

Temporal RDF

Claudio Gutierrez, Carlos Hurtado

*Department of Computer Science
Universidad de Chile*

and

Alejandro Vaisman,

*Department of Computer Science
Universidad de Buenos Aires*

Best paper award, ECSW 2005

Outline

- RDF Basics
- Introducing Time Into RDF
- Temporal RDF Graphs
- Syntax and semantics of TRDF Graphs
- Querying temporal RDF
- Conclusion

RDF Basics

- Web needs semantic information => the semantic web proposal.
- Machine readable semantics for data on the web.
- RDF proposed by W3C in 1998 as a metadata language for the web.
- Simple data model, extensible vocabulary based on URIs.

RDF Basics (cont.)

- In RDF, the universe to be modeled is a set of resources.
- RDF Graph: A set of RDF triples of the form $(v1, v2, v3)$ in $(U \cup B) \times (U) \times (U \cup B \cup L)$
(U,B,L: URI references, blanks, literals)
- A graph is *ground* if has no blank nodes
- A map is a function $\mu: UBL \rightarrow UBL$ preserving URIs and literals

RDF Basics (cont.)

- RDF is extended with a vocabulary, defined in RDFS (RDF Schema).
- Typical words: class, subclassof, property, subPropertyOf, type, range, domain ,
- reification : allows making statements about statements.

RDF Basics (cont.)

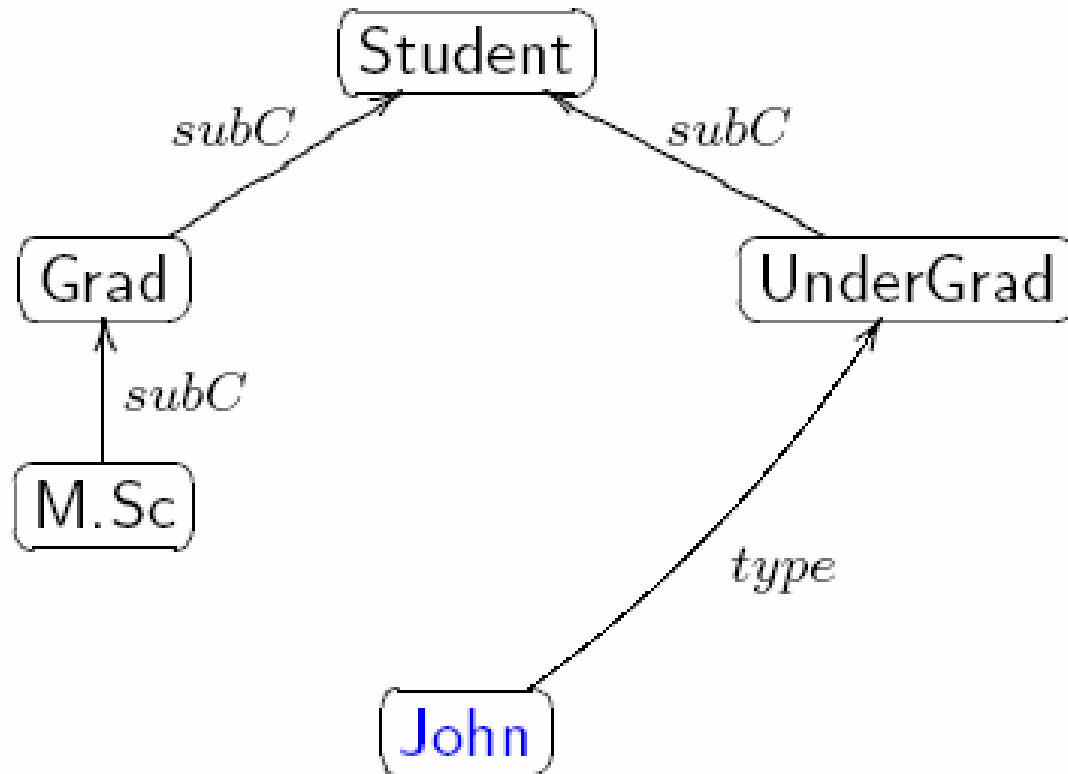
- First formal semantics given for RDF : Gutierrez, Hurtado & Mendelzon (PODS 2004).
- Gives a set of rules defining the semantics of RDF graphs.
- Defines the closure and the core of RDF graphs
=> a normal form.

RDF Basics (cont.)

