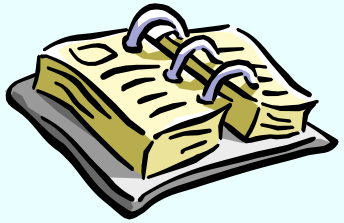
Ontology Development

Πληροφοριακά Συστήματα Διαδικτύου

Contents

- Ontology Development Background
- Methodology:
  - Kick-off
  - Refinement
  - Evaluation
  - Maintenance & Evolution



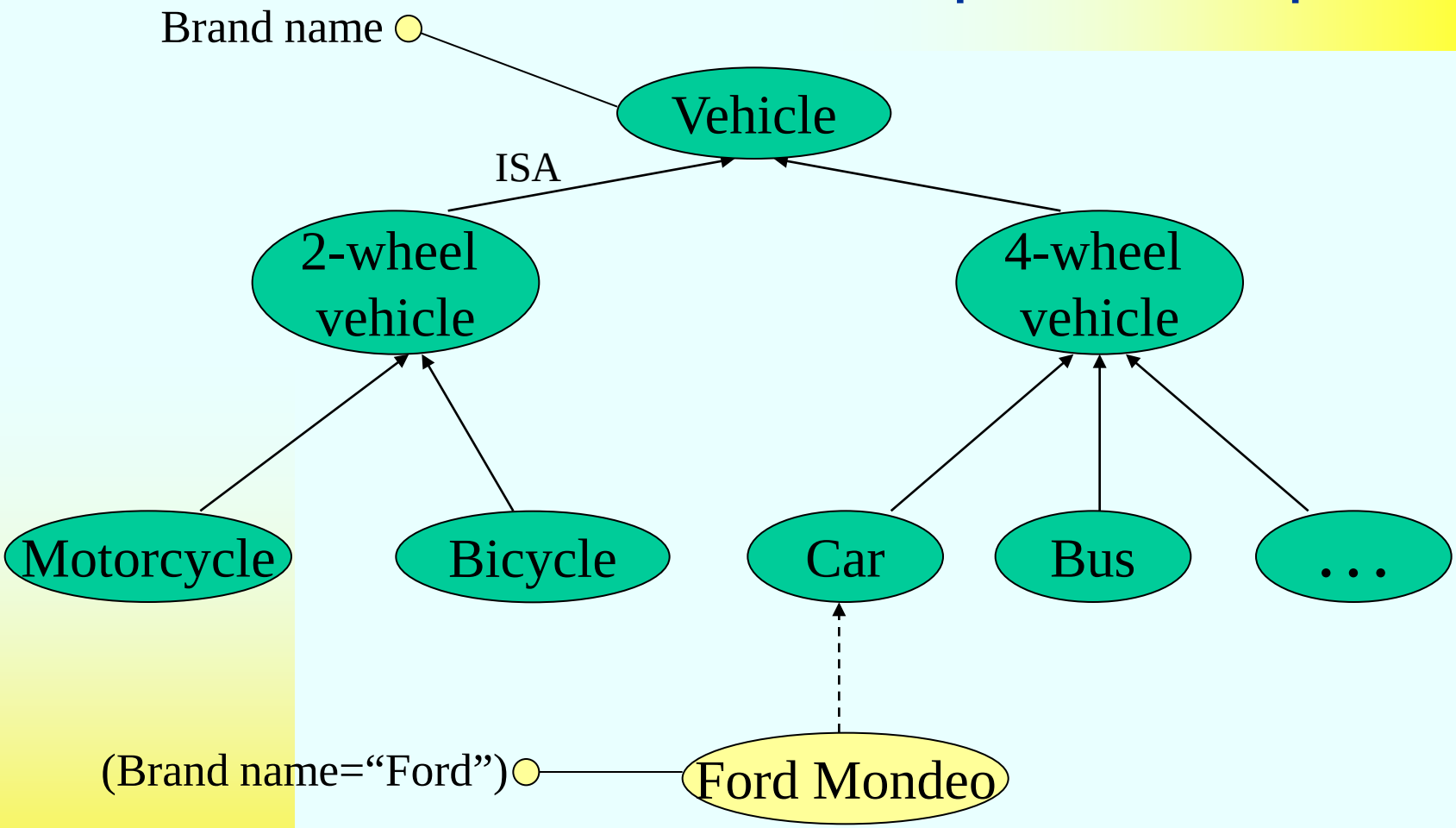
Ontology development: Motivation

- Why creating an ontology?
  - Sharing a common understanding of the information in a knowledge domain
  - Improving interoperability among applications that use the domain
  - Making domain assumptions explicit so that applying changes as these assumptions evolve becomes easier
  - Enabling re-use of the knowledge domain

