Towards Unified Semantics for RDF Stream Query Processing

Daniele Dell'Aglio, **Jean-Paul Calbimonte**,
Emanuele Della Valle and Oscar Corcho

Outline

- Notion of Time-varying and Instantaneous graphs
- RSPQL datasets
- Evaluation at a time instant
- Continuous evaluation
- Challenges and future directions

Query using SPARQL on Streams

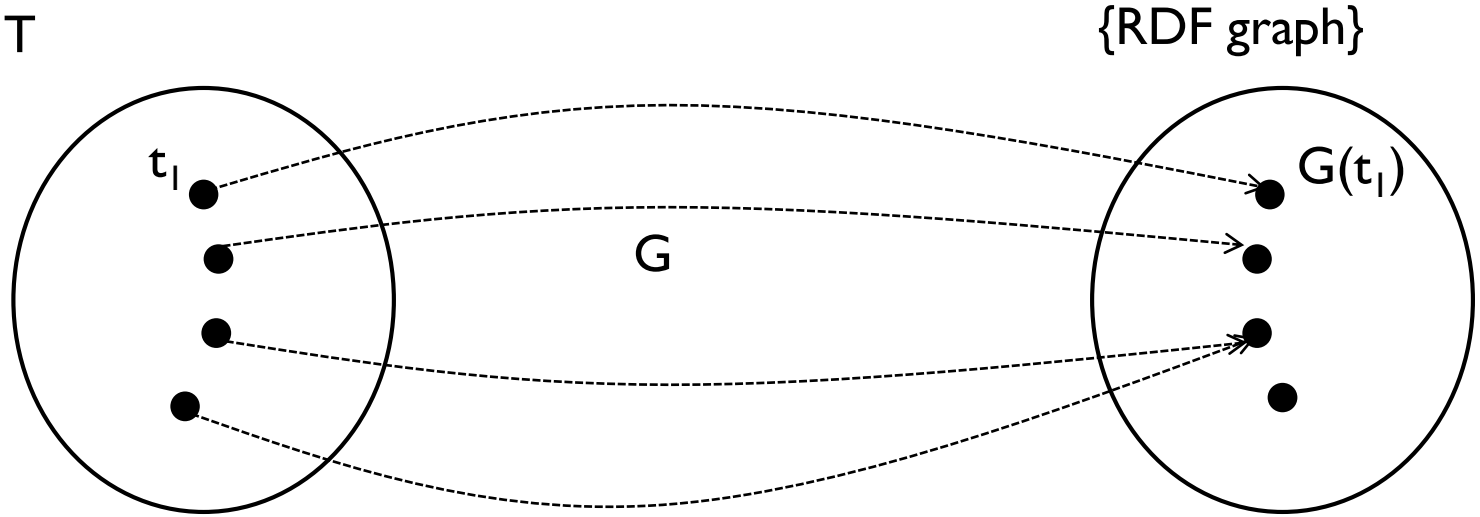
	Model	Continuous execution	Union, Join, Optional, Filter	Aggregates	Time window	Triple window	R2S operator	Sequence, Co-ocurrence	Time function
TA-SPARQL	TA-RDF	✗	✓	Limited	✗	✗	✗	✗	✗
tSPARQL	tRDF	✗	✓	✗	✗	✗	✗	✗	✗
Streaming SPARQL	RDF Stream	✓	✓	✗	✓	✓	✗	✗	✗
C-SPARQL	RDF Stream	✓	✓	✓	✓	✓	✗	✗	✓
CQELS	RDF Stream	✓	✓	✓	✓	✓	✗	✗	✗
SPARQLStream	(Virtual) RDF Stream	✓	✓	✓	✓	✗	✓	✗	✗
EP-SPARQL	RDF Stream	✓	✓	✓	✗	✗	✗	✓	✗
Instans	RDF	✓	✓	✓	✗	✗	✗	✗	✗

Powerful languages for continuous query processing

W3C RSP

→ review features in existing systems
→ agree on fundamental operators
→ discuss on possible semantics
https://www.w3.org/community/rsp/wiki/RSP_Query_Features

Time-varying and Instantaneous graphs

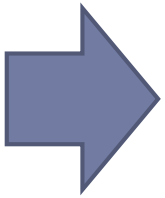


Time-Varying Graph $G:T \rightarrow \{\text{RDF graph}\}$

Instantaneous Graph $G(t_1) \in \{(s,p,o)\}$ (i.e., an RDF graph)

From SPARQL to RSP-QL

```
CONSTRUCT { …}
FROM NAMED GRAPH g:SG
FROM GRAPH g:P
WHERE {
  ?poi rdf:type e:bar .
  GRAPH g:SG {
      FILTER NOT EXIST { ?sb
e:knows ?se }
  }
  FILTER (?sb != ?se)
 }
}
```

