

SPARQL

Shima Dastgheib
Mehdi Allahyari

CSCI 8370 Advanced Database Systems
Spring 2012

Presentation outline

- Brief introduction to RDF triple pattern
- SPARQL
 - > Introduction to SPARQL
 - > How to write a SPARQL query? (Demo)
- Comparing some different approaches in RDF Storage and query processing based upon

RDF triple pattern

Subject

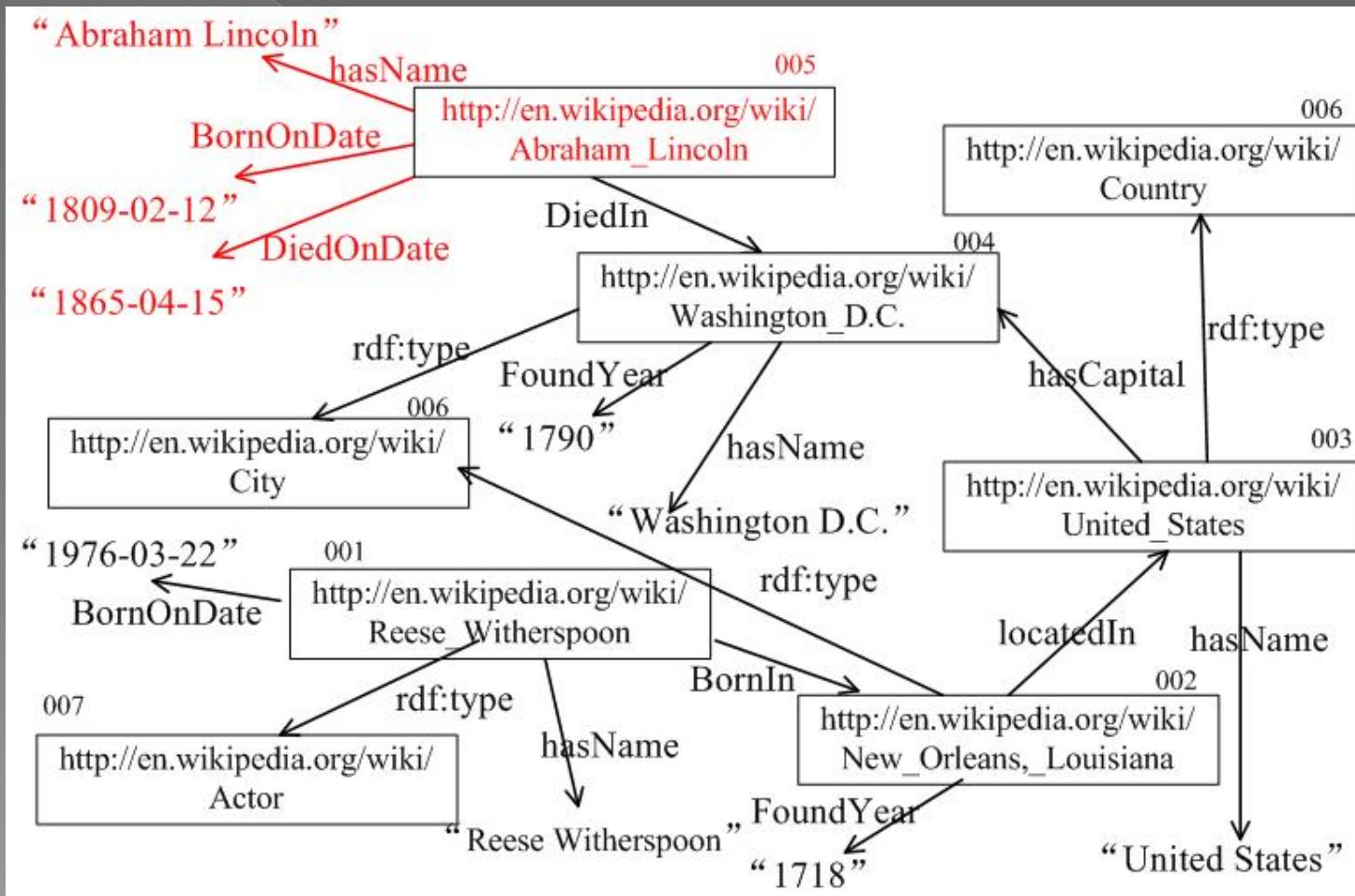
Predicate

Object



we will shed light on the RDF graph next week

SPARQL is about matching triple patterns



Something to know beforehand :

