DOMAIN ONTOLOGY DESIGN

Boriana Deliiska, University of Forestry, Sofia, Bulgaria

Abstract**:** *With the advent of the Web and data interchange intensive applications, ontologies are becoming a key issue. They are used for ensuring software interoperability and data sharing. Moreover, the ontologies are essential element of knowledge bases. An ontology basically provides a controlled vocabulary whose terms are precisely defined by lexicographic units such as dictionary entries or encyclopedia entries.*

In the article classification and structure of ontologies is given as well as strategies of domain ontology design. An overview of languages and tools for formal ontology description and management is made. The intention of this research is to design ontology of forestry on the base of multilingual dictionary and the respective database in this area.

Keywords*: domain ontology, knowledge base, ontology modelling tools, forestry*

1. Introduction

The most widely cited definition of ontology is that of Gruber (1993) who said that ontology was “a formal, explicit specification of a shared conceptualization, used to help programs and humans share knowledge.”

In the other hand [12], the ontology is defined as O = <X,R,F>, where X — end set of entities about a domain, R — end set of relations between entities, and F — end set of interpretation functions.

The ontology is based on raw text or on controlled vocabulary (also known as scheme or dictionary) with a limited set of precisely and carefully defined terms [11].

2. Ontology classification

Being applied in the context of developing knowledge-based systems, ontologies were classified into:

- *task ontologies* (including method ontologies) – provide a systematic vocabulary of the terms with definitions of the relevant concepts and relations used to solve problems associated with particular tasks (either domain independent or domain dependent);
- *application ontologies* – contain the necessary knowledge for modeling a particular application;
- *domain ontologies* – reusable in a given domain;
- *knowledge representation ontologies* (or representation ontologies) – capture the representation primitives used to formalize knowledge in knowledge representation paradigms;
- *top-level ontologies* (upper-level ontologies or foundation ontologies) **–** provide general notions under which with all the terms in existing ontologies should be linked to. At the moment the main problem is that there does not exist a unified top-level ontology in the community that cover the different aspects that are relevant when modeling knowledge-based systems [10];
- *core ontologies* (generic ontologies or meta-ontologies) – a very basic and minimal ontology consisting only of the minimal concepts required to understand the other concepts;
- *general (common)* ontologies include vocabulary related to things, events, time, space, causality, behavior, function, etc.

3. Ontology structure

An ontology representation includes two groups of elements [10]:

3.1. Basic elements:

- **classes** (**concepts** or nouns). A class is linked to superclass(es) and has **subclasses;**
- **properties** of each class (sometimes called **roles** or **slots).** Slots describe properties of classes and instances;
- **restrictions** on slots (sometimes called **role restrictions** or **facets**);
- **instances** of classes.

3.2. Advanced elements:

- **relations** (or verbs) between classes which represent a type of interaction between concepts of the domain (for example *is-a*, *subclass-of* and *connected-to*, *has-part*, etc). There are some special case of relations named **functions**;
- **axioms** – the definitions of terms in the ontology and constraints on their interpretation;
- **rules** – behaviour prescriptions of events.

An ontology together with a set of individual constitutes a **knowledge base**. There is small difference between the ontology ends and the knowledge base begins.

4. Methods and principles of domain ontology design

There is no one way or methodology for developing ontologies [6]. The main design criteria [1] are:

- **clarity and objectivity** – an ontology should provide the meaning of defined terms by objective definitions along with a natural language documentation;
- **completeness** – a definition expressed in terms of necessary and sufficient conditions is preferred over a partial definition (defined only through necessary or sufficient conditions);
- **coherence** – to permit inferences that are consistent with the definitions;
- **minimization** of the semantic distance between sibling concepts – similar concepts are usually grouped and represented as subclasses of one class and should be defined using the same set of primitives.

There are several main activities in ontology building [2]:

4.1. Determination the domain and scope of the ontology.

4.2. Consideration reusing existing ontologies.

4.3. Enumeration important terms in the ontology.

4.4. Definition the classes and arranging the classes in a taxonomic (subclass–superclass) hierarchy.

4.5. Definition the properties of classes–slots.

4.6. Definition the facets of the slots. Here slot cardinality, slot-value type, domain and range of a slot are defined.

4.7. Instance creation.

4.8. Knowledge base building by defining individual instances of the classes, filling in specific property value information and additional property restrictions.

5. Ontology development tools

In [7] a survey of ontology development tools is given. They are two main types: ontology modeling languages and software platforms for ontology building and management.

