

FOUNDATIONS OF SEMANTIC WEB TECHNOLOGIES

SPARQL Algebra

Sebastian Rudolph





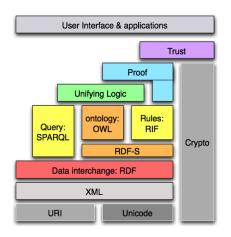
Content

Outside O VIVI

Overview & XML	9 APR DS2	Hypertableau II	/ JUN DS5
Introduction into RDF	9 APR DS3	Tutorial 5	11 JUN DS2
RDFS – Syntax & Intuition	12 APR DS5	SPARQL Syntax & Intuition	11 JUN DS3
RDFS – Semantics	16 APR DS2	SPARQL – Semantics	14 JUN DS2
RDFS Rule-based Reasoning	16 APR DS3	SPARQL Algebra	14 JUN DS3
Tutorial 1	19 APR DS5	SPARQL 1.1	14 JUN DS5
OWL – Syntax & Intuition	26 APR DS5	Tutorial 6	18 JUN DS2
Tutorial 2	3 MAY DS5	SPARQL Entailment	18 JUN DS3
OWL – Syntax & Intuition ctd.	7 MAY DS2	Ontology Editing	2 JUL DS2
OWL & Description Logics	7 MAY DS3	Ontology Engineering	2 JUL DS3
OWL 2	10 MAY DS5	Tutorial 7	9 JUL DS2
Tutorial 3	14 MAY DS2	Linked Data	9 JUL DS3
Tableau I	14 MAY DS3	Applications	12 JUL DS5
Tableau II	17 MAY DS5	Test Exam	16 JUL DS2
Tutorial 4	7 JUN DS2	Test Exam Evaluation	16 JUL DS3
Hypertableau I	7 JUN DS3	Q&A Session	19 JUL DS5

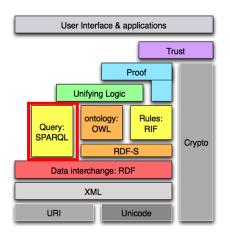


The SPARQL Query Language





The SPARQL Query Language





Agenda

- 1 Recap
- 2 Evaluation of the SPARQL Algebra
- 3 SPARQL Algebra Transformation
- Operators for the Modifiers
- 5 Summary



Agenda

- 1 Recap
- 2 Evaluation of the SPARQL Algebra
- 3 SPARQL Algebra Transformation
- Operators for the Modifiers
- 5 Summary



Recap: Introduced SPARQL Features

Basic Structure

PREFIX

Graph Patterns

Basic Graph Patterns

 $\{\dots\}$

OPTIONAL

UNION

Filter

BOUND

isBLANK

isLITERAL

STR

LANG

DATATYPE sameTERM

langMATCHES

REGEX

Modifiers

ORDER BY

LIMIT

DISTINCT

Output Formats

SELECT

ASK

ASK

DESCRIBE



```
{ ?book ex:price ?price .
FILTER (?price < 15)
OPTIONAL { ?book ex:title ?title }
  { ?book ex:author ex:Shakespeare } UNION
  { ?book ex:author ex:Marlowe }
}</pre>
```

Semantics of a SPARQL query:

- 1 Transformation of the query into an algebra expression
- 2 Evaluation of the algebra expression



```
{ ?book ex:price ?price
FILTER (?price < 15)
OPTIONAL { ?book ex:title ?title }
  { ?book ex:author ex:Shakespeare } UNION
  { ?book ex:author ex:Marlowe }
}</pre>
```

Attention: Filters apply to the whole group in which they occur



```
{ ?book ex:price ?price
  OPTIONAL { ?book ex:title ?title }
  { ?book ex:author ex:Shakespeare } UNION
  { ?book ex:author ex:Marlowe }
  FILTER (?price < 15)
}</pre>
```