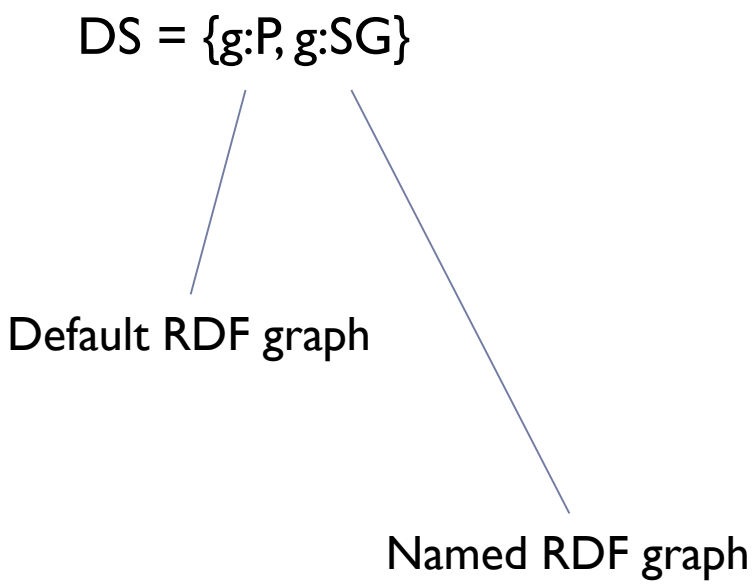


```
CONSTRUCT ISTREAM { … }
 FROM NAMED WINDOW :vLW
          ON s:I [RANGE PT4H
          STEP PT1H]
…
FROM NAMED GRAPH g:SG
FROM GRAPH g:POIs
 WHERE {
  WINDOW : vLW {
      …
  }
}
```

- (see http://bit.ly/1GLRStF)

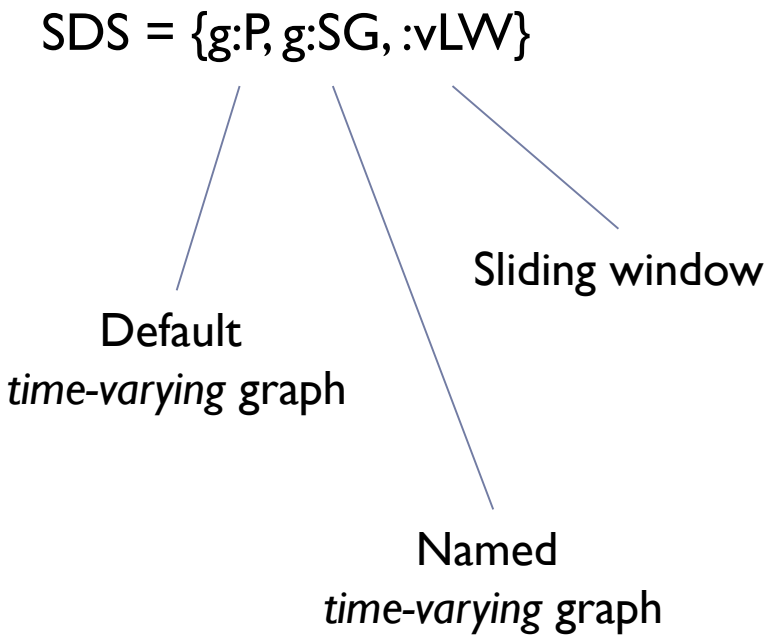
From RDF Dataset…

```
CONSTRUCT { …}
FROM GRAPH g:P
FROM NAMED GRAPH g:SG
WHERE {
…
}
```



… to RSPQL dataset

```
CONSTRUCT { …}
FROM GRAPH g:P
FROM NAMED GRAPH g:SG
FROM NAMED WINDOW :vLW
          ON s:I [RANGE PT4H
          STEP PT1H]
WHERE {
  …
 WINDOW : vLW {
     …
  }
  …
}
```



Instantaneous evaluation

- Graph patterns can be evaluated over instantaneous graphs (being RDF graphs).
- Fixed a time instant t and given a Basic Graph Pattern BGP:

$$eval(SDS, BGP, t) = eval(SDS(G, t), BGP)$$

- where SDS(G,t) =

$$SDS(G(t)) \quad \text{if G is a time-varying graph}$$
$$SDS(W(S,t)) \quad \text{if G is from a sliding window W}$$

Continuous evaluation

- The continuous evaluation is a sequence of instantaneous evaluations
  - Report is SECRET
  - Report policies:
    - CC Content Change: the window reports if the content changes.
    - WC Window Close: the window reports if the active window closes.
    - NC Non-empty Content: the window reports if the active window is not empty.
    - P Periodic: the window reports only at regular intervals
- Results of the instantaneous evaluations are combined w.r.t. the streaming operator involved in the query (Rstream, Istream, Dstream)

Conclusions

- The model captures the semantics of the window-based RSP engines (i.e., C-SPARQL, CQELS and $SPARQL_{stream}$)
- Future directions
  - Capture other RSP engines, e.g., EP-SPARQL, INSTANS
  - (Stream Reasoning?)

Thank you! Questions?

Towards Unified Semantics for RDF Stream Query Processing

Daniele Dell'Aglio (DEIB, Politecnico di Milano)

Jean-Paul Calbimonte (Faculty of Computer Science and Communication Systems, EPFL)

Emanuele Della Valle (DEIB, Politecnico di Milano)

Oscar Corcho (OEG, Universidad Politécnica de Madrid)