Introduction to AI

Francesca Toni

Semantic Web

- Russell & Norvig (sections 12.1, 12.2, 12.5)
- Poole & Mackworth (chapter 13)
- A Semantic Web Primer (2nd edition, 2008)
  Grigoris Antoniou, Frank van Harmelen

Outline

1. Today’s Web
2. The Semantic Web Impact
3. Semantic Web Technologies
4. A Layered Approach
5. Description Logic Programs

Today's Web

- Most of today's Web content is suitable for human consumption
  - seeking and making use of information,
  - searching for and getting in touch with other people,
  - reviewing catalogues of online stores and ordering products by filling out forms
- Current Web activities are not particularly well supported by software tools
  - except for keyword-based search engines (e.g. Google, AltaVista, Yahoo), but several problems (high recall/no precision, low or no recall, etc)

The Key Problem of Today's Web

- The meaning of Web content is not machine-accessible: lack of semantics
- It is simply difficult to distinguish the meaning between these two sentences:

  I am a professor of computer science.

  I am a professor of computer science, you may think. Well, . . .

The Semantic Web Approach

- Represent Web content in a form that is more easily machine-processable.
- Use intelligent techniques to take advantage of these representations.

- See www.w3.org/standards/semanticweb/

On HTML

- Web content is currently formatted for human readers rather than programs
- HTML is the predominant language in which Web pages are written (directly or using tools)
- Vocabulary describes presentation

HTML Example

<h1>Agilitas Physiotherapy Centre</h1>
Welcome to the home page of the Agilitas Physiotherapy Centre. Do you feel pain? Have you had an injury? Let our staff Lisa Davenport, Kelly Townsend (our lovely secretary) and Steve Matthews take care of your body and soul.

<h2>Consultation hours</h2>
Mon 11am - 7pm
Tue 11am - 7pm
Wed 3pm - 7pm
Thu 11am - 7pm
Fri 11am - 3pm<p>
But note that we do not offer consultation during the weeks of the State Of Origin games.

Problems with HTML

- Humans have no problem with this
- Machines (software agents) do:
  - How to distinguish therapists from the secretary,
  - How to determine exact consultation hours
  - They would have to follow the link to the State Of Origin games to find when they take place.

A Better Representation - XML

```
<company>
    <treatmentOffered>
        Physiotherapy
  </treatmentOffered>
    <companyName>
        Agilitas Physiotherapy Centre
  </companyName>
    <staff>
            <therapist>Lisa Davenport</therapist>
            <therapist>Steve Matthews</therapist>
            <secretary>Kelly Townsend</secretary>
    </staff>
</company>
```

Explicit Metadata

- This representation is far more easily processable by machines
- Metadata: data about data
  - to capture part of the *meaning of data*
- Semantic Web does not rely on text-based manipulation, but rather on machine-processable metadata

Drawbacks of XML

- XML provides a uniform framework for interchange of data and metadata between applications
- However, XML does not provide any means of talking about the semantics (meaning) of data
- E.g., there is no intended meaning associated with the nesting of tags:

```
<course name="Discrete Maths">
		<lecturer>David </lecturer>
</course>
<lecturer name="David ">
		<teaches>Discrete Maths</teaches>
</lecturer>
```

Opposite nesting, same information!

