Converting OBR-Scolio Ontology in OWL DL

Vanja Luković¹, Danijela Milošević², Goran Devedžić³, and Saša Ćuković⁴

¹ University of Kragujevac, Faculty of Technical Sciences, Svetog Save 65, 32000 Čačak, Serbia
vanja.lukovic@ftn.kg.ac.rs

² University of Kragujevac, Faculty of Technical Sciences, Svetog Save 65, 32000 Čačak, Serbia
danijela.milosevic@ftn.kg.ac.rs

³ University of Kragujevac, Faculty of Engineering, Sestre Janjić 6, 34000 Kragujevac, Serbia
devedzic@kg.ac.rs

⁴ University of Kragujevac, Faculty of Engineering, Sestre Janjić 6, 34000 Kragujevac, Serbia
cukovic@kg.ac.rs

Abstract. We developed the OBR-Scolio ontology that models scoliosis as pathological state of spine using the method of extracting from the FMA reference ontology in anatomical domain and the principal ontology framework of the OBR reference ontology, which is spread over anatomy, physiology and pathology domains. Following the FMA modeling framework, OBR-Scolio ontology initially is created in Protégé using its frame based representation. In order to enable more powerful reasoning support, ontology visualization, and more precise concepts' definition and description, we converted OBR-Scolio in OWL DL language, due to its higher expressiveness. The paper addresses and discusses the key conversion principles, as well as our experience in such conversion, and the results obtained from Racer reasoner.

Keywords: Medical Ontologies, Ontology Conversion, Frame-based representation, Description Logic, OWL DL.

1. Introduction

Significant amount of efforts employed so far in developing reference and application ontologies in the domain of medicine and biology, including anatomy structures, brought essential advances in both deep fundamental understanding of the issue, as well as the application capabilities and power deep fundamental understanding of the issue [1-9] as well as the application capabilities and power [10-12]. Widely accepted and applied frame [13], [14] and semantic network [15] based representations of the global medical concepts and terminology turns to demonstrate the lack of formal semantics for precise definition of ontology’s concepts, particularly in the pathology (sub)
domain and cannot benefit from the reasoning support provided by description logics (DL) [16], [17]. This is recognized as the essential need for conversion frame-based ontology representations into more expressive forms that provide more accurate and meaningful reasoning with subtle and delicate anatomical and pathological concepts.

We have developed the OBR-Scolio ontology that models scoliosis as pathological state of spine [18], [19]. It is extracted (pruned) from the FMA reference ontology [20] using de-novo method [21]. Having the FMA ontology constructed using Protégé frame-based representation, OBR-Scolio ontology consequently inherits the same encoding formalism. Although Protégé as frame based representation tool is equipped with complex knowledge capturing and ontology construction capabilities, the knowledge representation formalism it uses lacks in expressiveness [5], [6]. The key semantic deficiency stems from the fact that medical terminology used in domain knowledge representation, together with very rich relationships between corresponding entities, requires modeling approach and methodology that ensures high precision, which in turn prevents from misclassification, i.e. imprecise or even incorrect reasoning. In that course, OWL DL as the most desirable dialect of OWL for ontology development [16], [22-25] lends its description logic mechanisms for more accurate concepts’ definition and more reliable reasoning for clinical application.

In order to overcome aforementioned deficiencies, as well as to improve interoperability with other OWL ontologies, we convert original OBR-Scolio frame-based representation into OWL DL. The key objective is to facilitate and empower the scoliosis ontology with the expressivity that supports sound and consistent anatomical and pathological semantics [16], [23], [24]. In addition, this conversion also enhances features of reusability, sharing, maintenance and integration with other individual (application) biomedical ontologies.

For the sake of terminological clarification, in our study we use term conversion, in the same sense as used in [16] in particular, as well as in [23], [26], adopting the fact that OBR-Scolio originates from the FMA ontology and has been created with the same encoding formalism. In many other works terms migration [26-30], translation [23], [31] and transformation [32] are used interchangeably with the term conversion. Although in deep fundamental meaning each of the terms (may) make the difference, the term conversion applied in our case supports the widely accepted approach of conversion of frame based encoding into OWL DL and does not affect the overall methodology and findings.

Similarly with many others biomedical ontologies (see Section 2 for details) we found that conversion process from frame-based ontology representation into the OWL DL representation is indispensable for our application ontology.

---

1 FMA ontology is also available in other formats as OWL full file and Clips archive files, but these formats are not recommended to use in Protégé and do not match completely with original Protégé frame based representation [20].
OBR-Scolio too, due to higher expressiveness of the OWL DL language, possibility to visualize the ontology and to invoke reasoners for subsumption and consistency testing, as well as the capability to integrate it with other OWL ontologies.

The paper is organized as follows: Section 2 briefly overviews the work related to converting frame-based biomedical ontologies with DL formalisms. The next section introduces frame-based representation of OBR-Scolio application ontology in scoliosis domain. Fourth section describes our experience in converting OBR-Scolio Ontology in OWL DL. Finally, we provide the conclusions and directions of future work.

