

Ontology Database: a New Method for Semantic Modeling and an Application to Brainwave Data

Paea LePendur¹, Dejing Dou¹, Gwen Frishkoff², Jiawei Rong¹

¹Computer and Information Science, University of Oregon

²Learning Research and Development Center, University of Pittsburgh

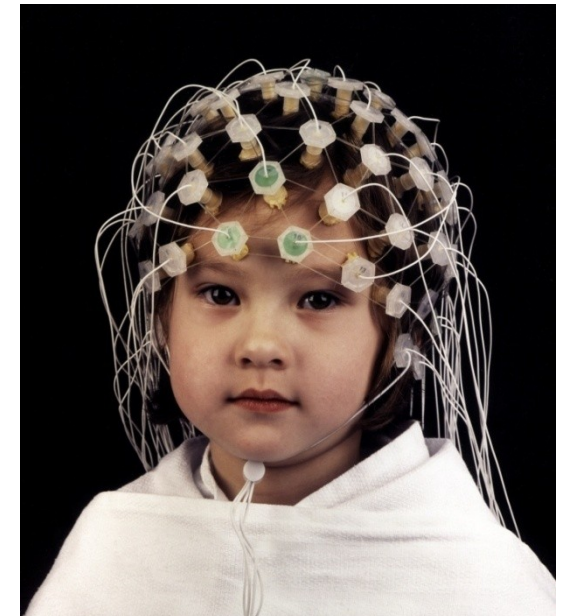
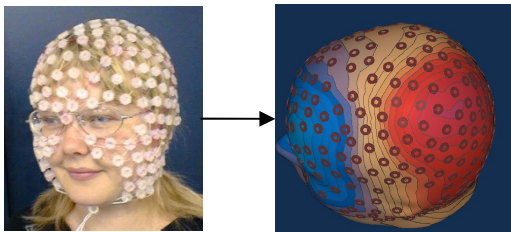
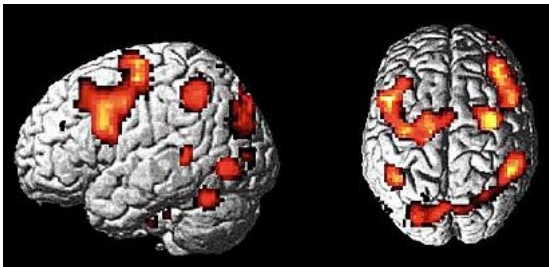
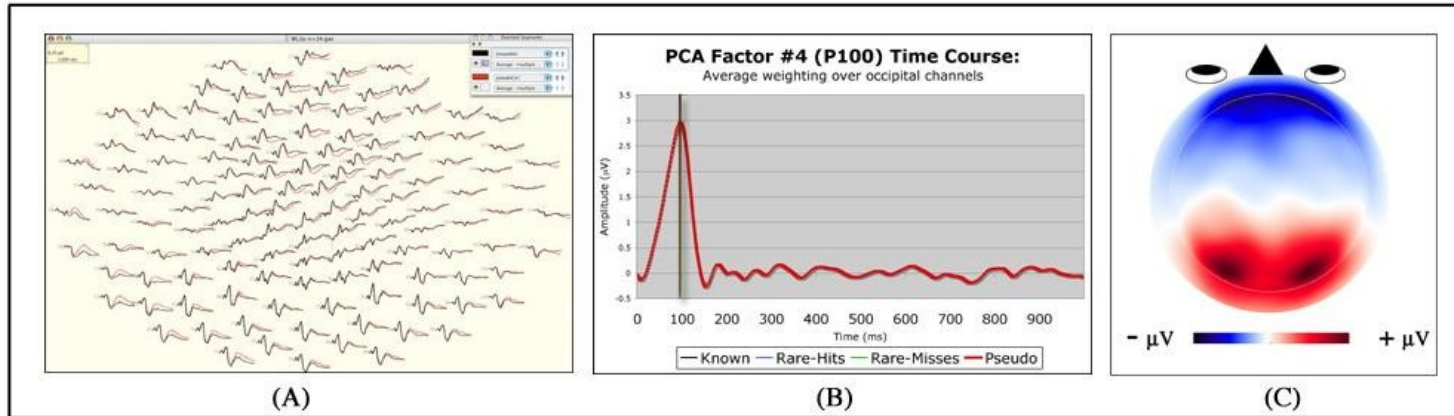


July, 2008 @ SSDBM '08

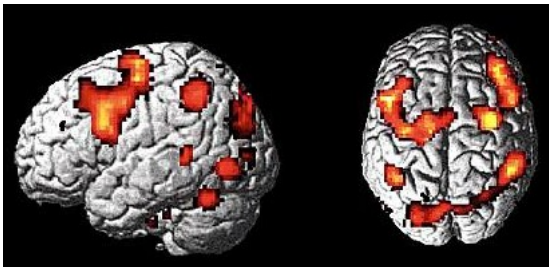
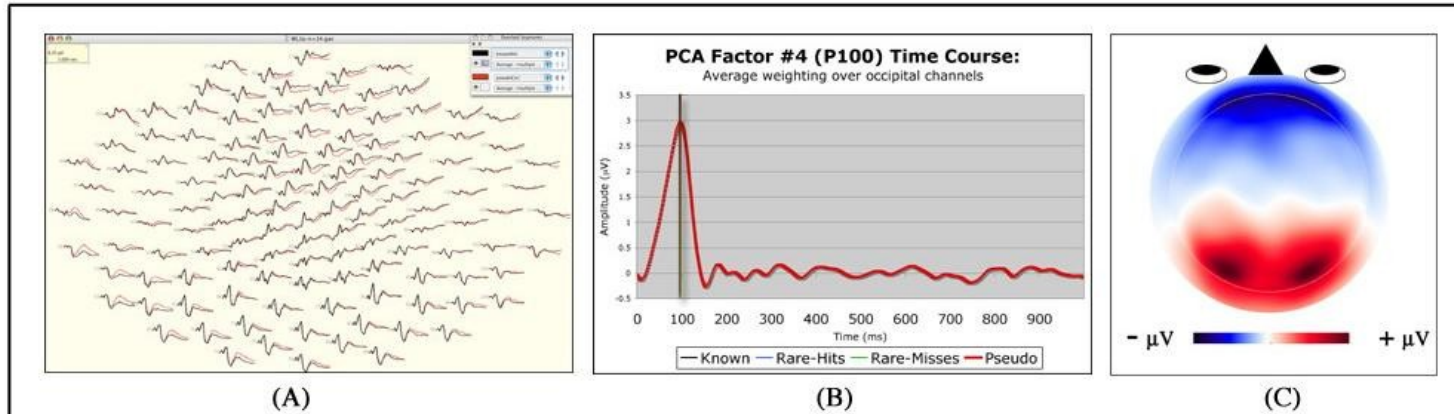
Outline

- Background and Related Work
 - Brainwave data and pattern analysis
 - The NEMO project as motivation
 - Domain ontologies
- Ontology Database Methodology
 - Existing, view-based technique
 - New, trigger-based technique
- Benchmarking Analysis
- Discussion and Future Work

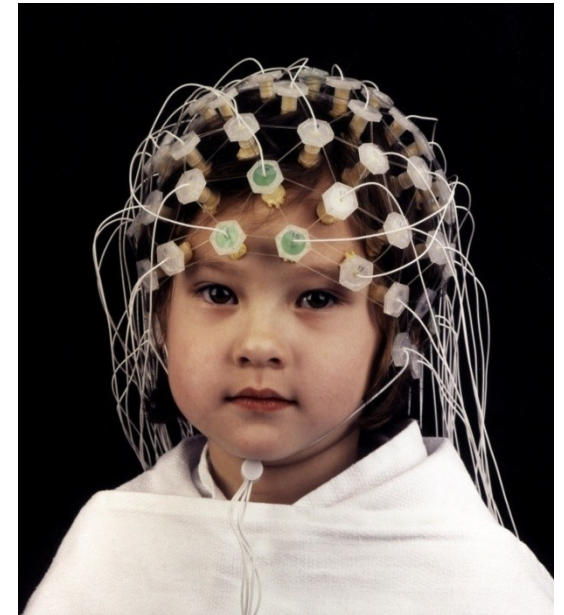
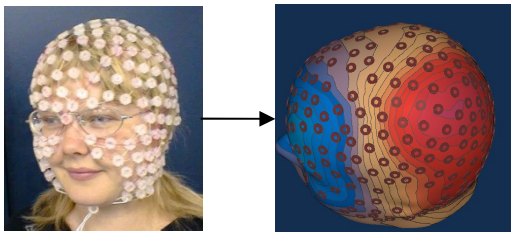
Brainwave Data



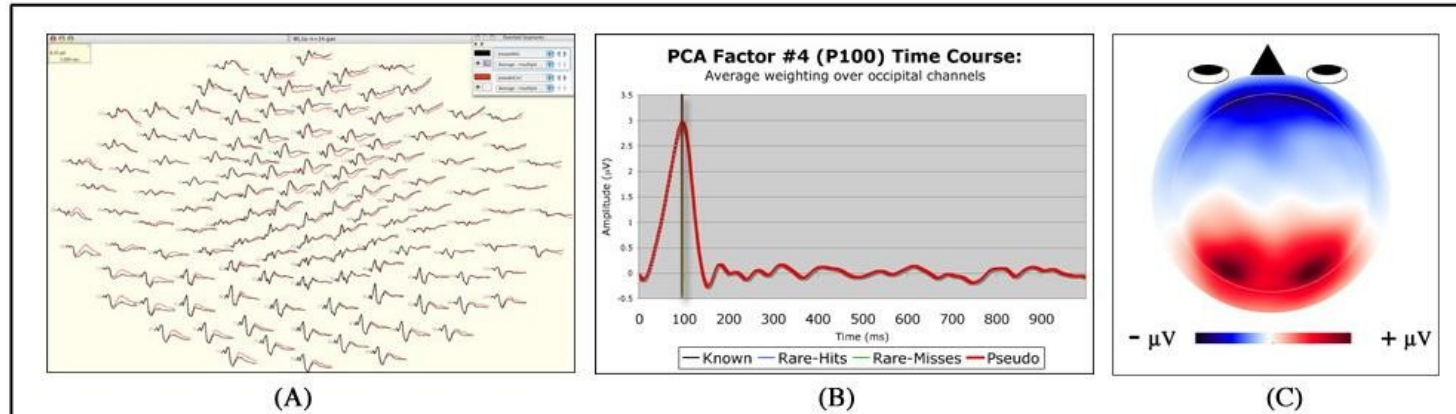
Brainwave Data



Talk about exponential growth!



Brainwave Data

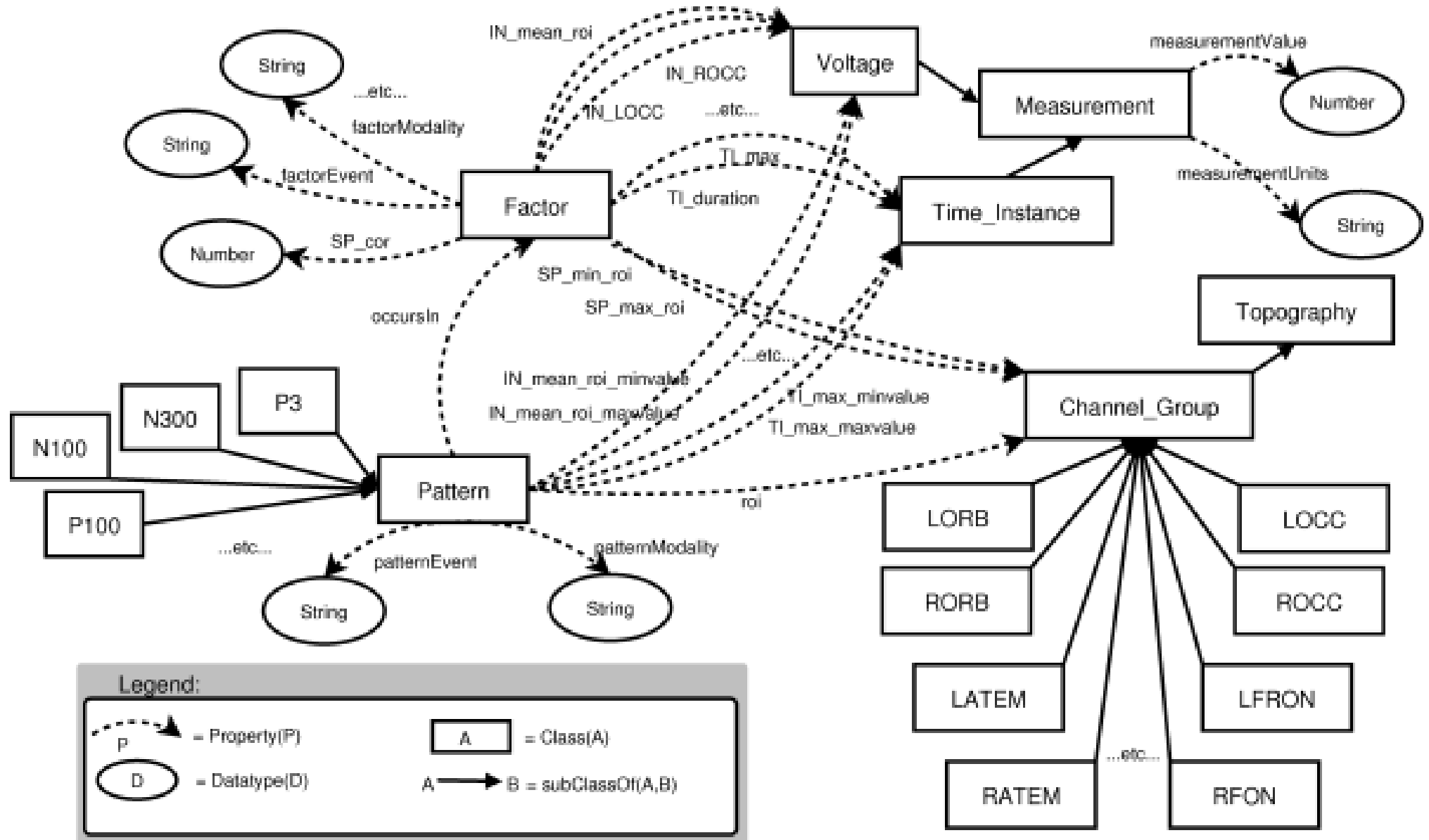


- Some problems with EEG/ERP data:
 - Complex dimensionality (spatial, temporal, functional)
 - Data sharing
 - Meta-analysis

Brainwave Ontologies

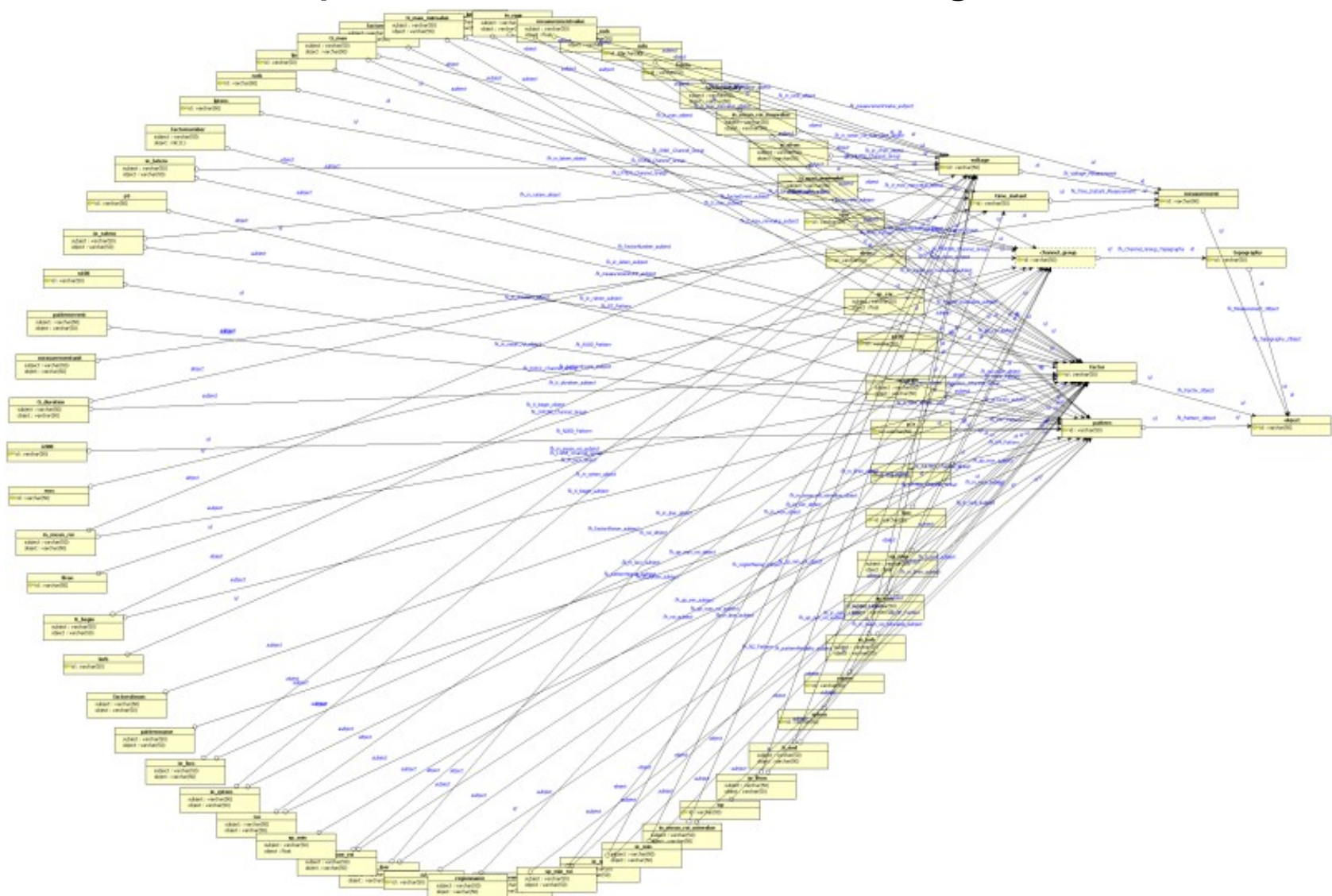
- To address these problems, ontologies are used:
 - Birnlex
 - NEMO (NeuroElectroMagnetic Ontologies)
- Distinct but inter-dependent models

NEMO (NeuroElectroMagnetic Ontology)



NEMO (NeuroElectroMagnetic Ontology)

Graphical View of the ER-Diagram



What are Ontologies?

- Machine processible models
- Logic-based formalisms
- Main communities:
 - Knowledge Representation and Reasoning (KRR)
 - Semantic Web

What does it have to do with databases?

- The problem of data scale (vs. model consistency)
 - Billion-triple challenge ISWC '08
- Views (Datalog) are coming back...
 - But databases have since evolved!
 - (e.g., Active Database technology)
- KRDB Group in Bozen-Bolzano, Italy
 - Reuniting Knowledge Representation and DataBases

Ontology Databases

- A Simple Problem Example
 - Some reasoning review
- Bridging the Gap
 - Ontologies and Databases
- Contrast Existing and Proposed Methodology

Example: a Simple Problem

This is what we know :

All sisters are siblings.

Hilary and Lynn are sisters.

This is what we want to know :

Who are siblings?

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

Obviously, the answer should be :

Hilary and Lynn are siblings.

$\{ \langle \text{Hilary}, \text{Lynn} \rangle \}$

A Goal Directed Search

Automated reasoning can solve this easily.

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\text{siblingOf}(x, y)$

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\frac{\emptyset}{\text{siblingOf}(x, y)} \text{ unify?}$

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\frac{\emptyset}{\text{siblingOf}(x, y)} \emptyset$

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\text{siblingOf}(x, y)$

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\text{sisterOf}(x, y) \Rightarrow \text{siblingOf}(x, y)$

$\text{sisterOf}(x, y)$

modus ponens

$\text{siblingOf}(x, y)$

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/x, y'/y\}$

$\text{sisterOf}(x, y) \Rightarrow \text{siblingOf}(x, y)$

$\text{sisterOf}(x, y)$

modus ponens

$\text{siblingOf}(x, y)$

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/x, y'/y\}$

$\text{sisterOf}(x, y) \Rightarrow \text{siblingOf}(x, y)$

unify?

$\text{sisterOf}(x, y)$

modus ponens

$\text{siblingOf}(x, y)$

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/x, y'/y\}$

$\text{sisterOf}(x, y) \Rightarrow \text{siblingOf}(x, y)$

$\text{siblingOf}(x, y)$

$\text{sisterOf}(\text{Hilary}, \text{Lynn})$

unify?

$\text{sisterOf}(x, y)$

modus ponens

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/x, y'/y\}$

$\text{sisterOf}(x, y) \Rightarrow \text{siblingOf}(x, y)$

$\text{siblingOf}(x, y)$

$\text{sisterOf}(\text{Hilary}, \text{Lynn})$

unify?

$\text{sisterOf}(x, y)$

modus ponens

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/x, y'/y\}$

$\text{sisterOf}(x, y) \Rightarrow \text{siblingOf}(x, y)$

$\text{siblingOf}(x, y)$

$\text{sisterOf}(\text{Hilary}, \text{Lynn})$

unify!

