Ontology-Based Computing

Kenneth Baclawski
Northeastern University and Jarg
The Onslaught

- Increasingly large amounts of information is becoming accessible electronically.
- The information sources are increasingly complicated.
- The diversity of types of information source is also increasing.

Technologies are emerging to cope with this onslaught: ontology-based computing.
Ontologies

- Shared understanding within a community of people
- Declarative specification of entities and their relationships with each other
- Constraints and rules that permit reasoning within the ontology
- Behavior associated with stated or inferred facts
Relational Database Schemas

- Well established technique for specifying the structure of shared data, not for communication between people or agents
- Declarative specification but of tables, not of entities and relationships
- Some constraints are expressible but no significant rules (such as inheritance)
- No explicit behavior
Object-Oriented Schemas

Emerging technology for communication between software components
Declarative specifications
Constraints and some rules
Several ways to specify behavior
The Unified Modeling Language (UML) is the standard OO modeling language.
Logic

- Very expressive but very difficult to use. Not designed for communication.
- Most logical languages are not based on entities and relationships.
- Very powerful inferencing capabilities.
- Do not usually have any associated behavior.
- Many examples: Prolog, KIF, ...
XML DTDs and XML Schema

- Defines a hierarchical document type. XML Schema defines data types. Designed for communication over the Web.
- Good support for entities and hierarchical relationships; awkward for others.
- Constraints can be imposed on the hierarchical structure and on data types.
- Behavior can be specified procedurally.
Knowledge Representations

- Very well developed branch of AI. Many tools, but mostly academic. Not yet used for communication over the Web.
- Powerful language for specifying entities and their relationships.
- Most are linked with inference engines.
- Behavior is typically handled in an ad hoc manner.
RDF and DAML

Resource Description Framework (RDF) is a knowledge representation language. It is a WWW Consortium Recommendation.

The DARPA Agent Markup Language (DAML) is an extension of RDF to serve as the basis for ontology-based computing over the Web: the Semantic Web.
Ontological Reasoning in RDF

1. Type constraint violation: The range of owns is Fish
2. There is no inconsistency: Wanda is a Fish (Mermaid?)
1. Cardinality constraint violation: George can’t have two majors
2. There is no inconsistency: Engineering = Arts & Sciences
Representing information

- Relational database: records
- OO database: instances
- Logic: facts
- XML: documents
- KR: annotations

All of these are graph structures: entities related to other entities by relationships.
Where is the meaning?

- Databases: select-project-join queries
- Logic: rules determined by unification
- XML: XSLT patterns
- KR: templates

All of these are forms of graph matching. The units of meaning are small connected subgraphs.
Coping with the onslaught

- **Volume of data**
  - Use distributed systems and indexes

- **Complexity of ontologies**
  - Graphical user interfaces
  - Consistency checking

- **Diversity of ontologies**
  - Ontology mediation