Photon: A Robust Cross-Domain Text-to-SQL System


The Chinese University of Hong Kong

salesforce research

ACL 2020 - System Demonstration
Motivation

End users access information system everyday, everywhere...

Live demo: http://naturalsql.com/
Motivation

Scenario 1: Everyone is a programming master

Live demo: http://naturalsql.com/
Motivation

Scenario 1: Everyone is a programming master

SELECT Quantity FROM Product WHERE Name = "Hoverboard x10"

SELECT Arriving_Time FROM Flights WHERE Flight_Number = "CZ327"

SELECT T2.name FROM Instructor AS T1 JOIN Department AS T2 ON T1.Department_ID = T2.ID GROUP BY T1.Department_ID HAVING AVG(T1.Rating) > (SELECT AVG(Rating) FROM Instructor)

SELECT Name FROM Country WHERE Continent = "Asia" ORDER BY LifeExpectancy LIMIT 1

Live demo: http://naturalsql.com/
Motivation

Scenario 2: Everyone simply talks to their information system

Live demo: http://naturalsql.com/
Motivation

**Scenario 2:** Everyone simply talks to their information system

- How many “Hoverboard x10” are left in stock?
- Give me the arriving time of “CZ327”.
- Which departments have instructors in general rated above average?
- Show Asian countries ordered by life expectancy.

Desiderata

Accurately map NL input to executable SQL queries

Live demo: http://naturalsql.com/
Desiderata

- Accurately map NL input to executable SQL queries
- Work across different databases
- Robustness - “don’t know” is better than mistakes
- Support user interaction

Photon: A Robust Cross-Domain Text-to-SQL System

Photon: A Robust Cross-Domain Text-to-SQL System

A SOTA neural text-to-SQL parser

A novel confusion detection approach

Template-based response generation for user interaction

Live demo: http://naturalsql.com/
Interaction Flow

Response Template

<table>
<thead>
<tr>
<th>Response</th>
<th>Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONFIRM_RESULT</td>
<td>“SQL: {PRED_SQL}: {NL RESPONSE}”</td>
</tr>
<tr>
<td>CONFIRM_CORRECTION</td>
<td>“Sorry, {CONF_TOKENS} is confusing in our scenario, do you mean {CORR_TOKENS}?”</td>
</tr>
<tr>
<td>NEED_REPHRASE</td>
<td>“Sorry, it is a confusing question for me, please rephrase your question and ask again.”</td>
</tr>
<tr>
<td>INVALID_QUERY</td>
<td>“Sorry, it is an invalidate query, please check the table names and associated fields of interest.”</td>
</tr>
</tbody>
</table>
Interaction Flow

Confusion Detection →

Text-to-SQL Parsing →

Response Generation

Response Template

<table>
<thead>
<tr>
<th>State</th>
<th>Response Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONFIRM_RESULT</td>
<td>“SQL: {PRED_SQL}. {NL RESPONSE}”</td>
</tr>
<tr>
<td>CONFIRM_CORRECTION</td>
<td>“Sorry, {CONF_TOKENS} is confusing in our scenario, do you mean {CORR_TOKENS}?”</td>
</tr>
<tr>
<td>NEED_REPHRASE</td>
<td>“Sorry, it is a confusing question for me, please rephrase your question and ask again.”</td>
</tr>
<tr>
<td>INVALID_QUERY</td>
<td>“Sorry, it is an invalid query, please check the table names and associated fields of interest.”</td>
</tr>
</tbody>
</table>
Interaction Flow

Confusion Detection → Text-to-SQL Parsing → Response Generation

Response Template

- **CONFIRM_RESULT**: “SQL: {PRED_SQL}. {NL_RESPONSE}”
- **CONFIRM_CORRECTION**: “Sorry, {CONF_TOKENS} is confusing in our scenario, do you mean {CORR_TOKENS}?”
- **NEED_REPHRASE**: “Sorry, it is a confusing question for me, please rephrase your question and ask again.”
- **INVALID_QUERY**: “Sorry, it is an invalidate query, please check the table names and associated fields of interest.”

Text-to-SQL Semantic Parsing

Spider Dataset (Yu et al. 2018)

Expert-annotated, cross-domain, complex text-to-SQL dataset

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th>Dev</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td># DBs</td>
<td>146</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td># Examples</td>
<td>8,659</td>
<td>1,034</td>
<td>2,147</td>
</tr>
</tbody>
</table>

Question: What are the name and budget of the departments with average instructor salary above the overall average?

