Cross-Domain SQL Query Generation for Context-Dependent Questions

Yale

Rui Zhang

Yale University

LILY Lab

Cross-Domain Context-Dependent Text-to-SQL

Semantic Parsing is the task of mapping natural language sentences into formal queries such as logical forms and SQL queries which has been studied for decades. Traditional benchmark semantic parsing datasets such as ATIS, Geoquery and WikiSQL only consider individual single-sentence sentences.
The model based on these data sets are thus constrained to performing semantic parsing over individual questions.

However, in a real-world application, users often access information with a system in a multi-turn interaction with a sequence of related questions. In this case, the meaning of questions strongly depends on the interaction history. Therefore, the ability to reason about the meaning of user utterances based on the context is crucial for the model to generate correct queries.

In this paper, we propose a novel model for mapping multi-turn context-based user questions into SQL queries for a range of domains. We evaluate our model on several datasets including (1) the ATIS dataset on flight-booking requests and (2) the SParC , a dataset for cross-domain semantic parsing in context, consisting of coherent question sequences annotated with SQL queries over 200 databases in 138 domains. Furthermore, we also show that our model can achieve state-of-the-art results on Spider, a cross-domain context-independent text-to-SQL dataset.

Our Model

Our model follows a sequence-to-sequence encoder-decoder framework with attention mechanisms. The encoder for natural language utterance and table schema makes use of bi-LSTM, self-attention, and cross-serial attention as illustrated in Figure 2. We also add a Table-aware BERT-encoder as shown in Figure 1.

To incorporate context within an interaction, it maintains and updates a discourse state encoder which encodes the history of previous utterances and is updated after each turn over the entire interaction, in addition to an utterance-level encoder for the current utterance. Also, positional encoding is used to take the position of each utterance relative to the current one into account.

During decoding, at each generation step, the decoder selects between a SQL keyword or a column and table name from the corresponding database of the current question.

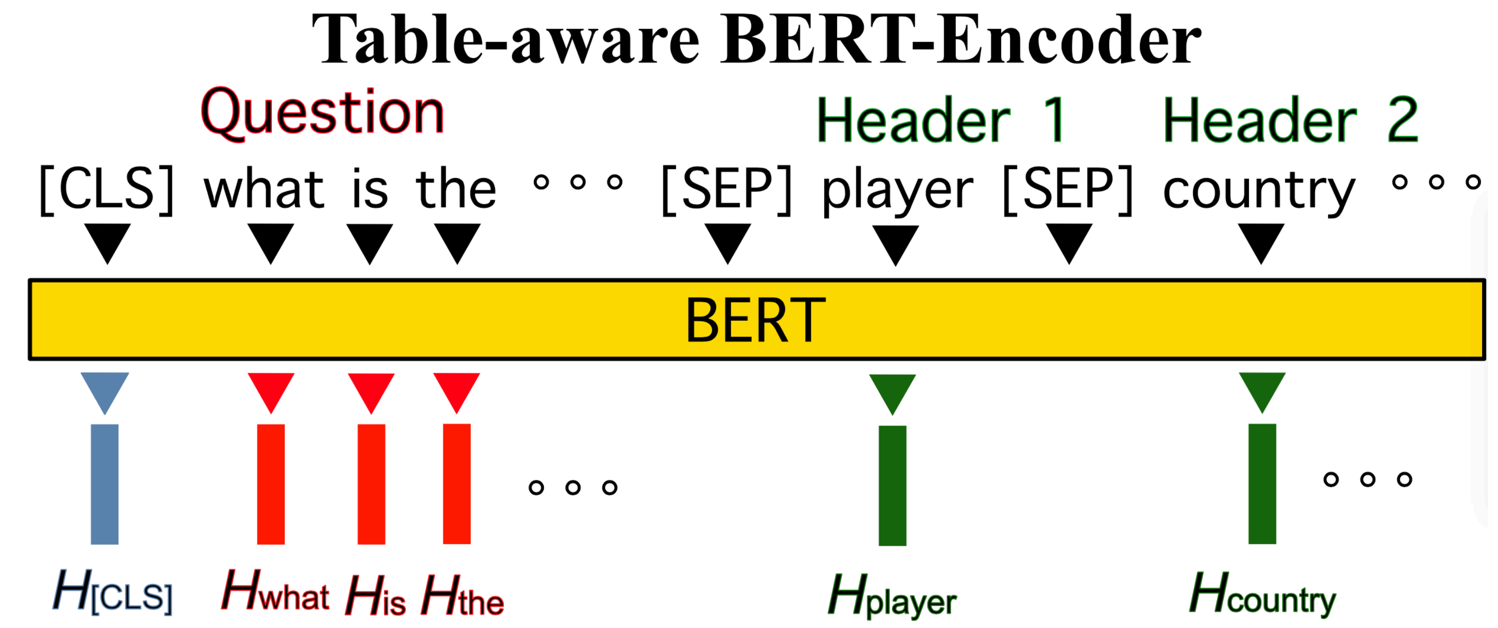

Figure 1. Table-aware BERT-Encoder (Hwang et al., 2019). We concatenate the utterance and the column names in table schema into a single sequence separated by [SEP] tokens as the input to the BERT model. The last layer of BERT model is used to the following encoder in Figure 2.

