Conceptual Modeling of Topic Maps with ORM Versus UML

Are D. Gulbrandsen
The XML group, Center for
Information Technology Services,
University of Oslo, Norway

1

Overview

- Introduction
- Topic Maps notation in UML and ORM
- Case Study - The Houston Ontology
- Model language evaluation criteria

2

UML: When I say UML in this context I mean UML2 class diagrams

Introduction

- There's a need for a graphical notation for ontology modeling
- There are a few people using UML Class diagrams for this
- Is ORM (Object Role Modeling) a good tool for Topic Maps ontology modeling?
  - Triggered by a discussion with Lars Marius Garshol at the first Norwegian Topic Map Conference in 2002

3

-We need a graphical model for easy communication when analysing or documenting

Goals

- To model and document a Topic Maps Knowledge Base for systems administration
  - The Houston topic map case
- Community building
  - To meet other people with common interests
- We have tools today that can generate a relational database from UML or ORM
  - My vision is a tool that can generate a TMCL schema from a conceptual model

4

-Community building: Haven't found many people interested in discussing this (Except Lars Marius Garshol)

-We want to use a formal model to capture the ontology, which later can be mapped to the TM metamodel.

Ontology

- The product of an attempt to formulate an exhaustive and rigorous conceptual schema about a domain
- Typically a hierarchical data structure containing all the relevant entities and their relationships and rules within that domain (e.g., a domain ontology).
- The computer science usage of the term ontology is derived from the much older usage of the term ontology in philosophy

» wikipedia.org

5

A little about meta models

- The TMDM standard defines a metamodel for Topic Maps
- The Tau model defines a TM metamodel
  - The foundation of the query language TMQL and the constraint language TMCL
- Work within OMG:
  - The Ontology Definition Metamodel RFP
    - Includes the metamodels for UML2, RDF Schema, OWL, SCL, ER and Topic Maps
  - The Business Semantics of Business Rules RFP
    - Includes ORM, UML2, and OWL

6

-Ref: The formal system, Jack Parks keynote

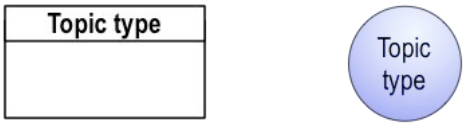
-It would have been interesting to discuss how the different metamodels corresponds to the Topic Maps metamodel, but I think TMCL and TMQL should be included in a discussion of this, and this has to be left for future research.

-Will take the easy route today: The pragmatic approach, With the root in the Natural system, the organization and people

Topic Maps notation

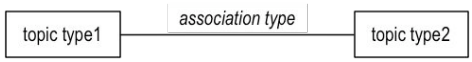
- The three main constructs:
  - Topics
  - Associations
  - Occurrences.
- Other important constructs:
  - Association Roles
  - Topic Names (Base Names and Variant Names)
  - Types
  - Scope (Theme)

Topics



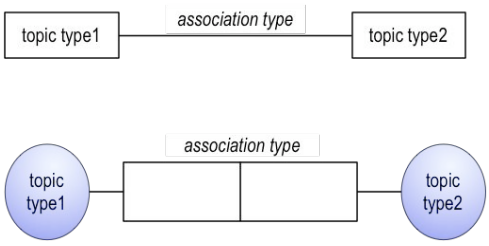
- UML: class
- ORM: object

Associations in UML

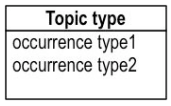


- A simple association in UML
  - Can also use association classes

Associations in UML and ORM

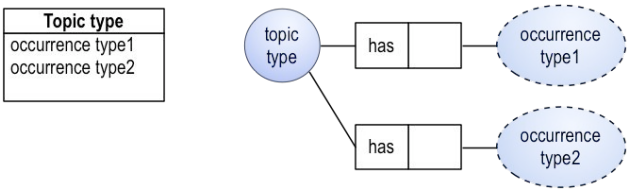


Occurrences in UML



11

Occurrences in ORM



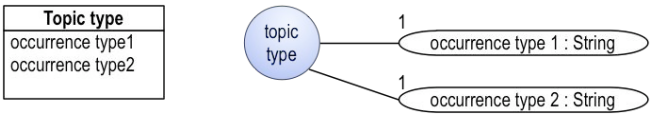
- UML: attributes
- ORM: object
  - ORM classifies objects into entity types (topic) and value types (occurrence)

12

-We can see that UML is more compact, but see the use of a simplified view in the next slide.

