SBVR Based Representation of SPARQL Queries and SWRL Rules for Analyzing Semantic Relations

Algirdas Sukys, Lina Nemuraite, Bronius Paradauskas, Edvinas Sinkевичius
Kaunas University of Technology
Kaunas, Lithuania
sukys.algirdas@gmail.com, lina.nemuraite@ktu.lt, bronius.paradauskas@ktu.lt, edvis.s@gmail.com

Abstract—Analyzing semantic relations is a cumbersome task because these relations often are distributed over different information sources and hidden in existing relational database structures. Even with Semantic Web ontologies describing semantic relations we need to explore them on the deep technological level that is not friendly for business users. Semantics of Business Vocabulary and Business Rules (SBVR) gives a possibility of representing OWL 2 ontologies, SWRL rules and SPARQL queries using concepts and semantic formulations expressed in SBVR Structured English language understandable for business users. We suggest formulating derivation rules and queries for analyzing semantic relations as SBVR rules and questions, and transforming them into SWRL and SPARQL.

Keywords—semantic relations; SBVR; SPARQL; OWL 2; SWRL.

I. INTRODUCTION

More and more business information becomes available in form of ontologies that are accessible to users by dedicated query languages as SPARQL [24][32]. However, users prefer querying using sentences in natural language [10]. Query languages are limited, therefore complex parsing and validating means are needed for ensuring flexibility of queries that should allow using synonyms and synonymous forms, and asking in various ways. Ontology reasoners could help to query using expressions that are not directly defined in ontology but could be derived on the base of existing formulations.

Semantics of Business Vocabulary and Business Rules (SBVR) is the OMG metamodel that defines the vocabulary and rules for describing the business semantics – business concepts, business facts, and business rules using SBVR Structured English (SSE) or other Controlled Natural Language [21]. The SSE language cannot represent all possible constructions of natural English language, however, it is understandable to business and information system experts, and is interpretable by computers. SBVR directly models business meaning and can support many controlled languages; consequently, it can be seen as an interface between real natural languages and business software systems [11].

SBVR presents a high-powered metamodel but a lot of work should be done for introducing SBVR based business semantics management into the enterprise architecture. SBVR is capable to represent the human understandable semantics beyond Web Ontology Language (OWL 2), Semantic Web Rules (SWRL), OWL for Web Services (OWL-S), Web Service Modeling Ontology (WSMO), and other ontological languages that were acknowledged as most expressive semantic representations before SBVR. SBVR directly models meaning and offers possibility to relate business semantics to software models and implementations; however, such a relation could be made through multiple transformations between various languages and architectural layers.

Despite the interest from the researchers and business side, until now only part of SBVR was used. SBVR questions provide a capability of querying business models and their implementations but they attained only little attention in the SBVR specification and related research. Kriechhammer in 2006 has noticed about possibilities of SBVR questions [12] for business people to query systems for business modeling without the support of programmers. To our knowledge, no further research in that direction was done. In [29], we applied the idea and presented the initial methodology for transforming SBVR questions into SPARQL.

In the current paper, we analyze further possibilities of querying business systems by supplementing previously described querying capabilities with the use of SWRL derivation rules together with OWL 2. We concentrate on querying some types of semantic relations: reflexive relations causing hierarchical links between individual concepts, and n-ary relations that cannot be directly represented in ontology and, consequently, are problematic for querying with SPARQL if we want to obtain SPARQL queries from SBVR questions acceptable for human.

The rest of the paper is organized as follows. Section 2 presents related works. Section 3 presents SBVR questions, corresponding SWRL rules and SPARQL queries for hierarchical relations, Section 4 – for relation properties allowing defining n-ary relations in OWL 2. Section 5 draws conclusions and outlines the future work.

II. RELATED WORK

SBVR metamodel and XMI schema may be used for developing software tools for managing business vocabularies as well as for automating development of software for managing business on the base of business semantics, i.e., in the way different of previously existed approaches, e.g., [18][22]. SBVR business vocabularies are
transformable into UML & OCL [3][17] and vice versa [2]; BPMN [1], RDB schemas [16], OWL [5], Web services [6], [9] etc. Besides automating the development of software models and code [14][15], SBVR Structured English may serve for creating semantic specifications of legacy information resources, integrating these resources, implementing contextualized and multilingual information systems, etc. The power of SBVR is disclosed by the fact that SBVR specification itself is formally written in SSE [13].

For practical usage, SBVR suffers from various limitations and the lack of tools as editors, validation mechanisms and inference engines. It would be desirable having the larger collection of data types and patterns for constructs needed for expressing arithmetic operations, data and time, past and future events, and similar. Spreeuwenberg and Anderson notice more deficiencies of SBVR: lack of inference; lack of references (rules should be stated in one sentence); necessity to introduce concepts before referencing to them; impossibility to express directives, etc. [26][27]. Individual researchers and their groups have proposed various SBVR extensions, but these extensions have not resulted in a new version of SBVR specification yet.

