CVPR 2026 · ABSTRACT VISUAL REASONING

Human-like Abstract Visual Reasoning via Understanding and Solving Reasoning Loop

Xinwang Chen[1], Xiuxing Li[1], Qing Li[1], Ziyue Zhuang[1], Yutong Wu[1], Ziyu Li[1], Zhuo Wang[1], Kai Li[2], Jianye Hao[2], Xia Wu[1]

[1] School of Computer Science and Technology, Beijing Institute of Technology · [2] Huawei Noah's Ark Lab

ARC-AGI-1 ACCURACY	PARAMETERS	IMPROVEMENT OVER TRM	MEAN REASONING STEPS
47.2%	7M	+2.6pt	12.4

Background & Motivation

ARC-AGI · few-shot abstract reasoning

ARC-AGI asks a model to infer the latent transformation rule from only **2–5 input/output demonstrations** , and apply it to a new problem input.

Existing models

1. One-shot, **static** example processing
2. Weak alignment between **understanding** and **solving**
3. Limited generalisation on ARC-AGI

Static · single pass

vs.

Human reasoning

1. Infer rules from **few** examples
2. **Draft** a tentative solution and verify
3. **Refine** hypotheses iteratively

Iterative · self-correcting

CORE GAP **Static processing misses the dynamic alignment between understanding and solving.**