Expand abbreviated IRIs





2. Replace triple patterns with operator Bgp(·)





3. Introduce the LeftJoin(·) operator for optional parts





- 4. Combine alternative graph patterns with Union(⋅) operator
- Refers to neighbouring patterns and has higher precedence than conjunction (left associative)





5. Apply Join(·) operator to join non-filter elements





6. Translate a group with filters with the Filter(·) operator





Online translation tool:

```
http://sparql.org/query-validator.html
```



Agenda

- 1 Recap
- 2 Evaluation of the SPARQL Algebra
- 3 SPARQL Algebra Transformation
- Operators for the Modifiers
- 5 Summary



Semantics of the SPARQL Algebra Operations

Bgp(P)	match/evaluate pattern P	
$Join(M_1, M_2)$	conjunctive join of solutions M_1 and M_2	
Union (M_1, M_2)	union of solutions M_1 with M_2	
LeftJoin (M_1, M_2, F)	optional join of M_1 with M_2 with filter	
	constraint F (true if no filter given)	
Filter(F, M)	filter solutions M with constraint F	
Z	empty pattern (identity for join)	



SPARQL Solutions

Solutions as partial functions

- Domain: variables from the query
- Range: IRIs ∪ blank nodes ∪ RDF literals
- Assignment σ for blank nodes in the query
- Evaluation $[Bgp(P)]_G$ of a BGP P over a graph G results in a multi set



Union of Solutions

Definition (Compatibility & Union)

Two solutions μ_1 and μ_2 are compatible if $\mu_1(x) = \mu_2(x)$ for all x, for which μ_1 and μ_2 are defined. The union of two compatible solutions μ_1 and μ_2 is defined as:

$$(\mu_1 \cup \mu_2)(x) = \begin{cases} \mu_1(x) & \text{if } x \in \text{dom}(\mu_1) \\ \mu_2(x) & \text{otherwise} \end{cases}$$

→ simple intuition: union of matching table rows



Union of Solutions

Definition (Compatibility & Union)

Two solutions μ_1 and μ_2 are compatible if $\mu_1(x) = \mu_2(x)$ for all x, for which μ_1 and μ_2 are defined. The union of two compatible solutions μ_1 and μ_2 is defined as:

$$(\mu_1 \cup \mu_2)(x) = \begin{cases} \mu_1(x) & \text{if } x \in \text{dom}(\mu_1) \\ \mu_2(x) & \text{otherwise} \end{cases}$$

- → simple intuition: union of matching table rows
- We now also define the evaluation of the other SPARQL algebra operators



Evaluation of Join(⋅)

For the evaluation of $\mathsf{Join}(A_1,A_2)$ over a graph G with A_1,A_2 algebra objects, we define:

- Let $M_1 = [A_1]_G$
- Let $M_2 = [A_2]_G$
- $\begin{array}{c} \bullet \ \ \text{Let} \, J(\mu) = \big\{ (\mu_1, \mu_2) \mid M_1(\mu_1) > 0, M_2(\mu_2) > 0, \\ \mu_1 \ \text{and} \ \mu_2 \ \text{are compatible and} \ \mu = \mu_1 \cup \mu_2 \big\} \end{array}$

 \rightarrow J defines compatible pairs of solutions from M_1 and M_2

The evaluation $[\![\mathsf{Join}(A_1, A_2)]\!]_{\mathsf{G}}$ results in

$$\left\{ \left(\mu, n \right) \mid n = \sum_{(\mu_1, \mu_2) \in J(\mu)} \left(M_1(\mu_1) * M_2(\mu_2) \right) > 0 \right\}$$



Example to Join(·)