Ontologies

The term ontology originates from philosophy

- The study of the nature of existence

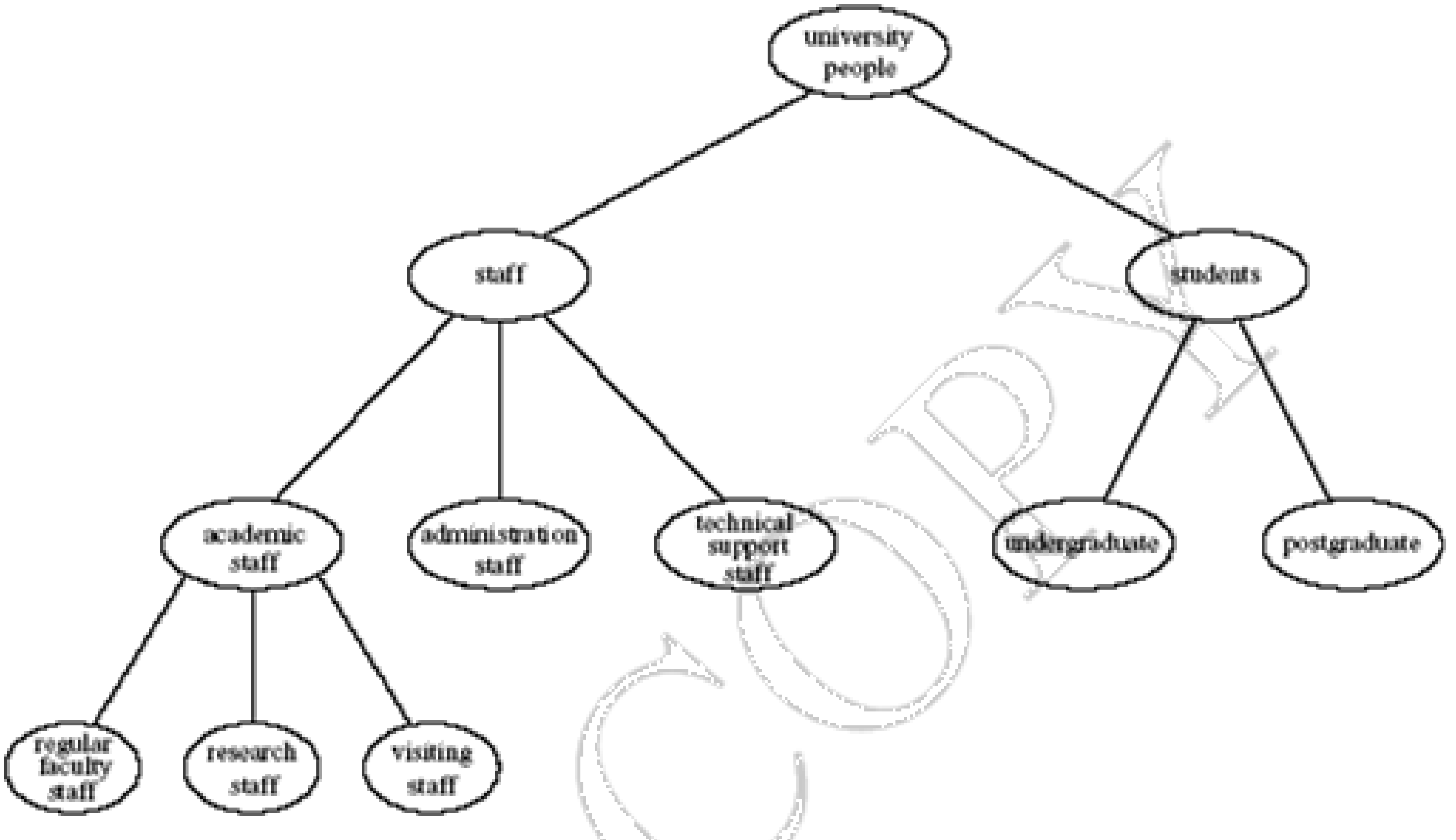
Different meaning from computer science

- An ontology is an explicit and formal specification of a conceptualization
- “the kind of things that exist in a given domain”

Typical Components of Ontologies

- Terms denote important concepts (classes of objects) of the domain
  - e.g. professors, staff, students, courses, departments
- Relationships between these terms: typically class hierarchies
  - a class C to be a subclass of another class C' if every object in C is also included in C'
  - e.g. all professors are staff members

Example of a Class Hierarchy



Further Components of Ontologies

- Properties:
  - e.g. X teaches Y
- Value restrictions
  - e.g. only academic staff members can teach courses
- Disjointness statements
  - e.g. academic and technical support staff are disjoint
- Logical relationships between objects
  - e.g. every department must include at least 10 academics

The Role of Ontologies on the Web

- Ontologies provide a shared understanding of a domain: semantic interoperability
  - overcome differences in terminology
  - mappings between ontologies
- Ontologies are useful for the organization and navigation of Web sites