The Human Reasoning Loop

Understanding ↔ Solving co-evolve

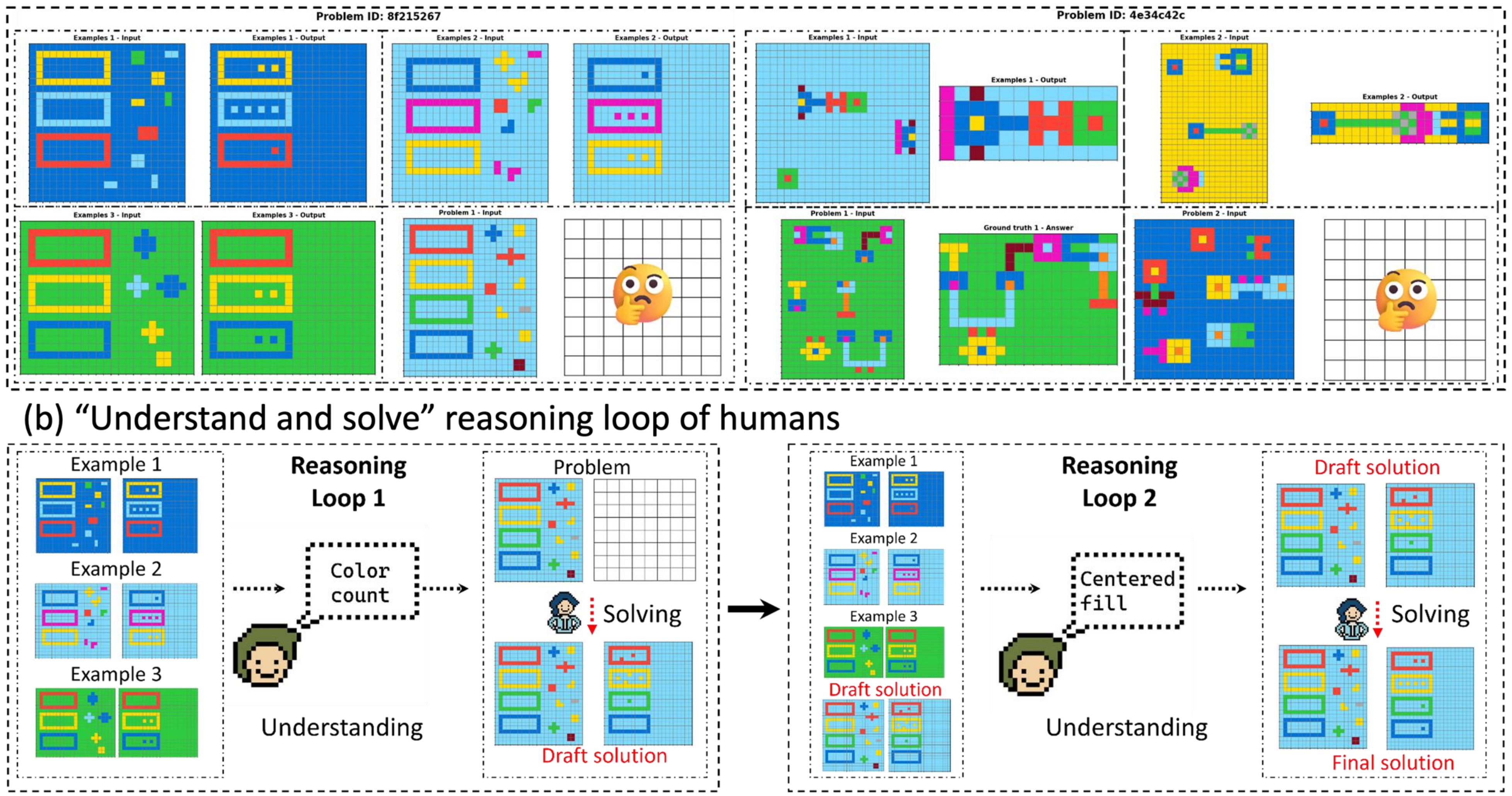


Fig. 1 — A human alternates between forming a rule (Understanding) and drafting an answer (Solving) across multiple loops.

1 Loop 1 · "Color count"

From three examples the solver forms a tentative rule and produces a **draft solution** — usable, but not yet consistent with every example.

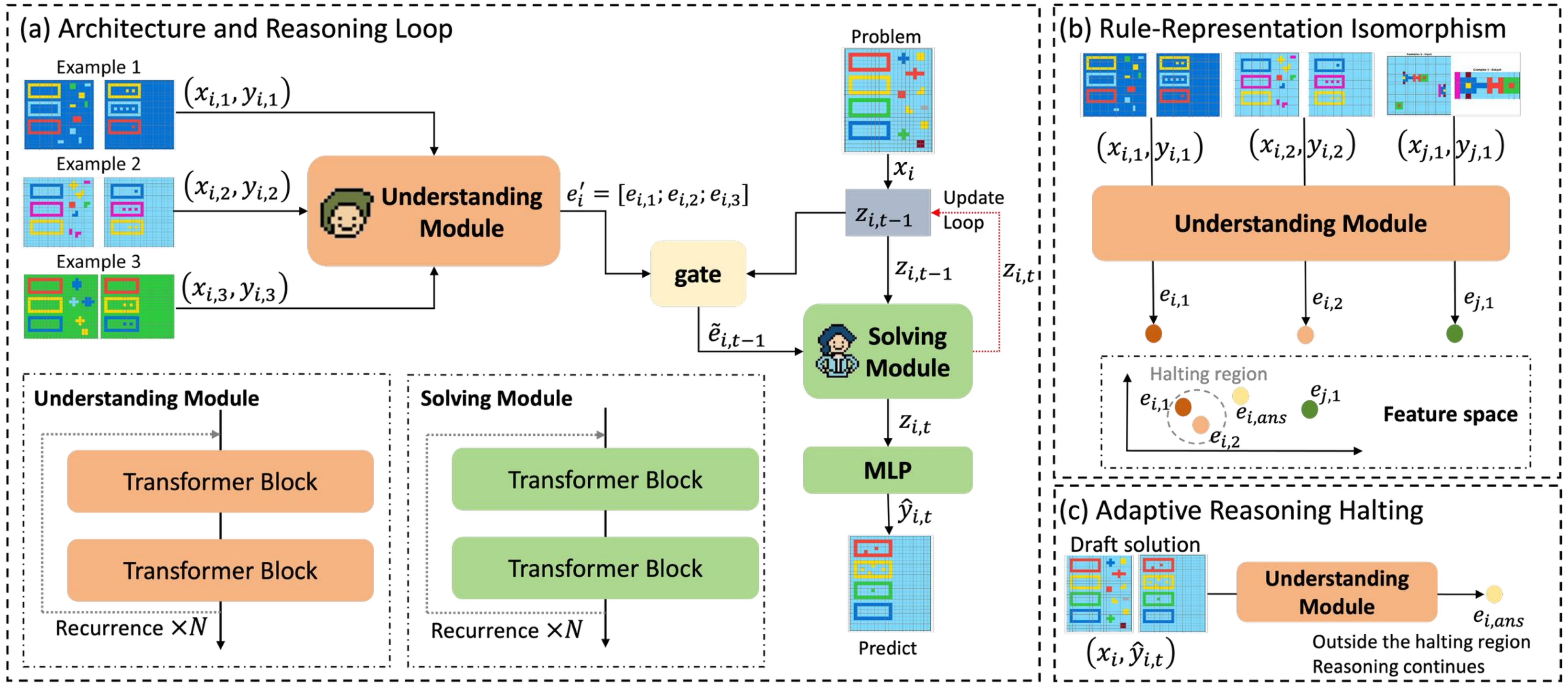
2 Loop 2 · "Centered fill"

