Knowledge Engineering Approach for constructing Ontology for e-Learning Services

M. Farida Begam¹ and Gopinath Ganapathy²
¹Department of Information Technology, Manipal University Dubai Campus, UAE, ²School of Computer Science and Engineering, Bharathidasan University, India

E-mail: farida@manipaldubai.com, gopinath.g@csbdu.in, drgg@bdu.ac.in

Abstract—Recent ongoing enhancement in web is providing Semantic Web based services a more significance. Ontologies are important technology for representing knowledge base and implementing Semantic web. Notion of pedagogy and learning has been recently changed from traditional instructor centric teaching-learning process to web based learner centric teaching-learning process. E-Learning is supported by various resources available over Internet such as on-line courses, lecture notes, slides, and webinars etc. Representing these resources using Meta data model and integrating and reusing these resources in order to provide Semantic web based e-Learning services is a challenging process. This paper outlines the knowledge engineering approach for constructing semantics for e-Learning domain entities using Web Ontology Language OWL. It also focuses on integration and reasoning out of these entities using intelligent agents in providing Semantic Web based e-Learning services.

I. INTRODUCTION

World Wide Web plays an important role in different kinds of societies such as education, health, business, commerce, government and etc. Drastic growth of information available on web is leading us to develop proper mechanism to organize and represent these data and information in meaningful way. Information retrieval, communication, storage, computing through clouds and many more processes can be done with the advent of current web services available. Advances in digital world and technologies help the e-learning community to learn without any barriers like space, and time. E-Learning has become the basic for the current rapidly growing education community when computation and information are ubiquitous and always available.

Web Services facilitate this e-Learning with the availability of applications through platform independent technology. Semantic web services present one more layer above existing web and allow the users to interact with the web in meaningful way. It provides information in a precise, machine interpretable form, ready for software agents to process, share, and reuse it, as well as to understand what the terms describing the data mean [1]. In e-Learning if processes are described with semantics; it will make the applications to interact in a meaningful way, in turn serves the heterogeneous learners to acquire the knowledge according to their perception. This approach will enhance the learner’s experience. Main objective of this semantic web is to extend the current web technology to allow the development of intelligent agents, which can automatically and unambiguously process the information available on millions of web pages [2]. These software based agents provide appropriate infrastructure and help in achieving development of semantic evolution of e-Learning systems. Conceptualization, Ontologies and data interchange formats RDF, XML, XTM (XML Topic Maps), OWL—Web Ontology Language, OWL-S, RDF Schema and Rule ML (Rule Mark up Initiative ) are the basic technologies to develop formal description of concepts, terms, and relationships within a given knowledge domain. Using OWL and OWL-S, the web resources are converted to proper knowledge representation such as semantic annotations. These annotations can be used to support the user in composing work flows for various business processes including e-Learning processes. Moreover, Semantic web services provide a number of different educational activities by transforming a static collection of information into distributed way on the basis of Semantic Web technology making content within the WWW machine-processable and machine-interpretable [5]. Converting the concepts into Ontologies can be done with various approaches. We have incorporated the knowledge engineering approach which eases the process of developing the Ontologies for an e-Learning domain. The main objective of this paper is to outline how to develop the Ontologies and how they can be integrated and reused in different e-Learning applications.

This paper is organized as follows: In Section 2, we discuss about the overview of related work in the field of constructing semantic entities and Section 3, we start with the structure of semantic web and description about semantic web technologies and in Section 4, Knowledge engineering approach for constructing Ontologies for e-Learning domain is discussed and section 5 presents how e-Learning related semantic entities can be developed using OWL constructs. Section 6 concludes the paper.
II. RELATED WORK
In recent years, lot of research is going on in the combined area of Semantic Web Services and Artificial Intelligence. The current standards for web services such as WSDL, UDDI and SOAP are used to describe, discover and communicate with syntactic representations where as RDF, RDF Schema and OWL are used for semantic description of entities. Service oriented architecture is becoming crucial with the emerging of Web Service Architecture (WSA) [3]. Semantic annotations exploit the full potential of the service-oriented approach. As stated by Sousan et al. [4] Ontology combines and relates number of ideas and concepts together in a hierarchical format. They suggested how to construct domain Ontology from texts using NLP software to identify the words and parts of the sentences. These parts are related and Ontologies are constructed. Semantics can be added to already available service descriptions using WSDL-S [6], OWL-S and Diane Service Description [7] which include four IOPE service description elements. Simon, Hornung and Lausen have provided the approach ViPER (Visual Perception-Based Extraction of Records) [8], a wrapper system for localizing and extracting structured data records from web pages. This is based on pattern searching and matching concepts. Same kind of work is done in TARTAR [9] which extracts proper tabular structures from arbitrary tables and used the concepts of F-logic frames which can be built from rich nested tables and F-logic Rules to obtain the semantic interpretation. TANGO (Table Analysis for Generation Ontologies) [11] is based on table analysis and used for generating Ontologies. OntoExtract and OntoWrapper which are part of CORPORUM OntoBuilder [15] are the tools to extract concepts from unstructured information and semi structured information respectively. We should consider three important axioms while generating Ontologies for any domain.
- There is more than one way to model a domain. The best model depends on the application of interest.
- Ontology generation is a repeated process. Always there is a way to refine and repeat the process of development.
- Ontology description that can be built using noun-verb and relationships between these two should be close to the physical and logical structure of the domain.

