

CS3025 Knowledge-Based Systems

# Standard Languages for the Semantic Web

**Jeff Z. Pan**

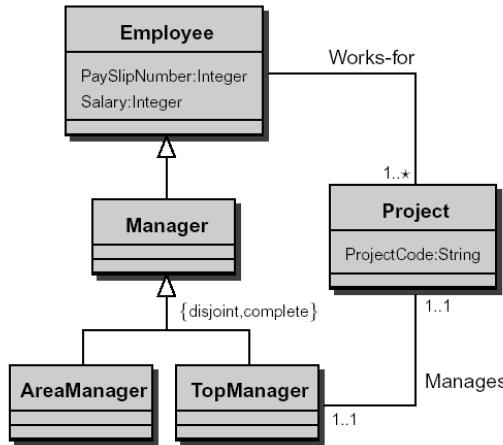
# History of Knowledge Representation



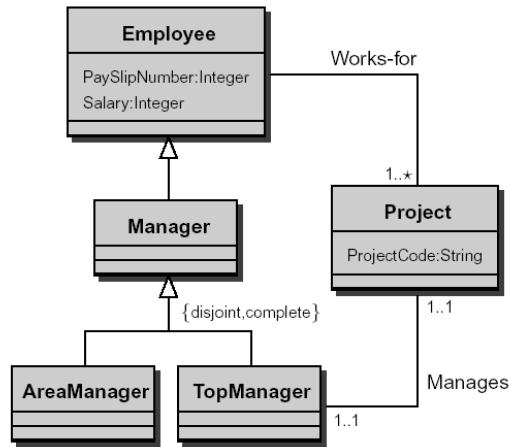
1. Natural language: long history
2. Semantic network
  - highlight **structured** information
  - no well-defined syntax and formal semantics
3. Frame
  - **class (concept), individual (object), property (relation)**
  - **subclass-of, instance-of, property**
  - defaults, procedures
4. First order logic
  - **formal**, (too) expressive
5. **KL-ONE/Description Logics/Ontology Language**
  - subset of FOL with good **implementability**
6. **Rules**
  - **intuitive**, with reasonable implementability

# Set Theory

- A set is an unordered collection of objects
  - Basic relation: member-of
  - Other relations: subset, superset, equivalent
  - Basic operations: union, intersection, difference, complements
- An **individual** is a name for an object
- A **class** “is” a set of objects
  - Employee: {E1, E2, E3, E4}
  - Project: {P1, P2}
- A **property** “is” a set to pairs (tuples) of objects
  - Works-for: {<E1,P1>, <E2,P1>, <E2,P2>, <E3,P1>, <E3,P2>, <E4,P2>}



# Set Theory and Interpretations



- Interpretations relate symbols with sets
- Interpretation is often written as  $(\Delta^I, \bullet^I)$ 
  - $\Delta^I$  : interpretation domain
    - similar to universal set
    - e.g.,  $\Delta^I = \{E1, E2, E3, E4, P1, P2\}$
  - $\bullet^I$  : interpretation function
- Interpretations of individuals, classes and properties
  - all individuals are members of the domain:  $o^I \in \Delta^I$
  - all classes are subsets of the domain  $A^I \subseteq \Delta^I$ 
    - e.g.,  $Employee^I = \{E1, E2, E3, E4\}$
  - all properties are subsets  $R^I \subseteq \Delta^I \times \Delta^I$ 
    - e.g.,  $Works\text{-}for^I = \{\langle E1, P1 \rangle, \langle E2, P1 \rangle, \langle E2, P2 \rangle, \langle E3, P1 \rangle, \langle E3, P2 \rangle, \langle E4, P2 \rangle\}$

