

NLPCC 2019 Shared Task:

Open Domain Semantic Parsing

In this year's NLPCC, we call for the *Open Domain Semantic Parsing* shared task. The goal of this task is to predict the correct logical form (in lambda-calculus) for each question in the test set, based on a given knowledge graph.

In this task, a **Multi-perspective Semantic ParSing** (or **MSParS**) dataset will be released, which can be used to evaluate the performance of a semantic parser from different aspects. This dataset includes more than 80,000 human-generated questions, where each question is annotated with entities, the question type and the corresponding logical form. We split MSParS into a train set, a development set and a test set. Both train and development sets will be provided to participating teams, while the test set will NOT. After participating teams submit their output files, we will evaluate their performances.

In MSParS, there is a total of 9 question types, including *single-relation*, *multi-hop*, *multi-constraint*, *multi-choice*, *aggregation*, *comparison*, *yes/no*, *superlative*, and *multi-turn*. Below table gives an annotation example for each question type.

Question Type	Question	Logical Form
single-relation	Who directed Forrest Gump	$\lambda x. \text{film_director}(\text{ForrestGump}, x)$
multi-hop	When was Google's founder born	$\lambda x. \exists y. \text{organization}(\text{Google}, y) \wedge \text{person_dateofbirth}(y, x)$
multi-constraint	What film was starred by Tom Hanks in 1994	$\lambda x. \text{film_actor}(x, \text{TomHanks}) \wedge \text{film_releasedate}(x, 1994)$
multi-choice	who invented iPod, Steve Jobs or Bill Gates	$\text{production_inventor}(\text{iPod}, x) \wedge (\{x == \text{SteveJobs}\} \vee \{x == \text{BillGates}\})$
aggregation	How many children does Bill Gates have	$\text{count}(\lambda x. \text{person_child}(\text{BillGates}, x))$
comparison	Which rivers are longer than 5,000 kms	$\text{compare}(\lambda x. \text{isa}(x, \text{river}), \lambda x. \lambda y. \text{river_length}(x, y), 5000)$
yes/no	Is iPod invented by Steve Jobs	$\text{production_inventor}(\text{iPod}, \text{SteveJobs})$
superlative	Which is the third longest river on earth	$\text{argmax}(\lambda x. \text{isa}(x, \text{river}), \lambda x. \lambda y. \text{river_length}(x, y), 3)$
multi-turn	Who founded Microsoft \rightarrow When was it founded	$\lambda x. \text{organization_dateof found}(\text{Microsoft}, x)$

For submission file, each participating team should predict the logical form of each question in the test set (i.e., MSParS.test) and fill the prediction to the field **<logical form id=XXX>**. Note, you can add your predictions to **<parameters id=XXX>** and **<question type id=XXX>** as well, but we will NOT evaluate them. Below is an example.

- **<question id=XXX>** chatfield reservoir average depth
- **<logical form id=XXX>** **add-your-prediction-here**
- **<parameters id=XXX>** will-NOT-evaluate
- **<question type id=XXX>** will-NOT-evaluate

Please send your final submission file to nanduan@microsoft.com before the submission deadline (2019-05-20).

Note, the test set will be released on 2019-05-15 (on GitHub).

The dataset is now downloadable on GitHub (<https://github.com/msra-nlc/MSParS>). Any question, please contact: nanduan@microsoft.com. We hope this dataset can further advance the development of the natural language understanding research.