In this paper we concentrate on how to construct Ontologies using OWL with the help of knowledge engineering approach and how intelligent agents are performing integration and reasoning to provide the better result in e-Learning processing.

III. SEMANTIC WEB – AN OVERVIEW
Semantic web is an extension of available web, which will unleash the revolution of new possibilities [2]. In many fields like medicine, bio informatics and web search, and data mining, it has got widespread usage. The strength of this technology is to be extended to e-Learning to get better outcome.

Fig. 1. illustrates the various technologies used in different layers of abstractions in the Semantic Web.

Unicode and URI: This is an international standard for encoding the text. This includes all scripts in active use today, many scripts known only by scholars, and symbols which do not strictly represent scripts, like mathematical and linguistic symbols. URI –Uniform Resource Identifier is used to identify any resource such web page, e-book, e-Learning material such as slide, presentation and etc.

XML+NS+rdfschema: XML (eXtensible Markup Language) XML Schema ensures that there is common syntax used in Semantic web. XML Name Spaces allow specifying different markup vocabularies in one document. XML Schema serves for expressing schema of a particular set of XML documents.

RDF + rdfschema: RDF (Resource Description Graph) is core data representation format for semantic web. It is based on triples subject-predicate-object that form graph of data. RDF Schema is used to define the vocabulary of RDF model.

Ontology Vocabulary: Ontology is a specification of conceptualization. It describes the knowledge base of the particular domain. It has set of components such as Classes, slots, facets, and instances. Concepts of the domain are represented as classes, properties of concepts called slots describing various features and attributes of the concepts and restrictions on slots are called facets. There are several higher level languages for generating Ontologies such as OIL and DAML + OIL and OWL. Wide usage and adoption of OWL makes this language as standard for Ontology representation language for Semantic web. It has set of XML elements and attributes which defines the terms and relationship among these terms in well structured manner. It also extends the elements of RDF and RDFS.OWL includes the concepts like union, intersection and complement of Ontology. Cardinality constraints on properties can be imposed on properties. Properties can also have characteristics in OWL. On the whole, OWL provides better way of representing knowledge base which is required in semantic web services. It helps us to obtain the shared understanding of the key concepts, ensures reusability of knowledge base, and makes distributed large scale machine processing automatic.
IV. METHODOLOGY

A. Knowledge Engineering Approach for the construction of Semantics for e-Learning Services

Semantic Web offers new technologies to the developers of Web-based applications aiming at providing more intelligent access and management of the Web information and semantically richer modeling of the applications and their users. [13]. Ontologies are base for defining semantics for any domain. Developing Ontologies involves 1) Defining classes in the Ontology 2) Arranging the classes in a taxonomic hierarchy 3) Defining slots and describing allowed values for these slots 4) Filling in the values for slots for instances [14]. Semantic web technology can significantly improve the effectiveness of digital resource sharing. By using an Ontology inference service, searching no longer need be constrained to matching the content only, but also by inferring the true meaning of the concept it is possible to retrieve all knowledge equivalent resources.

Knowledge engineering involves in following systematic/scientific approach for developing, maintaining and using knowledge based systems. This approach which has the base of software engineering steps provides us to construct shared knowledge bases in efficient manner. Ontologies are important tool which ensures reusability in this development process. This Ontology construction gains more and more importance in library science, e-commerce and search engine. Several learning objects ranging from simple text documents, power point slides, applets, and PDF exist over Internet. These objects should be represented with proper technique so that they can be exposed and shared or reused by e-Learning applications in turn used by versatile and dynamic learners. The main challenge in constructing educational subject Ontology is in adequately defining its scope and granularity, and how it should be represented to the user. In any e-Learning system, learners are in necessitating of learning content and assessment content. These content have set of services to be provided to the learners. Consider the training domain which may consist of user types Learner, Facilitator and resources such as web page, slides, oral presentation, web site and activity type such as taking on line exam, merging more than one resources and composition of work flow for personalizing the services to Learners, and semantic search. Currently there is wide spread use of Web Ontology language to develop Ontologies for e-Learning. As discussed by Natalya F. Noy and Deborah L. McGuinness, the following knowledge engineering methodology steps [14] are used for constructing Ontologies:

