Plug-and-Play Bi-CoT:
설명가능한 Multi-Hop QA 를 위한 자기인식 정방향 추론과 역방향 검증

강지헌 [1,*,©], 정수환 [2,*]
연세대학교 [1], 부경대학교 [2]

danieljiheon@yonsei.ac.kr[1], suhwan0823@gmail.com[2]

Plug-and-Play Bi-CoT: Self-Aware Forward Reasoning and Reverse Verification for Explainable Multi-Hop QA

Kang Ji Heon[1,*,©], Jeong Su Hwan[2,*]
Yonsei Univ.[1], Pukyong National Univ. [2]

요 약

본 논문은 Multi-Hop QA (MHQA)에서 추가적인 학습이 필요 없는 플러그인(Plugin) 모듈을 제안한다. 하위 질문에서 Self-Aware QA 를 적용해 단계별 Reasoning 품질을 높이고, 최종 단계에서 Reverse Verification(역방향 검증)을 통해 논리 경로와 증거 집합의 구조적 타당성을 확인하는 구조이다. 이는 기존 연구의 방향성을 강화하면서도 다른 도메인에서의 적용 가능성을 지닌다. HotpotQA 데이터셋에서 유의미한 성능 향상을 관찰할 수 있었고, 단순 출력만을 반환하는 기존 접근법과 다르게 각 단계별 질문·답변·근거를 명시적으로 제공함으로써 설명가능성을 크게 강화하였다.

Ⅰ. 서 론

Multi-Hop QA(MHQA)는 단일 문서나 사실로는 답할 수 없는 복합 질의에 대응하기 위해, 대규모 데이터 속에서 서로 다른 출처의 증거를 선택적으로 검색하고 연결하는 다단계 추론 과제이다. 최근 MHQA 연구들은 복합 질의를 해결하기 위해 질문을 하위 질문(Sub-question)으로 분해하여 Reasoning 품질을 높이는 전략을 취해왔다.[1]

그러나 이 방안은 두 가지 중요한 한계를 가진다. 첫째, 추론 경로의 불투명성이다. 대부분의 모델은 최종 답과 일부 근거 문서 전체를 반환하기 때문에, 사용자가 답변이 도출된 구체적 논리 구조를 확인하거나 오류 원인을 추적하기 어렵다. 둘째, Forward-only 구조에 따른 오류 전이다. 하위 질문에서 잘못된 답이 나오면 이후 단계에서 이를 바로잡을 기회가 거의 없어 최종 답변이 크게 왜곡될 수 있다.

본 연구는 이러한 한계를 해결하기 위해, 설명 가능성과 오류 완화를 핵심 목표로 하는, 어느 Retriever 과 QA 모델에도 결합해 사용할 수 있는 플러그인(Plugin) 모듈을 제안한다. 구체적으로 각 하위 질문의 답변과 근거를 명시적으로 기록하는 **Self-Aware QA** 단계를 통해 Reasoning 과정을 그대로 노출하고(설명 가능성 강화), Forward 추론 후 **Reverse Verification** 단계를 추가하여 논리 경로와 증거 집합을 다시 검증함으로써 오류 전이를 완화한다.

Ⅱ. 본론

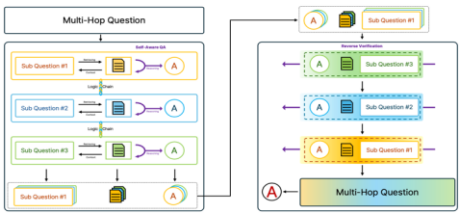

그림 1. 제안한 모듈의 파이프라인

본 모듈의 파이프라인은 하위 질문 분해로 시작해, 선택한 모듈의 Retrieval 과 1 차 QA, 그 결과를 바탕으로 한 Self-aware QA 로 순방향 Reasoning 강화, 그리고 Reverse Verification 을 통해 일관되고 신뢰성 있는 답변을 산출하는 네 단계로 구성된다.

(가) Sub-question 분해

원 복합 질문을 하위 질문(Sub-question)으로 분해하여 논리 과정을 구조화하는 단계이다. 이는 복합 질문을 단계별로 나누어 검색과 추론을 안정화한다.

