Querying RDF, RDFS, OWL

Part 1
What is a Graph Query Language?

- A Graph Query language *should* allow us to
  - Retrieve any query-specified portion of some graph data
  - Create a new graph by combining different pieces of retrieved subgraphs in a query-specified way
  - Compute a set of graph properties
    - Diameter
    - Distance between two nodes
    - Centrality of nodes
    - ...

- We will discuss SPARQL
  - Standard RDF Query Language
  - SPARQL only allows us to do a few of the operations an ideal graph query language should
A Single Variable Graph Pattern

\[
\text{http://xtech.2008.org}
\]

\[
\text{http://eve/}
\]

\[
\text{foaf:interest} \ ?x
\]

\[
\text{http://xtech2008.org}
\]

\[
\text{http://www2008.org}
\]

\[
\text{"Alice rules"}
\]
SELECT ?x, ?y where {?x foaf:interest ?y}

<table>
<thead>
<tr>
<th>?x</th>
<th>?y</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://eve/">http://eve/</a></td>
<td><a href="http://xtech.org">http://xtech.org</a></td>
</tr>
</tbody>
</table>
Basic Graph Patterns

What has Alice written?
- BGP

Who has common interests?
- BGP

Matching Literals
- Consider the data
  - Will it match
    - {?x hasPet “cat”} ?
    - {?x hasAge 29} ?

```turtle
{?x dc:creator http://alice/ .
 ?x dc:title ?y}

{?x dc:creator http://alice/;
 dc:title ?y}

{?x foaf:interest ?y .
 ?z foaf:interest ?y }

http://alice/ hasAge 29^^xsd:integer
http://alice/ hasPet “cat”@en
```
Structure of a SPARQL Query

- **Prologue:**
  - Prefix definitions are references in the query
  - No period ("." characters) character to separate (as in N3)
  - If we said `PREFIX : http://example.org/Hackers`
    - We could drop the FROM clause
    - We have to say :?x etc.

```sparql
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?x, ?z, ?y
FROM <http://example.org/Hackers>
WHERE {
  {?x dc:creator ?z .
   ?x dc:title ?y}
}
ORDER BY ?y
```
Structure of a SPARQL Query

- **Result form specification:**
  - **SELECT, DESCRIBE, CONSTRUCT, or ASK**
  - **SELECT:** - Variable list or asterisk ("*") character for all
    - **DISTINCT** for disjoint results
Structure of a SPARQL Query

Dataset specification:
- Specify the datasets to be queried
- FROM and FROM NAMED clauses (each with a URI)
- When multiple datasets are specified, the system assumes an RDF merge of the two graphs
- FROM NAMED is discussed later

```sparql
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
SELECT ?x, ?z, ?y  
FROM <http://example.org/Hackers>  
WHERE {  
{?x dc:creator ?z .  
 ?x dc:title ?y  
}  
ORDER BY ?y
```
Structure of a SPARQL Query

- **Solution modifiers:**
  - Modify the result set, but not the single results
  - ORDER BY, LIMIT, or OFFSET
    - **LIMIT** gets a query-specified number of results
    - **OFFSET** $k$ gets results starting from the $k$-th result record

```sparql
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  

SELECT ?x, ?z, ?y  
FROM <http://example.org/Hackers>  
WHERE {
  {?x dc:creator ?z .  
    ?x dc:title ?y}
}
ORDER BY ?y
```
Graph Patterns in SPARQL

- Basic graph pattern (BGP)
- Optional graph pattern
- Union graph pattern
- (Constraints)
- Graph graph pattern
- Group graph pattern
More on BGPs

- Using Blank Nodes in Queries
  - Blank nodes in graph patterns act as variables, not as references to specific blank nodes in the data being queried.
  - Permitted as subject and object of a triple pattern
    - Non-selectable variables
    - Indicated either as _:abc or as [ ]

Blank node identifiers can appear in query results
Optional Graph Patterns

- Who commented on “trouble_with_bob”?

```sql
select ?p, ?t
where {Trouble_with_bob blog:comment ?y .
  ?y dc:title ?t}
```

- Does not report eve

```sql
select ?p, ?t
where {Trouble_with_bob blog:comment ?y .
  optional {?y dc:title ?t}
}
```

- Reports eve
Union Graph Patterns

- Who is interested in the conferences Xtech 2008 OR WWW 2008?

```sparql
select ?x
where {
  {?x foaf:interest http://xTech2008/}
  UNION
  {?x foaf:interest http://www2008/}
}
```

- Union patterns are used to query for alternatives

```sparql
select ?x, ?y
where {
  {John foaf:interest ?x }
  UNION
  {John likes ?y}
}
```
Constraints – Filters

- Constraints filter solutions
  - Keyword **FILTER** followed by expression
  - Filter expressions contain operators and functions

```sparql
select ?y
where {
  _b20 dc:title ?y
  filter regex(?y "rule")
}
```

<table>
<thead>
<tr>
<th>?y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alice Rules</td>
</tr>
</tbody>
</table>
## Built-in Constraints

### Unary Operators

<table>
<thead>
<tr>
<th>Operator</th>
<th>Type(A)</th>
<th>Result type</th>
</tr>
</thead>
<tbody>
<tr>
<td>! A</td>
<td>xsd:boolean</td>
<td>xsd:boolean</td>
</tr>
<tr>
<td>+ A</td>
<td>numeric</td>
<td>numeric</td>
</tr>
<tr>
<td>- A</td>
<td>numeric</td>
<td>numeric</td>
</tr>
<tr>
<td>BOUND(A)</td>
<td>variable</td>
<td>xsd:boolean</td>
</tr>
<tr>
<td>isURI(A)</td>
<td>RDF term</td>
<td>xsd:boolean</td>
</tr>
<tr>
<td>isBLANK(A)</td>
<td>RDF term</td>
<td>xsd:boolean</td>
</tr>
<tr>
<td>isLITERAL(A)</td>
<td>RDF term</td>
<td>xsd:boolean</td>
</tr>
<tr>
<td>STR(A)</td>
<td>literal / URI</td>
<td>simple literal</td>
</tr>
<tr>
<td>LANG(A)</td>
<td>literal</td>
<td>simple literal</td>
</tr>
<tr>
<td>DATATYPE(A)</td>
<td>literal</td>
<td>simple literal</td>
</tr>
</tbody>
</table>
Find me all landlocked countries with a population greater than 15 million with the highest population country first.

```sparql
PREFIX type: <http://dbpedia.org/class/yago/>
PREFIX prop: <http://dbpedia.org/property/>
SELECT ?country_name ?population
WHERE {
  ?country a type:LandlockedCountries ;
    rdfs:label ?country_name ;
    prop:populationEstimate ?population .
  FILTER (?population > 15000000 && langMatches(lang(?country_name), "EN")).
} ORDER BY DESC(?population)
```

Try this at [http://dbpedia.org/sparql](http://dbpedia.org/sparql)
Homework

- Find everything about the country whose name is Afghanistan in language English
  - everything means all properties of the country
- Who is Barak Obama?
- Where is Greece?
- What is the capital of Nepal?
- What is the area of work of Albert Einstein?
- How is India related to “Indira Gandhi”?