



Weaver: Interweaving SQL and LLM for Table Reasoning

Rohit Khoja^{1*}, Devanshu Gupta^{1*}, Yanjie Fu¹, Dan Roth², Vivek Gupta^{1†}

¹ Arizona State University, ² University of Pennsylvania







Why is Table QA still Challenging?

- Tables contain structured (numbers, fields) and unstructured (long text/images) data
- **SQL**: great for logic, fails at semantic inference
- LLMs: handle semantics, struggle at structured logic

Example: "Which country had most competitors?" in the Figure,

 $SQL fails \rightarrow LLM infers nationality$

Existing SQL–LLM integration is rigid or shallow, lack adaptability to complex queries

Weaver Framework

LLM-generated dynamic execution plan

Weaver generates a **flexible plan** that adapts to query complexity, then executes through:

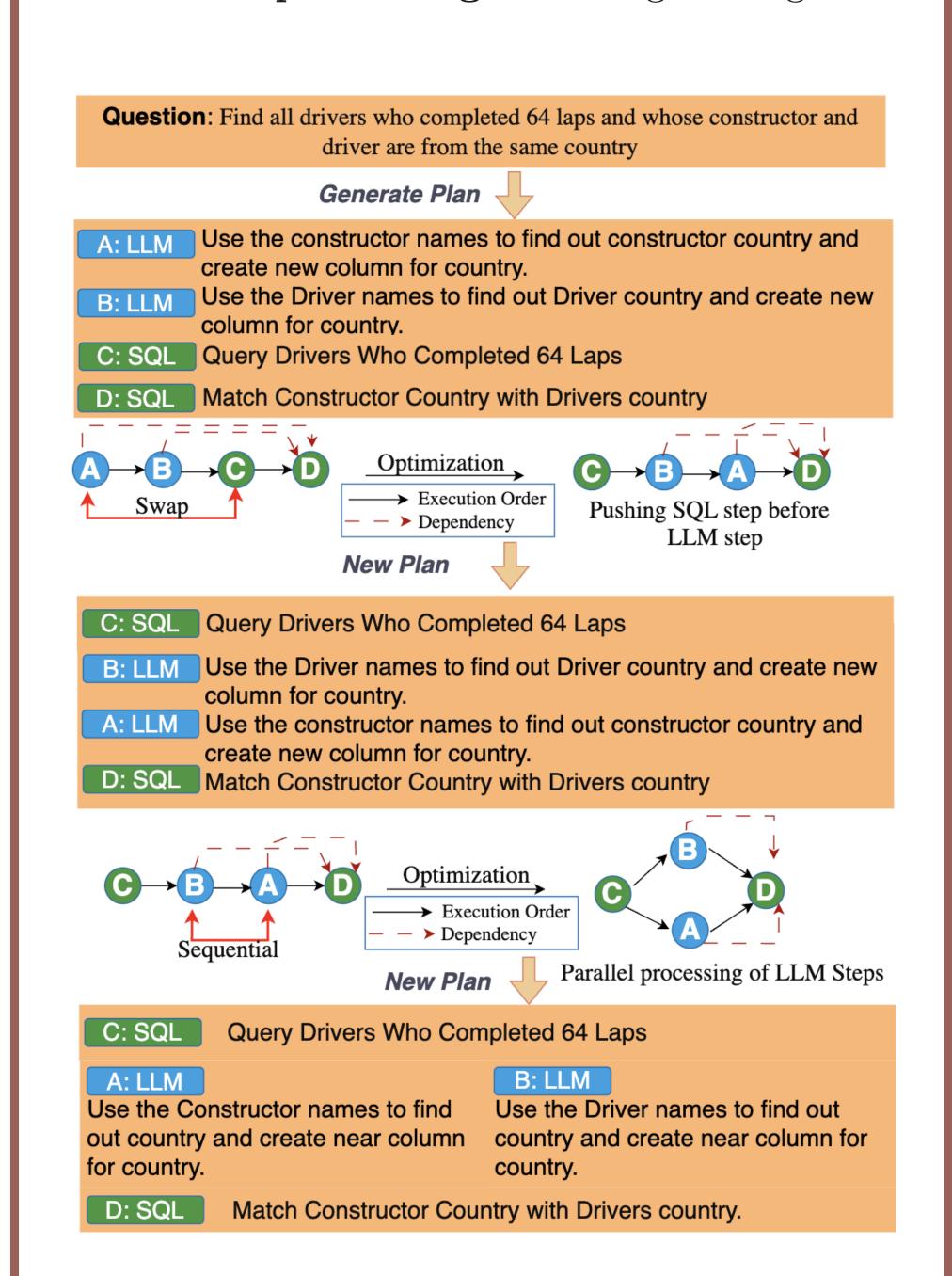
- 1. \mathbf{SQL} step \rightarrow Structured operations
- 2. LLM step \rightarrow Semantic reasoning
- 3. Verification \rightarrow Ensures correctness