2. Related Work

Numerous researchers and developing teams used Protége [33], modeling environment and knowledge representation tool for building frame-based medical ontologies [34], due to its powerful functionalities and capabilities. On the other side, this is not the only platform and tool in the field of medical ontology development that employs frame-based representations for modeling the domain's concepts and associated properties and relationships. Ontolingua Server [35], Onto Edit Professional [36], WebODE [37] and WebOnto [38] ontology development tools allow representing knowledge following a hybrid approach, based on frames and first order logic. However, these all demonstrate the same kind of deficiency — the lack of expressiveness required for derivation of the new facts resulting from reasoning over a single or multiple representations. The same holds for checking consistency and hierarchical organization of the classes in the ontology, in the sense of existence of classes' instances and classes' hierarchy. This is of the crucial importance for proper inferring process of the reasoner.

Description Logic (DL) adds a formal, logic-based semantics to structured knowledge representation paradigm, which is employed in frame and semantic network knowledge representation [15] and provides the theoretical foundation for ontology web languages (OWL). A DL knowledge base can be divided conceptually into three components: the Tbox, the Rbox and the Abox. The Tbox contains assertions about concepts (such as subsumption and equivalence axioms). The Rbox contains role inclusion axioms and constraints on the roles. The Abox contains role assertions between individuals and membership assertions. Beside formal, logic-based paradigm in knowledge representation another advantage of description logic is expressed in possibility to provide reasoning services. Main reasoning tasks concern: satisfiability, subsumption and instance checking. Subsumption and satisfiability reasoning tasks are description logic reasoning services which are typical for a terminology (TBox). Subsumption DLs reasoning service that supports the classification of the ontology is intended to indicate whether one concept is more general than another one, that is, whether the first subsumes
the second [39]. At the other hand, a concept makes sense and is said to be 
*satisfiable* if there is some interpretation (meaning) that satisfies the axioms of 
the terminology of the domain such that the concept denotes a nonempty set 
in that interpretation (for more formal definitions of satisfiability refer to [39]). A 
class is deemed to be unsatisfiable (inconsistent) if it cannot possibly have 
any instances [39], [40]. The basic reasoning task in an ABox is instance 
checking, which verifies whether a given individual is an instance of (belongs 
to) a specified concept. Instance checking reasoning support also *realization*, 
which finds the most specific concept an individual object is an instance of; 
and *retrieval*, which finds the individuals in the knowledge base that are 
instances of a given concept.

There are many description logic based languages like: DAML (DARPA 
Agent Markup Language), OIL (Ontology Inference Layer), DAML+OIL [41] 
and OWL [42, 43], which is based on DAML+OIL language. Description 
logics-based languages have a precisely and formally defined semantics. 
Some generic reasoning tools have been developed to leverage this 
semantics. Thus, an application can reason about an ontology represented in 
description logics without having to implement any inference function.

The most popular description logic formalism is currently the OWL (Web 
Ontology Language) that provides three increasingly expressive 
sublanguages: OWL Lite, OWL DL and OWL Full. Particularly, OWL-Lite and 
OWL-DL belong to description logics. OWL Lite possesses minimum and 
limited expressivity, without losing computational completeness and 
decidability (all computations will finish in finite time). OWL DL sublanguage is 
high expressive and decidable. OWL Full is maximally expressive but 
undecidable, with no computational guarantees (reasoning support is unsure). 
For the purpose of converting our application ontology we chose OWL DL 
sublanguage, because it offers a trade-off between expressivity and 
decidability.

There are many description logic based ontology development tools such 
as: OILEd [44] and OntoSaurus [45]. Although the new Protégé releases 
provide Protégé-OWL editor for representing ontologies using DL based 
formalisms, the need for conversion previously built frame-based Protégé 
ontologies into highly expressive OWL (W3C's Web Ontology Language) 
continually attracts the attention.

Biomedical ontologies are increasingly taking advantage of Description 
Logic (DL) based formalisms in representing knowledge. GALEN [46] and 
SNOMED CT (SNOMED Clinical Terms) [3] ontologies are two important 
examples of ontologies which were both developed in a native DL formalism. 
Other terminologies have been converted into DL formalism.

In [16] and [23] the conversion process of the FMA reference ontology from 
its frame-based representation in Protégé into OWL DL is described in detail. 
The conversion relies on translation and enrichment rules, implemented with 
flexible options. In [26] authors propose the conversion of the FMA reference 
ontology into OWL Full language, in other to directly represent metaclasses 
and to form application independent representation of the ontology. For the 
applications specific needs, OWL Full model of the FMA ontology needs to be
Converting OBR-Scolio Ontology in OWL DL

further restricted to the fragment that consist only application relevant concepts, which is then simplified into OWL DL. This is achieved by deleting all the OWL Full constructs (typically metaclasses and some relationships) that are not used in OWL DL representation of the application.

Unlike methods which are proposed in [16], [23] and [26] in [31] authors suggest complete OWL transformation of the Protégé frame based canonical representation of the FMA reference ontology. This is achieved by representing only the information that is explicitly present in the frames representation of the FMA or that can be directly inferred from the semantics of Protégé frames and by representing all the information that is present in the frames representation of the FMA. Complete representation of the FMA in OWL consists of two components: an OWL DL component that contains the FMA constructs that are compatible with OWL DL; and an OWL Full component that imports the OWL DL component and adds the FMA constructs that OWL DL does not allow.

