SpeakQL - Archaeologist

Jacob Yim

SpeakQL Summary

- Introduces SpeakQL, a speech-based system for creating SQL queries

  - Direct translation: accepts exact spoken SQL queries

  - Uses existing ASR

  - To correct ASR errors: uses SQL-specific similarity search algorithm to identify structure and database-specific voting algorithm to fill in literals

  - Additional SQL keyboard for touch modification of generated query
- SpeakQL achieves improved translation accuracy over existing natural language tools

Influenced by: Seq2SQL

- Original paper: Seq2SQL: Generating Structured Queries from Natural Language using Reinforcement Learning by Victor Zhong, Caiming Xiong, and Richard Socher

SEQ2SQL: GENERATING STRUCTURED QUERIES FROM NATURAL LANGUAGE USING REINFORCEMENT LEARNING

Victor Zhong, Caiming Xiong, Richard Socher
Salesforce Research
Palo Alto, CA
{vzhong,cxiong,rsocher}@salesforce.com

Influenced by: Seq2SQL

- Introduces Seq2SQL, a deep neural network for translating natural language into SQL queries
    - Leverages structure of SQL to prune output space of generated queries
    - Uses three smaller components corresponding to aggregation, SELECT clause, and WHERE clause
- Also contributes WikiSQL, a dataset of 80654 hand-annotated examples of natural language questions translated to SQL queries
    - Collected by crowdsourcing via Amazon Mechanical Turk
- Compared on WikiSQL with a modern semantic parser, Seq2SQL improves execution accuracy from 36% to 59%

Connection to SpeakQL

- Seq2SQL input is text-based, not spoken

    - Not optimized towards spoken input like SpeakQL

    - Large queries are more difficult for spoken systems, while many joins are more difficult for text-based systems

- Both systems leverage the structure of SQL to optimize translations

- SpeakQL is compared against Seq2SQL in evaluation

    - SpeakQL is 50% more accurate in translating spoken inputs

Influenced: Speech-to-SQL

- Original paper: Speech-to-SQL: Towards Speech-driven SQL Query Generation From Natural Language Question by Yuanfeng Song, Raymond Chi-Wing Wong, Xuefang Zhao, and Di Jiang

Speech-to-SQL: Towards Speech-driven SQL Query Generation From Natural Language Question

Yuanfeng Song[1,2], Raymond Chi-Wing Wong[1], Xuefang Zhao[2], Di Jiang[2]

[1]The Hong Kong University of Science and Technology, Hong Kong, China [2]AI Group, WeBank Co., Ltd, China

{songyf,raywong}@cse.ust.hk

Influenced: Speech-to-SQL

- Proposes SpeechSQLNet, an end-to-end neural architecture for translating human speech into SQL without an external ASR step
    - Fully bypasses a text representation for the natural language query, removing one source of error in translation (the ASR step)
    - Uses an encoder for linking natural language query with database schema
    - Also encodes SQL structure
- Achieves better performance on a set of benchmarks (WikiSQL + Spider)

Connection to SpeakQL

- Less structured input: SpeakQL is designed for exact spoken SQL statements, while SpeechSQLNet accepts unstructured ("fuzzy") queries
  - SpeechSQLNet tackles a similar but broader, more difficult problem
- SpeechSQLNet uses a similar general approach (DNNs, taking advantage of SQL structure) but bypasses ASR

References

- Seq2SQL: https://arxiv.org/pdf/1709.00103.pdf

- Speech-to-SQL: https://arxiv.org/pdf/2201.01209.pdf