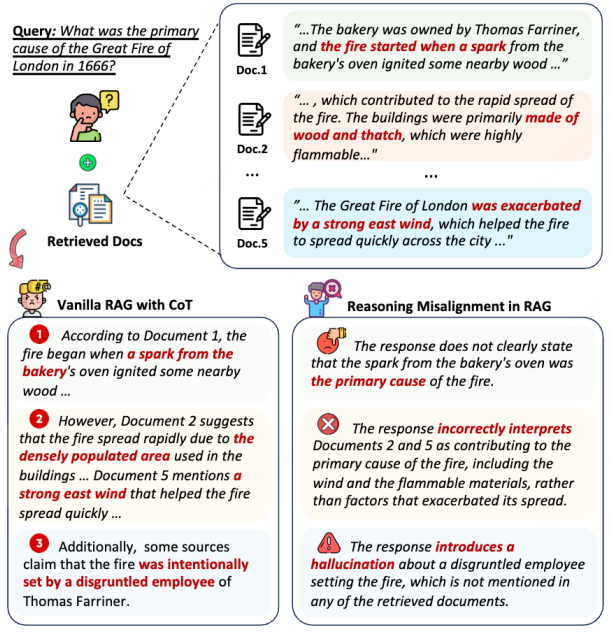
Retrieval is Not Enough: Enhancing RAG Reasoning Through Test-Time Critique and Optimization

Jiaqi Wei, Hao Zhou, Xiang Zhang, Di Zhang, Zijie Qiu, Wei Wei, Jinzhe Li, Wanli Ouyang, Siqi Sun

Motivation

- **Noisy retrieval**: RAG systems exhibit notable fragility, particularly when confronted with irrelevant or noisy retrieved evidence
- **Misalignment**: Standard RAG pipelines often fail to ensure that model reasoning remains consistent with the evidence retrieved, leading to factual inconsistencies.

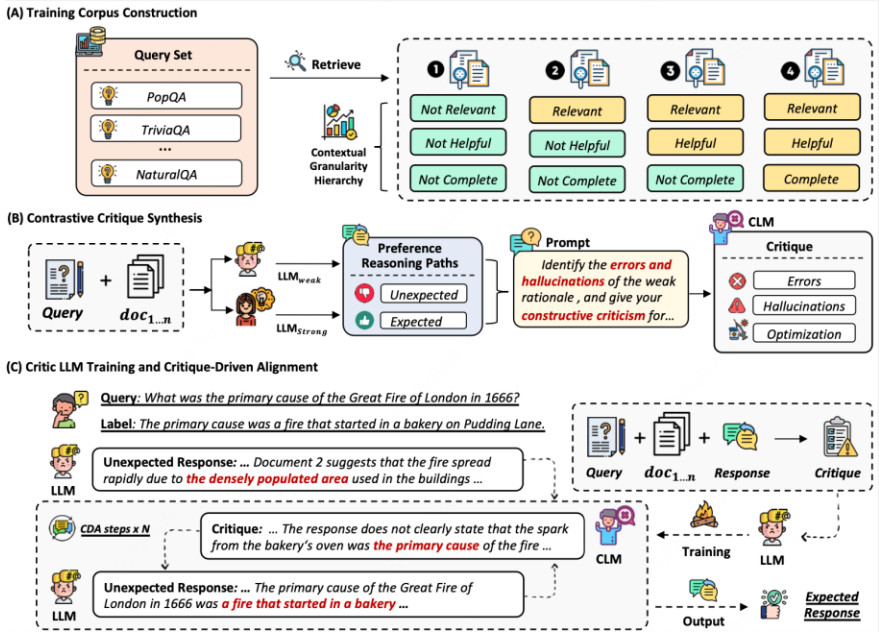


Reasoning Misalignment in RAG

We decompose RAG reasoning into **three interdependent phases**, each susceptible to misalignment despite ideal retrieval:

- Phase 1: Relevance Assessment
- Phase 2: Query-Evidence Mapping
- Phase 3: Evidence-Integrated Synthesis

Overview of AlignRAG Framework



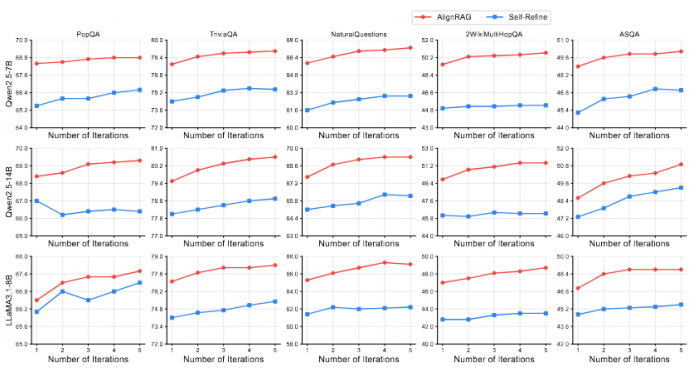
Key Contributions :

- ✓ Reconceptualize RAG as Retrieval-Augmented Reasoning, exposing **Reasoning Misalignment**.
- ✓ We introduce **Critique Learning**, a novel training pipeline enabling CLMs to generate retrieval-augmented critiques while mitigating self-preference bias via **contrastive synthesis**.
- ✓ We propose **AlignRAG**, a test-time reasoning optimization framework using **CDA** steps for iterative refinement
- ✓ **AlignRAG-auto**, a autonomous extension that adaptively determines the optimal refinement depth for improved efficiency and usability.