$\text{sisterOf}(x, y)$

modus ponens

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/x, y'/y\}$

$\text{sisterOf}(x, y) \Rightarrow \text{siblingOf}(x, y)$

$\text{siblingOf}(x, y)$

$\text{sisterOf}(\text{Hilary}, \text{Lynn})$

$\{x/\text{Hilary}\}$

$\text{sisterOf}(x, y)$

modus ponens

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/x, y'/y\}$

$\text{sisterOf}(x, y) \Rightarrow \text{siblingOf}(x, y)$

$\text{siblingOf}(x, y)$

$\text{sisterOf}(\text{Hilary}, \text{Lynn})$

$\{x/\text{Hilary}, y/\text{Lynn}\}$

$\text{sisterOf}(x, y)$

modus ponens

A Goal Directed Search

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/\text{Hilary}, y'/\text{Lynn}\}$
 $\text{sisterOf}(\text{Hilary}, \text{Lynn}) \Rightarrow \text{siblingOf}(\text{Hilary}, \text{Lynn})$

$\text{sisterOf}(\text{Hilary}, \text{Lynn})$

modus ponens

$\text{siblingOf}(\text{Hilary}, \text{Lynn})$

A Goal Directed Search

$\langle \text{Hilary}, \text{Lynn} \rangle \in \{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\forall x', y'. \text{sisterOf}(x', y') \Rightarrow \text{siblingOf}(x', y')$

$\forall_E \{x'/\text{Hilary}, y'/\text{Lynn}\}$
 $\text{sisterOf}(\text{Hilary}, \text{Lynn}) \Rightarrow \text{siblingOf}(\text{Hilary}, \text{Lynn})$

$\text{sisterOf}(\text{Hilary}, \text{Lynn})$

modus ponens

$\text{siblingOf}(\text{Hilary}, \text{Lynn})$

Key Question #1

If data storage and querying is our main goal...

Key Question #1

...do we really need all this reasoning?

Ontology Databases

Bringing ontologies and databases together.

Ontology Databases

Class

Property

Datatype

Axioms

Objects

Facts

Relation

Attribute

Datatype

keys

constraints

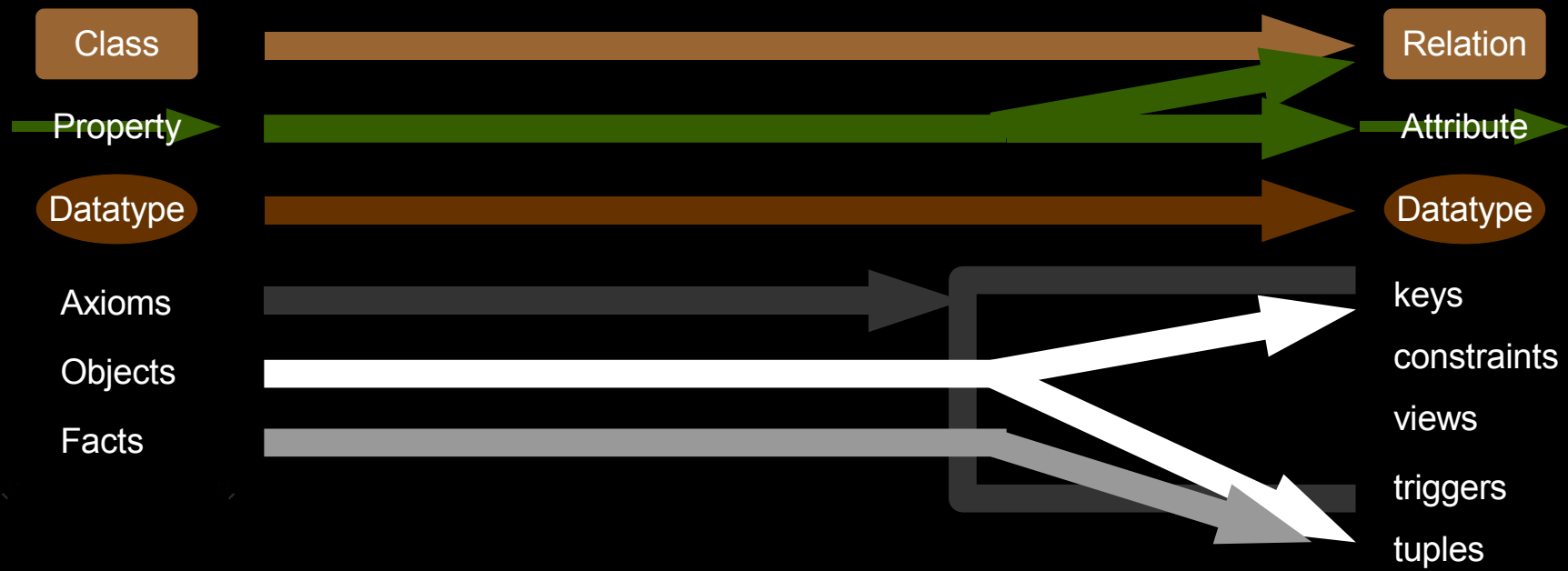
views

triggers

tuples

How do we bridge these?

Ontology Databases



Ontology Databases

Class

Property

Datatype

Axioms

Objects

Facts

Relation

Attribute

Datatype

keys

constraints

views

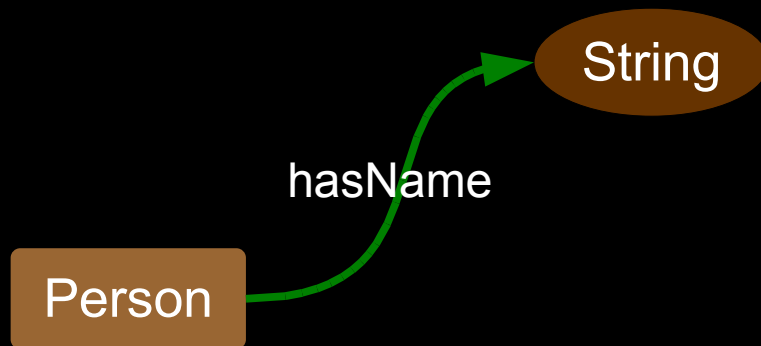
triggers

tuples

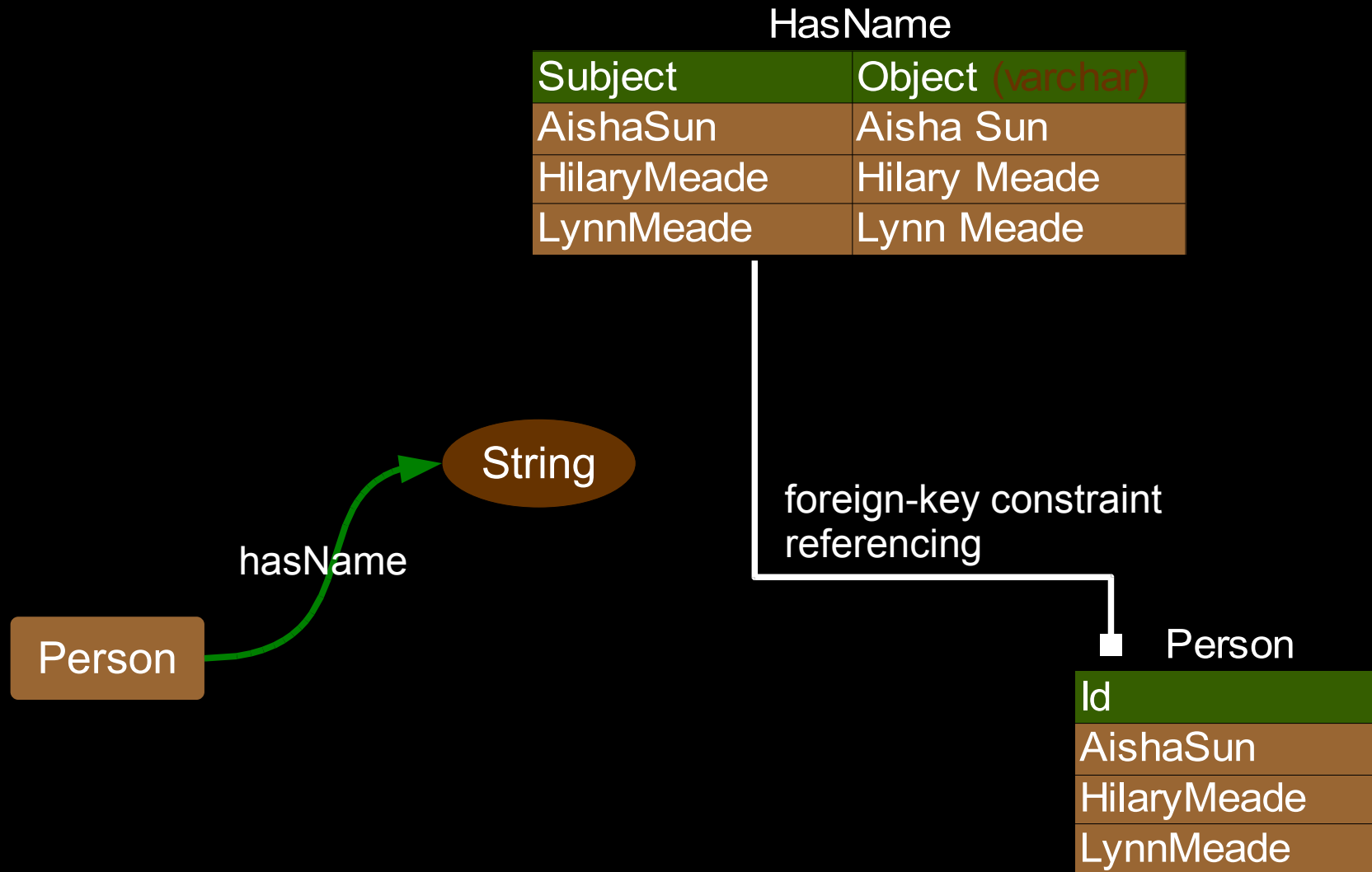
Here's an example.

Ontology Databases

datatype-properties

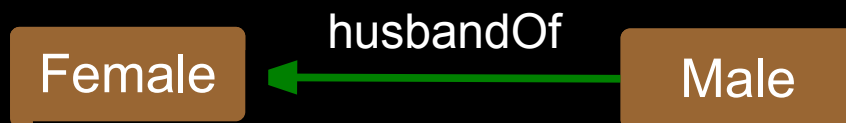


Ontology Databases



Ontology Databases

object-properties



Ontology Databases

Female

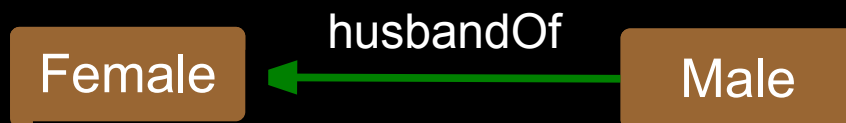
Id
AishaSun
HilaryMeade
LynnMeade

HusbandOf

Subject	Object
MahmudReece	LynnMeade

Male

Id
MahmudReece



Ontology Databases

Female

Id
AishaSun
HilaryMeade
LynnMeade

HusbandOf

Subject	Object
MahmudReece	LynnMeade

foreign-key constraint
referencing

::domain

Male

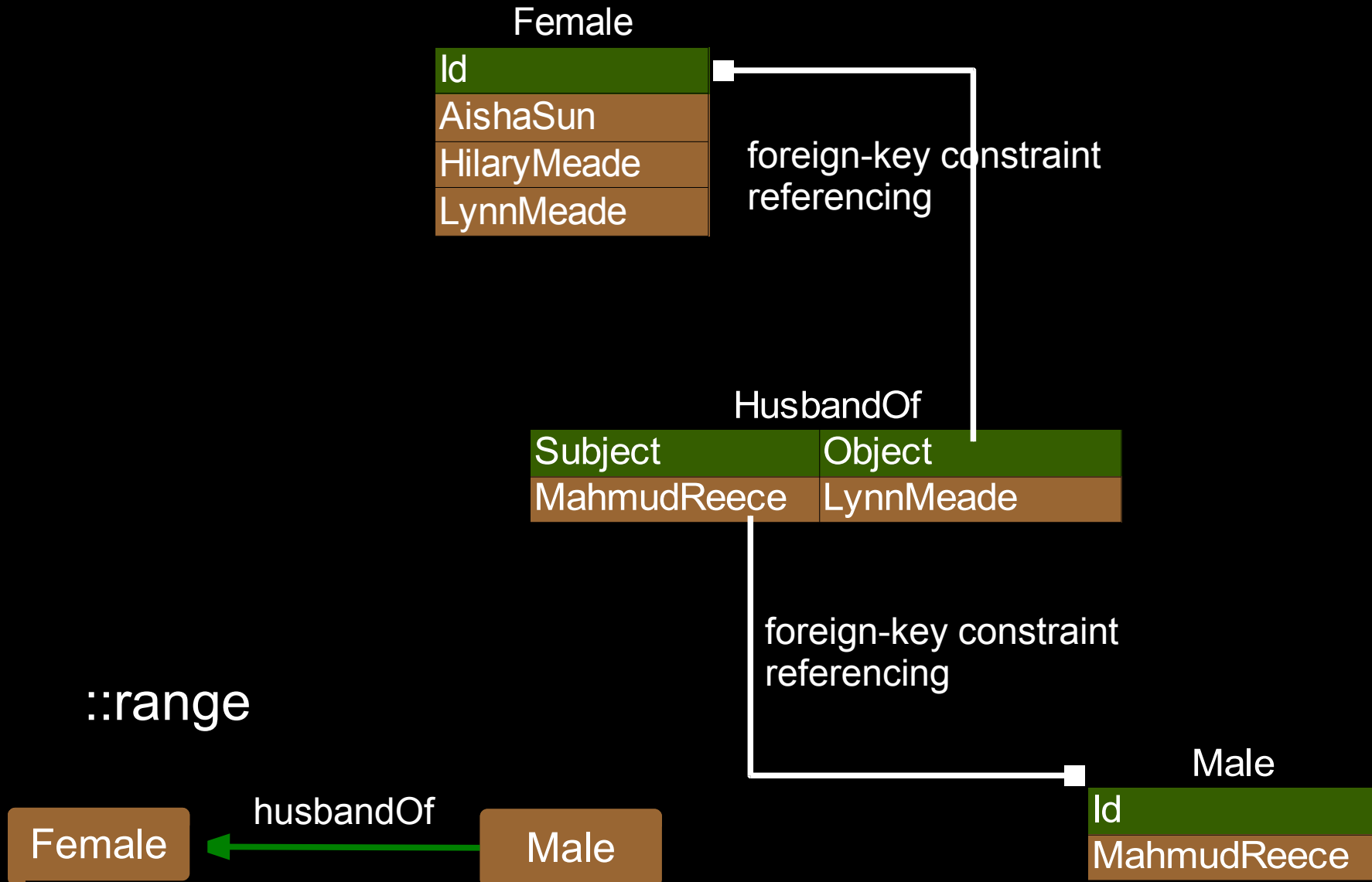
Id
MahmudReece

husbandOf

Female

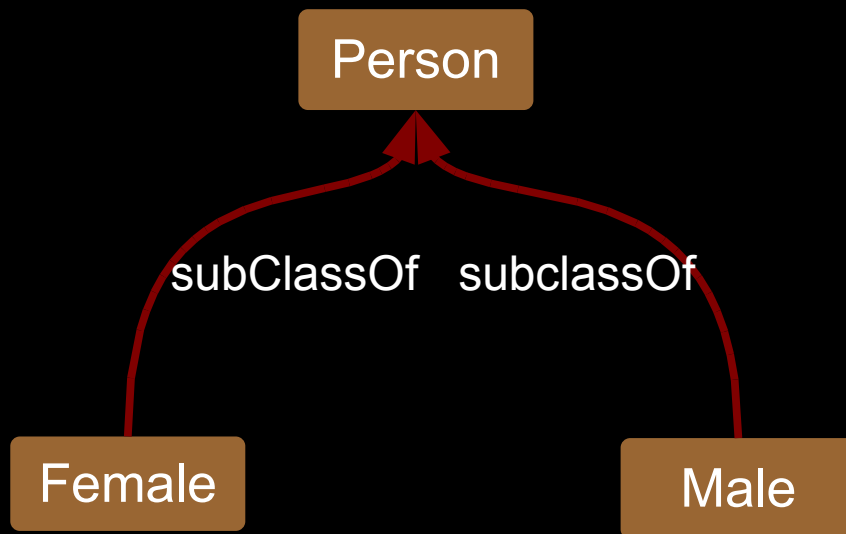
Male

Ontology Databases



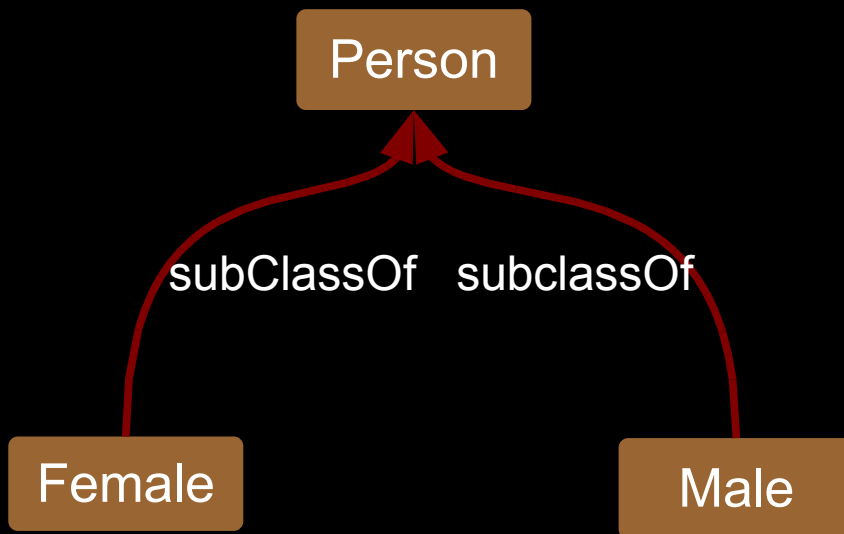
Ontology Databases

subClass axioms



Ontology Databases

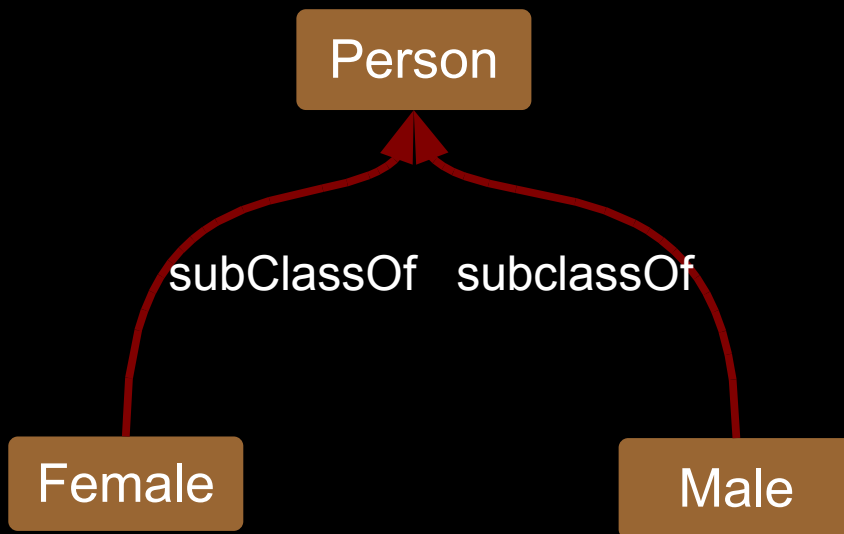
subClass axioms



Two approaches.

