

INFS3101 / 7100 Ontology and the Semantic Web

Module 8 Quality

Last lecture: Key Terms



- ❖ **BWW and Dolce are formal upper ontologies.** Include **systems, endurants, perdurants, subclasses**, etc. Formal upper ontologies provide a rich meta vocabulary to help develop ontologies and suggest **abstract data types** which can be supported by ontology servers to make it easier to build rich ontologies.

Principles for the Design of Ontologies

- ❖ An ontology is an explicit specification of a conceptualization
- ❖ We say that an agent commits to an ontology if its observable actions are consistent with the definitions in the ontology.
- ❖ Pragmatically, a common ontology defines the vocabulary with which queries and assertions are exchanged among agents
- ❖ Want ontology to be comprehensible to humans

Design Criteria

- ❖ **Clarity** - effectively communicate intended meaning
- ❖ **Coherent** - sanction consistent inferences
- ❖ **Extendible** - anticipate uses of shared vocabulary
- ❖ **Minimal encoding bias** - as free as possible from implementation decisions
- ❖ **Minimal ontological commitment** - as few claims as possible about the world being modeled.

Clarity

- ❖ "There's glory for you", [said Humpty]. "I don't know what you mean by "glory"; Alice said. Humpty Dumpty smiled contemptuously. "Of course you don't -- till I tell you. I meant "there's a nice knock-down argument for you!" "But "glory" doesn't mean "a nice knock-down argument"; Alice objected. "When I use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean -- neither more nor less."

▪ *Alice Through the Looking Glass*, Chapter VI

- ❖ If programs are to interoperate, they must give up this power and commit to an ontology

Clarity

- ❖ Ontology is a list of words intended to be interpreted by more than one agent
- ❖ Ontology has agreed interpretation of each
- ❖ How to exclude unintended interpretations?
 - Australian immigration student visa ontology includes "full-time enrolment"
 - Soccer includes "offside" and "goal"
 - Railway ontology includes "on-time/late"
- ❖ Contentious terms generally defined in terms of easily measured terms.

Clarity

- ❖ Definitions in more or less formal ways.
- ❖ Most formal is predicate calculus
- ❖ More formal definitions exclude more unintended interpretations.
- ❖ But not possible to exclude all
 - Skolem-Lowenheim theorem
- ❖ Possibility of clarity increases with formality of representation, but comprehensibility suffers
- ❖ Highly formal systems need tools to help with comprehension

Coherence

- ❖ Ontology should sanction consistent inferences
- ❖ Distinction between conceptualization and specification.
- ❖ Specification must sanction inferences at all, so needs high degree of clarity for coherence to be relevant
- ❖ Formal representations support computer-based reasoning
- ❖ Bib-1 sanctions no inferences. Defined subclasses sanction more than declared. Partially defined are in between

Coherence

- ❖ Possible reasoning tasks
 - Check for circular definitions "Catch-22"
 - Residence permit possible if have a job, but need a job to qualify for a residence permit
 - Inconsistency
 - You are a full-time student by the rules of the University but not by the rules of the immigration department
 - Representation allows inferences normal in conceptualization.
 - Conclude that one year is longer than 6 months
 - Tools that support formal representations also support automated reasoning

Extendibility

- ❖ Want it to be easy to extend or adapt the ontology
- ❖ Not possible to predict the future
- ❖ But change is easier if every aspect of the conceptualization is represented in one place in the representation, and one place only.
- ❖ Extendibility related to database normalisation
- ❖ Normalisation prevents update anomalies by storing dependencies once only

Extendibility

- ❖ Complex set of regulations will have clauses which apply in many cases. If repeated in all cases, then to change must change in many places. More extendible to factor common clauses out.
- ❖ Cost of lack of redundancy is difficulty of comprehension.
- ❖ Toolset can help by creating complete representation on the fly, like a database view.

Extendibility

- ❖ Can have potential redundancy
- ❖ A single concept can have a complex representation
- ❖ US president must be born in US and more than 35 years old.
- ❖ More extendible of ontology includes concepts
 - Native-born citizen: born in US
 - Mature person: age greater than 35 years
 - US president must be conjunction of the two
- ❖ Extendibility more concerning as clarity increases, as more complex concepts become more possible to represent

Encoding Bias

- ❖ Unnecessary incorporation of implementation decisions in ontology
- ❖ Instance of software engineering principle of delaying design decisions as long as possible since deeper decisions are more costly to change.
- ❖ Encoding bias depends on what ontology used for
 - Software design component further from implementation than ontology supporting interoperation

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

13

Encoding Bias

- ❖ Area of circle is $\pi \times \text{radius squared}$
- ❖ Applies to radius measured in angstrom units as well as radius measured in parsecs
- ❖ Applies in carpentry where metres accurate to a few millimetres, as well as precision machinery where metres need to be accurate to a few micrometres. Need more precision in π in latter case.
- ❖ Basic geometry ontology least encoding bias
- ❖ Incorporating units increases encoding bias
- ❖ Specifying precision of numbers increases more

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

14

Encoding Bias

- ❖ Specifying Unicode increases encoding bias
- ❖ Fixed-length text more encoding bias than variable-length
- ❖ Interoperating systems need more implementation decisions.
 - Need to know whether text in Unicode
 - Need to know precision of numbers
 - But internal representations like IEEE double float are encoding bias even here, since internals of agents not visible to ontology.