Ontology development vs. O.O. Software design

- Ontology development is different from object-oriented software design
  - In OO software design the operational properties of a class are important
  - In ontology design the structural properties of a class – how each class relates to other classes in the hierarchy – are important

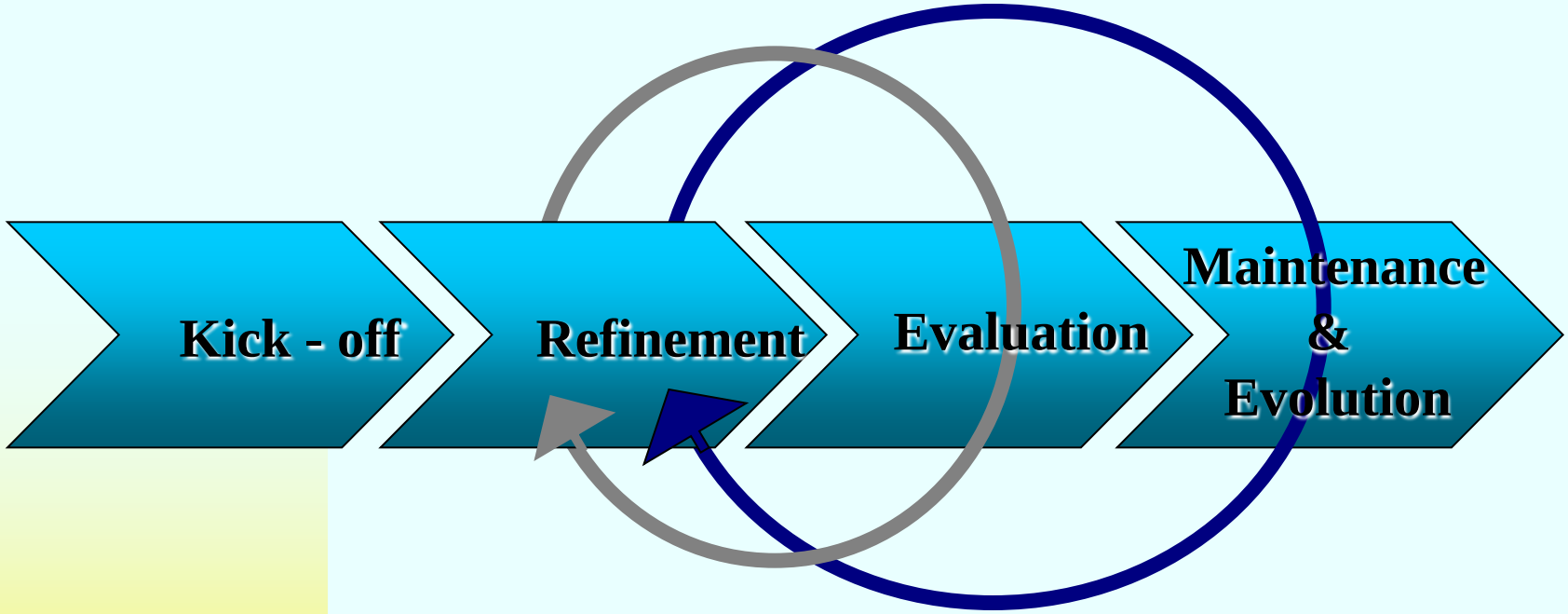
Ontology development: A simple example



Ontology development: A complex process

- Ontology construction is not yet a well understood process
  - Size
  - Complexity
  - Shared understanding within a group of people with different backgrounds
- There is not a single correct methodology for building an ontology
  - Applications using the ontology
  - Anticipated extensions

OntoKnowledge methodology



http://www.ontoknowledge.org/

OntoKnowledge methodology

Ontology development is an iterative process

1. Kick-off phase: capturing the ontology
   a. Define the requirements specification
   b. Build a first draft version in a semi-formal description
2. Refinement phase: formal representation
3. Evaluation phase
4. Maintenance and Evolution phase

Kick-off phase

- Step 1: Obtain the concepts and the relations between them
- Step 2: Build a classification of the concepts in the ontology

Kick-off phase: Step 1 - Brainstorming

- What is the domain the ontology will cover?
- What is the ontology going to be used for?
- Brainstorming session
  - All involved must collectively possess sufficient domain expertise
  - Quickly capture potentially relevant terms and phrases of the domain

Kick-off phase: Step 1 – Sources & Usage scenarios

- Identify knowledge sources
  - Existing ontologies
  - Domain experts (interviews, CQs, etc.)
  - Dictionaries, internal documents, etc.
- Potential users and usage scenarios
  - Identify roles
  - Identify how the ontology will be used
  - Identify terms and potential relations between them

Kick-off phase: Step 1 - Competency questions

- Competency questions (CQs): basic questions the ontology should be able to answer
  - Constitute expressiveness requirements
  - Indicate scope and content of the ontology
  - Used as reference during the evaluation
  - Ideally defined in a modular manner (higher level questions require the answer of lower level questions)

Kick-off phase: Step 1 – CQ examples

- Example competency questions
  - Is “Ford Mondeo” a 2-wheeled or a 4-wheeled vehicle ?
  - What is the brand name of “Ford Mondeo” ?
  - What brands manufacture busses ?
  - What brands manufacture both 2-wheeled and 4-wheeled vehicles ?