Another approach assumes the FMA transformation into a description logic-based representation different than OWL based [47]. The special emphasis is given to the representation of partitive relations using special modeling technique, called ePI (extended part/include), which is an extension of the SEP (Structure Entity Part) triplet [48] and the PI modeling scheme [49].

The ontology migration process has been conducted for other OWL DL sublanguages, as well. Namely, in [27] the authors present migration and enhancement process of the existing Catalogue and Index of French-speaking Medical Sites (CISMeF) terminology into OWL DL sublanguage. CISMeF “encapsulates” the MeSH (Medical Subject Headings) ontology, which is a component of the UMLS (Unified Medical Language System) Metathesaurus [50] in French version. The CISMeF terminology is automatically transformed from the previous relational database into OWL ontology, using Java and SQL queries.

Gene Ontology (GO) [17] migrated into DAML+OIL ontology language from XML version using five well defined staged approaches [28]. The key intention is to take the advantage of the richer formal expressiveness and the reasoning capabilities of DAML+OIL description logic language. The conversion capability is achieved using OilEd ontology development tool for editing DAML+OIL ontologies. Consistency checking and automatic concept classification is performed using FaCT description logic reasoner [51].

In order to provide more formal representation of the previously frame-based ontology intended for the use in intensive care units DICE (Diagnosis in Intensive Care Evaluation) [4] developed in Protégé, Cornet and Abu-Hanna proposed migration to DL formalism [29]. Their work stress that the DL formalization process of the ontology representation revealed implicit ambiguities and deficiencies in concepts’ definitions and that the result of the process is the representation that can support automatic inference.

In [30] authors describe the conversion process of Semantic Network component of the UMLS Metathesaurus ontology [7], [52] in OWL language. National Cancer Institute Thesaurus ontology (NCI Thésaurus) [1] is public domain description logic-based terminology firstly developed in XML format.
Afterwards the ontology is converted from XML format to computationally complete and decidable OWL Lite language. In this paper we describe our own experience in converting the OBR-Scolio ontology from the frame-based representation into the OWL DL representation.

3. OBR-Scolio Ontology

3.1. Definition

OBR-Scolio application ontology in scoliosis domain of a spine is created using the so-called “de novo” method that extracts the application ontology from the reference one(s). That is, OBR-Scolio ontology is created from the complete Protégé frame based realization of the FMA reference ontology [2], [20], [53], and the principal ontology framework of the OBR reference ontology [54]. The former is the ontology of human anatomy, while the latter spreads over anatomy, physiology and pathology domains. In [18] we present detailed description of the OBR-Scolio ontology derivation process.

The two reference biomedical ontologies we used for creation of our OBR-Scolio application ontology do not possess taxonomy of pathological structures. The FMA is the reference ontology of anatomy of the idealized human body, and the OBR ontology has also incomplete taxonomy that does not code information about human body pathology. From the clinical point of view the both, anatomy and pathology, are of the crucial importance in the sense of faithful communication, diagnosis, decision making, and adequate treatment of disorders and dysfunctions. Therefore, we primarily proposed the way for deriving the complete taxonomy of spinal disorders pathology based on FMA taxonomy of anatomical structures. Then we applied “de novo” method and deleted classes from the FMA anatomical class hierarchy that are not relevant for scoliosis domain, as well as from obtained pathological class hierarchy. Additionally, the essential OBR-Scolio Pathological_vertebral_column class hierarchy is created using selective propagation of pathological relation “curvature-of” through part-hood hierarchy of spine, which is in [18] represented using SEP triplet and adapted SEP triplet modeling methodologies. Finally, the elaborating taxonomy of the Pathological_vertebral_column class is created by grouping previously obtained basic curvature types into the class Basic_curvatures_of_vertebral_column, which is further classified depending on the curvatures localizations and flexibility [55], [56] (Fig. 1). Beside these basic curvatures type, scoliosis ontology also possess classes that represent curvatures type according Lenke [57], since this system is the most reliable system for idiopathic scoliosis classifications that takes into account deformities of spine in both frontal and sagittal plane, as scoliosis is three dimensional deformity of the spine.
As the Figure 1 displays, the Lenke_type_curvatures_of_vertebral_column class has six subclasses, which represent Lenke’s curves types depending on localization, degree and flexibility of manifested curvatures. These subclasses are composed of appropriate subclasses of the class Basic_curvatures_of_vertebral_column. In order to incorporate the degree of lumbar spinal deformity in the frontal plane and thoracic deformity in the sagittal plane, classes that represent adequate Lenke’s curve type are further classified into classes, whose names end with appropriate lumbar frontal.
spine modifier label (A, B or C) from CSVL\textsuperscript{2} line and together with appropriate thoracic sagittal spine modifier label (+, - or N) (Fig. 1).

