Ontological Engineering Process For Computerized System Subject Areas

KHAROUBI NAOUFEL GERASIMOV I.V DMITREVICH G.D

Saint Petersburg Electrotechnical University "LETI" Russia

Summary
This article represents an attempt to expose the benefits of using computerized system subject area ontologies. It also provides the major design phases of such ontologies.

Key words:
Ontological engineering, Ontology, knowledge base, semantic network

1. Introduction

Ontological engineering consists in design and development of ontology (specific type of knowledge base). It requires professional skills in knowledge engineering from developers – from knowledge extraction to structuring and formalization. Ontological engineering combines two major technologies of large system engineering – object-oriented analysis and structural analysis.

Ontology is designed as tree or network consisting of concepts and relationships between such concepts. These relationships can be of different types, for example, "is part of", "is a", "has property" etc. Concepts and relationship have universal character for a certain group of subject area concepts. A specific concept from such group can be selected for "filling in" its ontology and assigning specific attribute values. Ontology is a model of subject area using all available means of representation of knowledge relevant to such area.

Ontological structure, i.e. ontological engineering is the most powerful cognitive tool visualizing the structures of individual and corporate knowledge in all areas.

Ontological approach to computerized representation of information has already offered quite accomplished application in various areas. Theoretical research developments dedicated to semantically correct document search (including data selection by network agents) and computerized formation of ontologies based on

available resources is also very important. Additionally, based on this approach specialists perform the following processes:

• Data mining – identification of relevant dependences and trends as a result of intellectual analysis of data banks and databases.

• Text mining – a complex of techniques for knowledge extraction from documents in natural languages and their representation in a user-required format.

• Content analysis – a method for objective qualitative analysis of data resource content.

• etc.

Being a sui generis general conceptual dictionary, ontologies significantly facilitate mutual understanding between human individuals in data sharing. Besides, data representation in ontological format, in addition to computerized data processing, provides the most illustrative and accurate way to formulate knowledge in the selected area. The latter fact has an enormous value for educational purposes. Education, being a directed process of knowledge communication, can serve as a rather natural area for application of ontologies. Specifically, the development of efficient computerized education system to a large extent depends on the results of data structuring and computerized representation.

Importantly, computerized systems always require communication representation of various types of knowledge (regulatory and reference data, guidances, conceptual model of subject area, mathematical models, mathematical methods, standard solutions, programs, limitations, approved designed solutions etc.).

Ontology description can be made on numerous level, and each such level can be used as a skeleton for the next level (Fig.1).

Ontological engineering process for computerized system subject areas (PCSSA)

In the first place certain fundamental rules for ontology development must be formulated. These rules may seem rather categorical. However, in most cases they will be useful for design solutions.

- *The only correct method of subject area modeling does not exist – there are always viable alternatives. The best solution almost always depends on expected application and extensions.*
- *Ontological development is always iterative.*
- *Ontological concepts must be close to objects (physical or logical) and relationships in the subject area in consideration. Most probably, these will be nouns (objects) or verbs (relationships) in the sentences describing your subject area.*

Manuscript received April 5, 2009
Manuscript revised April 20, 2009

This means that the awareness of ontology application and its degree of detail or generalization will determine many modeling solutions. Among several viable alternatives we need to identify the one which will be the most efficient for task solution, as well as the most illustrative, upgradable and easiest in maintenance. We also need to remember that ontology is a model of the real world, and ontological concepts must reflect this reality. After determining the initial version of the ontology we can evaluate and adjust it by using it in applications or problem solution methods and/or by discussing it with the subject area experts. As a result, a revision of the initial ontology will be required most probably. This iterative design process will probably continue through the whole life cycle of the ontology.

In this article we focus on general aspects which must be taken in consideration and offer one of the possible methods for PCSSA development.

The logical and linguistic nature of the offered algorithm for ontological engineering of automated system subject area consists of several steps:

1) Task identification. Data engineer must identify the scope of issues to be supported by the ontology and the types of facts which will be accessible specifically to each individual task occurrence.
2) Collection of data related to the subject area. Data engineer either can be already an expert in the subject area in consideration, need to communicate with real experts for identification of all their knowledge as well as exclusion of contradictions and inaccuracies.
3) Vocabulary definition, i.e. identification of concepts – basic notions of the specific subject area.
4) Construing important concepts within the verbal model.
5) Determination of relations and interactions between the basic concepts.
6) Formation taxonomic hierarchy
7) Registration of general knowledge about the subject area in a formal format. Possibility of re-using the existing ontologies and standardized dictionaries.
8) Ontological coordination test.
9) Description of the specific task occurrence.
10) Query transmission to logical output procedure and answer receipt.

Debugging ontologies.

