

FOUNDATIONS OF SEMANTIC WEB TECHNOLOGIES

Semantics of SPARQL

Sebastian Rudolph



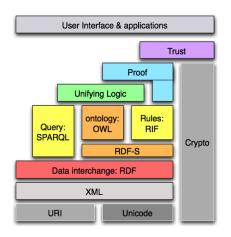


Content

Overview & XML	9 APR DS2	Hypertableau II	7 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

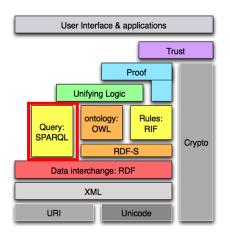


The SPARQL Query Language





The SPARQL Query Language





Agenda

- 1 Recap
- Output Formats
- 3 SPARQL Semantics
- 4 Transformation of Queries into Algebra Objects
- 5 Evaluation of the SPARQL Algebra
- 6 Summary



Agenda

- 1 Recap
- Output Formats
- 3 SPARQL Semantics
- Transformation of Queries into Algebra Objects
- 5 Evaluation of the SPARQL Algebra
- 6 Summary



Recap: Introduced SPARQL Features

Ba	 0.	 - 4 -	

PREFIX

Graph Patterns Basic Graph Patterns

{...}

OPTIONAL

UNION

Filter

BOUND

isBLANK

isLITERAL

STR

LANG DATATYPE

sameTERM

langMATCHES

REGEX

Modifiers

ORDER BY

OFFSET

DISTINCT

Output Formats

SELECT



Agenda

- 1 Recap
- Output Formats
- 3 SPARQL Semantics
- Transformation of Queries into Algebra Objects
- 5 Evaluation of the SPARQL Algebra
- 6 Summary



Output Format SELECT

So far all results have been tables (solution sequences): Output format SELECT

Syntax: SELECT <VariableList> or SELECT *

Advantage

Simple sequential processing of the results

Disadvantage

Structure/relationships between the objects in the results is lost



Output Format CONSTRUCT

CONSTRUCT creates an RDF graph for the results

Example Query



Output Format CONSTRUCT

CONSTRUCT creates an RDF graph for the results

Example Query

Advantage

Structured result data with relationships between the elements

Disadvantages

- Sequential processing of the results is harder
- No treatment of unbound variables (triples are omitted)



CONSTRUCT Templates with Blank Nodes

```
Data
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
_:a foaf:firstname "Alice";
    foaf:surname "Hacker" .
_:b foaf:firstname "Bob";
    foaf:surname "Hacker" .
```



CONSTRUCT Templates with Blank Nodes

```
Data

@prefix foaf: <http://xmlns.com/foaf/0.1/> .
_:a foaf:firstname "Alice";
   foaf:surname "Hacker" .
_:b foaf:firstname "Bob";
   foaf:surname "Hacker" .
```

Query

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX vcard: <http://www.w3.org/2001/vcard-rdf/3.0#>
CONSTRUCT {
    ?x vcard:N _ :v .
    _:v vcard:givenName ?gname;
    vcard:familyName ?fname
} WHERE {
    ?x foaf:firstname ?gname .
    ?x foaf:surname ?fname }
```



CONSTRUCT Templates with Blank Nodes

Resulting RDF graph

```
@prefix vcard: <http://www.w3.org/2001/vcard-rdf/3.0#> .
_:v1 vcard:N _:x .
_:x vcard:givenName "Alice";
    vcard:familyName "Hacker" .
_:v2 vcard:N _:z .
_:z vcard:givenName "Bob";
    vcard:familyName "Hacker" .
```



Further Output Formats: ASK & DESCRIBE

SPARQL supports two additional output formats:

- ASK only checks whether the query has at least one answer (true/false result)
- DESCRIBE (informative) returns an RDF description for each resulting URI (application dependent)

Example Query

```
DESCRIBE ?x WHERE { ?x <http://ex.org/emplID> "123" }
```

Possible Result (prefix declarations omitted):

