Exercise Sheet 11: Knowledge Graph Quality and Validation

Maximilian Marx, Markus Krötzsch

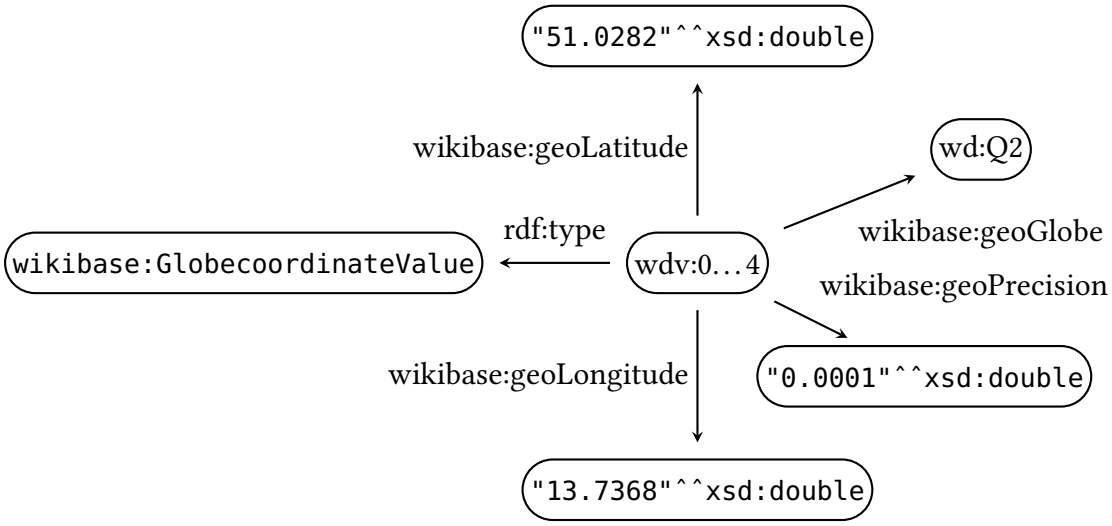
Knowledge Graphs, 2026-01-28, Winter Term 2025/2026

Exercise 11.1. Wikidata Property Constraints[1] are a mechanism to specify how properties should be used on Wikidata. As an example, a Symmetric Constraint[2] specifies that every statement for a given property must have a matching statement in the reverse direction (e.g., every "spouse" statement must have a matching "spouse" statement).

Use Nemo[3] and the Wikidata Query Service[4] to find statements violating the Symmetric Constraint for "spouse":

- write a SPARQL query trying to find all "spouse" statements lacking a symmetric counterpart.
- write a Datalog program that imports all "spouse" statements and uses stratified negation to find the statements lacking a symmetric counterpart.

Exercise 11.2. The following extract from Wikidata shows how geographic coordinates are encoded. Develop a SHACL schema that validates statement values for geographic coordinates in Wikidata.



Hint: Refer to the RDF Dump Format description[1] for details on the encoding.

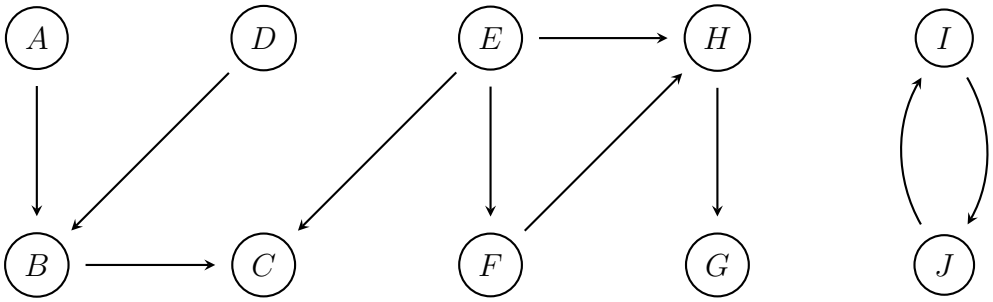
[1] https://www.wikidata.org/wiki/Help:Property_constraints_portal
[2] https://www.wikidata.org/wiki/Help:Property_constraints_portal/Symmetric
[3] https://tools.iccl.inf.tu-dresden.de/nemo/next
[4] https://query.wikidata.org
[1] https://www.mediawiki.org/wiki/Wikibase/Indexing/RDF_Dump_Format#Globe_coordinate

Exercise 11.3. Show that deciding whether a given RDF graph is valid with respect to some fixed ShEx schema is NP-hard by reducing from 3-colourability.

Hint: You can use the RDF Shape playground[2] to test ShEx validation.

Exercise 11.4. Let $G = \langle V, E \rangle$ be a directed graph. For two vertices $v, w \in V$, the *distance* $d(v, w)$ is the length of the shortest directed path from v to w (or ∞ if there is no such path). For a set $S \subseteq V$ of vertices, the *reachable set* is $R(S) \coloneqq \{\, v \in V \mid \exists w \in S.\ d(w, v) < \infty \,\}$. A *point base* of G is a minimal set $B \subseteq V$ such that $R(B) = V$.

Find a point base for the following graph. How does the point base change when adding the edge $\langle B, E \rangle$?



[2] http://rdfshape.weso.es