Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 37 (2014) 48 – 55

Procedia Computer Science

The 5th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN-2014)

A new distributed expert system to ontology evaluation

Djellali Choukri[a,b]

[a]LANCI UQAM, Local W-5353 Case postale 8888, Succ. Centre-Ville Montréal (Québec) H2X 3Y7, Canada
[b]LATECE UQAM, 201, PK 4470, Président Kennedy Montréal (Québec) H2X 3Y7, Canada

Abstract

This paper addresses the increasingly encountered challenge of ontology evaluation. The best known approaches to ontology evaluation focus on the used criteria and model domain. In the present study, we propose a new approach that uses a reasoning tool as distributed expert system based on symbolic structures of facts and rules. The semantic model found is used to verify the ontology consistency and unexpected relationships between the ontological artefacts. The evaluation is based on formal system defined in description logic SHOIN. The results show that our evaluation approach is independent of the conceptualization of domain model and considers the main features of ontology structure. Good experimental studies demonstrate the multidisciplinary applications of our approach.

Peer-review under responsibility of the Program Chairs of EUSPN-2014 and ICTH 2014.

Keywords: Semantic Web, Ontology, logic, owl, consistency, subsumption, instantiation, DIG.

1. Introduction

The ontology is proving to be the best solution for communication and knowledge sharing. It provides a deeper level of semantics with significant metadata and ontological commitments to share knowledge between the interaction partners. The underlying aim is to clarify the semantics of Web resources through metadata or annotations.
Ontology evaluation remains a significant problem in Semantic Web. The literature contains many definitions of ontology evaluation; many of these contradict one another.

djellali.choukri@courrier.uqam.ca

Peer-review under responsibility of the Program Chairs of EUSPN-2014 and ICTH 2014.
doi:10.1016/j.procs.2014.08.011

However, the best known and most cited is (Gómez-Pérez, see [9]), which is also the definition we adopt in our paper « A *technical judgment of the content of the ontology with respect to a frame of reference during every phase and between phases of their lifecycle* ».

Several ontology evaluation approaches have been proposed in the ontologies evaluation literature. The choice of an appropriate approach depends on the evaluation purpose; application and the criteria used to evaluate the ontology. These criteria focus on the characteristics of ontology and they are independent of the application domain. It is difficult to choose the best user who judges the ontology quality and the proposed criteria for the evaluation. Moreover, the choice of a suitable approach depends on the model domain.
In order to overcome these drawbacks, we propose a new evaluation approach which is based on the property of soundness, completeness and decidability.

The paper is organized as follows: In Section 2, we present the current state of the art in ontology evaluation, our research questions and the problematic of ontology evaluation. The conceptual architecture of our approach is given in Section 3. Before we conclude, we give in Section 4 a short evaluation with benchmarking model for our conceptual model. Then, a conclusion (Section 5) and future work (Section 6) end the paper.

2. State of the art, Problem and Research Questions

There are several approaches that have been presented in the literature for ontology evaluation. The best known approaches fall into one of the following categories:

- *Gold Standard*: following a benchmark comparison using the Gold Standard reference ontology.
- *Evaluation based on data*: using a data-driven method to evaluate the degree of fit between the ontology and a Data Set describing the problem domain to which the ontology refers. Therefore, this approach requires traceability mechanisms to describe the concrete relations between the ontology entities and the Data Set.
- *Evaluation based on criterion*: a qualitative approach following the ontology characteristics regardless of the application domain [10]. The most known criteria are grouped as shown in Table (1):

Table 1. Summary of work on the criterion of ontology evaluation.

Author	Year	Total criteria	Objective
Gruber	95	5	Design criteria
Uschold &Grüninger	96	3	Design criteria
Noy & Hafner	97	28	Design criteria
Hovy	97	36	Compare linguistic ontologies
Uschold	98	10	Identify the roles of ontology in applications

- *Evaluation based on tasks*: accomplishing a given task using ontology and evaluating the results [1].