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

15

Ontological Commitment

- ❖ Relates to the conceptualisation rather than to the specification
- ❖ How easy is it to re-use a conceptualisation in developing the conceptualisation for another similar world?
- ❖ The Periodic Table has a low ontological commitment because it can be used in any application relevant to the chemical elements.
- ❖ Enterprise packages like SAP have a high ontological commitment, because to adopt them an organisation must change many business practices.

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

16

Ontological Commitment

- ❖ Depends partly on what it means to be similar
- ❖ The rules for chess have a low ontological commitment because if you don't follow the rules you aren't playing chess
- ❖ The rules for American Major League Baseball have a high ontological commitment because lots of people play baseball in different ways.
- ❖ Chess has no similar games, while American Major League Baseball has many

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

17

Ontological Commitment

- ❖ Ontologies of institutional facts
- ❖ Framing rules for speech acts and contexts for institutional facts are governed by particular bodies in an institutional network
- ❖ The University has general rules about what it means to offer, take, and be credited for a course.
- ❖ Individual courses differ on how they are run
- ❖ But they all use the university's ontology
- ❖ Ontological commitment low within a university
- ❖ But hard to re-use one university's system in another
- ❖ Ontological commitment high outside the scope of the governing body

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

18

Ontological Commitment

- ❖ Sometimes the conceptualisation is a design issue
- ❖ Z39.50 originally included Bib-1 in its ontology, so an application had to commit to Bib-1 in order to commit to Z39.50
- ❖ Made it difficult to re-use Z39.50 outside libraries,
- ❖ So ontological commitment high.
- ❖ Present formulation separates the Use attributes into separate modules
- ❖ Ontological commitment now low

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

19

Ontological Commitment

- ❖ What we mean by ontological commitment depends on what we use the ontology for
- ❖ Gruber's ontologies are used in software engineering. The ontologies are essentially standard class libraries embodying scientific or engineering concepts
- ❖ Particular ontologies for particular applications incorporate these libraries without necessarily interoperating with each other
- ❖ To turn ontology into application, must refine to implementation

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

20

Ontological Commitment

- ❖ But we can reduce ontological commitment by modularisation, like Z39.50
- ❖ Can package a dimension ontology horizontally
 - Basic dimensional concepts: distance, area, time
 - Systems of units: KMS, CGS, FPS
 - Number precision arithmetic
- ❖ Can package a movie product/review/showing ontology vertically into its components
- ❖ Both of these strategies reduce ontological commitment

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

21

Ontological Commitment

- ❖ Another strategy is judicious abstraction
- ❖ Bibliographic ontology
 - Reference to publication
 - Publication in document
 - Publication has author.name:string, title: string, date:**timepoint**
 - Document has author.name:string, title:string, publisher.name:string
 - Author is agent, agent can be a person
 - Document can be book or journal
 - Functions:**timepoint** -> year, date, season, etc

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

22

Ontological Commitment

- ❖ Timepoint is primitive point in historical time
- ❖ Allows varying degrees of precision
- ❖ Multiplicity of entities can be more than application needs, but can be used to state integrity constraints
- ❖ Timepoints rather than dates to avoid encoding bias
- ❖ Names distinguished from objects because so distinguished in the world
- ❖ Timepoint is like a dimension with units year, season etc.
- ❖ Can add module with Gregorian year/month or Muslim year/month
- ❖ Or year/season or date or date/time
- ❖ Depending on the publications the ontology is intended to be used for
- ❖ Reduces ontological commitment by horizontal modularisation

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

23

Use of Principles

- ❖ Want to maximise clarity, coherence and extendibility
- ❖ Subject to cost
- ❖ Compensate loss of comprehensibility by tool support
- ❖ Want to minimise encoding bias and ontological commitment
- ❖ Subject to cost
- ❖ Tool support can help

INFS3101/INFS7100 week 6, 22 November, 2005

Bob Colomb

24

Cost/Benefit Considerations

- ❖ Sometimes cost is infinite. Defined subclasses for SNOMED Diagnoses is beyond current medical knowledge
- ❖ Benefit comes from reduced cost of implementations
- ❖ Improvement must be implementable.
- ❖ Ontologies typically implemented many times
- ❖ So benefit can repay considerable cost.

Summary: Key Terms



- ❖ Quality principles for ontologies include **clarity, coherence, extendibility, encoding bias** and **ontological commitment**. Want to maximise the first three and minimise the last two.
- ❖ Quality involves **cost-benefit tradeoffs**.

Resources

- ❖ **Essential**
 - Notes chapter 8 Quality
- ❖ **Further**
 - ❖ Gruber, T.R. (1993) *Toward Principles for the Design of Ontologies Used for Knowledge Sharing* Technical Report KSL 93-04 Knowledge Systems Laboratory Stanford University (on web site)