5.1. Languages for ontology modelling

Existing languages for ontology representation are [1, 10]: Prolog, XML, OKBC (Open Knowledge Base Connectivity), KIF (Knowledge Interchange format), LOOM, DAML+OIL (Ontology Inference Layer+Darpa Agent Markup Language), OWL (Web Ontology Language), XOL (XML-based Ontology exchange Language), RDF, RDF Schema, SHOE (Simple HTML Ontology Extensions), CKML (Conceptual Knowledge Markup Language) etc.

5.2. Platforms for ontology management

In the last years, a high number of environments for creating, editing, valuation and managing of ontologies written in the various languages have appeared. Survey of ontology editing tools summarizes more than 50 software tools. The most popular are [7]: Protégé 2000; OWL Ontology Validator; IBM Ontology Management; RACER System; OilEd; Chimaera Ontology Environment; Open; F-OWL; Ontolingua; Pellet OWL Reasoner; pOWL; WordMap Taxonomy Management; EON2002; Apollo; DOE; WebODE; Dumpont; Network Inference etc.

6. Domain ontology of forestry

There are some publications about terminology [8] and ontology of forestry. The most extensive is NEFIS project [10] which continues the activities of EFIS (European Forest Information System) and contributes to the activities coordinated by IUFRO in building a Global Forest Information System (GFIS). The metadata standard of this project is a part of Dublin Core adapted to forest information.

The intention of the current research is to project ontology of forestry on the base of multilingual dictionary and the respective database in this area [4]. The dictionary database contains more of 17 000 entries structured in main classes and subclasses. Moreover, in the database well-described definitions and synonym references exist. The leading language is English and each entry has its German, French and Russian equivalents. The Latin name of forest species is included also.

With these preconditions the building of current domain ontology of forestry includes following tasks:

- addition of missing entry definitions;
- design of taxonomy of forestry, including identifying concepts;
- defining of relation set;
- development of appropriate structure of the ontology database;

A part of forestry taxonomy is given on fig. 1. The main classes are in bold. Subclasses, slots and instances are in the consecutive levels of hierarchy.

The relation set include the following relations: *is-a*, *is-part-of*, *is-kind-of*, *is-subclass-of*, *is-located-in*, *is-instance-of*, *is-member-of*, *join-with*, *is-synonym-of* etc.