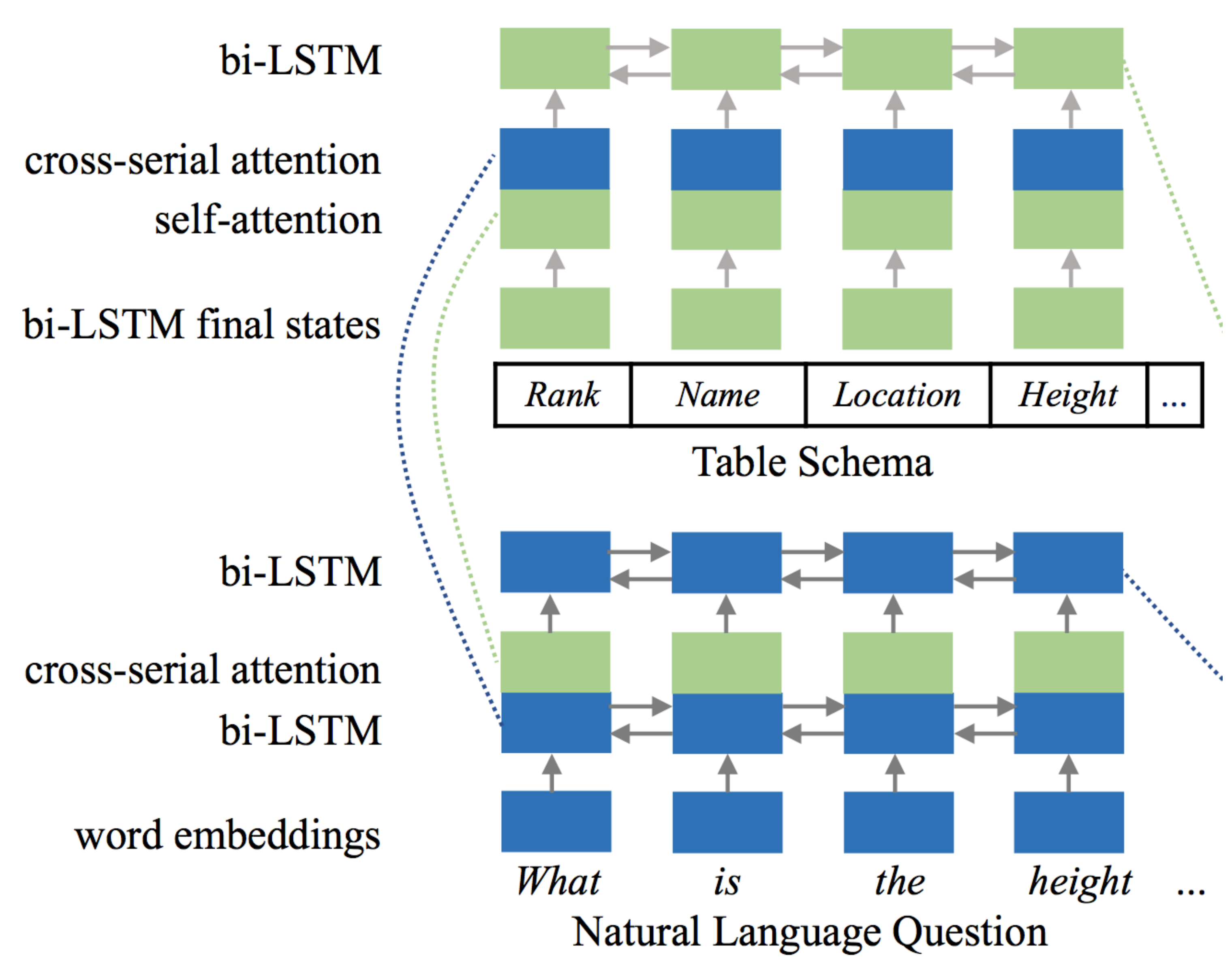

Figure 2. Encoder for Natural Language Utterance and Table Schema (Shi et al., 2018).

Datasets and Experimental Results

	Interaction #	Question #	Avg. turn #	Database #	Table #	Avg. Q token #	Q Vocab #
ATIS	1658	11,653	7.0	1	27	10.2	1582
SParC	4298	12,726	3.0	200	1020	8.1	3794

Table 1: Dataset Statistics.

	WHERE	AGG	GROUP	ORDER	HAVING	SET	JOIN	Nested
ATIS	100	16.6	0.3	0	0	0	99.9	99.9
SParC	42.8	39.8	20.1	17.0	4.7	3.5	35.5	5.7

Table 2: Distribution of SQL components.

	Test Set					Dev Set
	Easy	Medium	Hard	Extra Hard	All	All
SQLNet (Xu et al., 2017)	26.2	12.6	6.6	1.3	12.4	10.9
SyntaxSQLNet (Yu et al., 2018b)	38.6	17.6	16.3	4.9	19.7	18.9
SyntaxSQLNet + data augmentation	48.0	27.0	24.3	4.6	27.2	24.8
Lee (2019)	–	–	–	–	24.3	28.5
Ours	56.9	30.4	24.5	9.4	31.3	34.2
Ours + Table-aware BERT Encoder	80.6	58.7	49.2	23.7	55.4	59.8

Table 3: Spider results on dev set and test set.

	Dev Set			Test Set		
	Query	Relaxed	Strict	Query	Relaxed	Strict
FULL (Suhr et al., 2018)	$37.5_{\pm 0.9}$	$63.0_{\pm 0.7}$	$62.5_{\pm 0.9}$	$43.6_{\pm 1.0}$	$69.3_{\pm 0.8}$	$69.2_{\pm 0.8}$
FULL (Our Replication)	38.8	63.3	62.8	44.6	68.3	68.2

Table 4: ATIS results on dev set and test set. We show mean and standard deviation.

Model	Question Match		Interaction Match	
	Dev	Test	Dev	Test
CD-Seq2Seq	17.1	18.3	6.7	6.4
SyntaxSQL-con	18.5	20.2	4.3	5.2
SyntaxSQL-inp	15.2	16.9	0.7	1.1
Ours	31.4	32.2	12.4	10.0
Ours + Table-aware BERT Encoder	43.8	47.0	22.1	19.1

Table 5: SParC results on dev set and test set.