Realizing the idea of querying business in SBVR requires a creation of a whole infrastructure including tools for authoring SBVR business vocabularies and rules, transforming them into various software models and code, including OWL, SQL, Web Services, business process execution languages, and so on. Several EU projects are devoted for this purpose: OPAALS (2006–2010, generating Web services and data models from SBVR specifications [16][25]), ONTORULE (2009–2012, aiming at integrating knowledge and technologies needed for extracting ontologies and business rules from various documents, including natural language texts; managing them and implementing in software systems). The commercial tool suite for Business Semantics Management Collibra presents capabilities for authoring SBVR vocabularies and rules, generate ontologies and various models of information systems.

In our previous work, we focused on generating UML&OCL models from SBVR specifications. We proposed the methodology for specifying information system requirements on the base of SBVR business vocabularies and business rules related with business process models, and implemented the prototype of tool VeTIS supporting that methodology [3][17]. VeTIS tool recognizes SBVR concepts (object types, roles, fact types, fact type roles, individual concepts) and business rules (various kinds of semantic formulations) that make foundations for conceptualizing business and correspond to knowledge and metaknowledge level [19]. For representing SBVR questions, the extension of the VeTIS tool was needed for including business facts – instances of fact types (ground facts) and propositions (instances of complex semantic formulations based on several fact types) comprising the bottom knowledge level – fact level [20]). Representing business facts is crucial for implementing the semantic enterprise where you are able not only tracing business requirements to their implementations in software, but also accessing business data using a single language and terminology. In [29], we considered the way of transforming SBVR questions into SPARQL queries that can be executed against OWL 2 ontologies obtained from SBVR vocabularies and business rules.

SBVR questions are based on logical projections and are much more comfortable for business people than various query languages that are platform-specific and intended for IT specialists. We can apply several slightly different methodologies for querying: to derive relations using SWRL rules and OWL 2 reasoning tools, and querying using asserted as well as inferred individuals and properties, or formulating direct SPARQL queries. Also, we can store ontology instances in a relational database and apply gradual querying methodology by formulating part of a query in SPARQL and retrieve instances from a relational database using SQL [30].

Here, we will analyze how to translate SBVR questions and business rules into corresponding SWRL and SPARQL expressions. We will concentrate on two aspects: querying facts based on hierarchical and n-ary relations. Taxonomy of semantic relations analyzed by Chaffin [4] and Storey [28] includes seven main relation types: inclusion (class inclusion, meronymic, spatial), possession, attachment, attribution, antonyms, synonyms, and cases. Some meronymic, possession, and attachment relations can arise between objects of the same type. Typical reflexive relations are kinship relations that can comprise complex trees and forests; their analysis requires recursive queries.

N-ary relations have no problems for representing them in UML or SBVR, but they cannot be directly represented in ontology. There are solutions proposed by the W3C for representing n-ary relations in OWL [31] but they are based on the creation of new object types that often seem unnatural for formulating related queries close to natural language. We agree with Hoekstra who argues that the disallowance of n-ary relations is rather an advantage than the drawback of the OWL 2 because n-ary relations can be expressed by binary ones [8]; the presence of n-ary relations demonstrates that the meaning of the corresponding associations is not well-understood. Furthermore, if n-ary relations are un-ordered the semantics of relation can be lost because we may miss the subject of the sentence.

In SBVR and UML, fact type roles and association ends are ordered. However, for keeping the required order of fact type roles in SBVR, one should respectively formulate fact types; in UML, the order of n-ary association ends is set by the order of introducing these ends into a model – this is not always understood by modelers. Our proposed way of representing n-ary relations by binary ones in OWL 2 would be possible if we could define properties for relations (properties of OWL 2 object properties). Then we could distinguish fact type roles of subject and object as roles of a binary relation, and the remaining fact type roles as properties of that relation. In Section IV, we show how to define the SBVR fact types with synonymous forms that allow formulating SSE queries, acceptable for humans, and how to represent these queries in SPARQL by using OWL 2
punning for defining classes and object properties having same IRIs.