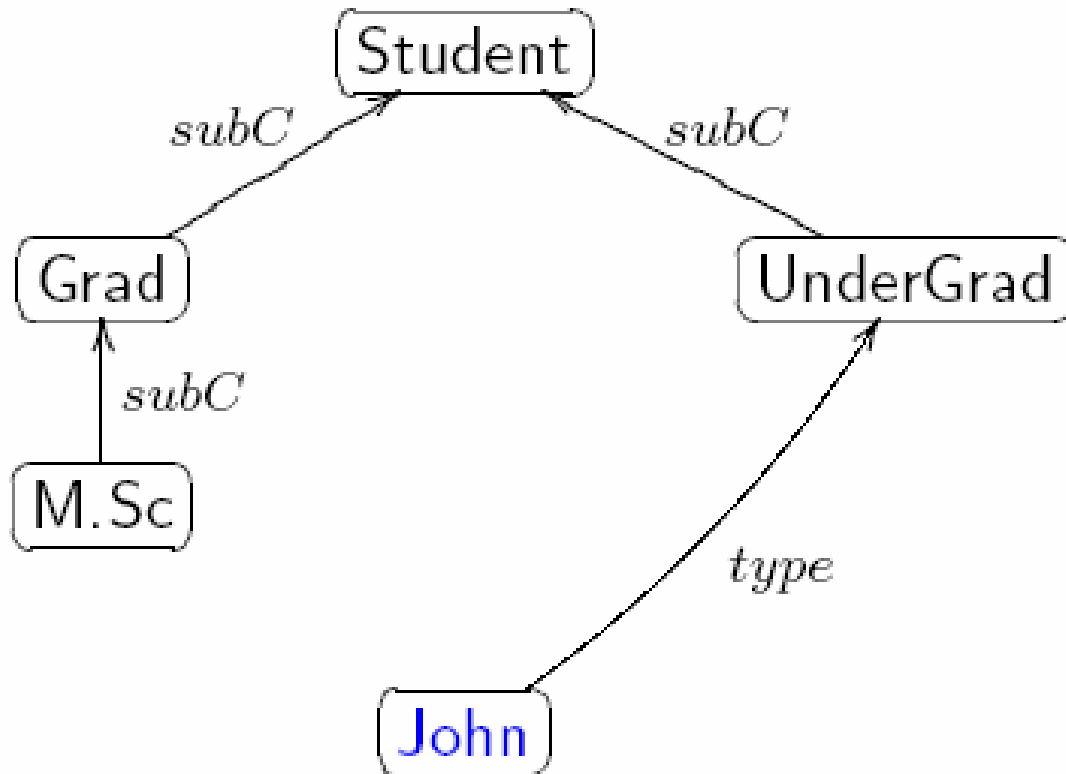
- Core : the unique minimal graph equivalent to an RDF graph G .
- Closure : a maximal set of triples G' over G , that contains G and is equivalent to G .
- Normal form (G) : the core of the closure.

(details in : “Foundations of Semantic Web Databases”, , GHM, PODS 2004)

