

Exercise Sheet 7: Expressivity of SPARQL

Maximilian Marx, Markus Krötzsch

Knowledge Graphs, 2023-12-05, Winter Term 2023/2024

Exercise 7.1. Which of the following graph patterns are expressible in SPARQL? Explain your answer by either giving a SPARQL query or by arguing why there is none.

1. Find nodes that are connected by an `eg:edge` path of length ≥ 100
2. Find nodes that are connected by an `eg:edge` path of length ≤ 100
3. Find nodes that are connected by an `eg:edge` path of length $\neq 100$
4. Find nodes that are not connected by an `eg:edge` path of length 100
5. In a graph with a `eg:parent` property, find nodes with a common ancestor
6. In a graph with a `eg:parent` property, find nodes that are cousins (of any degree)
7. Find nodes that are connected by `eg:propA` but not by `eg:propB`
8. Find nodes that are connected by an `eg:propA` path, but not by an `eg:propB` path
9. Find nodes that are connected by a path of nodes as in 7
10. Find nodes connected by an arbitrary path
11. Find nodes connected by an arbitrary path of even length
12. Check if the graph contains an even number of nodes

Exercise 7.2. Find a family of SPARQL queries that produce solutions where a variable name is mapped to a value that requires an exponential number of characters to write down (measured in the size of the query and RDF graph). What can you say about the growth of the result's size with respect to the size of the RDF graph when keeping the query fixed?

Exercise 7.3. Wikidata also contains lexicographic information: *Lexemes* are entities that have a language, a *Lemma* (the actual character sequence), and support claims the same way that other Wikidata entities do. In the Wikidata query service, lexemes are encoded using an `rdf:type` of `ontolex:LexicalEntry`. The language is identified by `dct:language`, and the lemma by `wikibase:lemma`. A lexeme can also have *senses* specifying their meaning, they can be reached by the `ontolex:sense` property, from which `wdt:P5137` connects to the corresponding Wikidata item.

Using this, write a query that uses the Wikidata query service¹ to find the top 10 languages by the number of lexemes that have at least one meaning corresponding to some kind of snow.

¹<https://query.wikidata.org>

Exercise 7.4. Consider the Datalog program P

$$\begin{aligned} \text{Parent}(x, y) &:- \text{father}(x, y) \\ \text{Parent}(x, y) &:- \text{mother}(x, y) \\ \text{Ancestor}(x, y) &:- \text{Parent}(x, y) \\ \text{Ancestor}(x, z) &:- \text{Parent}(x, y), \text{Ancestor}(y, z) \\ \text{Result}(y) &:- \text{Ancestor}(\text{alice}, y) \end{aligned}$$

and the facts

mother(alice, barbara)	father(alice, bob)
mother(barbara, christine)	father(barbara, charles)
mother(dave, emmy)	father(bob, dave)

Compute all query results for $\langle P, \text{Result} \rangle$.