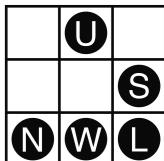


# A Query Model for Ontology Based Event Processing on RDF Streams

DEBS - 26.06.2019

Darmstadt, Germany, Europe, Earth, Milky Way,  
Universe.....42

Pieter Bonte, Riccardo Tommasini,  
Filip De Turck, Femke Ongenae, Emanuele Della Valle



# Authors



# That Asks Complex Questions



Who is driving event in the news right now?

Is public transportation where the people are?

Who are the best agents to route all these unexpected contacts about the tariff plan launched yesterday?

What is the expected time to failure when a turbine's barring starts to vibrate as detected in the last 10 minutes?

# Enrich the Streams



# Crossing the Streams



# Issues

- Need to manually produced the code to query databases that produced the complex results
- Need to manually sources and hard-code the logic for each application adapting to different databases
- Need to maintain such complex systems once you have it!

Scalability

Alert

## STREAM REASONING RESEARCH QUESTION

Is it possible to **make sense** in **real time** of **multiple, heterogeneous, gigantic** and inevitably **noisy** and **incomplete data streams** in order to support the **decision** processes of extremely large numbers of concurrent users?

E. Della Valle, S. Ceri, F. van Harmelen & H. Stuckenschmidt, 2010

# Stream Reasoning

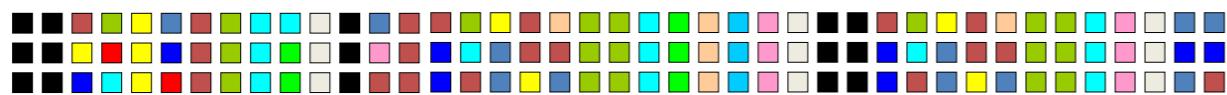
Is there a **primary cool** colour followed by a **secondary warm** one in the last minute

(, 13), (, 8), (, 8)

Which are the top-2 most frequent **cool** colours in the last minute?

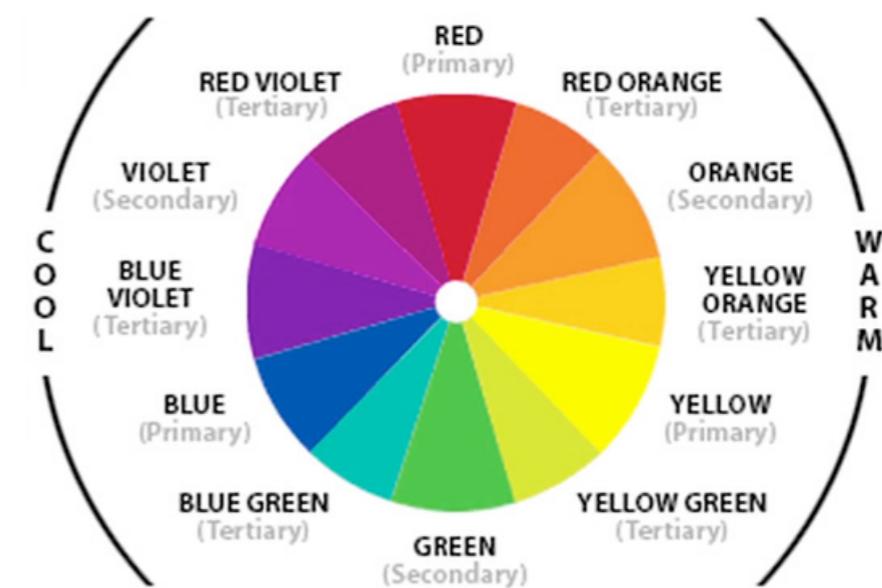
yes,  followed by 

1 minute wide window

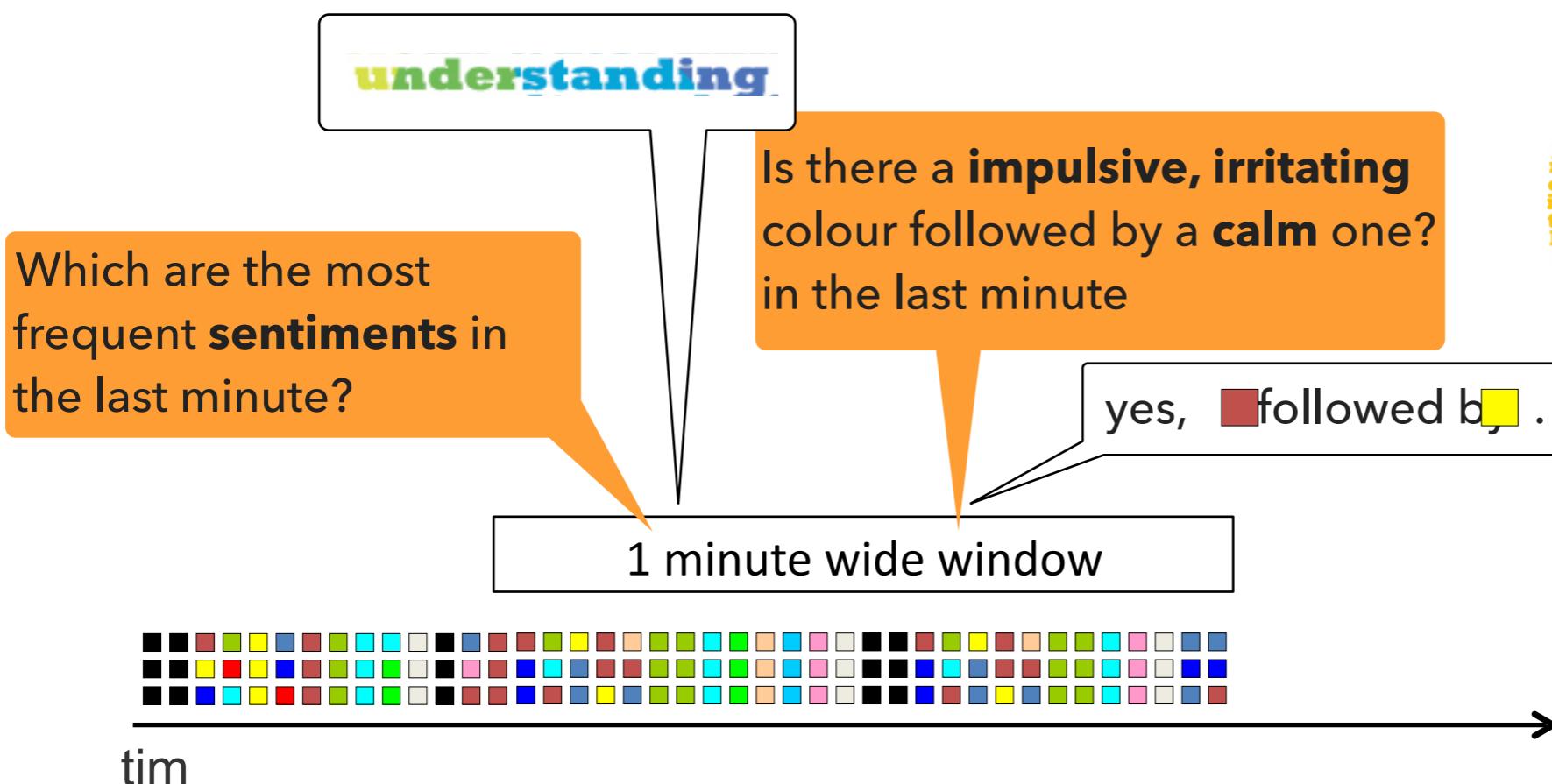


tim

An ontology of



# Stream Reasoning



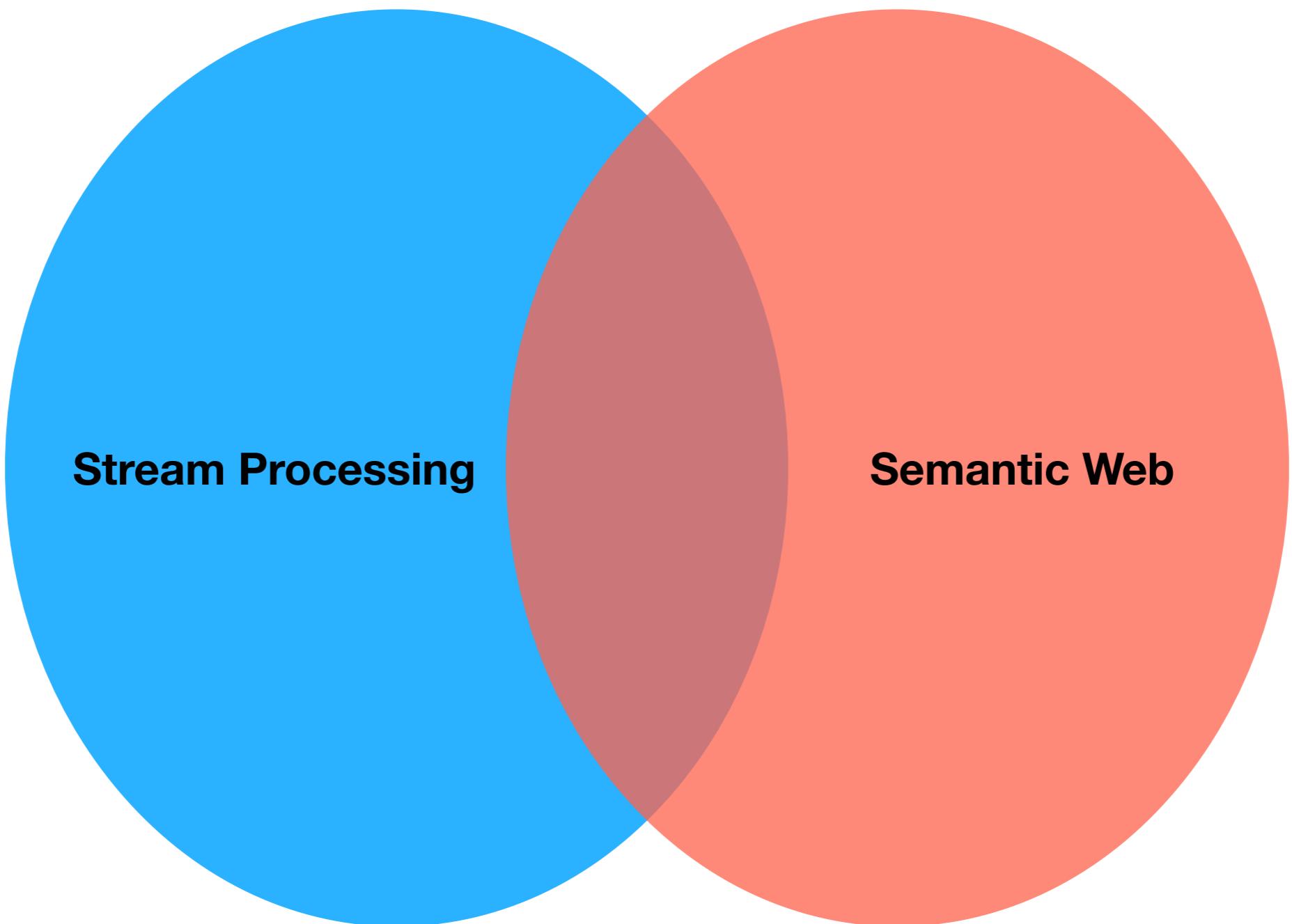
A different ontology (of colours)