```
_:a exOrg:emplID "123";
foaf:mbox_shalsum "ABCD1234";
vcard:N
[ vcard:Family "Smith";
vcard:Given "John" ] .
foaf:mbox_shalsum a owl:InverseFunctionalProperty .
```



Agenda

- 1 Recap
- Output Formats
- 3 SPARQL Semantics
- 4 Transformation of Queries into Algebra Objects
- 5 Evaluation of the SPARQL Algebra
- 6 Summary



Semantics of Query Languages

So far only informal presentation of SPARQL features

- User: "Which answers can I expect for my query?"
- Developer: "Which behaviour is expected from my SPARQL implementation?"
- Marketing: "Is our product already conformant with the SPARQL standard?"

--- Formal semantics should clarify these questions . . .



Logic-based Semantics

Semantics of formal logics:

- Model-theoretic semantics: Which interpretations do satisfy my knowledge base?
- Proof-theoretic semantics: Which derivations can be build from my knowledge base?
- ...



Logic-based Semantics

Semantics of formal logics:

- Model-theoretic semantics: Which interpretations do satisfy my knowledge base?
- Proof-theoretic semantics: Which derivations can be build from my knowledge base?
- ..



Semantics of Programming Languages

- Axiomatic semantics: Which logical statements hold for my program?
- Operational semantics: What happens during the processing of my program?
- Denotational semantics: How can we describe the input/output function of the program in an abstract way?



Semantics of Programming Languages

- Axiomatic semantics: Which logical statements hold for my program?
- Operational semantics: What happens during the processing of my program?
- Denotational semantics: How can we describe the input/output function of the program in an abstract way?

What to do with query languages?



Semantics of Query Languages (1)

Query Entailment

- Query as description of allowed results
- Data as set of logical assumptions (axiom set/theory)
- Results as logical entailment
- → OWL DL and RDF(S) as query languages.
- → conjunctive queries



Semantics of Query Languages (2)

Query Algebra

- Query as instruction for computing the results
- Queried data as input
- Results as output
- → Relational algebra for SQL
- → SPARQL Algebra



Agenda

- 1 Recap
- 2 Output Formats
- 3 SPARQL Semantics
- 4 Transformation of Queries into Algebra Objects
- 5 Evaluation of the SPARQL Algebra
- 6 Summary



```
{ ?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





Translation into SPARQL Algebra

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



Translation into SPARQL Algebra



Translation into SPARQL Algebra

Online translation tool:

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



Agenda

- 1 Recap
- 2 Output Formats
- 3 SPARQL Semantics
- Transformation of Queries into Algebra Objects
- 5 Evaluation of the SPARQL Algebra
- 6 Summary



Semantics of the SPARQL Algebra Operations

Now we have an algebra object, but what do the algebra operations mean?

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)



Semantics of the SPARQL Algebra Operations

Now we have an algebra object, but what do the algebra operations mean?

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)

- Only Bgp(·) matches or evaluates graph patterns
- We can use entailment checking rather than graph matching



Definition of the SPARQL Operators

How can we define that more formally?

Output:

"solution set" (formatting irrelevant)

Input:

- Queried (active) graph
- Partial results from previous evaluation steps
- Different parameters according to the operation
- → How can we formally describe the "results"?



SPARQL Results

Intuition: Results coded as tables of variable assignments

Result:

List of solutions (solution sequence)

→ each solution corresponds to one table row



SPARQL Results

Intuition: Results coded as tables of variable assignments

Result:

List of solutions (solution sequence)

--- each solution corresponds to one table row

Solution:

Partial function

Domain: relevant variables.

Range: IRIs ∪ blank nodes ∪ RDF literals

→ Unbound variables are those that have no assigned value (partial) function)



Evaluation of Basic Graph Patterns

Definition (Solution)

Let P be a basic graph pattern. A partial function μ is a solution for Bgp(P) over the queried (active) graph G if:

- 1 the domain of μ is exactly the set of variables in P,
- 2 there exists an assignment σ from blank nodes in P to IRIs, blank nodes, or RDF literals such that:
- **3** the RDF graph $\mu(\sigma(P))$ is a subgraph of *G*.