Kick-off phase: Step 1 – Ontology reuse

- Reusing existing ontologies
  - Worth considering what others have done
  - Due to compatibility requirements
  - If applicable in the problem at hand:
    - Extend and refine to fit specific needs
    - Possibly integrate with other ontologies

Kick-off phase: Step 1 - Guidelines

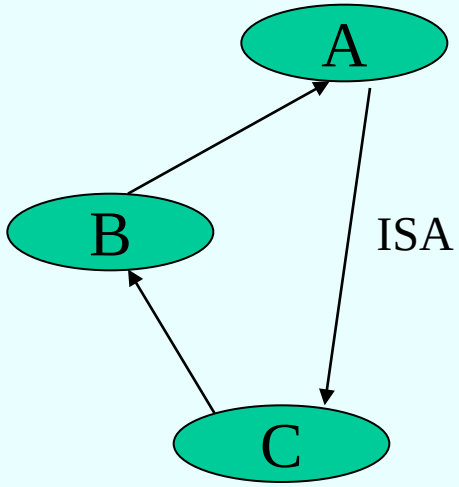
- Design guidelines: provide guidelines for users not familiar with ontology modelling
  - Estimate on the number of concepts
  - Level of granularity (how many classes in each level)
  - Naming rules for concepts and relations
  - Concepts in the ontology should be close to objects (physical or logical) and relations in the target domain

Kick-off phase: Step 1 – Guidelines (Cntd)

- Clarity: the ontology should minimize ambiguity
- Coherence: the ontology should be internally consistent
- ISA modelling: a subclass represents a concept that is “kind of” the concept represented by the superclass
- Synonyms for the same concept should not represent different classes
- Siblings in the hierarchy must be in the same level of generality

Kick-off phase: Step 1 – Guidelines (Cntd)

- Cycles in the hierarchy should be avoided



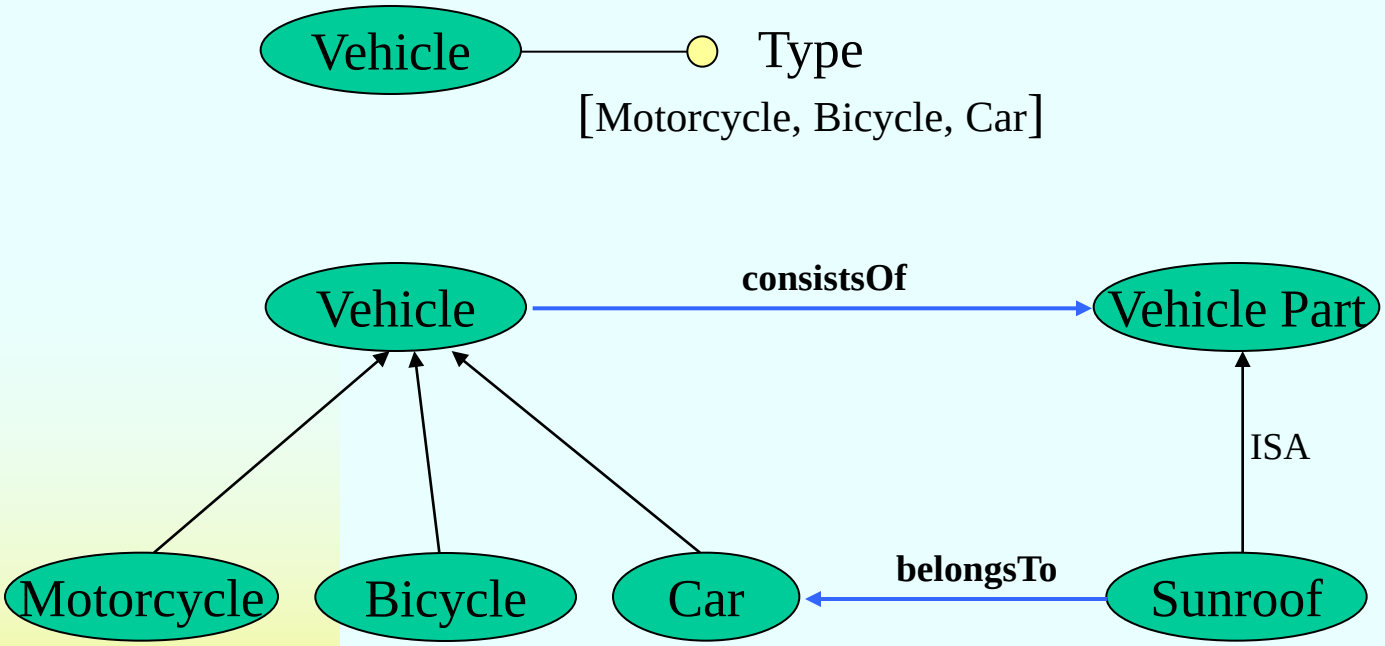
Kick-off phase: Step 1 – Guidelines (Cntd)