SQL:

```sql
SELECT T2.name, T2.budget
FROM Instructor AS T1 JOIN Department AS T2
ON T1.Department_ID = T2.ID
GROUP BY T1.Department_ID
HAVING AVG(T1.salary) >
(SELECT AVG(Salary) FROM Instructor)
```
Text-to-SQL Semantic Parsing

Serialize DB schema

<table>
<thead>
<tr>
<th>T</th>
<th>Instructor</th>
<th>C</th>
<th>ID</th>
<th>C</th>
<th>Name</th>
<th>C</th>
<th>Department_ID</th>
<th>C</th>
<th>Salary</th>
<th>C</th>
<th>...</th>
<th>T</th>
<th>Department</th>
<th>C</th>
<th>ID</th>
<th>C</th>
<th>...</th>
</tr>
</thead>
</table>

Live demo: http://naturalsql.com/
**Text-to-SQL Semantic Parsing**

**Text-table joint encoding**

<table>
<thead>
<tr>
<th>CLS</th>
<th>What is cost of a hoverboard?</th>
<th>SEP</th>
<th>T</th>
<th>Product</th>
<th>C</th>
<th>...</th>
<th>T</th>
<th>Order</th>
<th>C</th>
<th>...</th>
</tr>
</thead>
</table>

Text-to-SQL Semantic Parsing

Text-table joint encoding

What is cost of a hoverboard?

Live demo: http://naturalsql.com/
Text-to-SQL Semantic Parsing

Text-table joint encoding

Bidirectional LSTM Text Encoder

Table encoding

Field encoding

Table encoding

Field encoding

Data type
Primary key
Foreign key
Output states of the Bi-LSTM

Bidirectional LSTM

BERT

CLS What is cost of a hoverboard?

SEP T Product C ... T Order C ...
Text-to-SQL Semantic Parsing

Text-table joint encoding

What is cost of a hoverboard?

Product
Order

CLS

Pointer-Generator Decoder

Cross-Entropy Loss

+Picklist Value

Live demo: http://naturalsql.com/
Text-to-SQL Semantic Parsing Evaluation

<table>
<thead>
<tr>
<th>Model</th>
<th>EM Acc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNN (Bogin et al., 2019a)</td>
<td>40.7</td>
</tr>
<tr>
<td>Global-GNN (Bogin et al., 2019b)</td>
<td>52.7</td>
</tr>
<tr>
<td>EditSQL + BERT (Zhang et al., 2019)</td>
<td>57.6</td>
</tr>
<tr>
<td>GNN+Bertrand-DR† (Kelkar et al., 2020)</td>
<td>57.9</td>
</tr>
<tr>
<td>EditSQL+Bertrand-DR† (Kelkar et al., 2020)</td>
<td>58.5</td>
</tr>
<tr>
<td>IRNet + BERT (Guo et al., 2019)</td>
<td>61.9</td>
</tr>
<tr>
<td>RYANSQL + BERT † (Choi et al., 2020)</td>
<td>66.6</td>
</tr>
<tr>
<td>PHOTON</td>
<td>63.2</td>
</tr>
</tbody>
</table>

† denotes unpublished work on arXiv.

Table 3: Experimental results on the Spider Dev set (%). EM Acc. denotes the exact set match accuracy.

Spider leaderboard (May 1st, 2020) https://yale-lily.github.io/spider
Confusion Detection

What is the total?

Show me homes with good schools

How many tourists visited all of the 10 attractions?

Hey, lovely weather

Live demo: http://naturalsql.com/
Confusion Detection

Underspecified

What is the total?

Show me homes with good schools

How many tourists visited all of the 10 attractions?

Hey, lovely weather

Live demo: http://naturalsql.com/
Confusion Detection

What is the total?

Show me homes with good schools

How many tourists visited all of the 10 attractions?

Hey, lovely weather

Live demo: http://naturalsql.com/
Confusion Detection

What is the total?

Show me homes with good schools

How many tourists visited all of the 10 attractions?

Hey, lovely weather

Out-of-scope

Live demo: http://naturalsql.com/
Confusion Detection

What is the total?

Show me homes with good schools

How many tourists visited all of the 10 attractions?

Hey, lovely weather

Not a query

Live demo: http://naturalsql.com/
Table 5: Examples of question-side and schema-side transformations for generating training data for untranslatable question detection. Let \( Q \) denote the question and \( S \) denote the schema. For each transformation, we provide two examples, i.e., \((Q_1, S_1)\) and \((Q_2, S_2)\). The italic and bold fonts highlight phrases before and after transformations.
## Confusion Detection Dataset (UTran-SQL)

### Table 5: Examples of question-side and schema-side transformations for generating training data for untranslatable question detection. Let $Q$ denote the question and $S$ denote the schema. For each transformation, we provide two examples, i.e., ($Q_1$, $S_1$) and ($Q_2$, $S_2$). The italic and bold fonts highlight phrases before and after transformations.