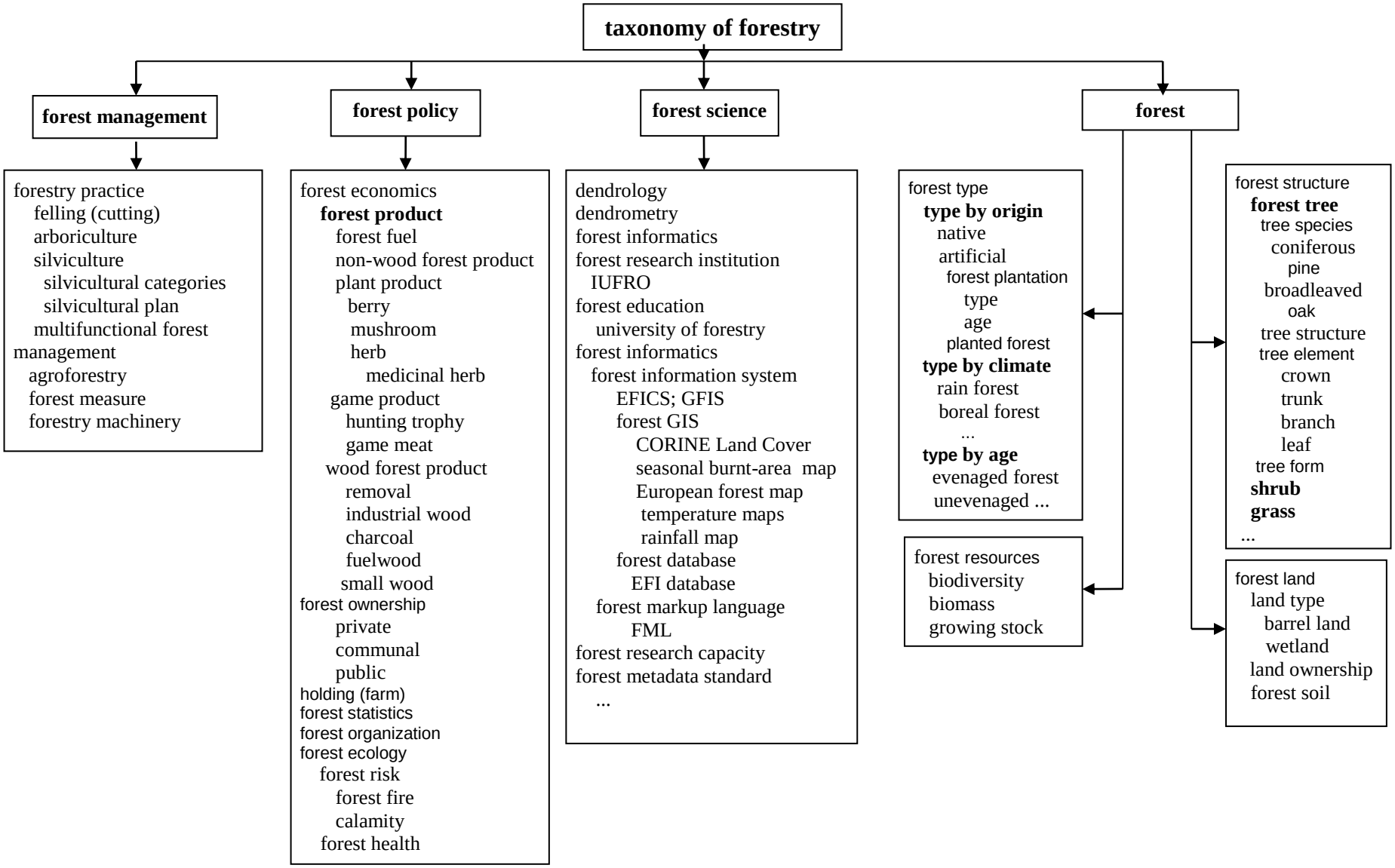
taxonomy of forestry
forest management
forestry practice
felling (cutting)
arboriculture
silviculture
silvicultural categories
silvicultural plan
multifunctional forest management
agroforestry
forest measure
forestry machinery
forest policy
forest economics
forest product
forest fuel
non-wood forest product
plant product
berry
mushroom
herb
medicinal herb
game product
hunting trophy
game meat
wood forest product
removal
industrial wood
charcoal
fuelwood
small wood
forest ownership
private
communal
public
holding (farm)
forest statistics
forest organization
forest ecology
forest risk
forest fire
calamity
forest health
forest science
dendrology
dendrometry
forest informatics
forest research institution
IUFRO
forest education
university of forestry
forest informatics
forest information system
EFICS; GFIS
forest GIS
CORINE Land Cover
seasonal burnt-area map
European forest map
temperature maps
rainfall map
forest database
EFI database
forest markup language
FML
forest research capacity
forest metadata standard
...
forest
forest type
type by origin
native
artificial
forest plantation
type
age
planted forest
type by climate
rain forest
boreal forest
...
type by age
evenaged forest
unevenaged ...
forest resources
biodiversity
biomass
growing stock
forest structure
forest tree
tree species
coniferous
pine
broadleaved
oak
tree structure
tree element
crown
trunk
branch
leaf
tree form
shrub
grass
...
forest land
land type
barrel land
wetland
land ownership
forest soil

fig.1

The formal description in XML of forest taxonomy and relations between dictionary entries looks like this:

```
<domain><entry name>forestry</entry name>
    <class><entry name>forest science</entry name>
        <subclass><language>English
            <entry name>dendrology</entry name>
              <definition> A branch of forestry which...</definition></language>
            <language>French<entry name>
            dendrologie</entry name></language>
            ...
        </subclass>
        ...
    </class> ...
</domain>
```

The structure of the dictionary database is given on fig.2.

