

# Least-to-Most 프롬프트 기법을 활용한 표 해석 연구

이정우, 이은수, 김지희  
동국대학교

jwlee0519@dgu.ac.kr, dmstn7432@dgu.ac.kr, jihie.kim@dgu.edu

## A Study on Table Analysis Using Least-to-Most Prompting

Jeongwoo Lee, Eunsoo Lee, Jihie Kim  
Dongguk Univ.

### 요약

본 논문은 Table Question Answering(TQA)에서의 복잡한 질문을 해석하기 위해 Least-to-Most 프롬프트 기법을 적용하는 방식을 제안한다. 해당 기법은 질문을 의미 단위의 세부 질문들로 분해한 뒤, 각 단계를 순차적으로 추론함으로써 모델이 복잡한 테이블 질문을 보다 효과적으로 처리할 수 있도록 유도한다. 모든 실험은 공개된 TQA 벤치마크인 WikiTableQuestions에서 수행되었으며, 추론에는 오픈소스 LLM인 LLaMA 3.3 70B 모델을 활용하였다. 실험 결과, 본 접근법은 다양한 테이블 구조와 질문 유형에 대해 높은 정확도를 보이며 WikiTableQuestion 벤치마크에서 71.18%의 우수한 성능을 달성하였다.

### I. 서론

Table Question Answering(TQA)은 구조화된 표를 기반으로 복잡한 자연어 질문에 응답하는 과제로, 대형 언어 모델(LLM)의 발전과 함께 다양한 프롬프트 기법을 활용한 해석 방식이 제안되고 있다. 대표적으로 Chain-of-Table [1]은 테이블에 조건을 순차적으로 적용해 중간 테이블을 생성하고 최종 답을 도출하는 방식으로, 구조는 명확하지만 초기 조건 분해가 부정확할 경우 추론 오류로 이어질 수 있다.

ReAcTable [2]은 단일 질문에 대해 추론(Reasoning)과 실행(Act)을 번갈아 수행하는 구조로, 모델이 자연어로 중간 판단을 서술한 후 SQL 등 실행 가능한 코드로 행동을 생성한다. 이 방식은 구조화된 문제 해결에 유리하지만, 별도의 코드 실행 환경이 필요하고 모델이 생성한 코드가 실행 오류나 구문 오류를 유발할 수 있다는 제약이 존재한다.

본 연구는 이러한 한계를 보완하고자 질문을 의미 단위의 세부 질문으로 분해하여 순차적으로 추론하는 Least-to-Most prompting 기법[3]을 Table QA에 적용하였다. 자연어 기반의 단계적 추론 과정으로 복잡한 질문을 효과적으로 처리할 수 있었으며, 특히 WikiTableQuestions 벤치마크[4]에서 높은 성능을 기록해 본 접근의 효과를 입증하였다.

### II. 본론

#### 2.1 방법론

본 연구는 Table Question Answering(TQA)에서 복잡한 자연어 질문을 효과적으로 해결하기 위해, Least-to-Most prompting 기법을 적용한 단계적 추론 방식을 제안한다. 이 방식은 하나의 질문을 의미

단위의 세부 질문으로 분해하고, 각 단계를 언어 모델이 순차적으로 해석하도록 유도함으로써 표 해석의 정확도와 추론 안정성을 동시에 확보하는 것을 목표로 한다.

##### 2.1.1 입력 형식의 전처리

기존 TQA 방식과 달리, 본 방법은 테이블을 [HEAD], [ROW]와 같은 명시적 형태로 모델에게 전달한다. 이를 통해 모델은 열 이름과 각 행의 셀 값을 명확히 구분할 수 있으며, 프롬프트 차원에서 테이블의 구조적 정보를 직접적으로 반영할 수 있도록 돕는다. 원본 테이블 예시는 표 1, 입력 형식의 예시는 그림 1에 제시되어 있다.