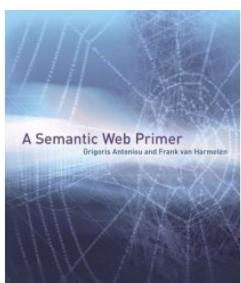
# Roadmap



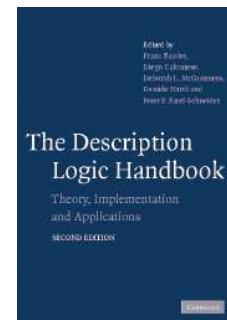
- Foundation
  - KR, ontology and rule; set theory
- Knowledge capture
- Knowledge representation
  - [Ontology: Semantic Web standards RDF and OWL, Description Logics](#)
  - [Rule: Jess](#)
- Knowledge reasoning
  - Ontology: formal semantics, tableaux algorithm
  - Rule: forward chaining, backward chaining
- Knowledge reuse and evaluation
- Meeting the real world
  - Jess and Java, Uncertainty, Invited talk

# Lecture Outline

- Motivation
- Introduction to RDF and OWL
- Some More Details of RDF and OWL
- Practical



[Section 3.1 – 3.5, Section 4.1 - 4.3]



[Chapter 14]

# Motivations



- SW requires meaning of web resources understandable by programs
- Other Web standards do not provide means of talking about semantics of data
  - HTML mixes content with presentation
  - XML only provides the standard syntax to structure data

# What is the Problem?

- Consider a typical web page:

The screenshot shows the homepage of the WWW 2002 conference. At the top, it features the URL <http://www2002.org> and the title "WWW 2002". Below this is the subtitle "THE ELEVENTH INTERNATIONAL WORLD WIDE WEB CONFERENCE". To the left is a logo for "W3C 2002 HAWAII" featuring a stylized figure. To the right is the "CONFERENCE ORGANIZERS" logo for the International World Wide Web Conference Committee. The main content area includes a sidebar with links like "Conference Proceedings", "Call for Participation Program", "Registration Information", "Hotel Accommodation", "Conference Committee", "Sponsorship/Exhibition Opportunities", "Volunteer Information", "Information about Hawaii", and "Previous & Future WWW Conferences". The main text discusses the conference's location at the Sheraton Waikiki Hotel in Honolulu, Hawaii, USA, from May 7-11, 2002. It highlights "1 LOCATION. 5 DAYS. LEARN. INTERACT." and lists "Registered participants coming from:" countries including Australia, Canada, Chile, Denmark, France, Germany, Ghana, Hong Kong, India, Italy, Ireland, Japan, Malta, New Zealand, The Netherlands, Norway, Singapore, Switzerland, The United States, Vietnam, and Zambia. A "REGISTER NOW" button is present. Below this, a section on "FEATURED SPEAKERS (CONFIRMED)" lists Tim Berners-Lee, Ian Foster, and Richard A. DeMillo, each with a small portrait and a brief bio.

- HTML Markup consists of:
  - rendering information (e.g., font size and colour)
  - Hyper-links to related content
- Content is accessible to humans but not (easily) to computers...

# What Information Can We See...

WWW2002

The eleventh international world wide web conference

Sheraton waikiki hotel

Honolulu, hawaii, USA

7-11 may 2002

1 location 5 days learn interact

Registered participants coming from

australia, canada, chile denmark, france, germany, ghana, hong kong, india, ireland, italy,  
japan, malta, new zealand, the netherlands, norway, singapore, switzerland, the united  
kingdom, the united states, vietnam, zaire

Register now

On the 7<sup>th</sup> May Honolulu will provide the backdrop of the eleventh international world wide web conference. This prestigious event ...

Speakers confirmed

Tim berners-lee

Tim is the well known inventor of the Web, ...

Ian Foster

Ian is the pioneer of the Grid, the next generation internet ...