The different approaches proposed in the ontologies evaluation literature consult experts who use their opinions and experiences in the evaluation process. However, they are subject to extreme evaluations. Furthermore, the choice of a suitable approach depends on the used criteria and the model domain.
In this paper, we propose a new approach that uses a reasoning tool as distributed expert system based on symbolic structures of facts and rules. The evaluation is based on formal system defined in description logic $\mathcal{SHOIN}$.

3. The architecture of our evaluation system

Consistency, subsumption and instantiation can accentuate the main features of the conceptual hierarchy as well as the ontological inference, i.e., decidable decisions about satisfiability, ontology hierarchy and completeness. Description logic is perfectly suited to this situation. It has a formal semantics based on logic and equipped by decidable decisions. In addition, the description logic differs from its predecessors, such as conceptual graphs, existential graphs, semantic networks and frames where it is governed by a formal semantics based on logic [6],[7].

As shown in figure (1), the descriptive inference system used to verify the consistency, soundness and completeness of an ontology is based on the inference engine RacerPro[2] (Renamed ABox and Concept Expression Reasoner) [8].
The ontology can be regarded as a T-Box/A-Box representation with a hierarchy of roles describing the domain in terms of classes (concepts) and properties (roles).
Thus, we can consider the evaluation system as distributed expert system based on structures of facts and rules. It provides a symbolic reasoning used to verify the ontology completeness and soundness.
The DIG protocol (XML standard) [3] is used to connect the applications to a semantic model based on ontology. The allocation of terminological knowledge base allows users to query the conceptual model. In this way, we can query the terminological knowledge base to ensure that all facts deduced from the ontology can be inferred from the Data Set. In other words, our main objective is to describe how the ontology realizes requirements and needs described in the requirements model, i.e., the specifications and ontological commitments.
The inference system is divided into three main components: inference engine, terminology T-Box and description of the world A-Box. In each of these components, knowledges are declared in the form of rules and facts. The rules are related to the terminology reasoning operations (subsumption and instantiation).

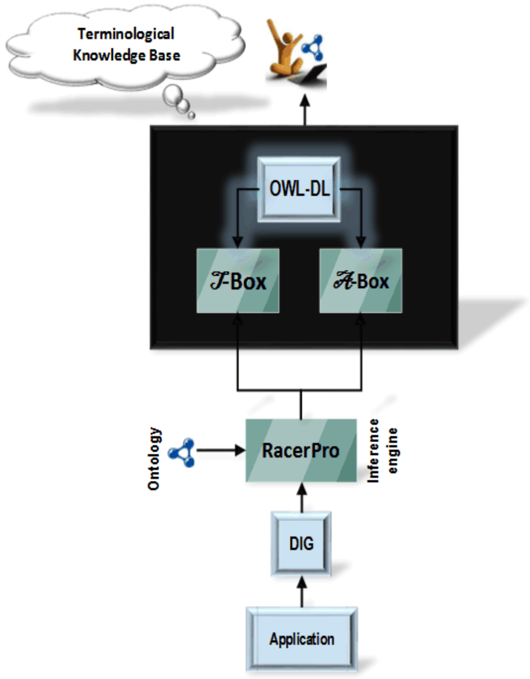

Fig. 1. The Descriptive Inference System.

4. Experimentation

4.1. Ontology

The CRISP-DM-OWL[1] ontology used in this project is integrated into a hybrid system DM, describing the artifacts and the basic rules to improve the intelligence level of the system. The ontology acts as a source of additional knowledge [4].

Figure (2) describes the CRISP-DM-OWL ontology with the artifacts involved in each section. The hierarchy of concepts is illustrated using the GraphViz tool [5]. In order to simplify the layout hierarchies of concepts and to allow a better understanding, we show only a few specializations.

[1] http://www.elmanahel.ca/ontology/crisp-dm-owl.owl

[2] http://racer.sts.tuhh.de/

Fig. 2. The different techniques of modeling.