III. ANALYZING HIERARCHICAL RELATIONS BETWEEN INDIVIDUAL CONCEPTS

Reflexive relations arise between objects of the same type playing different roles in the relation. For example, in SBVR terminology, fact type “father has son” represents associative fact type where fact type roles are played by different persons that are instances of object type “person” (here and in the following, we will use SBVR style “term” for terms, representing noun concepts; verbs, representing fact type symbols; “Names” for individuals, and “keywords” for keywords, articles etc making SSE sentences more understandable). We take a genealogy tree as an example of reflexive relations. OWL 2 ontology of a genealogy tree is presented in Figure 1 in UML notation where the association ends correspond to OWL 2 object properties and SBVR fact types e.g., has_son (not fact type roles e.g., “son” as it should be in actual UML models).

In the tree ontology, the asserted classes are “Person”, “Marriage” and “Sex”; they have asserted data properties and object properties “has_parent”, “has_sex”, and “has_marriage”. Remaining classes and object properties can be inferred from OWL 2 axioms and SWRL rules, e.g., “Man” can be defined as “Person” who has sex of male:

```
EquivalentClasses( Man Person (ObjectHasValue(has_sex value male_sex)))
```

A couple is defined as an object property and can be derived by SWRL rule:

```
DifferentFrom(?x,?z)→couple(?x,?z)
```

We also use OWL 2 punning and create a class “couple” that is understood as a pair of male and female that have at least one child and, optionally, may be married. Figure 2 presents individuals that are analyzed in the following.

For defining an object property “has_kin” we can use the classical SWRL rules:

```
has_parent(?x,?y)→has_antecedent(?x,?y)
has_parent(?x,?y),has_antecedent(?y,?z)→
has_antecedent(?x,?z)
```

That means kin of a person are her antecedents and descendants, as well as descendants of her antecedents except the person herself. We can define rules for all object properties of the genealogy tree in a similar way.

![Figure 1. OWL 2 ontology of the genealogy tree in UML notation](image-url)
Representing SWRL rules in SBVR is straightforward with the use of implication formulation. On the contrary to SWRL, SBVR rules can use disjunctive logical formulations that should be transformed into separate SWRL rules:

It is necessary that person1 has_antecedent person2 if person1 has_parent person2.

\[
\text{...}
\]

It is necessary that person1 has_kin person2 if person1 has_antecedent person2 or person1 has_descendant person2 or person1 has_antecedent person3 and person3 has_descendant person2 and not person1 equals person2.

Using OWL 2 reasoner (Pellet or Hermit), we can infer kin of all persons and formulate simple SBVR questions as:

“What are kin_of person Vita_Binkiene?”

“Are kin person Ema_Griniene and Vytas_Grinius?”

that are translated into simple SPARQL queries:

**SELECT DISTINCT ?kin**

{gen:Vita_Binkiene gen:has_kin ?kin},

which gives the results presented in Figure 3. For executing queries, we have used the ARQ 2.8.4 query engine that supports SPARQL 1.1 extensions, and Pellet 2.2.2 OWL Reasoner.

**Figure 3.** Results of the query “What are kin_of person Vita_Binkiene?”

The second query is of ASK type (the “kin” is the synonymous noun form of “has_kin”):

**ASK**

{gen:Vita_Binkiene gen:has_kin ?kin}

and gives the result false.

Without the use of reasoner, formulating SPARQL queries from SBVR questions can become complicated as SPARQL recursive queries are problematic to identify from SBVR rules:

**select distinct ?kin(select ?antecedent( gen:Vita_Binkiene gen:has_parent* ? antecedent . not exists (? antecedent gen:has_parent ?x))).**

**filter(?kin != gen:Vita_Binkiene)**

Another way of transforming SBVR questions into SPARQL is to formulate derivation rules in CONSTRUCT query form [23]. However, CONSTRUCT queries inefficiently work with large rule sets. Therefore we transform SBVR rules into separate CONSTRUCT queries that give results in separate triple graphs, and apply Jena Union function that dynamically relates result graphs through their common elements. Finally, we execute SELECT query in the united graph:

**CONSTRUCT {?x gen:has_antecedent ?y .} WHERE{?x gen:has_parent ?y .}**

... **CONSTRUCT {?x gen:has_kin ?y .} WHERE {?x gen:has_antecedent ?y .}**

**CONSTRUCT {?x gen:has_kin ?y .} WHERE {?x gen:has_descendant ?y .}**

**CONSTRUCT {?x gen:has_kin ?z .} WHERE {?x gen:has_antecedent ?y . ?y gen:has_descendant ?z . filter(?x != ?z)}.**

**SELECT DISTINCT ?kin {gen:Vita_Binkiene gen:has_kin ?kin}.**

The SELECT gives the same results as in Figure 3.

