Learning to Refine with Fine-Grained Natural Language Feedback

Manya Wadhwa, Xinyu Zhao, Jessy Li, Greg Durrett
manya.wadhwa@utexas.edu

Detect, Critique, and Refine (DCR) enables better post-hoc refinement

Input (LLM response)

A California judge has issued a temporary injunction against the state's high … has caused controversy since its implementation in 2005. … allowed thousands of students who did not pass to potentially graduate.

Step 1: ***Detect*** M_{detect} marks incorrect sentences

A California judge has issued a temporary injunction against the state's high…

… has caused controversy since its implementation in 2005.

In response to a lawsuit by 10 students who failed the exam, a judge has allowed thousands of students who did not pass to potentially graduate.

Step 2: ***Critique*** $M_{critique}$ generates feedback per error sentence

The document does not specify the year … The error span is:… To fix this consider removing the span …

The document does not specify the judge's … The error span is:… To fix this consider adding …

Step 3: ***Refine*** M_{refine} edits the input using combined per-sentence feedback

A California judge has issued a temporary injunction against the state's high ….. 10th-grade level English, has caused controversy since its implementation ~~*in 2005*~~*. In response to a lawsuit by 10 students who failed the exam, a judge* ~~*has allowed thousands of students who did not pass to potentially graduate*~~ ***issued a temporary injunction against California's high school exit exam***

LLM response → Fine-Grained Detect Model M_{detect} → Error Sentences → Critic Model $M_{critique}$ → Per-sentence Feedback → LLM response + Combined Feedback → Refine Model M_{refine} → Refined Output

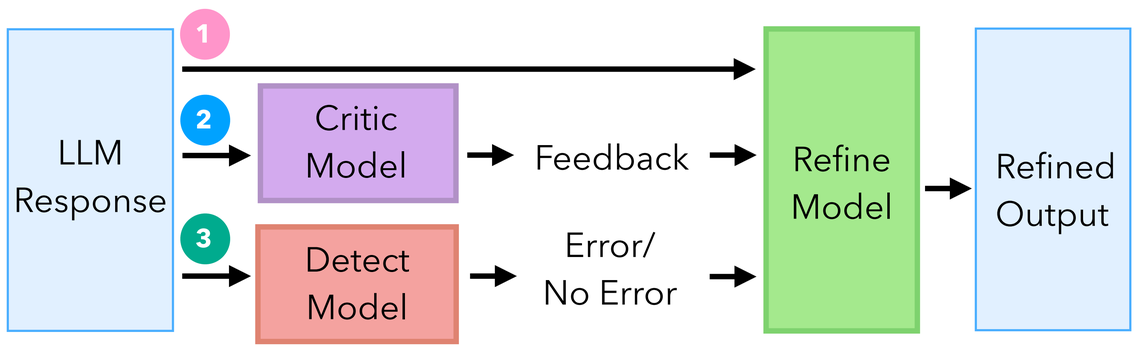
Experiments

Task: Improve factual consistency of document grounded summaries
Datasets: UltraChat (Ding et al, 2023), TofuEval (Tang et al, 2024)
Evaluation Metrics: Δ AlignScore, GPT-4 Win Rate on Factuality
(more metrics in the paper!)

Detect Model	MiniCheck (Tang et al, 2024)
Critic Model	Llama-2-7b-Chat (FT) Llama–3-8b-Instruct (FT) GPT-4 (prompted)
Refine Model	Llama-2-7b-Chat (FT) Llama–3-8b-Instruct (FT) GPT-4 (prompted)

Baseline Refinement Methods

1 Direct Refine 2 Feed + Refine 3 Detect + Refine

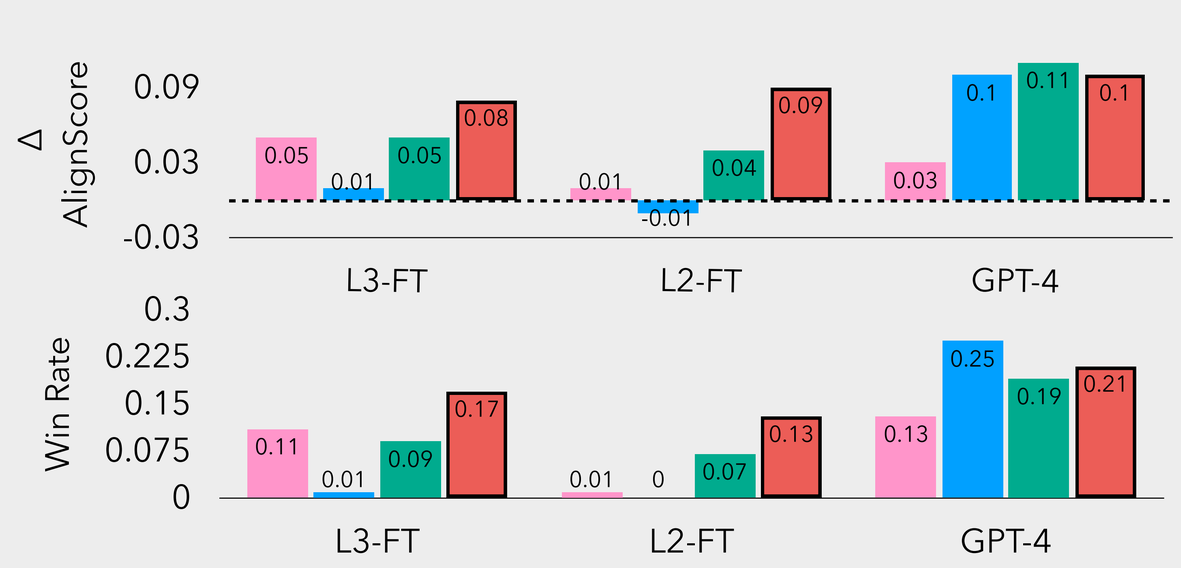


Results

Result 1: DCR refinement outperforms standard refinement strategies

Direct Refine
Feed+Refine
Detect + Refine
DCR (proposed)

Downstream evaluation of refinement on TofuEval



Result 2: Proposed M_{critic} *in* DCR outperforms refining with off-the-shelf critic models

Shepherd
UltraCM
Selfee-13b
DCR (proposed)

Downstream evaluation of refinement on TofuEval

Takeaways

1. DCR does better than baseline refinement techniques
2. Decomposing *Critique* into *Detect* and *Critique* helps train targeted and more accurate critic models
3. Fine-tuned models can perform at-par with stronger models using DCR

More results on critique correctness and faithful refinements in the paper!