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Original data</th>
<th>Transformed data</th>
<th>Confusing text span</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Swap</strong></td>
<td>Q1: How many conductors are there? S1:</td>
<td></td>
<td>Conductor ID</td>
</tr>
<tr>
<td></td>
<td>Q2: What are the maximum and minimum values of area code types? S1:</td>
<td></td>
<td>Area Code</td>
</tr>
<tr>
<td><strong>Question</strong></td>
<td>Q1: How many are there? S1:</td>
<td></td>
<td>Country</td>
</tr>
<tr>
<td><strong>Schema Drop</strong></td>
<td>Q1: How much surface area do the countries in the Caribbean cover to? S1:</td>
<td></td>
<td>Name</td>
</tr>
<tr>
<td></td>
<td>Q2: Find the name and age of the visitors S2:</td>
<td></td>
<td>Visitor</td>
</tr>
</tbody>
</table>
# UTran-SQL Data Statistics

<table>
<thead>
<tr>
<th></th>
<th>Spider</th>
<th>Spider_UTran</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>Dev</td>
</tr>
<tr>
<td># Q</td>
<td>8,659</td>
<td>1,034</td>
</tr>
<tr>
<td># UTran Q</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td># Schema</td>
<td>146</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>13,392</td>
<td>1,631</td>
</tr>
<tr>
<td></td>
<td>4,733</td>
<td>597</td>
</tr>
<tr>
<td></td>
<td>918</td>
<td>112</td>
</tr>
</tbody>
</table>

Table 1: Data split of Spider and Spider\_UTran. Q represents all the questions, UTran Q represents the untranslatable questions.
# UTran-SQL Data Statistics

<table>
<thead>
<tr>
<th></th>
<th>Spider</th>
<th>Spider_UTran</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>Dev</td>
</tr>
<tr>
<td># Q</td>
<td>8,659</td>
<td>1,034</td>
</tr>
<tr>
<td># UTran Q</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td># Schema</td>
<td>146</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 1: Data split of Spider and Spider_UTran. Spider represents all the question and Spider_UTran represents the untranslatable question.

Confusion Detection Model

• **Translatability prediction**: binary classification based on [CLS] representation of the BERT text-table encoder

• **Confusion span detection**: predicting the start and end token indices
Question Rephrasing Model

Original input: How many candidates are registered in statistics?
Processed input: How many [MASK] are registered in statistics? [TABLE NAMES]

System: candidates is confusing here, do you mean students?

Table & Column Names

<table>
<thead>
<tr>
<th>Table &amp; Column Names</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>students</td>
<td>.289</td>
</tr>
<tr>
<td>teachers</td>
<td>.017</td>
</tr>
<tr>
<td>courses</td>
<td>.013</td>
</tr>
<tr>
<td>names</td>
<td>.009</td>
</tr>
<tr>
<td>student details</td>
<td>.008</td>
</tr>
</tbody>
</table>
Limitations

• We assume only one confusion span per sentence
• We assume the confusion span is a column mention
• The transformation rules can introduce errors
• Alternatives for confusion detection in text-to-SQL are worth exploring
  - Yao et al. 2019
  - Yao et al. 2020 (concurrent)
• Limited set of user actions are considered

Live demo: http://naturalsql.com/
Additional Demo Features

- Upload your own DBs for testing
- Effective DB schema visualization and data browsing
- Rate your experience and provide feedback

Live demo: http://naturalsql.com/
Related Work

- **ATIS Corpus collection**
- Hemphill et al. 1990
- Dahl et al. 1994

- **Learning logical-form based semantic parsers for NLIDBs**
- Zelle and Mooney 1996
- Popescu et al. 2003
- Zettlemoyer and Collins 2005

- **Seq2Seq-style neural semantic parsing**
- Sutskever et al. 2014
- Bahdanau et al. 2015

- **Neural networks widely adopted in NLP**
- Dong and Lapata 2016

- **WikiSQL: a large-scale, cross-domain text-to-SQL corpora**
- Zhong et al. 2017

- **TypeSQL, column attention, sketch-based, execution guided, RL, meta-learning**
- Xu et al. 2017
- Dong and Lapata 2018
- Wang et al. 2018
- Yu et al. 2018a

- **Syntax-guided, GNN, schema linking, SemQL**
- Yu et al. 2018b
- Bogan et al. 2019
- Shin et al. 2019
- Guo et al. 2019
- Wang et al. 2020

- **Spider: expert-annotated, large-scale, cross-domain, complex**
- Yu et al. 2018c

- **Table-Aware BERT Encoder, surpassed human-performance on WikiSQL**
- Devlin et al. 2018
- Hwang et al. 2019

- **LM pre-training: BERT**
- ...
Related Work

• Most state-of-the-art cross-domain, complex text-to-SQL semantic parsers are not well packaged for user test and interaction

• Most existing NLIDB systems are DB-specific or non-interactive
Live Demo: http://naturalsql.com/

Join us at the Q&A sessions

Tuesday July 7, 2020 UTC+0 17:00-17:45
Tuesday July 7, 2020 UTC+0 20:00-20:45