TOWARDS PRINCIPLES FOR THE DESIGN OF ONTOLOGIES USED FOR KNOWLEDGE SHARING

Presented by Thomas R. Gruber
@ International Workshop on Formal Ontology, March’93.
Introduction

- Analysis of design requirements for shared ontologies.
- Proposal for design criteria to guide the development of ontologies for knowledge-sharing purposes.
Conceptualization: (defn.)

- It is an abstract, simplified view of the world that we wish to represent for some purpose.

- It is the objects, concepts and other entities that are assumed to exist in some area of interest and their relationships.
Ontology:

- It is an **explicit** specification of conceptualization. Defines the vocabulary with which set of agents can communicate.

- **Ontological commitments are agreements to use the shared vocabulary in coherent and consistent manner.**

A commitment to a common ontology is a guarantee of consistency but not completeness.
Design criteria for ontologies:

- **Clarity**: An ontology should effectively communicate the intended meaning of defined terms.

- **Coherence**: An ontology should give inferences that are consistent with the definitions.

- **Extendibility**: An ontology should have the capability to add new terms for special uses without revision of existing vocabulary.
**Minimal encoding bias:** An ontology should be independent to the issues of implementing language.

**Minimal ontological commitments:** An ontology should require the minimal ontological commitment that is sufficient to support knowledge sharing.
Tradeoffs:

- If particular about clarity that may result in limitation to extendibility.

- If we decide on extensible ontology then ontological commitment will be increased.

- The more clarity in definition will lead to encoding bias.

- The extensible or general ontology may result in incoherent in the system. So require careful coding.
Knowledge interchange format (KIF):

- *It* is a computer-oriented language for the interchange of knowledge among disparate programs. It has declarative semantics, and it is logically Comprehensive.

It provides
- the representation of knowledge about the representation of knowledge;
- the representation of nonmonotonic reasoning rules;
- the definition of objects, functions, and relations.

Overview of KIF:

- KIF is a prefix notation for predicate calculus with functional terms and equality.

  - Free variables: `prefix ? instance of the class C; he class C; =`  
    - Implication: `=>,<=,=>`
Case study:

- How to represent the mathematical models of engineered systems.

Aim is to define the vocabulary and conceptual foundation for sharing this model among the systems.

Important concepts in the ontologies are
  - physical-quantity (5 meter)
  - Physical-dimension (length, length/time),
  - unit-of-measure (meters, kilometers/hour),
  - magnitudes
Description of physical-quantity in KIF form:

(defrelation PHYSICAL-QUANTITY
  (<=> (PHYSICAL-QUANTITY ?q)
    (and (defined (quantity.magnitude ?q))
      (double-float (quantity.magnitude ?q))
      (defined (quantity.unit ?q))
      (member (quantity.unit ?q)
        (setof meter second kilogram ampere Kelvin mole candela)))))

Constructor function (instance) for the above class:

(deffunction THE-QUANTITY
  (<=> (and (defined (THE-QUANTITY ?m ?u))
    (= (THE-QUANTITY ?m ?u) ?q))
  (and (physical-quantity ?q)
    (= (quantity.magnitude ?q) ?m)
    (= (quantity.unit ?q) ?u))))
Analysis of version: 1

- **Encoding bias:**
  - Double float
  - Set of units

- **Limit on extendibility:**
  - No possibility to add other units
(defrelation PHYSICAL-QUANTITY
  (<= (PHYSICAL-QUANTITY ?q)
    (and (defined (quantity.magnitude ?q))
      (magnitude (quantity.magnitude ?q))
      (defined (quantity.unit ?q))
      (unit-of-measure (quantity.unit ?q))))))

Now, define magnitude class of real numbers, comes from KIF’s number ontology

  (defrelation MAGNITUDE
    (<= (MAGNITUDE ?x)
      (real-number ?x)))

To accommodate alternate sets of units,

  (defrelation UNIT-OF-MEASURE
    (class UNIT-OF-MEASURE))
Contd…

Ex:
Getting new unit (velocity) by defining the basic units (length, time)

(defobject METER (basic-unit METER))
(defobject SECOND (basic-unit SECOND))
(defobject METER/SECOND
  (= METER/SECOND (unit* meter (unit^ second -1))))
(= REAL-FAST (the-quantity 10000 meter/second))
If the representation allows one to infer something that is not true in the conceptualization then the ontology is **incoherent**.

**EX:**

If

\[(= \text{(quantity.unit (the-quantity 6 foot)) foot})\]
\[(= \text{(quantity.unit (the-quantity 2 yard) yard})\]
and

\[(\text{not} \ (= \text{foot yard}))\]
then

\[(\text{not} \ (= \text{(the-quantity 6 foot) (the-quantity 2 yard)}))\].
Avoiding coherence

(deffunction QUANTITY.MAGNITUDE

  (=> (and (defined (QUANTITY.MAGNITUDE ?q ?u))
          (= (QUANTITY.MAGNITUDE ?q ?u) ?m))
      (and (physical-quantity ?q)
           (unit-of-measure ?u)
           (magnitude ?m))))
If we introduce the concept of physical-dimension in the representation we can have compatibility among quantities and units.

- (defrelation PHYSICAL-DIMENSION
  (class PHYSICAL-DIMENSION))

- (defrelation UNIT-OF-MEASURE
  (=> (UNIT-OF-MEASURE ?u)
   (and (defined (unit_dimension ?u))
    (physical-dimension (unit_dimension ?u)))))

- (deffunction QUANTITY.DIMENSION
  (=> (physical-quantity ?q)
   (and (defined (QUANTITY.DIMENSION ?q))
    (physical-dimension (QUANTITY.DIMENSION ?q)))))
Analysis:

- Separating the core ontology from the specific conventions minimizes the **ontological commitment** of participating agents.

  set of all agents need to commit to the core theory, they can commit to differing standards of measure.

Since commitment to an ontology does not require completeness of inference, agents can “understand” the conditions under which a value exists without knowing how to compute the value.
Case study :2

- Design the ontology for sharing bibliographic data.

Goal of bibliographic ontology is

- Facilitate automatic translation among existing database.
- To support the specification of reference-formatting styles independently of DB or program.
- To provide knowledge level vocabulary for interacting (bibliographic data base search)
(defrelation REFERENCE
  (=> (REFERENCE ?ref)
      (and (defined (ref.document ?ref))
           (defined (ref.title ?ref))))

(deffunction REF.DOCUMENT
  (=> (and (defined (REF.DOCUMENT ?ref))
      (= (REF.DOCUMENT ?ref) ?doc))
     (and (reference ?ref) (document ?doc))))

(deffunction REF.TITLE
  (=> (and (defined (REF.TITLE ?ref))
      (= (REF.TITLE ?ref) ?title))
     (and (reference ?ref) (title-name ?title))))
Analysis:

- Provides integrity constraints.
- Addition of conceptual entities with data will reduce encoding bias.
- Improves the identifiers to objects.
Summary

- The evaluation of design depends on available knowledge and application anticipated for domain.

In the engineering domain

- If encoding is intrinsic to the conceptualization encoding bias can be avoided by representing both class and concepts.
- Studied how to extend the representational vocabulary without over commitment.
- Ontology related the concepts and specifies the terms more clearly.
- Provides the useful inferential services.
Thank you