3.2. Protégé frame based representation of the OBR-Scolio ontology

For creation of OBR-Scolio application ontology we have chosen the Protégé 3.4.4 ontology development and knowledge acquisition environment \[\text{[33]}\]. In Protégé all concepts are represented as frames, which are data structures that contain all the information in the ontology about a given concept: attributes (properties) of the entity that concept refers to and the relationships of the entity to the other entities. Each concept’s frame is modeled both as a class and as an instance of corresponding metaclass. Its role as a class allows for propagation of its set of attributes to corresponding subclasses. On the other side, its role as an instance of the metaclass is to express all attributes of the metaclass as template slots with specific values assigned to them.

The main disadvantage of the Protégé frame based representation of the OBR-Scolio ontology is lack of adequate semantics for accurately defining ontology’s relevant concepts that represent Lenke type curvatures classification. Therefore, the need for representing the OBR-Scolio ontology in OWL DL higher expressive language becomes indispensible. In the next Section we present our experience in converting the OBR-Scolio ontology from Protégé frame based representation into OWL representation in detail.

4. OBR-Scolio Ontology in OWL DL

Three important types of ambiguities of a frame-based ontology representation according to [29] are:

- **Specification (in)completeness.** The concepts may have incomplete specification due to the fact that specified slot-values contain necessary but not sufficient conditions. Furthermore, frame-based ontology representation suffers from the lack of possibility to stress whether the specification of a concept is complete or incomplete, which can lead to improper placement of a class in the hierarchy and complicate the process of automatically adding or removing classes from the class hierarchy.

- **Semantic vagueness.** A slot value may be interpreted as a restriction, as a possible value, or as a necessary value. Consequently, the process of querying for classes meeting certain complex criteria, as well as consistency checking in class specifications meets practical drawbacks.

- **Multi valuing.** Due to three possible interpretations of multi-valued slots, only-one, one or more, or all of the values, the process of automatic

---

\(2\) CSVL is central sacral vertical line, drown vertically through the midpoint of first sacral vertebra S1 on standing posterioanterior radiograph film [56].
Converting OBR-Scolio Ontology in OWL DL

detection of redundancies and inconsistencies in classes’ specification may turn to be highly complicated.

Table 1. Correspondence between frames-based, Description Logic and OWL DL formalisms of ontology representation. Adapted from [29]

<table>
<thead>
<tr>
<th>Frames</th>
<th>Description Logic</th>
<th>OWL DL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>Concept</td>
<td>Class</td>
</tr>
<tr>
<td>Slot</td>
<td>Role</td>
<td>Property</td>
</tr>
<tr>
<td>Slot-value</td>
<td>Role-filler</td>
<td>Property value</td>
</tr>
<tr>
<td>Value type</td>
<td>Range of role</td>
<td>Range of property</td>
</tr>
<tr>
<td>/</td>
<td>Role restriction</td>
<td>Property restriction</td>
</tr>
<tr>
<td>Subclass</td>
<td>Child, subsume</td>
<td>Subclass</td>
</tr>
<tr>
<td>Superclass</td>
<td>Parent, subsumer</td>
<td>Superclass</td>
</tr>
<tr>
<td>Instance</td>
<td>Individual</td>
<td>Object, Individual</td>
</tr>
<tr>
<td>Incomplete specification</td>
<td>Primitive definition</td>
<td>Primitive definition</td>
</tr>
<tr>
<td>Complete specification</td>
<td>Definition</td>
<td>Definition</td>
</tr>
</tbody>
</table>

As opposed to frame-based ontology representation OWL DL formalism of ontology representation possesses many advantages:
- Higher expressiveness of the precisely defined semantics and richer set of operators that make possible for concepts to be clearly defined and described.
- Capability to invoke reasoners for subsumption testing, automatic computation of the inferred ontology class hierarchy against to asserted ontology class hierarchy, and consistency checking, whether or not it is possible for a class to have any instance.
- Integration of the OWL DL represented ontology with other OWL ontologies.
- Possibilities to visualize essential classes and their relations with other classes.

Another difference between frame-based and Description Logics based representation is that the latter relies on the “open world assumption” [39] whereas the former assumes a closed world. In a closed world, everything that is not explicitly said is assumed to be false. Therefore, there is a need to introduce a closure axiom [58] for all relationships in the ontology represented in Description Logic formalism.

OWL, Description Logic and Protégé frame based ontologies have similar components. Basically, OWL ontology consists of: Classes, Properties and Objects (Individuals) that roughly correspond to Protégé’s frame based ontology’s components: Classes, Slots and Instances, while adequate Description Logic ontology’s components are: Concepts, Roles and Individuals (Table 1). As opposed to OWL and Description Logic formalism, semantically poor frame based formalism lacks in defining restriction for slots values (Table 1).
In converting the OBR-Scolio ontology we have applied Export function from the frame based Protégé project into an empty OWL/RDF newly created Protégé project. Given that in generated OWL ontology every class has been represented also as an instance of another appropriate metaclass, we can conclude that using the Export function we have got OWL ontology, which hasn’t been represented in OWL DL language form, as we intended, but in OWL Full language form. Thence, in order to convert the generated OWL ontology in OWL DL language form, we have to define that every class is instance of a general owl:class, by changing metaclass of all subclasses of the owl:Thing class to owl:class (Fig. 2), and to delete all properties values information of metaclass instantiation. In addition, we had to manually apply some conversion rules for slots and classes according to [16], [23] and [26], which will be in detail described in following Subsections 4.1 and 4.2, respectively. Also, we converted some specific properties into annotation properties (Subection 4.2) and defined some basic ontology classes (Subection 4.3).
Converting OBR-Scolio Ontology in OWL DL