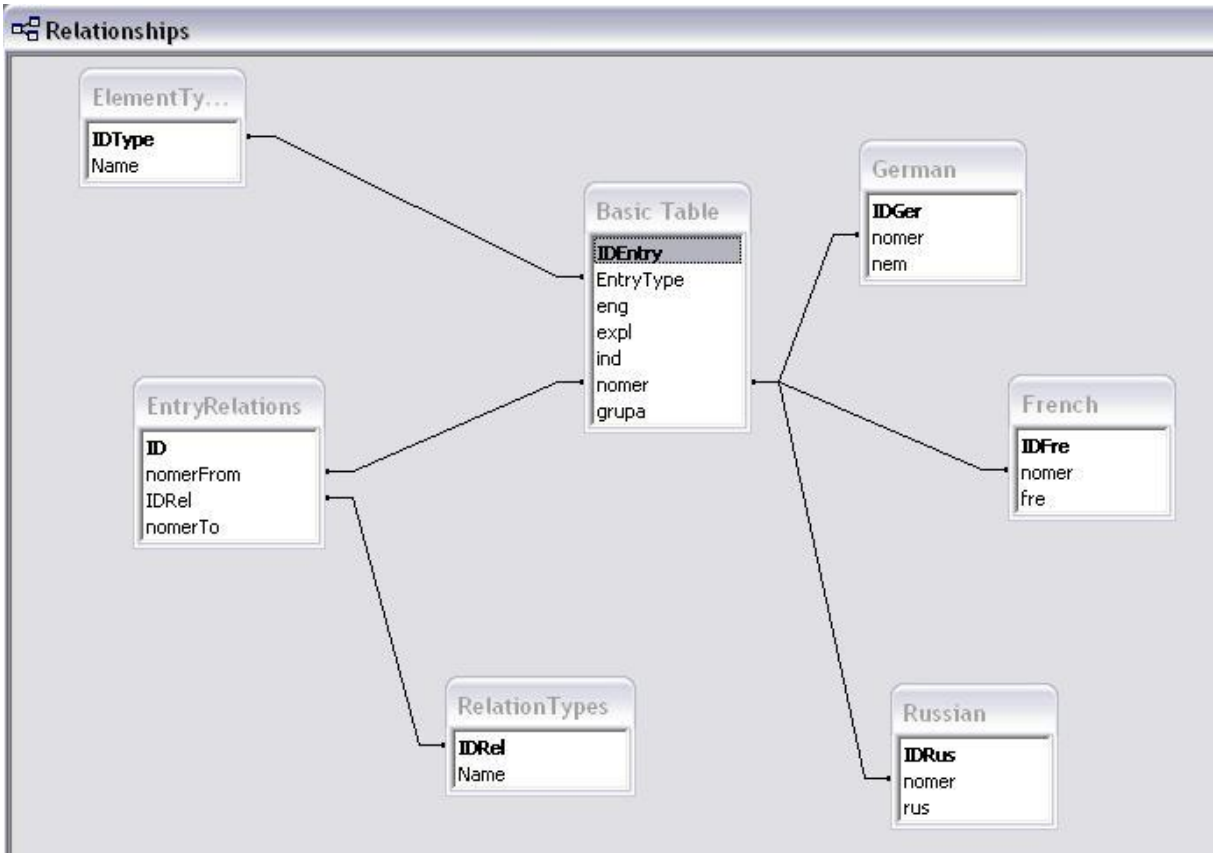

fig.2

7. Conclusion

The current domain ontology of forestry is created on the respective MS Access database and well-structured dictionary of forestry. It is used in web-based tool for modelling of sustainable development of forest plantation [3].

This method can be used for building of another domain ontologies.

The future works on the domain ontology includes the development of user instruments for information retrieval. Moreover, publication in Internet of online ontology is provided for the purpose of e-learning in area of forestry.

REFERENCES

[1] Benjamins V. R., Perez A. G., Knowledge-System Technology: Ontologies and Problem-Solving Methods, http://hcs.science.uva.nl/usr/richard/pdf/kais.pdf.

[2] Blázquez M., Fernández M., García-Pinar J.M., Gómez-Pérez A., Building Ontologies at the Knowledge Level using the Ontology Design Environment, http://ksi.cpsc.ucalgary.ca/KAW/KAW98/blazquez/

[3] Danchev A., N.Hadjidimitrov, B.Deliiska, Information Modeling of the Forestry Sustainable Development in Bulgaria, Third GIS in Turkey Conference, 6 – 9 October, 2004, Fatih University, Istanbul, Turkey

[4] Delijska B., P. P. Manoilov, Elsevier Dictionary of Forestry, Amsterdam, Elsevier, 2004

[5] Hirst G., Ontology and the Lexicon, info.tuwien.ac.at/cf2005/caad/posters/cf2005_227_poster.pdf

[6] Noy N. F., McGuinness D. L., Ontology Development 101: A Guide to Creating Your First Ontology, http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html

[7] Ontology-based information exchange for knowledge management and electronic commerce, IST-2000-29243, 31 May 2002

[8] Proceedings of third expert meeting on harmonizing forest-related definitions for use by various stakeholders, FAO, Rome, 17-19.01.2005

[9] Reenrolls K., Lee F., Ibrahim M., Fedorec F., NEFIS Technology Review, http://cms1.gre.ac.uk/research/cassm/NEFIS/Papers/Technology%20Review_10.pdf, oct. 2003

[10] Ribière M., Charlton P., Ontology Overview, *Motorola Labs*, http://www.fipa.org/docs/input/f-in-00045/f-in-00045.pdf.

[11] Staab S., Studer R. (ed.) On-To-Knowledge Methodology, Handbook on Ontologies Springer, 2003

[12] *Wikipedia*, http://ru.wikipedia.org/wiki.html