Learning Minimal Sufficient Evidence Graphs for GraphRAG via Nash-Guided Optimization

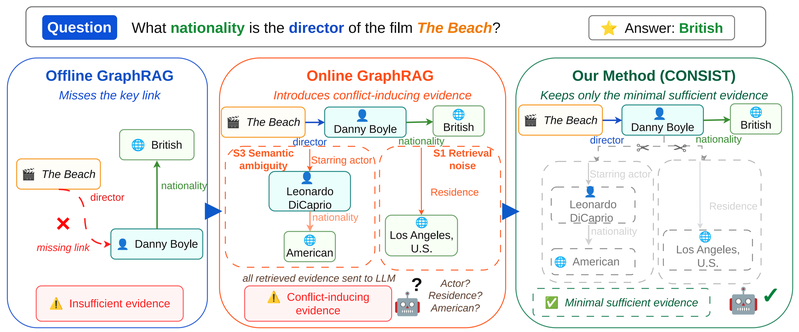
Hao Wu[1,2], Xinguo Yu[1,†], Wanqing Li[2], Hongru Sun[2], Xiao Luo[3], Wenbin Zhang[4], Sasha Nikolic[2], Shirui Pan[5], Yi Guo[6,†], Jie Yang[2,†]
[1]University of Wollongong [2]Central China Normal University [3]University of Wisconsin-Madison
[4]Florida International University [5]Griffith University [6]University of Western Sydney

NEURAL INFORMATION PROCESSING SYSTEMS
NeurIPS 2026

How can GraphRAG construct an evidence graph that is sufficient for answering a query while remaining compact and conflict-free?

1. Motivation, Gap and Contributions

Motivation. GraphRAG organizes retrieved knowledge into connected evidence for multi-hop reasoning. Offline graphs may miss query-critical links, while online expansion may introduce redundant evidence, counterfactual facts, or semantic ambiguity. The goal is to retain sufficient evidence while limiting misleading or unnecessary structures.



2. Graph Construction and Conflict-Aware Utility

Learning objective. Given query q, search the reachable graph space $\mathcal{G}(q)$ for evidence that supports answering q, avoids incompatible knowledge, and remains compact. The weight λ balances Consistency and Minimality.

$$G_q^* = \arg\max_{G_q \in \mathcal{G}(q)} \; S(G_q, q) + \lambda C(G_q, q) + (1-\lambda)\mathcal{M}(G_q). \tag{1}$$

Candidate construction. Extract query entities, align them with document nodes, and select top-m LLM-ranked seeds. Start with $G_0 = \{\mathcal{V}_{\text{seed}}, \varnothing\}$ and frontier $\mathcal{F}_0 = \mathcal{V}_{\text{seed}}$. Expand structural neighbors and semantic relations whose similarity is at least τ_{sem}:

$$\begin{cases} \mathcal{V}_{t+1} = \mathcal{V}_t \cup \bigcup_{v \in \mathcal{F}_t} \big(\mathcal{N}_{\text{str}}(v) \cup \mathcal{N}_{\text{sem}}(v)\big), \\ \mathcal{E}_{t+1} = \mathcal{E}_t \cup \big\{(v, v_{\text{str,sem}}) \mid v \in \mathcal{F}_t, \\ \qquad v_{\text{str,sem}} \in \mathcal{N}_{\text{str}}(v) \cup \mathcal{N}_{\text{sem}}(v)\big\}. \end{cases} \tag{2}$$

Sufficiency. The answer-support scorer rewards evidence useful for answering the query.

$$S(G, q) = \text{LLM}_{\text{sup}}(G, q). \tag{3}$$

Consistency. LLM_{pair} scores conflict between evidence edges. Subtracting its average rewards more consistent graphs.

$$C(G, q) = 1 - \text{Avg}_{(e_i, e_j) \in \mathcal{E}(G)} \text{LLM}_{\text{pair}}\big(q, (e_i, e_j)\big). \tag{4}$$

Minimality. With G' denoting the graph before expansion, the edge-count ratio acts as a compactness proxy and discourages unnecessary growth.