4.1. Converting Protégé Slots of the OBR-Scolio Ontology into Properties in OWL DL

In Protégé frame based ontology representation, slots have a Value Type specification (e.g., Integer, String, Boolean, Symbol with Allowed Values, Instance with Allowed Classes and Class with Allowed Superclasses) that corresponds to the range of properties in Protégé OWL ontology representation (Fig. 3, Table 1). Using Export function OBR-Scolio slots with Integer, String or Boolean Value Type specification have been automatically converted in OBR-Scolio datatype properties with appropriate range. Yet Class and Instance Value Type have been automatically converted in OBR-Scolio object properties without range specification. Thereby, according to [16], [23] and [26] we manually added the range specification for Class and Instance Protégé frame Value Types, as union of all Allowed Superclasses or all Allowed Classes (Fig. 4).

![Slot Editor](image)

**Fig. 3.** Slot editor of the adjacent slot with Boolean Value Type in Protégé frame based ontology representation

At the other hand, during this direct translation Symbol Value Type specification has been automatically converted to Datatype property, which is only appropriate for True and False allowed values. However, for other allowed values we convert this automatically created Datatype property in Object property and specialize its range as the enumerated class of all previously created OWL:Thing individuals [16], [23], [26] (e.g. for individuals: 1-dimension, 2-dimension and 3-dimension, which are allowed values for Symbol Value Type of the dimension slot, we have created corresponding OWL:Thing individuals: individual_1-dimension, individual_2-dimension, individual_3-dimension and defined the range for the dimension property as union of such created individuals.
Fig. 4. Range specialization of attributed part property as the union of all Allowed Classes in Protégé OWL based ontology representation

Furthermore, during this direct translation slots with cardinality at most one are automatically converted into functional properties, while slots having inverses (Inverse-Slot specification) are automatically converted in a owl:inverseOf relation in OWL DL.

4.2. Converting Protégé frame Classes of the OBR-Scolio Ontology into Classes in OWL DL

In OWL ontology representation all classes are subclasses of the class owl:Thing, although it is the general class that represents the set containing all individuals. OBR-Scolio ontology is organized as a hierarchy of mutually disjoint concepts. Nevertheless, in OWL ontology representation classes are assumed to “overlap” and are not disjoint by default, which is the key distinction from the frame based ontology representation that is based on “unique name assumption” (everything named different is different). Hence, we have to specify that all direct subclasses of a class are mutually disjoint, which ensures that an individual that has been asserted to be a member of one class in the group of a class’ siblings cannot be a member of any other class in that group.
Every class of the OBR-Scolio Protégé frame based ontology representation is represented both as a metaclass and an instance of another metaclass. As a metaclass it has name specification, specification of slots introduced in this class and all inherited slots from its superclasses (the upper part of the Fig. 5). As an instance of another metaclass the class has slots values specifications of that metaclass (lower part of the Fig. 5). In directly obtained OWL representation of the OBR-Scolio ontology by using the Export function, specifications of class properties and all inherited properties have been lost (Fig. 6). However, in this direct translation all metaclass instantiations information has been retained, which is customary for the Protégé OWL Full ontology representation. Considering that it is not possible to automatically compute the classification hierarchy and check for inconsistencies in the OWL Full ontology representation, we convert the ontology in OWL DL representation. Therefore, we have to delete this metaclass instantiations information, thus specifications of all slots values (properties values in OWL) have thereby also been lost. Afterwards, we had to manually add all these missing specifications using rules that we will explain in detail in the in the sequel.

Every slot in a class with Class or Instance Value Type in Protégé OWL DL based ontology representation is according to [16], [23] and [26] represented as universal property restrictions (owl:allValuesFrom) on union of all allowed
slot’s superclasses or classes. Accordingly, beginning from the top level classes and analyzing all slots with Class or Instance Value Type specification, we have added universal property restrictions owl:allValuesFrom in all domain classes containing these slots in the whole OBR-Scolio ontology.

Fig. 6. Class editor of Organsimal_continuant class in Protégé OWL based ontology representation after direct translation from the Protégé frame based ontology representation

For example, in Organsimal_continuant class exist only located in and location of slots with Class Value Type (Fig. 5). These slots have the same allowed superclass Organismal_continuant, for which we had to create owl:allValuesFrom restriction (Fig. 7).

Every slot introduced in a class with Symbol Value Type and allowed slot’s values different from True and False in Protégé OWL DL based ontology representation is also represented as universal property restrictions (owl:allValuesFrom) on union of all OWL:Thing individuals that correspond to these allowed slot’s values.

Introducing universal property restrictions (owl:allValuesFrom) for Class, Instance and Symbol value, we have retrieved specifications of all Object properties in complete classes hierarchy of the OBR-Scolio ontology.