-An occurrence is essentially a specialized kind of binary association, where one participant in the association must be a topic and the other an information resource

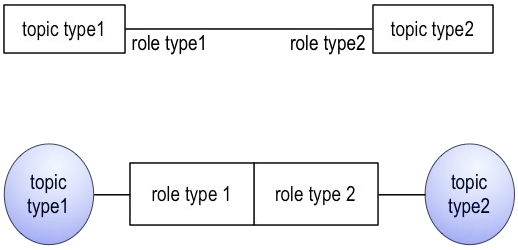
Occurrences ORM view



13

A simplified view is an abstraction mechanism in ORM.

Association Roles



Type, scope and names

- In a conceptual model everything is types
- Scope (set of themes)
  - I have not found any standard way of modeling this with UML
  - Have suggested modeling this in ORM as a textual notation, inspired by LTM
- Names
  - Will not go into details. Modeled with a rectangular shape

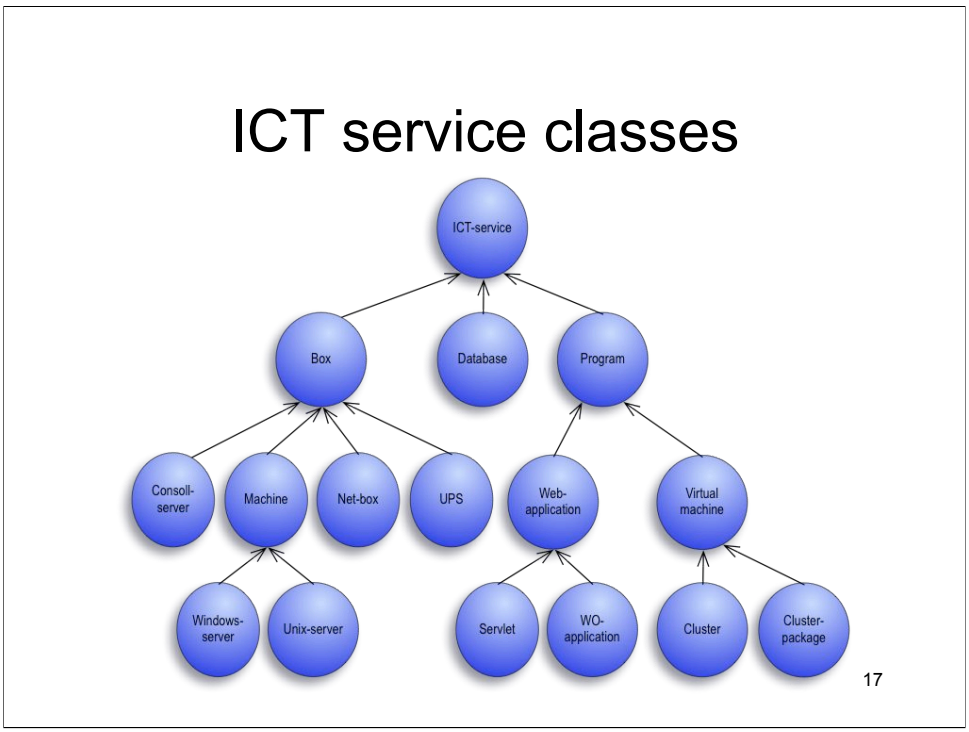
The Houston Ontology

- The University of Oslo is one of the largest universities in northern Europe, with approximately 30000 students and 4600 employees
- ICT control centre called Houston
- A Topic Maps Knowledge Base for systems operation, administration and maintenance documentation
  - An online version of parts of the ontology: http://folk.uio.no/areg/topicmaps/HoudiniOntology/

16

-Houston monitors and administers a wide range of ICT-service types, spanning from gateways to Digital Library services.