b a l a n c e  
g r o w t h v i b r a n t c o l o u r  
c o n t r o v e r s y f l a m b o y a n t  
e n e r g y a c t i v i t y a p p e t i t e s o c i a l-  
i z a t i o n b l o o d h e a t v i g o r p a s s i o n a t e  
i n t e n s e n e r c e l o v e d a n g e r e x c i t i n g  
s t r o n g h i r r i t a t i n g l i p s h e a r t s s e x y r o-  
m a n c e s e n s u a l i t y i m p u l s i v e l e a d e r s h i p  
c o u r a g e c o m p e t e n c e i n d e p e n d e n c e o r-  
g a n i z a t i o n s e l f -m o t i v a t i o n s p i r i t u a l i t y p l ea-  
s u r e v i t a l i t y w i l l t o w i n s u r v i v a l i n s t i n c t i n-  
t u i t i o n e n t r e p r e n e u r i a l d e s i r e f i r e s t i m u l a t i o n  
j o y r a g e s u n s h i n e t r o p i c a l e n t h u s i a s m f a s-  
c i n a t i o n h a p p i n e s s c r e a t i v i t y a t t r a c t i o n s u c c e s s  
c i t r u s e n d u r a n c e i l u m i n a t i o n w i s d o m w e a l t h  
i n t e l l e c t l o y a l t y f r e s h n e s s g r o w t h h a r m o n y f e-  
r i t y s a f e t y m o n e y v i s i o n e x p e r i e n c e n o v i c e  
h o p e n a t u r e f i n a n c e a m b i t i o n g r e e d j e a l o u s y  
h e a l i n g p r o t e c t i o n p e a c e s k y s e a d e p t h t r u s t  
c o n f i d e n c e f a i t h t r u t h h e a v e n m i n d t r a n q u i-  
l i t y c a l m s i n c e r i t y c l e a n w a t e r m i n e r a l p r eci-  
s i o n e x p e r t i s e u n d e r s t a n d i n g s o f t n e s s  
k n o w l e d g e p o w e r r o y a l t y n o b i l i t y l u x u r y  
e x r a v a g a n c e d i g n i t y m y s t e r y m a g i c a r-  
t i f i c a l n o s t a l g i a g l o o m f r u s t r a t i o n l i g h t  
g o o d n e s s i n n o c e n c e p u r i t y p e r f e c-  
t i o n p o s i t i v e b e g i n n i n g c o o l s i-  
m p l i c i t y c h a r i t y a n g e l s s t e r i l i t y e l-  
e g a n c e f o r m a l i t y e v i l f e a r  
u n k n o w n f e e l i n g a u t h o-  
r i t y p r e s t i g e g r i e f  
h a r m o n y

# Stream Reasoning

(in practice)

- Addresses data variety employing semantic technologies,
  - RDF, SPARQL, and reasoning methods (materialisation, or query rewriting)
- Addresses data velocity employing stream processing technologies
  - Data Stream Management Systems, Complex Event Processing, and. Big Streaming Systems (check out our tutorial)



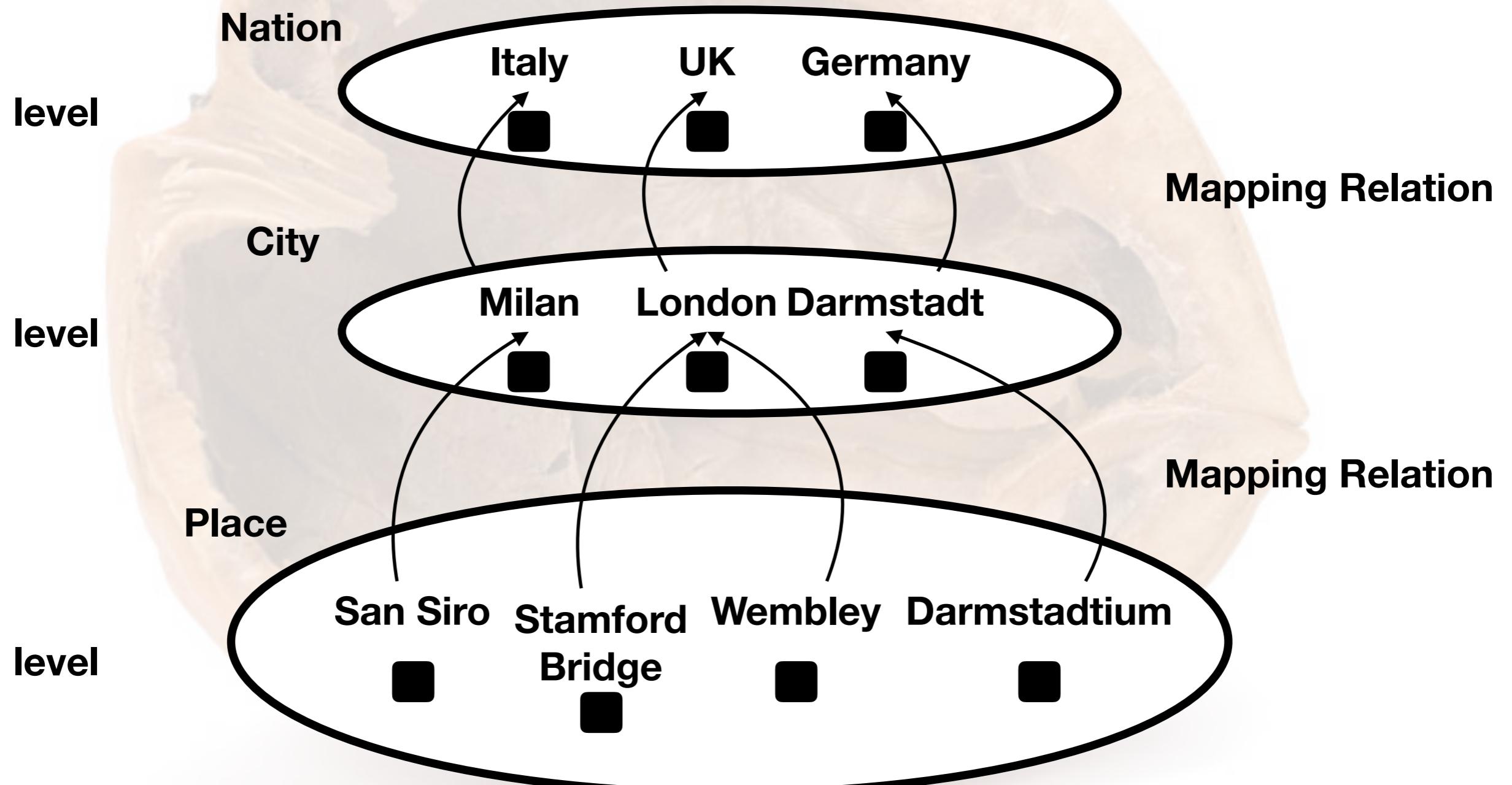
# Research Question

- (Mezzo) Leverage advanced language features in streaming systems to speed up throughput of stream reasoning problems
  - (micro) leveraging inheritance in streaming languages to speed up hierarchical stream reasoning