Back-and-forth: $SQL \leftrightarrow LLM \leftrightarrow SQL \leftrightarrow LLM$

Planning Optimization

Strategies:

- SQL reordering: Push SQL steps earlier
- Parallelization: Independent LLM calls in parallel
- Batch processing: Batching for large data



23% reduction in steps with 1% accuracy loss

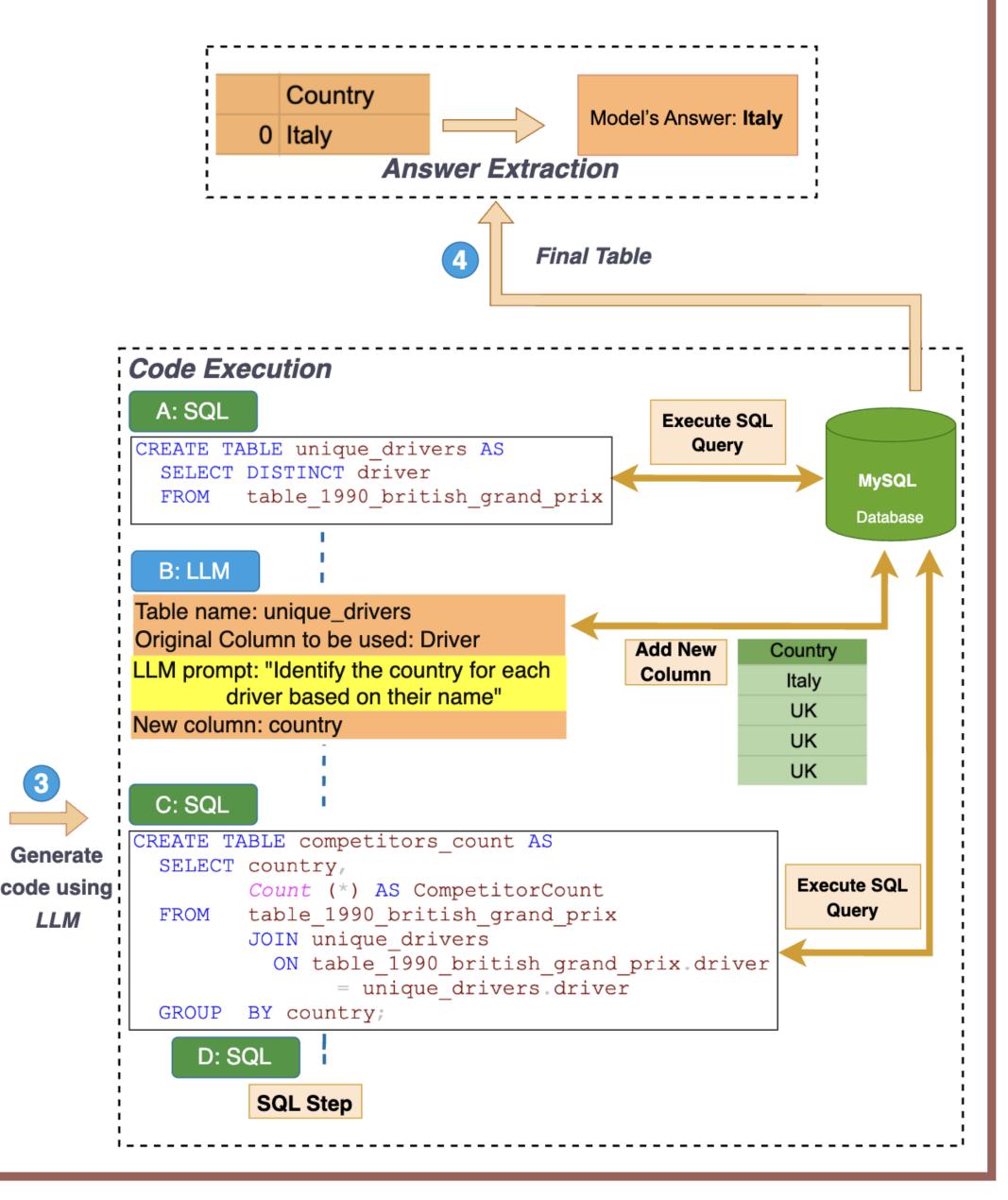
Weaver Architecture

Four-phase modular pipeline:

- 1. Preprocessing: Parse table/question, extract metadata
- 2. Planning: LLM generates step plan, determines SQL vs. LLM operations