Using the draft as new evidence, the solver **re-understands** and emits a refined **final solution**.

INSIGHT **Understanding and solving must co-evolve until they agree.**

Method · USRL Framework

Understanding–Solving Reasoning Loop



UM
Understanding

Gate
Modulation

SM
Solving

1. **UM** reads the examples and emits a raw rule representation e'.
2. **Gate** modulates e' with the current solution state z_{t-1}.
3. **SM** updates the state to z_t and produces a draft solution $\hat{y}_t$.
4. Loop until understanding & solving are **consistent**.

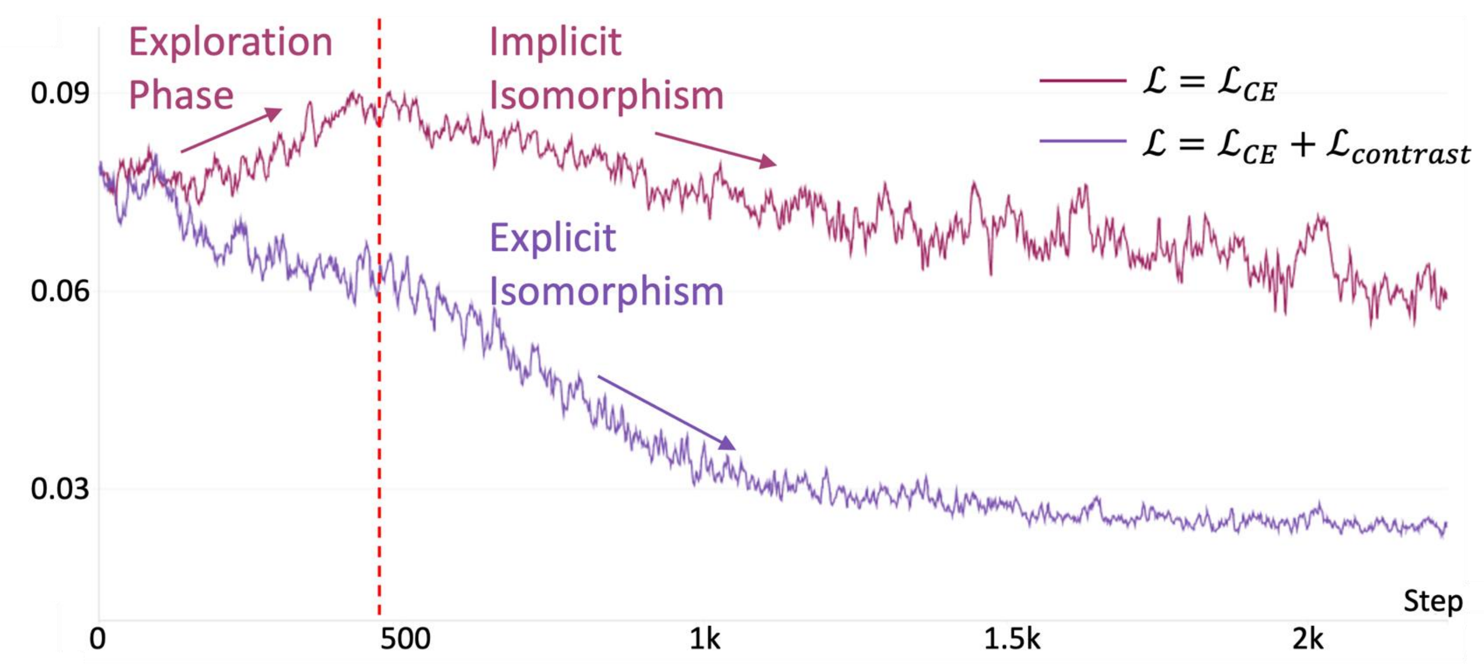
UM · Understanding Module Gate · MLP & scaler SM · Solving Module

