

Nabla-Reasoner: Gradient Descent at Inference Time Makes LLMs Think Harder

Kabui, Charles

2026-03-14

[Read at ToKnow.ai](#)



Researchers at UT Austin, UC San Diego, TTIC, and Georgia Tech introduced [-Reasoner](#), a framework that applies gradient descent to a language model’s token outputs during inference to improve reasoning. Instead of generating many candidate answers and picking the best one (“Best-of-N”), -Reasoner uses Differentiable Textual Optimization (DTO): it takes the model’s initial output logits (the raw scores before token selection), then runs gradient descent

using signals from both the base model and a reward model to push the output toward better answers. It refines one token at a time, accepting a change only if it leads to a higher-scoring response. On [MATH-500](#), AIME, and AMC benchmarks, -Reasoner scored 80.4% on MATH-500 with Qwen-2.5-7B-Instruct, up from 71.2% with greedy decoding. It matched training-based methods like GRPO while using [up to 40% fewer model calls](#) than Best-of-N.

Current reasoning methods scale by generating more tokens: more serial model calls. -Reasoner scales with more parallel computation per call, since gradient updates across all tokens happen in a single forward-backward pass. Better answers at lower serving cost. The [paper](#) also proves that optimizing outputs via gradient descent at inference time is mathematically equivalent to KL-regularized reinforcement learning, just applied per-sample instead of across the dataset.

This connects to how the field thinks about [making LLMs reason better](#). Training-time and inference-time improvements may not be separate strategies but two views of the same optimization. If gradient-based inference scaling keeps closing the gap with RL-trained models, the question becomes less “how do we train better reasoners?” and more “how much compute do we spend at inference?”

Sources:

- [-Reasoner Paper \(arXiv\)](#)
- [GitHub Repository \(VITA-Group\)](#)
- [Hugging Face Paper Page](#)
- [MATH Benchmark \(Hendrycks et al.\)](#)

Disclaimer: For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation. Read more: [/terms-of-service](#)