Ontologies and Rules for the Semantic Web

G. Antoniou

Computer Science Department, University of Crete
Institute for Computer Science - FORTH

Three Generations of the Web

❶ Hand-coded Web content
- huge authoring and maintenance effort
- hard to deal with dynamically changing content

❷ Automated on-the fly content generation
- based on templates filled with database content
- later extended with XML document transformations

❸ Automated processing of content
- The Semantic Web !

Today's Web

- Meant primarily for human consumption
- Keyword-based search critical for its success

Problems with keyword-based retrieval:

- High recall, low precision
- Low or no recall
- Results highly sensitive to vocabulary
- Results are single web pages; integration of results from various pages requires human involvement

The Semantic Web Vision

- Machine-processable Web content
- Meta-data: Web content carries its meaning on its sleeve
- Ontologies: shared understanding of the semantics
- Enables agent communication at a Web-wide scale

Can it work? The Semantic Web versus Artificial Intelligence

- It uses many AI techniques, but
- Partial solutions work!

From Today's Web to the SW: Knowledge Management

Today:

- Keyword-based search
- Human involvement for browsing retrieved information
- Inconsistent, outdated information

Knowledge Management on the Semantic Web:

- Knowledge organized in conceptual spaces
- Automated tools support maintenance (e.g. inconsistencies)
- Query answering based on semantics

From Today's Web to the SW: B2C E-Commerce

Today:

- Manual browsing is typical (but how many?)
- Shopbots can help, but they rely on wrappers (error-prone, re-programming may be necessary)

B2C E-Commerce on the Semantic Web :

- Pricing and product information, delivery, refund and privacy policies will be retrieved and compared to the user requirements
- In the future, sophisticated shopping agents will negotiate with shop agents on the buyer's behalf

From Today's Web to the SW: B2B E-Commerce

Today:

- Relies on exchange formats like EDI
  - Complex, difficult to program, isolated technology
- Business partnerships are static
- B2B portal business model is not well-supported

B2B E-Commerce on the Semantic Web:

- Businesses will enter partnerships without much overhead
- Abstract domain models, translation services between terminologies
- Auctioning, negotiations and drafting contracts

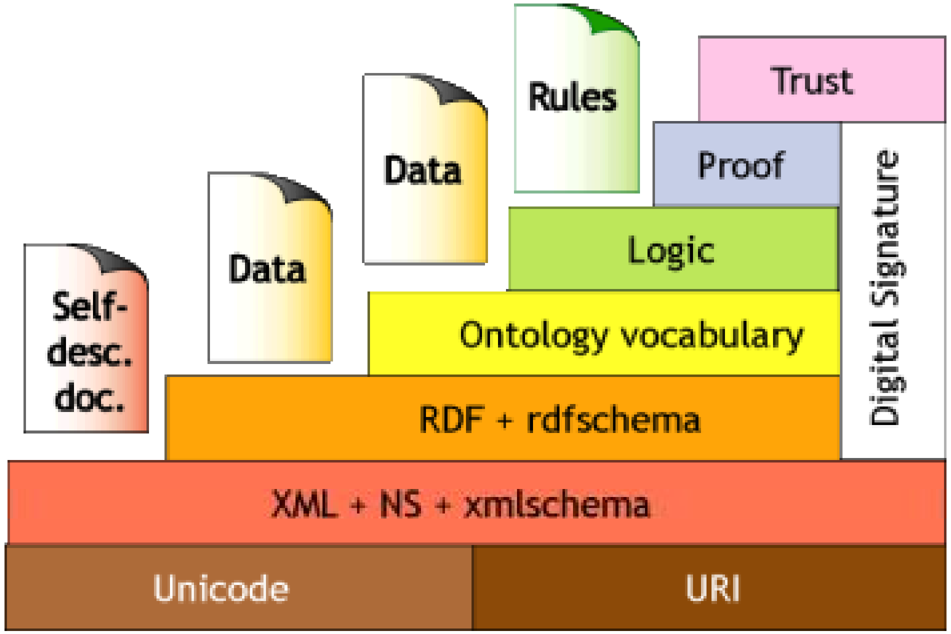
Vision: Agents in Knowledge-Based E-Markets

Coming soon to a world near you:

- Billions/trillions of agents
- Cognitive agents:
  - Knowledge gathering, Reasoning, Economic optimization
- … doing our bidding
  - but with some autonomy

A 1st step: Ability to communicate with sufficiently shared meaning… via the Semantic Web

The Semantic Web Design: The SW Tower



- Downward compatibility
- Upward partial understanding

What is an Ontology?