- Determine the domain and scope of the Ontology
  - Based on certain basic and competency questions, the domain and scope of Ontologies are defined clearly
- Consider reusing existing Ontologies
  - If one team develops Ontologies for a domain, other group can reuse or extend for their purpose.
- Enumerate important terms in the Ontology
  - Helps to find out the relationships among various concepts used in the domain
- Define the classes and the class hierarchy
  - Different approaches available such as top down, bottom up and combination of both. Whichever suits the requirements choose that approach to develop the hierarchy.
- Define the properties of classes
  - This step is to provide information about classes, in turn helps to develop the proper internal structure of an Ontologies
- Define the facets of the properties
  - Consider the different facets about the property. For instance, Course may be delivered through multiple presentations. The cardinality, value type, participation in the relationship can be different facets of the Properties.
- Create instances
  - Creating instances of the classes in the hierarchy

Basically the above mentioned steps are similar way of working in Unified Modeling Language to model the system under study. We divide the Ontologies to be generated for training domain into five categories.

- Curriculum Ontology
  - Used to define the educational goals related to curriculum of the training
- Pedagogy Ontology
  - Used to represent the approaches or methodologies of Pedagogy
- Learning Ontology
  - Used to represent the Learning objects/materials Concepts
- Assessment Ontology
  - Used to represent the assessment resources Concepts
- Feedback Ontology
  - Used to get the findings/results concepts

To illustrate the creation of Ontologies for e-Learning training domain, we apply the steps mentioned above and construct Ontologies.

Step 1: Defining the domain and scope of the Ontology

Domain: It is a set of possible entities which are used by the well defined processes in semantic web services. It includes demonstration of e-Learning Ontologies such as Learner, Course, Web Page, Slides, Presentation, and etc.

Scope: It clearly defines the Setting up to use these Ontologies in E-Learning related applications such as searching for learning content, assessment content or searching the topic or course.

Step 2: Finding Important Terms: Learner, Course, Web Page, location, topic, Course’s components such teaching plan, Slides, Presentation, books and books
Step 3: Defining Classes and Class Hierarchy: The following classes and class hierarchy have been identified and is to be transformed to Ontology. UML class diagram (Fig. 3) is used to depict the Ontologies for the training domain.

IV. CONSTRUCTION OF ONTOLOGIES FOR E-LEARNING TRAINING DOMAIN

There has been lot of interest and investment shown in personalizing e-Learning based on knowledge representation in terms of Ontologies. Well defined procedure needs to be incorporated in Ontology construction. Transforming UML class diagram into OWL-Ontology involved several steps. The following four steps illustrate how to apply knowledge engineering approach to construct Ontologies for e-Learning training domain. The example Ontology list provided in each example helps to understand how to create relevant Ontology for the domain specified.

A. Define the classes and class hierarchy

In this step, Ontology vocabulary owl:Class is used to define classes in OWL documents. This subclass of rdfs:Class. Sub classes can be created using rdfs:SubClassof. The following example will help to understand how to create classes and class hierarchy. All the owl listing given in this section based on the UML diagram given in the Fig. 2.

```xml
<owl:Class rdf:ID="Person">
  <owl:Class rdf:ID="Learner"><rdfs:subClassOf rdf:resource="#Person"/>
</owl:Class>
</owl:Class>
```

B. Define class properties, restrictions and cardinality constraints.

Owl vocabulary used to define properties owl:ObjectProperty. Consider that Learner registers a course. At least one Course must be registered by a Learner. How this restriction can be imposed on the property? The following list shows this:

```xml
<owl:Class rdf:ID="Course">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="registered_by"/>
      <owl:someValuesFrom rdf:Resource="Learner"/>
    </owl:Restriction>
    <owl:Class>
      <owl:Class rdf:ID="Person">
        <owl:Class rdf:ID="Learner"><rdfs:subClassOf rdf:resource="#Person"/>
      </owl:Class>
    </owl:Class>
  </rdfs:subClassOf>
</owl:Class>
```

specifies that at least one value of registered_by property is an instance of Learner.

```xml
<owl:Class rdf:ID="Course">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="registered_by"/>
      <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:cardinality>
    </owl:Restriction>
    <owl:Class>
      <owl:Class rdf:ID="Person">
        <owl:Class rdf:ID="Learner"><rdfs:subClassOf rdf:resource="#Person"/>
      </owl:Class>
    </owl:Class>
  </rdfs:subClassOf>
</owl:Class>
```