Ontology Databases

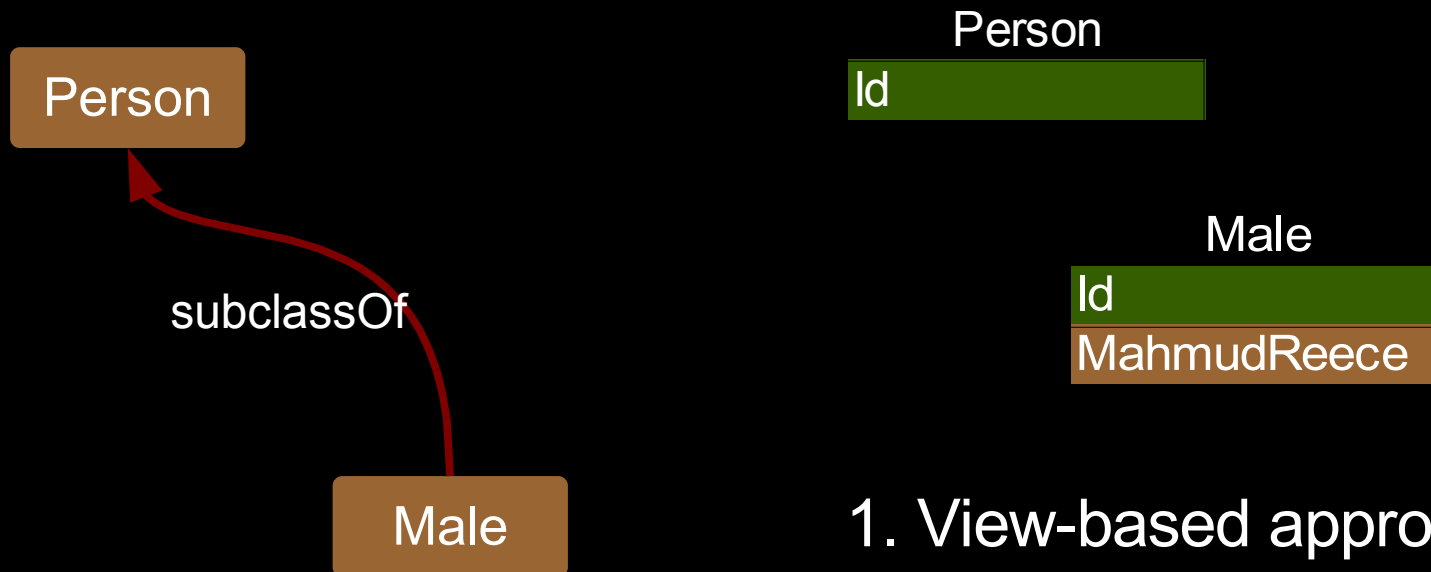
subClass axioms



1. View-based approach.

Ontology Databases

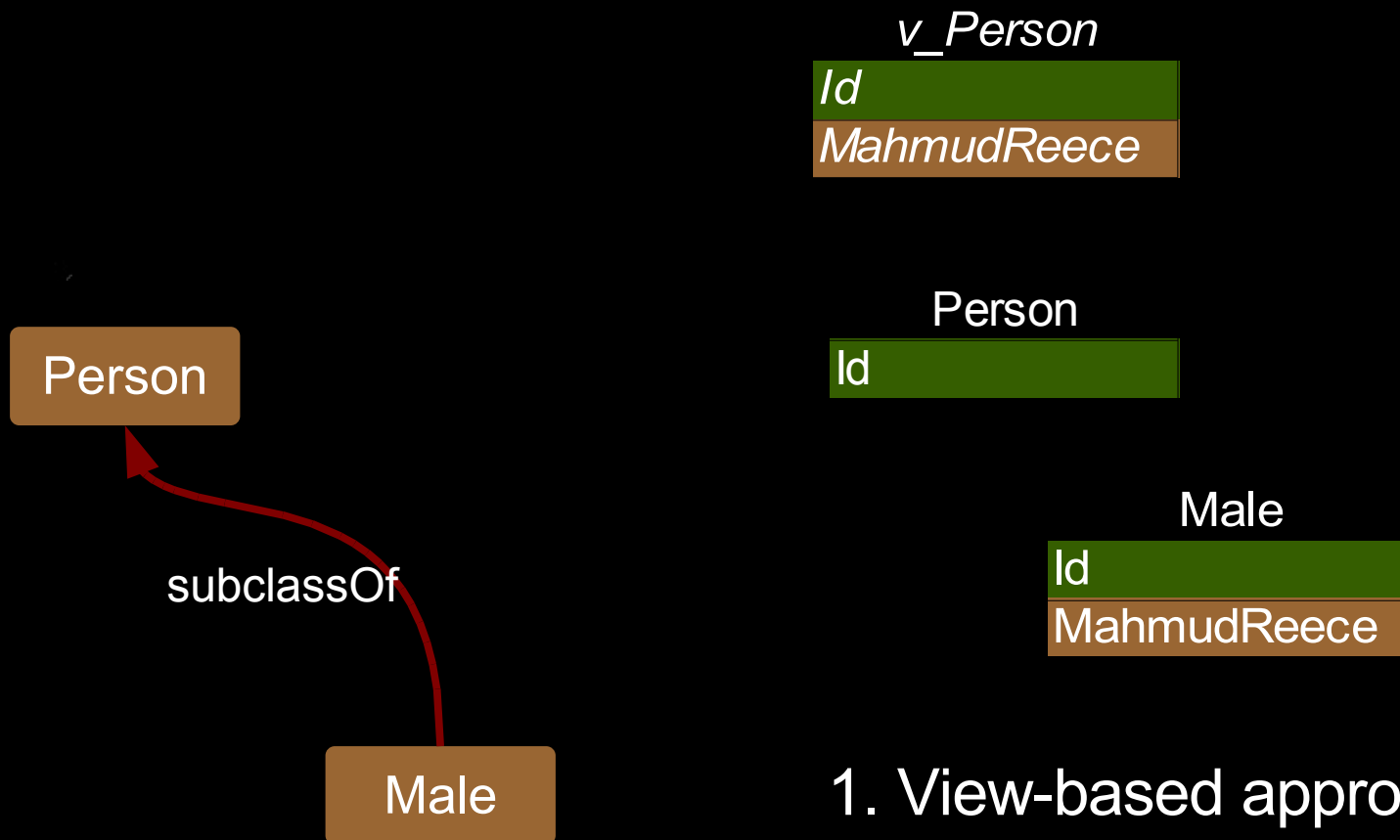
```
CREATE VIEW v_Person(id) AS  
  SELECT id FROM Person  
  UNION  
  SELECT id FROM Male
```



1. View-based approach.

Ontology Databases

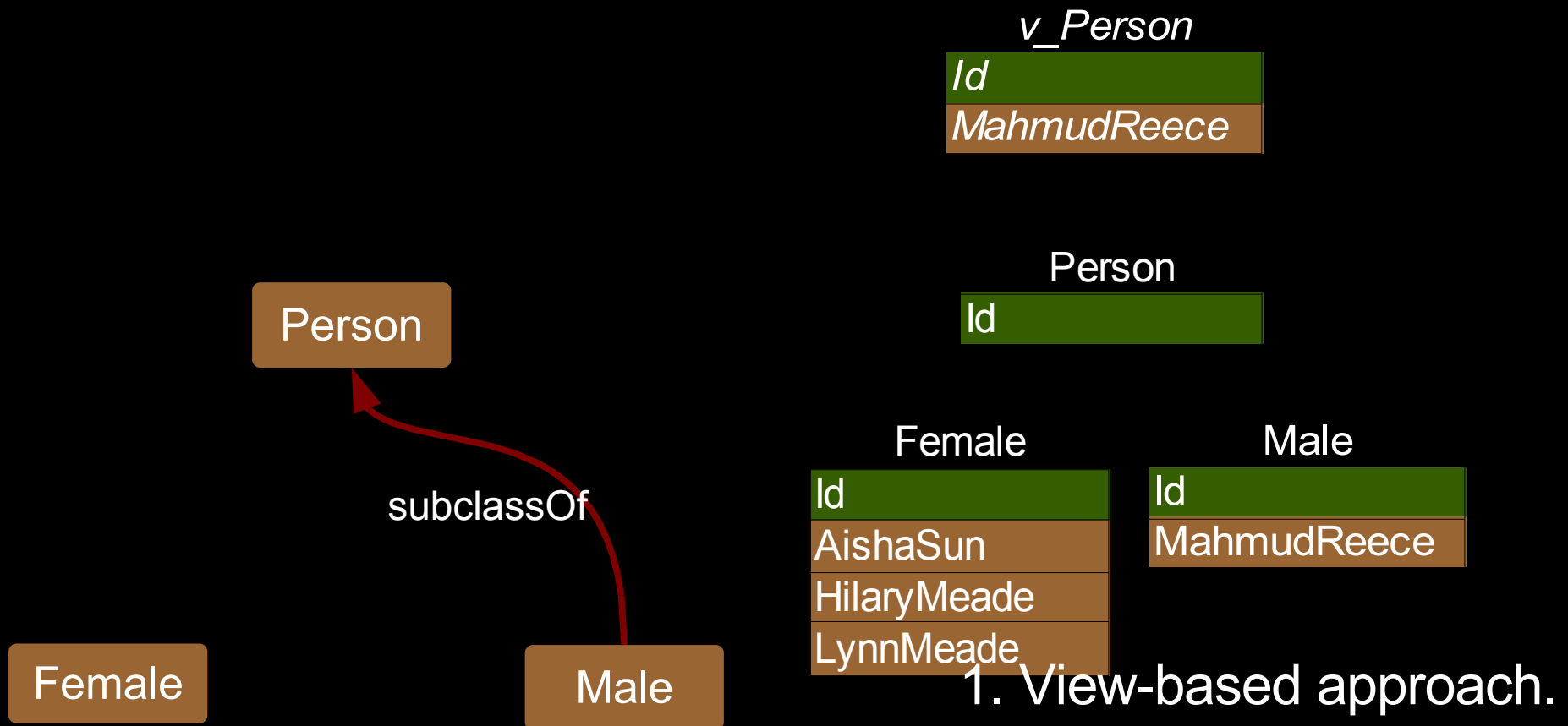
```
CREATE VIEW v_Person(id) AS  
  SELECT id FROM Person  
  UNION  
  SELECT id FROM Male
```



1. View-based approach.

Ontology Databases

```
CREATE VIEW v_Person(id) AS  
  SELECT id FROM Person  
  UNION  
  SELECT id FROM Male
```



1. View-based approach.

Ontology Databases

```
CREATE VIEW v_Person(id) AS
  SELECT id FROM Person
  UNION
  SELECT id FROM Male
  UNION
  SELECT id FROM Female
```

v_Person

<i>Id</i>
<i>MahmudReece</i>
<i>AishaSun</i>
<i>HilaryMeade</i>
<i>LynnMeade</i>

Person

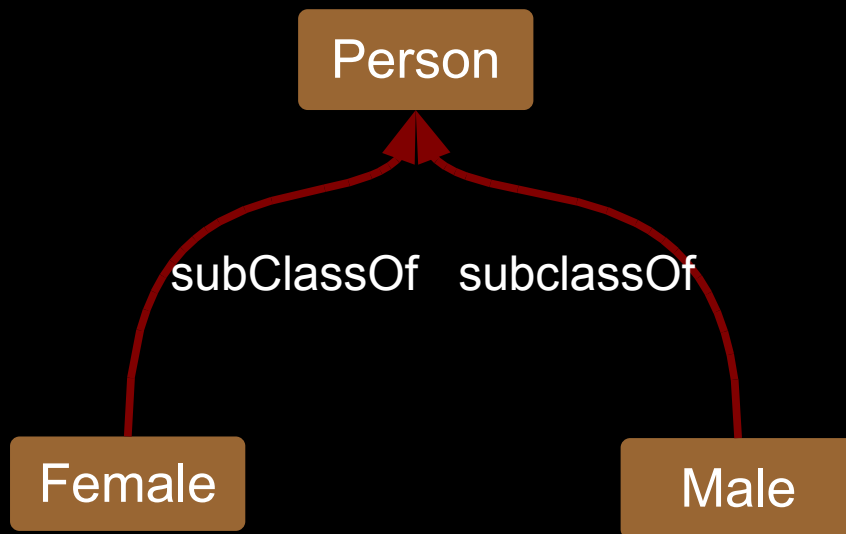
Id

Female

Id
AishaSun
HilaryMeade
LynnMeade

Male

Id
MahmudReece



1. View-based approach.

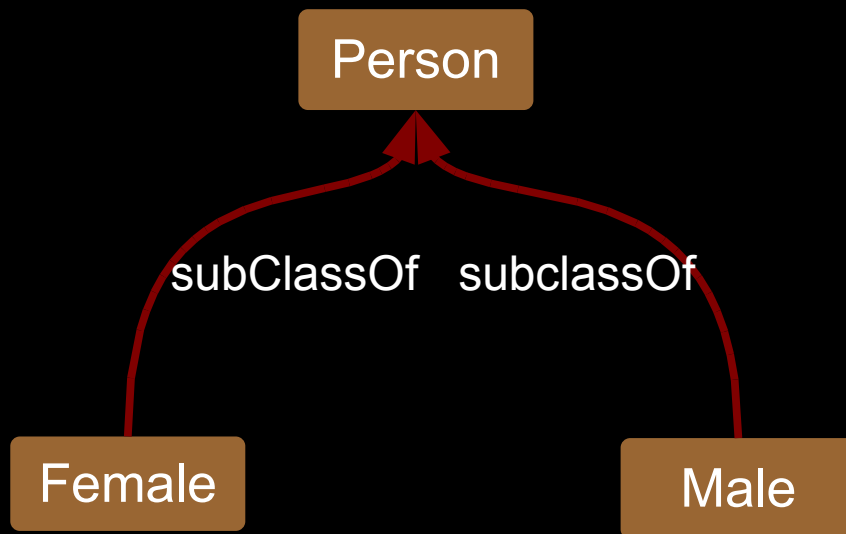
Ontology Databases

DLDB [Pan & Heflin, 2003] implements the view-based approach to store and retrieve voluminous Semantic Web data.

1. View-based approach.

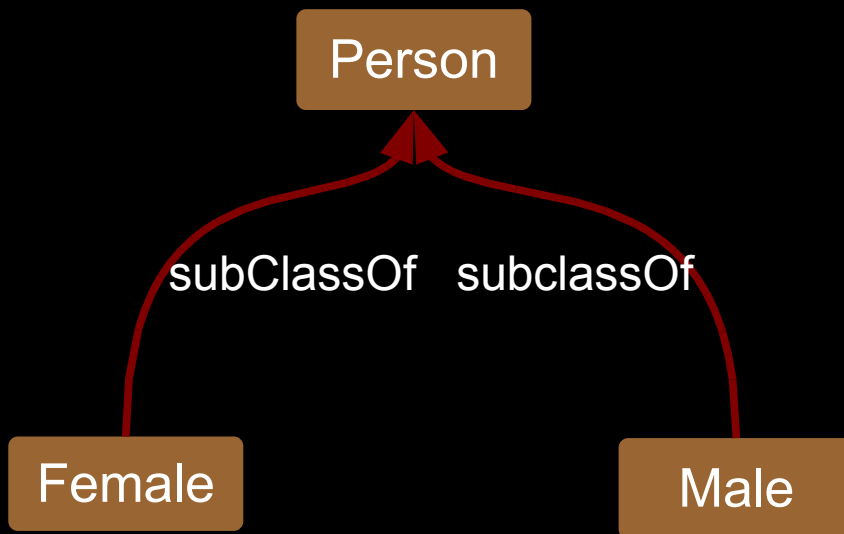
Ontology Databases

subClass axioms



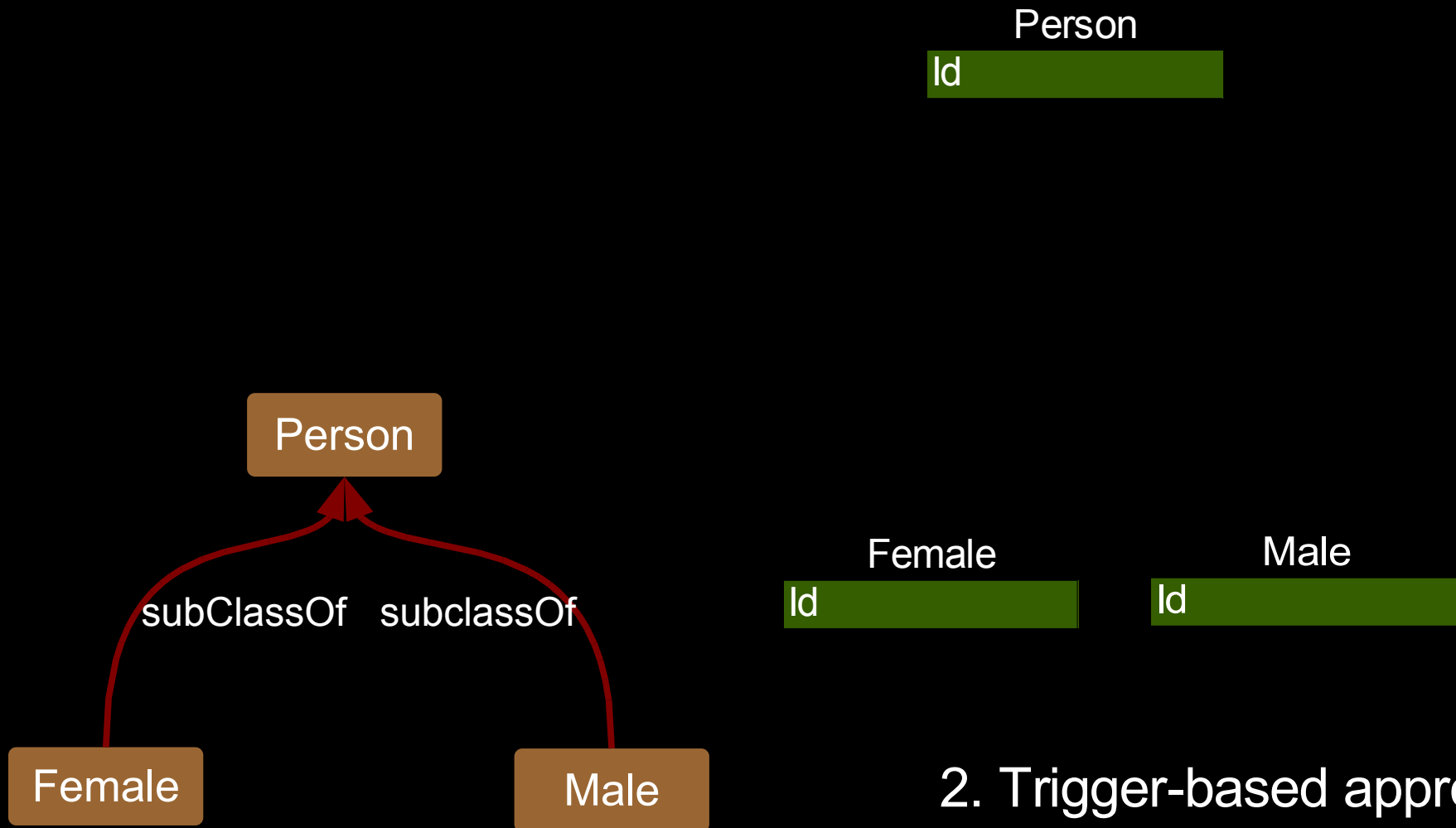
Ontology Databases

subClass axioms



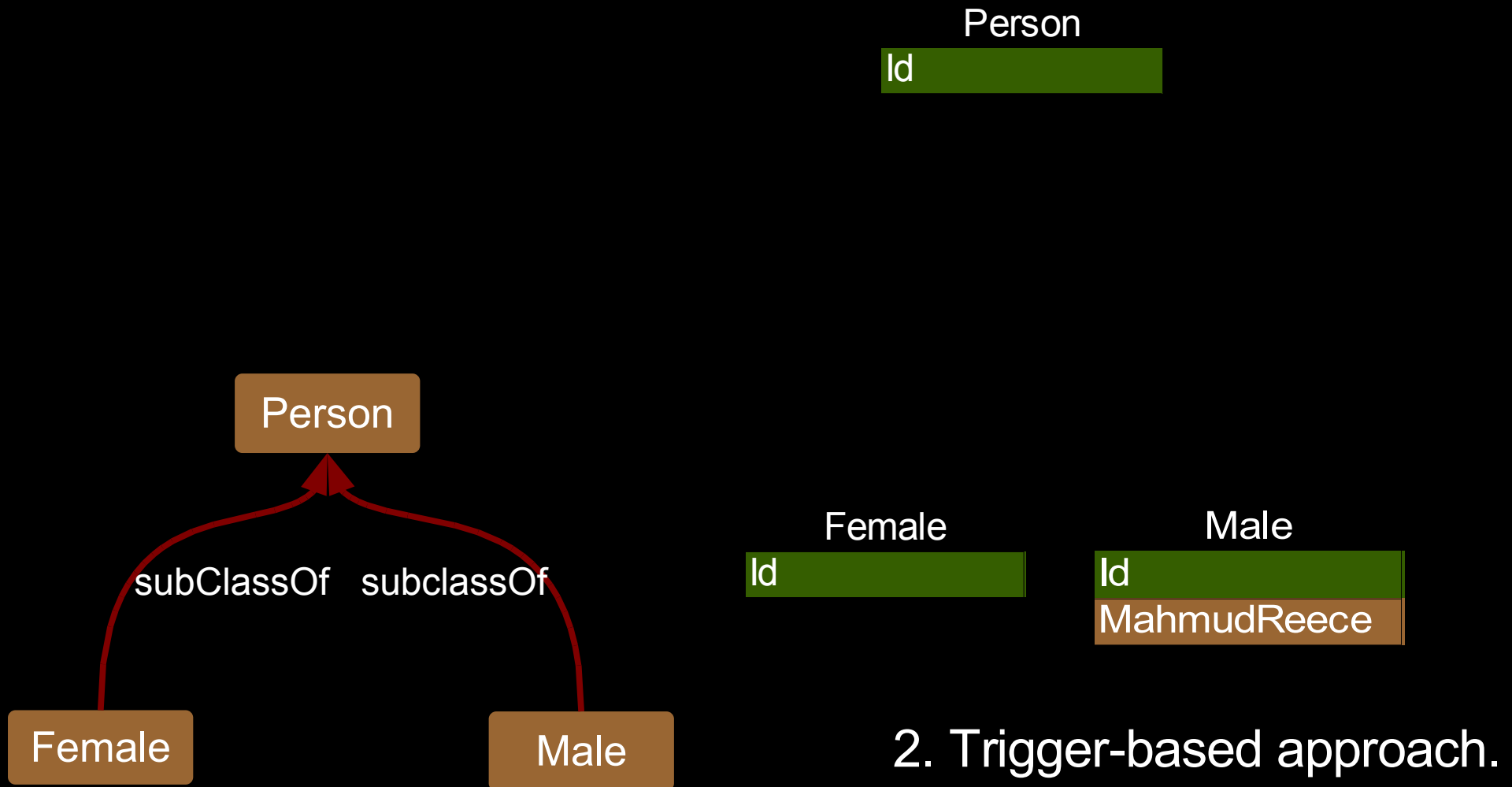
2. Trigger-based approach.