**IV. DEFINING N-ARY RELATIONS IN OWL 2**

SBVR metamodel allows formulating fact types with more than two fact type roles and preserves the ordering of these roles. However, for further retaining the meaning of relations it is desirable to express them as binary relations because in the fact type there always are two main fact type roles – subject and object; the remaining fact type roles mean either properties of the fact type relating subject and object, either properties of object. We should formulate fact types in such a way that it would be possible to identify what fact type roles mean properties of fact type, and what ones mean properties of object. That is, instead of defining SBVR fact type as

“person works_in organization in_position_from_date till_date”,

it is desirable to represent it in such a way:

“person works_in organization_from_date till_date”.

Synonymous form: position,

“person occupies_position_of_position”,

“position performs_office_of_office”.

Here the last fact type is introduced for representing object type “office” that is a role type of the role “position”. Figure 4 presents example ontology for modeling properties of relation in OWL 2. Role “Position” is explicitly presented as a class. Dependency with stereotype **<EquivalentClasses>>** means that EquivalentClasses axiom
is defined for classes “works_in” and “Position”. We cannot define properties of relations (i.e., object properties) in OWL 2, but we can use punning and define a class with the same name as the object property “works_in”. For flexibility of formulating propositions and questions, we create the additional class “Position” equivalent to class “works_in”, and attribute properties of the relation for that class.

![Figure 4. Example of modeling properties of relations in UML.](image)

Guizardi and Wagner suggested the similar solution [7] (without the role type) for representing properties of relations in UML despite that UML has the association class for this purpose. They have criticized the ambiguity of UML in UML despite that UML has the association class for representing a formal relation, and a separate class for materializing that relation.

Having such ontology, we can simply ask

“What are organizations that person Jonas Grinius works_in from_date 2000-01-01 till_date 2003-01-01?”

Here we cannot use current reasoners (Pellet, HermiT) because they do not derive relation properties on the base of punning but we can attain these properties using SPARQL. For this, we should supplement the question with a business rule

“It is necessary that person works_in organization from_date till_date if the person occupies_position_of position from_date till_date and the position is_employed_in the organization.”

Then we instantiate the rule for person “Jonas Grinius” and transform the question and the rule into the SPARQL query (bound is added to optional variables):

```sparql
select distinct ?org
?position org:from_date ?from .
optional {?position org:till_date ?till} .
filter (?from <= "2003-01-01"^^xsd:date &&
?till >= "2000-01-01"^^xsd:date ||
!bound(?till))
```

For the given ontology example of organizations and persons we also can define the OWL 2 object property chain:

```
SubObjectPropertyOf(ObjectPropertyChain
occupies_position_of performs_office_of has_office_of)
```

that is represented as SBVR structural business rule:

It is necessary that person has_office_of office if the person occupies_position_of position and the position performs_office_of the office.

Current OWL 2 reasoners understand object property chains. As previously, we will consider two cases: 1) when we use a reasoner and formulate a simple question; 2) when we do not use a reasoner and formulate a question together with derivation rule.

Using reasoner, we can give the SBVR question “What person has_office_of Professor” that transforms into the SPARQL query:

```
SELECT ?persons
(?persons org:has_office_of gen:Professor) .
```

SBVR question without reasoning

“What person occupies_position that performs_office_of Professor?” transforms into SPARQL query:

```
SELECT ?persons ?position
?position org:performs_office_of org:Professor
```

V. CONCLUSIONS AND FURTHER WORKS

The paper presented some formulations of SBVR fact types and questions for recursive and n-ary relations, together with representing them in OWL 2, SWRL and SPARQL, and executing obtained queries in current ontology reasoning tools and SPARQL engines. For analyzing recursive relations, we represent them by SBVR derivation rules and propose two scenarios for transforming them: 1) using SWRL rules, when we obtain simple SPARQL queries but should apply OWL 2 reasoner in addition to SPARQL engine; 2) using sequences of SPARQL CONSTRUCT queries. Both cases are practicable in real life applications. For analyzing n-ary relations, we propose to represent them by using SBVR synonymous forms. Transformation of these forms into OWL 2 results in introducing a new class and a binary relation with the same name (allowed by OWL 2 punning) for defining the main relation between subject and object; and object properties of that class representing remaining roles of n-ary relation. Such representation allows avoiding unnatural names that appear in traditional solution.

The limitations of the approach are in the fact that OWL 2, even supplemented with SWRL, is not enough to model the complete semantics of SBVR vocabularies, and, consequently, SPARQL is not enough to model SBVR questions. From the other side, it still is not clear how to represent some OWL or SPARQL constructs in SBVR in a natural way. Our future work will focus on analysis of more patterns for improved flexibility of SBVR questions translatable to SPARQL queries as well as providing experiments for integrating the proposed solutions into the realistic enterprise context. Also, we have some initial prototype of SBVR editor for Lithuanian language and are willing to work for embodying the SBVR Structured Lithuanian.

REFERENCES


