

NSWI144 – Linked Data – Lecture 6 – 19 November 2012

SPARQL

Martin Svoboda

Faculty of Mathematics and Physics
Charles University in Prague



Outline

- SPARQL 1.0
 - Semantics
 - Assignments

Basics

- Idea
 - Matching subgraphs and substitution of variables
- Syntax
 - PREFIX ...
 - **SELECT** ...
 - **FROM** ...
 - **WHERE** { ... }
 - **ORDER BY** ... **LIMIT** ... **OFFSET** ...

Basics

- Patterns
 - **Basic**... when a set of triple patterns must all match
 - **Group**... when a set of graph patterns must all match
 - **Optional**... when additional patterns may extend solutions
 - **Alternative**... when two or more possible patterns are tried
 - **Graph**... when particular datasets should be queried

Basics

- **Filters**

- Impose constraints on values of variables

- **Modifiers**

- **DISTINCT**... when duplicates should be removed
- **REDUCED**... when duplicates may be removed
- **ORDER BY**... used to order the solution sequence
- **LIMIT**... used to limit the number of returned solutions
- **OFFSET**... used to specify the first returned solution

Basics

- Forms
 - **SELECT**... standard solutions sequence
 - **ASK**... test of a solution existence
 - **DESCRIBE**... graph about selected resources
 - **CONSTRUCT**... graph constructed from a pattern

Semantics

- Basics
 - Variable mapping
 - **Compatible mappings**
 - Query semantics
- Operators
 - **join**_{rdf}... union of compatible mappings
 - **union**_{set}... standard union of mappings
 - **minus**_{rdf}... mappings incompatible with the others
 - **opt**_{rdf}... union of **join**_{rdf} and **minus**_{rdf}

Semantics

- Patterns
 - **Triple...** all possible mappings from data
 - **AND...** conjunction based on join_{rdf}
 - Used for basic and group graph patterns
 - **UNION...** disjunction based on $\text{union}_{\text{set}}$
 - **OPTIONAL...** optionality based on opt_{rdf}
- Filters
 - **FILTER...** satisfiability of constraints
 - Predicates, connectives

Sample Data

- Data assumed in the assignments...
 - **Vocabulary**
 - @prefix i: <http://is.cuni.cz/is#> .
 - **Datasets**
 - **Students**
 - @prefix s: <http://is.cuni.cz/students/> .
 - **Teaching**
 - @prefix t: <http://is.cuni.cz/teaching/> .
 - **Faculty**
 - @prefix f: <http://is.cuni.cz/faculty/> .

Sample Data

- Vocabulary
 - **Classes**
 - Student, Teacher, Name
 - Course, Enrollment, Event, Semester
 - Room, Building, Department
 - **Properties**
 - ...

Sample Data

- Student dataset

- **s:s2** **i:number** "2" ;
i:name [**i:first** "Martin" ; **i:last** "Svoboda"] ;
i:email "svoboda@ksi.mff.cuni.cz" ;
i:web "http://www.ksi.mff.cuni.cz/~svoboda/" .
- **s:s4** **i:number** "4" ;
i:name [**i:first** "Tomas" ; **i:last** "Knap"] ;
i:email "knap@ksi.mff.cuni.cz" .
- **s:s6** **i:number** "6" ;
i:name [**i:first** "Jakub" ; **i:last** "Klimek"] ;
i:web "http://www.ksi.mff.cuni.cz/~klimek/" .

Sample Data

- Student dataset

- `s:s2 i:enroll [i:course t:c7 ;
 i:term t:term1112W ;
 i:result "2"] .`
- *... and data about other students and enrollments*

Sample Data

- Teaching dataset

- `t:c7 i:title "XML Technologies" ;
i:code "NPRG036" ;
i:prerequisite t:c5 ; i:prerequisite t:c3 .`
- `t:c5 i:title "Internet" ;
i:code "NSWI096" ;
i:prerequisite t:c3 .`
- `t:c3 i:title "Programming I" ;
i:code "NPRG030" .`
- `t:term1112W i:year "2011/12" ; i:period "W" .`
- `t:t5 i:name [i:first "Irena" ; i:last "Holubova"] ;
i:affiliation f:d8 .`