Every metaclass instantiation specification of a slot value, when a slot is of Class Value Type is represented as property restriction with owl:someValuesFrom constraint on a value, which is a subclass of an allowed superclass. Moreover, every metaclass instantiation specification of a slot value, when a slot is of Instance, Symbol, Integer, Float, String or Boolean...
Value Type is represented as property restriction with \textit{owl:hasValue} constraint on the value.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig7}
\caption{Creating \textit{owl:allValuesFrom} restriction in Protégé OWL based ontology representation}
\end{figure}

Slots such as \textit{OBRID} and \textit{definition} that represent concepts identifiers and descriptions are manually converted from datatype properties with \textit{String} range into \textit{annotation} properties: \textit{OBRIDa}, \textit{definitionA}. Moreover, all slots representing concepts names that are instances of the class \textit{Concept\_name} such as: \textit{Preferred name}, \textit{Synonym} and \textit{Non-English equivalent} have also been manually converted from object properties into following annotation properties: \textit{Preferred\_nameA}, \textit{SynonymA} and \textit{Non-English\_equivalentA}. This is done by deleting these object properties and all instances of the \textit{Concept\_name} class as well as by converting all slots values of an instance of the \textit{Concept\_name} class into OWL DL \textit{data literal}. For example, for the \textit{Preferred name} slot in the \textit{Organismal\_continuant} class we have created annotation property \textit{Preferred\_nameA} and converted all values of appropriate instance of the \textit{Concept\_name} class into following data literal (Fig. 8):

\begin{quote}
“autor(Cornelius Rosse, MD) authority(Rosse MD) date_entered(Fri Mar 5 3:34:47 AM 2010) name(Organismal\_continuant)".
\end{quote}
4.3. Defining Classes in OWL DL by Necessary and Sufficient Conditions

Applying all previously described conversion rules, frame-based classes’ specifications (slots and slots values) have been converted into OWL’s properties restrictions. These define anonymous super-classes of the class being specified. Interpretation of these properties’ restrictions is two-fold: Interpreting them as the sets of necessary conditions or interpreting them as the sets of necessary and sufficient conditions. In the first case all classes would have primitive definitions. These lead to consistent ontology but inhibit
the process of automatic classification of classes. At the other hand, if we interpret frame-based classes’ specifications as the sets of necessary and sufficient conditions, all classes will have complete definitions. However, in such case all classes that were solely defined as subclasses of exactly one other class, as well as many other classes with the same sets of necessary and sufficient conditions will be regarded as equivalent in the ontology.

Considering this explanation and the fact that our application OBR-Scolio ontology was created using the method of extracting from the FMA and OBR reference ontologies (see [18] for details), using the process of deleting all classes that were not relevant for our application domain, our further approach was to define only representative classes of the OBR-Scolio ontology in order to have full reasoning support.

Using structural and non-structural criteria for all minor curvatures of vertebral column [18] all classes representing basic curvatures of vertebral column were defined. For instance, the Figure 9 illustrates definition of the class *Lumbar_structural_curvature*.

**Fig. 9.** Definition of the class *Lumbar_structural_curvature*

Furthermore, all classes representing Lenke’s curvatures types of vertebral column were defined on the basis of previously defined classes that represent basic curvatures of vertebral column. At this point it is important to note that these concept definitions do not violate tree model property. As example definition of the class *Lenke_type4_curvatures_of_vertebral_column* is shown in Figure 10.

**Fig. 10.** Definition of the class *Lenke_type4_curvatures_of_vertebral_column*
4.4. Reasoning over OBR-Scolio in OWL with Racer and Visualizing the Ontology

OWL DL ontology representation provides the maximum expressiveness without losing computational completeness and decidability of reasoning systems. The reasoners have a wide specter of functionalities, e.g. whether or not it is possible for the classes to have instances. While processing the consistency and hierarchical organization upon the aimed ontology, the reasoner can derive the inferred ontology class hierarchy as opposed to the asserted one. Additionally, classification checking strongly demands that classes should be defined with both, the necessary and sufficient conditions. In order to check the consistency and hierarchical organization of the classes, the OBR-Scolio ontology is converted into OWL DL, as proposed in [18].

For the reasoning purpose over OBR-Scolio ontology in Protégé OWL, we have used the Racer (Version 1.9.2 beta), a DIG (Description Logic Implementers Group) compliant reasoner [59], [60], available for variety of platforms [61]. The Racer comes with a number of features, such as following:
- Discovering inconsistent concepts or relationships, and faulty subclass or sub-property relationships
- Concept equivalence detection in terminology creation and merging, as well as matching of search queries and document annotations
- Recognizing the parent/child relations in directed acyclic graph concept
- Concept position determining in the specific hierarchy for enabling vocabulary merging into DAG structure
- Semantic distance estimation between concepts to limit the navigation of DAG, etc.

The Racer was launched from Protégé-OWL, but the classification processing has failed because Racer could not handle the entire OBR-Scolio OWL ontology. Thus, we choose to test smaller portions of the ontology to detect the errors and incrementally added more features subsequently while analyzing the results.