- A new class or a new property value?
  A new class when
  - The concept is important in the domain
  - The concept becomes a restriction for some property in some other class

  A new property value when
  - The concept represents an extrinsic property of some class (number, color, location, etc.)

...

Kick-off phase: Step 1 – Guidelines (Cntd)



Kick-off phase: Step 1 – Guidelines (Cntd)

- When to introduce a new class? When the concept to model
  - Has additional properties
  - Has different restrictions
  - Participates in different relations
- Modelling a concept as an instance or as a class?
  - Individual instances represent the most specific concepts in the knowledge domain
  - Depends on applications using the ontology

Kick-off phase: Step 1 – Guidelines (Cntd)

- Minimal ontological commitments: limiting the scope of the ontology
  - Domain assumptions
    - Make as few claims as possible
    - Make as much claims as necessary
  - Domain information
    - Do not include all possible information
    - Include what is necessary for the applications that will use the ontology

Kick-off phase: Step 2 – The different approaches

- Define the classes of the ontology
  - Top-down approach: Model the most general concepts first and refine them
  - Bottom-up approach: Find the most specific concepts and generalize to obtain higher level concepts
  - Middle-out approach: Identify the most important concepts and obtain the rest by specialization and generalization

Kick-off phase: Step 2 – Pros and cons

- Defining classes – pros and cons of different approaches
  - Top-down: high quality, fine grained ontology but may not cover all domain information
  - Bottom-up: less quality, more complete ontology addressing all available domain information
  - Middle-out: more natural, concentrates on what is important and builds on from there, better controls the desired level of detail

Kick-off phase: Step 2 – Defining properties & relations

- Define the relations of the ontology
  - Identify which domain terms represent relations or properties of some classes
  - A property should be attached to the most general class that can have this property
  - Kinds of properties
    - Intrinsic (flavor, color)
    - Extrinsic (name, location)
    - Relations to other classes

Kick-off phase: Step 2 – Defining restrictions

- Restrictions on properties
  - Cardinality: how many values a property has
  - Value type: String, Number, Boolean, Class, etc.
  - Domain: the classes to which the property is attached
  - Range: the allowed classes, if the value-type is Class

Utilizing Facets

- Facets can be viewed as different perspectives, viewpoints, or dimensions of a particular domain
- Based on synthesis of classes from subject statements of the domain – conceptual clustering
- A faceted scheme provides a controlled vocabulary in the form of terms arranged systematically by facets and a set of rules on how to combine such terms to define conceptual categories

Facets Example

- Consider the following statements
  - A comfortable, luxurious red Mercedes
  - An expensive fast Maserati
  - A cheap slow Punto
- Conceptual grouping will result in
  - Brand: Mercedes, Punto, Maserati
  - Price: expensive, cheap
  - Performance: slow, fast
  - Quality: comfortable, luxurious

Refinement phase

- Produce a formal representation of the ontology described in the previous phase
  - Refine the ontology
  - Choose a representation language (RDFS, DAML+OIL, OWL, etc.)
    - Expressive power
    - Tool support for reasoning
  - Formalize the representation and produce a mature application-oriented ontology

Evaluation phase

- Evaluate the resulting ontology
  - Using as reference
    - Requirements specifications
    - Competency questions
    - Usage scenarios
  - Test the ontology in the target application environment
    - Obtain usage patterns by monitoring user navigation
    - Recheck less frequently accessed parts for relevance
    - Possibly expand parts accessed with high frequency

Maintenance and Evolution phase

- Apply changes to the ontology while it evolves over time
  - Who is responsible for maintaining the ontology?
  - How the maintenance is going to be performed
- Before each major change all possible effects must be thoroughly examined
- Thoroughly documented according to purpose and scope

The importance of documentation

One of the main barriers to effective knowledge sharing is the inadequate documentation of knowledge bases and ontologies

- All important assumptions should be documented
- Tools that facilitate both formal and informal documentation are necessary

Formal procedure for the generation and evaluation of the conceptualization