Journal of Theoretical and Applied Information Technology
20th February 2014. Vol. 60 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PATTERN REPRESENTATION MODEL FOR N-ARY RELATIONS IN ONTOLOGY

[1]VINU P.V., [2]SHERIMON P.C., [3]RESHMY KRISHNAN, [4]YOUSSEF SAAD TAKRONI
[1, 2] Research Scholar, M.S. University, India
[3]Associate Professor, Muscat College, Oman
[4]Assistant Professor, Arab Open University, Oman
E-mail: [1]vinusherimon@yahoo.com , [2]sherimonpc@yahoo.com , [3]reshmy_krishnan@yahoo.co.in , [4]yst@aou.edu.om

ABSTRACT

The vision of Semantic Web is to make web resources more accessible to automated resources. Here the role of ontology is to provide vocabulary for metadata description with computer-understandable semantics. The main components of ontology are concepts, relations and individuals. The most common type of relation is binary relation that maps between a single subject and a value. Sometimes there exist n-ary relations in ontology. W3C provides several patterns to represent n-ary relations. In this paper we discuss the issues in n-ary relations, the concept of RDF reification and provide an appropriate pattern to represent the n-ary relations. The examples of n-ary relations are taken from Seafood Ontology we developed earlier.

Keywords: *N-Ary Relations, Reification, Semantic Web, OWL, RDF*

1. INTRODUCTION

The Semantic Web is an alternative to the current World Wide Web, which primarily presents data as HTML pages that are easily interpretable to human eyes [5]. It is all about making web data machine readable. Currently in Web, the data present in HTML pages are connected to each other using hyperlinks. In contrast, Semantic Web is collection of semantically defined data that can generate dynamic web pages easily understandable by both machines and humans. Here more concentration is given on the data itself rather on the presentation of the data. In Semantic Web the information contained in documents is given an explicit meaning, making it easier to be processed by applications [8]. Web Ontology Language (OWL) [8] can be used to represent the meaning of terms and the relationships between those terms. It is more expressive than XML [Extensible Markup Language], RDF [Resource Description Framework] and RDF-S [Resource Description Framework Schema], making it easier to represent machine interpretable content on the Web [8, 9].

A well-known limitation of OWL 2 is that only binary relations between classes can be represented [1, 2, 3, 4]. It is not possible to represent a relation which involves more than two dimensions in OWL. But using the concept of reification, n-ary relations can be represented in RDF. RDF provides a built-in vocabulary intended for describing RDF statements. A description of a statement using this vocabulary is called a reification of the statement. RDF directly supports reification. But we don't want to use RDF reification, since it stores additional information about a triple, thus affecting the reasoner engines. So currently W3C is not supporting the concept of reification. In this paper, we discuss the issues in modeling n-ary relations in Seafood Ontology using reification and implement an alternate pattern to represent these relations. All the case studies are taken from Seafood Ontology [11].

The remainder of the paper is organized as follows: - Section 2 explains the ontology languages followed by the description of n-ary relations in Section 3. The issues in n-ary relations, reification and its drawbacks are explained in this section. In section 4, an appropriate pattern to represent the n-ary relations are outlined followed by the implementation in Protégé in Section 5. Conclusion and Future is written in Section 6.

2. ONTOLOGY LANGUAGES

Ontology languages allow users to write explicit, formal conceptualizations of domain models. XML, RDF, RDFS, OIL, DAML+OIL and OWL are the basic ontology languages for Semantic Web [10]. XML is a tag based language that allows users to define their own tags to describe the structure of a document [15]. But it

Journal of Theoretical and Applied Information Technology
20th February 2014. Vol. 60 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

failed to describe the semantics of the document. An XML application, RDF is the next data model used by the Semantic Web based on triples, <subject, predicate, object> [5]. For example, <Protégé, *hasCreator* Stanford University> RDF uses URLs [Uniform Resource Locators] to identify the resources on Web and a form of XML, RDF/XML is used to represent and process the information on Web. RDF is the basic framework on which RDFS, OWL, and the Semantic Web are based [5]. RDF data are often stored in databases so-called RDF stores, and are retrieved using the query language, SPARQL [SPARQL Protocol and RDF Query Language] [5]. RDFS is intended as a framework for interpreting the meaning of data expressed in RDF [5]. It extends the vocabulary of RDF. Later more expressive ontology languages such as OIL, DAML+OIL and OWL were developed. OWL extends RDFS and has more powerful inference rules and built-in constructs [5].