In philosophy:

- Theory of what exists in the world

In IT:

- Consensual and formal description of shared concepts in a domain
- Aid to human communication and shared understanding
- Machine-processable for agent communication

What is an Ontology? (2)

Typically an ontology contains:

- The objects of a domain (classes)
- Relations between classes
- Including a hierarchical organization

Useful for:

- Improving accuracy of Web searches
- Exploiting generalization/specialization information
- Establishing semantic middleware between applications

What is an Ontology? (3)

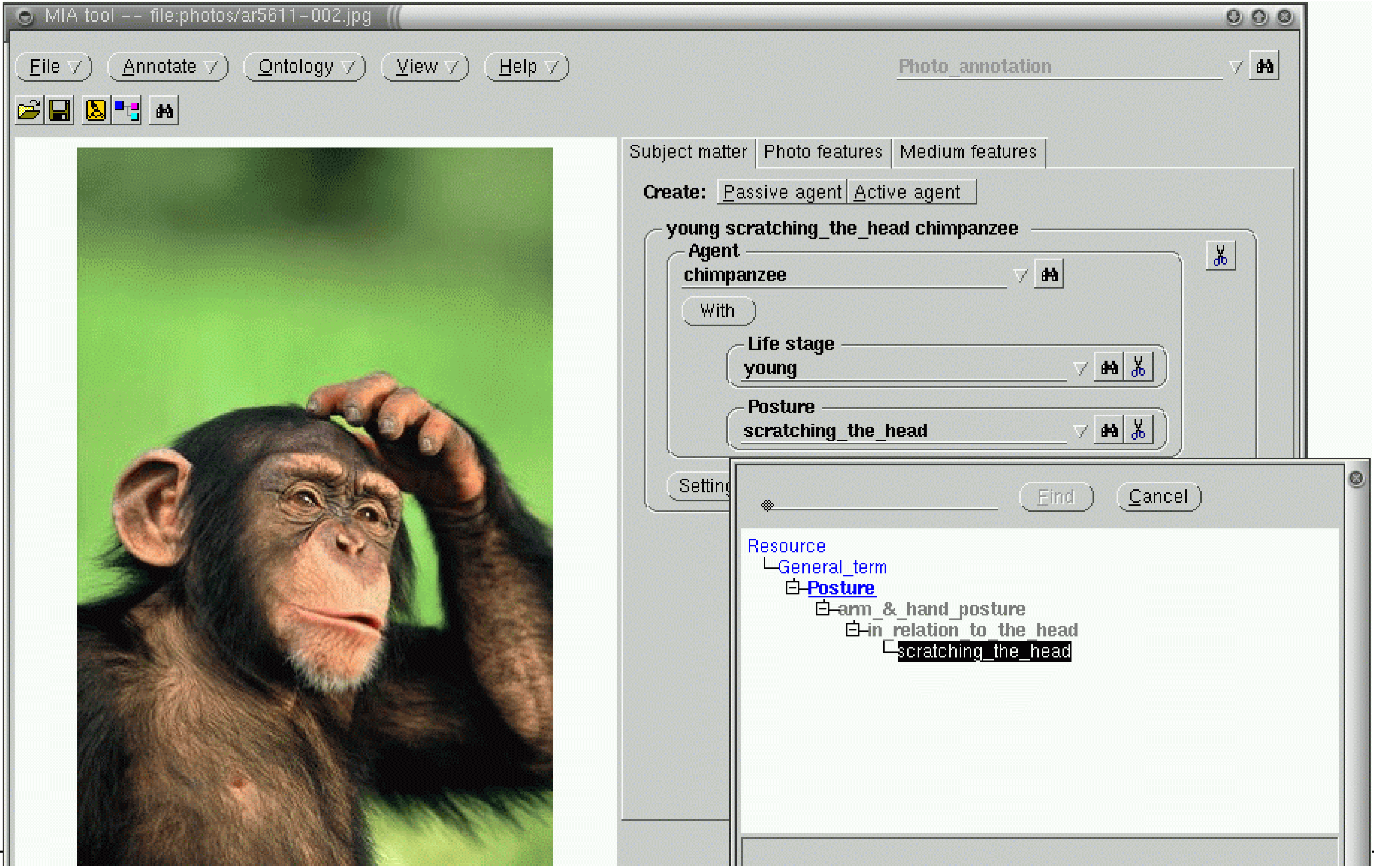
Ontologies are a key technology in semantic information processing