- URI

http://dbpedia.org/resource/Barack_Obama

- Prefix : for abbreviating URLs

Prefix dbpedia: <http://dbpedia.org/resource/>

dbpedia : Barack Obama

one useful property from RDF

- `rdf:type` → a property between an instance and the class it belongs to

Example: UGA `rdf:type` University

SPARQL Recipe

Ingredients & Formulation

prefix declarations

PREFIX foo: <<http://example.com/resources>>

...

dataset definition

FROM ...

result clause

SELECT ...

query pattern (one or more triple patterns)

WHERE {

...

}

solution modifiers

ORDER BY / LIMIT/OFFSET

Do more than SELECT

- ASK
- CONSTRUCT
- DESCRIBE

SPARQL

By Example

<http://www.cambridgesemantics.com/2008/09/sparql-by-example/#q12>

SPARQL 1.1

- Some new features:

- > Update (INSERT/DELETE)
- > Subqueries
- > Aggregation (MAX,MIN,SUM,COUNT,etc)
- > Federation
- > And more

Demo

SPARQLing ProkinO

[http://gumbo.cs.uga.edu/prookino2/about/
browser](http://gumbo.cs.uga.edu/prookino2/about/browser)

Comparing some approaches regarding RDF Storage and Querying based upon

Based on gstore paper

Naïve Triple Store

SPARQL Query:

```
Select ?name Where { ?m <hasName> ?name. ?m  
<BornOnDate> "1809-02-12". ?m <DiedOnDate>  
"1865-04-15". }
```

Prefix: y= http://en.wikipedia.org/wiki/

Subject	Predict	Object
y:Abraham_Lincoln	hasName	"Abraham Lincoln"
y:Abraham_Lincoln	BornOnDate	"1809-02-12"
y:Abraham_Lincoln	DiedOnDate	1865-04-15
y:Abraham_Lincoln	DiedIn	y:Washington_D.C
y:Washington_D.C	hasName	"Washington D.C."
y:Washington_D.C	FoundYear	1790
y:Washington_D.C	rdf:type	y:city
y:United_States	hasName	"United States"
y:United_States	hasCapital	y:Washington_D.C
y:United_States	rdf:type	Country
y:Reese_Witherspoon	rdf:type	y:Actor
y:Reese_Witherspoon	BornOnDate	"1976-03-22"
y:Reese_Witherspoon	BornIn	y:New_Oreleans,_Louisiana
y:Reese_Witherspoon	hasName	"ReeseWitherspoon"
y:New_Oreleans,_Louisiana	FoundYear	1718
y:New_Oreleans,_Louisiana	rdf:type	y:city
y:New_Oreleans,_Louisiana	locatedIn	y:United_States

Too many
Self-Joins

SQL:

```
Select T3.Subject  
From T as T1, T as T2, T as T3  
Where T1.Predict="BornOnDate"  
and T1.Object="1809-02-12" and  
T2.Predict="DiedOnDate" and  
T2.Object="1865-04-15" and T3.  
Predict="hasName" and  
T1.Subject = T2.Subject and T2.  
Subject= T3.subject
```

Property Table

SPARQL Query:

```
Select ?name Where { ?m <hasName> ?name. ?m  
<BornOnDate> "1809-02-12". ?m <DiedOnDate>  
"1865-04-15". }
```

Reducing # of
join steps

Prefix: y= http://en.wikipedia.org/wiki/

People

Subject	hasName	BornOnDate	DiedOnDate	DiedIn	BornIn	rdf:type
y:Abraham_Lincoln	"Abraham Lincoln"	1809-02-12	1865-04-15	y:Washington_D.C		
y:Reese_Witherspoon	"Reese Witherspoon"	1976-03-22		y:Washington_D.C	y>New_Orleans,_Louisiana	y:Actor

City

Subject	FoundYear	rdf:type	locatedIn	hasName
y>New_Orleans,_Louisiana	1718	y:city	y:United_States	
y:Washington_D.C	1790	y:city	y:United_States	"Washington D.C."