Ontology Databases

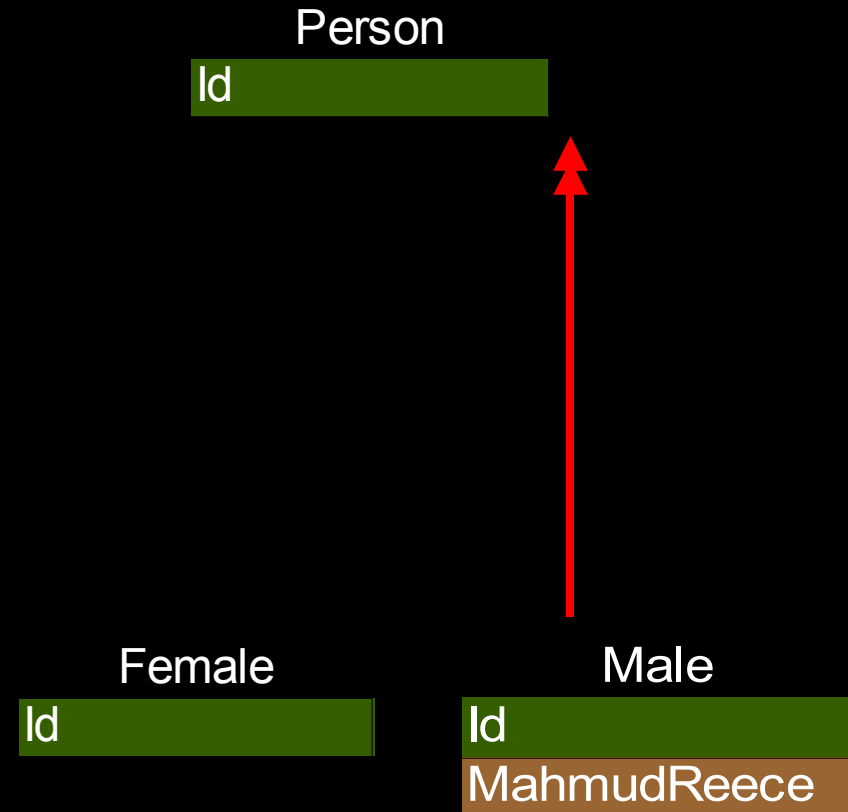
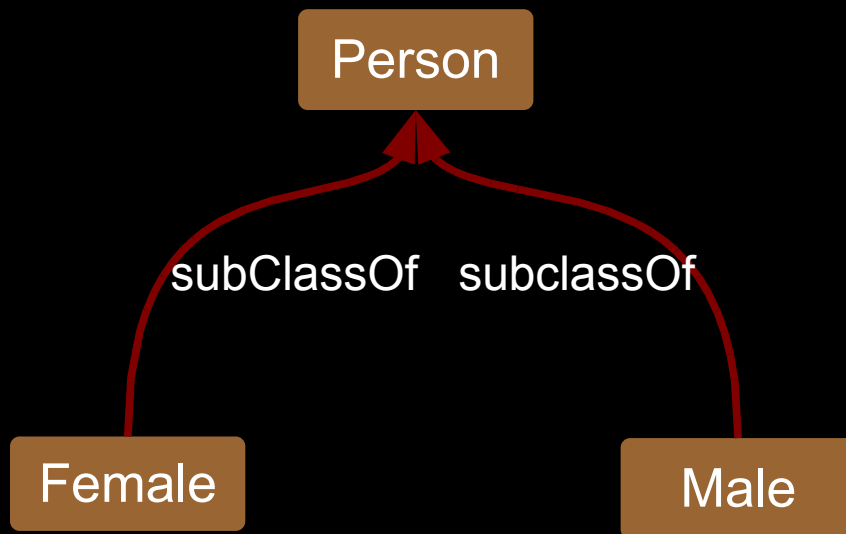


2. Trigger-based approach.

Ontology Databases

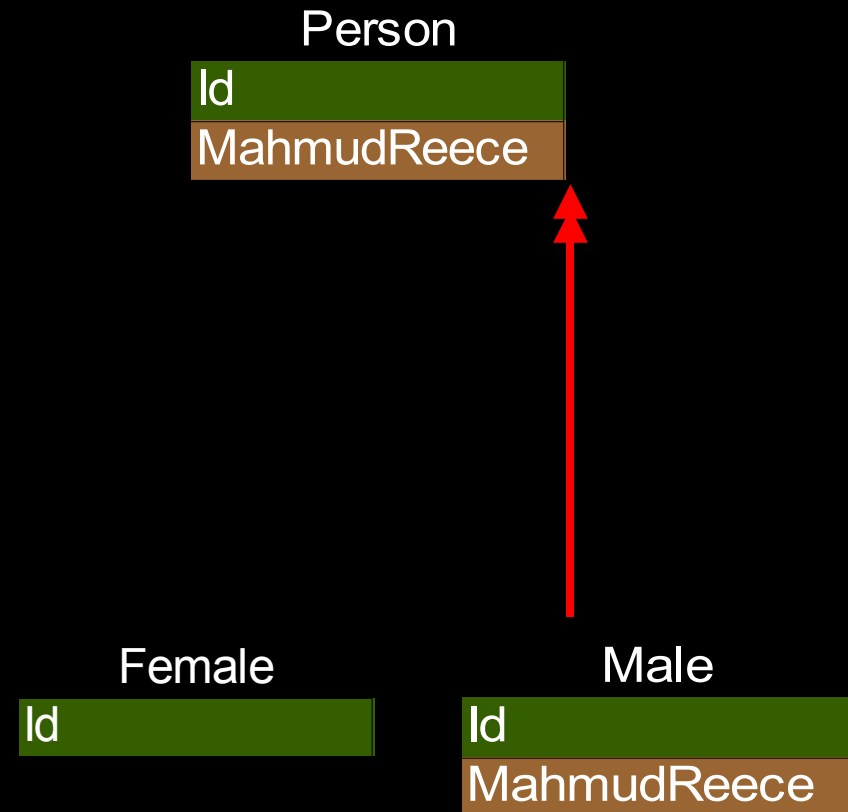
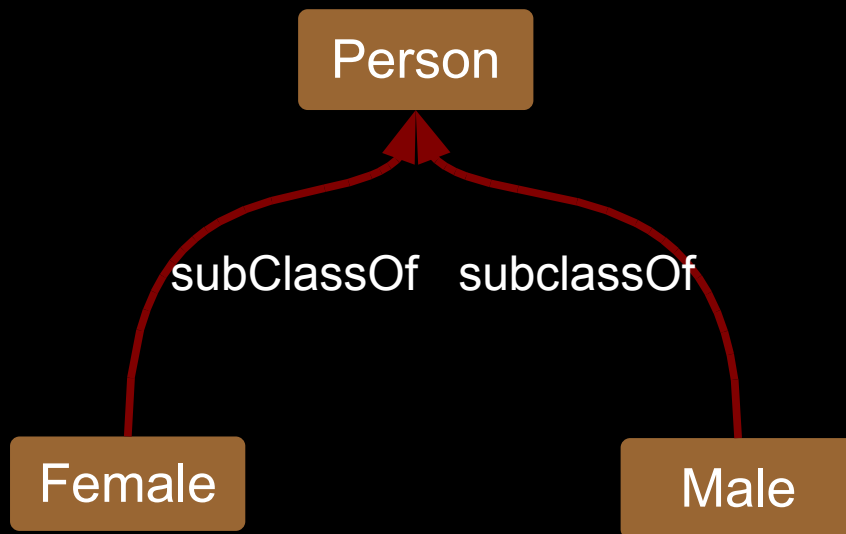


Ontology Databases



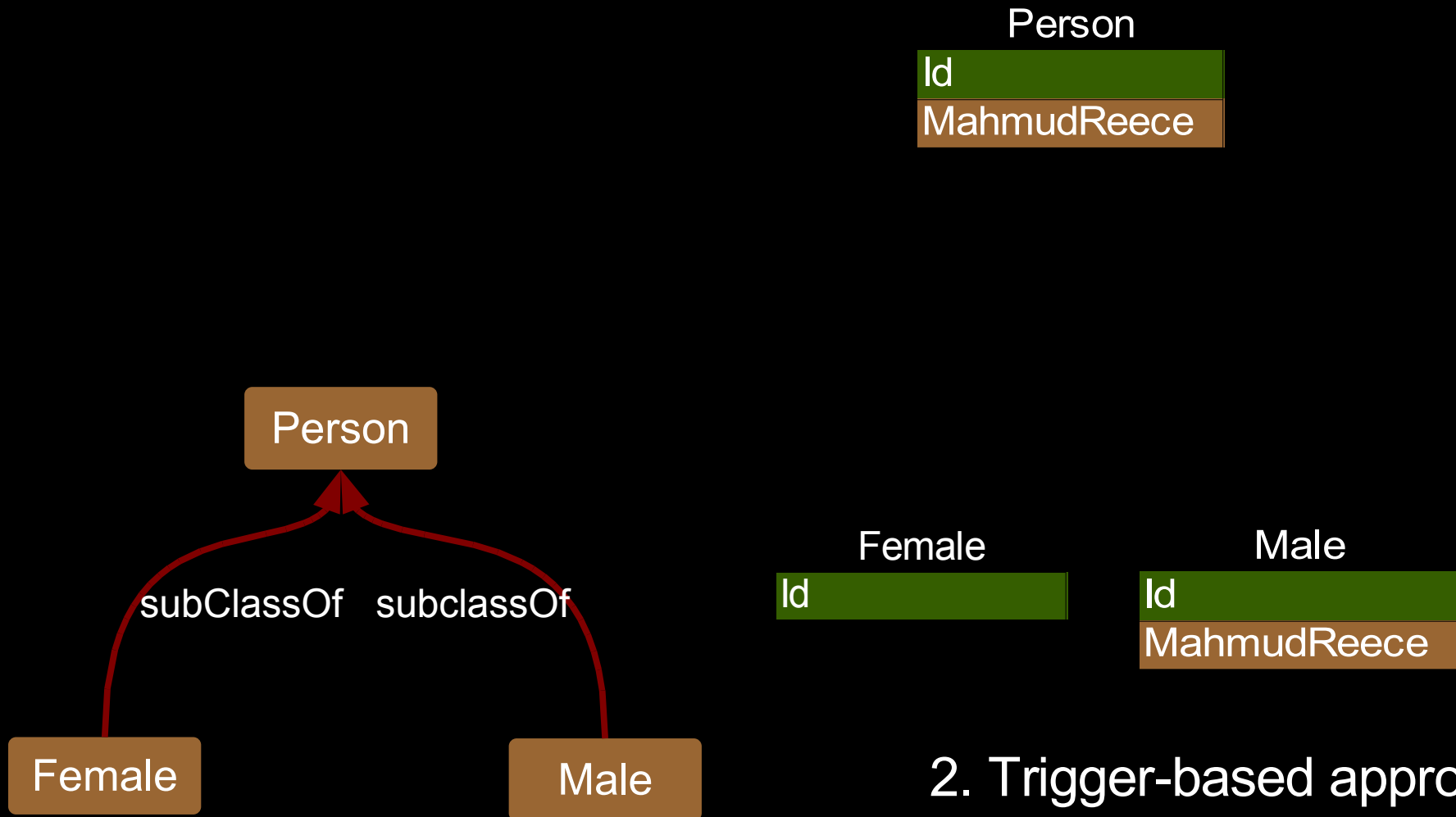
2. Trigger-based approach.

Ontology Databases



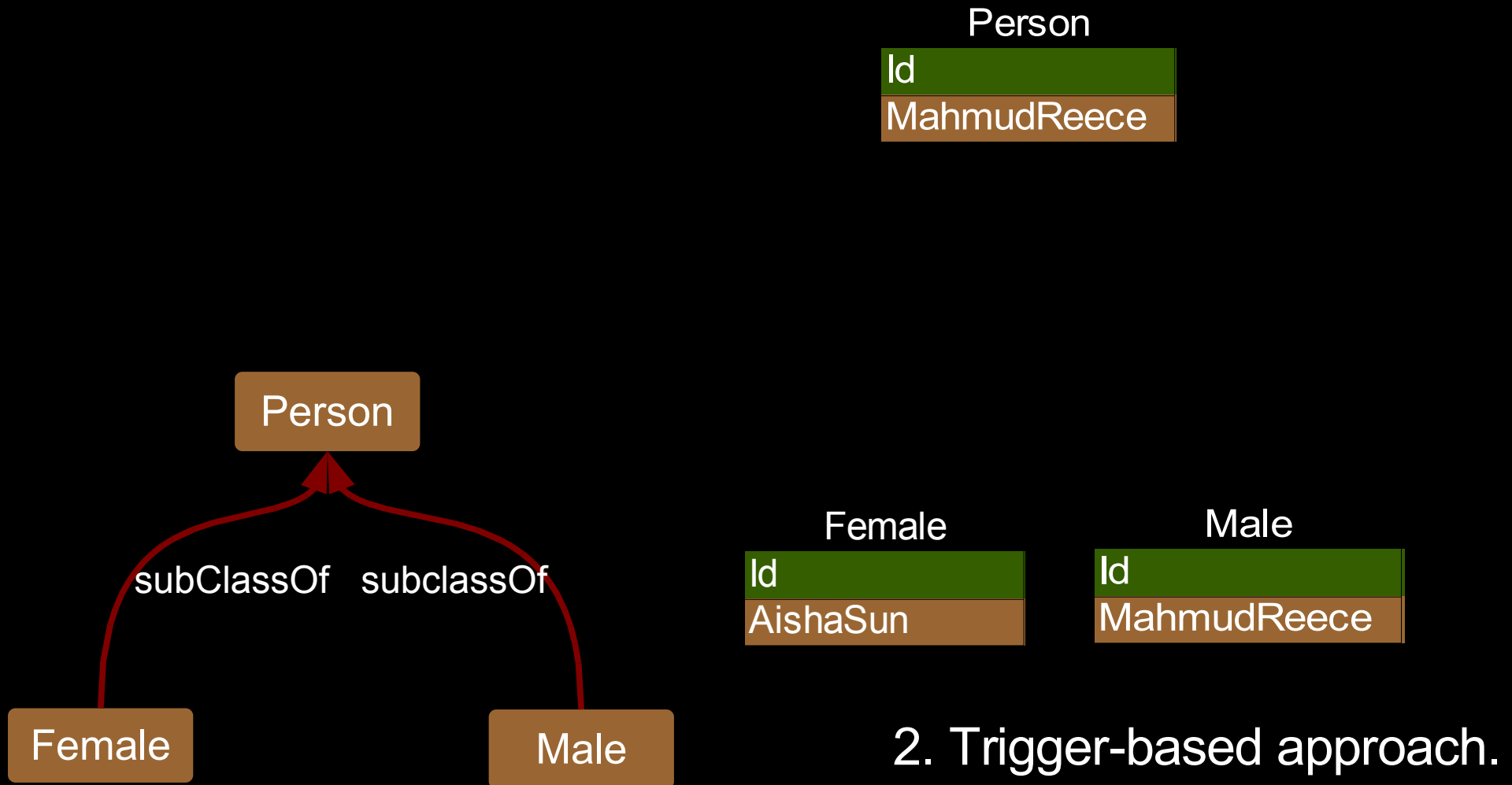
2. Trigger-based approach.