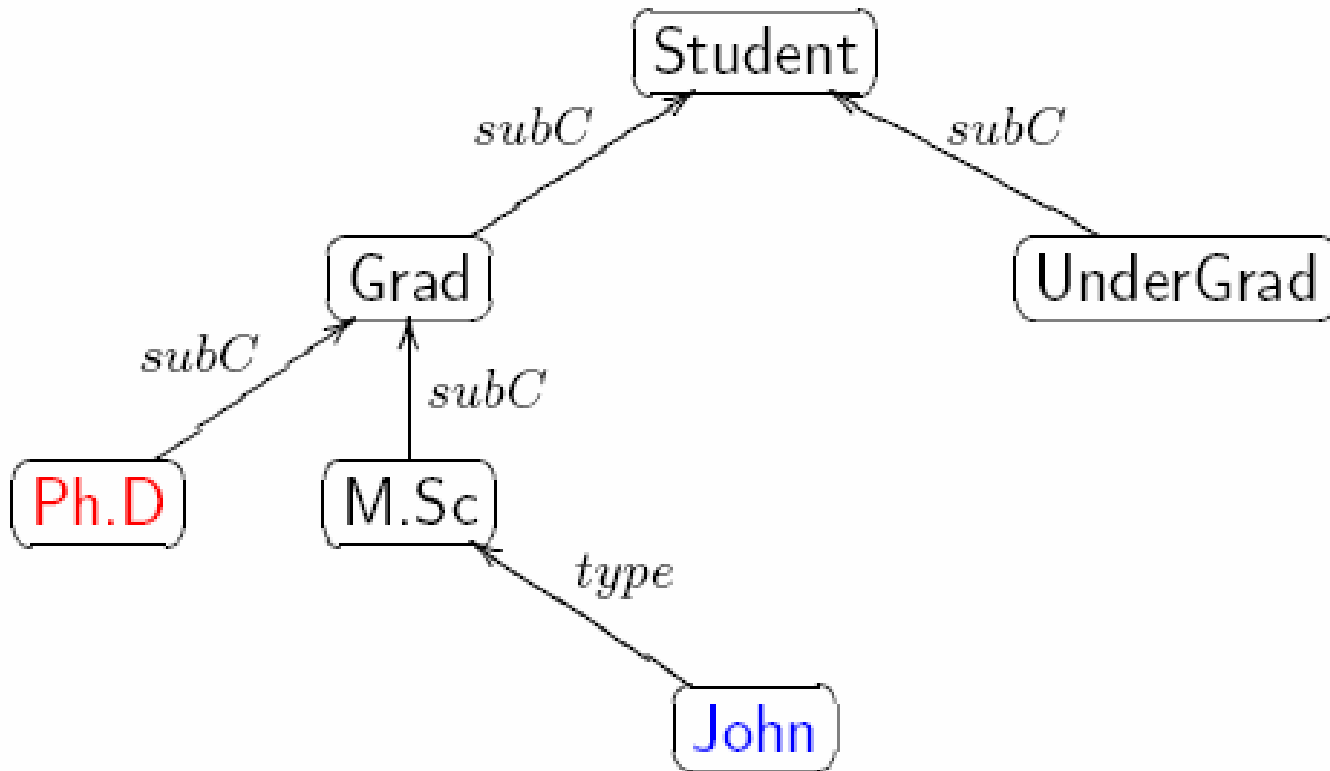
RDF Graph Representation



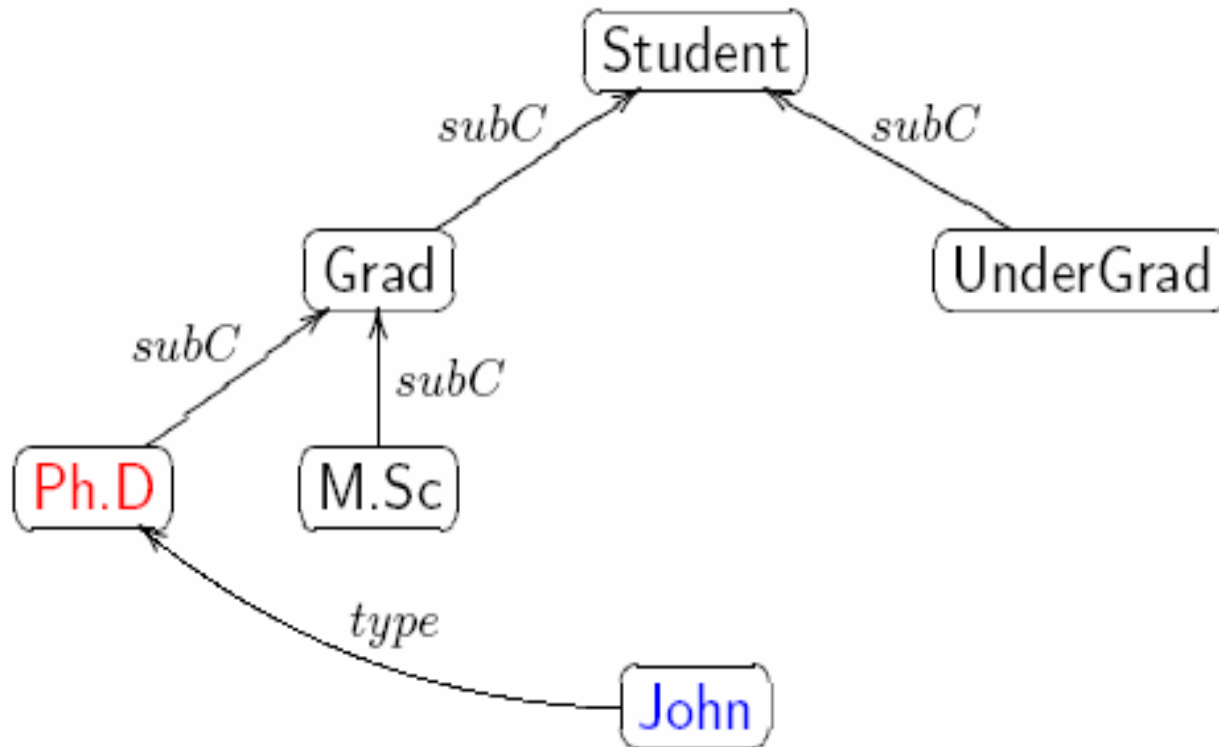
Updates in RDF Graphs



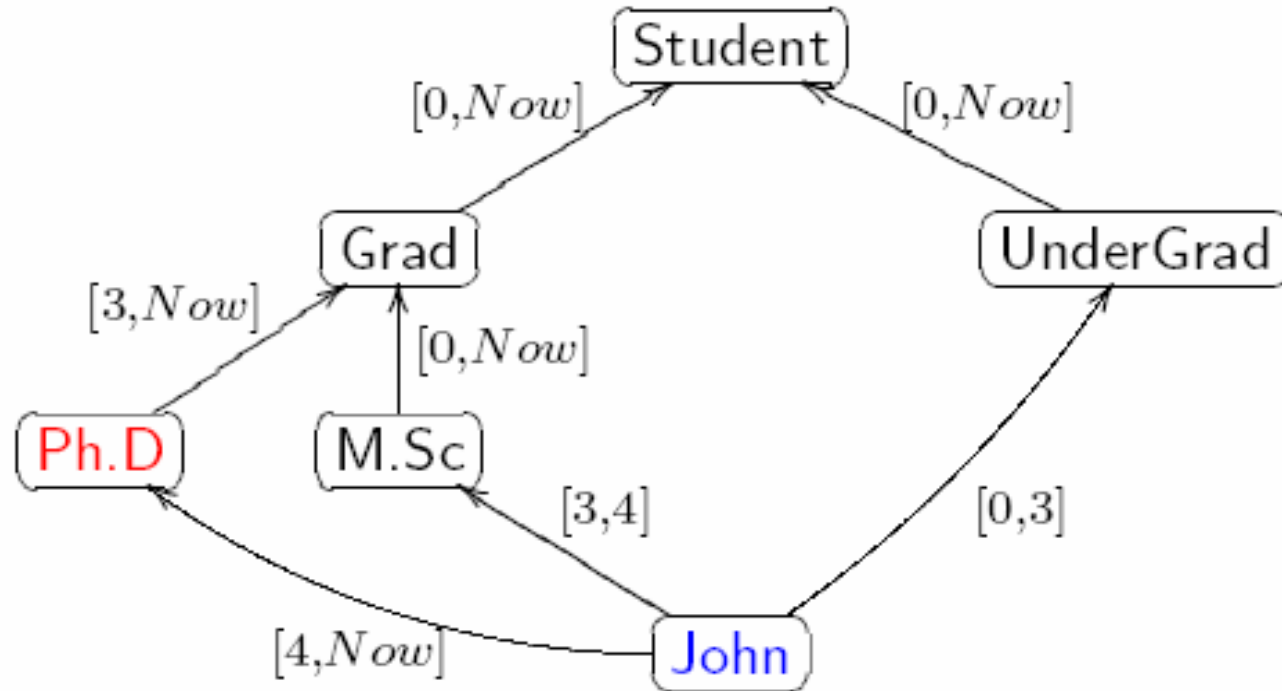
Updates in RDF Graphs



Updates in RDF Graphs



Temporal RDF Graph



General Issues

- Versioning versus Labeling
 - Label elements subject to change
 - Maintain a snapshot of each state of the graph

General Issues

- Versioning versus Labeling
 - Label elements subject to change
 - Maintain a snapshot of each state of the graph
- Time Points versus Time Intervals.

$$[4, 31] = [4] \cup [5] \cup \dots \cup [30] \cup [31]$$