# What Information Can a Program See...

# Solution: XML Markup with “Meaningful” Tags?

# But What About...

<conf> 

❀❀❀ m o m ♦ m ■ ♦ w ❀■♦m□■♦m■♦m●♦□□●o♦♦m♦m♦m□■</conf>

<place>♦❀□♦♦□♦ ♦❀&❀&❀ ❀□♦❀•

¶□■□●◆●◆▣ முக்குள்ளங்கள் +  
+  
</place>

<date> 05</date>

<slugan> ● പ്രഥമം മാത്രം ഒരു സ്ക്രിപ്റ്റ് ദിനം വരെ ഉപയോഗിക്കണം  
</slugan>

<participants> የዚህ እ♦♦ወጪ በቃርቃ♦እውነዎች በቃርቃ♦♦ ገዢዎች እውነዎች

<introduction> ☀ m. n. ☀ X ☀ ♦ m. ☀ □ ☀

አዲስ አበባ የኢትዮጵያ ማኅበር ቤት

<speaker> \*

<bio>❀○ ❁♦ ♦❀❀ ♦❀●● &■□♦■ ❁■❀❀■♦□□ 口❀ ♦❀❀ ❀❀❀

# Programs See...

<■○Ⓜ>◆◆◆◆◆

※※※※ **※**●**※** ♦**※** □**※** ☀♦**※**□♦**※**♦**※**□♦**※**♦**※**● **♦**□**※**●**※** **♦**□**※**●**※** **♦**□**※**●**※** **♦**□**※**●**※**

<●□❖❖◆❖□>滴❖□❖◆❖□◆❖❖&❖❖&❖滴❖□◆❖●

¶□■□●♦●♦■　❀❀♦❀❀❀　+滴</●□❀❀♦❀□>

<ଓଡ଼ିଆ>କେବଳ ଫିଲେ ଏବଂ ଫିଲ୍‌ମ୍‌ର ପାଇଁ ଓଡ଼ିଆ ଲାଙ୍ଘନିକାରୀ

◆□▨▨&▨□◆ ▨□■▨▨□○▨▨▨

<♦□ⓂⒶ&;Ⓜ□>✿Ⓜ○ ⓁⓂ□■Ⓜ□♦Ⓜ●Ⓜ.Ⓜ.</♦□ⓂⒶ&;Ⓜ□>

<ঠ খ ক প> \*খো খু দেশ মু </ঠ খ ক প>

<◆□Ⓜ.&Ⓜ.□>☞◆■ ☞◆□◆Ⓜ.□</◆□Ⓜ.&Ⓜ.□>

# Lecture Outline

- Motivation
- Introduction to RDF and OWL
  - The big picture
- Some More Details of RDF and OWL
- Practical

# Basic Ideas of RDF

- Basic building block: Subject-property-value triple
  - It is called a statement
  - E.g. Jeff teaches CS3025 is a statement
- RDF (Resource Description Framework) has been given a syntax in XML
  - This syntax inherits the benefits of XML
  - Other syntactic representations of RDF possible
    - such as Notation 3 (N3) syntax: [Subject property value .]

# Basic Ideas of RDF (2)

- The fundamental concepts of RDF are:
  - resources
  - properties
  - statements

# Resources

- We can think of a resource as an object, a “thing” we want to talk about
  - E.g. authors, books, publishers, places, people, hotels
- Every resource has a **URI**, a Universal Resource Identifier
- A URI can be
  - a URL (Web address) or
  - some other kind of unique identifier



The screenshot shows a Google search results page for the query "Obama". The search bar at the top contains "Obama". Below it, there are tabs for "Web", "Images", "News", "Videos", "Books", "More", and "Search tools". The "Web" tab is selected. The results section starts with a snippet for "Barack Obama - Wikipedia, the free encyclopedia", which includes a link to [www.barackobama.com/](http://www.barackobama.com/). It also features a quote from President Obama: "Sign the petition to say it's time to give America a raise. Sign the petition. 'You and I, as citizens, have the power to set this country's course.' —President Obama." To the right of this snippet is a grid of four small images of Barack Obama. Below this is a larger image of Barack Obama smiling, with a caption "Barack Obama" and a link to his Google+ profile. The profile shows 4,269,608 followers. The bio states: "Barack Hussein Obama II is the 44th and current President of the United States, and the first African American to hold the office. Born in Honolulu, Hawaii...". It also lists his parents (Ann Dunham, Barack Obama, Sr.), siblings (Malia, Sasha), and education (Harvard Law School). A "Recent posts" section shows a tweet about the European Reassurance Initiative and another about Putin's visit to Poland. A "People also search for" section includes links to Michelle Obama, Vladimir Putin, George W. Bush, and Mitt Romney.