Furthermore, the inferred class hierarchy with defined and primitive classes, as well as the annotation and all datatype properties, without object properties, was computed in about 1.8 seconds on Pentium 4 with 512MB. The OBR-Scolio consistency checking was also successfully performed on this portion of the ontology. Although we encountered the warnings on the range of the float datatype for cardinality property we have changed the Range of the property to string datatype. Additionally we encountered the warnings on the hasValue restriction on some boolean datatype properties. Therefore we converted these properties in object properties with range as enumerated class of two previously created OWL Thing individuals: individual_True and individual_False. Accordingly, the OBR-Scolio ontology has been successfully checked for consistency without any warnings.

Afterwards, we have introduced object properties with restrictions in subclass axioms without any inverse object properties, and the Racer successfully processed consistency of OBR-Scolio ontology and provided inferred class hierarchy in less than 1.9 seconds.
The next step was adding the pairs of inverse properties: constitutional_part, constitutional_part_of, located_in, location_of, member, member_of, regional_partition_1, regional_partition_1_of and regional_part, regional_part_of with assigned restrictions. This part of OBR-Scolio ontology was also successfully checked for consistency and computed inferred class hierarchy was obtained in 2 minutes 20.4 seconds. As the result of classifying the ontology, it is shown that asserted and inferred taxonomies are the same.

However, when inverse pairs of properties were added, such as: part, part_of, systemic_part or systemic_part_of, the reasoner has failed to generate any results. Hence, we concluded that the reason for failure comes from the complexity of generated OWL ontology, due to the presence of complex restrictions on part, part_of, or systemic_part, systemic_part_of pairs of inverseOf ontology properties.

Upon completion of the step by step reasoner testing along with detailed analyses of the obtained results, the usefulness of DLs reasoning techniques and benefits of representing the ontology in OWL became quite obvious.

The visualisation of asserted and inferred taxonomy is enabled by converting the OBR-Scolio ontology in OWL. The visualisation of asserted and inferred taxonomy of the OBR-Scolio class Basic_curvatures_of_vertebral_column is presented on Figure 11. For this purpose we used OWLViz plugin [62] in Protégé version 4.1_beta [33]. Additionally, the OntoGraf plugin provides the functionality of visualisation of concepts and all corresponding relations. The Figure 12 shows visualisation of taxonomy of class Lenke_type_curvatures_of_vertebral_column and its relations with concept Basic_curvature_of_vertebral_column.
Fig. 12. Visualization of taxonomy of class Lenke_type_curvatures_of_vertebral_column and its relations with concept Basic_curvature_of_vertebral_column OntoGraf plugin

5. Discussion and Future Work

For generating the application ontology OBR-Scolio we have employed FMA version 2.0, which was released in the native Protégé frame based representation with a MySQL database backend. Newer versions of FMA reference ontology (FMA 3.0, FMA 3.1 and FMA 3.2) beside native frame based representation, possess also representations based on OWL 1.0 language [31] and are available from the original FMA site [20]. These are actually OWL Full representations of the FMA reference ontology that were generated using the conversion script written by the Natasha Noy [31]. This script can generate two OWL language versions of the FMA ontology: OWL DL and OWL Full components. However, we couldn’t employ these available OWL components to generate our application ontology for these two reasons: OWL Full component, although complete representation of the FMA reference ontology, is not amenable to automated reasoning, while the OWL DL component is incomplete representation of the FMA reference ontology, because it is obtained in omitting the metaclasses information from the native frame based representation of the FMA ontology.

For all these described reasons, for the conversion purpose of our application ontology in OWL DL language (OWL 1.0 language version) we have employed the general approaches described in [16, 23], which are also applied for converting the whole FMA reference ontology from its canonical
Converting OBR-Scolio Ontology in OWL DL

frame based representation into the OWL DL language. This conversion requires not only a syntactic “translation” of the original frame based ontology representation, but also a semantic “enrichment” of the generated ontology. The latter implies adding the properties restrictions and logical definitions of the classes described in detail in this work, and which are someway specific, depending on the target application ontology.

Beside general applied approach in [16, 23], in this paper we propose some solutions regarding definitions of the main concepts of our application ontology. For this purpose the used OWL 1 (DL) semantic approved to be enough expressive. In [16], [22] and [23] the class definitions, which are obtained by selecting the constitutional_part property, were not “semantically” satisfying for all classes, because all anatomical entities cannot be uniformly defined solely in terms of their constitutional parts. Our solution overcomes this limitation by using combination of two or more properties in the combination of the overall OWL DL language expressiveness to define representative anatomical concepts.

Our future direction will be to employ the newest released version of the FMA-OWL 2 (DL) ontology [24] for extracting OBR-Scolio application ontology and to compare the obtained ontology with current OWL 1 (DL) version. The results and conclusions from this comparative analysis will bring further enhancements, since this version of the FMA ontology is in OWL DL language that include metaclass knowledge, thanks to OWL 2 new metamodeling features. In addition, unlike the previous FMA-OWL 1 (DL) realization [22], this FMA version possesses additional definitions of FMA classes and axioms based on lexical patterns, which are semantically correct and reliable from an anatomical viewpoint. Having in mind very specific application domain, employment of other reasoners (such as HermiT [63]) on OWL 2 version of the OBR-Scolio ontology, and corresponding analysis of the reasoning results will be carried out in order to improve overall performance of the OBR-Scolio ontology.