-ICT - Information and Communication Technology. We are responsible for network and telephone system as well as other IT-services



-The different types of ICT-services form a class-hierarchy.

-This is used both for the classification of services and for object oriented specialization, where a subclass can add required documentation attributes.

-Hardware, for instance, may have a geographical address and picture, while a web application is addressable by a URI.

-This current class hiererchy is very pragmatic

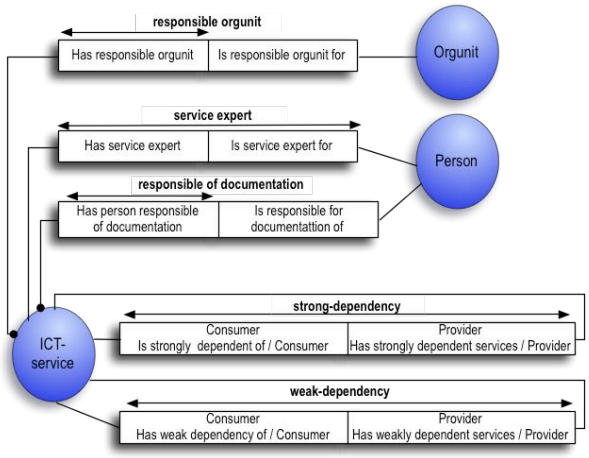
-Based on what the different groups need to document for different classes of services

-Inheritance and specialization

-The system is gradually introduced for new service types

-We are using agile development methods, so it's important that the ontology is easy to change

Ontology overview



18

-Using abstraction, to hide details for the main entities

-The arrows are Uniqueness constraints, probably the hardest part for people new to ORM

-The black dots are mandatory role constraints

-Can use cardinality constraints too

-Oganisational Unit is (usually) a group responsible for a service class

-for instance the unix group is responsible for system administration and documentation of unix servers

-At the core of the knowledge base are the dependencies between different instances of services, modelled as Topic Map associations.

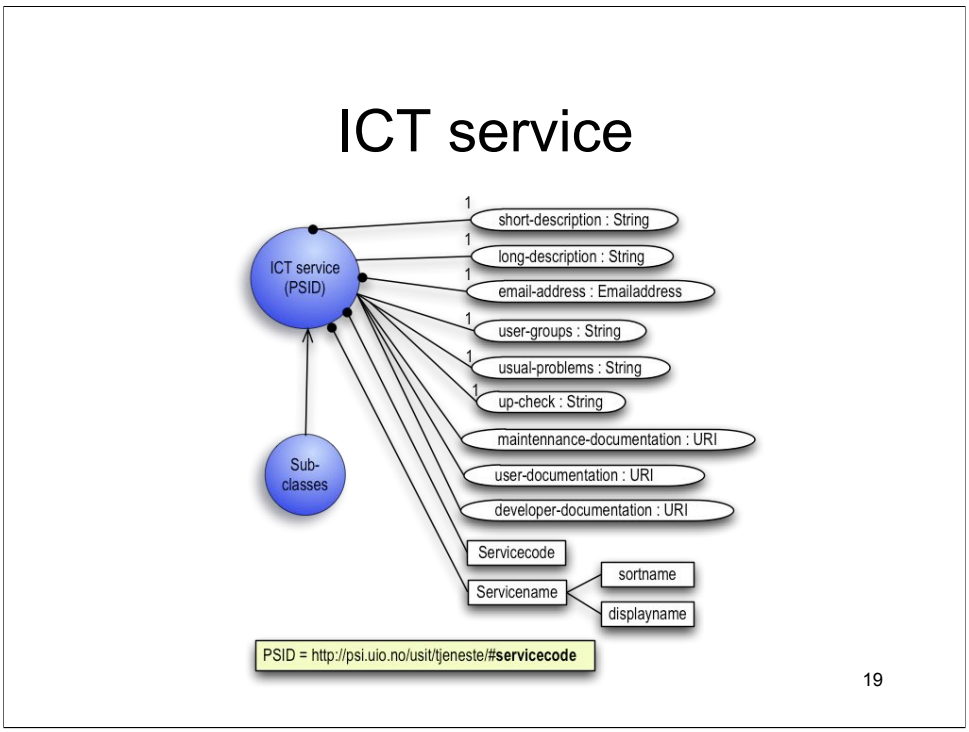
-We use the terms 'strong dependency' and 'weak dependency'.