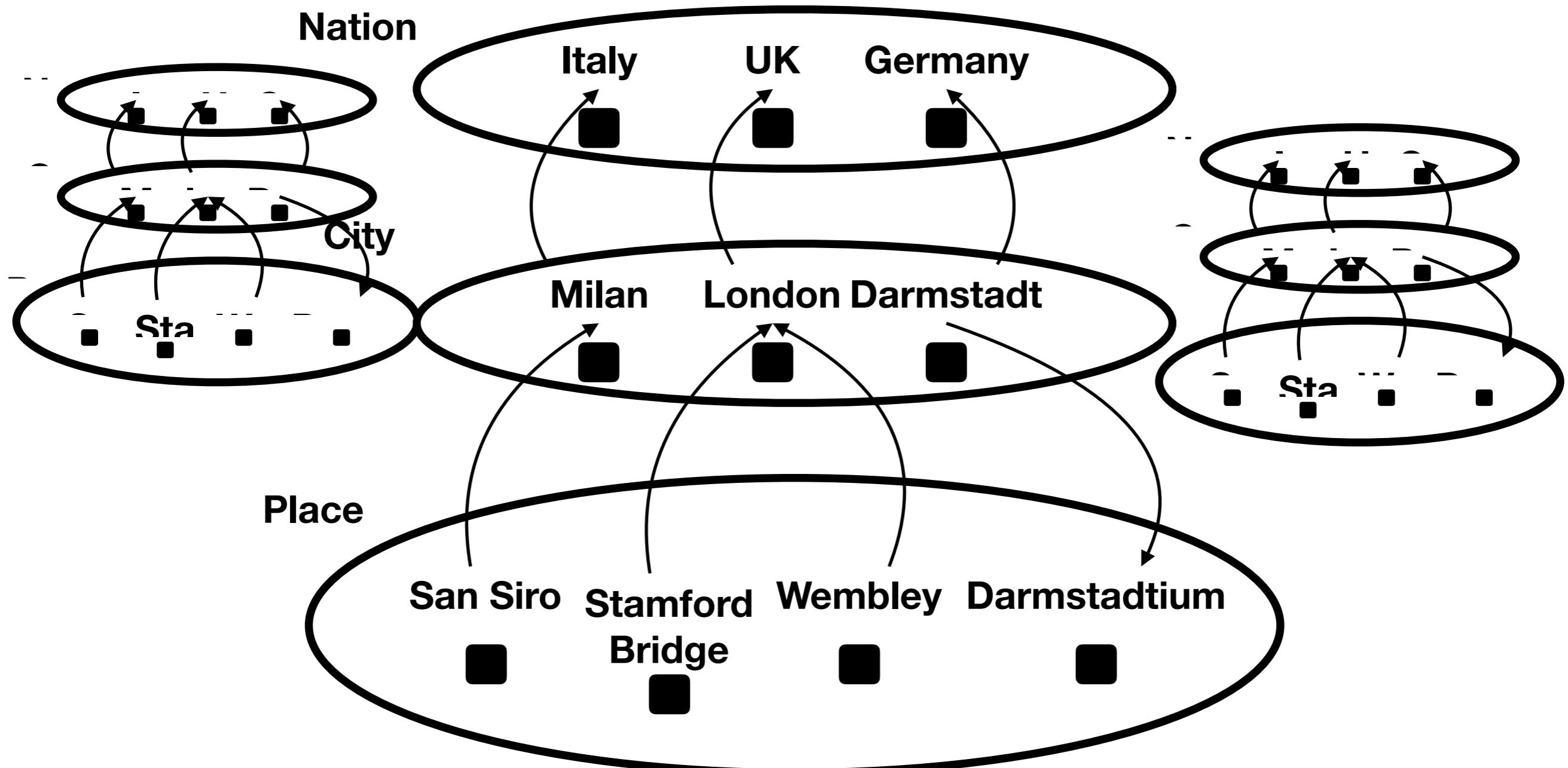
# Contributions

- C-TRA, an extension of Taxonomy-Relational Algebra [20] to the continuous semantics
- Alignment of query answering in CTRA to query answering in (continuous) SPARQL under RDFS entailment regime
- CSPRITE: two algorithms to efficient hierarchical reasoning of RDF streams.

# h-domain



# taxonomy



# T-Schema & T-Relation

Type table

	<b>Subject</b>	<b>obs:l2</b>
$t_a$	obs_x	WeatherObservation
$t_b$	obs_y	GeometryObservation

$\varepsilon^{obs:l1}$   
 $\xrightarrow{obs:l2}$

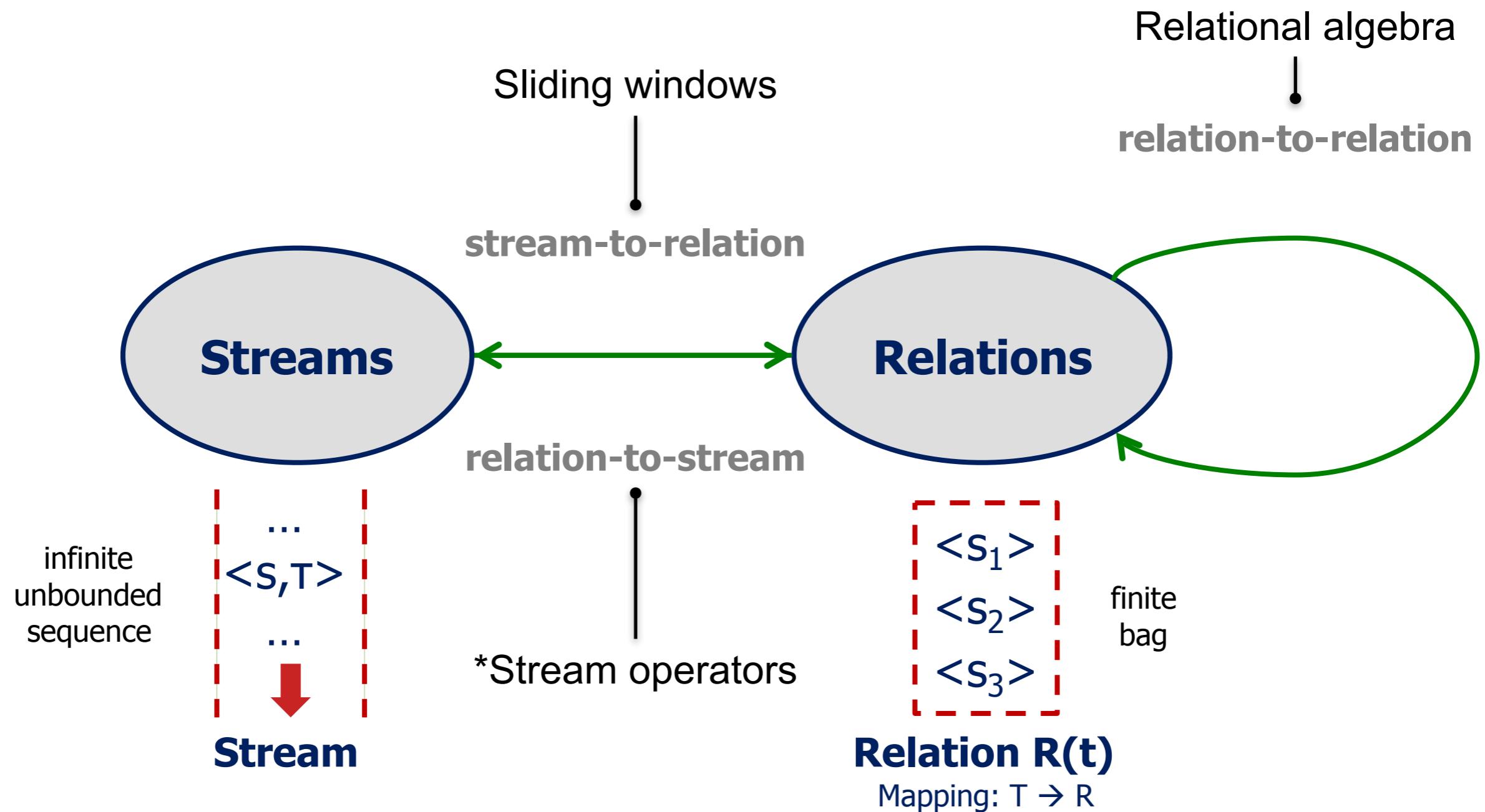
	<b>Subject</b>	<b>obs:l2</b>	<b>obs:l1</b>
$t_a$	obs_x	WeatherObservation	Observation
$t_b$	obs_y	GeometryObservation	Observation

a)

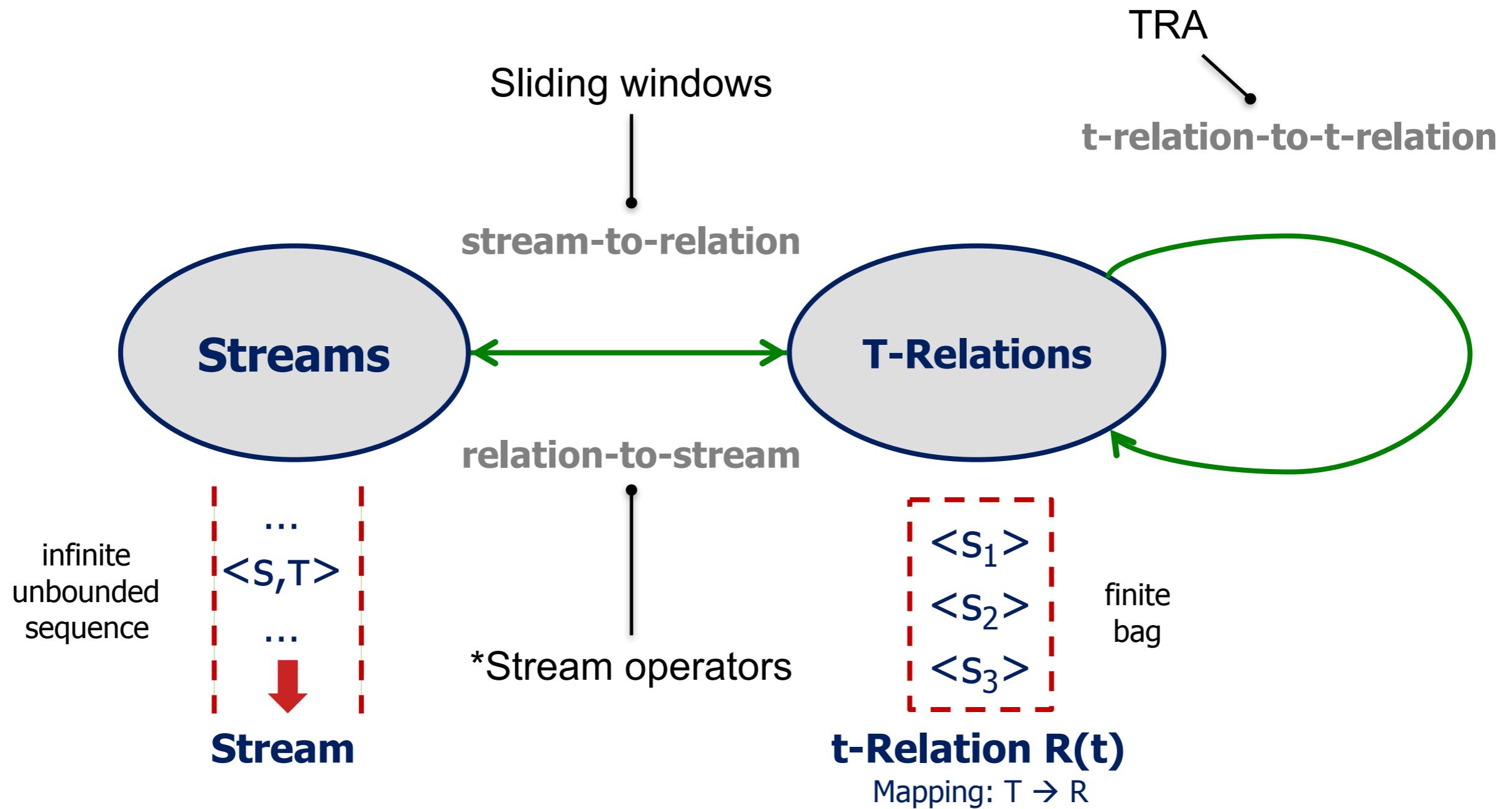
b)