Evaluation of Basic Graph Patterns

- The result of evaluating Bgp(P) over G is written $[Bgp(P)]_G$
- The result is a multi set of solutions μ
- \bullet The multiplicity of each solution μ corresponds to the number of different assignments σ



Multi Sets

Definition (Multi Set)

A multi set over a set S is a total function $M: S \to \mathbb{N}^+ \cup \{\omega\}$

- N⁺ denotes the positive natural numbers
- $\omega > n$ for all $n \in \mathbb{N}^+$
- M(s) is the multiplicity of $s \in S$
- ω : countably infinite number of occurrences
- We represent a multi se over the set *S* also with the set $\{(s, M(s)) \mid s \in S\}$
- We write $(s, n) \in M$ if M(s) = n
- We assume that M(s) = 0 if $s \notin S$
- Alternative notation: $\{a,b,b\}$ corresponds to the multi set M over the set $\{a,b\}$ with M(a)=1 and M(b)=2



```
ex:Birte ex:gives [
   a ex:Lecture ;
   ex:hasTopic "SPARQL" ] .
ex:Sebastian ex:gives [
   a ex:Lecture ;
   ex:hasTopic "DLs and OWL" ] .

Bgp(?who ex:gives _:x . _:x ex:hasTopic ?what)
```



```
ex:Birte ex:gives _:a .
_:a rdf:type ex:Lecture .
_:a ex:hasTopic "SPARQL" .
ex:Sebastian ex:gives _:b .
_:b rdf:type ex:Lecture .
_:b ex:hasTopic "DLs and OWL" .
Bgp(?who ex:gives _:x . _:x ex:hasTopic ?what)
```





```
ex:Birte ex:gives _:a .
_:a rdf:type ex:Lecture .
:a ex:hasTopic "SPAROL" .
ex:Sebastian ex:gives _:b .
:b rdf:type ex:Lecture .
_:b ex:hasTopic "DLs and OWL" .
Bgp(?who ex:gives _:x . _:x ex:hasTopic ?what)
           ?who \mapsto ex: Birte, ?what \mapsto "SPAROL"
    \mu_1:
    \sigma_1:
          \_:x \mapsto \_:a
    \mu_2:
           ?who → ex: Sebastian. ?what → "DLs and OWL"
           :x\mapsto :b
    \sigma_2:
```



```
ex:Birte ex:gives _:a .
_:a rdf:type ex:Lecture .
:a ex:hasTopic "SPAROL" .
ex:Sebastian ex:gives _:b .
:b rdf:type ex:Lecture .
_:b ex:hasTopic "DLs and OWL" .
Bgp(?who ex:gives _:x . _:x ex:hasTopic ?what)
            ?who\mapstoex:Birte, ?what\mapsto"SPAROL"
    \mu_1:
    \sigma_1:
           \underline{\phantom{a}}:x\mapsto\underline{\phantom{a}}:a
    \mu_2: ?who \mapsto ex: Sebastian, ?what \mapsto "DLs and OWL"
            :x\mapsto :b
    \sigma_2:
```

Two solutions each with multiplicity 1



Exercise Solution Sets

```
ex:Birte ex:gives [
   a ex:Lecture ;
   ex:hasTopic "SPARQL" ] .
ex:Birte ex:gives [
   a ex:Lecture ;
   ex:hasTopic "SPARQL Algebra" ] .

Bgp(?who ex:gives _:x . _:x ex:hasTopic _:y)
```





```
ex:Birte ex:gives _:a .
_:a rdf:type ex:Lecture .
_:a ex:hasTopic "SPARQL" .
ex:Birte ex:gives _:b .
_:b rdf:type ex:Lecture .
_:b ex:hasTopic "SPARQL Algebra" .

Bgp(?who ex:gives _:x . _:x ex:hasTopic _:y)
```





```
ex:Birte ex:gives _:a .
_:a rdf:type ex:Lecture .
_:a ex:hasTopic "SPARQL" .
ex:Birte ex:gives _:b .
_:b rdf:type ex:Lecture .
_:b ex:hasTopic "SPARQL Algebra" .
Bgp(?who ex:gives _:x . _:x ex:hasTopic _:y)
    \mu_1:
            ?who \mapsto ex:Birte.
                                 :v → "SPAROL"
           :x → :a
    \sigma_1:
         ?who \mapsto ex:Birte,
    \mu_2:
                                 :y → "SPARQL Algebra"
    \sigma_2:
          :x → :b
```



```
ex:Birte ex:gives _:a .
_:a rdf:type ex:Lecture .
_:a ex:hasTopic "SPARQL" .
ex:Birte ex:gives _:b .
_:b rdf:type ex:Lecture .
_:b ex:hasTopic "SPARQL Algebra" .
Bgp(?who ex:gives _:x . _:x ex:hasTopic _:y)
    \mu_1:
           ?who \mapsto ex:Birte.
                                 :v → "SPAROL"
    \sigma_1: x \mapsto a
            ?who \mapsto ex:Birte,
    \mu_2:
                                 :y → "SPARQL Algebra"
    \sigma_2:
          :x\mapsto :b
```