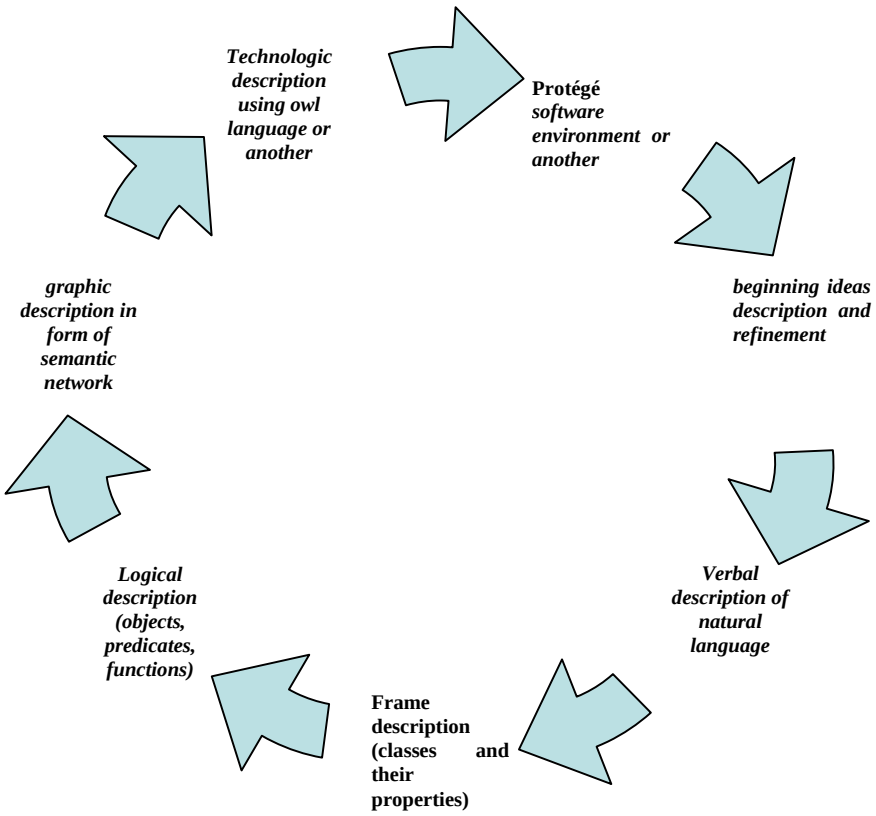
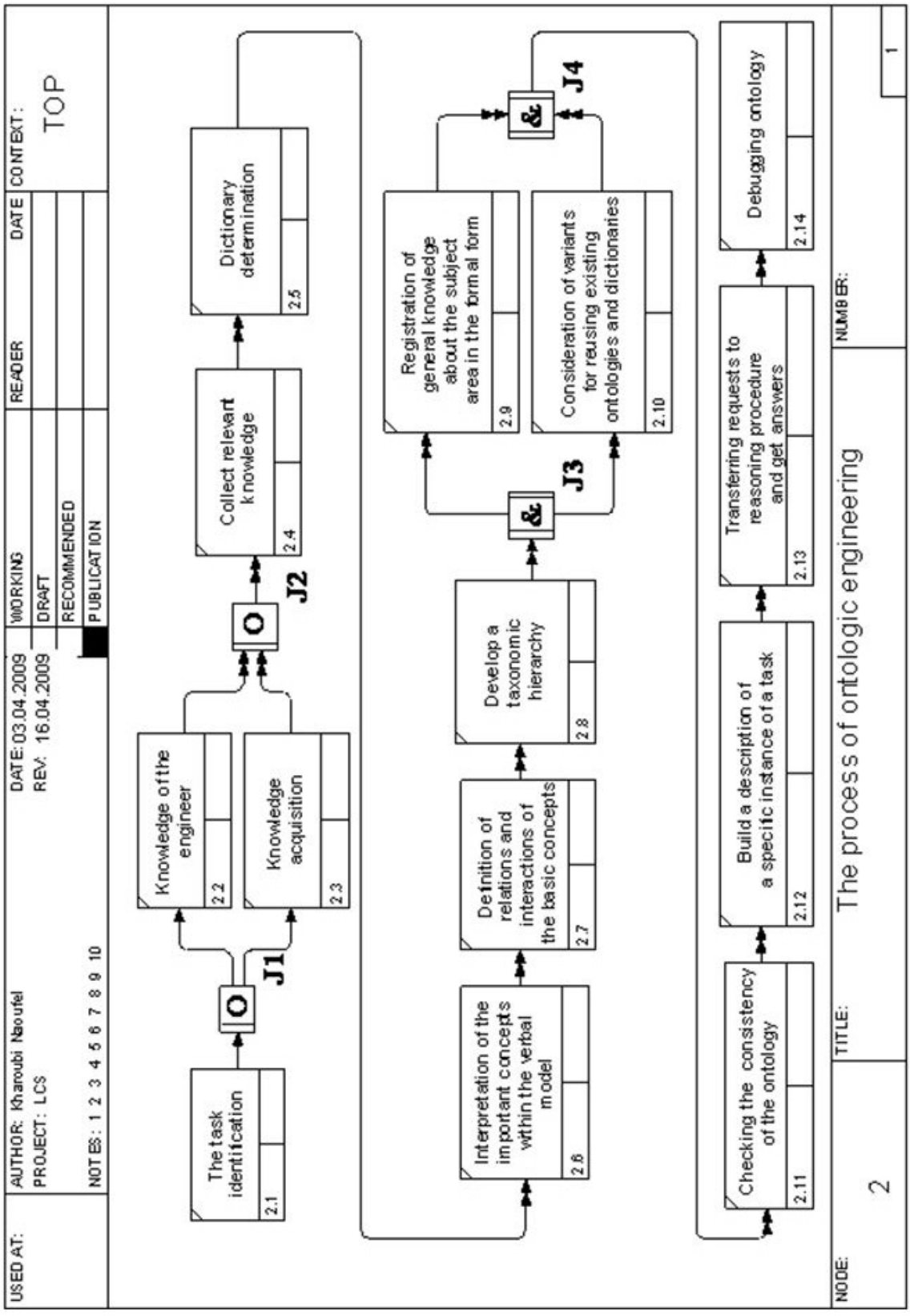