upward extension

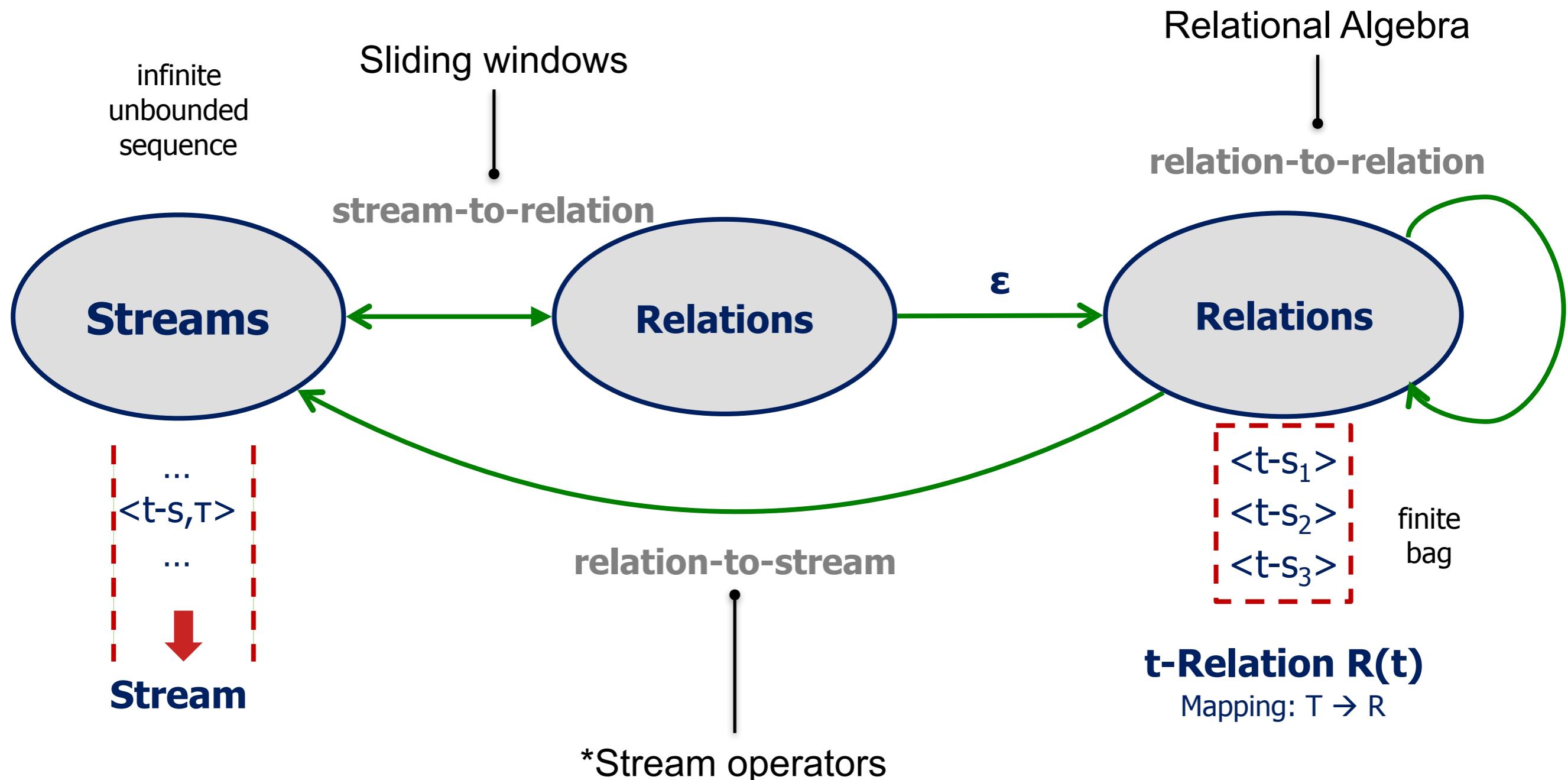
# The CQL model

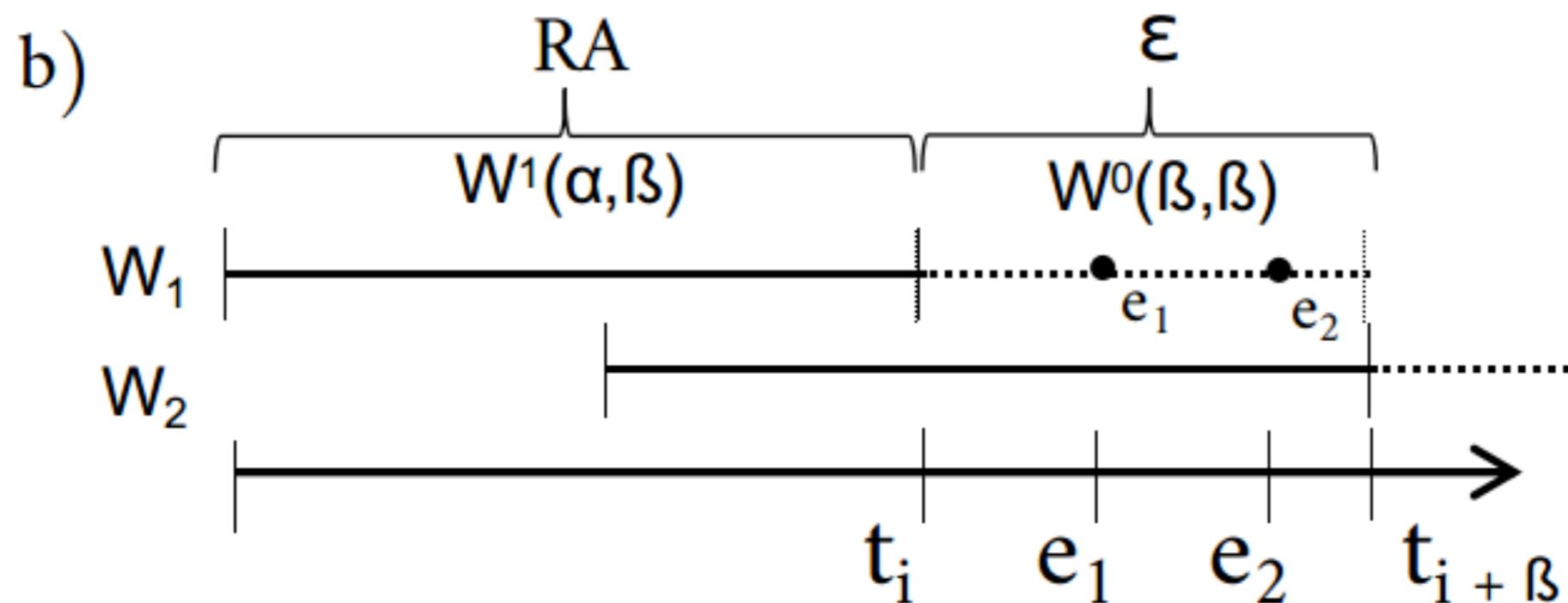
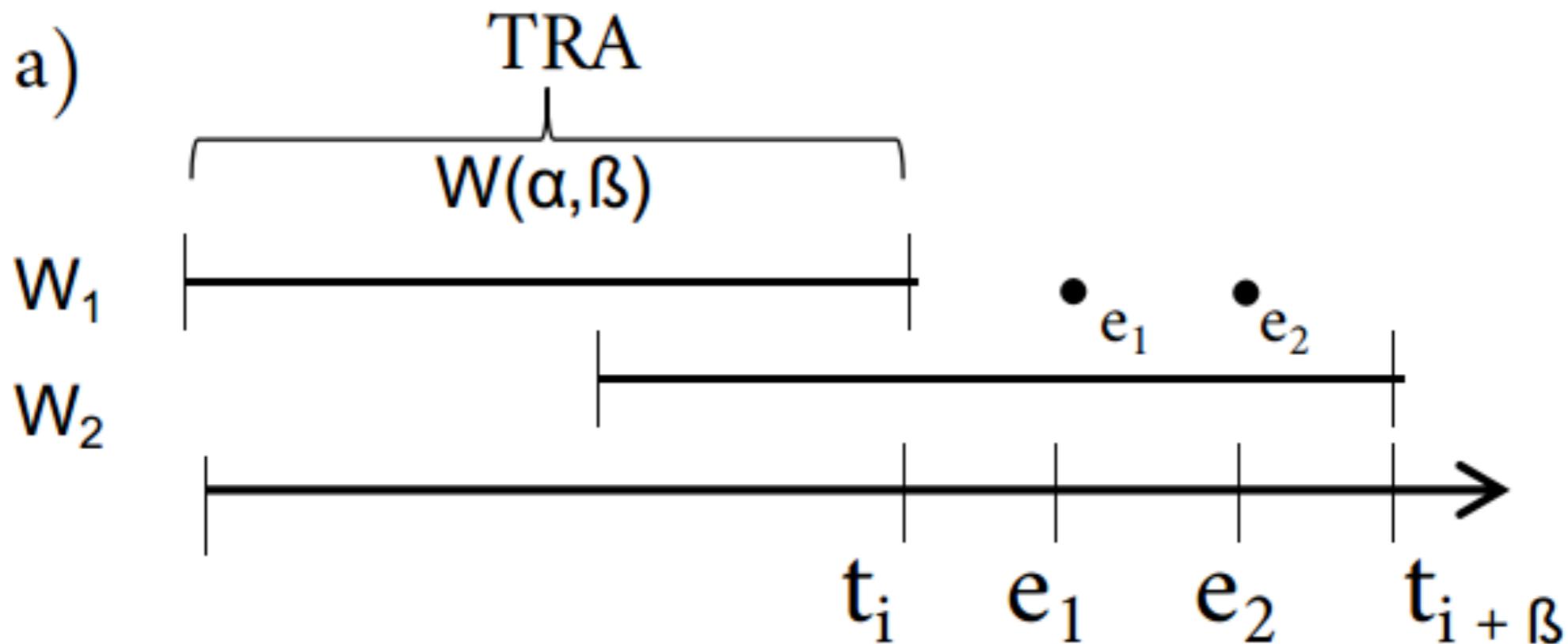