One solution with multiplicity 2



Definition (Compatibility)



Definition (Compatibility)

```
\mu_1: ?x \mapsto ex : a, ?y \mapsto ex : b

\mu_2: ?y \mapsto ex : b, ?z \mapsto ex : c
```



Definition (Compatibility)

```
\mu_1: ?x \mapsto ex: a, ?y \mapsto ex: b

\mu_2: ?y \mapsto ex: b, ?z \mapsto ex: c \checkmark
```



Definition (Compatibility)

```
\mu_1: ?x \mapsto ex: a, ?y \mapsto ex: b

\mu_2: ?y \mapsto ex: b, ?z \mapsto ex: c \checkmark

\mu_1: ?x \mapsto ex: a, ?y \mapsto ex: b

\mu_2: ?x \mapsto ex: b, ?z \mapsto ex: c
```



Definition (Compatibility)

```
\mu_1: ?x \mapsto ex: a, ?y \mapsto ex: b
\mu_2: ?y \mapsto ex: b, ?z \mapsto ex: c \checkmark
\mu_1: ?x \mapsto ex: a, ?y \mapsto ex: b
\mu_2: ?x \mapsto ex: b, ?z \mapsto ex: c \checkmark
```



Definition (Compatibility)

```
\begin{array}{l} \mu_1\colon ?\mathbf{x}\mapsto \mathbf{ex}:\mathbf{a}, ?\mathbf{y}\mapsto \mathbf{ex}:\mathbf{b} \\ \mu_2\colon ?\mathbf{y}\mapsto \mathbf{ex}:\mathbf{b}, ?\mathbf{z}\mapsto \mathbf{ex}:\mathbf{c} \quad \checkmark \\ \mu_1\colon ?\mathbf{x}\mapsto \mathbf{ex}:\mathbf{a}, ?\mathbf{y}\mapsto \mathbf{ex}:\mathbf{b} \\ \mu_2\colon ?\mathbf{x}\mapsto \mathbf{ex}:\mathbf{b}, ?\mathbf{z}\mapsto \mathbf{ex}:\mathbf{c} \quad \mathbf{f} \\ \mu_1\colon ?\mathbf{x}\mapsto \mathbf{ex}:\mathbf{a} \\ \mu_2\colon ?\mathbf{y}\mapsto \mathbf{ex}:\mathbf{b} \end{array}
```



Definition (Compatibility)

```
\mu_1: ?x \mapsto ex: a, ?y \mapsto ex: b
\mu_2: ?y \mapsto ex: b, ?z \mapsto ex: c
\checkmark
\mu_1: ?x \mapsto ex: a, ?y \mapsto ex: b
\mu_2: ?x \mapsto ex: b, ?z \mapsto ex: c
\checkmark
\mu_1: ?x \mapsto ex: a
\mu_2: ?y \mapsto ex: b
```



Union of two compatible solutions μ_1 and μ_2 :

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

- → simple intuition: union of matching table rows
- Next lecture: Evaluation of the main algebra operators



Agenda

- Recap
- Output Formats
- 3 SPARQL Semantics
- Transformation of Queries into Algebra Objects
- 5 Evaluation of the SPARQL Algebra
- 6 Summary



Summary

- SPARQL queries are translated into algebra objects
- The BGPs generate solutions
- Other algebra operators combine solutions
- Details in the next lecture