OWL 2.0, an extension of OWL, published in 2004 became aW3C (World Wide Web Consortium) recommendation in 2009 [12]. In OWL, the basic unit of ontology is a class, which represents a set of individuals, and its properties, which represent (binary) relations between individuals [7]. OWL (Web Ontology language) supports two different types of properties – data properties and object properties. An object property specifies a relation-ship or link between two individual objects. Data properties relate individuals to predefined data type values or literals. Predefined data types include types such as strings, integers, Boolean etc. Consider two classes namely seafood and country. An object property '*hasCountryOfExport*' links instances of seafood class to one or many instances of the country class. A seafood individual may be linked by a data property '*hasName*' to a string value that represents the name of the seafood. All these properties represent the relationship between two individuals. OWL provides three expressive sub-languages, OWL Lite, OWL DL, OWL Full. The major limitation of OWL is no direct support for n-ary relationships.

3. N-ARY RELATIONS

The most common type of relation is binary relation that maps between a single subject and a value. But sometimes, we have to define such relationships that hold between one subject and two or more values. Relations that link an individual to more than one individual or value are called N-ary relations. An n-ary relation on sets $A_{1,}A_{2,}A_{3,} \ldots \ldots \ldots . . A_{n,}$ is a set of ordered n-tuples, $a_1, a_2, a_3, \ldots \ldots \ldots . . a_n$, $a_i \in A_i$ for all i, $1 \leq i \leq n$.

For example, consider the sentence: Land Cruiser is a four wheeler type vehicle which is manufactured by Toyota Corporation, founded in 1937. Many individuals are involved in this relation – 'Land Cruiser', 'Four Wheeler', 'Toyota Corporation', and 1937.

3.1 Issues in n-ary Relations

In Semantic Web languages, such as RDF and OWL, a property is a binary relation: in-stances of properties that link two individuals [3]. RDF and OWL only allows defining properties to link two individuals at a time. So if properties can link only two individuals how can be represent a relation that involves more than two individuals (n-ary relation-ships)? An n-ary relation cannot be split up into 'n' binary relations, because the relations it defines are all interconnected in some way [6].

For example, consider the scenario from seafood ontology. 'Sardine is exported to China and it requires Cadmium and Salmonella test'. Here there is a relation in which the Fish 'Sardine', the Country, 'China' and the TestType 'Salmonella' and 'Cadmium' participate. This is a 3-ary relation because it involves the individuals that belong to Fish, Country and TestType classes. The above sentence is split into two parts:-

'Sardine is exported to China'

'China requires Cadmium and Salmonella Test'

These relations are modeled in Figure 1. But this type of modeling is incorrect. It can be proved using set theory as follows:-

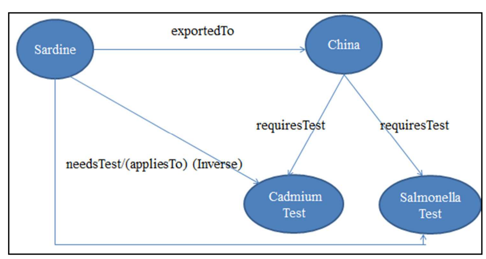

Figure1. Test Specification of Sardine to China

Let C, F, T and TS represents the set of countries, fishes, tests and test specifications.

C = $\{C_1, C_2, C_3, \ldots \ldots \ldots . . C_i$

F = $\{F_1, F_2, F_3, \ldots \ldots \ldots . . F_l\}$

T = $\{T_1, T_2, T_3, \ldots \ldots \ldots . . T_i\}$

Journal of Theoretical and Applied Information Technology
20th February 2014. Vol. 60 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$TS \subset F \times C \times T \quad (1)$$

$$TS_i \in F \times C \times T \quad (2)$$

Consider a mandatory test, $x_i \in T$ required for Sardine fish to export to China, then

$$x_i \in P \cap Q \quad (3)$$

where P is the set of all tests required for China and Q is the set of all tests required for Sardine fish. It is modeled in Figure 1.