- Applications: knowledge management, e-business, industrial engineering, etc.

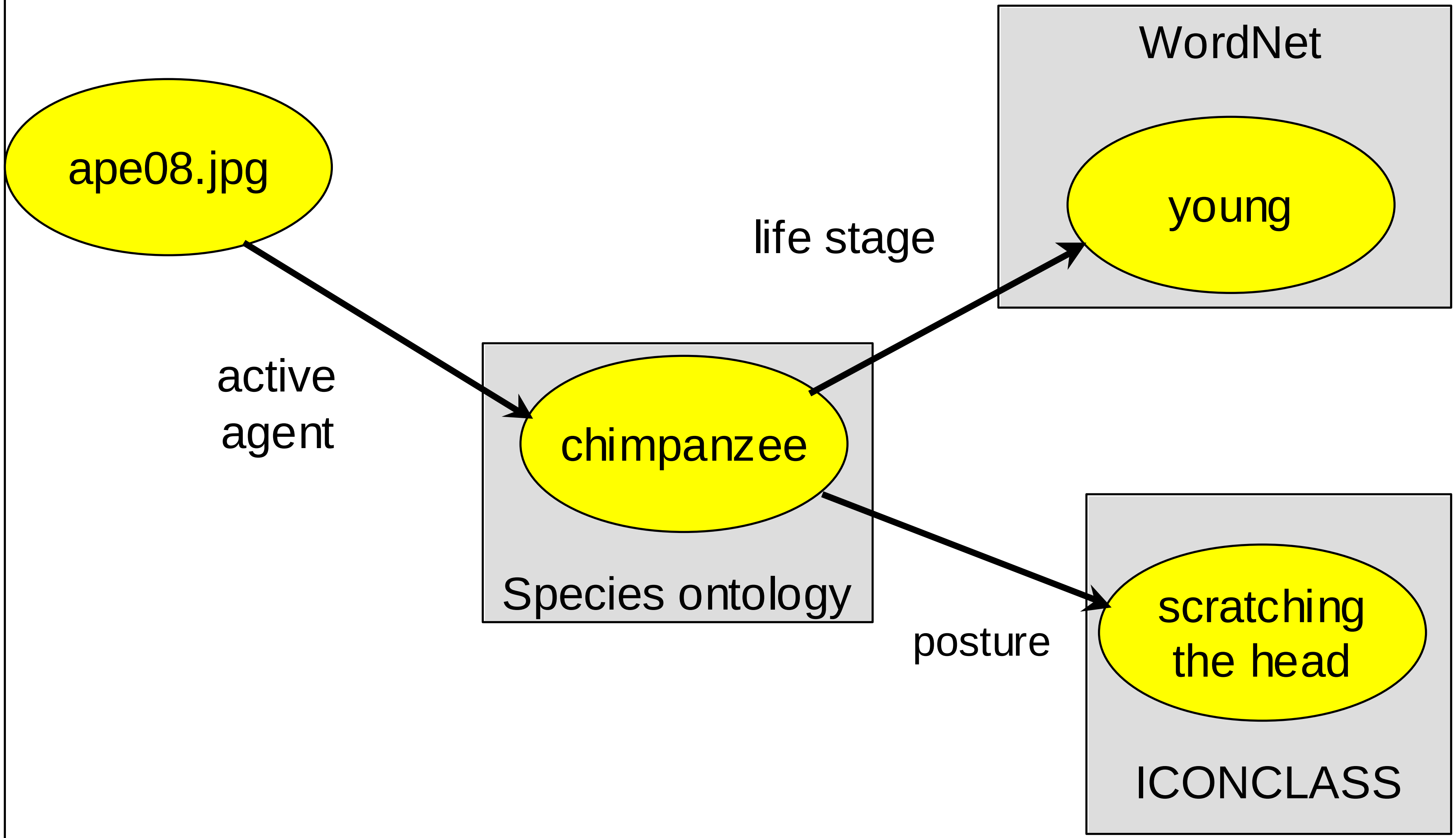
Typical Ontology Use Case: Image Search

- A person searches for photos of an “orange ape”
- An image collection of animal photographs contains snapshots of orang-utans.
- The search engine finds the photos, despite the fact that the words “orange” and “ape” do not appear in annotations