-If a service X has a strong dependency on service Y and service Y is not functioning, then service X will not function.

-Sauron is for example a Solaris-server that has a strong dependency on uio-gateway01, which is a gateway.

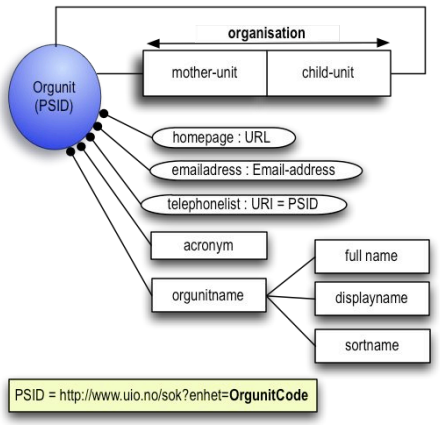
-One of the key features for the control centre is dynamic browsing of dependency graphs for the different services.

-This is a key feature when planning maintenance of a service, or in getting an overview of the side effects if a service is failing.



This is translated rather direct from Norwegian, so some topic type names may look a bit strange.

OrgUnit


organisation
mother-unit
child-unit
Orgunit
(PSID)
homepage : URL
emailadress : Email-address
telephonelist : URI = PSID
acronym
orgunitname
full name
displayname
sortname
PSID = http://www.uio.no/sok?enhet=OrgunitCode

Evaluation criteria

- This short presentation will focus on
  - Expressibility
  - Clarity
  - Semantic stability
  - Abstraction mechanisms

21

Will focus on the four most interesting criteria

Expressibility

- ORM has a richer set of restrictions, but both notations opens up for extensions
  - For example the Object Constraint Language - an OMG standard for specifying invariants, preconditions, postconditions and other kinds of constraints on UML models
- It is planned that TMCL can use TMQL queries as restrictions
- With graphical notations we will have to use the 80/20 principle

Clarity

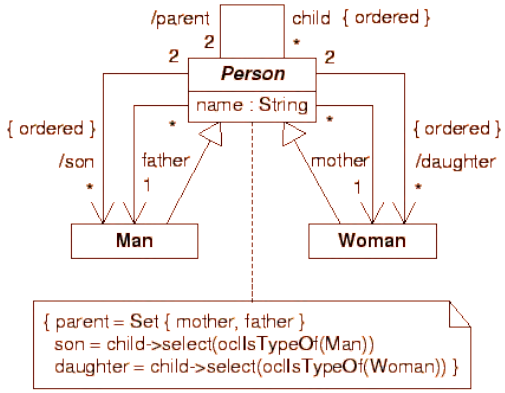
- Clarity is quite subjective, and depending on what people are used to
- There are a few things I think help to make ORM clearer than UML
  - The focus on relationship roles
    - I suggests an extension to ORM that allow contextual role names according to scope
  - Mandatory role constraint
  - Uniqueness constraints

Clarity 2

- There's a few things I think help to make UML clearer
  - The compact notation
    - In ORM it's possible to generate a more compact view of the model, even an UML-like table view
  - More people know the notation

24

Family ontology as UML



25

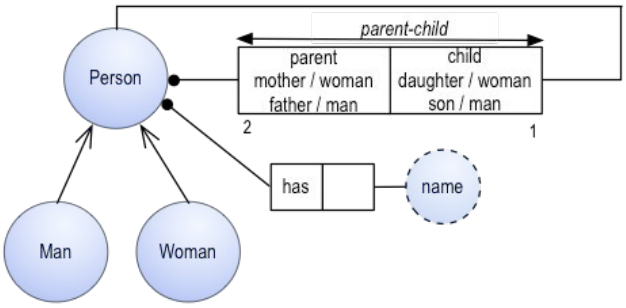
If we open up for using a separate textual constraint language and too many extensions, the model might be correct, and it might be possible to automatically generate a schema, but as a tool for communication the model will be less clear.