4.2. Inference engine

In order to ensure the ontology consistency, RacerPro provides several inference rules that deduce implicit knowledge from the explicitly represented knowledge. These inference rules are decidable and with low complexity. The inference engine transforms the ontology artefacts as descriptive rules to support formal reasoning. The rules are categorized and arranged according to the level of complexity.

It should be noted that the ontology $O = (C, H_C, R_C, H_R, I, R_I, A)$ is considered in two parts:

- The extensional part C, R_C ou $\mathcal{T}$-Box: representation and manipulation of concepts and roles in terminological level.
- The intensional part I, R_I ou $\mathcal{A}$-Box: representation and manipulation of individuals in a factual level.

In this symbolic definition, ontology is represented by a hierarchy of symbols H_C (H_R) connected by the subsumption relation $\subseteq$.

C : concepts are arranged in a schema hierarchy H_C.
R_C : the set of relationships between concepts which are also arranged in a hierarchy H_R.
I : the instances that are interconnected by all instances of properties R_I.
A : the set of axioms used to express other relationships between concepts and to constrain their interpretations.

RacerPro adopts Fixed Point Semantics in order to avoid the definitions of terminology, which contain terminological cycles (rewriting termination).

We choose RacerPro reasoning as a tool for our approach because it offers more features and more graphics editors such as OilEd and RacerPorter (RACER Interactive Client Environment). It also offers a manipulation of symbolic reasoning for application developers. In addition, RacerPro offers several optimization techniques with proof tools:

- RacerPro includes several optimization techniques to ensure good performance of search, in particular, the dependency-directed backtracking and DPLL-style semantic branching.
- The verification of conceptual consistency.
- The search for inconsistent concepts in terminology $\mathcal{T}$-Box.
- The determination of parents and children of a concept.
- The verification of the consistency of the description of the world $\mathcal{A}$-Box.
- Testing the description of the world $\mathcal{A}$-Box and the terminology $\mathcal{T}$-Box.
- Find the subsumed/subsuming ontological artefact in the terminology $\mathcal{T}$-Box and description of the world $\mathcal{A}$-Box.
- Calculate the direct types of individuals [2], [7].

4.3. Terminology $\mathcal{T}$-Box

This component contains terminological axioms that describe the relationship between concepts and roles. Thus, RacerPro define terminological axioms which have the following form:

$$A \subseteq B(r_1 \subseteq r_2) \vee A \equiv B(r_1 \equiv r_2) \quad \forall A, B \in O; \forall r_1, r_2 \in R_I .$$

The axioms of the first type are called inclusions, while axioms of the second type are called equalities. The semantic description of the terminological description $\mathcal{T}$-Box is interpreted as a subset of a domain of interpretation.

The subsumption $\subseteq$ is the terminological inference based on concept expressions in the terminological description $\mathcal{T}$-Box [7].

Several rules can be present in the terminology $\mathcal{T}$-Box:

- **R1** : $(A \equiv B) \leftrightarrow (B \subseteq A \wedge A \subseteq B)$.
- **R2** : $\neg(A \equiv B) \leftrightarrow (A \cap B \subseteq \perp)$.
- **R3** : $(A \subseteq B) \leftrightarrow ((A \cap \neg B)^I = \phi)$.
- **R4** : $(A \equiv B) \leftrightarrow ((A \cap \neg B)^I = \phi \wedge (\neg A \cap B)^I = \phi)$.
- **R5** : $\neg(A \equiv B) \leftrightarrow ((A \cap B)^I = \phi)$.
- **R6** : $(A^I = \phi) \rightarrow (A \subseteq \perp)$.
- ….. $\forall A, B \in O$.

For reasons of readability, we cannot present all the rules.

Figure (3) illustrates the localhost connection with RacerPro server, which shows the consistency verification of terminological axioms $\mathcal{T}$-Box.