Key Mechanism · Isomorphism & Halting

What makes the loop converge

A Rule-Representation Isomorphism

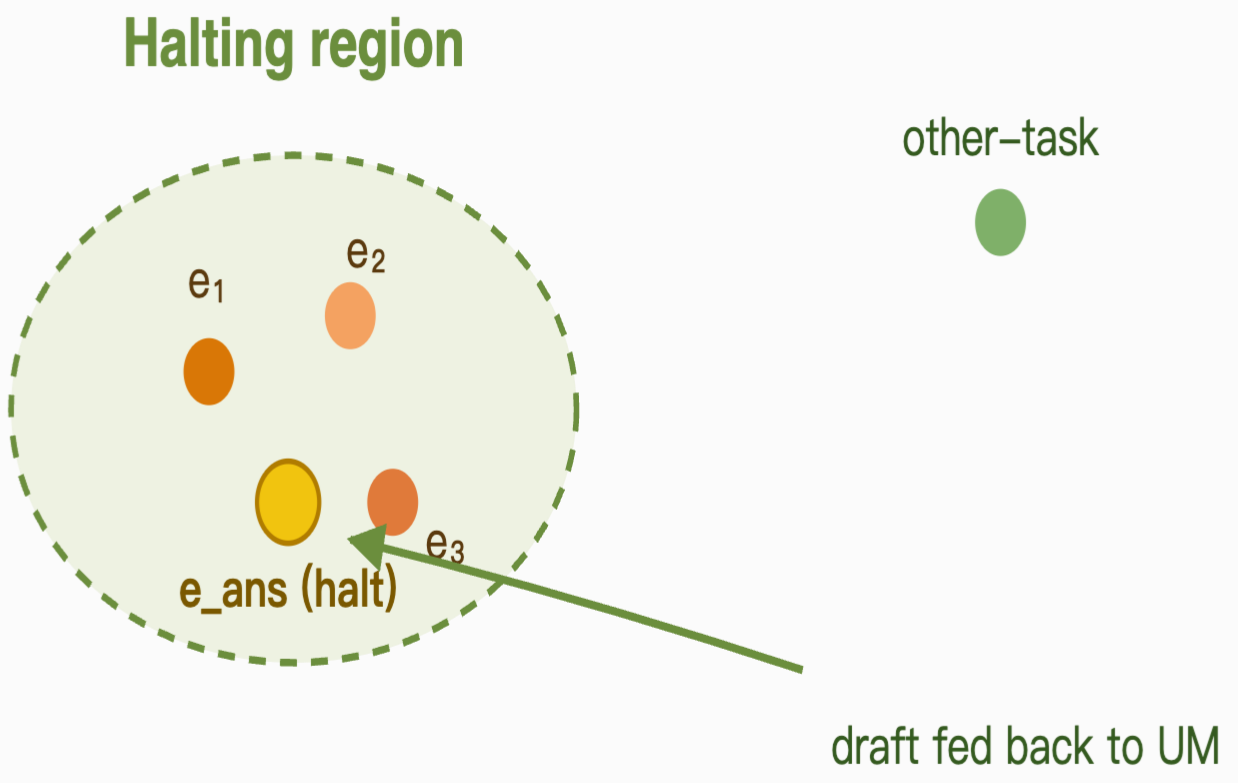
A **contrastive loss** pulls rule embeddings from the same task together and pushes different tasks apart.



OBJECTIVE $\mathcal{L} = \mathcal{L}_{CE} + \mathcal{L}_{contrast}$

B Adaptive Reasoning Halting

Feed the draft $(x_i, \hat{y}_t)$ back to **UM** as a pseudo-example; halt when its embedding lands inside the task cluster.

