Methodological Considerations in Building Business Ontologies for Decision Support

Necula, Sabina-Cristiana

Alexandru Ioan Cuza University of Iasi

October 2012

Methodological Considerations in Building Business Ontologies for Decision Support

SABINA-CRISTIANA NECULA
Department of Research
Alexandru Ioan Cuza University of Iasi
Carol I Blvd, no. 22, Iasi, 700505
ROMANIA
sabina.mihalache@gmail.com http://wwww.knowledgedecisionmaking.ro

Abstract: - This paper tries to identify some methodological consideration in building business ontologies for decision support. We treat the problems that business domain poses for knowledge engineers. We also discuss the main semantic web technologies involved in developing ontologies. We consider also the problems that decision support presents for ontologies. From these considerations we identify and discuss some methodological aspects that building ontologies for business decision support must take into considerations.

Key-Words: - decision support, semantic web, ontology, business, methodology, RDF

1 Introduction

Modern decision support systems (DSS) provide their users with a broad range of capabilities. Current DSS facilitate a wide variety of decision tasks including information gathering and analysis, model building, sensitivity analysis, collaboration, alternative evaluation, and decision implementation. Often, DSS are built and used for ad hoc analyses, but increasingly, decision support is integrated into business processes and information systems. The global Internet and the World Wide Web are now the primarily enabling technologies for delivering computerized decision support. Due to the growing interest in the Web, there are many on-going efforts to develop and implement Web-based DSS in various areas, such as health care, private companies, government, and education. [4]

A number of articles [5, 6, 7, 8] have also addressed issues associated with Web-enabling DSS.

Web-based technologies are having a major impact on design, development, and implementation processes for all types of decision support systems (DSS) [9, 3].

Semantic web efforts can be characterized as follows: [1]

1) Closed domains: in contrast to Web 2.0, most Semantic Web applications have assumed closed domains of manageable size, such as the proteins domain, digital libraries and corporate intranets.

2) Complex and comprehensive modeling. The Semantic Web community has aimed to mdel as much of the underlying complexity of a domain as possible while covering the domain comprehensively. This is reflected in RDF and RDF(S) and in the standardized Web Ontology Language OWL. In addition, academic research has contributed methodologies for ontology engineering, evolution, debugging, and modulariation, aiming for a thorough understanding of the complexity of common ontology language.

3) Design for knowledge engineers. The complexity of the modeling and modelling languages has meant that trained engineers are required for domain modelling and are implicitly assumed to be involved in the design and maintainance of the ontology.

4) Sophisticated reasoning. Due to the complex domain modeling, there has been a need for sophisticated inferencing methods and scalable reasoners. This has led to the development of increasingly scalable reasoning solutions.

5) Complex specifications and heavy-weight tools. The documentation and specifications of Semantic Web languages are well-known to be complex and often inscrutable to the average Web developer. Similarly, in comparison to Web 2.0, the vast majority of tools for the Semantic Web are heavy-weight to the point of being unwieldy. There are some improved modelling tools like Protege or Swoop, but those tend to focus on the knowledge engineer as opposed to a Web developer.

There are several techniques to design and develop ontologies [9,11]. Each of them encompasses a number of fundamental rules and specifications to guide developers in the design of proper ontologies.

The term mashup involves an easy, fast integration, and frequently used open APIs and data sources to produce results [10].

2 Problem Formulation

We try to present the idea of using ontologies in achieving management objectives.

Considering business decisions in a broad sense namely that of financial-accounting decision transforms the information model belonging to accounting into a management model: namely, more resources management models oriented to organization. The objective of management is the maximization of business value by optimal use of resources and by monitoring the indicators of effectiveness and efficiency.

In a restricted sense, the necessary information needed to make business decisions create a metamodel to organize financial information, which organized and structured optimally provide adequate information to responsible with resource management.

Thus, management models overlap with models of economic and financial analysis with models of management accounting, financial and economic models. Metamodel is interwoven with organizational management models and does not require tracking of indicators but setting constraints recognition, asset classification and valuation for financial accounting attachment values.

Information necessary to make decisions comes from heterogeneous sources: the organization and outside environment. More information reduces uncertainty and organization reduce uncertainty.

Models and business rules of the organization and organizing processes contain control specification decisions.

Thus, we identified two research questions:

1) What are the main methodological guidelines for building ontologies?
2) What are the restrictions that business fields and decision support imposes.

We try to treat all these questions into the following paragraphs.

3 Problem Solution

Ontology is the study of the existence of all sorts of abstract entities from a specific domain. The two sources of ontological categories are observation and reasoning. The result of this study called ontology is a catalog of the types of things that are assumed to exist in the D domain from the perspective of a person who uses the language L to talk about D.

The observation provides real world knowledge and the reasoning makes sense to the observation, providing a framework called metaphysical abstractions. Philosophers are building ontologies from top to bottom. Programmers tend to build from bottom to up, managing to build applications for so-called “small” world. The main advantage of a limited area is the ease of analysis, design and implementation. The disadvantage is the impossibility of the distribution and use of data by other application programs.