# The CTQL model

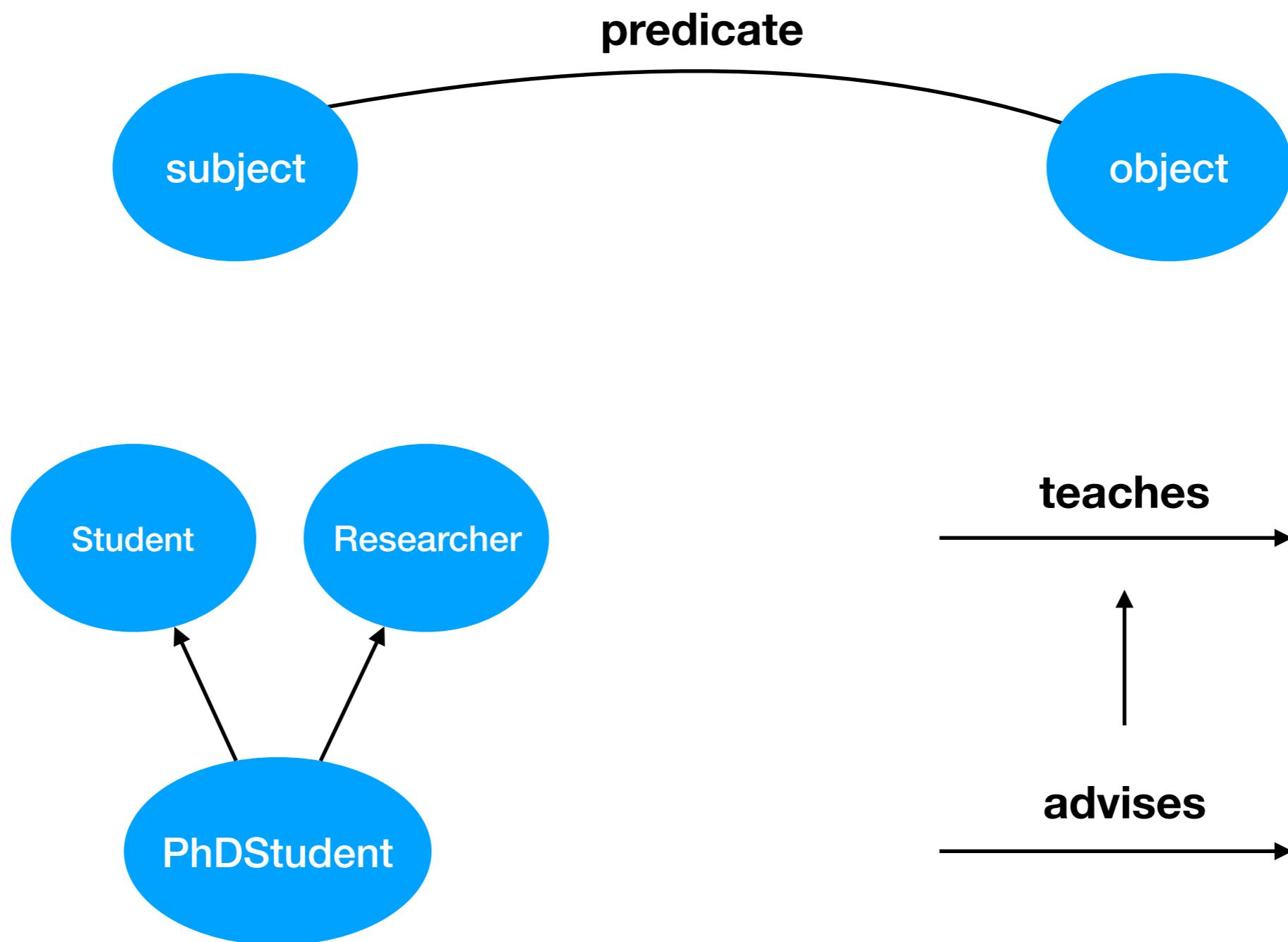


# The CTQL model

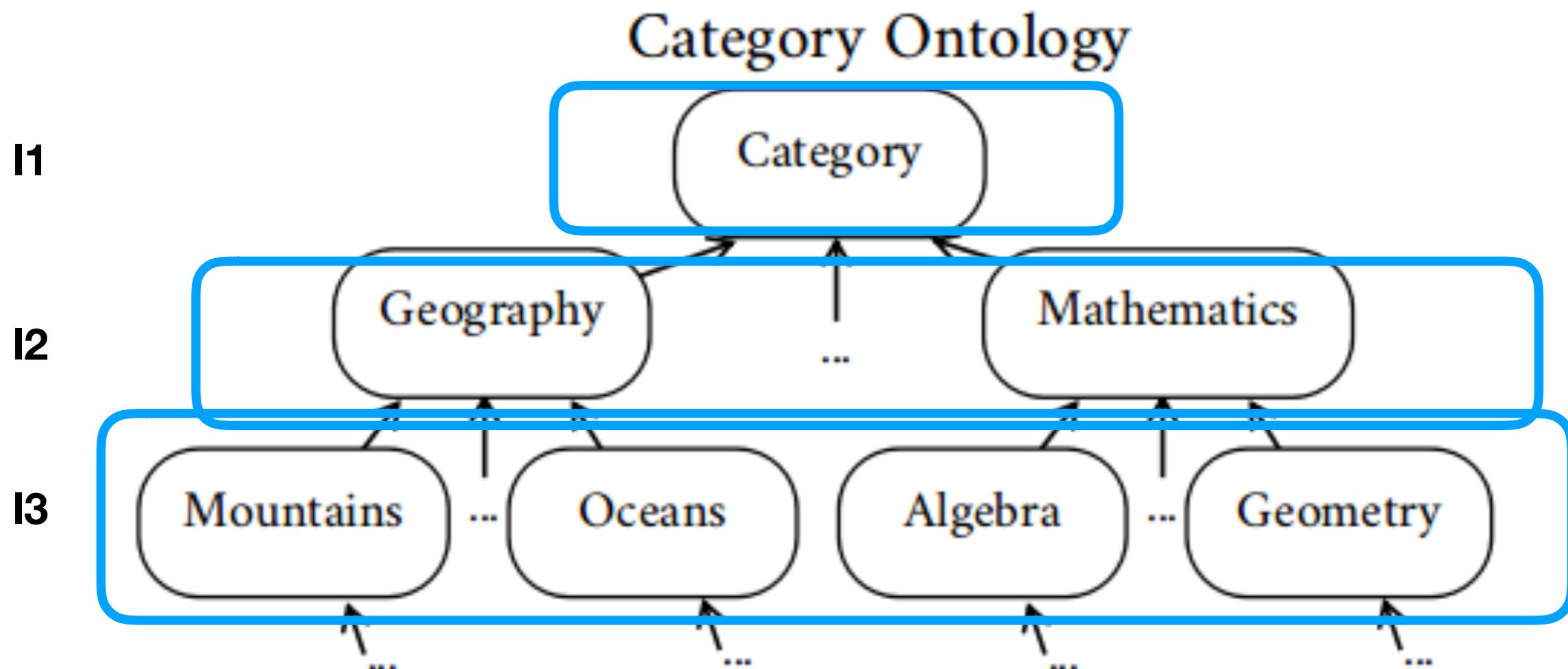




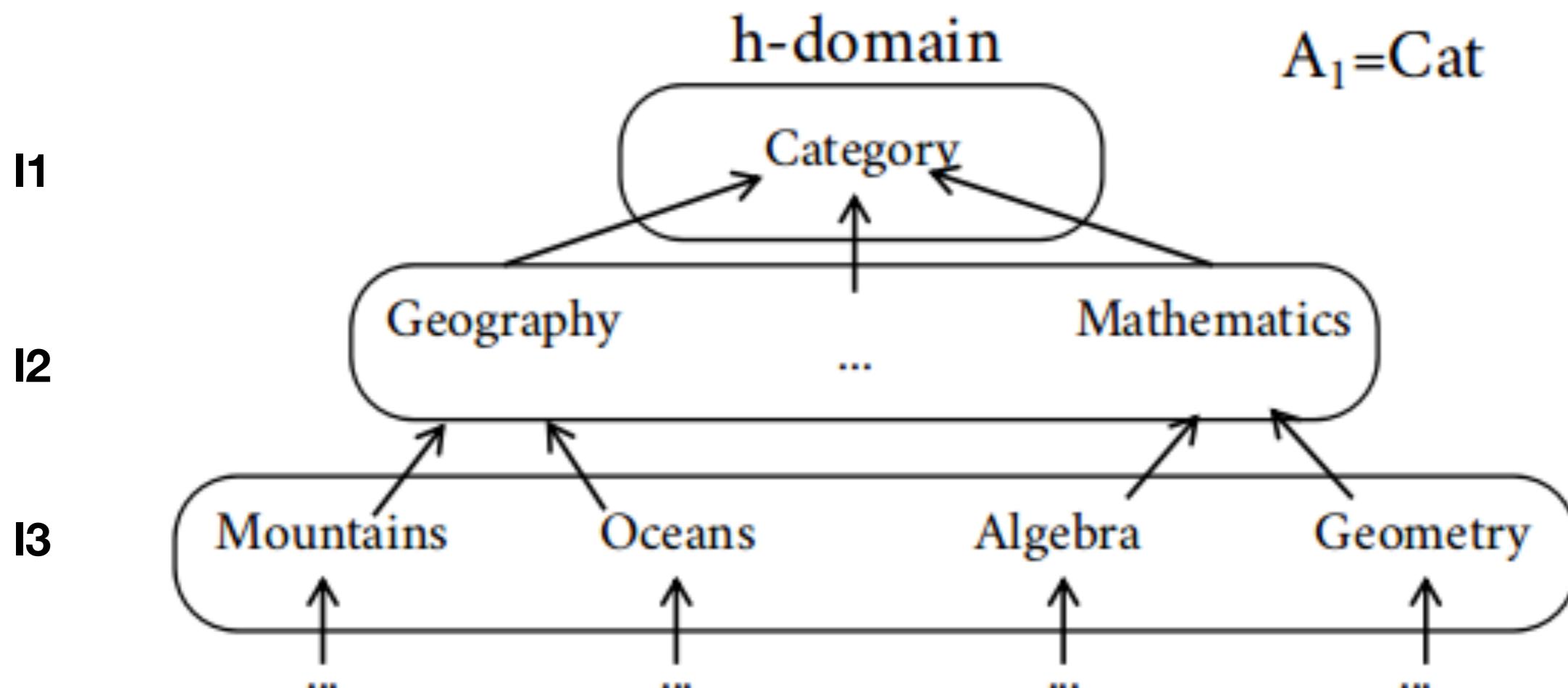
# RDF and RDFS Ontologies



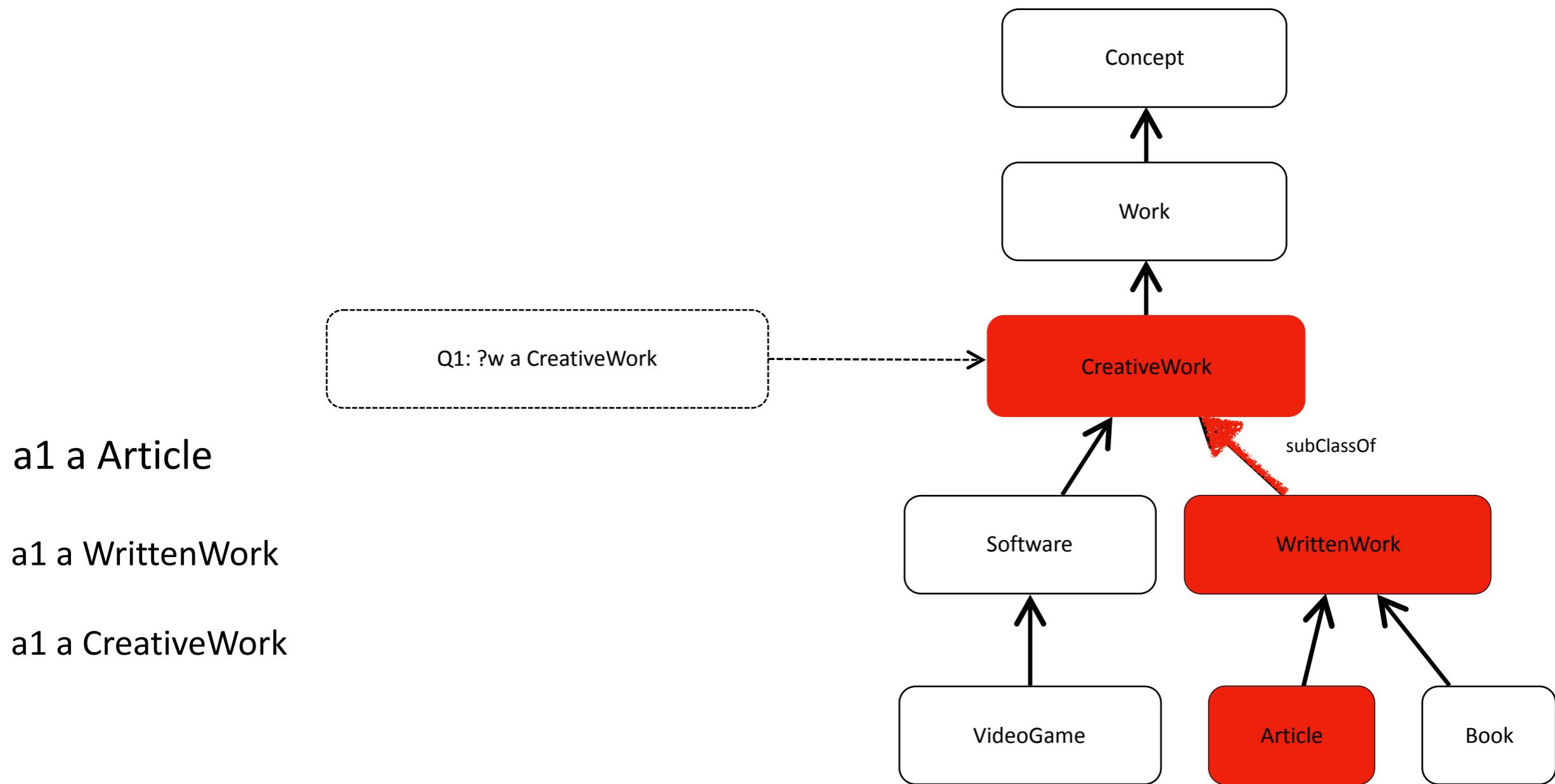
# TRA and RDFS Ontologies



# TRA and RDFS Ontologies

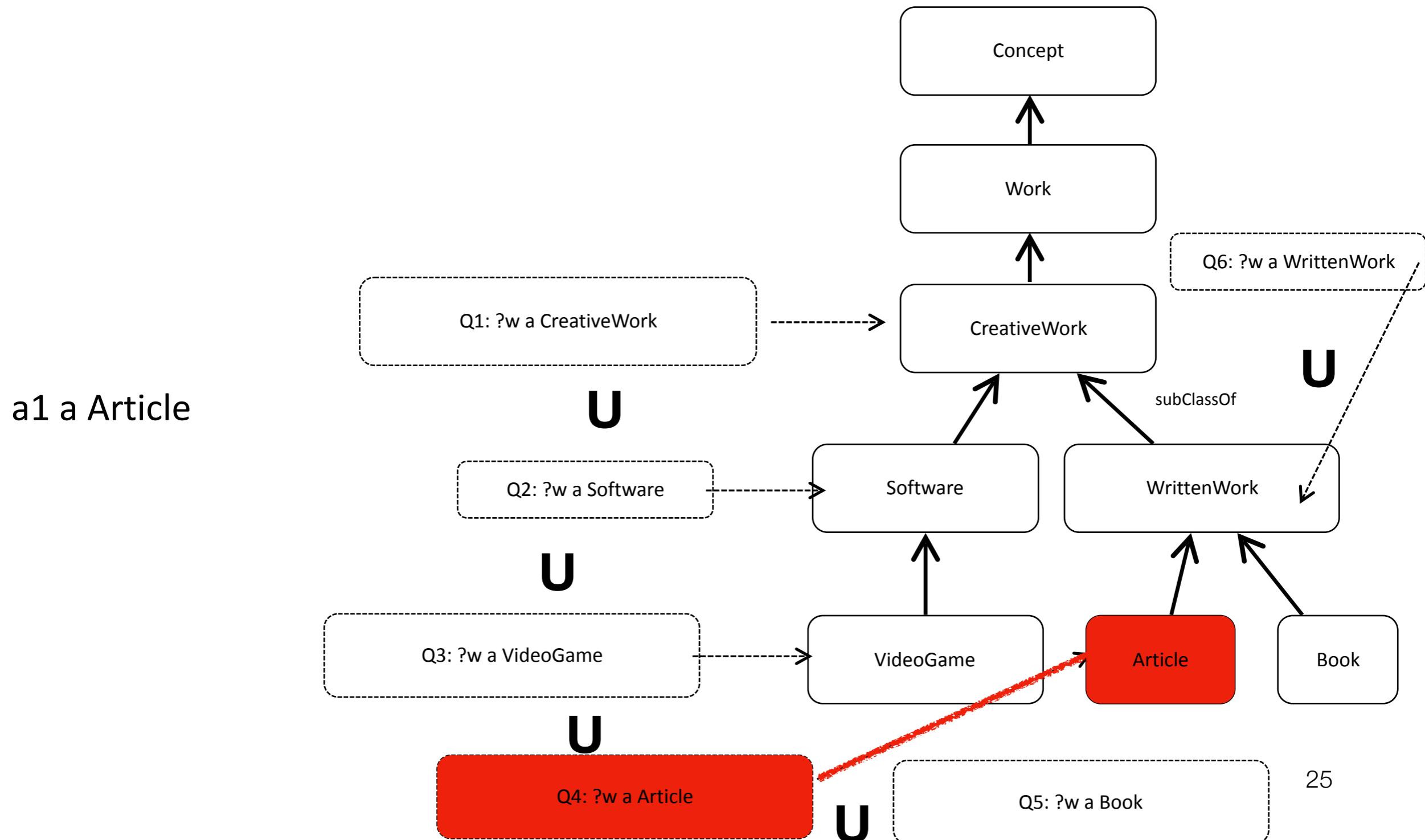


# RDF Stream Processing UE



**Materialisation!**

# RDF Stream Processing UE



# C-Sprite

---

**Algorithm 1** Query registering

---

**Precondition:**  $Q$  a collections of queries, each interested in one or more types.

```
1  $H \leftarrow ConvertToHierarchy(O)$   $\triangleright$  Stores parents for each class in the Ontology  $O$ 
2 function PREPAREHIERARCHY( $H, Q$ )
3    $H' \leftarrow []$ 
4   for  $q \in Q$  do
5     for (concept, parents)  $\in H$  do
6       if  $q \in parents$  then
7          $H'[concept].append(q)$ 
8       end if
9     end for
10    end for
11    return  $H'$ 
12 end function
```