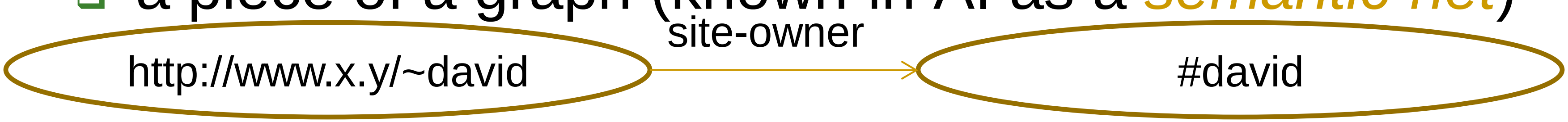
Web Ontology Languages (1)

RDF Schema

- RDF is a *data model* for objects and relations between them
- RDF Schema is a *vocabulary description language*
- Describes properties and classes of RDF resources
- Provides semantics for generalization hierarchies of properties and classes

RDF

- Statements assert the properties of resources
- A statement is an object-attribute-value triple
  - It consists of a resource, a property, and a value
  - E.g. *(http://www.x.y/~david,*
    http://www.k.z/site-owner, #david)
- Can be also seen as
  - a piece of a graph (known in AI as a *semantic net*)



  - a piece of XML code
  - an atom: *site-owner(http://www.x.y/~david, #david)*

From RDF to RDF Schema

- RDF is a universal language that lets users describe resources in their own vocabularies
  - RDF does not assume, nor does it define semantics of any particular application domain
- The user can do so in RDF Schema using:
  - Classes and Properties
  - Class Hierarchies and Inheritance
  - Property Hierarchies

RDF Schema - relationships

- The relationship between instances and classes is through **type**
  - **type(a,C)** states that a is instance of class C
- Classes can be organised in hierarchies through **subClassOf**
  - **subClassOf(C,D)** states that class C is a subclass of class D
- Properties can be organised in hierarchies through **subPropertyOf**
  - **subPropertyOf(P,Q)** states that property Q is true whenever property P is true