We consider $Join(A_1, A_2)$ over a graph G with $[\![A_1]\!]_G = M_1$, $[\![A_2]\!]_G = M_2$ and:

$$\begin{split} M_1 = & \{ ((\mu_1 \colon ?x \mapsto ex : a, ?y \mapsto ex : b), 2), \\ & ((\mu_2 \colon ?x \mapsto ex : a, 1) \} \\ M_2 = & \{ ((\mu_3 \colon ?y \mapsto ex : b, ?z \mapsto ex : c, 3) \} \\ & \mu = ?x \mapsto ex : a, ?y \mapsto ex : b, ?z \mapsto ex : c \\ & J(\mu) = & \{ (\mu_1, \mu_3), (\mu_2, \mu_3) \} \end{split}$$

$$Join(M_1, M_2) = & \{ (\mu, n) \mid n = \sum_{(\mu_1, \mu_2) \in J(\mu)} (M_1(\mu_1) * M_2(\mu_2)) > 0 \} \\ = & \{ (\mu, 9) \} \\ n = & 2 * 3 + 1 * 3 = 6 + 3 = 9 \end{split}$$



Evaluation of Union(⋅)

The evaluation of $\mathsf{Union}(A_1,A_2)$ over a graph G, written $[\![\mathsf{Union}(A_1,A_2)]\!]_{\mathsf{G}}$, with A_1,A_2 algebra objects results in:

$$\left\{ (\mu,n) \mid M_1 = \llbracket A_1 \rrbracket_{\mathsf{G}}, M_2 = \llbracket A_2 \rrbracket_{\mathsf{G}}, n = M_1(\mu) + M_2(\mu) > 0 \right\}$$



Evaluation of Filter(⋅)

The evaluation of Filter(F,A) over a graph G, written $[\![Filter(F,A)]\!]_G$, with F a filter condition and A an algebra object results in:

$$\Big\{(\mu,n)\mid M=[\![A]\!]_{\mathsf{G}}, M(\mu)=n>0 \text{ and } [\![\mu(F)]\!]=\mathsf{true}\Big\}$$

 $\llbracket \mu(F) \rrbracket$ is the Boolean result of evaluating the filter condition



Evaluation of LeftJoin(·)

The evaluation of LeftJoin(A_1, A_2, F) over a graph G with F a filter condition and A_1, A_2 algebra objects is defined as:

- $M_1 = [A_1]_G$
- $M_2 = [A_2]_G$

The evaluation $[\![\operatorname{LeftJoin}(A_1,A_2,F)]\!]_G$ of $\operatorname{LeftJoin}(A_1,A_2,F)$ over G results in

$$\begin{split} & \llbracket \mathsf{Filter}(F,\mathsf{Join}(A_1,A_2)) \rrbracket_{\mathsf{G}} \cup \\ & \Big\{ \big(\mu_1, \mathit{M}_1(\mu_1) \big) \mid \mathsf{for} \ \mathsf{all} \ \mu_2 \ \mathsf{with} \ \mathit{M}_2(\mu_2) > 0 : \mu_1 \ \mathsf{and} \ \mu_2 \ \mathsf{are} \\ & \mathsf{incompatible} \ \mathsf{or} \ \llbracket (\mu_1 \cup \mu_2)(F) \rrbracket = \mathsf{false} \Big\} \end{split}$$



Example

```
@prefix ex: <http://eq.org/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
ex • Hamlet
                 ex:author ex:Shakespeare :
                 ex:price "10.50"^^xsd:decimal .
                 ex:author ex:Shakespeare .
ex:Macbeth
ex:Tamburlaine
                 ex:author ex:Marlowe;
                 ex:price "17"^^xsd:integer .
ex:DoctorFaustus ex:author ex:Marlowe;
                 ex:price "12"^^xsd:integer;
     ex:title "The Tragical History of Doctor Faustus" .
ex · Romeo.Tulia
                 ex:author ex:Brooke ;
                 ex:price "9"^^xsd:integer .
{ ?book
         ex:price ?price . FILTER (?price < 15)
 OPTIONAL { ?book ex:title ?title . }
  { ?book ex:author ex:Shakespeare . } UNION
  { ?book ex:author ex:Marlowe . }
```



```
@prefix ex: <http://eg.org/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
ex:Hamlet
                  ex:author ex:Shakespeare;
                  ex:price "10.50"^^xsd:decimal .
                  ex:author ex:Shakespeare .
ex:Macbeth
ex:Tamburlaine ex:author ex:Marlowe:
                  ex:price "17"^^xsd:integer .
ex:DoctorFaustus ex:author ex:Marlowe ;
                  ex:price "12"^^xsd:integer;
     ex:title "The Tragical History of Doctor Faustus" .
ex · Romeo.Tulia
                  ex:author ex:Brooke ;
                  ex:price "9"^^xsd:integer .
Filter(?price < 15,
 Join(LeftJoin(Bap(?book <http://eg.org/price> ?price),
            Bgp(?book <http://eg.org/title> ?title), true),
     Union(Bgp(?book <http://eg.org/author>
                <http://eg.org/Shakespeare>).
          Bap(?book <http://eg.org/author>
                <http://eg.org/Marlowe>))))
```



