

Three Diverse Tasks:

- 1 Language Modeling
- 2 Multimodal Understanding
- 3 Visual Generation

Stage1: Pretraining

Mixed Long-CoT Data:

Reasoning-intensive tasks:

DeepSeek-R1
LMM-R1 ...

World knowledge-aware T2I Gen:

GPT-4.1

Primary
CoT DataVerify and
ReformatMixed Long-CoT Data
with Unified Format

Stage2: Mixed Long-CoT Finetuning

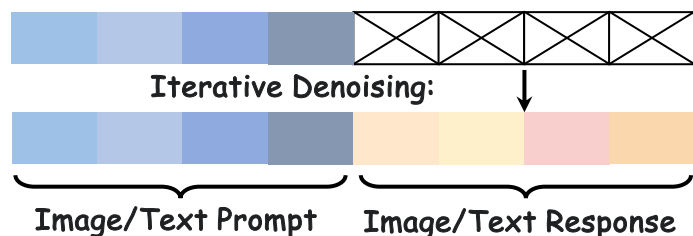
UniGRPO Diversified Rewards:

$$J_{UniGRPO}(\theta) = \mathbb{E}_{o \sim \pi_{\theta}(\cdot|q)}[\mathcal{F}(R^{Uni}(o)) - \beta P(o)]$$

- 1 Textual Reasoning
 R^{LM} : Correctness, Format, ...
- 2 Multimodal Reasoning
 R^{MMU} : Correctness, CLIP-Score, Format, ...
- 3 Text-to-Image Generation
 R^{T2I} : CLIP-Score, ImageReward, ...

Stage3: UniGRPO Training

Discrete Diffusion Generation



Remask Response for Masked Log Probs



Masked Token Log Probs:

$$\pi'_{\theta}(o_i|q, p) = \frac{1}{|M|} \sum_{o_{i,t} \in M} \log p_{\theta}(o_{i,t}|q)$$

Uniformly Random Masking in Iterations

Iteration 1, $p = 0.25$ $\rightarrow \pi'_{\theta}(o, p_1)$ Iteration 1, $p = 0.50$ $\rightarrow \pi'_{\theta}(o, p_2)$ Iteration 1, $p = 0.75$ $\rightarrow \pi'_{\theta}(o, p_3)$

...

Sequence Log Probs by Averaging $\pi'_{\theta} = \frac{1}{\mu} \sum_{p=1}^{\mu} \pi'_{\theta}(o, p_i)$

Input Token Completion Token Mask Token Remask Token

Textual Reasoning

Q: How many one-fourths are there in 7/2?

Text
Tokenizer

Multimodal Reasoning

Q: What is the cat in the image doing?

Image
TokenizerText
Tokenizer

Text-to-Image Generation

Q: Generate an image about a famous landmark building in Europe.

Text
TokenizerReason and Generate with
World Knowledge

MLLM (Discrete Diffusion)

A: <think> Asking this question is the same as ... </think> 14

A: <think> Okay, there's a small white cat pushing... </think> The cat is pushing a small shopping cart.

A: <think> The Eiffel Tower is ...</think> Eiffel Tower is reaching for the sunny sky.