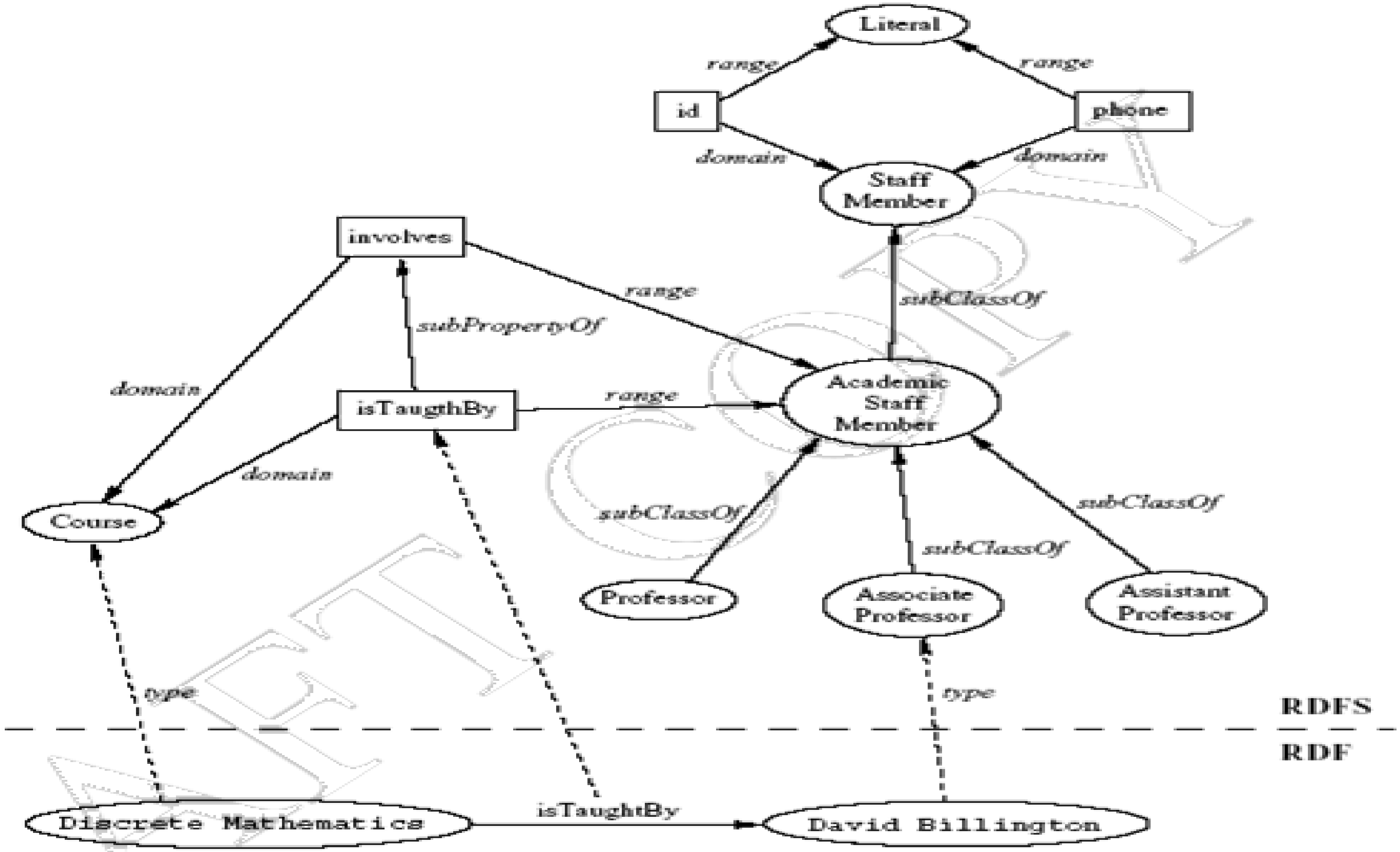
RDF Schema – examples of relationships

- **type(mary,researchStaff)**
  - Mary is a member of the class “research staff”
- **subClassOf(researchStaff,academicStaff)**
  - All research staff members are also academic staff members
- **subPropertyOf(isTaughtBy,involves)**
  - If a course C is taught by an academic staff member A, then C also involves A

RDF Schema – range/domain restrictions

- Discrete Maths is taught by Concrete Maths??
  - We want courses to be taught by lecturers only
  - Restriction on values of the property “is taught by”
  - **range (P,C)**: class C is the range of property P
- Room MZH5760 is taught by David??
  - Only courses can be taught
  - This imposes a restriction on the objects to which the property can be applied
  - **domain(P,C)**: class C is the domain of property P

RDF "Layer" vs RDF Schema "Layer"



Limitations of RDF Schema

- Disjointness of classes not supported
  - e.g. **male** and **female**
- No Boolean combinations (union, intersection, complement) of classes
  - e.g. **person** is the disjoint union of the classes **male** and **female**
- No Cardinality restrictions
  - e.g. a person has exactly two parents, a course is taught by at least one lecturer
- Special characteristics of properties not supported
  - Transitive property (like “greater than”)

Web Ontology Languages (2)

OWL

- A richer ontology language
  - relations between classes
    - e.g., disjointness
  - cardinality
    - e.g. “exactly one”
  - richer typing of properties
  - characteristics of properties (e.g. transitivity)

Three Species of OWL

- W3C'sWeb Ontology Working Group defined OWL as three different sublanguages:
  - OWL Full (undecidable)
  - OWL DL (description logic, efficient reasoners)
  - OWL Lite (sub-language of OWL DL, easier to grasp/implement)
- Each sublanguage geared toward fulfilling different aspects of requirements

Outline revisited

1. Today's Web
2. The Semantic Web Impact
3. Semantic Web Technologies
4. A Layered Approach
5. Description Logic Programs

