NTT

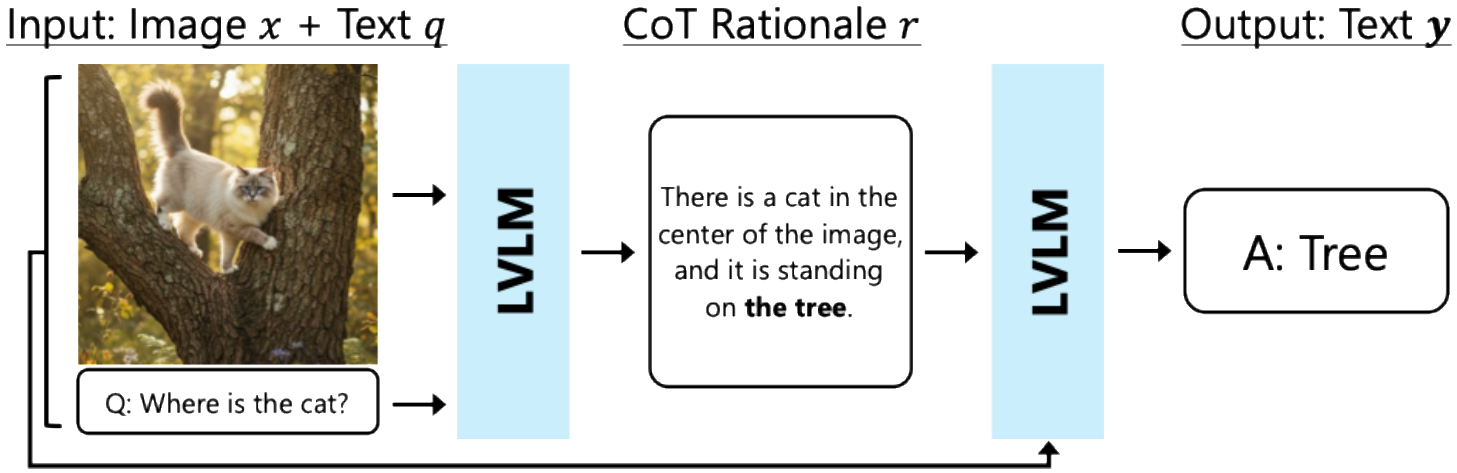
CVPR JUNE 3-7, 2026 DENVER COLORADO

Rationale-Enhanced Decoding for Multi-modal Chain-of-Thought

Shin'ya Yamaguchi, Kosuke Nishida, Daiki Chijiwa

Motivation | Multi-modal CoT with LVLM

- **Large Vision-Language Model (LVLM)** is a multi-modal LLM, which integrate visual input into token space, e.g., Gemma-3, Qwen2.5-VL
- **Multi-modal Chain-of-Thought (CoT)** enhances LVLMs' multi-modal reasoning capability by output via rationale from image and query



Challenge | Ignorance of Rationale in CoT

Finding: LVLMs unlikely use rationales in CoT

① Contributions of inputs in LVLMs

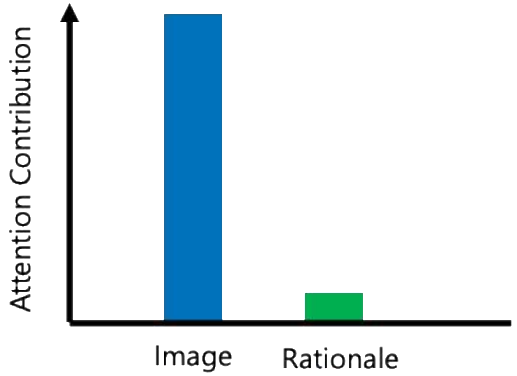
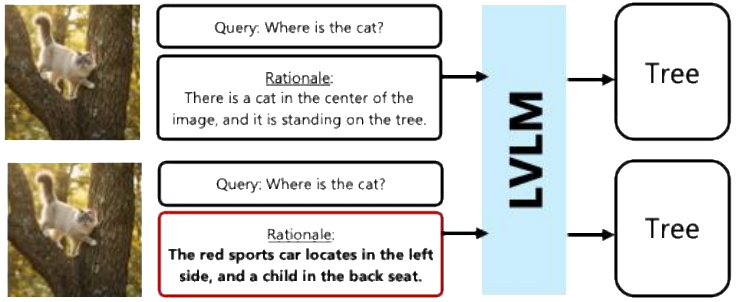