General Issues

- Versioning versus Labeling
 - Label elements subject to change
 - Maintain a snapshot of each state of the graph
- Time Points versus Time Intervals.

$$[4, 31] = [4] \cup [5] \cup \dots \cup [30] \cup [31]$$

- Temporal Query Language
 - Point based (variables refer to point times)
 - Interval based (variables refer to intervals)

Definitions

Temporal Triple: an RDF triple with a temporal label,
e.g. $(a, b, c)[t]$

Temporal Graph: set of temporal triples

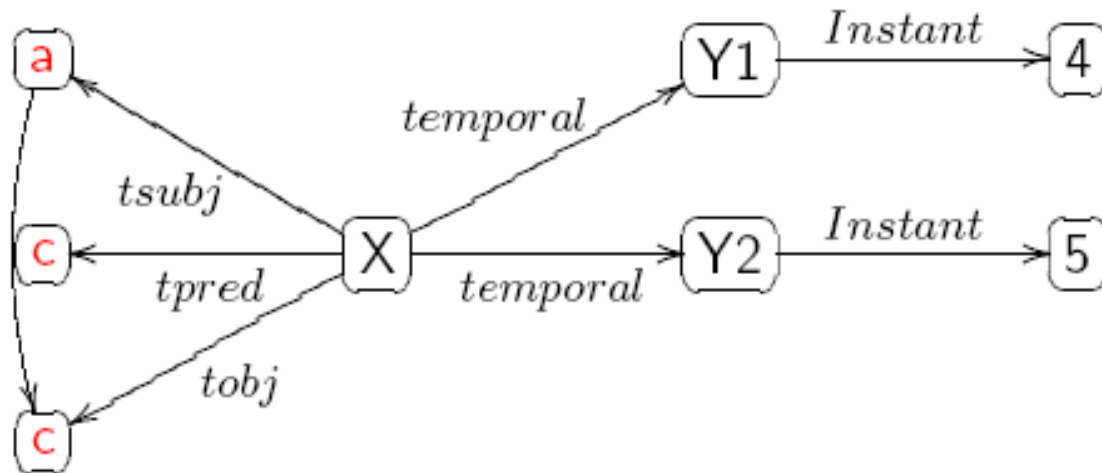
Snapshot of graph G at time t :