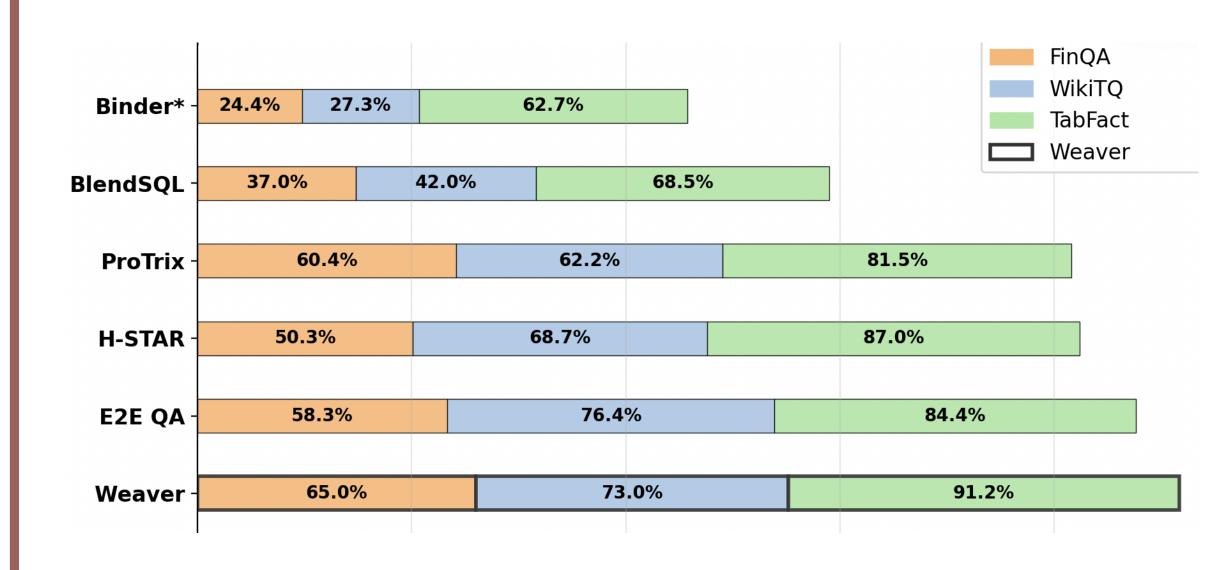
				<i>Table</i> 1990 E	· <i>QA</i> British Grand Prix					
	Rank Driver					TimeRetired	d			
	1 Alain Pro		Prost	rost Ferrari		64		1:18:31		
	2	2 Thierry Boutse		Williams-Renault		64		39.092		
	3 Ayrton Senna		Senna	McLaren-Honda		64		43.088		
	4 Éric Bernard		ernard	Lola-Lamborghini		64		401:03:00		
Question: which country had the most competitors? Gold Answ					d Answer: Ital	y				
Pre-processed table Relevant Columns: {Driver}										
Со	lumn Na	me Data Type	Formatting N	leeded	Co	olumn De	script	ion		
*		* Cardinal	None		Position of the driver at the end of the race					
Rank_1 Driver		String	None		(e.g., 1, 2, 3, Ret, DNS, DNQ, DNPQ) Name of the driver					
Constructor		_	None		Name of the constructor (team) the driver is racing for					
			Convert to delta form		Time taken to complete the race or reason for retirement (if applicable)					
-		Genera	te Plan usi	na LL	м 👢 2 * 1	Rename	d SQI	L reserved keyv	vor	
D	lannin									
		Evtr	act the uniqu	ıe driv	ers to identify the	country	,			
	A:	SOL	ach driver .		ore to identify the					
	R·	Use	the driver n	ames	to determine the c	ountry	of e	ach driver		
and create a new column for country.										
C: SQL Count the number of competitors from each country by joining previous table with the original.										
D: SQL Find the country with the highest number of competitors.										
		Verify	/ Plan usino	g <i>LLM</i>	1					
]	New Plan					
	D: SQL Find the country with the highest number of competitors by selecting									

- 3. Code Execution: Execute SQL queries & LLM inference
- 4. Answer Extraction: Combine results, generate final answer



Performance Results & Efficacy

Weaver Outperforms SOTA:



Multimodal Table QA:

Dataset	Modalities	Accuracy Gain
MMTabQA	Text + Images	+6.6%
FinQA-MM	Tables + Passages	+17.3%
OTT-QA-MM	Tables + Passages	+2.9%

Key: +5-10% accuracy gain

Efficiency:

• Avg 6 API calls/query

Method	Calls
Binder	50
H-STAR	8
Weaver	5.5

Efficacy:

• 28.1% gain on large tables

Method	Acc.
H-STAR	35.9%
ProTrix	37.5%
Weaver	65.6%

Interpretability:

Transparent, debuggable

Conclusion

Dynamic SQL-LLM weaving enables accurate, interpretable, and efficient Table QA

Key Takeaways:

- Modular, interpretable pipeline for hybrid table reasoning
- 5–10% accuracy gain over state-of-the-art methods
- Multimodal support (text, image, table)
- Flexible planning adapts to query complexity
- Efficient execution with fewer API calls (avg. 6 calls/query)

Impact: Weaver bridges the gap between structured data processing and semantic understanding, enabling more robust and adaptable table question answering systems for real-world applications.



Scan for paper