# Properties

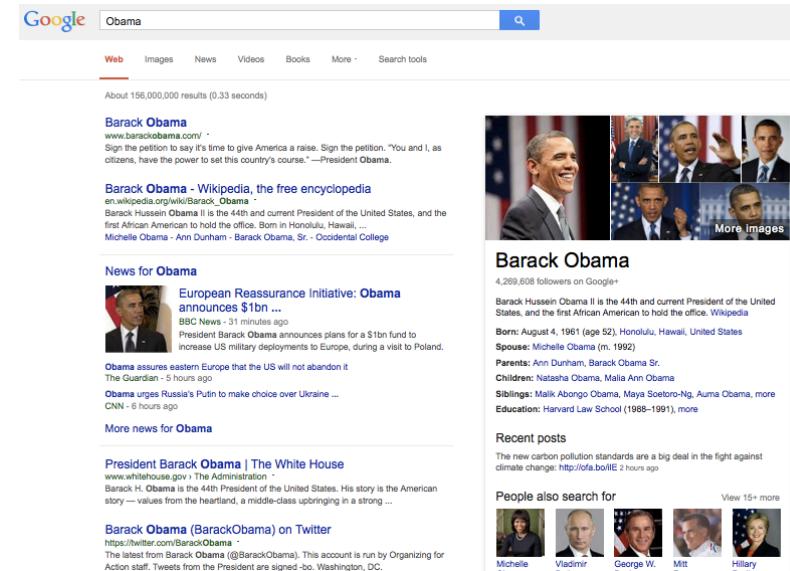
- Properties are a special kind of resources
- They describe relations between resources
  - E.g. “written by”, “age”, “title”, etc.
- Properties are also identified by URIs
- Advantages of using URIs:
  - A global, worldwide, unique naming scheme
  - Reduces the homonym problem of distributed data representation



The screenshot shows a Google search results page for the query "Barack Obama". The top result is a link to the official White House website, which includes a quote from Obama: "Sign the petition to say it's time to give America a new...". Below this, there is a summary box for "Barack Obama - Wikipedia, the free encyclopedia", followed by a list of news articles and other links. To the right of the search results, there is a sidebar with a large portrait of Barack Obama, several smaller thumbnail images of him, and a section titled "Barack Obama" with a brief biography and links to his social media profiles.

# Statements

- Statements assert the properties of resources
- A statement is an resource-property-value triple
  - It consists of a resource, a property, and a value
- Values can be resources or **literals**
  - Type literals: “15”^^xsd:integer
  - Plain literals (strings): “Tom”



The screenshot shows a Google search results page for the query "Barack Obama". The top result is a link to the official White House website. Below it are news snippets from BBC News and The Guardian. To the right, there's a sidebar with a profile picture of Barack Obama, his biography, and a "Recent posts" section featuring a tweet about carbon pollution.