---

---

**Algorithm 2** Calculate the query matches on a hierarchical level

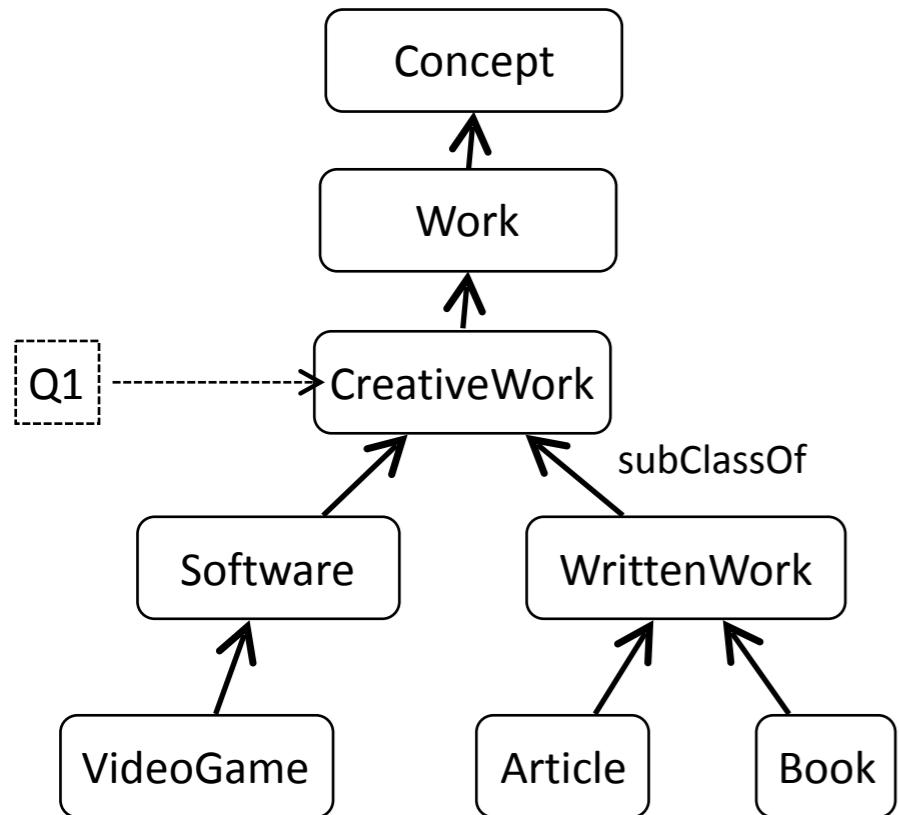
---

**Precondition:**  $Q$  a collections of queries, each interested in one or more types.

```
1  $H \leftarrow ConvertToHierarchy(O)$   $\triangleright$  Stores parents for each class in the Ontology  $O$ 
  (preprocessing step)
2  $H' \leftarrow PrepareHierarchy(H, Q)$   $\triangleright$  (preprocessing step)
3  $triple \leftarrow ClassAssertion(type, subject)$ 
4 function CHECKHIERARCHYMATCH( $H', triple$ )
5    $QueryMatches \leftarrow H'(types(triple))$   $\triangleright$  types extracts the type assertions of a
  triple
6   return  $QueryMatches$ 
7 end function
```

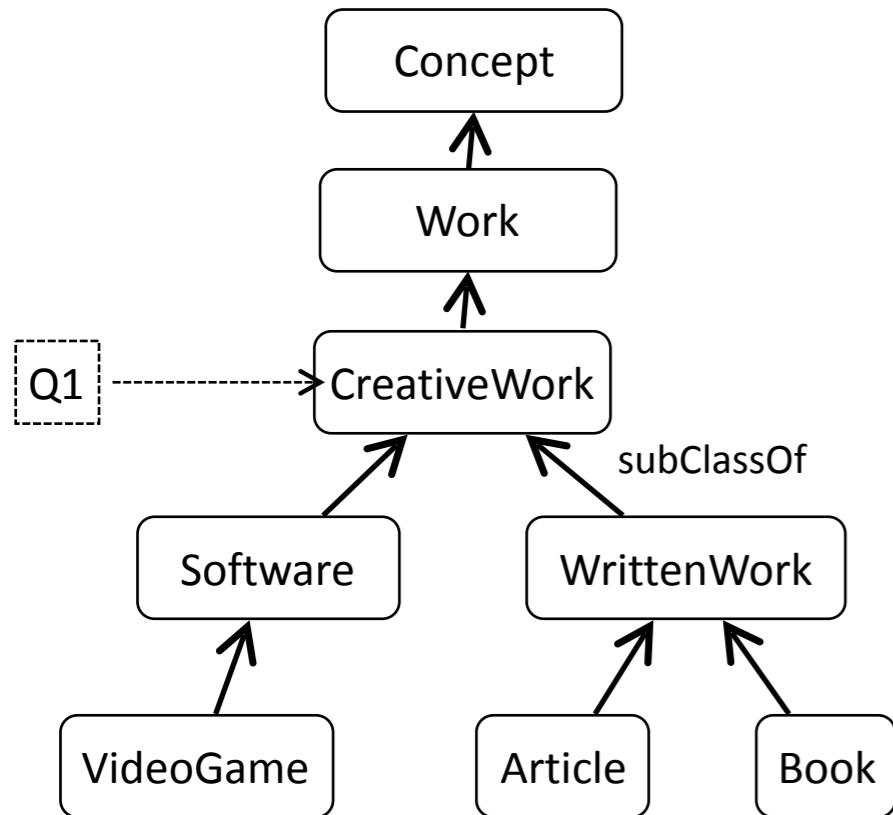
---

# C-Sprite



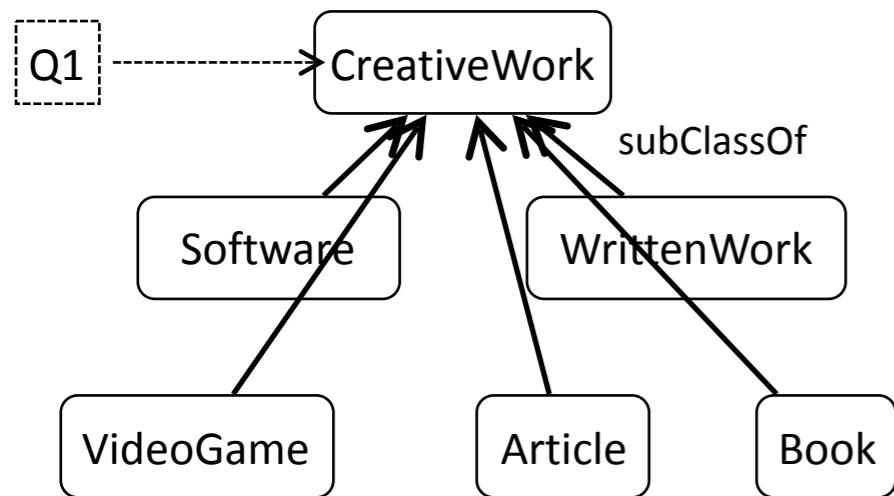
Concept: [Concept]  
Work: [Work, Concept]  
CreativeWork: [CreativeWork, Work, Concept]  
WrittenWork: [WrittenWork, CreativeWork, Work, Concept]  
Article: [Article, WrittenWork, CreativeWork, Work, Concept]  
Book: [Book, WrittenWork, CreativeWork, Work, Concept]  
Software: [Software, CreativeWork, Work, Concept]  
VideoGame: [VideoGame, Software, CreativeWork, Work, Concept]

# C-Sprite



Concept: [Concept]  
Work: [Work, Concept]  
CreativeWork: [CreativeWork, Work, Concept]  
WrittenWork: [WrittenWork, CreativeWork, Work, Concept]  
Article: [Article, WrittenWork, CreativeWork, Work, Concept]  
Book: [Book, WrittenWork, CreativeWork, Work, Concept]  
Software: [Software, CreativeWork, Work, Concept]  
VideoGame: [VideoGame, Software, CreativeWork, Work, Concept]

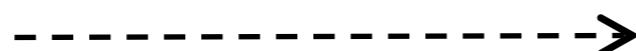
# C-Sprite



Concept: [Concept]  
Work: [Work, Concept]  
CreativeWork: [CreativeWork, Work, Concept]  
WrittenWork: [WrittenWork, CreativeWork, Work, Concept]  
Article: [Article, WrittenWork, CreativeWork, Work, Concept]  
Book: [Book, WrittenWork, CreativeWork, Work, Concept]  
Software: [Software, CreativeWork, Work, Concept]  
VideoGame: [VideoGame, Software, CreativeWork, Work, Concept]

# Evaluation

Change stream

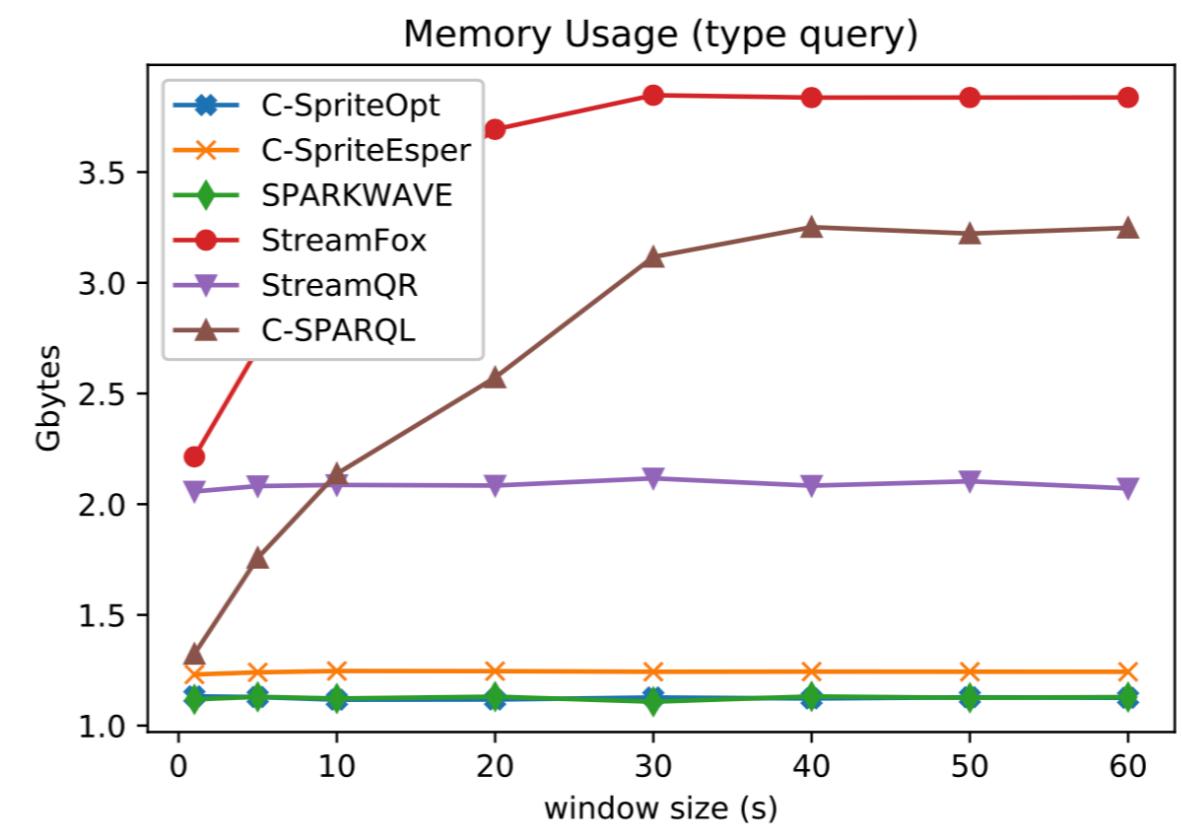
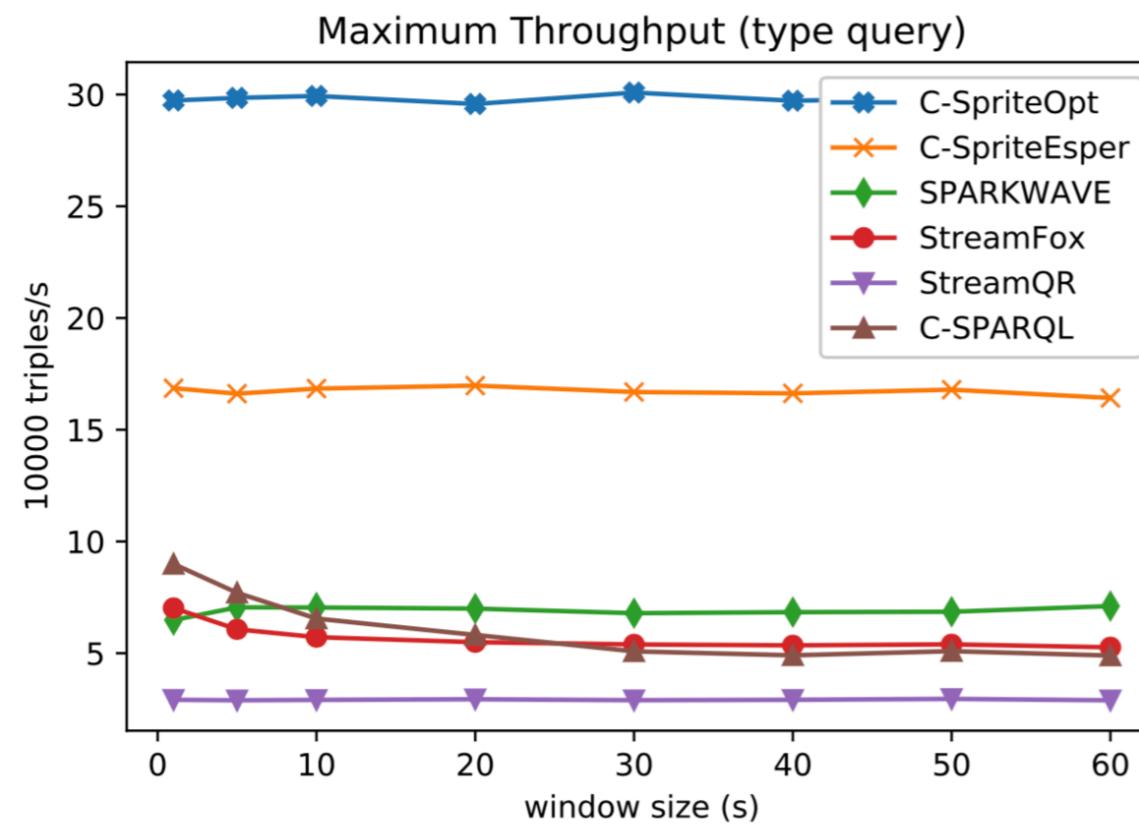


