

CySpider: A Neural Semantic Parsing Corpus with Baseline Models for Property Graphs

Conference Publishing

Ziyu Zhao

Authors: Ziyu Zhao, Wei Liu, Tim French & Michael Stewart

2023-11-27

Publication

Springer Link AI 2023: AI 2023: Advances in Artificial Intelligence pp 120-132

Part of the Lecture Notes in Computer Science book series (LNAI, volume 14472)

AI 2023: Advances in Artificial Intelligence

36th Australasian Joint Conference on Artificial Intelligence, AI 2023, Brisbane, QLD, Australia, November 28-December 1, 2023, Proceedings, Part II

Quality Indicators

Peer Reviewed

Relevance to the Centre

Enterprise knowledge graphs are gaining increasing popularity in industrial applications, with a pressing demand for natural language interfaces to support non-technical end-users. For natural language queries to relational databases, the neural semantic parsing task Text-to-SQL achieves strong performance in translating text inputs to SQL queries. However, very few public corpora are available for the training of neural semantic parsing models that convert textual queries to graph query languages. In this research, we develop a generic SQL2Cypher algorithm that can map a SQL query to a set of Cypher clauses, where Cypher is a query language used by a popular property graph database Neo4j. The converted Cypher statement is then combined with the original natural language query to create a parallel corpus that enables end-to-end training of neural semantic parsing models for Text-to-Cypher.

To evaluate the dataset quality, we construct a corresponding graph database to obtain execution accuracy. In addition, the Text-to-Cypher corpus features four transformer-based baseline models. The availability of such corpus and baseline models is critical in developing and benchmarking new machine learning methods in advancing natural language interfaces for fact retrieval from large graph-based knowledge repositories. The source code and dataset are available at github(<https://github.com/22842219/SemanticParser4Graph>)

Link to Publication