Example Evaluation (1)

book

ex:Tamburlaine

ex:DoctorFaustus



Example Evaluation (1)

book		
ex:Tamburlaine		
ex:DoctorFaustus		





Example Evaluation (2)

book
ex:Hamlet
ex:Macbeth
ex:Tamburlaine
ex:DoctorFaustus



Example Evaluation (3)

book	price
ex:Hamlet	10.5
ex:Tamburlaine	17
ex:DoctorFaustus	12
ex:RomeoJulia	9



Example Evaluation (3)

book	price		
ex:Hamlet	10.5		
ex:Tamburlaine	17		
ex:DoctorFaustus	12		
ex:RomeoJulia	9		

	book	title
	ex:DoctorFaustus	"The Tragical History
1		of Doctor Faustus"



Example Evaluation (4)

book	price			title			
ex:Hamlet	10.5						
ex:Tamburlaine	17						
ex:DoctorFaustus	12	"The	Tragical	History	of	Doctor	Faustus"
ex:RomeoJulia	9						



Example Evaluation (5)

	book	price			title			
	ex:Hamlet	10.5						
€	ex:Tamburlaine	17						
ех	:DoctorFaustus	12	"The	Tragical	History	of	Doctor	Faustus"



Example Evaluation (6)

book	price			title			
ex:Hamlet	10.5						
ex:DoctorFaustus	12	"The	Tragical	History	of	Doctor	Faustus"



Agenda

- 1 Recap
- 2 Evaluation of the SPARQL Algebra
- 3 SPARQL Algebra Transformation
- Operators for the Modifiers
- 5 Summary



Formal Algebra Transformation

- During parsing of a query, a parse tree is constructed
- The parse tree contains objects that correspond to the grammar
- For the transformation, we traverse the parse tree and recursively build the algebra objects
- The query pattern is a GroupGraphPattern consisting of the following elements:
 - TriplesBlock
 - Filter
 - OptionalGraphPattern
 - GroupOrUnionGraphPattern
 - GraphGraphPattern



Part of the SPARQL Grammar

```
∷= '{'TriplesBlock?
GroupGraphPattern
                                ((GraphPatternNotTriples
                                  |Filter)". ? TriplesBlock?)*
                              '}'
GraphPatternNotTriples
                           ::= OptionalGraphPattern
                               GroupOrUnionGraphPattern
                               GraphGraphPattern
                          ::= 'OPTIONAL' GroupGraphPattern
OptionalGraphPattern
                          ::= GroupGraphPattern ('UNION'
GroupOrUnionGraphPattern
                               GroupGraphPattern )*
                           ::= 'FILTER' Constraint
Filter
```



Transformation of

GroupOrUnionGraphPattern

translate(GroupOrUnionGraphPattern G)

```
Input: a GroupOrUnionGraphPattern G with elements e_1, \ldots, e_n
Output: a SPARQL algebra expression A
1: for i=1,\ldots,n do
2: if A is undefined then
3: A := translate(e_i)
4: else
5: A := Union(A, translate(e_i))
6: return A
```



Transformation of GraphGraphPattern

translate(GraphGraphPattern G)