Q1: ?w a CreativeWork

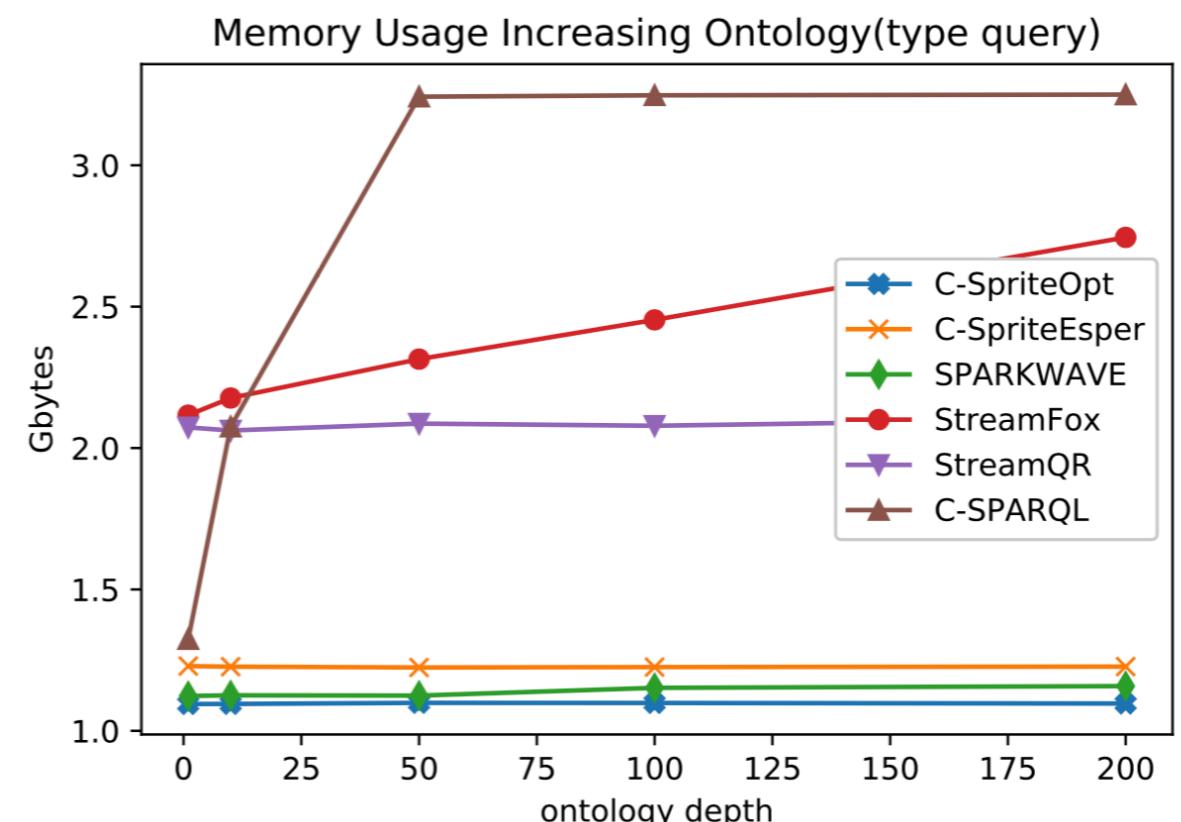
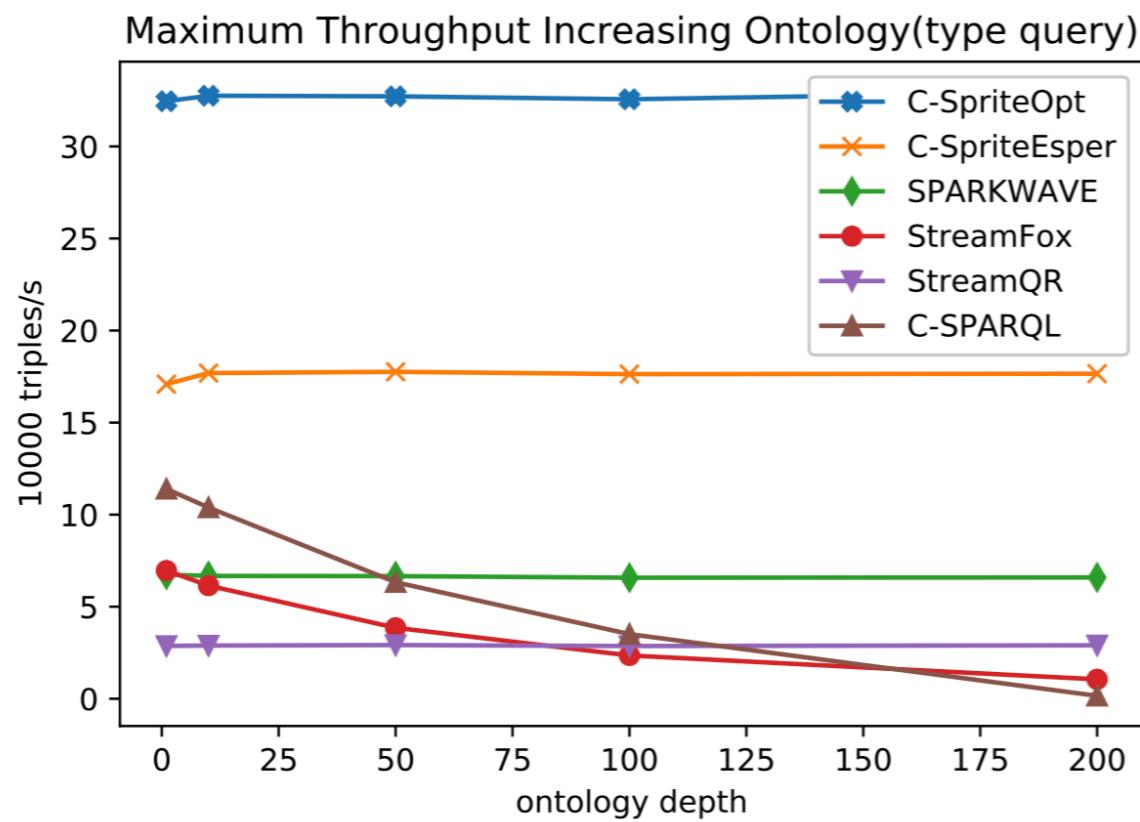


	Absolute Number	Relative Number
all triples	3.511.629	100%
Creative Works	56.581	1,61%
Top 5 Creative Works:		
MusicalWork	21.438	0,61%
Film	13.890	0,40%
WrittenWork	6.814	0,19%
TelevisionShow	4.579	0,13%
Software	4.493	0,13%

# Evaluation: Increasing window size

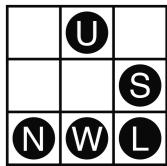


# Evaluation: Increasing ontology depth



# Future Work

- TRA @ Apache Spark
- Query Containment



# Questions?

- **Email:** riccardo.tommasini@polimi.it
- **Twitter:** @rictomm
- **Github:** riccardotommasini
- **Web1:** [riccardotommasini.com](http://riccardotommasini.com)
- **Web2:** [streamreasoning.org](http://streamreasoning.org)
- **Web3:** [streaminglang.io](http://streaminglang.io)