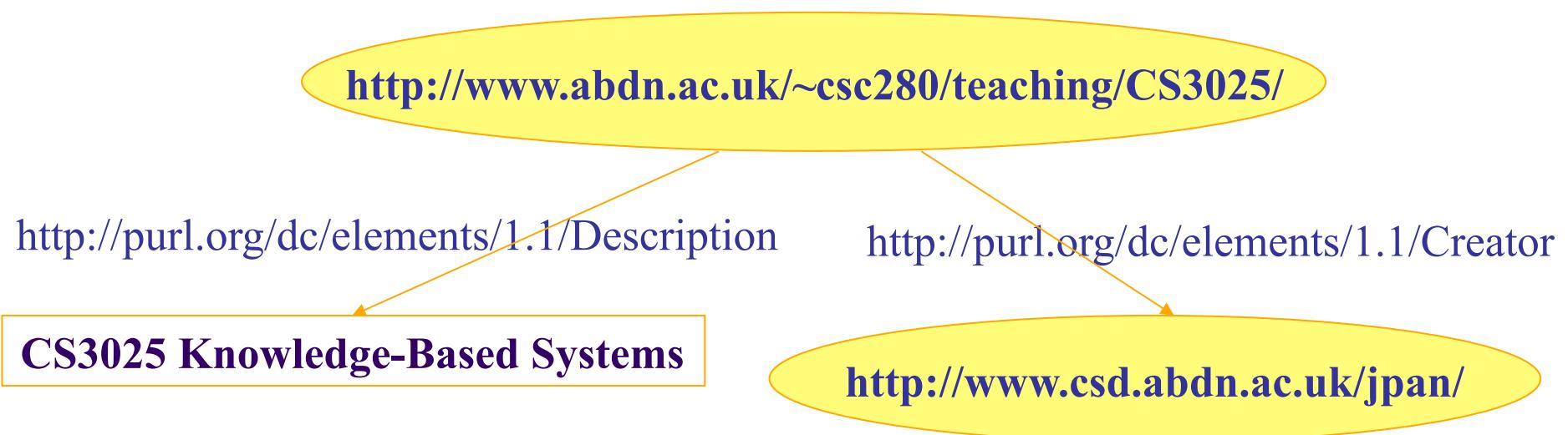
# How to Enable Machine/Program Understandability



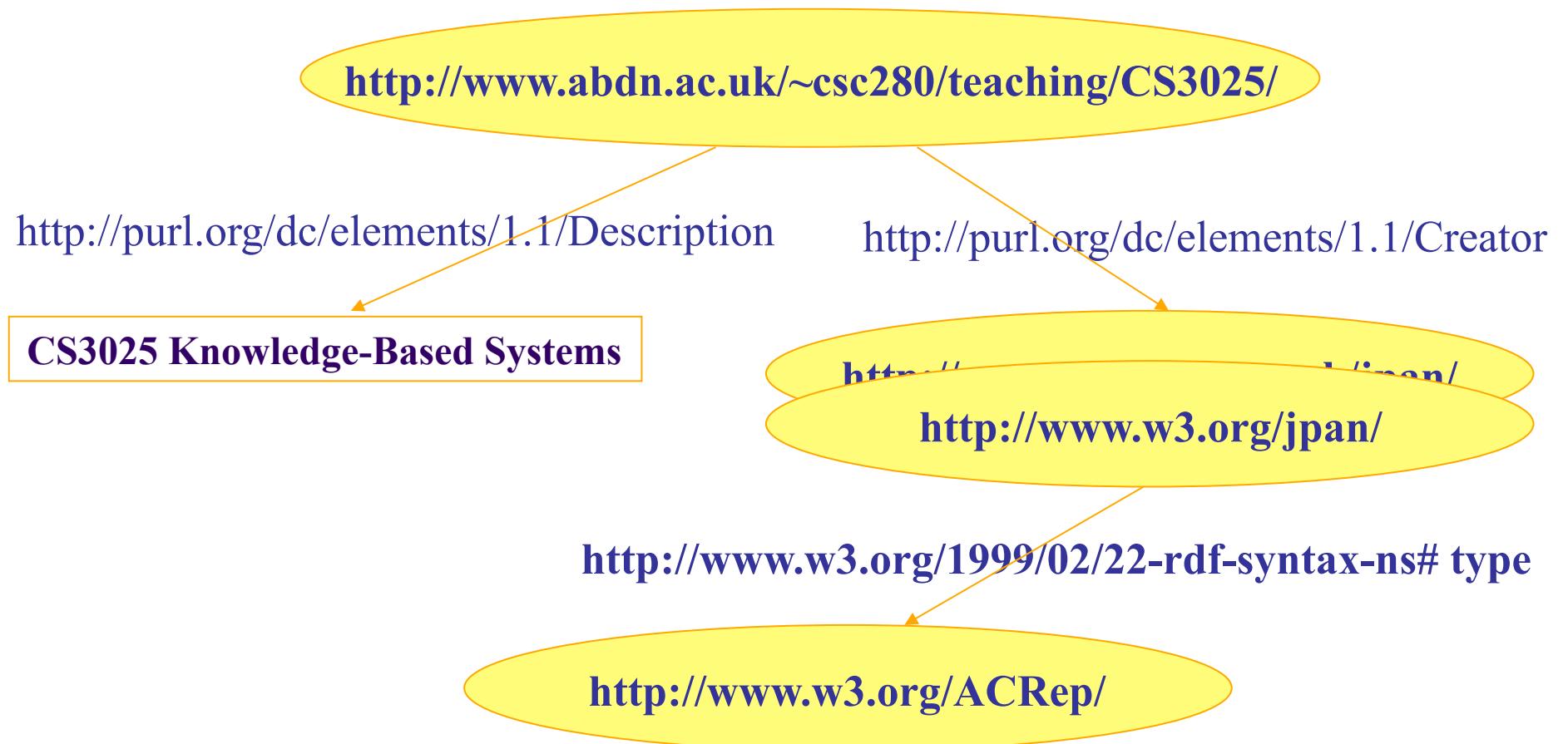
- To add annotations to Web resources
  - Describing their contents and functionalities
- To provide semantics for annotations