Ontology Databases

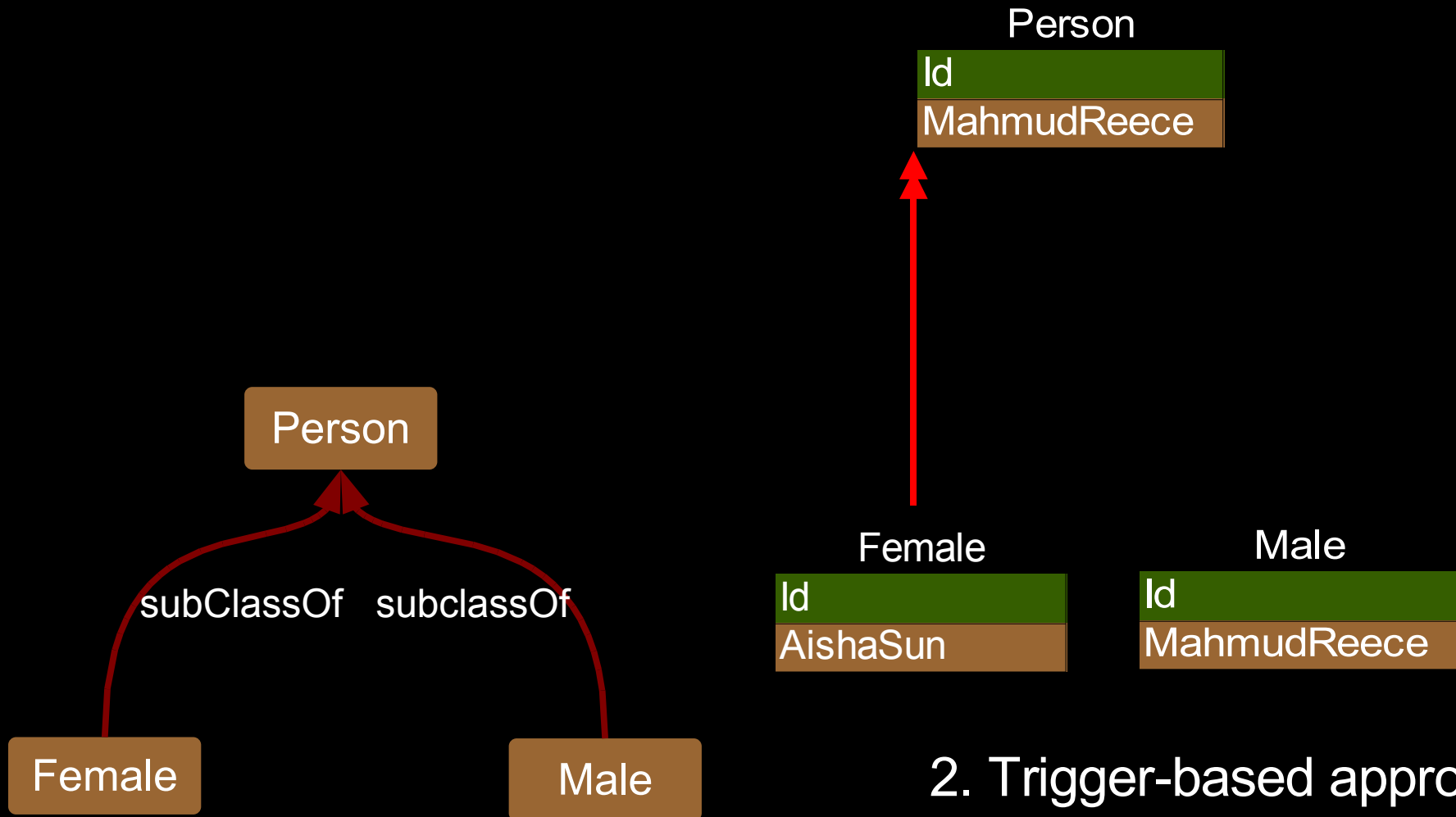


2. Trigger-based approach.

Ontology Databases

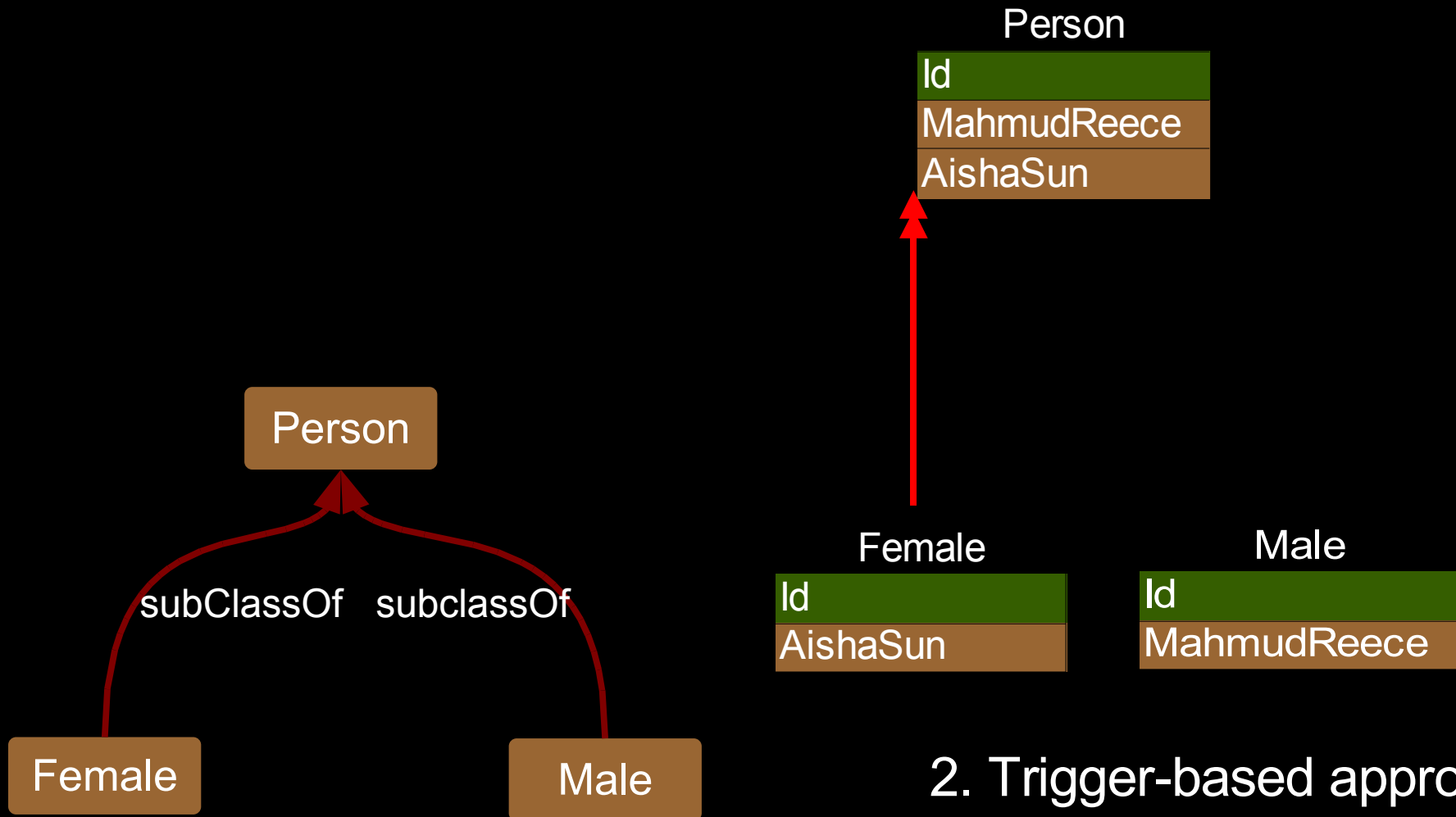


Ontology Databases



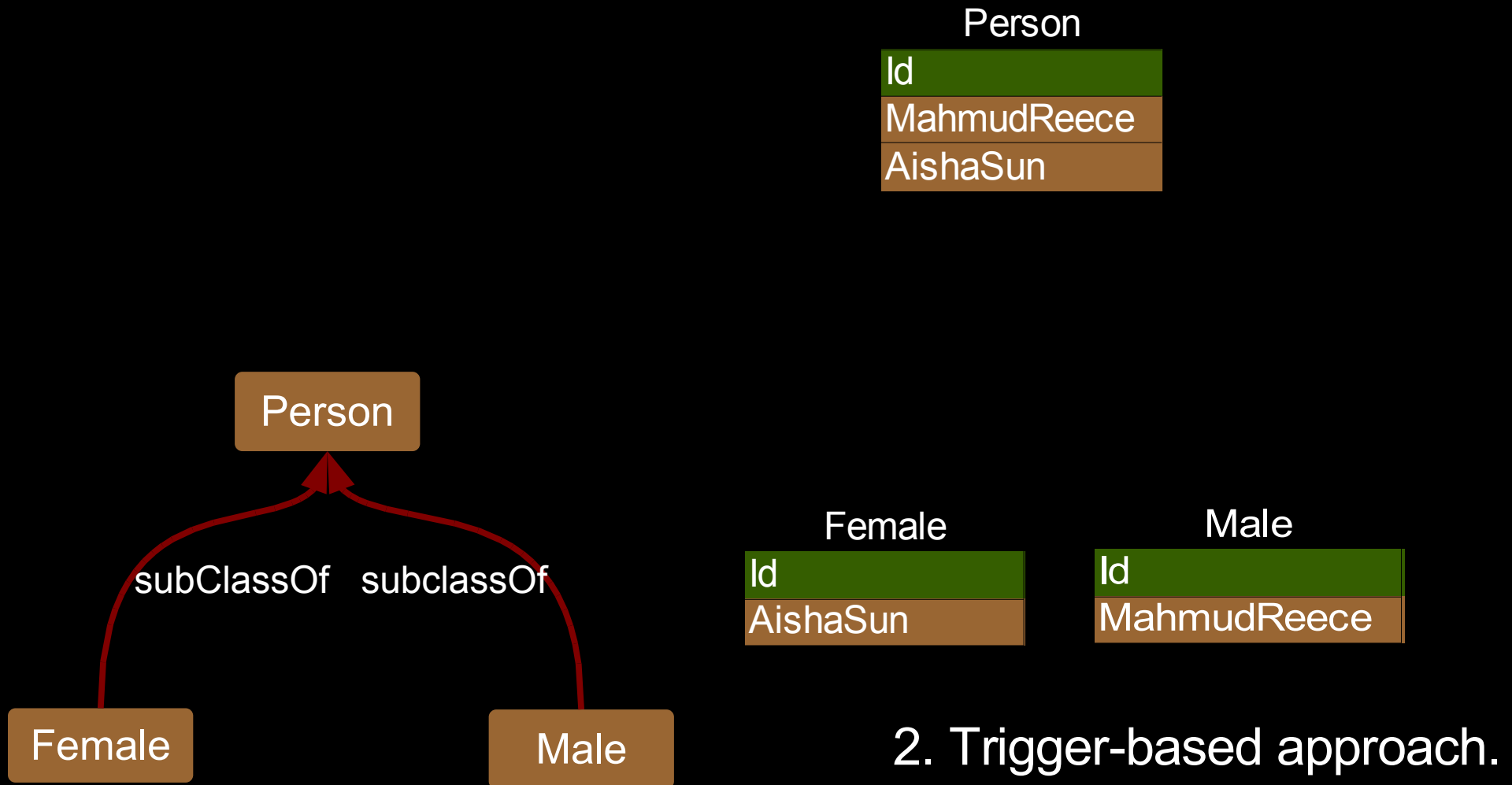
2. Trigger-based approach.

Ontology Databases

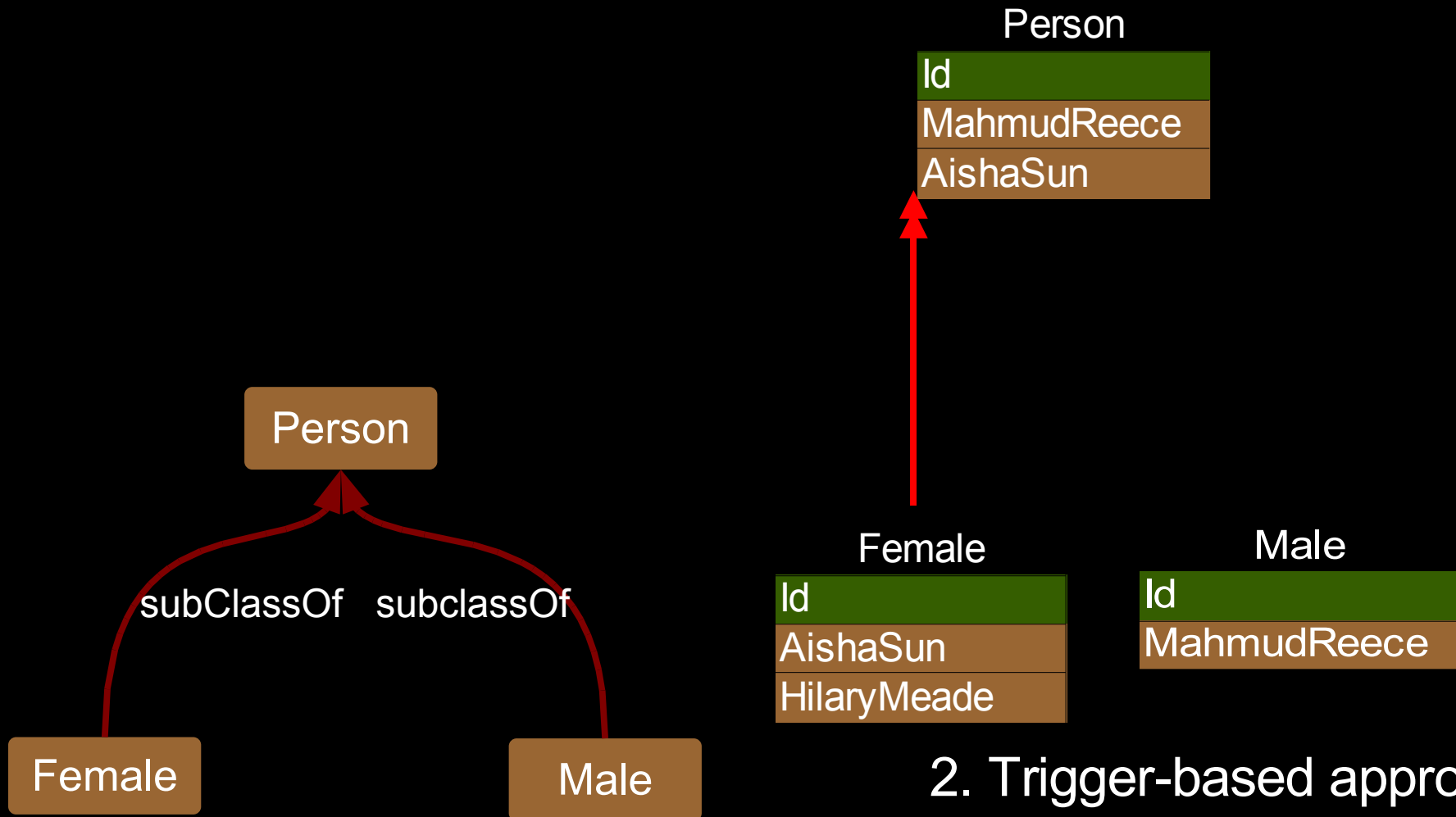


2. Trigger-based approach.

Ontology Databases

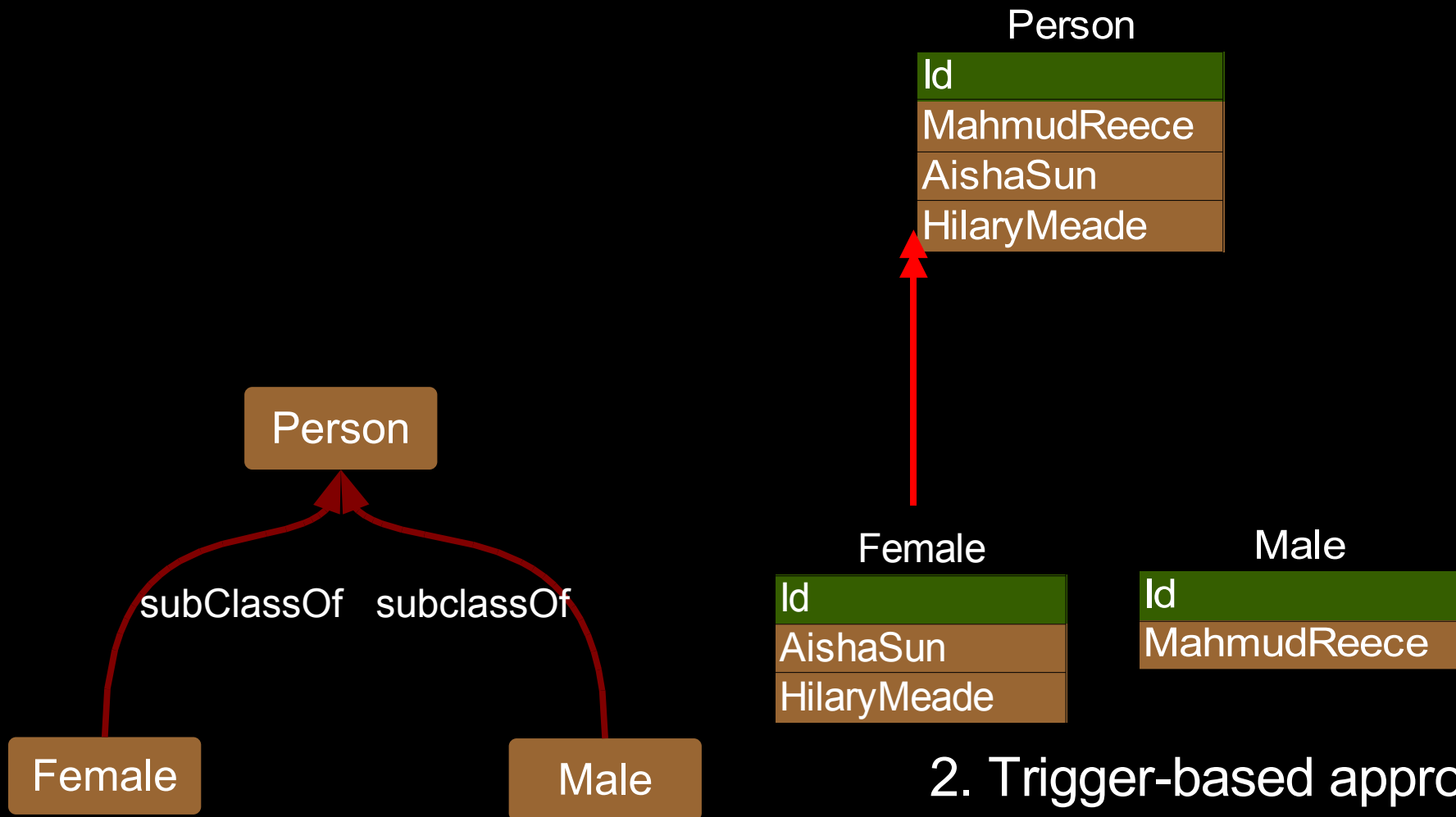


Ontology Databases



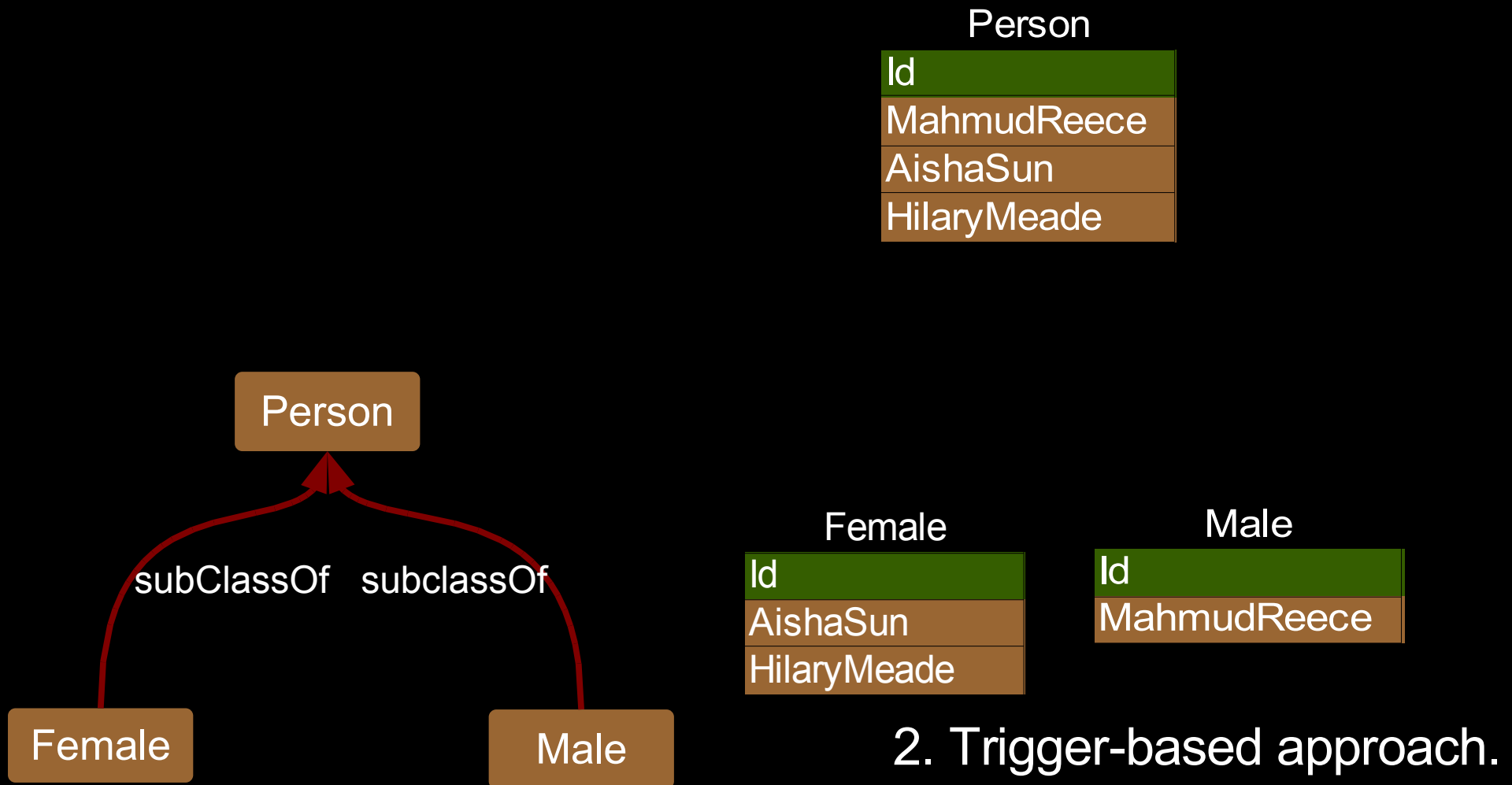
2. Trigger-based approach.