Country

Subject	hasName	hasCapital	rdf:type
y:United_States	"United States"	y:Washington_D.C	Country

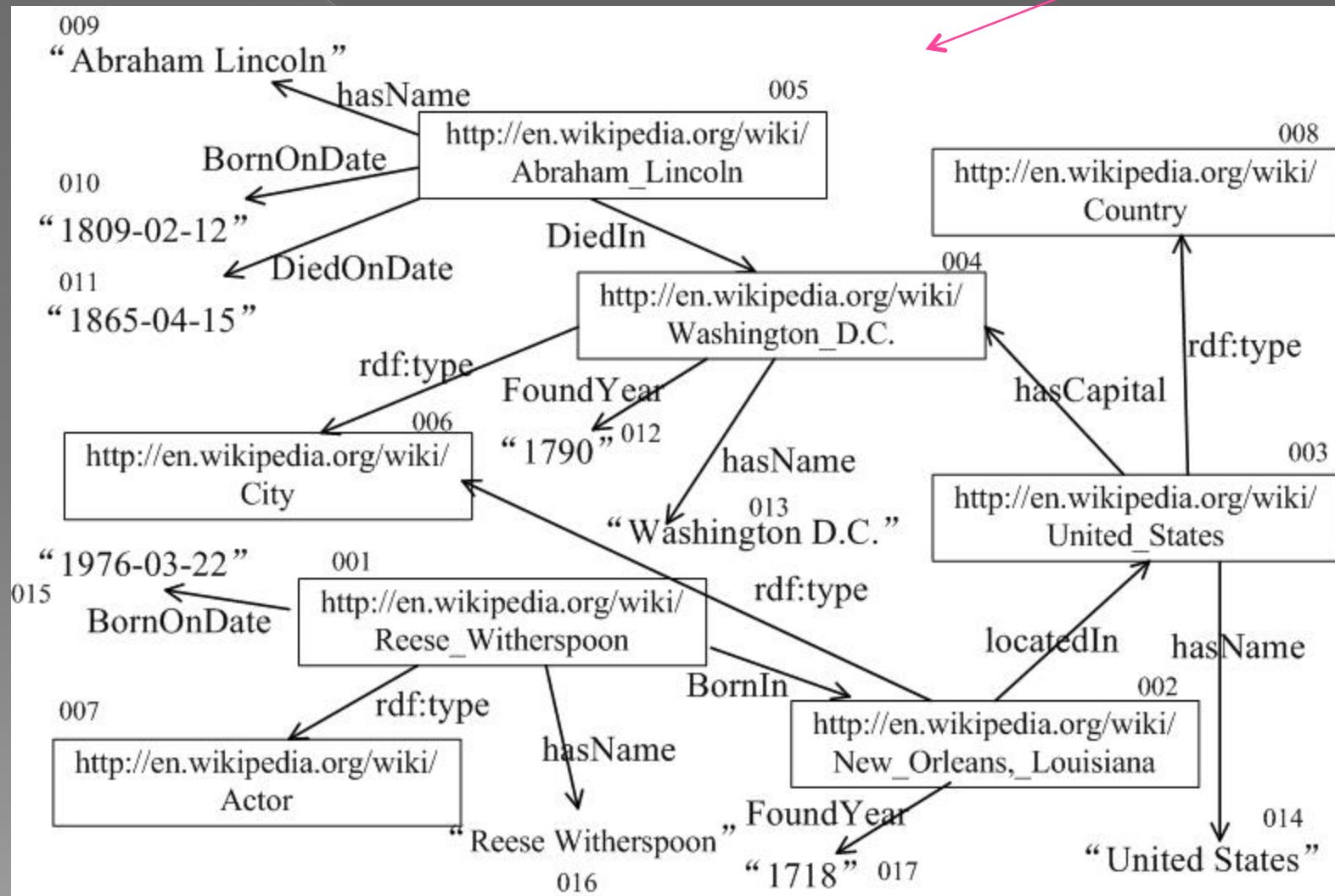
SQL:

```
Select People.hasName from People where  
People.BornOnDate = "1809-02-12" and People.DiedOnDate = "1865-04-15".
```