# What do Annotations Look Like?

- Annotations are represented as RDF triples



# A Network of Annotations



# Example: RDF/XML Syntax



```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:csd="http://www.csdl.ac.uk/"
  >

<rdf:Description rdf:about="http://www.abdn.ac.uk/~csc280/
  teaching/CS3025/">
  <dc:Creator rdf:resource="http://www.csdl.ac.uk/jpan/ ">
    <dc:Description> CS3025 Knowledge-Based Systems
  </dc:Description>
</rdf:Description>

</rdf:RDF>
```

# Example: Notation 3 (N3) Syntax



```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns>
@prefix dc: <http://purl.org/dc/elements/1.1>
@prefix csd: <http://www.csd.abdn.ac.uk/>
```

```
[http://www.abdn.ac.uk/~csc280/teaching/CS3025/
dc:Creator csd:jpan .]
```

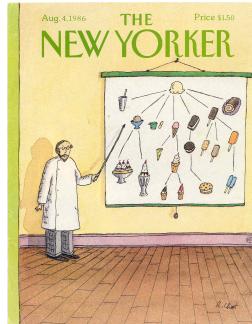
```
[http://www.abdn.ac.uk/~csc280/teaching/CS3019/
dc:Description “CS3025 Knowledge-Based Systems” .]
```

# What do Annotations Mean?



- So far we have used the following in our annotations
  - <http://purl.org/dc/elements/1.1/Description>
  - <http://purl.org/dc/elements/1.1/Creator>
  - <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
  - <http://www.w3.org/ACRep/>
- What do they mean?

# Ontology



- An ontology is a model of (some aspects of) the world:
  - It is constituted by a key vocabulary for the domain, plus
  - definition and constraints about the vocabulary, to capture the intended meaning of the vocabulary

# Requirements for Ontology Languages



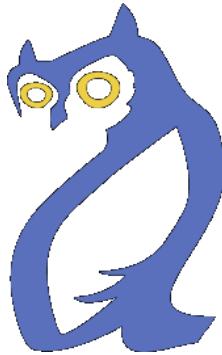
- Ontology languages allow users to write explicit, formal conceptualisations of domain models
- The main requirements are:
  - a well-defined syntax
  - a formal semantics
  - efficient reasoning support
  - sufficient expressive power
  - convenience of expression