Ontology Databases

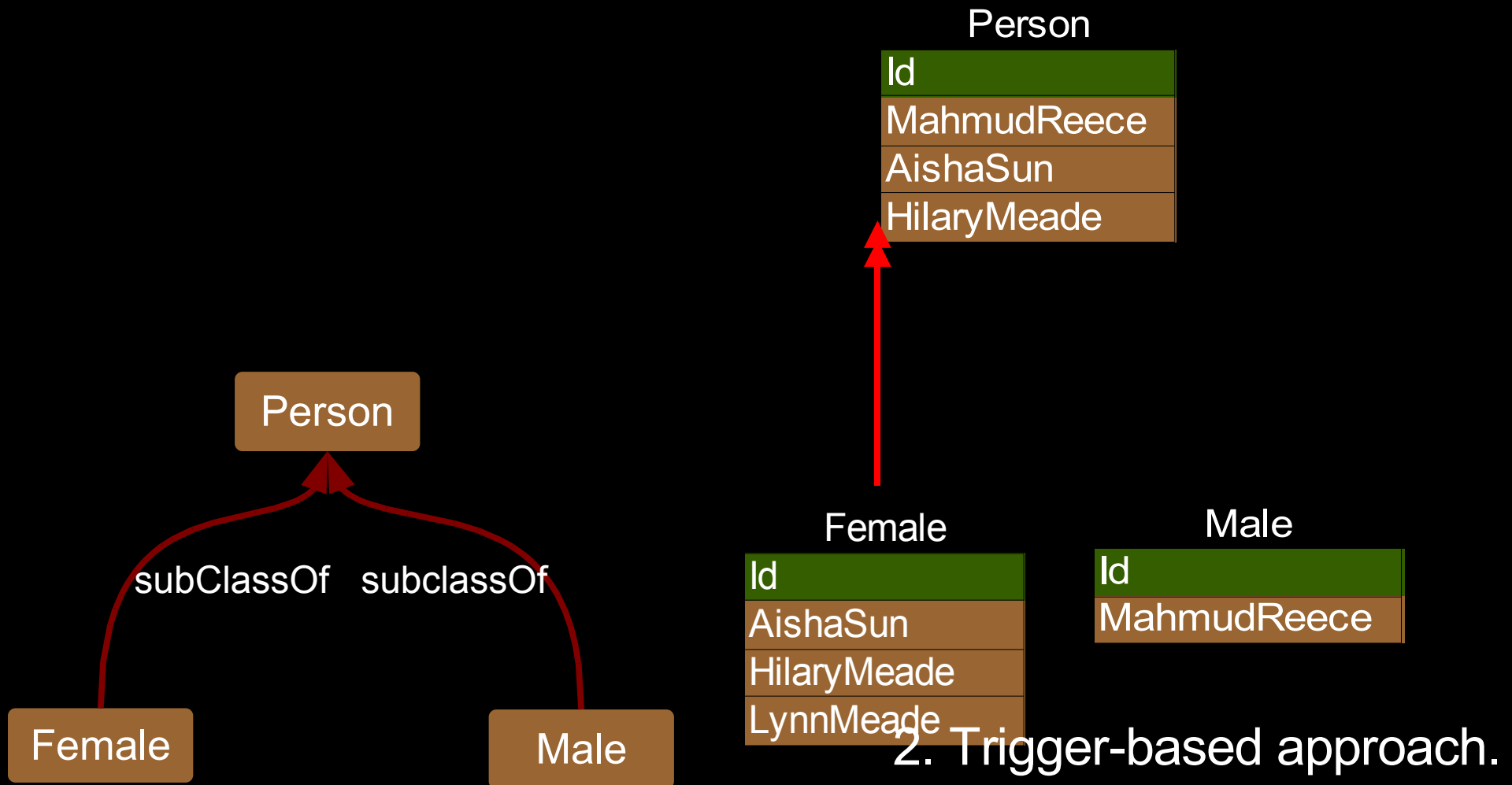


2. Trigger-based approach.

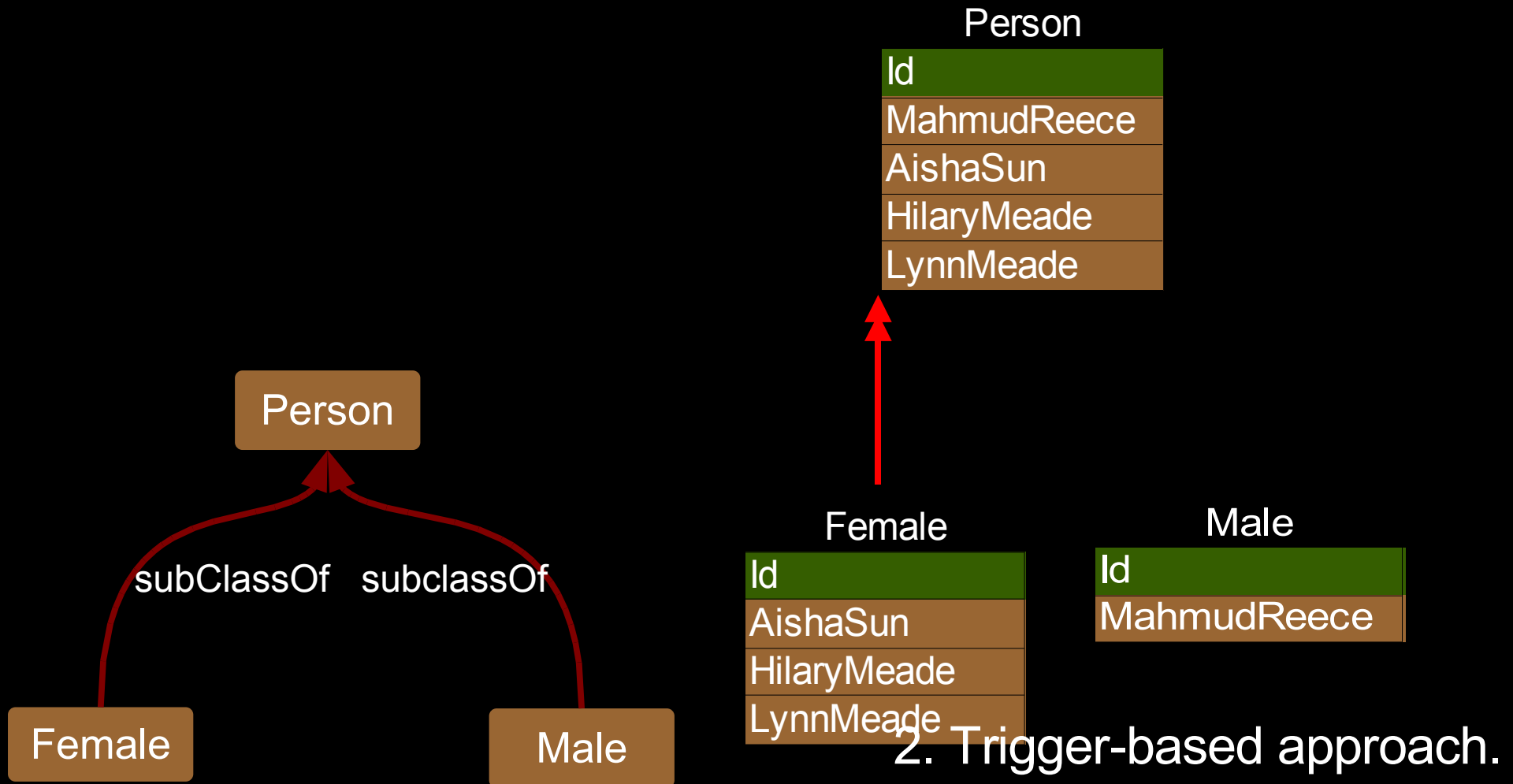
Ontology Databases



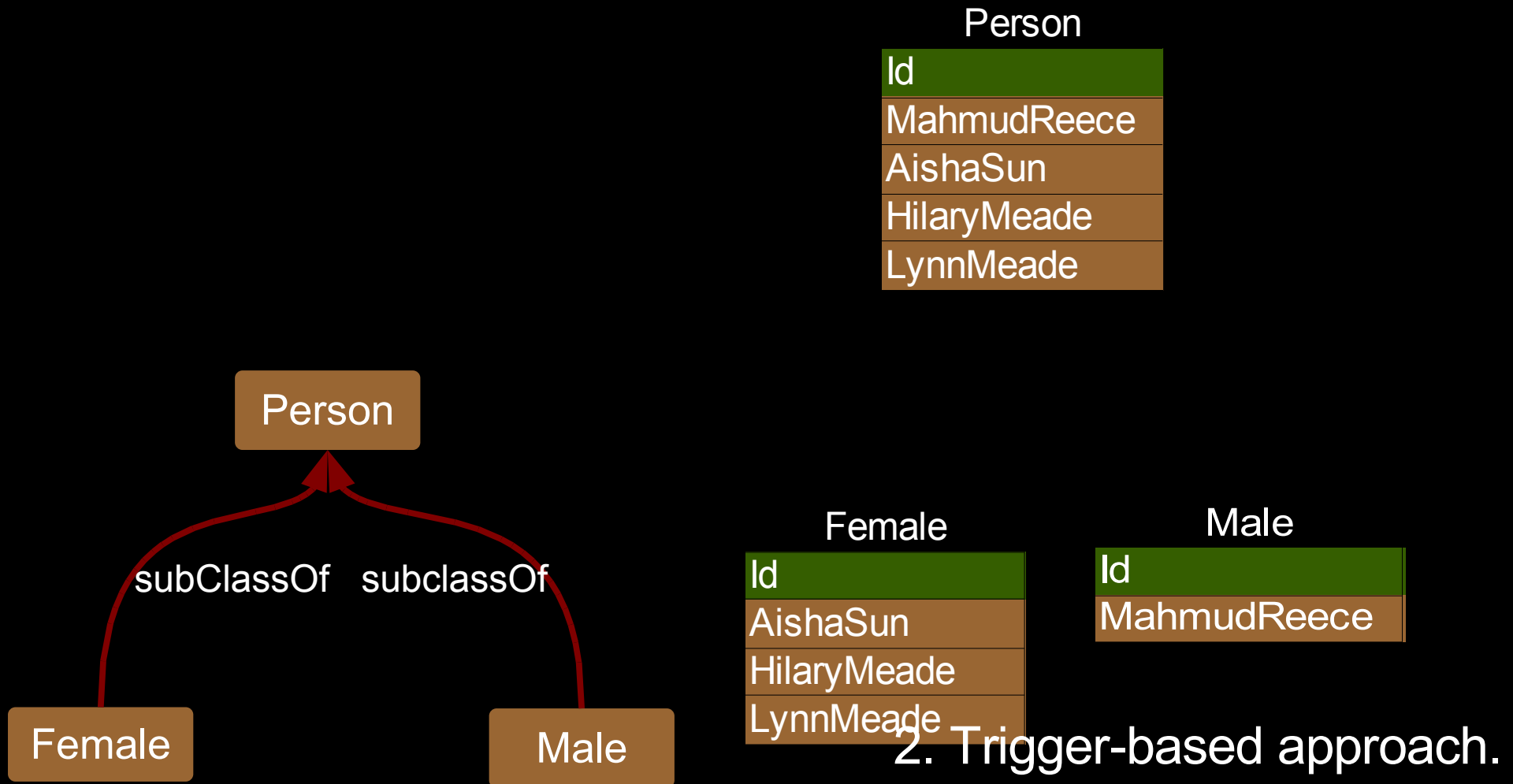
Ontology Databases



Ontology Databases



Ontology Databases



Ontology Databases

subProperty axioms

2. Trigger-based approach.

Ontology Databases

subProperty axioms

(basically the same idea)

2. Trigger-based approach.

Ontology Databases

OntoDB [SSDBM '08] implements the trigger-based approach.

2. Trigger-based approach.

Ontology Databases

Class

Property

Datatype

Axioms

Objects

Facts

Relation

Attribute

Datatype

keys

constraints

triggers

tuples

Ontology Databases

Class

Property

Datatype

Axioms

Objects

Facts

Relation

Attribute

Datatype

keys

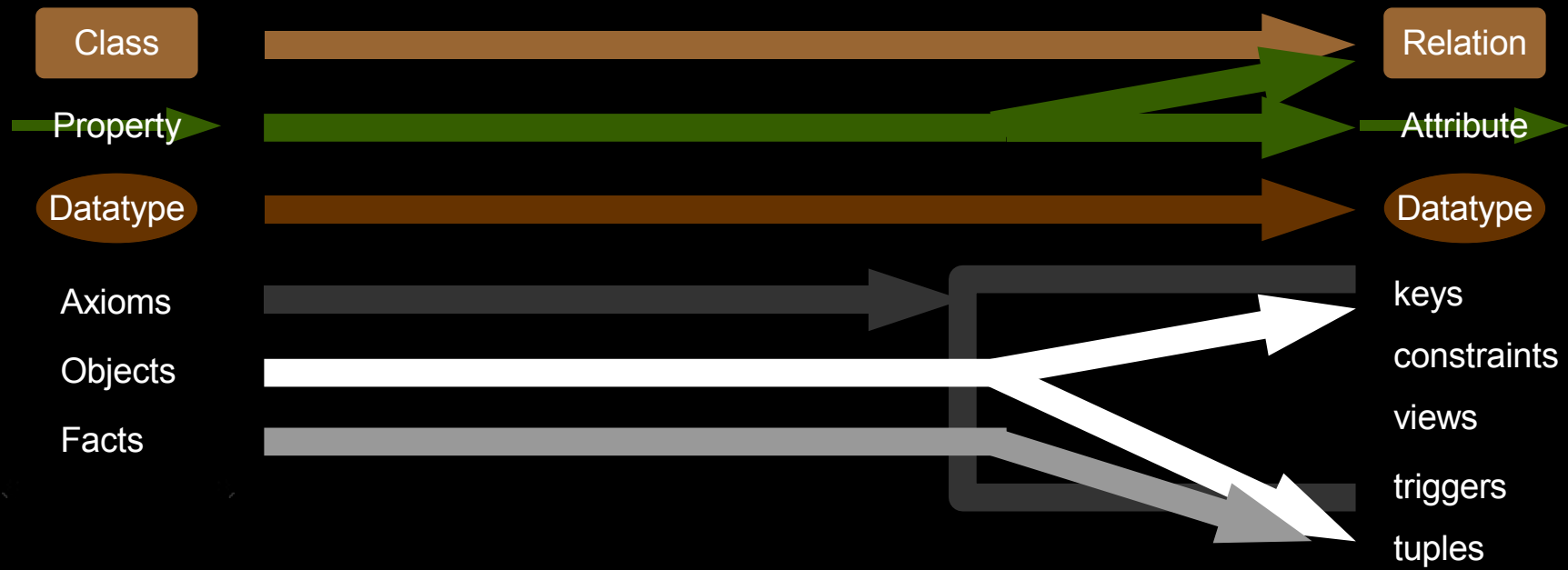
constraints

triggers

tuples

Now we have bridged these.

Ontology Databases



Ontology Databases

Class

Property

Datatype

Axioms

Objects

Facts

Relation

Attribute

Datatype

keys

constraints

triggers

tuples

So what?

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.

Hilary and Lynn are sisters.

This is what we want to know :

Who are siblings?

Obviously, the answer should be :

Hilary and Lynn are siblings.

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.

Hilary and Lynn are sisters.

SiblingOf	
Subject	Object

SisterOf	
Subject	Object

This is what we want to know :

Who are siblings?

Obviously, the answer should be :

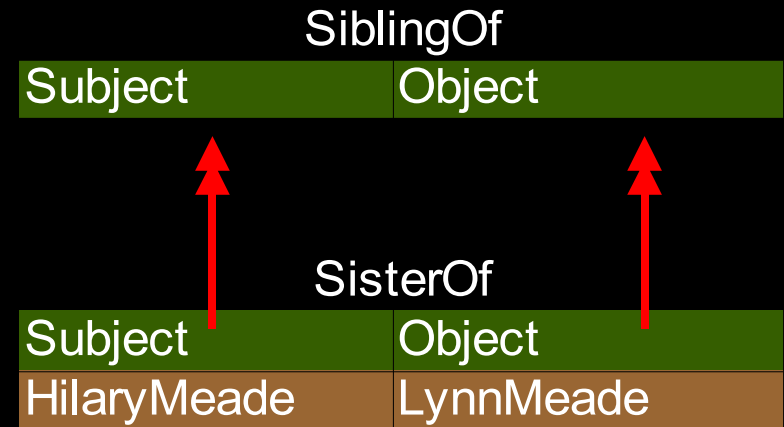
Hilary and Lynn are siblings.

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.

Hilary and Lynn are sisters.



This is what we want to know :

Who are siblings?

Obviously, the answer should be :

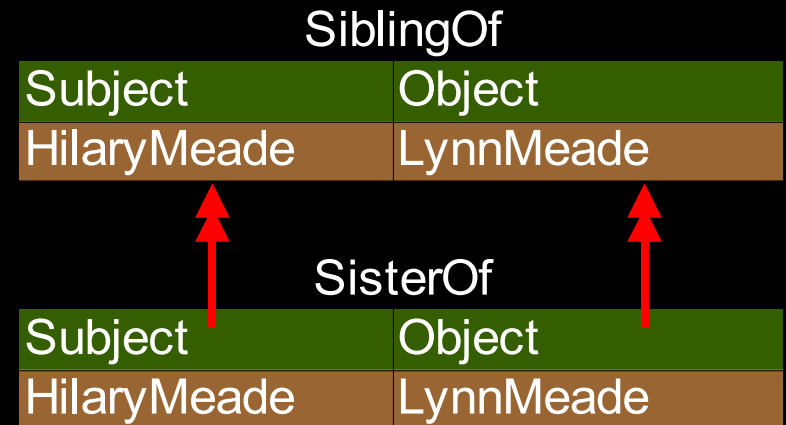
Hilary and Lynn are siblings.

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.

Hilary and Lynn are sisters.



This is what we want to know :

Who are siblings?

Obviously, the answer should be :

Hilary and Lynn are siblings.

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.

Hilary and Lynn are sisters.

SiblingOf

Subject	Object
HilaryMeade	LynnMeade

SisterOf

Subject	Object
HilaryMeade	LynnMeade

This is what we want to know :

Who are siblings?

Obviously, the answer should be :

Hilary and Lynn are siblings.

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.

Hilary and Lynn are sisters.

SiblingOf

Subject	Object
HilaryMeade	LynnMeade

SisterOf

Subject	Object
HilaryMeade	LynnMeade

This is what we want to know :

Who are siblings?

$\{ \langle x,y \rangle \mid \text{siblingOf}(x,y) \}$

Obviously, the answer should be :

Hilary and Lynn are siblings.

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.

Hilary and Lynn are sisters.

SiblingOf

Subject	Object
HilaryMeade	LynnMeade

SisterOf

Subject	Object
HilaryMeade	LynnMeade

This is what we want to know :

Who are siblings?

$\{ \langle x,y \rangle \mid \text{siblingOf}(x,y) \}$

Obviously, the answer should be :

Hilary and Lynn are siblings.

Just look it up!

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.

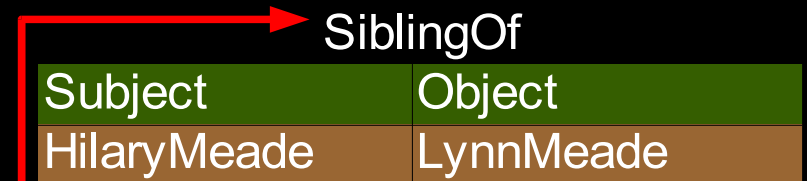
Hilary and Lynn are sisters.

This is what we want to know :

Who are siblings?

Obviously, the answer should be :

Hilary and Lynn are siblings.



SiblingOf	
Subject	Object
HilaryMeade	LynnMeade

SisterOf	
Subject	Object
HilaryMeade	LynnMeade

$\{ \langle x,y \rangle \mid \text{ siblingOf}(x,y) \}$

Just look it up!

A Simple Problem (revisited)

This is what we know :

All sisters are siblings.


Hilary and Lynn are sisters.

This is what we want to know :

Who are siblings?

Obviously, the answer should be :

Hilary and Lynn are siblings.



SiblingOf	
Subject	Object
HilaryMeade	LynnMeade

SisterOf	
Subject	Object
HilaryMeade	LynnMeade

$\{ \langle x, y \rangle \mid \text{siblingOf}(x, y) \}$

$\{ \langle \text{Hilary}, \text{Lynn} \rangle \}$

A Data-Driven Search

This process is data-driven,
loosely based on forward chaining.

A Data-Driven Search

This process is data-driven,
loosely based on forward chaining.

Clearly, we are trading space for query time.

A Data-Driven Search

This process is data-driven,
loosely based on forward chaining.

Clearly, we are trading space for query time.
(We eagerly propagate data.)

Key Question #2

In eagerly propagating data, do we incur a significant load-time cost?

Key Question #2

In eagerly propagating data, do we incur a significant load-time cost?

Probably?

Key Question #3

Do we actually improve query time?

Key Question #3

Do we actually improve query time?

Most likely.

Lehigh University Benchmark

A standard benchmarking suite, which includes:

- the university ontology (department, faculty, student...)
- standard dataset generator
- a set of 14 queries testing various features:
 - subsumption depth
 - instance checking
 - meta features (subProperty, inverse)
 - completeness
 - stars and chains (kinds of joins)

[Lehigh University, SWAT lab, under Jeff Heflin's direction]

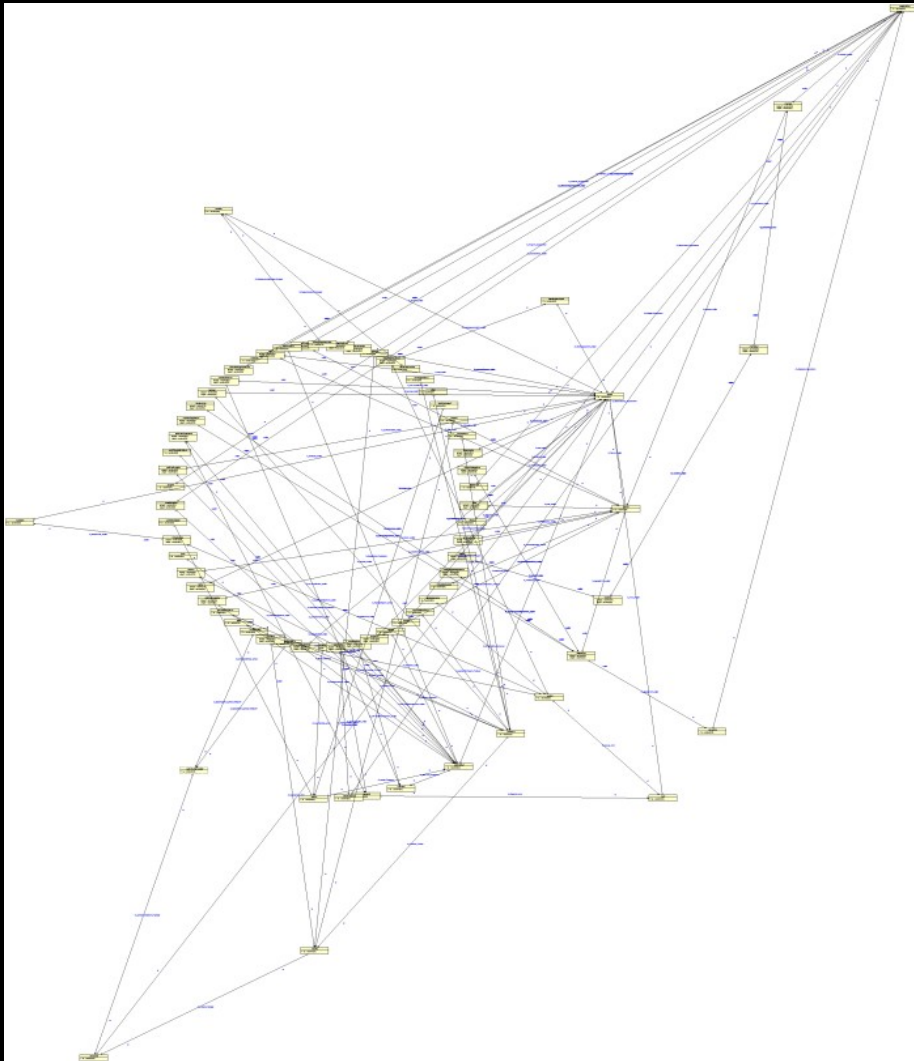
Lehigh University Benchmark

ParentClass	Class
AdministrativeStaff	SystemsStaff
Course	GraduateCourse
Employee	Faculty
Faculty	Lecturer
Faculty	PostDoc
Faculty	Professor
Object	Director
Object	Employee
Object	Organization
Object	Person
Object	Publication
Object	Schedule
Object	Student
Object	TeachingAssistant
Object	Work
Organization	Department
Organization	ResearchGroup
Organization	University
Person	GraduateStudent
Professor	AssistantProfessor
Professor	AssociateProfessor
Professor	FullProfessor
Publication	Software
Publication	Specification
Student	ResearchAssistant
Student	UndergraduateStudent
Work	Course
Work	Research
...etc...	