$$G(t) = \{(a, b, c) : (a, b, c)[t] \in G\}$$

Notion of temporal entailment $G_1 \models_{\tau} G_2$

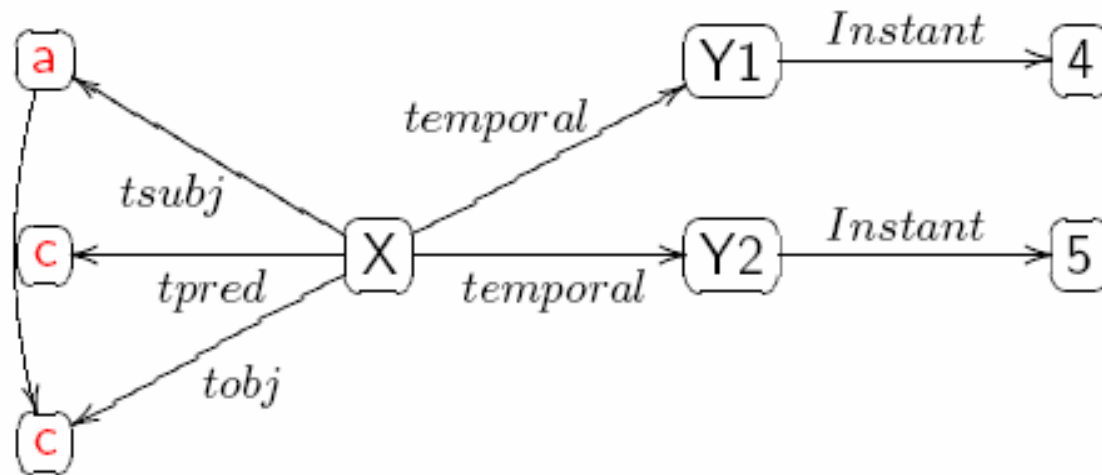
Syntax for (a,b,c)[4,5]

- Point version

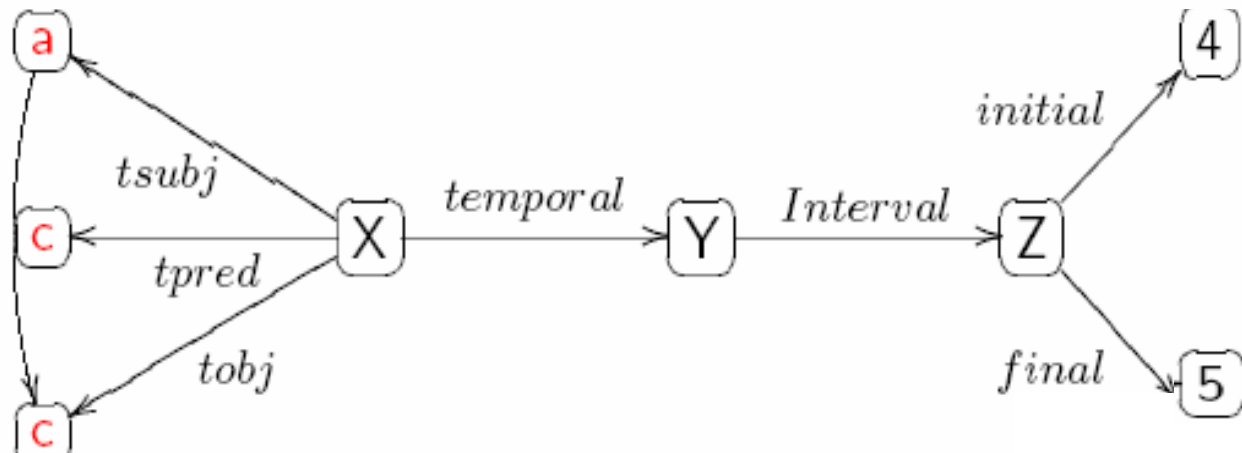


Syntax for (a,b,c)[4,5]