Fig. 1 Ontology description and work levels


USED AT:
AUTHOR: Kharoubi Naoufel
PROJECT: LCS
DATE: 03.04.2009
REV: 16.04.2009
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
TOP
NOTES: 1 2 3 4 5 6 7 8 9 10
NODE: 2
TITLE: The process of ontologic engineering
NUMBER: 1
The task identification 2.1
Knowledge of the engineer 2.2
Knowledge acquisition 2.3
Collect relevant knowledge 2.4
Dictionary determination 2.5
Interpretation of the important concepts within the verbal model 2.6
Definition of relations and interactions of the basic concepts 2.7
Develop a taxonomic hierarchy 2.8
Registration of general knowledge about the subject area in the formal form 2.9
Consideration of variants for reusing existing ontologies and dictionaries 2.10
Checking the consistency of the ontology 2.11
Build a description of a specific instance of a task 2.12
Transferring requests to reasoning procedure and get answers 2.13
Debugging ontology 2.14
O
J1
J2
&
J3
&
J4

Conclusion

This article represents an attempt to expose the benefits of using ontologies in computerized system subject areas. It lists the steps of designing such ontologies. Naturally, the implementation of this approach in the existing computerized systems will require significant revisions. However, the new benefits of computerized systems are cost effective. Moreover, the whole or part of ontology of the same subject area of computerized systems can be re-used in the ontological design of another subject area of computerized systems, which reduces development and implementation costs.

References

[1] Gerasimov I. V MAIGA A.I Lozova L.N ontology engeenering .Tools and specifications ontology modeling. Manual. St. Petersburg Publishing SPb «LETI» 2009

[2] Russell Stewart, Norwig Peter. P24 Machine intellect: modern approach, 2nd ed. : Translation from English. – M. : Williams Publishers, 2006. – 1408 p. : il. – Paral. Tit. Engl.

[3] Gavrilova T.A., Khoroshevsky V.F. Intellectual system databases. College textbook. – SPb: “Piter”, 2000.

[4] Giarratano Joseph, Ryley Gary. D40 Expert Systetms: Principles and programming, 4th ed.: Translation from English. – M. : Williams Publishers, 2007. – 1152 p. : il. – Paral.tit. Engl.

[5] Luger J. Machine intellect: strategies and methods for complicated problem solution, 4th ed.: Translation from English. – M. Williams Publishers, 2003. – 864 p.

[6] The Protégé Ontology Editor and Knowledge Acquisition System [HTML] (http://protege.stanford.edu/)

Kharoubi Naoufel was born in 1979 in Tunisia . Finished Masters degree in computer science , from Saint Petersburg Electrotechnical University "LETI" ,Russia in 2007, then joined PhD program in computer science in LETI.
His research interest includes automated systems, artificial intelligence, ontological engineering.

Gerasimov.Igor Vladimirovich Professor, member of the Academy of Sciences of Higher Professional Education of the Russian Federation, Honored Worker in Higher Education of Russia. Dean Faculty of Computer Technology and Information (FKTI), Professor and Department Head of computer-aided design systems (CAD). Electrical Engineer; well-known specialist in the field of computational physics processes, means of functional transformation of the form in different signals and processes, topological methods of structural synthesis of component-based framework for automation and computer engeenering, information technology design.
http://www.etu.ru/faculty/fkti.htm
Author of numerous books, some of the most recent of which are:
1. Gerasimov I. V MAIGA A.I Lozova L.N ontology engeenering .Tools and specifications ontology modeling. Manual. St. Petersburg Publishing SPb «LETI» 2009
2. Gerasimov I.V, Kalmychkov V.A, Lozova L.N complementary modeling in CAD environments: Virtualization quantum of information. SPb «LETI», 2007, 208p.

Dmitrevich Gennadi Daniilovich Professor, Honored Worker of Higher Professional Education of Russian Federation , Professor University "LETI" Faculty of Computer Technology and Information (FKTI), Interest Area - optimization in CAD