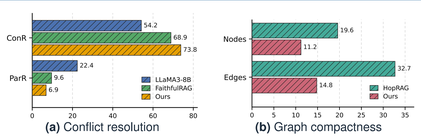
$$\mathcal{M}(G) = |\mathcal{E}(G')|/|\mathcal{E}(G)|. \tag{5}$$

Overall utility. The three terms jointly evaluate candidate edits and future rollouts:

$$\begin{aligned} U(G, q) &= \text{LLM}_{\text{sup}}(G, q) \\ &\quad + \lambda\Big(1 - \text{Avg}_{(e_i, e_j) \in \mathcal{E}(G)} \text{LLM}_{\text{pair}}\big(q, (e_i, e_j)\big)\Big) \\ &\quad + (1-\lambda)\big(|\mathcal{E}(G')|/|\mathcal{E}(G)|\big). \end{aligned} \tag{6}$$

Both scorers use the same frozen backbone, given only the query and retrieved graph, without ground-truth answers. If no prior edges exist, candidate expansion relies solely on semantic relations.

3. Graph Quality: Conflict Resolution and Compactness

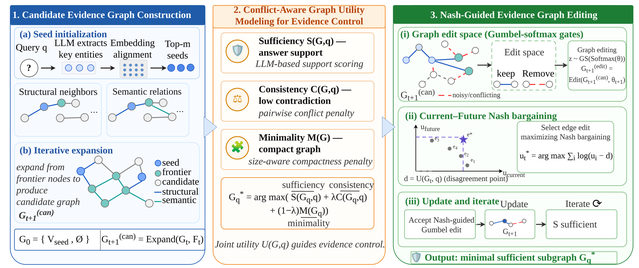


(a) Conflict resolution (b) Graph compactness

We assess context alignment (ConR), parametric-knowledge alignment (ParR), and retained graph size. CONSIST increases ConR and reduces ParR on ConFiQA. On 2Wiki, it retains 11.2 nodes and 14.8 edges, versus HopRAG's 19.6 and 32.7: **42.9% fewer nodes and 54.7% fewer edges**.

4. CONSIST: Exploration and Editing

CONSIST alternates candidate expansion with conflict-aware editing. Structural and semantic connections improve coverage; utility-guided pruning controls which evidence remains in the graph.



Current-future coordination retains useful evidence while limiting unreliable reasoning paths after further expansion.

5. Nash-Guided Evidence Graph Editing

Stochastic edge editing. Each gate determines whether a candidate edge is retained. A Gumbel-based relaxation enables gradient-based editing, with uniform noise $\epsilon \sim \mathcal{U}(0,1)$, temperature Δ, and sigmoid σ:

$$\text{GS}(\boldsymbol{\theta}) = \sigma\left(\frac{\log \epsilon - \log(1-\epsilon) + \boldsymbol{\theta}}{\Delta}\right). \tag{7}$$

$$G_{t+1}^{(\text{edit})} = \text{Edit}\big(G_{t+1}^{(\text{can})}, \boldsymbol{\theta}_{t+1}\big). \tag{8}$$

$$\theta_{t+1}^{e} = \begin{cases} \theta_t^e, & e \in \mathcal{E}_t, \\ 1, & \text{otherwise.} \end{cases} \tag{9}$$

Nodes are removed when all incident edges are pruned. Equation (9) preserves existing edge parameters within a query and initializes new edges at 1, giving them a temporary retention preference before utility-based editing.

Current and future players. Immediate improvement can overlook conflicts introduced by later expansion. We evaluate the edited graph and its rollout after H further expansion steps **without additional editing**:

$$\begin{aligned} u_{\text{cur}}\big(G_{t+1}^{(\text{edit})}; \boldsymbol{\theta}\big) &= \mathbb{E}\big[U\big(G_{t+1}^{(\text{edit})}(\boldsymbol{\theta}), q\big)\big], \\ u_{\text{fut}}\big(G_{t+1+H}^{(\text{edit})}; \boldsymbol{\theta}\big) &= \mathbb{E}\big[U\big(G_{t+1+H}^{(\text{edit})}(\boldsymbol{\theta}), q\big)\big]. \end{aligned} \tag{10}$$