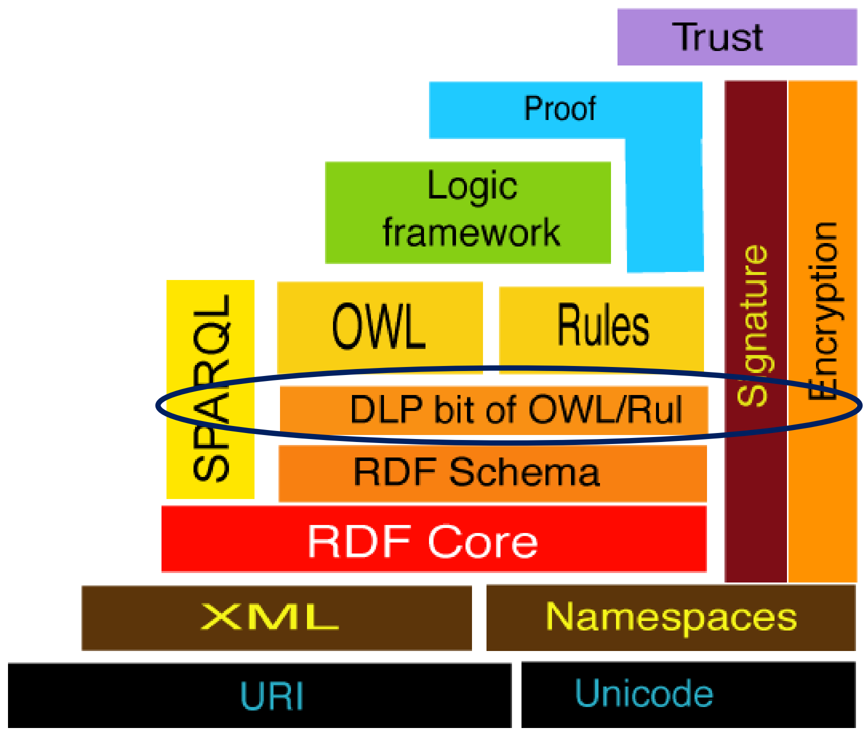
A Layered Approach

- The development of the Semantic Web proceeds in steps
  - Each step building a layer on top of another

Principles:

- Downward compatibility
- Upward partial understanding

The Semantic Web Layer Tower



Description Logic Programs

- Description Logic Programs (DLP) can be considered as the intersection of definite logic programs and description logic (DL)
- DLP allows to combine advantages of both approaches. For example:
  - A modeler may take a DL view, but
  - the implementation may be based on DLP technology

RDF and RDF Schema in DLP

- A triple of the form **(a,P,b)** in RDF can be expressed as a fact **P(a,b)**
- An instance declaration of the form **type(a,C)** (stating **a** is instance of class **C**) can be expressed as **C(a)**
- The fact that **A** is a subclass (or subproperty) of **B** can be expressed as **A(X)** $\rightarrow$ **B(X)**
- The fact that **C** is the domain of property **P** can be expressed by **P(X,Y)** $\rightarrow$ **C(X)**

OWL expressed in DLP (1)

- **sameClassAs(C,D)** (or samePropertyAs)
  - **C(X) → D(X)**
  - **D(X) → C(X)**
- Transitivity of a property **P**
  - **P(X,Y),P(Y,Z) → P(X,Z)**

OWL expressed in DLP (2)

- The intersection of $\mathbf{C_1}$ and $\mathbf{C_2}$ is a subclass of **D**
  - $C_1(X), C_2(X) \rightarrow D(X)$
- **C** is subclass of the intersection of $\mathbf{D_1}$ and $\mathbf{D_2}$
  - $C(X) \rightarrow D_1(X)$
  - $C(X) \rightarrow D_2(X)$
- The union of $\mathbf{C_1}$ and $\mathbf{C_2}$ is a subclass of **D**
  - $C_1(X) \rightarrow D(X)$
  - $C_2(X) \rightarrow D(X)$

Restrictions in OWL using DLP (1)

- **C** subClassOf **allValuesFrom(P,D)** can be expressed as
  - **C(X),P(X,Y) → D(Y)**
  - Where **P** is a property, **D** is a class and **allValuesFrom(P,D)** denotes the anonymous class of all **x** such that **y** must be an instance of **D** whenever **P(x,y)**
  - The opposite direction cannot in general be expressed

Restrictions in OWL using DLP (2)

- Cardinality constraints and complement of classes cannot be expressed using definite clauses in the general case

Summary

- Today's Web: a critical analysis
- The Semantic Web Impact: some examples
- Semantic Web Technologies:
  - Metadata (XML)
  - Ontologies (RDF schema, OWL)
- A Layered Approach
  - Semantic Web technologies
- Description Logic Programs
  - Versus RDF/RDF Schema
  - Versus OWL