Knowledge Graph Evaluation System

Assignment 8 - April 3rd, 2017
Amar, Ian & Sabbir
Project Updates

- Modeling of Individuals in Magicdraw
  - Individuals modeled based on results from Stardog Query
  - Incorporate Predicates
    - Examples: hasMentionType & fromSentence
  - Entity Disjointness
    - Modeling inconsistencies as subClassOf relationships to existing hierarchy

- Utilized reasoning tools to assure that the current ontology is valid

- Use Case Updates - Competency Questions
  - Revision of competency questions with outdated terminology

- Website Revisions
P200 Instances

- P200 is extracted as different types i.e. site, family and protein
- Instances of P200 are caught by the maxCardinality restriction
- site and protein are asserted to be disjoint.
- However protein and family aren’t disjoint
## P200 Justifications

<table>
<thead>
<tr>
<th>Modification</th>
<th>URI:</th>
<th>Label:</th>
<th>Justification:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entity Mention</td>
<td>ment-PMC524479-UAZ-r1-524479-55-73</td>
<td>P200</td>
<td>Added entity mention individual with the label and URI</td>
</tr>
<tr>
<td>Entity Mention</td>
<td>ment-PMC524479-UAZ-r1-524479-55-74</td>
<td>P200</td>
<td>Added entity mention individual with the label and URI</td>
</tr>
<tr>
<td>Entity Mention</td>
<td>ment-PMC507879-UAZ-r1-507879-9-34</td>
<td>P200</td>
<td>Added entity mention individual with the label and URI</td>
</tr>
<tr>
<td>Protein Mention</td>
<td>protein-ment-PMC524479-UAZ-r1-524479-55-73</td>
<td>P200</td>
<td>Added same as protein individual</td>
</tr>
<tr>
<td>Site Mention</td>
<td>site-ment-PMC524479-UAZ-r1-524479-55-74</td>
<td>P200</td>
<td>Added same as site individual</td>
</tr>
<tr>
<td>Family Mention</td>
<td>family-ment-PMC507879-UAZ-r1-507879-9-34</td>
<td>P200</td>
<td>Added same as family individual</td>
</tr>
</tbody>
</table>
P57 Instances

- Predicates `hasMentionType` to check the type of these individuals for inconsistencies such as disjointness of incompatible concepts

- Used `fromSentence` to check the context of the extraction for redundancy
<table>
<thead>
<tr>
<th>Modification Instance Type</th>
<th>URI:</th>
<th>Label:</th>
<th>Justification:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entity Mention</td>
<td>ment-PMC522874-UAZ-r1-522874-180-333</td>
<td>p57</td>
<td>Added entity mention individual with the label and URI</td>
</tr>
<tr>
<td>Entity Mention</td>
<td>ment-PMC520749-UAZ-r1-520749-204-682</td>
<td>p57</td>
<td>Added entity mention individual with the label and URI</td>
</tr>
<tr>
<td>Entity Mention</td>
<td>ment-PMC520749-UAZ-r1-520749-204-681</td>
<td>p57</td>
<td>Added entity mention individual with the label and URI</td>
</tr>
<tr>
<td>Entity Mention</td>
<td>ment-PMC522874-UAZ-r1-522874-180-331</td>
<td>p57</td>
<td>Added entity mention individual with the label and URI</td>
</tr>
<tr>
<td>Family Mention</td>
<td>family-ment-PMC522874-UAZ-r1-522874-180-333</td>
<td>p57</td>
<td>Added same as family individual</td>
</tr>
<tr>
<td>Site Mention</td>
<td>site-ment-PMC520749-UAZ-r1-520749-204-682</td>
<td>p57</td>
<td>Added same as site individual</td>
</tr>
<tr>
<td>Protein Mention</td>
<td>protein-ment-PMC520749-UAZ-r1-520749-204-681</td>
<td>p57</td>
<td>Added same as protein individual</td>
</tr>
<tr>
<td>Protein Mention</td>
<td>protein-ment-PMC522874-UAZ-r1-522874-180-331</td>
<td>p57</td>
<td>Added same as protein individual</td>
</tr>
</tbody>
</table>
How Individuals are Being Used

- Individuals contain context regarding the point of extraction
- Individuals provide semantics to create disjointness rules
- Individuals contain mention types which can be checked for inconsistencies
- Individuals have an rdf:type associated with a supporting ontology, which can be leveraged for domain specific inference
Question: Which entity extractions are mapped to disjoint types?

Example Instances:

- `<ment-PMC524479-UAZ-r1-524479-55-73> rdfs:label <P200> kgcs:hasMentionType <Protein> kgcs:fromSentence <sent-PMC524479-UAZ-r1-524479-55>`
- `<ment-PMC524479-UAZ-r1-524479-55-74> rdfs:label <P200> kgcs:hasMentionType <Site> kgcs:fromSentence <sent-PMC524479-UAZ-r1-524479-55>`

- Two individuals extracted from the same sentence have different values for kgcs:hasMentionType
- Disjointness rule: disjointWith(Protein, Site)
  - Individual entity mentions will trigger corresponding disjointness inconsistency
- This example assumes that a protein cannot be a site, which may not necessarily be true
- It is possible to consider more obvious cases, such as disjointWith(Site, Family)
Future Work

- Addition of new individuals based on results of further queries
- Further updating of use case competency questions to align them with individuals listed in the ontology
  - Future iterations of application may result in different results
- Update Jena code to generate automatic knowledge graphs

Any Questions?