Difference hat zwei listen: triples hinzugefügt und gelöscht

Diese werden dann vom ursprünglichen graphen hinzugefügt oder gelöscht. Dann wird qse ganz normal ausgeführt. Laufzeit:

Laufzeit normal: O(2n) + O (2m)

Laufzeit-Differenz: O(2\*n\*m) + O(2n) (??) weiß nicht wie lange difference von jena braucht

Mit sets:

Zwei mal converten O(n)+O(m), Zwei mal remove all O(n)+O(m) + durchlaufen O(2n)

Kann nicht schneller sein, außer es gibt eine Variante, die Change sets schneller zu bekommen als 2m

Könnte schneller sein, wenn man changeset hinauflädt (added triples und removed triples)

Könnte auch das changeset nehmen und dann direkt in den algorithmus eingreifen mit den daten

Allgemein nachfragen

Allgemein nachfragen: ist input immer zwei graphen oder auch nur changeset? Ist es dann rdfpatch oder einfach 2 versionen?

Interim resulsts: default shapes, die man nicht wieder konstruieren muss dann.

Shactor mit film.nt:

Wenn Datenset ausgeführt wird, werden zuerst mittels type die klassen herausgefiltert.

* Parser, entityExtraction
  + Befüllt entitydatahashmap: Map<Node, EntityData>
  + Befüllt Map<Integer, Integer> classEntityCount
  + (Schritt 1)
* Nach Klassenauswahl (geht graph zum zweiten mal durch)
  + Parser, entityConstraintsExtraction
  + Befüllt entityDataHashMap weiter (Schritt 2)
  + Befüllt classToPropWithObjTypes
* extractSHACLShapes
* ShapesExtractor: constructDefaultShapes

Welche Results kann ich verwenden, außer die shapes, die schon konstruiert worden sind? Kashif fragen?

Wie funktionieren Changesets für große Graphen? Ist es dann nicht komplett unperformant, zuerst die genaue Differenz zu berechnen, wenn später eh gesamplet wird?

Exakt und approx nochmal besprechen

Kashif fragen

Qualitative evaluation: reicht expert interview oder muss ich auch fragebogen machen, um die usablity zu messen?

Fragen:

<https://afs.github.io/rdf-delta/cmds.html>

Passt state of the art für sie?

Proposal Review Problem Statement: ♣ What is the problem that does not exist anymore once the thesis is (successfully) finished? Try to describe it in one or two sentences. ♣ Does the proposal specify any measurable characteristics of the current situation that need to be solved? Are metrics (KPIs) for these characteristics defined? Are there even values for these characteristics given? If so, outline them.

Goals and Expected Outcome: ♣ What is the proposed solution (to overcome the problem)? Try to describe it in one or two sentences. ♣ Are the characteristics of the solution well defined? Are metrics (KPIs) for these characteristics specified? If so. Outline them. ♣ When is the solution considered a success?

Research Questions: ♣ Is it possible to provide scientific evidence for all answers to the research questions? ♣ Are there any research questions that are not precise (Yes/No questions, “How to develop sth”, etc.)? ♣ If the goal is to develop something appropriate/suitable/etc., does it become clear what appropriate/suitable/etc. means?

Research Methods: ♣ Are the envisioned research method(s) appropriate to reach the expected research results? ♣ Are the methodological steps to be followed clear? If so, outline them. ♣ Is each methodological step based on a research method known from the literature? If so, mention the corresponding paper from the literature describing the underlying research method for each methodological step. ♣ [Only for a thesis that follows Design Science: Design Science requires by definition the application of rigorous research methods. Is it clear which research methods are used within the Design Science framework?]

Evaluation: ♣ Which research method is used for the evaluation? ♣ Are the metrics to be used in the evaluation clear? Which are they? ♣ Is the evaluation appropriate? Does the evaluation provide scientific evidence by coming up with results based on KPIs that demonstrate that the solution solves the identified problem? Provide arguments. [Note, an approach could also fail - so after the thesis, the instantiation of the proposed evaluation may show that the problem is still unsolved.]

State of the Art: ♣ Does the elaboration on state of the art show that the student knows the relevant literature in the field? Provide arguments. ♣ Is there any literature missing?

Relevance to the Curriculum: ♣ Do you consider the thesis relevant for the given curriculum? Provide arguments. ♣ Is the relevance argued by a somewhat arbitrary list of courses (without arguing why each course is relevant), or does the section elaborate on the knowledge necessary to conduct the thesis and, then, show in which courses this know-how is taught?

Overall review ♣ Do you consider the proposal concise? ♣ Is the proposal of excessive length and, thus, should be shortened? First, note a proposal should be limited to about five pages (9000 characters including whitespaces) without the literature list. ♣ Is the style of writing appropriate? Is there a clear line of arguments? ♣ What are the strengths of the proposal? ♣ What are the limitations of the proposal? ♣ What should be revised?

Seaborn und rdfpatch ist eher für changesets, wenn man selber changes durchführt.

Apache jena difference compared die zwei graphen ganz einfach.