-Ref: Networked Knowledge Representation and Exchange using UML and RDF

-Stephen Cranefield

-Journal of Digital Information

Family ontology as ORM



26

-About The use of scopes for role-name

-The role-name of parent is mother in the context where the role is played by a woman (subclass of person)

-The role-name of parent is father in the context where the role is played by a man (subclass of person)

-The role-name of child is daughter in the context where the role is played by a woman (subclass of person)

-The role-name of child is son in the context where the role is played by a man (subclass of person)

-Mandatory role contraint: A person always has parents (But a person does not always have children)

Semantic stability

- When modeling with UML It's not unusual to have to change an attribute into a class
  - Can be a quite big change with some side effects
- ORM is an attribute free model
  - A ORM model is essentially a connected network of object types and relationship types
  - We only have to change the object from a value type to an entity type
  - It is more semantically stable

27

-I think this would be important in the case of using reification in a topic map

Abstraction mechanisms

- Allows unwanted details to be removed from presentation
- UML models tend to more compact than ORM models
- ORM use abstraction mechanisms
  - Split model in several parts
  - Hide unnecessary details
  - Show information in different views
- Abstraction mechanisms improve clarity
  - Makes the model easier to understand

28

See examples in the case ontology

Conclusion

- It seems like ORM is usable for Topic Maps modeling
  - ORM is more semantically stable if changes occur
  - The notation has more standard constraints
  - You have to make up your own mind about clarity
  - UML is more compact
    - But It's possible to generate the same view in ORM
  - UML has more widespread use and tool support
  - The suggested syntax for scope on association roles can give ORM good expressibility and clarity
    - It remains to see if it corresponds well with the ORM metamodel, and can be made into an extension

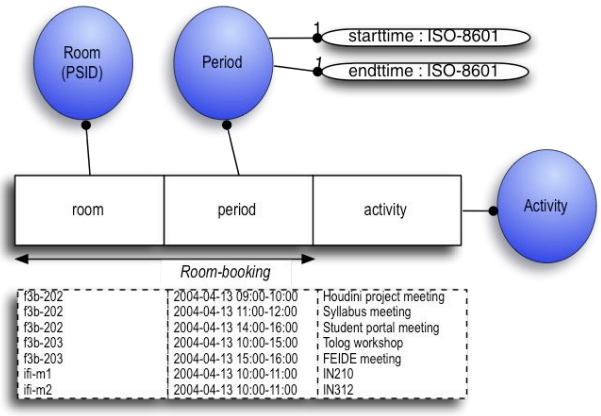
29

Questions?

30

Question from Steve Pepper about modeling n-ary associations with ORM. - See exampleof this in the first extra slide.

Example of 3-ary association



31

The table under the roles is an example of the standard way of showing examples, instances of an association, with ORM.

The application

- Harvests information from authoritative sources, and use Topic Maps to merge it all together.
- A topic map browser/editor
  - Standard forms based web editing
  - Based on the Ontopia Knowledge Suite (www.ontopia.net)
  - Spring J2EE application framework (www.springframework.org)
  - Freemarker template engine (freemarker.sourceforge.net)

32

-Much of the information needed already exists in different information systems, and much of our challenge has been systems integration.

-Our strategy is to harvest as much information as possible from authoritative sources, and use Topic Maps as semantic glue to merge it all together.

-Scripts are made for doing batch conversion of authoritative data to LTM (Linear Topic Map Notation) and saving the documents in a CVS repository.

-Published Subjects and Merging are inherent parts of the Topic Map standard, and make Topic Maps an ideal solution when trying to achieve semantic integration of decentralized and heterogeneous information sources.

-Are D. Gulbrandsens PSID: http://www.uio.no/sok?person=areg