Input: a GraphGraphPattern

Output: a SPARQL algebra expression A

1: if G GRAPH IRI GroupGraphPattern then

A := Graph(IRI, translate(GroupGraphPattern))

3: else if G GRAPH Var GroupGraphPattern then A := Graph(Var, translate(GroupGraphPattern))

5: return A



Transformation of GroupGraphPattern

translate(GroupGraphPattern G)

```
Input: a GroupGraphPattern G = (e_1, \ldots, e_n)
Output: a SPARQL algebra expression A
1: A := Z { the empty pattern}
2: F := \emptyset \{ \text{ filter} \}
3: for i = 1, ..., n do
4:
        if e_i is FILTER(f) then
5:
6:
7:
8:
9:
            F := F \cup \{f\}
        else if e_i is OPTIONAL { P } then
            if translate(P) is Filter(F', A') then
                A := LeftJoin(A, A', F')
            else
10:
                 A := LeftJoin(A, translate(P), true)
11:
         else
12:
             A := Join(A, translate(e_i))
13: if F \neq \emptyset then
         A := Filter(\bigwedge_{f \subset F} f, A)
15: return A
```



Simplification of Algebra Objects

- Groups with just one pattern (without filters) result in Join(Z, A) and can be substituted by A
- The empty pattern is the identity for joins:
 - Replace Join(Z, A) by A
 - Replace Join(A, Z) by A



Agenda

- 1 Recap
- 2 Evaluation of the SPARQL Algebra
- 3 SPARQL Algebra Transformation
- 4 Operators for the Modifiers
- 5 Summary



Operators for Representing the Modifiers

ToList(M)	Constructs from a multi set a sequence with the same elements and multiplicity (arbitrary order, duplicates not necessarily adjacent)
OrderBy(<i>M</i> , comparators)	sorts the solutions
Distinct(M)	removes the duplicates
Reduced(M)	may remove duplicates
Slice(M, o, l)	cuts the solutions to a list of length <i>l</i> starting
	from position o
Project(M, vars)	projects out the mentioned variables



Transformation of the Modifiers

Let q be a SPARQL query with pattern P and corresponding algebra object A_P . We construct an algebra object A_q for q as follows:

- 2 $A_q := \text{OrderBy}(A_q, (c_1, \dots, c_n))$ if q contains an ORDER BY clause with comparators c_1, \dots, c_n
- 3 $A_q := \text{Project}(A_q, \text{vars})$ if the result format is SELECT with vars the selected variables (* all variables in scope)
- $A_q := \operatorname{Distinct}(\mathsf{A}_\mathsf{q})$ is the result format is SELECT and q contains DISTINCT
- $oldsymbol{\delta} A_q := \mathsf{Reduced}(\mathsf{A_q})$ if the result format is <code>SELECT</code> and q contains <code>REDUCED</code>
- **6** $A_q := \operatorname{Slice}(\mathsf{A}_\mathsf{q}, \operatorname{start}, \operatorname{length})$ if the query contains <code>OFFSET</code> start or <code>LIMIT</code> length where start defaults to 0 and <code>length</code> defaults to $(\| \llbracket A_q \rrbracket_\mathsf{G} \| \operatorname{start})$



Evaluation of the Modifiers

The algebra objects for the modifiers are recursively evaluated

- Evaluate the algebra expression of the operator
- Apply the operations for the solution modifiers to the obtained solutions



Agenda

- 1 Recap
- 2 Evaluation of the SPARQL Algebra
- 3 SPARQL Algebra Transformation
- 4 Operators for the Modifiers
- 5 Summary



Summary

- We learned how to evaluate SPARQL gueries
- The query is transformed into an algebra object
- The query basic graph patterns generate solutions
- The other operators combine solutions
- The result format determines how the solutions are presented



Outlook

- Next lecture: SPARQL 1.1 features
- Non-Query parts of the specification (Protocol, Service Descriptions, Update, ...)
- Then: Entailment Regimes (SPARQL with inferred results)