6. Conclusions

OBR-Scolio ontology initially has been developed using frame-based representation framework, since its key cornerstones are the FMA and the OBR reference ontologies that play crucial role in medicine. That is, the OBR-Scolio’s modeling approach extracts from both relevant entities of biomedical reality in the domain of spinal deformities, and consequently inherits the same representation. Its key intention is to bring together and merge essential anatomical and pathological concepts and knowledge that intrinsically supports faithful clinical communication in diagnosis and treatment of spinal disorders, primarily scoliosis. However, frame-based representation demonstrates numerous deficiencies that fundamentally stems from poor or no semantics. Since reliable and faithful clinical communication and decision making assumes high level of precision and relevancy, unambiguous and
consistent information and data integration is absolutely indispensable. Therefore, we have chosen OWL DL language as it is based on description logic modeling framework that provides highly expressive capabilities of semantic reasoning.

The Protégé frame-based mode has an “export to OWL” option. However, this option only performs a straightforward translation that ignores all the features that do not have a direct equivalent. This paper presents detailed additional steps needed to transform the OBR-Scolio ontology into OWL DL representation using Protégé ontology editor and knowledge acquisition system that supports both formalisms. Although virtually similar, these two approaches rely on fundamentally different modeling assumptions. The differences between description logics and frames are not only syntactic, but also semantic. Frames’ semantics is not as precisely defined as description logics’ one. A description logics representation of the OBR-Scolio application ontology in the OWL DL would allow developers to combine it with other OWL ontologies and to employ advanced inference capabilities: satisfiability, subsumption, classification, consistency checking, instantiation, realization and retrieval using generic reasoning tools. All these contribute to maintenance of a consistent terminological system and improve results of queries.

The conversion process from frame-based to OWL DL ontology representation is accomplished using the Export function in Protégé and by applying many additional rules that include: deleting all information of metaclass instantiation, converting Protégé slots into OWL DL properties and converting Protégé classes in OWL DL classes. The result of conversion process provides ontology specification completeness, semantic consistency, and possibility to employ reasoners for subsumption testing and automatic checking for consistency, ambiguity, and redundancy in classes’ hierarchy and specification. For the reasoning purpose over OBR-Scolio ontology we have used the Racer, a DIG compliant reasoner. The Racer was launched from Protégé-OWL, but the classification processing has failed because Racer could not handle the entire OBR-Scolio OWL ontology. Due to the presence of complex restrictions on inverse pairs of properties such as: part, part_of and systemic_part, systemic_part_of checking the consistency and computing inferred classes hierarchy were successful only without these properties.

The OBR-Scolio ontology is aimed as guide for scoliosis professionals and for other medical staff in health institutions and also for educating orthopedic students. OBR-Scolio ontology is due to its OWL-DL representation currently employed in ScolioMedIS, a web-based information system for visualization and monitoring of idiopathic scoliosis [19]. The system provides innovative programming access to orthopedics and physicians for convenient and reliably diagnose, treat and monitor patients with idiopathic scoliosis.

**Acknowledgement.** This research work is supported by the Serbian Ministry of Science and Technology under the grant III-41007: “Application of Biomedical Engineering in Preclinical and Clinical Practice”.
Converting OBR-Scolio Ontology in OWL DL

References


Vanja Luković is teaching and research assistant at Faculty of Technical Sciences in Čačak, University of Kragujevac (Serbia), since 2001. She received the Master degree at School of Electrical Engineering in Belgrade, Serbia in 2007. Currently she is working on her PhD theses in the area of ontological modeling in medicine at Faculty of Technical Sciences, within the project “Application of Biomedical Engineering in Preclinical and Clinical Practice” of the Ministry of Education, Science and Technological Development of Republic of Serbia. Her main research interests include development and evaluation of biomedical ontologies, object-oriented programming and information systems design. She has authored/co-authored more than 20 papers.

Danijela Milošević is associate professor at the Faculty of Technical Sciences in Čačak, University of Kragujevac, Serbia. She holds a PhD in Computer Science and her research interests include ontologies, knowledge representation, intelligent tutoring systems, user modeling and bioengineering. Author/co-author of over 70 research papers, two textbooks, and has been actively involved in a number of research projects addressing e-Learning tools and technologies, adaptive instructional design and development of user model ontologies. She is a member of IEEE computer society.

Goran Devedžić is professor at Faculty of Engineering, University of Kragujevac, Serbia. His research interests focus on the advanced product and process development, industrial and medical application of soft computing techniques, and bioengineering. He has authored/co-authored more than 100 research papers, published in international and national journals or presented at international and national conferences, as well as three books on CAD/CAM technology and 3D product modeling.

Saša Ćuković is teaching and research assistant and PhD candidate at Faculty of Engineering, University of Kragujevac, Serbia. He was scholarship holder of Ministry of Education, Science and Technological Development of Republic of Serbia and DAAD grant holder at Technical University of Munich. His main research interests include CAD/CAM-systems, reverse engineering and noninvasive 3D reconstruction and modeling in engineering and medicine, augmented reality and computer vision. He has authored/co-authored more than 20 papers.

Received: June 11, 2012; Accepted: January 15, 2013