하위 질문을 생성할 때 의존성(dependency)을 명확히 반영하고, 선행 답변이 필요한 경우에는 해당하는 질문에 치환을 적용한다. 예를 들어, “[Answer of SubQ1] is what kind of food?”로 질문을 생성하여 하위 질문간의 논리 구조를 종속성에 맞게 형성한다. 그리고 검색을 진행할 경우에는 실제 하위 질문 1 번의 답으로 치환하여 질문을 완성한다.

(나-0) Retrieval & QA *(Optional Choice)*

해당 부분은 기존의 MHQA 를 진행할 수 있는 어떤 Retriever 과 QA 모델도 사용 가능하다. 즉, 기존 MHQA 가 진행될 수 있도록 하는 기본 구조이며, 우리의 플러그인 모듈에 해당하지 않는다.

어떤 질문이 들어오면, 방대한 양의 데이터 속에서 관련된 정보를 가져오는 것을 Retriever 모델이 진행한다. 그리고 관련된 정보 속에서 질문의 대답을 찾는 것을 QA 모델이 진행하게 된다.

이 과정을 통해 추후 Reasoning 강화를 위해 필요한 기초 검색 데이터와 1 차 답변을 준비한다.

(나) Self-Aware QA (Forward CoT)

각각의 하위 질문에 대해 Retrieval 결과와 QA 답변을 기반으로, 모델이 스스로 답의 타당성을 점검하고 Reasoning 을 강화하고 경로를 명시적으로 기록할 수 있다.

---

* These authors contributed equally to this work.
© Corresponding author.

질문 속에서 필요한 조건을 추출하고, 그 조건을 가지고 온 데이터에 충족하는지 평가한다. 그 과정에서 구한 답변이 맞는지, 아니라면 새로운 답을 도출한다. 답을 찾는 과정에서 사용한 핵심적인 하나의 데이터, Evidence 를 얻게 된다.

(다) Reverse Verification (Reverse CoT)

Forward-only 구조는 초기 단계의 오류가 최종 답으로 전이된다. 이를 보완하기 위해 원질문에서 역방향으로 논리구조를 재구성하는 Reverse Verification 을 진행한다.

먼저 원질문과 하위 질문을 검토해 답 유형·비교 유형·구조적 문제를 판별한다. 여기서 불필요한 단계를 제거하고, 필수 정보만 남겨 역순으로 단순화한다. 역방향으로 검증하는 과정을 통해 하위 질문 분해에서 놓친 거시적 관점을 회복하고, 질문의 범위와 답 유형의 일치를 점검할 수 있다.

이후 검색 Evidence 를 다시 확인하면서 필요한 경우 단계를 병합하거나 롤백해 오류를 교정한다. 최종적으로 재구성된 논리 구조를 바탕으로 최종 답과 핵심 Evidence 를 함께 제시하여 Reasoning 의 일관성과 타당성을 높인다.

실험 결과

본 연구는 다량의 위키백과 문서가 들어있는 HotpotQA Full Wiki[2] 데이터셋에서 검증한다. 기본적인 MHQA 를 위한 Retriever 과 QA 모델은 각각 MDR[3]과 allenai/unifiedqa-t5-large[4]를 사용한다. MDR 의 성능 향상을 위해 각 하위 질문의 의미를 유지한 단일 Paraphrase 를 추가 생성하여 Retrieval 시 다양한 질의 표현을 확보하였다.

이를 Baseline 으로 잡고 우리의 플러그인 모듈을 추가하여 성능 향상을 비교한다. 플러그인 모듈은 GPT-4.1 과 GPT-4o 를 이용해 구성한다.

평가 지표는 정답과 완전히 일치한 비율인 EM(Exact Match)과 정답과 토큰 겹침 비율인 F1 을 사용한다.

1 번 실험

	EM (%)	F1 (%)
Our method	**42.77**(+18.61%p)	**55.40**(+22.25%p)
only Self-Aware QA	32.67(+8.51%p)	42.62(+9.47%p)
Baseline	24.16	33.15

표 1. Baseline 대비 성능 향상 지표

표 1 에서 보이듯, Baseline 대비 Self-Aware QA 만 적용해도 EM/F1 이 향상되었다. 그리고 Reverse Verification 까지 결합한 최종 방법은 성능이 모두 추가로 크게 향상되었다. 각 단계별 질문·답변·근거를 명시적으로 제공하여 논리 구성을 수정할 수 있게 되어, 오류 진단·교정이 된 덕분에 생긴 성능 향상으로 판단된다.

