

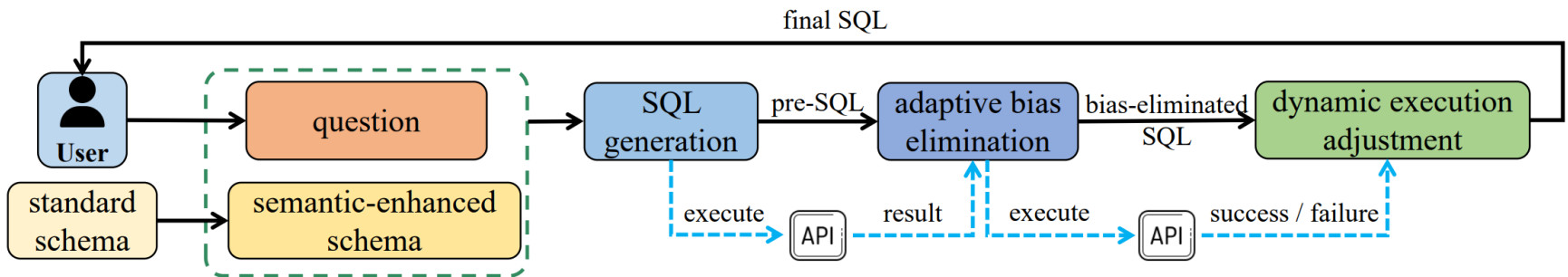
SEA-SQL: Semantic-Enhanced Text-to-SQL with Adaptive Refinement

Chaofan LI, Yingxia SHAO, Yawen LI, Zheng LIU

Frontiers of Computer Science, DOI: [10.1007/s11704-025-41136-3](https://doi.org/10.1007/s11704-025-41136-3)

Problems & Ideas

- Problems of current Text-to-SQL approaches:
 - **Inherent model bias:** When generating SQL queries, models may exhibit biases due to patterns present in their training data.
 - **Unexecutable SQL:** Although SQL is generated by the LLM, they lack the capability to verify the executability of the final SQL query.
 - **Expensive inference cost:** Recent works focus on leveraging GPT-4 for SQL generation. However, it comes at a significant expense.
- Ideas: SEA-SQL incorporates a semantic-enhanced schema to enrich database information and refine SQL queries through adaptive bias elimination and dynamic execution adjustment.



An overview of the SEA-SQL framework: The SQL query is generated using the semantic-enhanced schema. Then, through adaptive bias elimination, the inherent biases caused by the LLM are eliminated. Finally, with dynamic execution adjustment, the SQL query is ensured to execute successfully to return the expected results to the user.

Main Experimental Results

- Leverging GPT-3.5, the results of SEA-SQL remain competitive with GPT-4 based methods, demonstrating SOTA performance based on GPT-3.5 and maintaining competitiveness relative to GPT-4.

Method	Model	Zero-Shot	Few-Shot	Fine-Tune	dev
GPT-4 [17]	GPT-4	✓			72.9
DIN-SQL [17]	GPT-4		✓		82.8
DAIL-SQL [18]	GPT-4		✓		83.5
MAC-SQL [20]	GPT-4		✓		86.8

C3 [16]	GPT-3.5	✓			81.8
DIN-SQL [17]	GPT-3.5		✓		74.9
DAIL-SQL [18]	GPT-3.5		✓		78.2
MAC-SQL [20]	GPT-3.5		✓		80.6

T5-3B [11]	T5-3B			✓	74.4
T5-3B + PICARD [11]	T5-3B			✓	79.9
Supervised Fine-Tuning [18]	Llama2-7b			✓	66.7
Supervised Fine-Tuning [18]	Llama2-13b			✓	67.0
SEA-SQL(ours)	GPT-3.5 + Mistral-7B	✓		✓	83.6

Method	Model	Zero-Shot	Few-Shot	Fine-Tune	dev		test	
					EX	VES	EX	VES
GPT-4	GPT-4				46.35	49.77	54.89	60.77
DIN-SQL	GPT-4		✓		50.72	58.79	55.90	59.44
DAIL-SQL	GPT-4		✓		54.76	56.08	<u>57.41</u>	<u>61.95</u>
MAC-SQL	GPT-4		✓		57.56	58.76	59.59	67.68

DIN-SQL	GPT-3.5		✓		44.92	46.73	-	-
DAIL-SQL	GPT-3.5		✓		44.65	46.99	-	-
MAC-SQL	GPT-3.5		✓		50.56	61.25	-	-
SEA-SQL(ours)	GPT-3.5 + Mistral-7B	✓		✓	<u>56.13</u>	<u>60.83</u>	53.66	57.00

SEA-SQL and baselines on Spider development set (EX result) (left) and BIRD (right), where the best outcomes are highlighted in bold, while the second highest results are underlined.

- The cost of SEA-SQL to generate a single SQL query is minimal, ranging from 9% to 58% compared to other GPT-3.5 based methods.
- SEA-SQL leverages GPT-3.5 and requires only 0.9% to 5.3% of the expenses associated with GPT-4 for generating a single SQL query, yet it achieves competitive results.

Method	Model	Pre (\$)	Generate / Query (\$)	All (\$)	GPU-Time / Query (s)	API-Time / Query (s)	Full-Time / Query (s)
GPT-4	GPT-4	-	0.0351	54	-	1.17	1.17
DIN-SQL	GPT-4	-	0.7867	1207	-	13.86	13.86
DAIL-SQL	GPT-4	-	0.1232	189	-	12.29	12.29
MAC-SQL	GPT-4	-	0.2151	330	-	15.84	15.84

DIN-SQL	GPT-3.5	-	0.0776	119	-	13.86	13.86
DAIL-SQL	GPT-3.5	-	<u>0.0113</u>	17	-	12.29	12.29
MAC-SQL	GPT-3.5	-	0.0209	<u>32</u>	-	15.84	15.84
SEA-SQL	GPT-3.5	44.78	0.0065	55	1.09	7.37	8.46

The cost of different methods on the BIRD development set.