Expectations use Monte Carlo sampling. The utility set $\mathcal{U}_{t+1}(\boldsymbol{\theta})$ is induced by candidate edits. Nash-guided selection favors joint improvement and penalizes one-sided gains:

$$u_{t+1}^* = \arg\max_{u_i \in \mathcal{U}_{t+1}(\boldsymbol{\theta})} \sum_{i \in \{\text{cur,fut}\}} \log(u_i - d_t). \tag{11}$$

With $d_t = \mathbb{E}[U(G_t, q)]$, both players must improve over the unchanged graph; otherwise, keep it. Positive utilities alone, as allowed by $d_t = 0$, do not ensure improvement. Stop when $S(G, q) \geq \tau_{\text{stop}}$, then generate $\hat{a} = \text{LLM}(G_q^*, q)$. Graphs and editing parameters are not reused across queries.

6. Main Experimental Results

Setup. We evaluate EM/F1 on three multi-hop QA datasets and four conflict-oriented settings. CONSIST uses GPT-4o, DeepSeek-V3, and LLaMA-3-8B; its results are averaged over 10 runs.

Table 1: Multi-hop QA with DeepSeek-V3 and GPT-4o.

Method	2Wiki		HotpotQA		MuSiQue		Avg.	
	EM(↑)	F1(↑)	EM(↑)	F1(↑)	EM(↑)	F1(↑)	EM(↑)	F1(↑)
DeepSeek-V3	31.2	36.5	29.4	40.2	7.8	16.4	22.8	31.0
+ RAPTOR	46.7	54.8	47.2	63.1	18.8	28.6	37.6	48.8
+ TOG	32.8	37.3	24.1	29.2	5.3	7.0	20.7	24.5
+ G-Retriever	36.2	45.5	36.9	49.3	9.2	16.9	27.4	37.2
+ LightRAG	30.0	36.4	38.4	49.8	10.4	17.1	26.3	34.4
+ GraphRAG	31.8	37.9	45.0	56.9	13.8	20.6	30.2	38.5
+ HippoRAG	57.8	68.4	49.8	64.7	19.0	29.8	42.2	54.3
+ ReLink	62.8	72.2	55.8	70.4	25.2	37.0	47.9	59.9
+ **CONSIST**	**65.3**	**73.9**	**62.6**	**77.8**	**43.5**	**55.9**	**57.1**	**69.2**
Std.	±1.0	±0.9	±1.3	±1.2	±0.7	±1.1	±1.0	±1.1
GPT-4o	29.2	35.8	33.0	42.4	10.6	21.4	24.3	33.2
+ SiRe/RAG	59.6	67.9	61.7	76.5	40.5	53.1	53.9	65.8
+ TCR-QF	59.8	68.0	55.8	70.8	30.3	43.2	48.6	60.7
+ HopRAG	61.1	68.3	62.0	76.1	42.4	54.9	55.1	66.4
+ **CONSIST**	**67.2**	**75.7**	**64.3**	**79.0**	**44.1**	**56.8**	**58.5**	**70.5**
Std.	±1.1	±0.8	±0.9	±0.8	±0.5	±1.0	±0.8	±0.9

CONSIST achieves the best EM/F1 on 2Wiki, HotpotQA, and MuSiQue. DeepSeek-V3 average EM improves from ReLink's 47.9 to **57.1**.

Table 2: Conflict-oriented QA with LLaMA-3-8B.

Method	FaithEval		ConFiQA (MC)		ConFiQA (MR)		ConFiQA (QA)		Avg.	
	EM(↑)	F1(↑)	EM(↑)	F1(↑)	EM(↑)	F1(↑)	EM(↑)	F1(↑)	EM(↑)	F1(↑)
LLaMA-3-8B	6.0	27.7	2.1	5.0	1.9	6.1	1.3	6.1	2.8	11.2
+ Opin(Instr)	15.1	34.9	57.3	67.4	54.0	65.9	67.4	76.9	48.5	61.3
+ CAD	42.7	59.4	11.4	16.0	31.3	40.0	38.1	48.3	30.9	40.9
+ COIECD	41.3	56.1	24.0	28.5	43.3	50.9	60.1	67.1	42.2	50.7
+ ADACAD	47.7	61.4	32.5	43.3	42.5	46.7	48.1	55.7	42.7	51.8
+ KRE	12.1	59.1	59.8	68.2	58.9	68.7	74.7	84.0	51.4	70.0
+ FaithfulRAG	49.1	62.4	61.2	70.3	53.1	62.4	65.2	70.4	57.2	66.4
+ **CONSIST**	**55.7**	**68.5**	**65.1**	**73.9**	**63.8**	**72.7**	**77.6**	**87.3**	**65.6**	**75.6**
Std.	±1.2	±1.5	±0.7	±0.8	±0.6	±0.6	±0.3	±0.2	±0.7	±0.8