Forward CoT 만 적용한 경우, 과도한 하위 질문 분해나 불완전한 답변으로 인해 원 질문의 의도에서 벗어나는 경우가 있는데, 이를 Reverse CoT 가 잡아주기도 하였다. 이는 Reverse CoT 가 Forward CoT 의 초기 오류 전이 문제뿐 아니라, 불필요한 단계나 불완전 답변으로 인한 결론 누락까지 효과적으로 완화함을 보여준다.

한편, 전체 질의 중 약 65.3%만이 Reverse CoT 단계로 진입했으며, 나머지는 대부분 Retrieval 단계에서 필수 Evidence 를 확보하지 못해 역검증이 불가능했다. 이는 초기 데이터 검색의 중요성을 시사한다. 더 나은 성능의 Retriever 모델을 선택한다면, 성능 향상 효과 역시 커질 것으로 전망된다. 이에 따라 Reverse Verification 에서의 논리 구조 생성이 얼마나 정확한지를 비교하기 위한 실험을 진행하였다.

2 번 실험

	EM (%)	F1 (%)
DrKIT[5]	42.13	51.72
Transformer-XH-final[6]	51.60	64.07
Our method	**56.52**	**72.88**
AISO(SOTA)[7]	67.46	80.52

표 2. Reverse Verification(Reverse CoT)의 정확도와 기존 모델의 성능비교

기존 연구인 DrKIT, Transformer-XH 와 비교 시 본 방법은 경쟁력 있는 성능을 보였다. 그럼에도 SOTA 와 차이가 있는데, 이는 본 연구의 학습 없는 Inference-only 방식이 가지는 한계로도 해석 가능하다. 데이터셋에 최적화가 되지 않았기 때문이다. 하지만 Inference-only 방식이기 때문에 다른 모델과 다르게 플러그인 모듈로써 사용 가능하기도 하다. 이는 성능 좋은 모델에 결합할수록, 자세하고 좋은 논리 구조를 얻을 가능성을 시사한다.

Ⅲ. 결론

본 연구는 Multi-Hop QA 에서 제기된 두 가지 핵심 한계, 즉 추론 경로의 불투명성과 Forward-only 구조의 오류 전이를 해결하기 위해, Self-Aware QA 와 Reverse Verification 을 결합한 추가 학습 없는(Training-free) 양방향 CoT 파이프라인을 제안하였다.

HotpotQA 실험을 통해 단순 Retrieval+ QA 대비 유의미한 성능 향상을 입증했으며, 일부 Fine-tuning 된 MHQA 모델에 근접하거나 우세한 결과를 달성함으로써 학습 없이도 경쟁력 있는 수준의 성능을 보여주었다.

제안한 모듈은 각 단계별 질문·답변·근거를 명시적으로 기록하여 추론 경로를 그대로 드러내고, 역방향 검증으로 오류를 교정함으로써 설명 가능성과 신뢰성을 동시에 강화한다.

또한 플러그인 형태로 설계되어 기존 MHQA 모델은 물론 더 강력한 Retriever·QA 모델과 결합 시 더욱 정교하고 견고한 논리 구조를 제공할 수 있으며, 도메인 전이와 복합 추론이 필요한 다양한 응용 과제에도 확장될 수 있는 실용적 가치를 지닌다.

참 고 문 헌

[1] Trivedi, H. et al. *Interleaving Retrieval with Chain-of-Thought Reasoning*. ACL, 2023.

[2] Yang, Z. et al. *HotpotQA: A Dataset for Diverse, Explainable Multi-hop QA*. EMNLP, 2018.

[3] Xiong, W. et al. *Answering Complex Open-Domain Questions with Multi-Hop Dense Retrieval*. ICLR, 2021.

[4] Khashabi, D. et al. *UnifiedQA: Crossing Format Boundaries With a Single QA System*. arXiv:2005.00700, 2020.

[5] Dhingra, B. et al. *Differentiable Reasoning over a Virtual Knowledge Base*. ICLR, 2020.

[6] Zhao, C. et al. *Transformer-XH: Multi-Evidence Reasoning with eXtra Hop Attention*. ICLR, 2020.

[7] Zhu, Y. et al. *Adaptive Information Seeking for Open-Domain QA*. EMNLP, 2021.