

SymISO

Overview

This tool aims to match any given query metagraph (i.e., to compute the instances of the metagraph) over a large input graph. For the definition and examples of metagraphs, refer to the citation below. Currently, the tool only works on undirected graphs, where nodes are typed, and edges are untyped (or rather, edge type is a function of the two node types).

Citation

Semantic Proximity Search on Graphs with Metagraph-based Learning.

Y. Fang, W. Lin, V. W. Zheng, M. Wu, K. C.-C. Chang and X. Li.
In *ICDE* 2016, pp. 277--288.

Code and Data

The provided code *symiso* is a compiled binary for linux environment.
Sample datasets are included. The source of these datasets can be found in the above citation.

Usage

Command line

```
./symiso data=<String> query=<String> maxfreq=<Integer> subgraph=<String>  
stats=<String>
```

Command line arguments

1. `data=<String>`
The input graph filename. The file is in the [Labeled Graph Format](#). The graph is treated as undirected, and edge types are not considered at the moment.
2. `query=<String>`
The input filename for a list of query metagraphs, in the [Metagraph Query Format](#). These query metagraphs can be mined from the input graph using a [modified version of GRAMI](#).
3. `maxfreq=<Integer>`
The maximum number of instances to match, for each query metagraph. The program immediately moves on to the next query after the specified maximum number of instances are found.
4. `subgraph=<String>`
The filename to output the metagraph database, which contains a list of processed metagraphs. The file is in the [Metagraph Database Format](#).
5. `stats=<String>`
The directory name to output matched instances of each metagraph. Make sure you manually create the the directory before running.
 - One instance file per metagraph.

- Instance filenames are named according to the ID of each processed metagraph (see subgraph above).
- Each line in an instance file representing one instance, containing tab delimited node IDs of the input graph, in the order according to the order of nodes in the processed metagraph (see subgraph above).

Sample command line

```
./symiso data=dblp.lg query=dblp.q maxfreq=1000000000 subgraph=dblp.gdb  
stats=output.dblp
```

Disclaimer

We provide any code and/or data on an as-is basis. Use at your own risk.