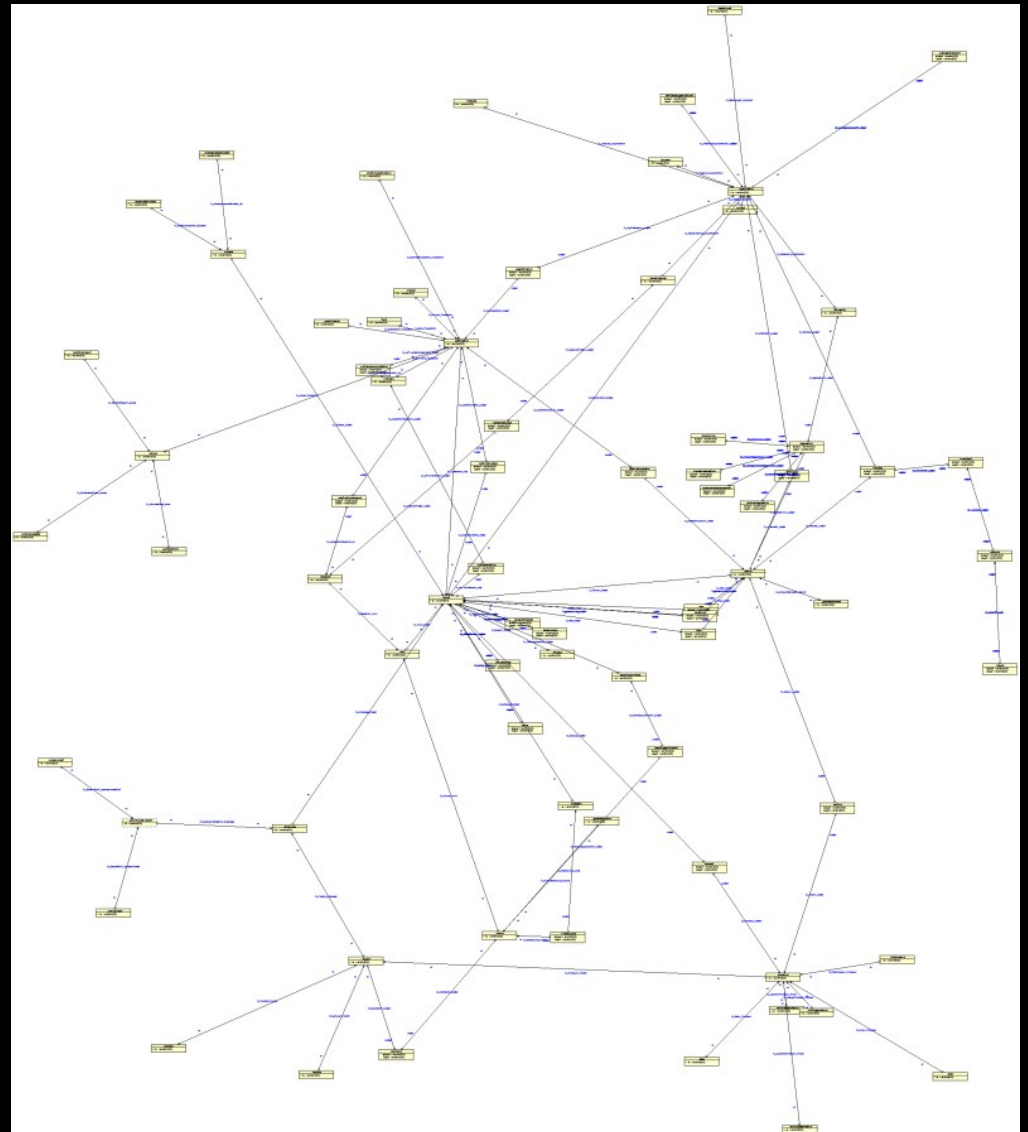
Property
advisor
affiliatedOrganizationOf
affiliateOf
degreeFrom
doctoralDegreeFrom
emailAddress
hasAlumnus
headOf
listedCourse
mastersDegreeFrom
member
memberOf
name
officeNumber
publicationAuthor
publicationDate
publicationResearch
researchInterest
researchProject
softwareDocumentation
softwareVersion
subOrganizationOf
takesCourse
teacherOf
teachingAssistantOf
title
undergraduateDegreeFrom
worksFor
...etc...

Lehigh University Benchmark

Radial Tree View

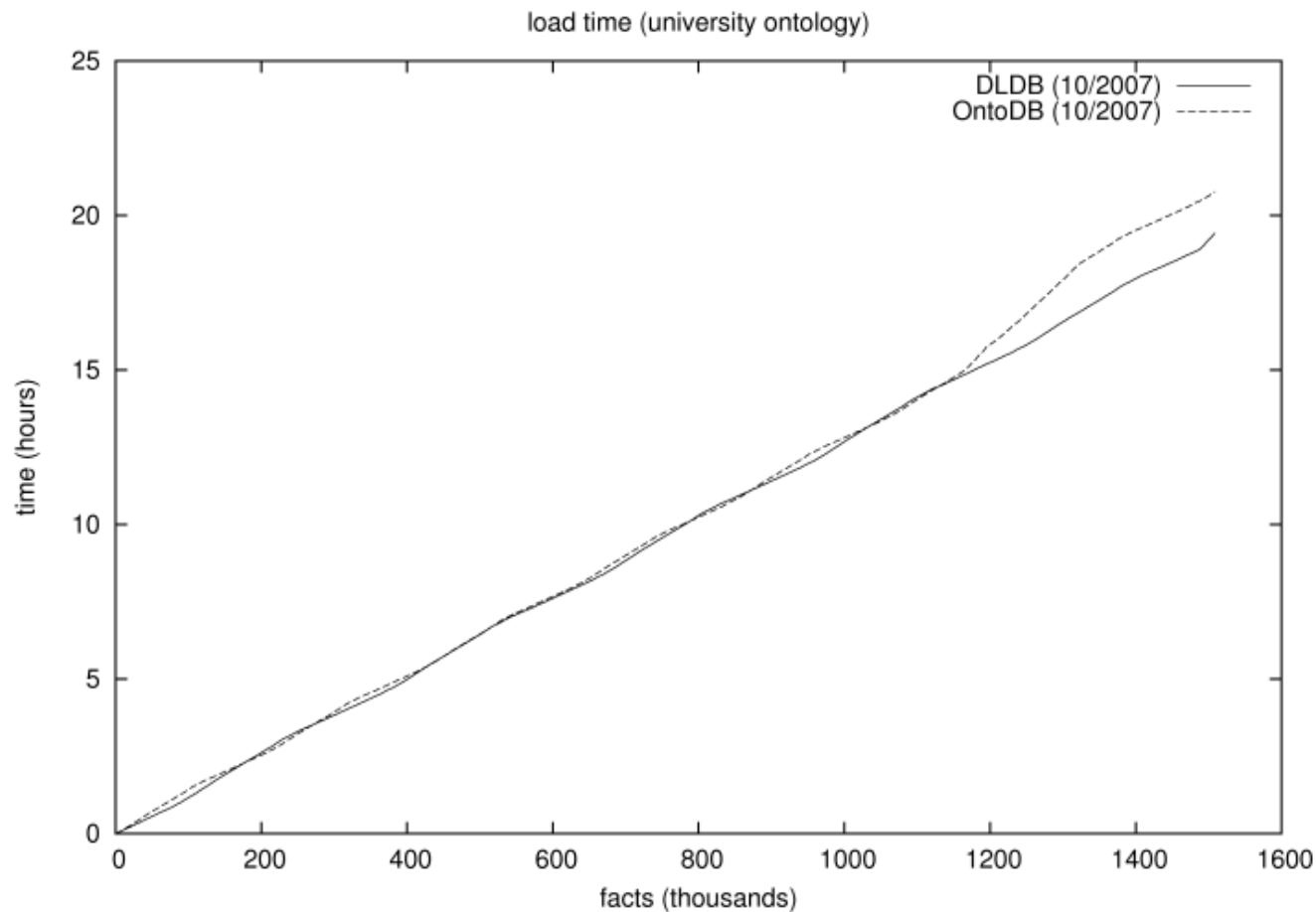


Radial Isometric View



Lehigh University Benchmark

Load Time (1.5 million facts)
(10 Universities, 20 Departments)



Lehigh University Benchmark

In trading space, do we incur a significant load-time cost?

No!

Lehigh University Benchmark

In trading space, do we incur a significant load-time cost?

No!

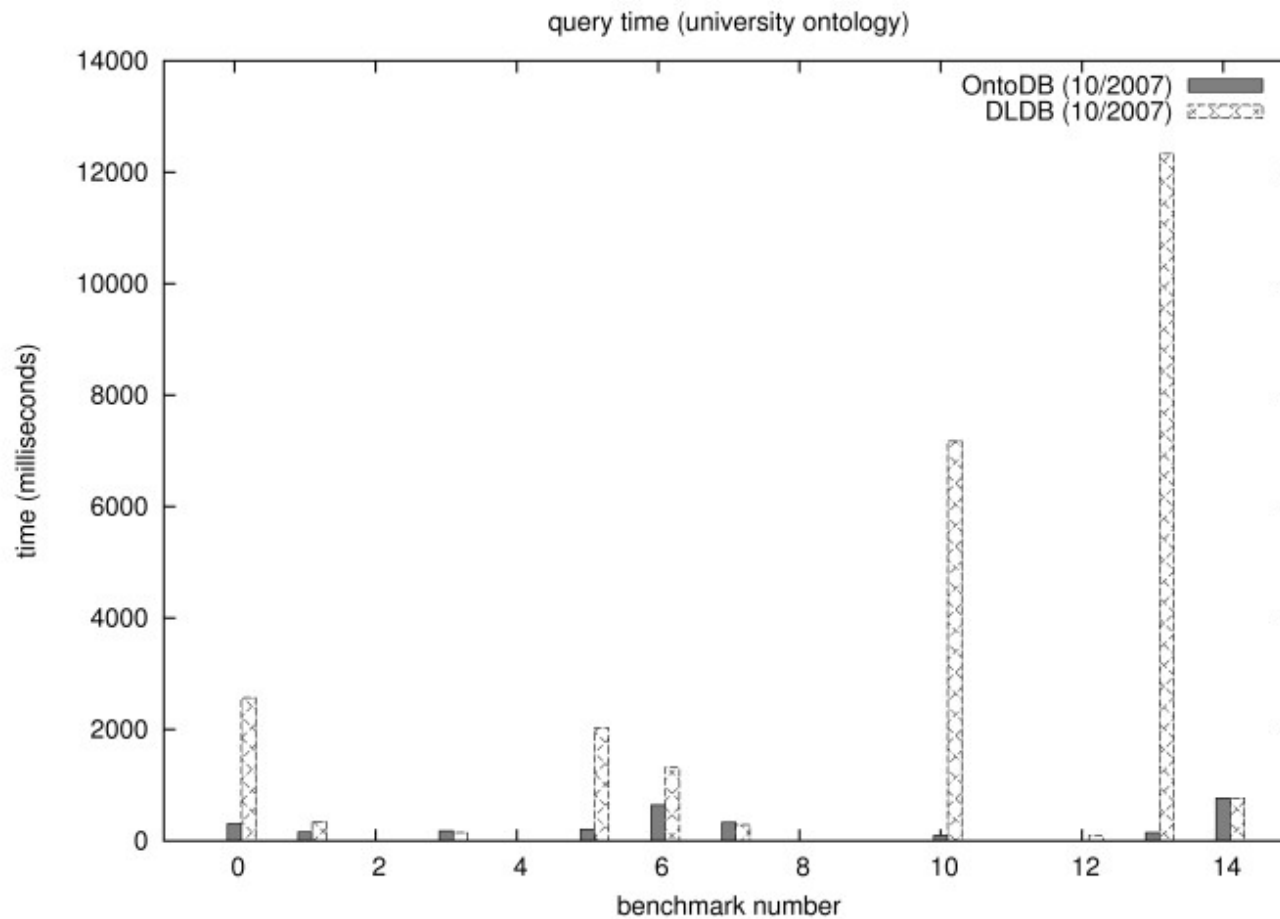
(This was surprising.)

Lehigh University Benchmark

Do we actually improve query time?

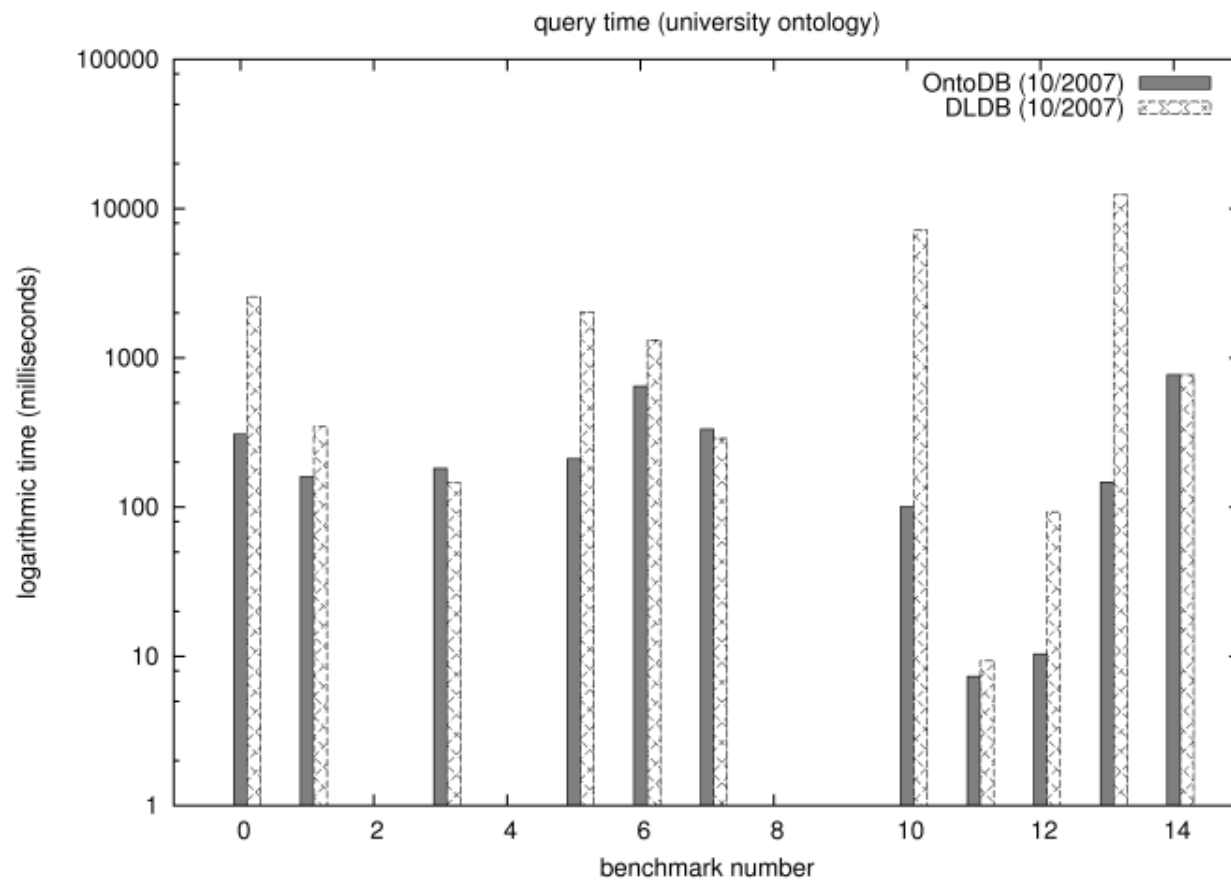
Lehigh University Benchmark

Query Performance



Lehigh University Benchmark

Query Performance
(logarithmic time)



Lehigh University Benchmark

Do we actually improve query time?

Yes!

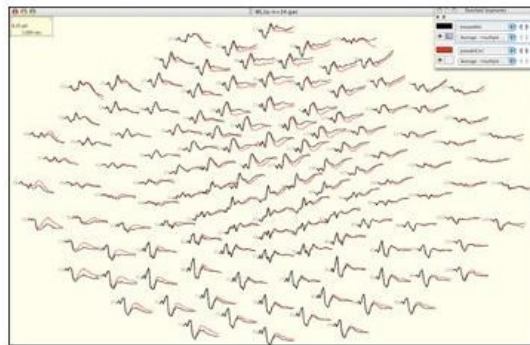
Lehigh University Benchmark

Do we actually improve query time?

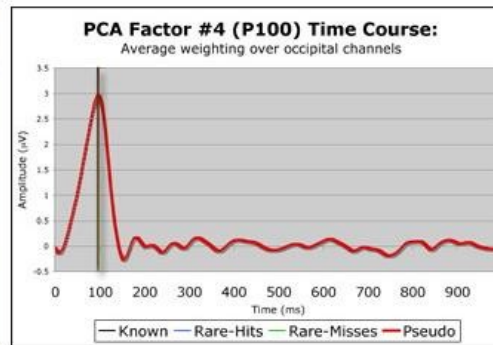
Yes!

As we expected.

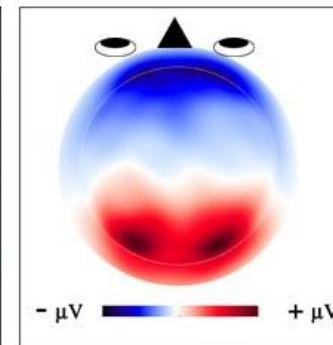
NEMO



(A)

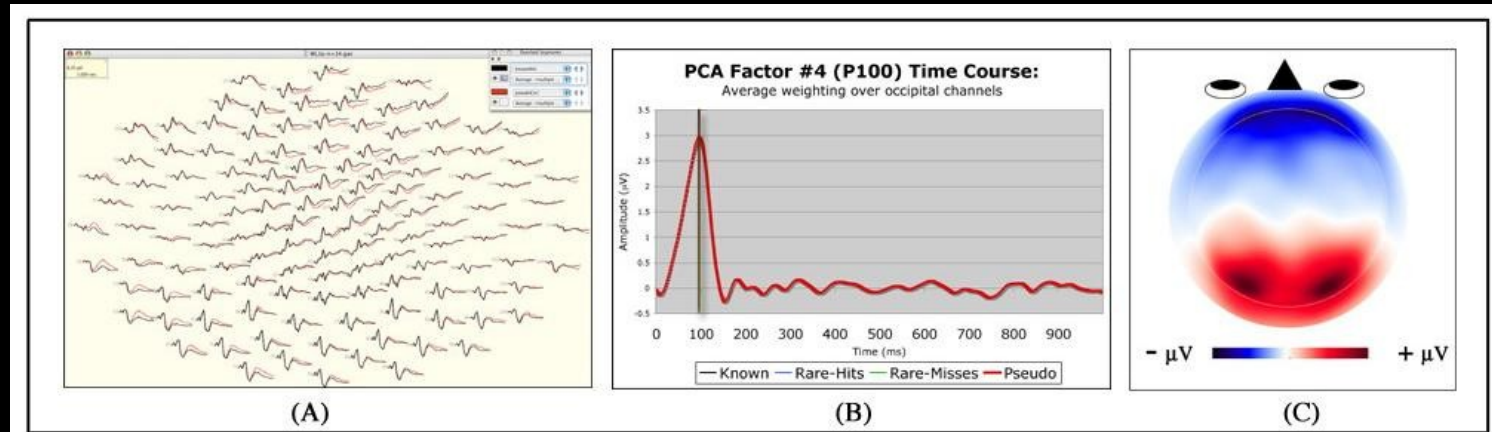


(B)



(C)

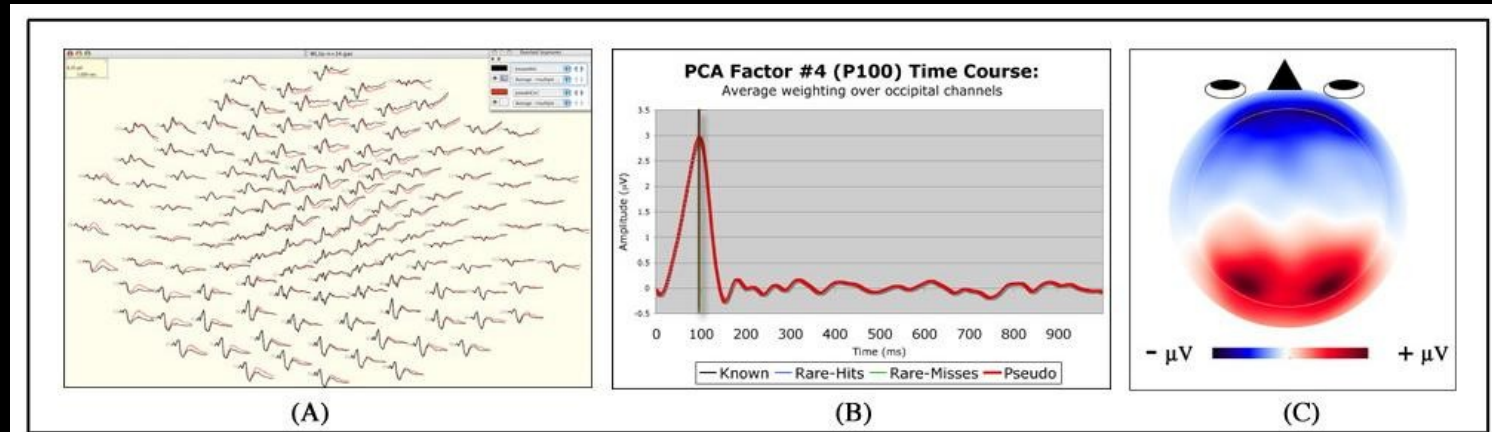
NEMO



Expert queries answered 100% correctly.