In terms of software, there are computer-based applications which are using two types of ontologies. The first type is that of ontologies for knowledge-based systems, ie a system of financial diagnosis. These ontologies are characterized by a relatively small number of concepts, but connected by a large number and variety of relationships. The concepts are grouped into complex conceptual schemes or scenarios. For each concept there can be one or more customizations. Unlike the first type of ontologies, the second type is the lexical ontology which includes a large number of concepts, connected by a small number of types of relationships. Concepts are represented, for example, in WordNet, through a set of synonymous words. Such ontologies are used in human language processing systems.

A knowledge base is considered to be composed from a facts base and from a rules base. The rule base contains knowledge represented through the IF-THEN-ELSE metamodel and it results in a file. The combined database file and the file containing the rules can be accessed by an inference engine that has implemented a control strategy in order to infer a solution. It is important to note that the database can be realized from a single table and not from many tables. Equally, the defined rules are stored in a single file.

Redundancy issues do not concern the knowledge bases. In a database a unique registration data should be provided from both physical as semantically point of view. A database as a data source for a knowledge base uses data structures permitted by the representation language of the developed application. Because implementing knowledge base is not possible with a database management system, the interface between data and the knowledge base will be supported by ODBC or

JDBC driver. Thus, data redundancy remains a problem to be solved by database designers.

3.1. Problems raised by decision support

Realizing a distinction between a database and a knowledge base refers to data organization. Organization must refer this time the concepts and not to data by applying a general knowledge or existing rules of concepts.

Organizing knowledge in an ontology representation refers to a class-property-instance (individual) formalism, which corresponds to a database (class-table name, property-the name of the field, instance-record) and contains existing rules of classes and properties.

The most difficult problem in development of any knowledge base consists in expanding the knowledge base through the acquisition of knowledge or through learning. In artificial intelligence there is a distinguish between acquiring which is the introduction of new pieces of knowledge into the knowledge base and learning which is the deduction of new knowledge from a range of existing knowledge through generalization.

The ontology is intended to promote communication between computers and between applications independent from a particular individual system's technologies or from information processing architecture.

3.2. Problems that business domain raises

If we take an example of operational management as the one of inventory, we can observe inconsistencies of the rational model in terms of context.

We took this example, just because the literature is considering that the development and use of inventories is a structural problem that can be addressed rationally. Inventory depends on consumer demand, inventory costs, quantity supplied, lot, time parameters. The theory is supported by Willson model and its derivation. Literature acknowledges that there are uncontrollable variables such as costs, demand during replenishment, quantity delivered and trying to control factors through estimates and assumptions. Since it is a management problem, not a decision, it may be presumed rational approach to the information provided by the models is useful for operating activities.

The information provided will be used in making management decisions concerning the identification of relevant costs and opportunity costs. Relevant costs and opportunity costs depend on the context; they cannot be expected (both in value and in nature) and is the information to be extracted from the decision maker in moments of decision.

Economic analysis, statistics, computer science, management accounting considers necessary to analyze historical data for forecasting. All economic indicators that follow a formula, a law of evolution can be predicted. However, the opportunity cost in decision making cannot be estimated, anticipated or projected. Determining cost behavior could be a problem to be solved by neural networks and data mining technology. Solving problems by using data mining tools has not been proven yet satisfactory: data sets are large and, as a result, researchers in the field today are looking to find algorithms for classification, segmentation, extraction of general classes by using formulas Math. Without exception, all are facing the same problem of scalability and, in fact, representative for the universe of model: more, all say that data must be linear or that the distances between the data must be consistent, when in fact it is known that in the real world it is impossible. Algorithms needed only to be used are the search by using artificial intelligence search phrase in the space state. The only solution that seems to be anticipating is estimating cost behavior by using the method of estimation and analysis of the life cycle cost of the product to achieve that part. Law development costs can be plotted and could be a real tool of analysis, control, planning and decision.

A second example concerned the classification of expenditure in the category of "capital" or "operating" expenses.

Management plays an important role when deciding whether a payment should be considered an asset or expense. It is capital expenditure classification payments or management charges and involves the principle of ensuring connectivity revenue expenditure. Because production presumes further benefits expense recording in one of two categories is a matter of professional judgment.

Computer could provide one solution to this: extract the response from the decision-maker. The decision maker chooses a category of expenditure and thus the record is made. Realization does not depend on accounting income and any mistake of judgment it will be difficult to prove. Respecting the principle of connectivity revenue to expenditure is part of the decision model for implementing knowledge.

The organization develops accounting policy for recognition as an asset or expense payment. Implementation of such rules is proposed to manage rule-based systems. We propose that such rules are implemented on concepts from an ontology rather than directly on the data, because there cannot be developed comprehensive policies for the classification of payments as an expense or asset. Thus it would be possible to trigger several rules regarding payment or asset concept and warning about the violation of a rule or a principle failure.