gStore

Literal Vertex

Entity Vertex

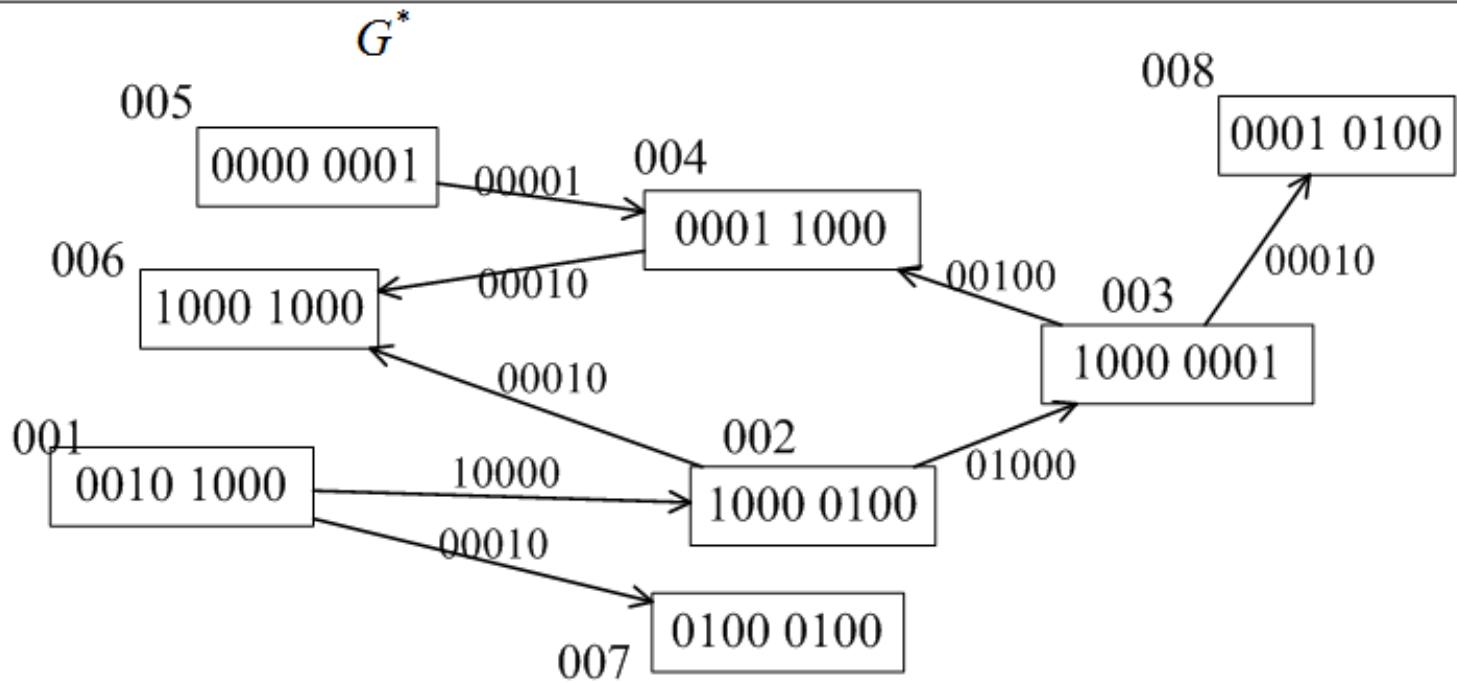


Adjacency List

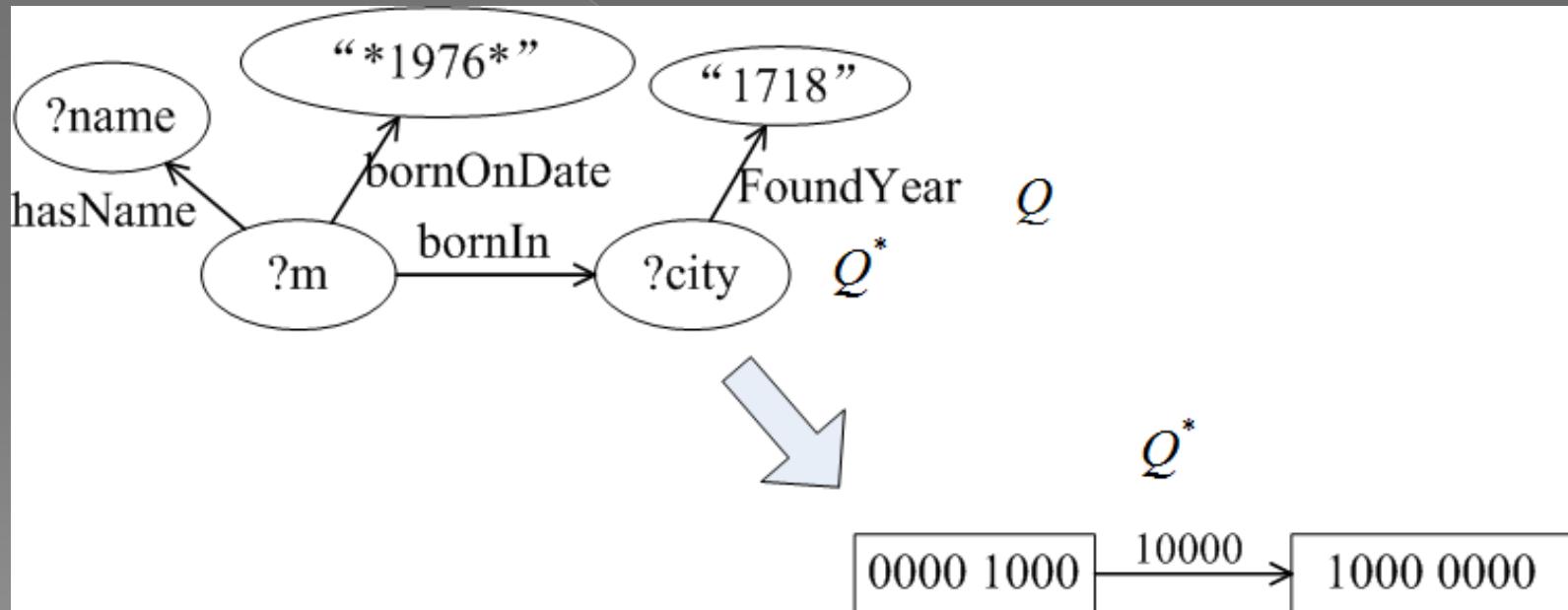
Prefix: y= http://en.wikipedia.org/wiki/

vID	vLabel	adjList {(eLabel, nLabel)+}
001	y:Abraham_Lincoln	(hasName, “Abraham Lincoln”), (BornOnDate, “1809-02-12”) (DiedOnDate, “1865-04-15”) (DiedIn, y:Washington_D.C)
002	y:Washington_D.C	(hasName, “Washington D.C.”), (FoundYear , “1790”) (rdf:type, y:city)
003	y:United_States	(hasName, “United States”), (hasCapital,y:Washington_D.C) (rdf:type, y:country)
004	y:Reese_Witherspoon	(hasName, “ReeseWitherspoon”), (BornOnDate, “1976-03-22”) (hasCapital, y:New_Orleans,_Louisiana) (rdf:type, y:Actor)
005	y:New_Orleans,_Louisiana	(FoundYear, “1718”), (locatedIn, y:United_States) (rdf:type, y:city)