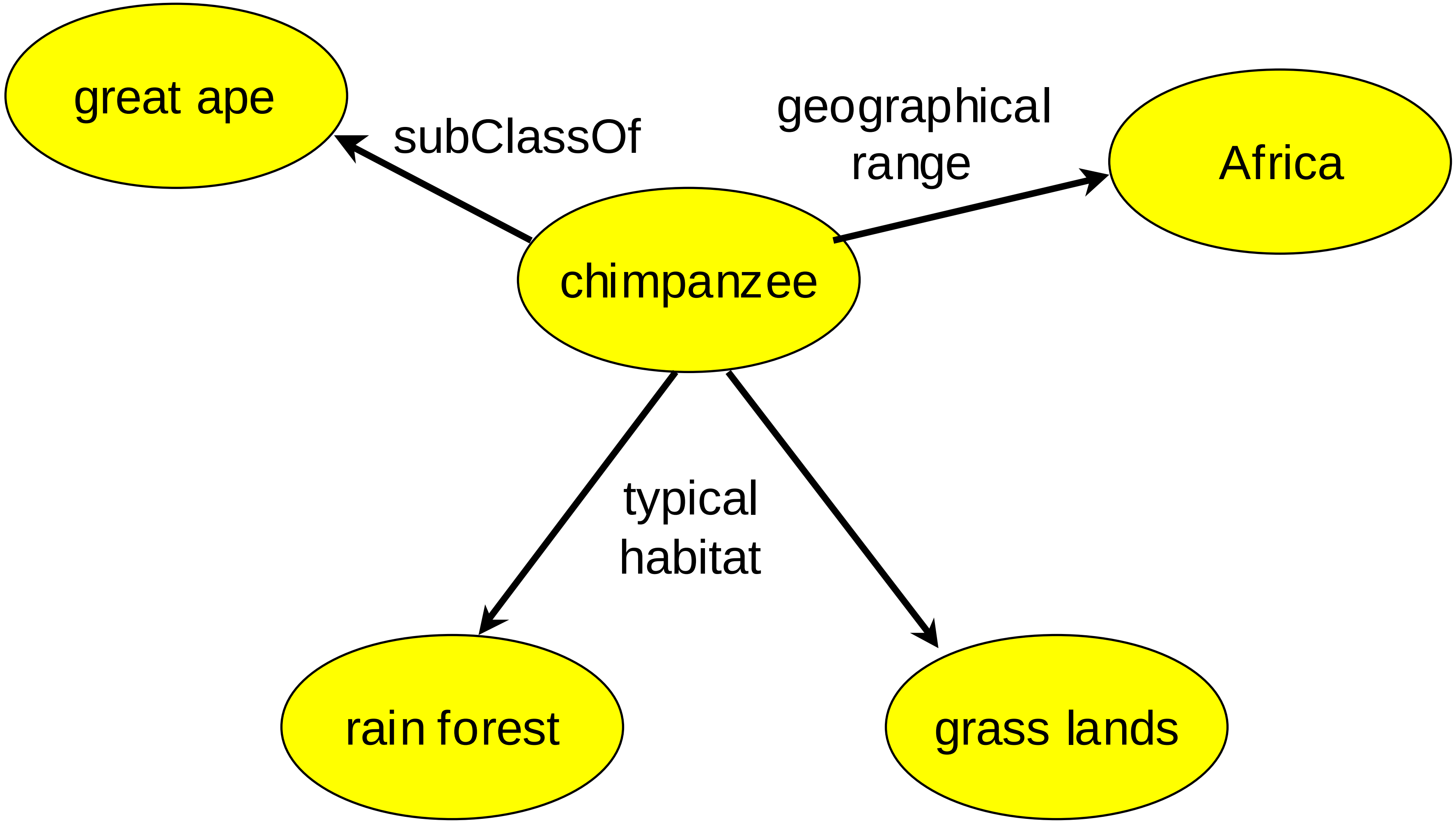
Example Semantic Annotation



RDF Annotation of A Web Resource



Ontologies Describe Concepts Used



Requirements for Ontology Languages

- Well-defined syntax
- Well-defined and intuitively clear semantics
- Efficient reasoning support
- Sufficient expressive power
- Convenience of expression

All are important, but there is tradeoff between:

- Efficient reasoning support
- Sufficient expressive power

Ontologies: The Role of Reasoning

Class membership

- x instance of C, C subclass of D, therefore x instance of D

Equivalence of classes

- A equivalent to be, B equivalent to C, therefore A equivalent to C

Consistency

- Uncovers errors in the ontology and its instantiation

Classification

- P a sufficient condition for C, x satisfies P, therefore x is an instance of C

Ontologies: The Role of Reasoning (2)

Benefits of reasoning:

- Useful for query answering
- As a design support tool
  - For large ontologies
  - With multiple authors
- For integrating and sharing ontologies

Benefits of Reasoning: An Example

Knowledge:

- herbivore ⇔ animal eats (plant or (part_of plant))
- tree ⇒ plant
- branch ⇒ part_of tree
- leaf ⇒ part_of branch
- giraffe ⇒ animal eats leaf
- part_of = transitive

Now we can derive:

- giraffe ⇒ herbivore

Web Ontology Languages

Main ontology languages for the Semantic Web:

- RDF Schema
- DAML+OIL
- OWL

The essence of RDF Schema:

- Classes and binary relations (properties)
- Subclass and subproperty relations
- Domain and range restrictions