It appears useful to stress that the only solution is the segmentation of knowledge, classification of concepts with knowledge representation formalisms and not the execution of algorithms. But how data is represented according to the needs of specific applications is appropriate to amend the scheme of representation. And then, the only possible solution is the use of ontologies, which currently is still cumbersome and incipient.

3.3. Methodological considerations

Using ontologies aims sharing common meanings and not knowledge management, handling of general knowledge. Meaning perceived by an individual or a group is computer automated and applications and information systems share common ontology in order to enhance human-computer interaction.

People observe objects in the environment and create, over time, a map of relationships between concepts that describe objects and situations. In creating these relationships, human concepts used in activities are related to the manifestation of the learning experience. Once learned, the relationship will be able to create knowledge in tandem with the concepts. Like objects, relations will create knowledge in three steps: recognition, classification and evaluation. Concepts and relations are declarative knowledge or declarative statements with truth value.

So far, we have shown that people use environmental objects and relations (information message) to create knowledge in three stages which we detail below.

Recognition refers to the admission of an element into a system, subject to certain conditions (relations), previously established. From the point of view of modeling in computer language we associate with admission recognition of an item in a superclass objects. For example, the general opinion that an asset may be recognized if its cost can be measured reliably and it is possible that in the future would, use quantifiable economic benefits. The two conditions are necessary and sufficient.

Classification is the classification of items previously recognized and accepted in a particular category, depending on the characteristics of the item. From the point of view of modeling in computer language we associate with the employment classification of an item in a class of objects. For example, we can consider that an asset is an asset that will be used more than 1 year.

Evaluation is the process of customizing the elements recognized and classified for individualization attachment element and a purpose. From the point of view of modeling in computer language, we associate evaluation with setting the properties of an object. Each asset has a return during the production process involving. Return on assets is closely related to the costs arising from its use. The relationship between cost / performance and decisions on assets is a causal relationship. Formalize relationship allows decision-making models provide interactivity with the decision-maker involved. Causal relationship can be used in applications alert / warning or decision models using information from other chains inferential trigger alert.

Decision making is characterized by intensive use of knowledge. In our opinion, the decision making is a process of applying knowledge in conjunction with information identifying the manifestation of reasoning.

Intensive use of knowledge is a sequential process of deriving new knowledge able to guide the decision makers or organize existing models.

Knowledge processes of decision-making process are:

- The production of information;
- The evaluation of information;
- The certification / validation information.

The production process involves the application of information techniques, rules, methods based on existing information needs. It is a process of transforming data provided by environmental events and transactions in information, using knowledge.

Information evaluation process involves the application conditions prior to decision on information products to evaluate decision alternatives.

The certification / validation information should verify prior experiences or checking existing decision models of decision alternatives in order to choose the best alternative and decision.

In all processes of knowledge during decision making people recognize, classify and evaluate information.

4 Conclusion

Issues so far were been considered in order to identify opportunities for formalization of data, information and knowledge by existing computer technology. Use of information technology has resulted in information systems, intelligent systems and their hybrid forms.

Very important is observing practical implications of which are:

- Integration of cognitive knowledge creation process with existing technologies can give decision makers the foundation intelligent decision systems;
- Declarative knowledge representation in ontologies may allow execution of procedural knowledge by an inference engine to improve the knowledge represented.

Acknowledgement

This work was supported by CNCSIS-UEFISCSU, project number PN II-RU code 188/2010.

References:

[1] A. Ankolekar, M. Krotzsch, T. Tran, D. Vrandecic, The two cultures: Mashing up Web 2.0 and the Semantic Web, *Journal of Web Semantics*, 6(2008) 70-75

[2] S. Ba, A.B. Kalakota, Whinston, Executable Documents DSS, *Proceedings of 3rd International Conference on DSS*, (Hong Kong, 22-23 June 1995).

[3] H.K. Bhargava, D.J. Power, Decision Support Systems and Web Technologies: A Status Report, *Seventh Americas Conference on Information Systems* (2001).

[4] H.K. Bhargava, D.J. Power, D. Sun, Progress in Web-based decision support technologies, *Decision Support Systems* 43(2007), Elsevier, 1083-1095

[5] M. Cohen, C.B. Kelly, A.L. Medaglia, Decision support with Web-enabled software, *Interfaces* 31(2001) 109-129

[6] J. Czyzyk, J. Owen, S.J. Wright, Optimization on the Internet, *OR/MS Today* 24 (1997) 48-51

[7] D. O'Leary, Internet-based information and retrieval systems, *Decision Support Systems* 27(3) (1999), 319-327

[8] D.J. Power, Web-based and model-driven decision support systems: concepts and issues, *Proceeding of the Americas Conferece on Information Systems* (AMCIS 2000), Long Beach, California, 10-13 August 2000

[9] S.S. Sure, Y.R. Studer, Methodology for development and employment of ontology based knowledge management applications, *SIGMOD Record* 31(4) (2002) 18-23

[10] Mashup (web application hybrid) wikipedia

[11] DARPA, Ontologies by keyword, DAL Ontology Library, The DARPA Agent Markup language Homepage, 2004