Sample Data

- Teaching dataset

- **t:t5 i:teach** [
 i:course t:c7 ; i:term t:term1112W ;
 i:day "MON" ; i:time "14:00" ;
 i:place f:roomS5 ; i:length "90" ;] .
- **t:t5 i:teach** [
 i:course t:c5 ; i:term t:term1112W ;
 i:day "WED" ; i:time "12:00" ;
 i:place f:roomS3 ; i:length "90" ;] .
- **t:t5 i:teach** [
 i:course t:c5 ; i:term t:term1112W ;
 i:day "FRI" ; i:time "10:00" ;
 i:place f:roomS1 ; i:length "90" ;] .
- ... *and data about other teachers, semesters and teaching*

Sample Data

- Faculty dataset

- `f:d8 i:name "Department of Software Engineering" ;
i:id "DSE" .`
- `f:roomS5 i:number "S5" ;
i:building f:buildingS ;
i:capacity "116" .`
- `f:buildingS i:address [
i:street "Malostranske namesti 25" ;
i:city "Prague"] .`
- *... and data about other departments, rooms and buildings*

Assignment 6.1

- Select all students
 - Return student numbers, first and last names

Assignment 6.1

- Statement

- PREFIX i: <http://is.cuni.cz/is#>
SELECT ?number ?first ?last
FROM <http://is.cuni.cz/students/>
WHERE {
 ?s i:number ?number ;
 i:name [i:first ?first ; i:last ?last] .
}

- Solutions

| ?number | ?first | ?last |
|---------|----------|-----------|
| "2" | "Martin" | "Svoboda" |
| "4" | "Tomas" | "Knap" |
| "6" | "Jakub" | "Klimek" |

Assignment 6.1

- Alternative statement

- ```
PREFIX i: <http://is.cuni.cz/is#>
SELECT ?number ?first ?last
FROM <http://is.cuni.cz/students/>
WHERE {
 ?s i:number ?number ;
 i:name ?n .
 ?n i:first ?first ; i:last ?last .
}
```

# Assignment 6.2

- Select all courses with prefix equal to *NSWI*
  - Use *regex(string, pattern)* function
  - Return course codes and titles
  - Order courses by their titles

# Assignment 6.2

- Statement

- PREFIX i: <http://is.cuni.cz/is#>  
SELECT ?code ?title  
FROM <http://is.cuni.cz/teaching/>  
WHERE {  
    ?c i:code ?code ; i:title ?title .  
    FILTER ( regex(?code, "^NSWI" ) )  
}  
ORDER BY ?title

- Solutions

| ?code     | ?title ▼           |
|-----------|--------------------|
| "NSWI096" | "Internet"         |
| "NPRG030" | "Programming I"    |
| "NPRG036" | "XML Technologies" |

# Assignment 6.3

- Select students and their emails
  - Existence of emails is not required
  - Return student numbers and emails

# Assignment 6.3

- Statement

- PREFIX i: <http://is.cuni.cz/is#>  
SELECT ?number ?email  
FROM <http://is.cuni.cz/students/>  
WHERE {  
    ?s i:number ?number .  
    OPTIONAL { ?s i:email ?email . }  
}

- Solutions

| <b>?number</b> | <b>?email</b>             |
|----------------|---------------------------|
| "2"            | "svoboda@ksi.mff.cuni.cz" |
| "4"            | "knap@ksi.mff.cuni.cz"    |
| "6"            |                           |

# Assignment 6.4

- Select students, their emails and web pages
  - Both emails and web pages are not required
  - Return student numbers, emails and web pages

# Assignment 6.4

- Statement

- PREFIX i: <http://is.cuni.cz/is#>  
SELECT ?number ?email ?web  
FROM <http://is.cuni.cz/students/>  
WHERE {  
 ?s i:number ?number .  
 OPTIONAL { ?s i:email ?email . }  
 OPTIONAL { ?s i:web ?web . }  
}

- Solutions

| ?number | ?email        | ?web            |
|---------|---------------|-----------------|
| " 2 "   | "svoboda@..." | ".../~svoboda/" |
| " 4 "   | "knap@..."    |                 |
| " 6 "   |               | ".../~klimek/"  |