- Instance declarations

Missing Features of RDF Schema

Local scope of properties

- Cows eat only plants, while other animals eat meat, too

Disjointness of classes

- Males and females are disjoint classes

Boolean combinations of classes

- Persons are union of males and females

Cardinality constraints

- An offered course is taught by at least one lecturer

Special characteristics of properties

- Transitivity, uniqueness, inverse of another property

Web Ontology Language: OWL

An upcoming W3C standard

- Evolved out of DAML+OIL
- Has RDF/XML syntax
- Datatypes are built using XML Schema
- Instances are built using RDF
- Ideally compatible with RDF Schema (see below)
- Correspond ideally with a description logic (see below)

The Layered Design of OWL

OWL Full

- Compatible with RDF Schema
- Problematic, for example: we can declare cardinality constraint on rdf:Class, thus restricting the number of classes used in any ontology using this code
- Undecidable

The Layered Design of OWL (2)

OWL DL

- Restricts application of constructors to each other
- Corresponds to a description logic
- Somewhat efficient reasoning support
- Not compatible with RDF/S: a legal RDF document may not be a legal OWL DL document

The Layered Design of OWL (3)

OWL Lite

- Further restriction
- E.g. no arbitrary cardinality restrictions, no disjointness
- Advantage: easier to grasp
- Advantage: easier to implement for tool builders
- Disadvantage: restricted expressivity

Rules for the Semantic Web

Next step in the development of the Semantic Web.

Why?