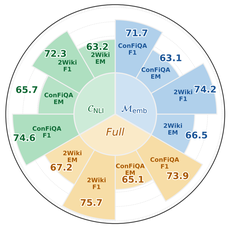
On FaithEval and three ConFiQA settings, CONSIST achieves the best EM/F1 with LLaMA-3-8B. Average EM/F1 reach **65.6/75.6**, compared with FaithfulRAG's 57.2/66.4, supporting reasoning under conflicting evidence.

7. Component and Utility Ablations

Table 3: Impact analysis from individual components for CONSIST, using 2Wiki (GPT-4o) and ConFiQA (LLaMA-3).

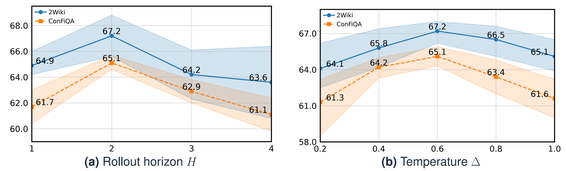
	2Wiki		ConFiQA	
	EM(↑)	F1(↑)	EM(↑)	F1(↑)
Backbone	29.2	35.8	2.1	5.0
+ EXP	56.2	65.1	49.4	55.6
+ GS-LLM	59.4	68.5	54.9	60.4
+ GS-U	63.9	72.8	61.7	67.3
+ GS-SUM	64.6	73.4	63.2	70.1
+ Full	**67.2**	**75.7**	**65.1**	**73.9**

Each component improves performance. Nash coordination further gains **2.3 F1 points on 2Wiki** and **3.8 on ConFiQA-MC** over direct utility summation.



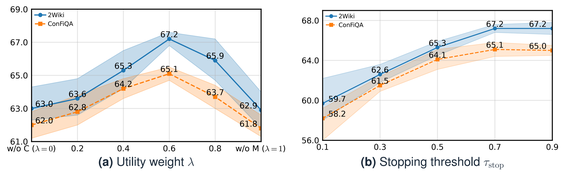
NLI-based Consistency and embedding-based Minimality remain effective. NLI-based Consistency slightly improves ConFiQA performance, supporting flexibility in utility design.

8. Sensitivity Analysis



(a) Rollout horizon H (b) Temperature Δ

Editing sensitivity. Both datasets achieve their highest EM at $H = 2$ and $\Delta = 0.6$.



(a) Utility weight λ (b) Stopping threshold τ_{stop}

Utility balance. EM peaks at $\lambda = 0.6$ and stabilizes around $\tau_{\text{stop}} = 0.7$.

9. Conclusion

This paper presented CONSIST, a conflict-aware Nash-optimized framework for learning minimal sufficient evidence graphs in GraphRAG. Instead of accepting all retrieved or expanded evidence for question answering, CONSIST explicitly regulates graph construction by jointly modeling sufficiency, consistency, and minimality, and further refines candidate graphs through Gumbel-based edge editing guided by current–future Nash bargaining. Experiments on multi-hop question answering and conflict-oriented benchmarks show that CONSIST consistently improves answer accuracy while producing substantially more compact evidence graphs.

10. Acknowledgments

The work of Jie Yang is supported by Australian Research Council Discovery Project (DP210101426). The work of Xinguo Yu and Hao Wu are supported by General Project of the National Natural Science Foundation of China (Grant No: 62277022). The work of Hao Wu is also supported by scholarship from the China Scholarship Council.

Correspondence: hwu@mails.ccnu.edu.cn jiey@uow.edu.au † Corresponding authors **CONSIST**: COnflict-aware NaSh-optimized MInimal SufficienT graph learning