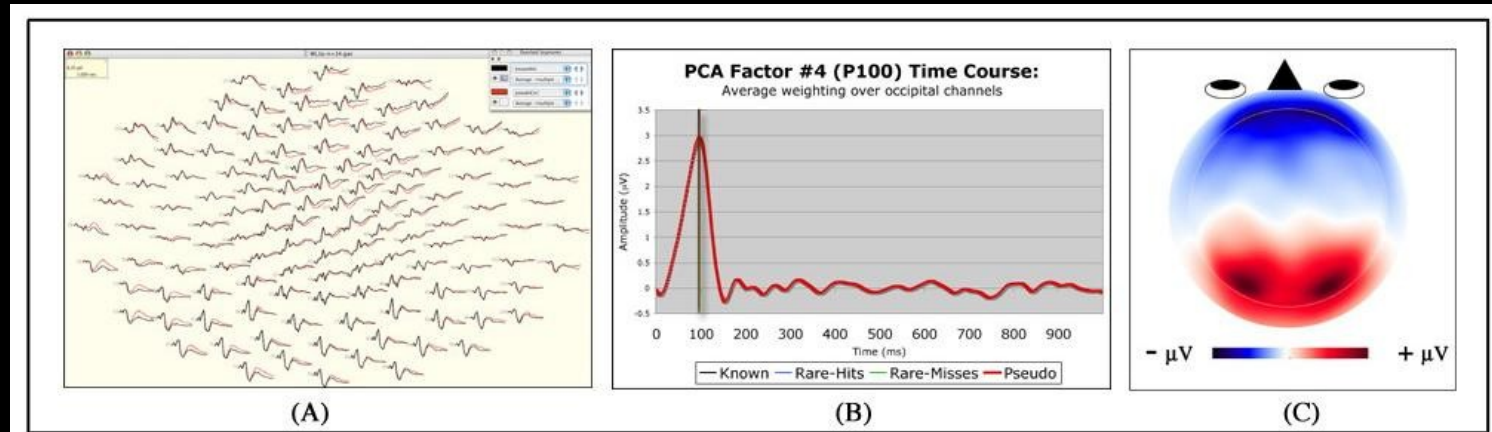
Less than 10 millisecond average response time, regardless of query complexity.

NEMO



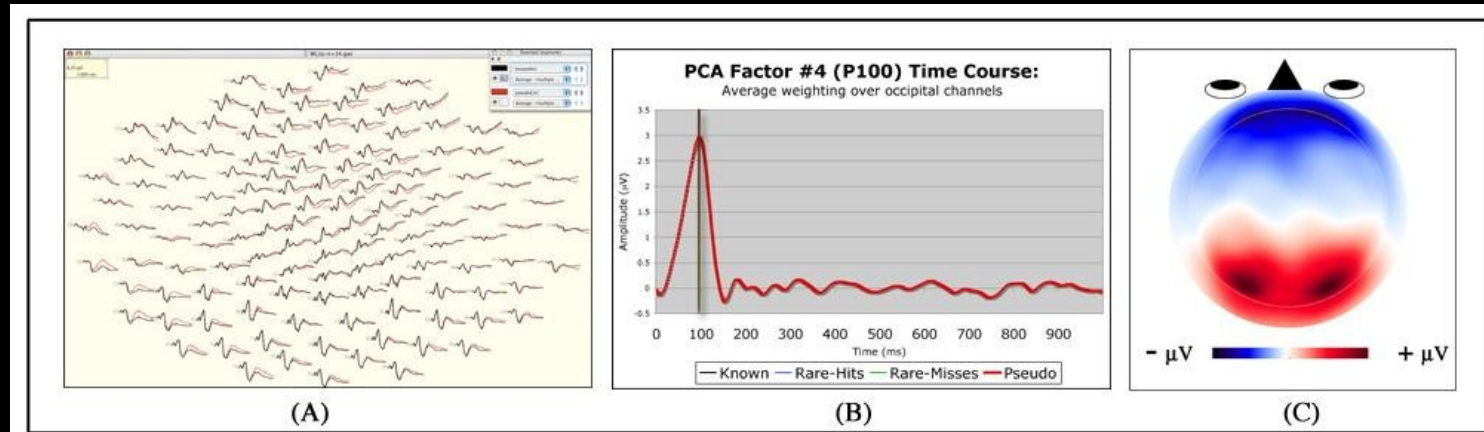
- Show the region of interest for all ERP patterns that occur between 0 and 300ms.
- Which PCA factor do P100 patterns most often appear in?
- What is the range of intensity mean for the region of interest for N100 patterns?
- Show the patterns whose region of interest is left occipital and occurs between 220 and 300ms.

NEMO



Main points:

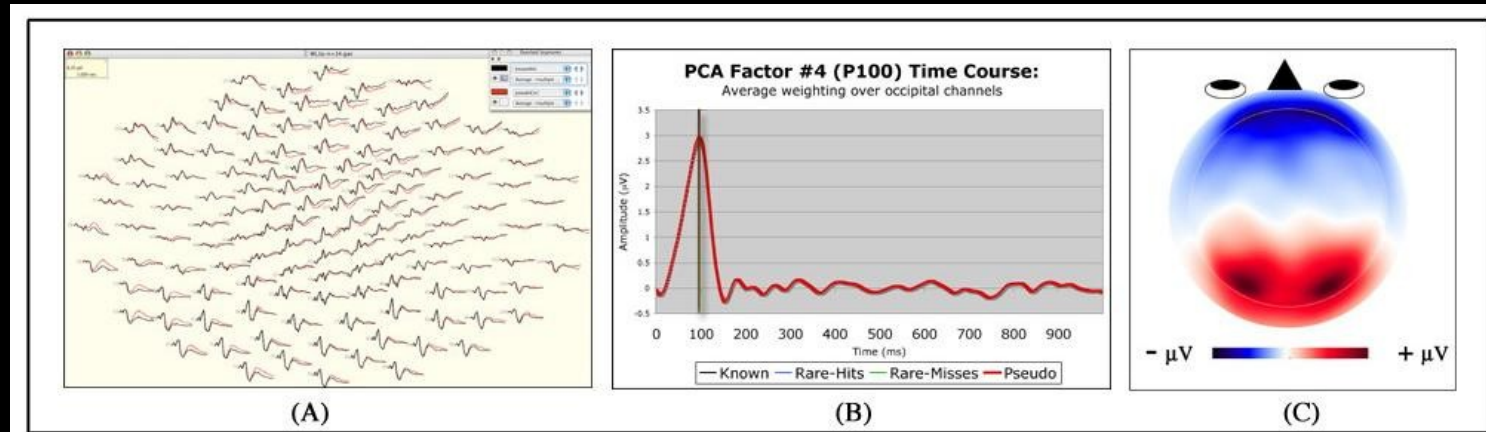
NEMO



Main points:

Ontology-based Modeling

NEMO

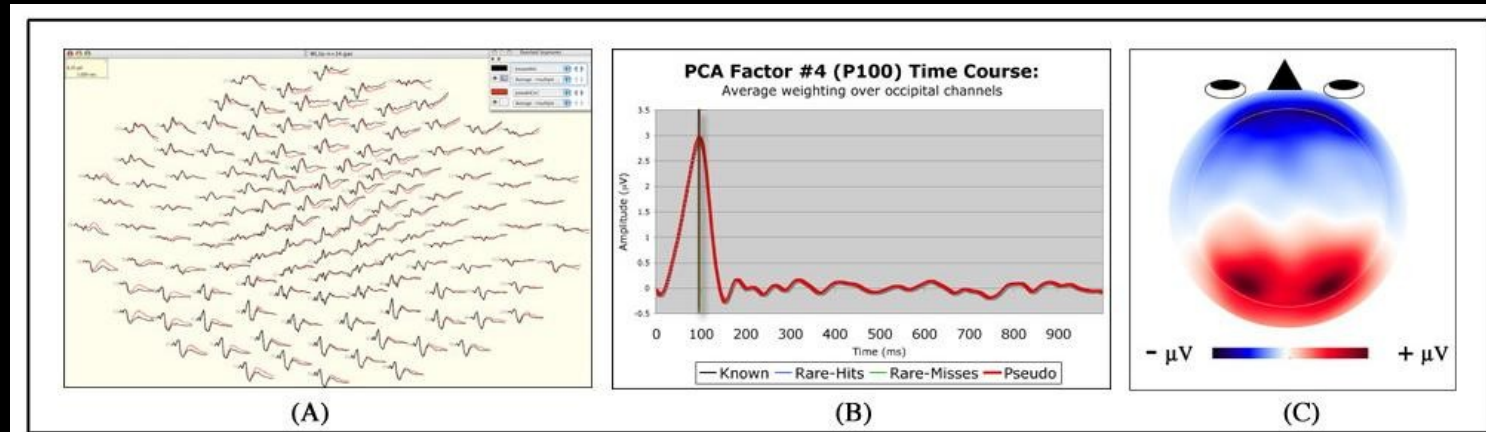


Main points:

Ontology-based Modeling

Ontology-based Query Answering Process

NEMO



Main points:

Ontology-based Modeling

Ontology-based Query Answering Process

Cross-lab information modeling, storage and analysis

Ontology Databases

Ongoing Work

- Disjunctive Logical Models

- Scalable T-Box Reasoning (model-based)

- Meta-analyses (cross-lab integration)

Thank you!

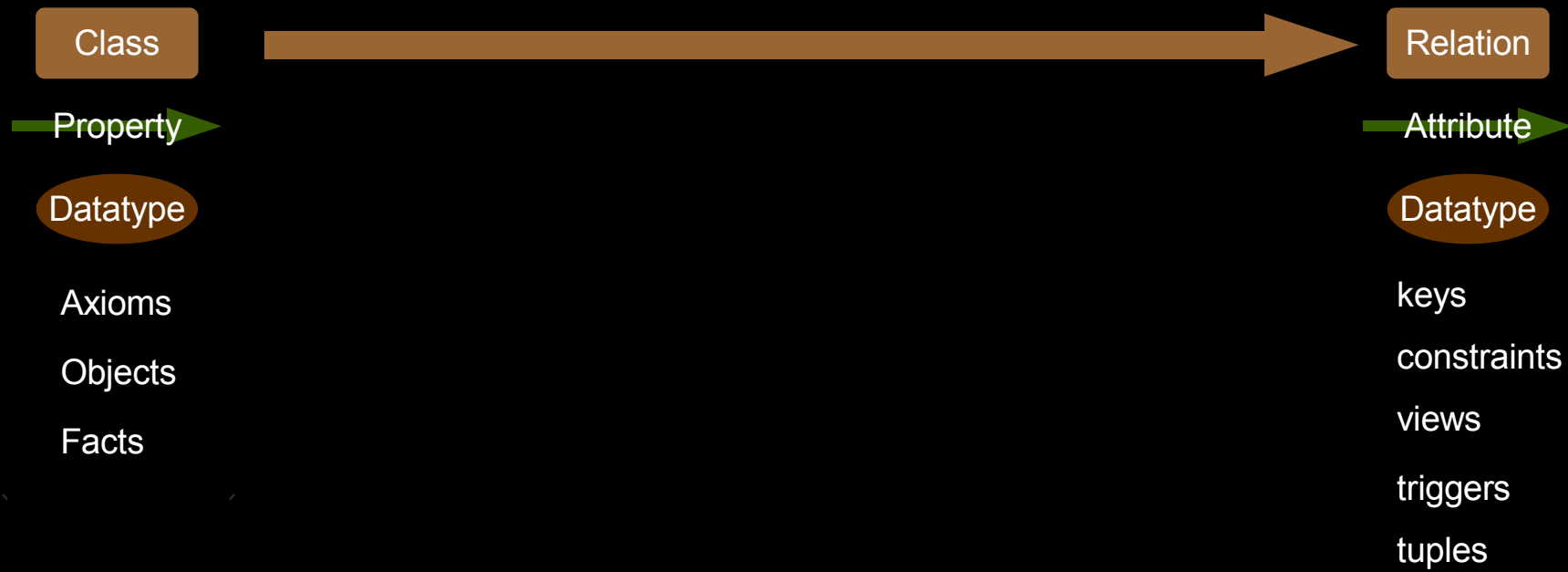
paea@cs.uoregon.edu

Ontology Databases

Questions?

paea@cs.uoregon.edu

Ontology Databases



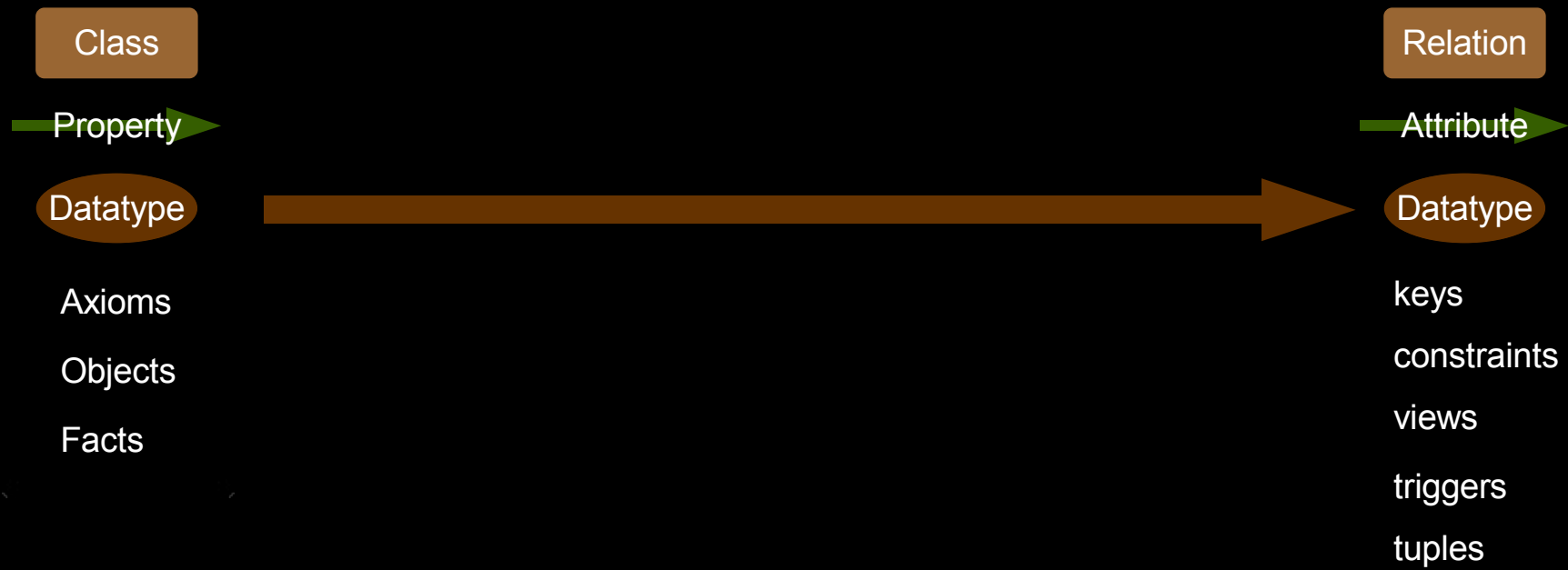
Ontology Databases



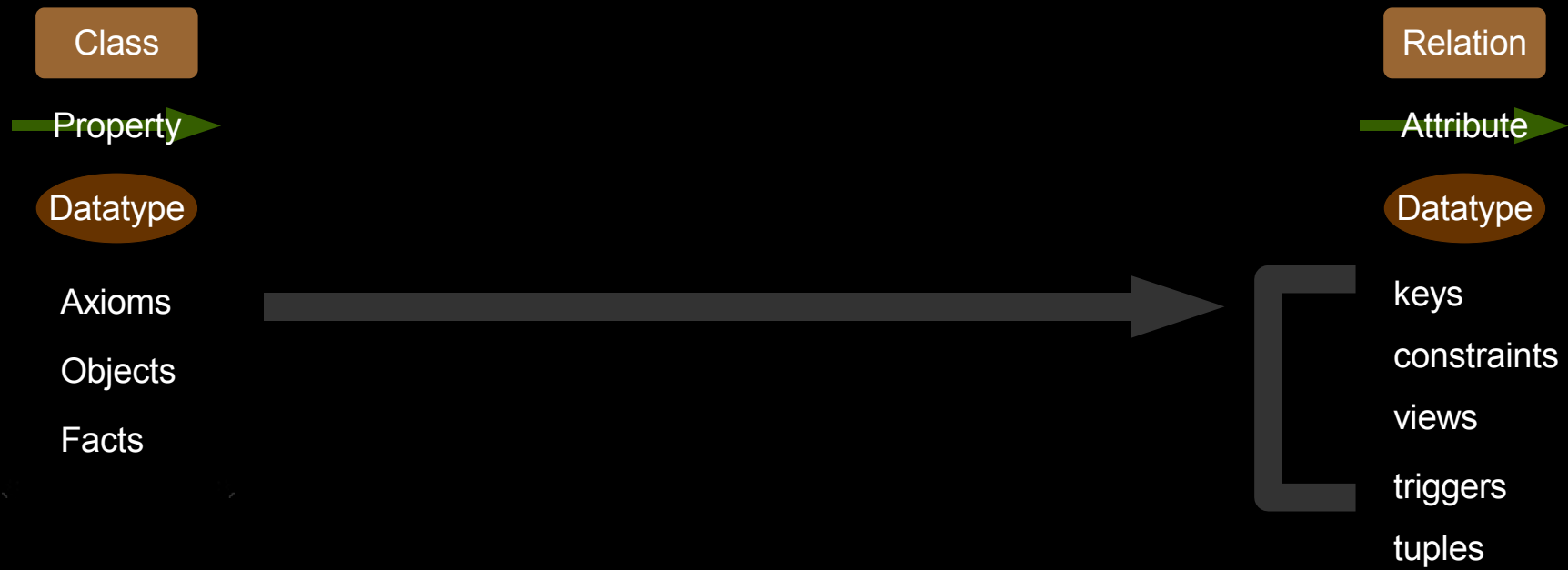
Ontology Databases



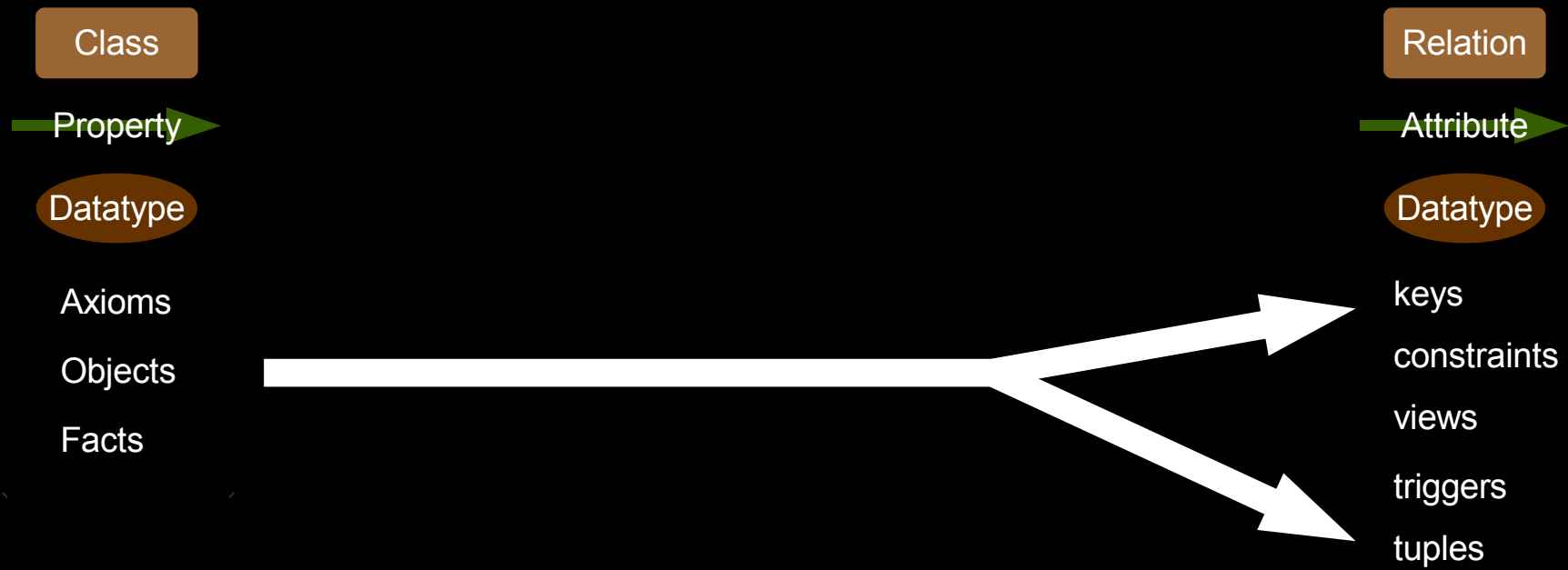
Ontology Databases



Ontology Databases



Ontology Databases



Ontology Databases