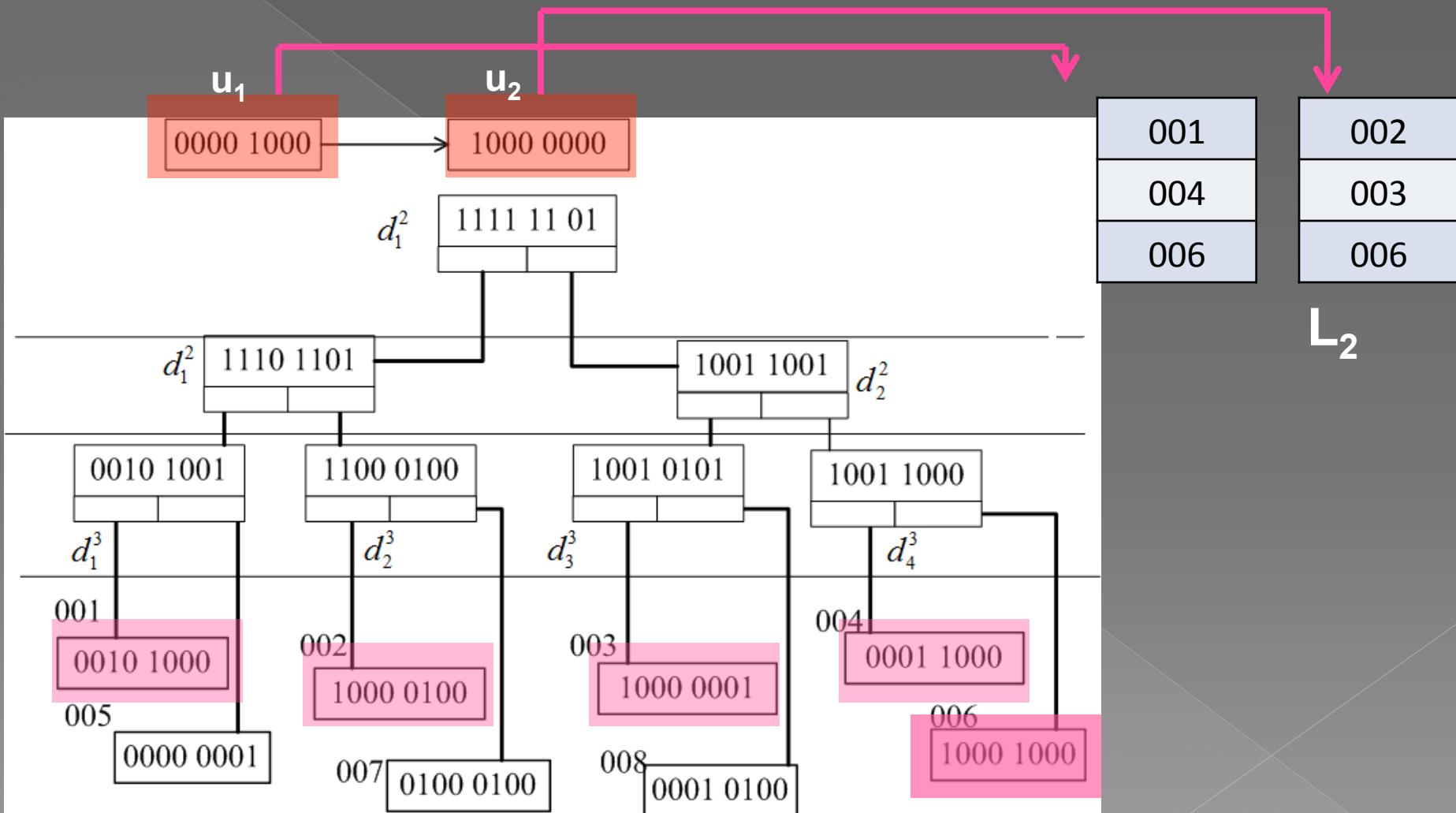
Signature Graph G^*



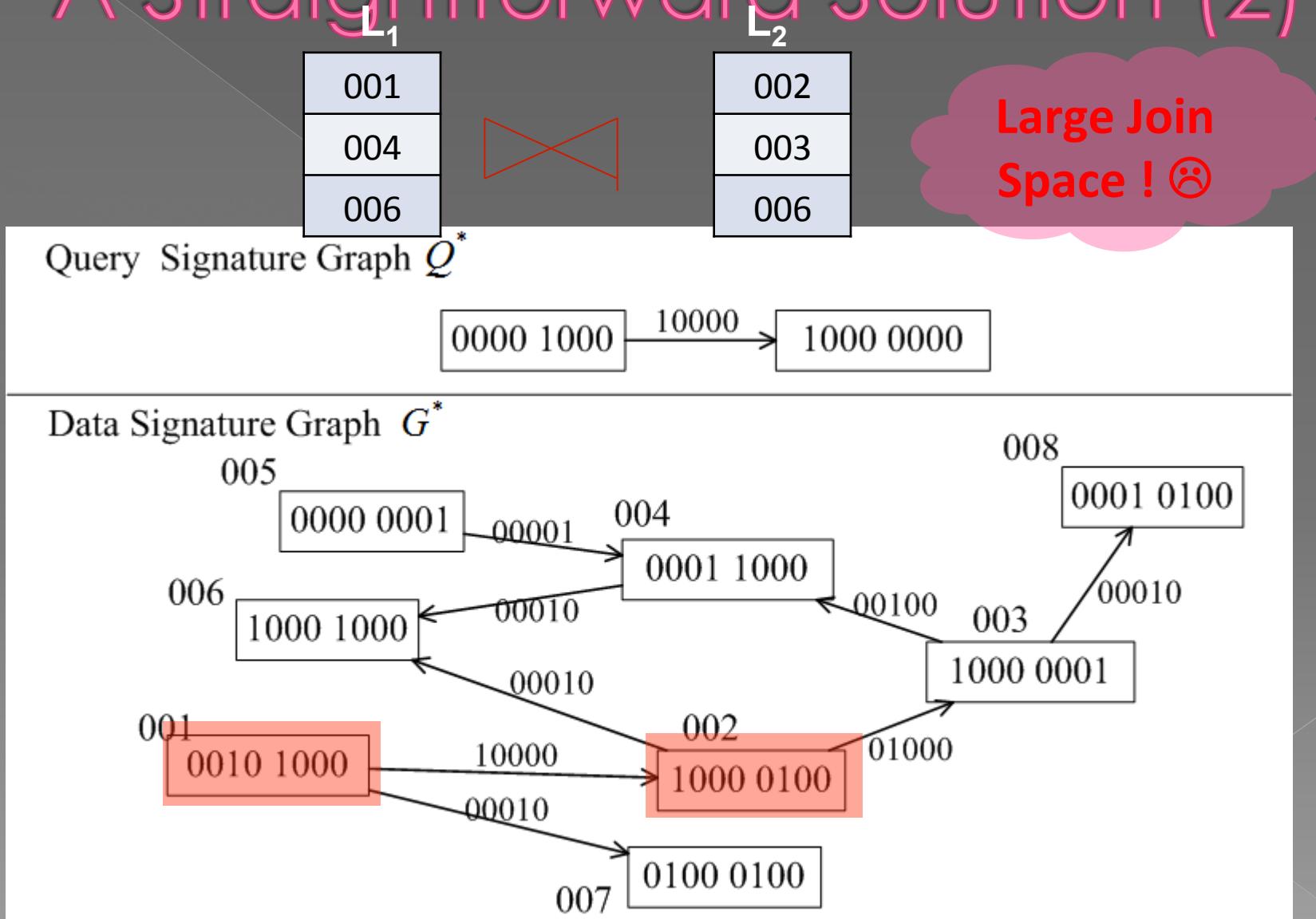
Query Signature Q^*



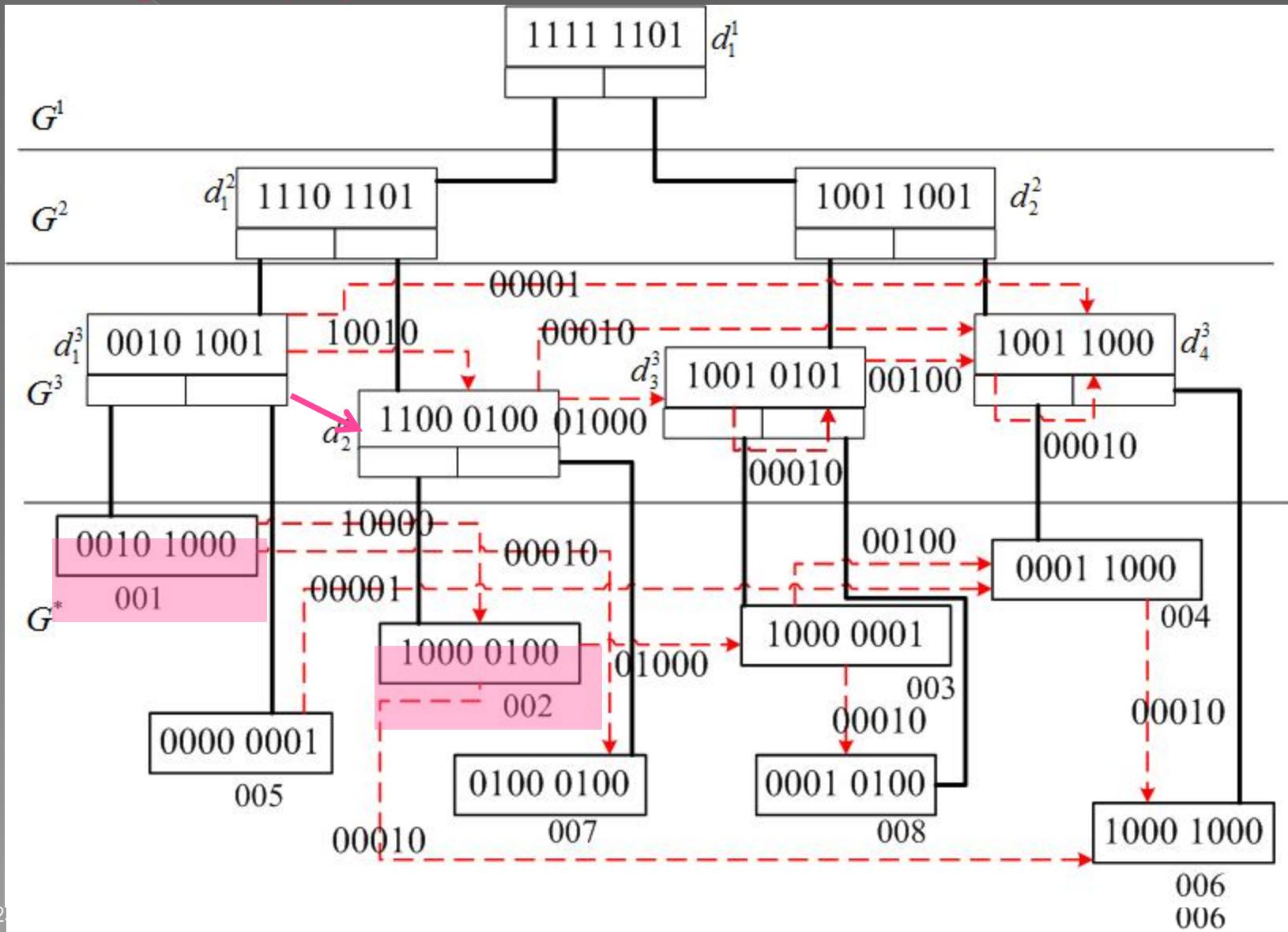
A Straightforward Solution (1)



A Straightforward Solution (2)



VS-Tree



References

1. <http://www.cambridgesemantics.com/2008/09/sparql-by-example/>
2. <http://gumbo.cs.uga.edu/prokino2/about/browser>
3. L. Zou, J. Mo, L. Chen, M. Ozsu, and D. Zhao.gStore: Answering SPARQL Queries via Subgraph Matching. PVLDB, 4(8), 2011.