Image tokens dominant in the attention layers

② Randomly Swapping Rationales



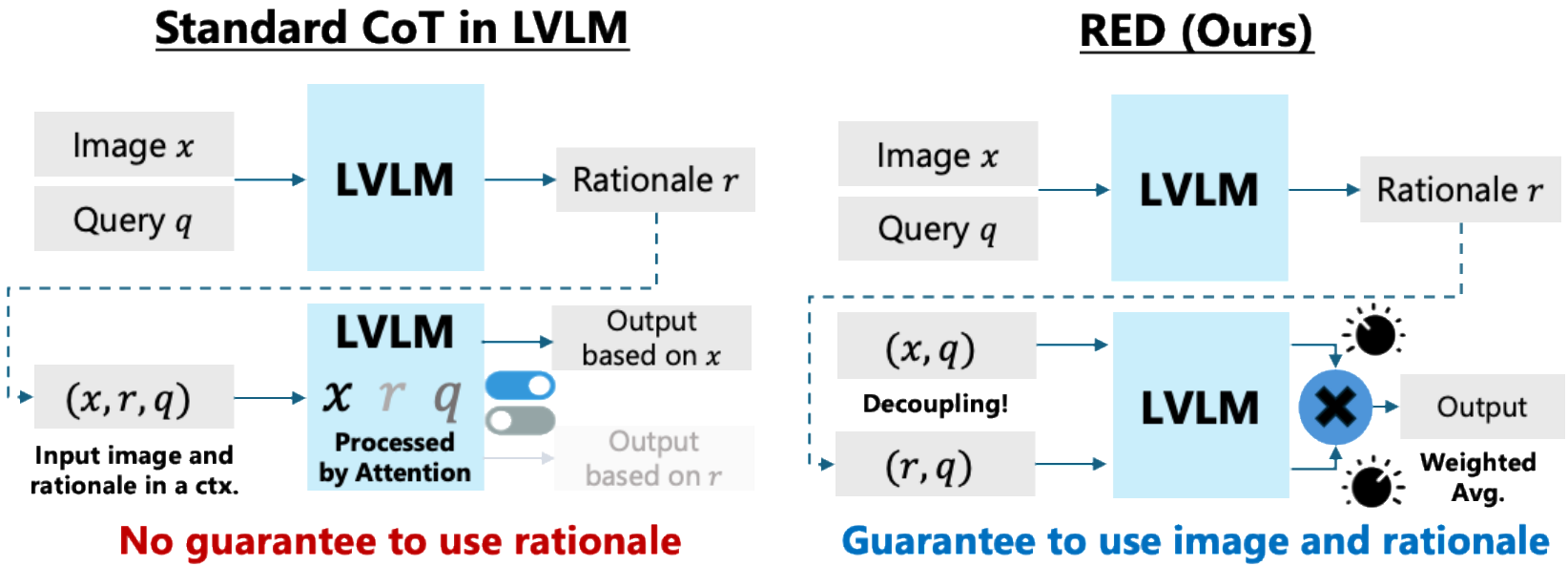
Rationale contents hardly affect the outputs

③ Next-token Distribution: $p_\theta(y_i) := p_\theta(y_i|\boldsymbol{y}_{<i}, x, r, q)$

There is no guarantee the model uses rationales in decoding

Method | Rationale-Enhanced Decoding

Idea: Decoupling rationales and image conditioning



Method | Rationale-Enhanced Decoding

❖ **Reformulating CoT as reward maximization on rationale-conditioning**

$$\max_{\pi} \mathbb{E}_{\pi}[\log p(y_i \sim \pi | y_{<i}, r, q)] - \frac{1}{\lambda} \mathbb{D}_{\mathrm{KL}}[\pi || p_\theta(y_i | \boldsymbol{y}_{<i}, x, q)]$$