Now consider a mandatory test, $y_i \in T$required for Sardine fish to export to Japan, then

$$y_i \in R \cap Q \quad (4)$$

where R is the set of all tests for Japan. Refer Figure2.

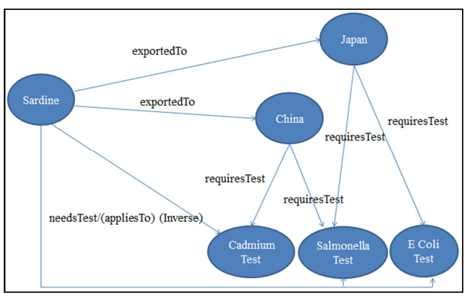

Figure2. Test Specification of Sardine to China and Japan

Assume that there is a mandatory test 'a' for exporting Shrimps to China. So,

$$a \in P \quad (5)$$

Also assume that the same test 'a' is specified as mandatory for Sardine Fish to export to Japan. So,

$$a \in Q \quad (6)$$

Now from (5) and (6), we know that

$$a \in P \cap Q \quad (7)$$

From (3) and (7), it derives that the test 'a' now becomes a mandatory test for Sardine Fish to export to China. But according to (3),

$$a \notin x_i \quad (8)$$

For example, the mandatory tests for China for exporting Sardine Fish were only Cadmium Test and Salmonella Test according to Figure 1. But now there is an additional test 'EColi' which is False since the test 'EColi' was not listed earlier under the mandatory tests of Sardine to China. This is shown in Figure3.

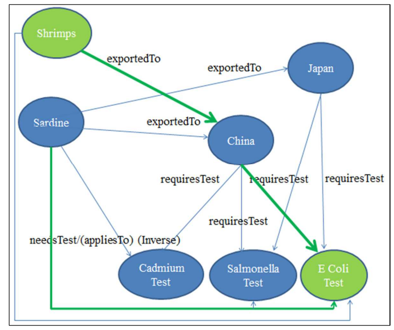

Figure3. Test Specification of Sardine and Shrimps

So it is proved that we cannot split an n-ary relation into 'n' binary relations. This means that an n-ary relation can only be represented in its reified form in OWL [1].

3.2 Reification

Reification is a concept in knowledge representation community where one statement becomes the subject of next statement. That is, a statement is made about a statement. It is a general purpose technique for representing n-ary relations using a language such as OWL that permits only binary relations [2]. This is useful for representing long sentences. RDF directly supports reification. The RDF/OWL approach to n-ary relations is to map those using binary relations, by creating an intermediate entity that serves as the subject for the entire set of relations; this entity is then in turn made the object for a relation in which the main subject is the subject [6].

For example, consider the sentence: Land Cruiser is a four wheeler type vehicle which is manufactured by Toyota Corporation which was founded in 1937. The above sentence can be split into 3 different sentences.

Statement 1 →(Land Cruiser, type, four-wheeler)

Statement 2 → (Land Cruiser, manufacturer, Toyota Corporation)

Statement 3 →(Toyota Corporation, foundedYear, 1937)

When we use reification, Statement 2 and Statement 3 can be expressed as follows:-

Statement 2 → (Statement 1, manufacturer, Toyota)

Journal of Theoretical and Applied Information Technology
20th February 2014. Vol. 60 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Statement 3 → (Statement 2, foundedYear, 1937)

One solution to the problem of modeling n-ary relations explained in Section 3.1 is applying reification. Consider the same Seafood ontology example. 'Sardine is exported to China and it requires Cadmium and Salmonella test'

A new class is reified to model the n-ary relationship. The relationships of each element to the reified class are defined. Applying the concept of Reification, the equation (2) for test specification is represented as

$$TS_i \in (F \times C) \times T . \qquad (9)$$

Since Cartesian products are non-associative, our constraints are satisfied. So in the above example, exportedTo is the reified relation.

