_1, r_1, r_2, e_3)$.
- Training with CoT: the bridge entity (e_2) can be extracted from the middle layer hidden states $E(\text{layer index}, r_2)$, for ID $\text{layer index} = 3$, for OOD, $\text{layer index} = 5$ (OOD is more challenging than ID).
- Training without CoT: The generalization circuit **fails to emerge** in OOD settings. In ID settings, $\text{layer index} \geq 5$, **higher than** that in Training with CoT.

Transformer limits in OOD:

- **Intuitively, smaller layer index imply that more layers remain available for processing the second hop, potentially leading to better performance.** Only when the intermediate result e_2 is resolved can the final result e_3 be correctly derived, traditional transformer cannot process subtasks in parallel, which is the limitations of traditional transformer architectures.
- The generalization circuit we discovered essentially corresponds to the **reuse of the model's weights** (effective depth increases).

Thank you!

ICLR
International Conference On Learning Representations

中国人民大學
RENMIN UNIVERSITY OF CHINA

高瓴人工智能学院
Gaoling School of Artificial Intelligence