| Year | Player                | Team      | Games Played | Points Per Game |
|------|-----------------------|-----------|--------------|-----------------|
| 2013 | James Harden          | Rockets   | 78           | 25.4            |
| 2014 | Stephen Curry         | Warriors  | 80           | 23.5            |
| 2015 | LeBron James          | Cavaliers | 76           | 25.3            |
| 2016 | Kawhi Leonard         | Spurs     | 74           | 21.2            |
| 2017 | Kevin Durant          | Warriors  | 70           | 25.1            |
| 2018 | Giannis Antetokounmpo | Bucks     | 82           | 26.9            |
| 2019 | Kyrie Irving          | Celtics   | 67           | 23.8            |
| 2020 | Luka Doncic           | Mavericks | 61           | 28.8            |
| 2021 | Joel Embiid           | 76ers     | 68           | 29.6            |
| 2022 | Jayson Tatum          | Celtics   | 74           | 30.1            |

표 1. 예시 원본 테이블

```
[HEAD]: Year | Player | Team | Games Played | Points Per Game
[ROW] 1: 2013 | James Harden | Rockets | 78 | 25.4
[ROW] 2: 2014 | Stephen Curry | Warriors | 80 | 23.5
[ROW] 3: 2015 | LeBron James | Cavaliers | 76 | 25.3
[ROW] 4: 2016 | Kawhi Leonard | Spurs | 74 | 21.2
[ROW] 5: 2017 | Kevin Durant | Warriors | 70 | 25.1
[ROW] 6: 2018 | Giannis Antetokounmpo | Bucks | 82 | 26.9
[ROW] 7: 2019 | Kyrie Irving | Celtics | 67 | 23.8
[ROW] 8: 2020 | Luka Doncic | Mavericks | 61 | 28.8
[ROW] 9: 2021 | Joel Embiid | 76ers | 68 | 29.6
[ROW] 10: 2022 | Jayson Tatum | Celtics | 74 | 30.1
```

그림 1. 테이블 프롬프트 구조

### 2.1.2 Least-to-Most 형식의 질문 분해

Table QA 에서 주어지는 복잡한 질문은 LLM 이 단일 단계에서 정확하게 처리하기 어려운 경우가 많다. 이에 본 연구는 질문을 테이블 속성 및 구조에 기반한 의미 단위의 하위 질문(subquestions)으로 분해하고, 각 질문에 대해 순차적인 추론 절차를 적용하는 방식으로 문제에 접근한다. 이러한 분해 방식은 사람이 사전에 설계한 Least-to-Most 프롬프트 템플릿 예시를 기반으로 few-shot 형태로 프롬프트에 포함되며, 모델은 이를 참조하여 새로운 테이블 질문에 대해서도 유사한 방식으로 추론을 수행할 수 있다. 질문 분해 구조는 그림 2 에 제시하였다.

|   |
|---|
| Question: Which player had the highest points per game?                                       |
| Let's think step-by-step.   |
| 1. What is the highest value in the 'Points Per Game' column?<br>→ The highest value is 30.1. |
| 2. Which player has that value?<br>→ Jayson Tatum has 30.1 points per game.                   |
| Answer: Jayson Tatum  |

그림 2. Least-to-Most 질문 분해 구조

Least-to-Most 방식은 Chain-of-Thought 나 ReAct 와 달리, 하나의 질문을 보다 단순한 하위 질문으로 나누고 이를 순차적으로 해결하는 과정을 통해 모델이 단계적으로 추론을 수행할 수 있다는 강점이 있다. 이와 같은 방식은 복잡한 테이블 기반 질의의 처리 난이도를 낮추고 모델의 표 해석 정확도와 추론 안정성을 높이는 데 기여할 수 있다.

### 2.2 실험 및 평가 결과

본 연구는 Meta 의 오픈소스 언어 모델 LLaMA 3.3 70B 를 기반으로, 기존 프롬프트 기반 접근법들과 제안하는 Least-to-Most prompting 기법을 비교 분석하였다. LLaMA 3[5]은 2024 년 4 월 Meta 에 의해 공개된 모델로, 다양한 자연어 처리 작업에서 GPT-4 에 준하는 성능을 보이며 주목받았다. 본 연구에 사용된 LLaMA 3.3 은 이 모델 계열의 후속 고성능 버전으로, 동일 파라미터 크기(70B)에서 추론 속도와 정확도가 향상된 개선 모델이다.