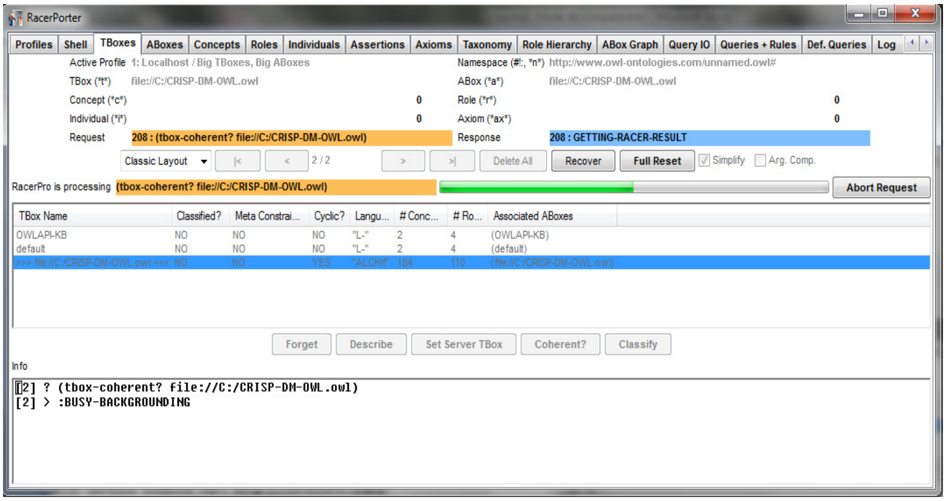

Fig. 3. The terminological reasoning $\mathcal{T}$-Box.

4.4. Description of the world $\mathcal{A}$-Box

The second component of the terminological knowledge base is a description of the world $\mathcal{A}$-Box. The ontological entities in this world can take the following form:

$$A(o_1),\ r(o_1, o_2),\ (\forall A \in O\ ,\ \forall o_1, o_2 \in I, \forall r \in R_I)\ .$$

The first type is called concept assertion and the second type is called the role assertion. In the description of the world $\mathcal{A}$-Box, RacerPro declares all instances I that are interconnected with all relations R_I .
The inference engine uses the terminological instantiation reasoning to determine if an object o is an instance of a concept A (also known as the validity of the assertion $A(o)$) [7].
Several instantiation rules may arise in the description of the world $\mathcal{A}$-Box:

- **R1** : $(B(o) \wedge B \subseteq A) \rightarrow A(o)$.
- **R2** : $((A(o) \wedge B(o)) \wedge (\neg(A \subseteq C)) \wedge (\neg(B \subseteq C)) \wedge (((\,and\ \ A \wedge B) \subseteq C)) \rightarrow C(o)$.
- **R3** : $r(o_1, o_2) \wedge ((all\ \ r\ A))(o_2) \rightarrow A(o_1)$.
- **R4** : $r(o_1, o_2) \wedge ((A \subseteq B) \wedge (all\ r\,B)(o_2) \rightarrow A(o_1)$.
- **R5** : $((1 \le i \le n),\ A(o_i) = B(o_i)) \rightarrow (A \equiv B)$.
- ….., etc. $\forall A, B, C \in O\ ,\ \forall o_1, o_2 \in I$.

Figure (4) shows the consistency treatment using the terminological instantiation reasoning in the description of the world $\mathscr{A}$-Box.