- Point version



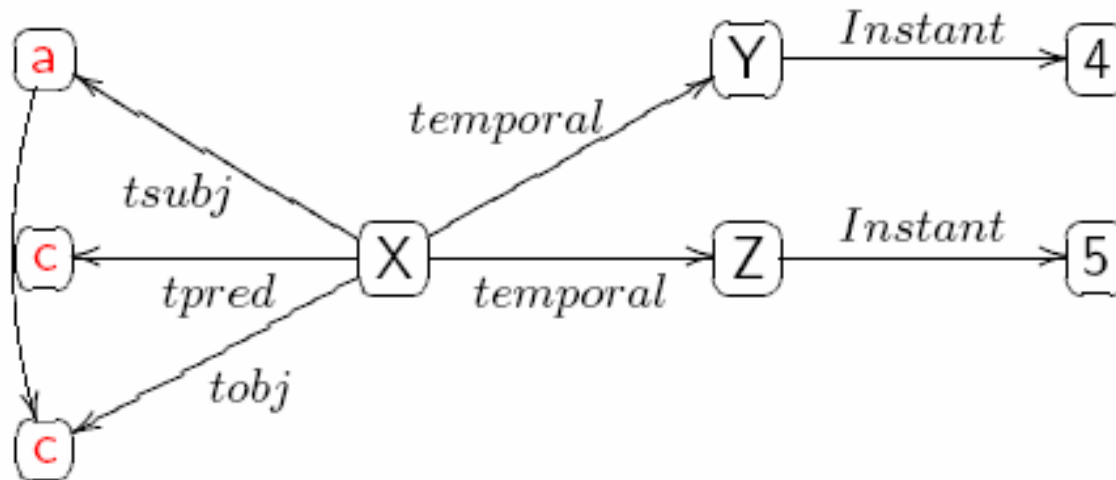
- Interval version



Syntax (rules)

Rule 1-2: Equivalence between point and interval versions

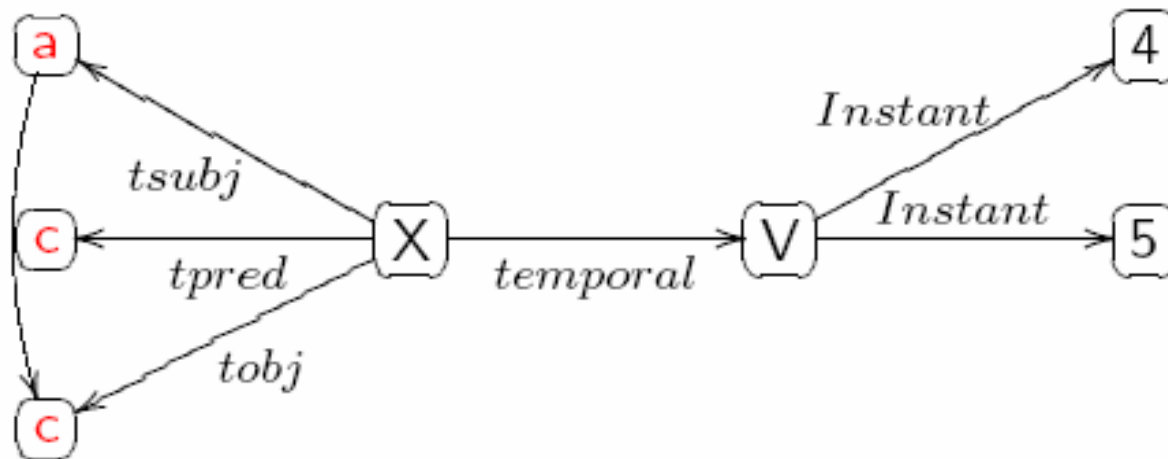
Rule 3: Normalization of point-version:



Syntax (rules) (cont.)

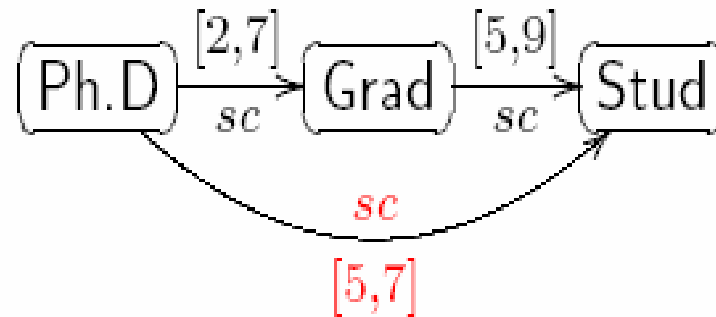
Rule 1-2: Equivalence between point and interval versions

Rule 3: Normalization of point-version:



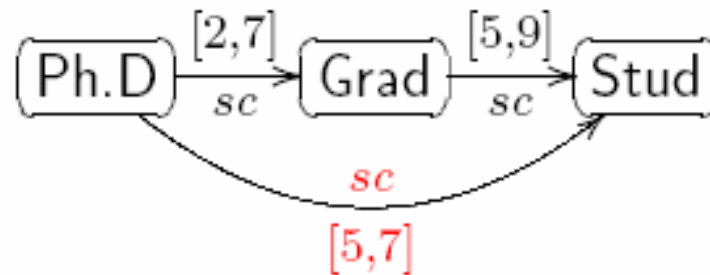
Temporal RDF Intrinsic Issues

- Notion of temporal Entailment \models_T

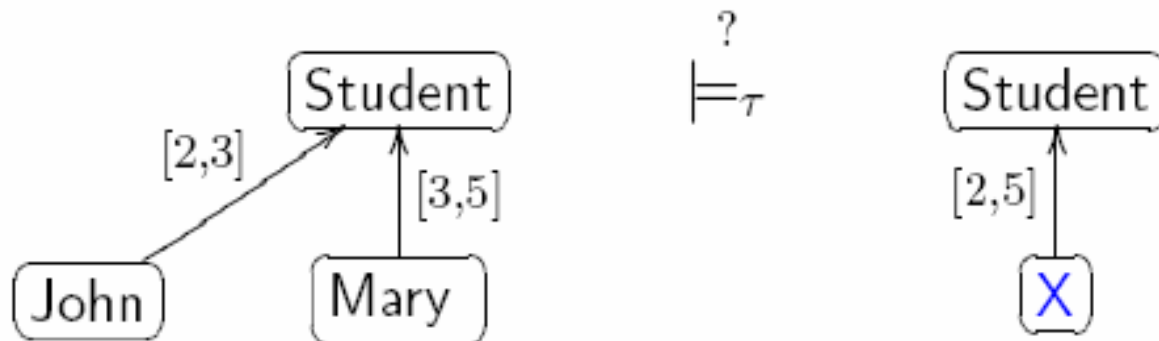


Temporal RDF Intrinsic Issues

- Notion of temporal Entailment \models_{τ}

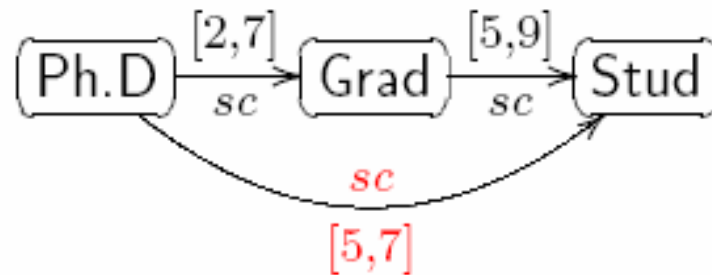


- Treatment of temporal Blank Nodes:

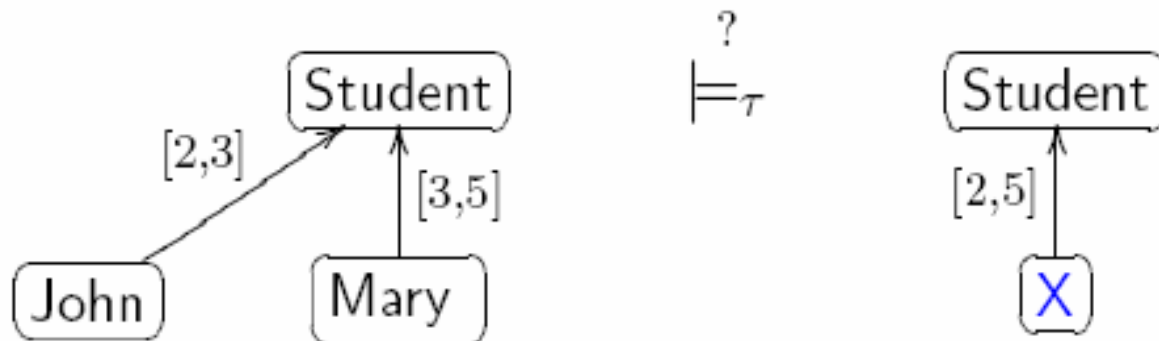


Temporal RDF Intrinsic Issues

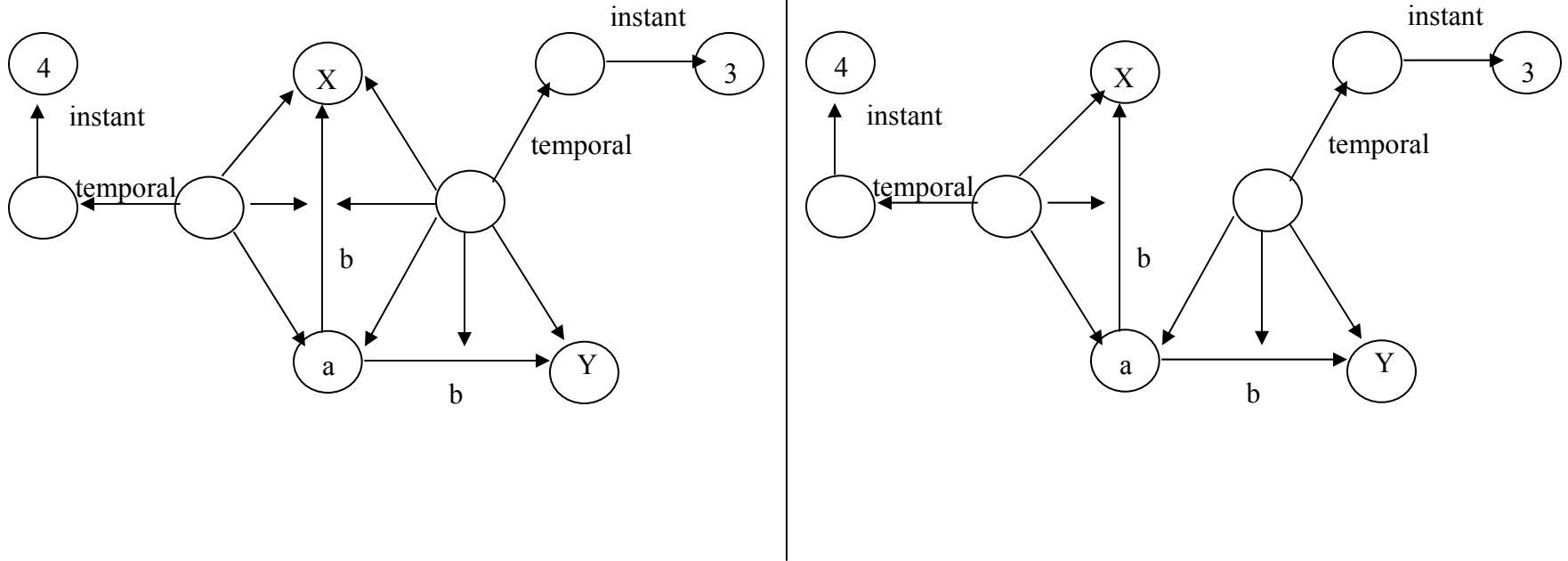
- Notion of temporal Entailment \models_{τ}



- Treatment of temporal Blank Nodes:



Temporal RDF Intrinsic Issues



Both Graphs have equivalent snapshots at any time t ; however, the graph on the LHS is not entailed by the Graph of the RHS. Note that for $t=3$, making $X = Y$ both Graphs are the same.

Semantics

Ground Case:

$$G_1 \models_{\tau} G_2 \text{ if for each } t, G_1(t) \models G_2(t)$$

Non Ground Case:

$G_1 \models_{\tau} G_2$ if there are ground instances $\mu_1(G_1)$ and $\mu_2(G_2)$ such that for each t :

$$\mu_1(G_1)(t) \models_{\tau} \mu_2(G_2)(t)$$

Proposition. For ground graphs, $G_1 \models_{\tau} G_2$ implies $G_1(t) \models G_2(t)$ for all times t .

Semantics

The **temporal closure** $tcl(G)$ is a maximal set of temporal triples G' such that:

- G' contains G
- G is equivalent to G'

Proposition.

$$G_1 \models_{\tau} G_2 \text{ iff } tcl(G_1) \models_{\tau} G_2$$

Proposition. Deciding if G' is the closure of G is DP-complete.

Querying Temporal RDF

- Proposal: Conjunctive fragment with
- interval and point variables
 - aggregate functions
 - constructor of graphs for answers

Querying Temporal RDF

Proposal: Conjunctive fragment with

- interval and point variables
 - aggregate functions
 - constructor of graphs for answers
- Students who have taken a Master course between year 2000 and the present time
 - Students taking Ph.D courses together and the time when this occurred
 - Time intervals when the IT Master program was offered
 - Students applying for jobs at time t after finishing their Ph.D program in no more than 4 years

Conclusion

1. Semantics for Temporal RDF graphs
2. Syntax to incorporate the framework into standard RDF
3. Sound and complete inference rules for temporal graphs
4. Complexity bounds showing temporal RDF preserves complexity of RDF
5. Sketch of Temporal RDF query language