Main Results

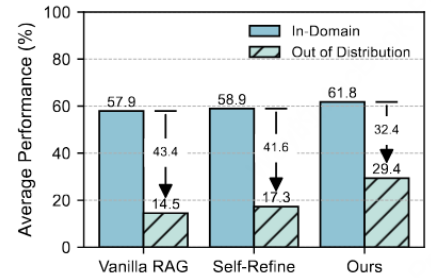
Method / Metric	NQ accuracy	MultiHopQA accuracy	TriviaQA accuracy	PopQA accuracy	ASQA str-em	Avg.
Baselines w/o Retrieval						
Chain-of-thought [48]						
Qwen-2.5-Instruct$_{7B}$	33.9	45.0	58.3	26.9	20.5	36.9
Qwen-2.5-Instruct$_{14B}$	48.1	49.3	72.8	25.4	31.6	45.4
Llama-3.1-Instruct$_{8B}$	42.1	41.9	61.8	26.9	25.1	40.0
Standard RAG with Reasoning						
Vanilla Reasoning						
Qwen-2.5-Instruct$_{7B}$	60.2	44.7	73.2	63.7	42.8	56.9
Qwen-2.5-Instruct$_{14B}$	63.6	44.8	77.0	65.3	45.2	59.2
Llama-3.1-Instruct$_{8B}$	62.0	43.0	73.4	65.0	45.2	57.7
RAG w/ Training-time Refinement						
RetRobust [33]						
Llama-2$_{13B}$*	39.6	51.5	–	–	–	–
Llama-3-Instruct$_{8B}$*	54.2	54.7	71.5	56.5	40.5	55.5
InstructRAG [13]						
Qwen-2.5-Instruct$_{7B}$	63.8	46.3	76.1	67.5	47.5	60.2
Qwen-2.5-Instruct$_{14B}$	66.3	47.3	78.7	67.8	48.5	61.7
Llama-3.1-Instruct$_{8B}$	66.3	45.1	76.6	66.9	47.2	60.4
RAG w/ Test-time Refinement						
Self-RAG [19]						
Llama-2$_{7B}$ + CLM$_{7B}$*	42.4	35.9	68.9	55.8	30.0	46.6
Llama-2$_{13B}$ + CLM$_{13B}$*	46.4	36.0	70.4	56.3	31.4	48.1
Llama-3-Instruct$_{8B}$ + CLM$_{8B}$*	42.8	32.9	71.4	55.8	36.9	48.0
Self-Refine						
Qwen-2.5-Instruct$_{7B}$ + SELF$_{7B}$	61.6(Δ)	45.0(Δ)	74.4(Δ)	65.5(Δ)	45.2(Δ)	58.3(Δ)
Qwen-2.5-Instruct$_{14B}$ + SELF$_{14B}$	65.1(Δ)	46.1(Δ)	78.0(Δ)	67.0(Δ)	47.3(Δ)	60.7(Δ)
Llama-3.1-Instruct$_{8B}$ + SELF$_{8B}$	61.4(Δ)	42.8(Δ)	74.1(Δ)	66.1(Δ)	44.7(Δ)	57.8(Δ)
AlignRAG-fixed						
Qwen-2.5-Instruct$_{7B}$ + CLM$_{8B}$	65.9 (↑ 4.3%)	49.5 (↑ 4.5%)	77.8 (↑ 3.4%)	68.4 (↑ 2.9%)	48.9 (↑ 3.7%)	62.1 (↑ 3.8%)
Qwen-2.5-Instruct$_{14B}$ + CLM$_{8B}$	67.7 (↑ 2.6%)	49.8 (↑ 3.7%)	79.5 (↑ 1.5%)	68.4 (↑ 1.4%)	48.6 (↑ 1.3%)	62.8 (↑ 2.1%)
Llama-3.1-Instruct$_{8B}$ + CLM$_{8B}$	65.3 (↑ 3.9%)	47.0 (↑ 4.2%)	77.0 (↑ 2.9%)	66.5 (↑ 0.4%)	47.1 (↑ 2.4%)	60.6 (↑ 2.8%)

Iterative Alignment in RAG

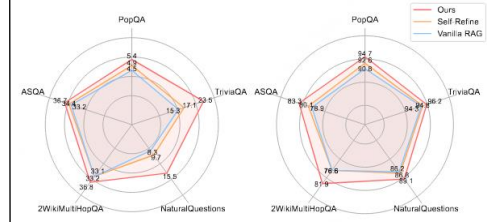


More Analysis

Generalization to OOD Scenarios



When Retrieval Falters, AlignRAG Thrives



(a) w/o answer. (b) w/ answer.

Integrate as a Plug-and-play Module

Method	ID (avg.)	OOD (avg.)
Qwen2.5-7B		
InstructRAG	59.5 (Δ)	28.0 (Δ)
w/ Alignment	61.5 (↑ 2.0%)	30.1 (↑ 2.1%)
w/ Alignment^{1}	63.0 (↑ 3.5%)	31.7 (↑ 3.7%)
Qwen2.5-14B		
InstructRAG	61.7 (Δ)	24.9 (Δ)
w/ Alignment	62.5 (↑ 0.8%)	33.4 (↑ 8.5%)
w/ Alignment^{1}	63.9 (↑ 2.2%)	34.3 (↑ 9.4%)
LLaMA3.1-8B		
InstructRAG	60.4 (Δ)	28.4 (Δ)
w/ Alignment	61.7 (↑ 1.3%)	29.4 (↑ 1.0%)
w/ Alignment^{1}	61.9 (↑ 1.5%)	30.5 (↑ 2.1%)