- Provides expressivity not offered by OWL
  - Rules are orthogonal to description logics
- Rule technology is more mature than description logics
- Non-experts are more familiar with rules
- Rules integrate well into the commercial mainstream software engineering, e.g. OO and DB

Rules for the Semantic Web (2)

Where do rules fit in the Semantic Web design?

- Rules within Web ontology languages
- Rules “on top” of ontology languages

Rules on top of ontologies

- Ontological knowledge used in rules (bodies)
- No new ontological knowledge is derived using rules

Rules for the Semantic Web (3)

Rules in Web ontology languages

- Full combination of rules and description logics leads to computability and intractability problems
- Currently: common subset of DL and rules
- An alternative: RDFS + Rules

Rule Standardization for the Semantic Web

RuleML Initiative (2000 -)

- Dozens of institutions, especially in EU and USA
- Mission: Enable semantic exchange of rules and facts between most commercially important rule systems
- First standards specifications ready, quite stable

Current state: Datalog. Planned extensions:

- Logically more expressive: full Horn logic programs
- Reactive rules
- Nonmonotonic rules

Vision: Uses of Rules in E-Business

Rule based business policies and business processes:

- Represent seller's offerings: Products and services, bids, capabilities, common catalogue of multiple suppliers
- Represent buyer's requests, interests, bids
- Matchmaking between the two
- Represent sales help, customer help, authorization/trust
- All at a high level of abstraction

Semantic Web Services

Web services: Procedures/methods which are invoked through a Web protocol interface, typically with XML inputs and outputs

Semantic Web Services = Semantic Web + Web Services

- Knowledge-based service descriptions
- Discovery/search
- Invocation
- Negotiation, deals, e-contracts
- Composition, planning

Semantic Web Services: Some High-Level Tasks

Web service discovery

- Find me a shipping service that will transport frozen vegetables from San Francisco to my place

Web service invocation

- Buy me “The Lord of the Rings” at www.amazon.com

Web service deals, and their negotiation

- Propose a price with shipping details for used Del laptops to Sue Smith

Web service selection, composition and interoperation

- Make the travel arrangements for the next WWW conference

Semantic Web Research at ICS-FORTH

Key researchers:

- Grigoris Antoniou,
- Vassilis Christophidis
- Panos Constantopoulos
- Martin Doerr
- Dimitris Plexousakis

Semantic Web Research at ICS-FORTH (2)

Main topics:

- RDF Suite
- Semantic Web Services
- Rules for the Semantic Web
- Integration of large, heterogeneous ontologies (SEMKOS)
- Applications: P2P, Web portals, E-Learning

The ICS-FORTH RDF Suite

One of the leading environments for maintaining and processing RDF/S documents

Main components:

- The Validating RDF Parser (VRP)
- The RDF Schema Specific Data Base (RSSDB)
- The RDF Query Language (RQL)
- The RDF/S View Language (RVL)

Semantic Web Applications at ICS-FORTH

E-Learning:

- Models of content, learner and learning strategies
- Learner-centric processes, individualized courses
- Free from time and location constraints
- Material can be reused

Cultural Heritage:

- CIDOC Conceptual Reference Model
- About to become an ISO standard

P2P for Knowledge Management

Research Agenda for the Semantic Web

Exploit Data and Knowledge Base technologies for supporting a global economy of Knowledge

- Formal models of resource descriptions and schemas
- Ability to define mappings between semantic representations
- Value-added services: querying and rule-based inferencing, belief revision, knowledge management, service brokering and composition
- Scalable infrastructures: peer-to-peer systems, semantic grids

Applications Agenda

Semantic Web Services

E-Commerce

E-Learning

Medicine

Knowledge Management

E-Government

Will the Semantic Web Succeed?

Big pluses:

- growing need for semantic search of information
- Web services will not work without ontologies

Potential initial successes:

- In closed environments with central authority (Knowledge Management, E-Government),
- Large virtual collections, e.g. cultural heritage or medicine

Will the Semantic Web Succeed? (2)

The make or break: Availability of large amounts of semantic content

Technically the Semantic Web can be built… today