# Assignment 6.5

- What the following patterns mean?

- ```
WHERE {  
  ?s i:number ?number .  
  OPTIONAL {  
    ?s i:email ?email .  
    OPTIONAL { ?s i:web ?web . }  
  }  
}
```

- ```
WHERE {
 ?s i:number ?number .
 OPTIONAL { ?s i:email ?email ; i:web ?web . }
}
```

# Assignment 6.5

- Alternative 1

| ?number | ?email        | ?web             |
|---------|---------------|------------------|
| " 2 "   | "svoboda@..." | ".../~svoboda/ " |
| " 4 "   | "knap@..."    |                  |
| " 6 "   |               |                  |

- Alternative 2

| ?number | ?email        | ?web             |
|---------|---------------|------------------|
| " 2 "   | "svoboda@..." | ".../~svoboda/ " |
| " 4 "   |               |                  |
| " 6 "   |               |                  |

# Assignment 6.6

- Select all courses that are taught on Mondays or Fridays
  - Assume only courses in winter semester 2011/12
    - `t:term1112W`
  - Return course references and codes

# Assignment 6.6

- Statement

- PREFIX i: <http://is.cuni.cz/is#>  
SELECT DISTINCT ?c ?code  
FROM <http://is.cuni.cz/teaching/>  
WHERE {  
  ?c i:code ?code .  
  ?e i:course ?c ; i:day ?day ;  
  i:term t:term1112W .  
  FILTER ( (?day = "MON") || (?day = "FRI") )  
}

- Solutions

| ?c   | ?code     |
|------|-----------|
| t:c7 | "NPRG036" |
| t:c5 | "NSWI096" |

# Assignment 6.7

- Select all courses that are **not** taught on Mondays and nor on Fridays
  - Assume only courses in winter semester 2011/12
  - Return course references and codes

# Assignment 6.7

- Statement

- ...

```
WHERE {
 ?c i:code ?code .
 OPTIONAL {
 ?e i:course ?c ; i:day ?day ;
 i:term t:term1112W .
 FILTER ((?day = "MON") || (?day = "FRI"))
 }
 FILTER (! bound(?e))
}
```

- Solutions

|           |              |
|-----------|--------------|
| <b>?c</b> | <b>?code</b> |
|-----------|--------------|

# Assignment 6.7

- The following statement would not work!

- ```
PREFIX i: <http://is.cuni.cz/is#>
SELECT DISTINCT ?c ?code
FROM <http://is.cuni.cz/teaching/>
WHERE {
  ?c i:code ?code .
  ?e i:course ?c ; i:day ?day ;
    i:term t:term1112W .
  FILTER ( (?day != "MON") && (?day != "FRI") )
}
```

- Solutions

| ?c | ?code |
|------|-----------|
| t:c5 | "NSWI096" |

Assignment 6.8

- Select all teachers that have conflicts in their teaching schedules
 - Teacher has a conflict in his/her schedule, if we can find at least two schedule events with at least partially overlapping times
 - Moreover, we need at least 60 minutes for transfers between different buildings and 10 minutes for breaks within the same building
 - Assume only courses in winter semester 2011/12
 - Return first and last names of these teachers

Assignment 6.8

- Statement

- ```
PREFIX i: <http://is.cuni.cz/is#>
SELECT DISTINCT ?first ?last
FROM <http://is.cuni.cz/teaching/>
FROM <http://is.cuni.cz/faculty/>
...
```

# Assignment 6.8

- Statement...

- WHERE {  
    ?t i:name [ i:first ?first ; i:last ?last ] ;  
    i:teach ?e1 ; i:teach ?e2 .  
?e1 i:term t:term1112W ; i:day ?d1 ; i:time ?t1 ;  
    i:place [ i:building ?b1 ] ; i:length ?l1 .  
?e2 i:term t:term1112W ; i:day ?d2 ; i:time ?t2 ;  
    i:place [ i:building ?b2 ] .  
FILTER (  
    (?e1 != ?e2) && (?d1 = ?d2) && (?t1 <= ?t2) &&  
    ( ((?b1 != ?b2) && (?t2 - (?t1 + ?l1) < 60)) ||  
      ((?b1 = ?b2) && (?t2 - (?t1 + ?l1) < 10)) )  
    )  
}