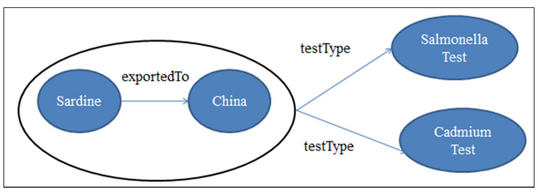

Figure4. Reification

3.3 Drawbacks of Reification

According to W3C, these types of complicated statements affect OWL reasoners. The major drawbacks of reification are increased no. of triples in the datasets leading to in-creased storage. It increases the complexity of the ontology [4]. A potential disadvantage of RDF reification is that there is no connection between the original statement and the reified statement. If one is modified, the other is not automatically modified. This leads to a situation where reified statement are reifying triples that are not actually asserted, and vice versa. So now W3C is not supporting reification.

4. ALTERNATE PATTERN

W3C has suggested several patterns to solve the problem of representing n-ary relations. One common solution is to represent the relation as a class rather than a property [13]. Individual instances of such classes correspond to instances of the relation [13]. Specifically, an n-ary relation is represented as a new object that has all the arguments of the n-ary relation as objects of properties [2]. An individual is created for representing the relation instance and all participants are linked to it.

To represent the seafood ontology example, a new class named Test Specification is created and 3 properties are defined to represent the 3-ary relation. It is represented in Figure5.

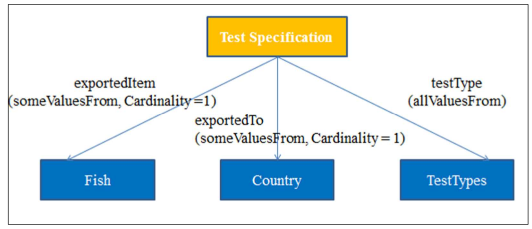

Figure5. Representing 3-ary relationship in Seafood Ontology

Restrictions describe the constraints on relationships that the individuals participate in for a given property. OWL provides three types of value restrictions for specifying the range of a property when it is used with an instance of a particular class: owl:allValuesFrom, owl:someValuesFrom, and owl:hasValue [14]. Existential restrictions ∃ or someValuesFrom restriction describe the set of individuals that have at least one specific kind of relationship to individuals those are members of a specific class. Universal restrictions ∀ or allValuesFrom describe the set of individuals that, for a given property, only have relationships to other individuals that are members of a specific class. hasValue restriction implies that for all instances, they must have an occurrence of the property with the specified value [14]. Cardinality restrictions specify the exact number of relationships that an individual must participate in for a given property.

The restrictions applied to Test Specification class are:

∃ exportedItem	Fish (Cardinality = 1)
∃ exportedTo	Country (Cardinality =1)
∀ testType	Test (minCardinality =1)

The above restrictions state that every individual of TestSpecification class must have at least one relationship with a member of the corresponding classes. This is the open world assumption (OWA). So each test specification consists of exactly one item [instance of Fish class], exactly one country [instance of Country class] and all the specified Tests [instance of Test].

Journal of Theoretical and Applied Information Technology
20th February 2014. Vol. 60 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

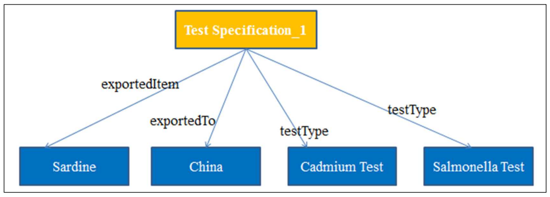

Figure6. Instance of Test Specification class

Test Specification_1 is an individual instance of the Test Specification class represent-ing an instance of a relation:

In Turtle format the class is represented as follows:-

: Test Specification_1

a : Test Specification;

: exportedItem Sardine;

: exportedTo China;

: testType Cadmium Test;

: testType Salmonella Test;

5. IMPLEMENTATION

Figure 7 shows the implementation of a subclass Test-Spec-UncookedFrozen-Crustaceans-China of the class Test Specification in Protégé. The values of the object property exportedItem can be one of the individuals of the Crustaceans class. The constraint defined for the exportedItem property is someValuesFrom with cardinality 1. There are many tests required for 'Crustaceans' to export to 'China'. All these mandatory tests are added to the testType object property and the constraint defined for the property is allValuesFrom.