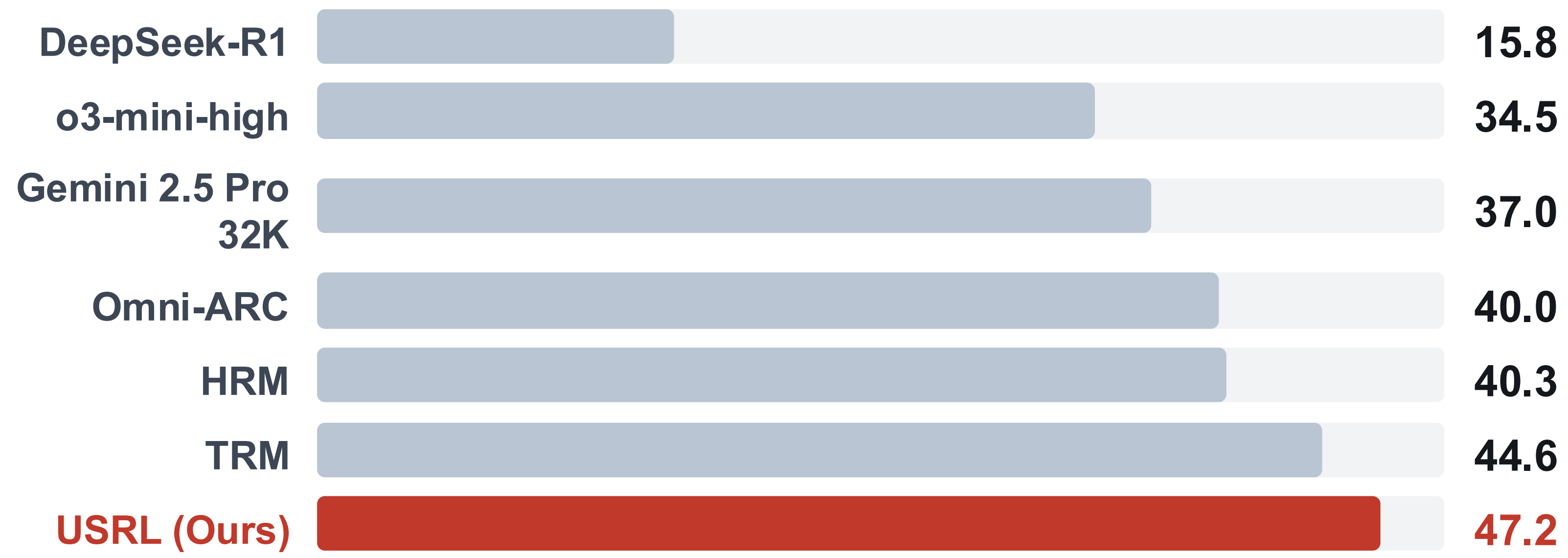


HALT IF $C_{ans} > C_{intra}$

Experiments · ARC-AGI-1

7M parameters · 47.2 % accuracy

Model comparison on ARC-AGI-1



Best ARC-AGI-1 accuracy among compared models — +2.6 pt over TRM

Ablation (cumulative)

SETTING	ACC.	Δ
① Base (SM only)	**36.5**	—
② + UM (w/o gate)	**40.8**	**+4.3**
③ + UM (w/ gate)	**43.2**	**+2.4**
④ + Contrastive loss	**46.4**	**+3.2**
⑤ + Stochastic halting	**47.2**	**+0.8**
⑥ + Adaptive halting (inf.)	**47.0**	**−0.2**

12.4 **mean reasoning steps** at T_{max}=16 — adaptive halting preserves peak accuracy while **the model thinks longer only when it needs to.**

Takeaway

01

Small models can reason.

A 7M-parameter network surpasses much larger baselines on ARC-AGI-1 — scale is not the only path to abstract reasoning.

02

Structure the loop, not the size.

Explicitly modelling the **understanding** ↔ **solving** loop, with rule-representation isomorphism and adaptive halting, is the key inductive bias.

03

Next: compositional sub-skills.

Residual failure modes — counting, comparison, multi-step composition — point to clear directions for future small-model reasoning research.

Thank you.

Xinwang Chen et al.
Beijing Institute of Technology · Huawei Noah's Ark Lab

CVPR JUNE 3-7, 2026 DENVER COLORADO