Reward: Log-likelihood of rationale-conditional distribution

Constraint: KL-div. from image-conditional distribution

❖ **Closed-form Solution: Power-of-Expert (PoE)**

$$p_\theta(y_i) \propto p_\theta(y_i | \boldsymbol{y}_{<i}, x, q) \times p_\theta(y_i | \boldsymbol{y}_{<i}, r, q)^{\lambda}$$

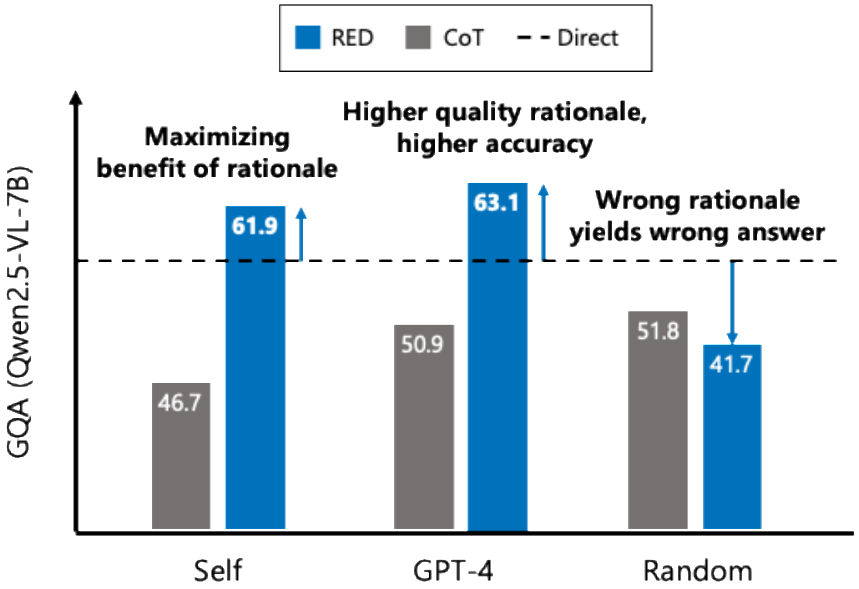
Balancing Weight

Image-conditional distribution

Rationale-conditional distribution

Results | Proof-of-Concept by Intervention

❖ **RED successfully enhances the dependency on the rationale**



	Self (↑)	GPT-4 (↑)	Random (↓)
Gemma-3-4B			
CCoT	44.54 (+4.54)	45.79 (+5.79)	44.35 (+4.35)
CCoT + RED	**45.87** (+5.87)	**48.93** (+8.93)	39.36 (-0.64)
Gemma-12-4B			
CCoT	44.50 (-0.84)	45.61 (+0.27)	44.30 (-1.04)
CCoT + RED	**47.50** (+2.16)	**50.04** (+4.70)	43.29 (-2.05)
Qwen2.5-VL-7B			
CCoT	46.69 (-14.19)	50.85 (-10.03)	51.76 (-9.12)
CCoT + RED	**61.92** (+1.04)	**63.11** (+2.23)	41.74 (-19.14)
Llama3-LLaVA-Next-8B			
CCoT	60.43 (-4.79)	57.84 (-7.38)	60.91 (-4.31)
CCoT + RED	**65.91** (+0.69)	**67.88** (+2.66)	58.70 (-6.52)

Conclusion

❖ Rationale-Enhanced Decoding (RED)

- ▹ An optimal, plug-and-play decoding method for LVLMs to harmonize visual and rationale information by PoE of image- and rationale-conditional distributions
- ▹ Opening a new direction to improve the faithfulness of LVLMs by decoding

❖ Personal Tips for Top-Venue Submission

- ▹ Solve real-world problems (≠ existing problems)
- ▹ Start from the simplest research question and method
- ▹ Read a lot and Write a lot (You CAN'T outsource your curiosity to LLMs)