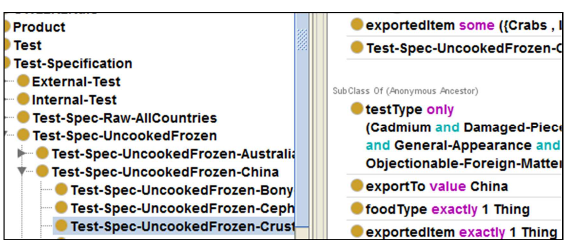

Figure7 Class Implementation in Protégé

6. CONCLUSION AND FUTURE

Ontology provides a shared and common understanding of a domain that can be communicated across people and application systems. Ontologies are able to define relationships, semantics, enhanced clarity, all of which collectively enable information retrieval in a meaningful way. Web Ontology Language (OWL) is used to represent the meaning of terms and the relationships between those terms in ontology. A well-known limitation of OWL 2 is that n-ary relations cannot be represented. The concept of reification was suggested in RDF to represent n-ary relations. But according to W3C, RDF reification is not a good practice to represent n-ary relations. Alternate patterns are suggested by W3C. This paper focusses on the issues of n-ary relations in Seafood Ontology and represents the same using a pattern suggested by W3C. In future we are planning to propose a new pattern to represent n-ary relations in Semantic Web.

.

ACKNOWLEDGEMENT

This work is published as part of a project funded by The Research Council [TRC], Oman under Agreement No. ORG/ AOU/ ICT/ 11/ 015, Proposal No ORG/ICT/11/004 and Arab Open University, Oman Branch

REFRENCES:

[1] Rinke Hoekstra. Ontology Representation – Design Patterns and Ontologies that Make Sense, volume 197 of Frontiers of Artificial Intelligence and Applications. IOS Press, Amsterdam, June 2009.
[2] Batsakis S., Petrakis E.: Representing and Reasoning over Spatio Temporal Information in OWL 2.0, 6th Workshop on Semantic Web Applications and Perspectives (SWAP' 2010), Bressanone, Italy, September 21-22, 2010.
[3] Natasha Noy, Alan Rector, Defining N-ary Relations on the Semantic Web, W3C Working Group Note 12 April 2006
[4] Robert Stevens, Phillip Lord (2010) Reification of properties in an ontology. Ontogenesis.
[5] Sebastian Bauer, Introduction to Bio-Ontologies, Chapman and Hall/CRC 2011, Print ISBN: 978-1-4398-3665-1, eBook ISBN: 978-1-4398-3666-8
[6] http://semanticweb.org/wiki/N-ary_relations
[7] Robert Stevens, et. al., Using OWL to model biological knowledge, International Journal of Human Computer Studies, Elsevier, 2007, pp. 583-594
[8] D.L. McGuinness and F. VanHarmelen. OWL Web Ontology Language Overview", W3C Recommendation, February2004.http://www.w3.org/TR/owl-guide/
[9] http://www.w3c.org
[10] http://www.w3.org/TR/2004/REC-rdf-primer-0040210/#reification

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[11] Vinu P.V., Sherimon P.C., Dr. Reshmy Krishnan, Development of Seafood Ontology for Semantically enhanced Information Retrieval, International Journal of Computer Engineering and Technology, IJCET, ISSN 0976 – 6367, Volume 3, Issue 1. January-June 2012, pp 154-162.
[12] Boris Motik, Peter F. Patel-Schneider, and Bijan Parsia. OWL 2 web ontology language document overview. URL, 2009. http://www.w3.org/TR/owl2-overview/.
[13] http://www.w3.org/TR/swbp-n-aryRelations/
[14] John Hebler. et.al, Semantic Web Programming Wiley Publishing, 2009
[15] Sinuhe Arroyo, et.al. Semantic Web Languages –Strengths And Weakness, IADIS International Conference 2004, Lisbon, Portugal