실험에는 대표적인 Table QA 벤치마크인 WikiTableQuestions(WTQ) 데이터셋을 사용하였다. WikiTableQuestions 는 위키피디아에서 수집된 실제 테이블과 자연어 질문 쌍으로 구성되어 있으며, 정답만 제공되고 중간 추론 과정은 포함되어 있지 않다. 따라서 모델이 복잡한 질문에 대해 스스로 추론을 진행하며 정답을 도출해야 하므로 end-to-end 추론 능력을 평가하기에 적합하다.

본 연구는 프롬프트 구조 방식에 초점을 둔 방법론 간의 비교를 위해 ReAcTable 과 Chain-of-Table 을 비교군으로 선정하고, 동일한 모델과 환경에서 성능을 비교 평가하였다.

그 결과, 제안한 Least-to-Most prompting 기법은 71.18%의 정확도를 기록하며 기존 방법들을 상회하는 성능을 보였다. 실험 결과는 표 2 에 명시되어 있다.

| Method                               | Accuracy(%) |
|--------------------------------------|-------------|
| Ours<br>(Least-to-Most)              | 71.18       |
| ReAcTable<br>(ReAct)                 | 62.55       |
| Chain-of-Table<br>(Chain of Thought) | 52.61       |

표 2. 성능 비교 결과

### III 결론

본 논문에서는 Table Question Answering(TQA) 과제에서 복잡한 질문을 단계적으로 해석하기 위한 방법으로 Least-to-Most prompting 기법을 적용하였다. 이 방식은 테이블을 명시적인 구조로 입력하고, 질문을 의미 단위의 세부 질문들로 분해한 뒤, 각 단계에 순차적으로 추론을 수행함으로써 외부 코드나 도구 없이도 높은 정확도를 달성할 수 있도록 설계되었다.

특히, LLaMA3.3 70B 기반으로 WikiTableQuestions 벤치마크에서 기존 프롬프트 기반 접근법들과 비교 실험을 통해 71.18%의 우수한 결과를 달성하였다. 이는 오픈소스 환경에서도 높은 성능을 보이며, 복잡한 테이블 질의 처리에 실용성과 확장성을 갖춘 접근법임을 시사한다.

### ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획재정부의 대학ICT 연구센터육성사업(ITP-2025-RS-2020-II201789) 과 인공지능융합혁신산재양성사업(ITP-2025-RS-2023-00254592) 의 연구결과로 수행되었으며, 한국과학기술정보연구원에서 운영하는 뉴런 슈퍼컴퓨터 계산자원을 이용해 수행되었음.

### 참고 문헌

- [1] Z. Wang, H. Zhang, C.-L. Li, J. M. Eisenschlos, V. Perot, Z. Wang, L. Miculicich, Y. Fujii, J. Shang, C.-Y. Lee, and T. Pfister, "Chain-of-Table: Evolving tables in the reasoning chain for table understanding," *arXiv preprint arXiv:2401.04398*, Jan. 2024.
- [2] Y. Zhang, J. Henkel, A. Floratou, J. Cahoon, S. Deep, and J. M. Patel, "ReAcTable: Enhancing ReAct for table question answering," *arXiv preprint arXiv:2310.08157*, Oct. 2023.
- [3] D. Zhou, N. Schärli, L. Hou, J. Wei, N. Scales, X. Wang, D. Schuurmans, C. Cui, O. Bousquet, Q. Le, and E. Chi, "Least-to-most prompting enables complex reasoning in large language models," *arXiv preprint arXiv:2205.10625*, May 2022.
- [4] P. Pasupat and P. Liang, "Compositional semantic parsing on semi-structured tables," in *Proc. 53rd Annu. Meeting Assoc. Comput. Linguistics (ACL)*, Jul. 2015, pp. 1470-1480.
- [5] Llama Team, "The Llama 3 herd of models," *arXiv preprint arXiv:2407.21783*, Jul. 2024.