# RDF Schema



- RDF Schema (RDFS)
  - Allow users to introduce vocabulary
- Key meta classes for declarations:
  - `rdfs:Class`, `rdfs:Resource`, `rdfs:Property`
- Key meta properties for basic axioms:
  - `rdf:type`, `rdfs:subClassOf`, `rdfs:subPropertyOf`,  
`rdfs:domain`, `rdfs:range`
- Examples:  
[csd:Lecturer `rdfs:subClassOf` csd:AssistantProf .]  
[csd:Lecturer `rdfs:subClassOf` csd:Staff .]  
[csd:jpan `rdf:type` csd:Lecturer .]  
[csd:AssistantProf `rdfs:subClassOf` csd:Lecturer .]

# OWL: Web Ontology Language



- OWL is the W3C Web Ontology Language
  - Based on Description Logics
  - Has several syntax
    - RDF/XML
    - Abstract
    - DL
- Motivation: more expressive power, such as
  - Disjointness of classes
    - [DisjointClasses \(csd:Male csd:Female\)](#)
  - Equivalent individuals
    - [SameIndividual \(csd:jpan w3:jpan\)](#)
  - Boolean combination of classes
    - [Class \(Person complete unionOf \(csd:Male csd:Female\)\)](#)
  - and many more ...

# Lecture Outline

- Motivation
- Introduction to RDF and OWL
- Some More Details of RDF and OWL
- Practical

# Status of RDF and OWL

- Both are W3C recommendations
- Resource Description Framework (RDF)
  - <http://www.w3.org/RDF/>
  - **Primer:** <http://www.w3.org/TR/REC-rdf-syntax/>
- Web Ontology Language (OWL) 2.0
  - <http://www.w3.org/TR/owl2-overview/>

# How many facts?

Schema-like

```
(deftemplate person
  (slot name)
  (slot age)
  (slot gender)
  (slot partner)
  (multislot children))
```

Data

```
(assert (person
           (name Michael)
           (gender male)
           (age 35)
           (partner Diana)
           (children Philip
                Julia)))
<fact-0>
```

# Class Hierarchies



- Classes can be organised in hierarchies
  - A is a sub-class of B if every instance of A is also an instance of B
  - Then B is a super-class of A
- Sub-class-of relations are transitive
- Example

[csd:Lecturer rdfs:subClassOf csd:Staff .]  
[csd:Staff rdfs:subClassOf csd:Person .]

[csd:jpan rdfs:type csd:Lecturer .]

# Property Hierarchies



- Properties can be organised in hierarchies
  - P is a sub-property of Q if  $Q(x,y)$  holds whenever  $P(x,y)$  holds
  - Then Q is a super-property of P
- Example

[**hasGoodFriend rdfs:subPropertyOf hasFriend .**]

[John hasGoodFriend Mary .]

# Domains of Properties



- A domain of a property P is
  - the class of those resources that
  - may appear as subjects
  - in a triple with property P
- If the range is not specified, then any resource can be the value

[csd:teach rdfs:domain csd:Lecturer .]

[csd:jpan teach csd:CS3025 .]

# Ranges of Properties



- A range of a property P is
  - the class of those resources that
  - may appear as values
  - in a triple with property P
- If the domain is not specified, then any resource can be the subject

[csd:teach rdfs:range csd:Course .]

[csd:jpan teach csd:CS3025 .]

# Practical



- How to construct OWL ontologies with ontology editors
- Go through tutorials of an ontology editor called **Protégé**
  - sub-class, super-class
  - disjoint classes
  - **Important:** report to the demonstrator about the your comments of the pizzas ontologies and any changes you want to make against them

# Summary



- RDF
  - Data model for the Semantic Web
    - and machine understandable data
  - Basic data language and schema language
- OWL
  - Web Ontology Language
  - Advanced data language and schema language
  - More expressive power but still decidable