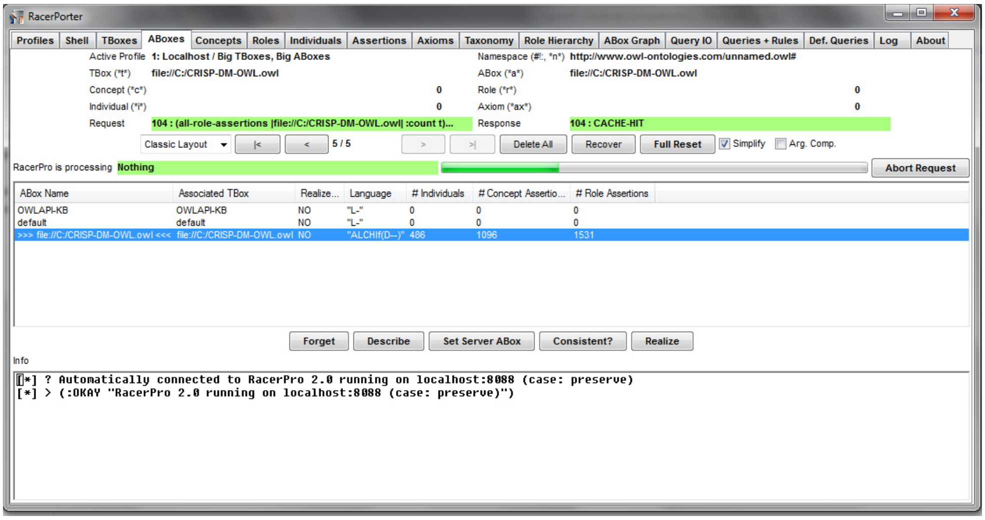

Fig. 4. The terminological instantiation reasoning of the world description $\mathscr{A}$-Box.

The inference rules mentioned above are very important because they are used to verify the ontology consistency, unexpected relationships between the ontological artefacts, terminological knowledge base unsatisfiability, determine subsumers and subsumed, avoid definitions of terminology that contain cycles and fixed points, etc.

5. Conclusion

In this paper, we have introduced a new approach providing standardized evaluation for ontologies. Our approach is independent of the conceptualization of the domain model and considers the main features of the ontology structure and its population (concepts, instances, axiom, relationship, etc.). It has a formal semantics based on description logic and equipped by decidable decisions. In addition, our system offers several advantages:

- Soundness, completeness and decidability.
- Ability to transform a descriptive representation to a first order predicate logic representation.
- Efficient for reasoning by classification.
- A well-defined semantics.
- The simpleness of modeling ontologies.
- Duality expressiveness versus complexity.
- The representativeness is described by two levels, i.e., the terminology description $\mathscr{T}$-Box and world description $\mathscr{A}$-Box.
- Classification and instantiation are operations that are at the base of the terminological reasoning.

6. Future work

The purpose of our next work is to propose a visualisation tool for representing the ontological knowledge. With the presentation of the ontology hierarchy we can basically understand the inheritance relationships between ontological artefacts. The visualization tool should at least provide an overview of the hierarchy or partial views, allowing the user to focus on a part of the ontology. These elements should be displayed so that you can discern the relevant information.

References

1. Brank, Janez, Marko Grobelnik et Dunja Mladenić. 2005. A survey of ontology evaluation techniques.
2. Haarslev, Volker, et Ralf Möller. *Racer: A Core Inference Engine for the Semantic Web*: *EON,* 2003.
3. S Bechhofer, R Moller, and P Crowther. The DIG description logic interface: DIG/1.1, 2003.
4. Shen Yanfen. A formal ontology for Data Mining: principles, design and evolution Thesis UQTR, 2007.
5. Gansner, E.R. Drawing graphs with Graphviz, Technical report, AT&T Bell Laboratories, Murray, 2009.
6. The description logic handbook: theory, implementation, and applications. Edited by. Franz Baader. Deborah L. McGuinness. Daniele Nardi. Peter F, 2003.
7. Napoli, Amedeo. An introduction to description logics. 1997.
8. Haarslev V, Hidde K, Mölr R, Wessel M. The RacerPro knowledge representation and reasoning system. Semantic Web 2011; 1-7.
9. Gómez-Pérez A, Fernández-López M, Chorcho O: Ontological Engineering, Springer-Verlag, London, UK, 2004.
10. Lozano-Tello, Adolfo, et Asunción Gómez-Pérez. «Ontometric: A method to choose the appropriate ontology». Journal of Database Management, 2004; 15: 1-18.