From this, owl:someValuesFrom provides the inference that Jhon and Joe are Learners of the training Course J2EE

Cardinality constraints also provide a way to impose restriction on properties. If we want to specify that exactly one Learner is to register for a course, the following list specifies that using cardinality vocabulary.

```xml
<owl:Class rdf:ID="Course">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="registered_by"/>
      <owl:minCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:minCardinality>
    </owl:Restriction>
    <owl:Class>
      <owl:Class rdf:ID="Person">
        <owl:Class rdf:ID="Learner"><rdfs:subClassOf rdf:resource="#Person"/>
      </owl:Class>
    </owl:Class>
  </rdfs:subClassOf>
</owl:Class>
```

```xml
<owl:someValuesFrom rdf=Resource="Learner">
<owl:Class rdf:ID="Person">
</owl:Class>
```

specifies maximum range for any restriction on a particular property.

If we need to describe the property, rdfs:domain, rdfs:range and rdfs:subPropertyOf Vocabulary can be used.

In OWL, two types of properties can be specified. Owl:ObjectProperty is used to connect a resource to another resource and owl:DatatypeProperty is used to connect a resource to an rdfs:Literal or an XML schema built-in data type value.
For example Object Property can be used to specify that the Course is taught by Facilitator. Here Course and Facilitator are resources (Classes).

```xml
<owl:DatatypeProperty rdf:ID="has_emailAddress">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
  <rdfs:domain rdf:resource="#Facilitator"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
```

In the above mentioned list using rdf:type vocabulary, we specify the characteristics of property. Here functional property specifies that there is at most one value for that property. It is used to specify that every Facilitator has unique mail id: there is at most one unique value for each instance.

There are some other types of properties can be specified such as symmetric property, transitive property, inverse property and inverse functional property. These characteristics help to enhance the expressiveness of the domain we represent in the knowledge base.

Symmetric Property is used when a resource or class R1 is related to other Resource or class R2 with property P, then R2 is related to R1 by the same property. For example, E-Learning web site is linked with another web site.

```xml
<owl:DatatypeProperty rdf:ID="#E-LearningSite"/>
```

Inverse property is used to specify that a resource or class R1 is related with resource or class R2 by the property P, then the inverse of Property P⁻¹ will connect to resource R1. For example, inverse property can be specified as follows: If the Course is taught by Facilitator then we can specify that Facilitator teaches the Course.

```xml
<owl:InverseOf rdf:resource="#taught_by"/>
```

For example DatatypeProperty can be used to specify that Facilitator has emailAddress.

```xml
<owl:ObjectProperty rdf:ID="taught_by">
  <owl:inverseOf rdf:resource="#taught_by"/>
  <rdfs:domain rdf:resource="#Facilitator"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:ObjectProperty>
```
Transitive Property is used to specify that a resource or class R1 is related to resource or class R2 by the property P, and R2 is related to a resource R3 by the same property, then R1 is related to R3 by the property P.

In e-Learning scenario consider the following example and its equivalent Ontology defined with transitive property.

If Learner L1 has better score than Learner L2 and If L2 has better score than L3, then L1 has better score than L3.

If a RDF file has the following instance of Learner resource,

```
 Lantern rdf:ID="John">
     <has_better_score rdf:resource="Joe">
     </Learner>
```

And in another RDF file, there is another instance of Learner resource as follows:

```
 <Learner rdf:about="http://www.training_site.com/people#Joe" xmlns="http://www.w3.org/2002/07/owl#">
     <has_better_score>
     </Learner>
```

From this transitive property, the software agent understands that John has better score than Jane.

Functional property states that for a given rdfs:domain value there is unique rdfs:range value. An InverseFunctional property is just the opposite of it. For a given rdfs:range value, the rdfs:domain value must be unique. For example, assume course is taught by only one Facilitator. This can be defined as shown below:

```
 <owl:ObjectProperty rdf:ID="taught_by">
     <rdfs:domain rdf:resource="#Course" />
     <rdfs:range rdf:resource="#Facilitator" />
 </owl:ObjectProperty>
```

C. Ensure maximum reusability

Semantic Web is all about how automatic data processing of distributed information over the web can be carried out for getting intelligent results. This information has to be handled in efficient and effective ways to get better inferences. OWL has some of the characteristics which are listed out to enhance the automatic data processing.

- Equivalent classes uses the vocabulary owl:equivalentClass to specify that they are same
- Disjoint classes uses the vocabulary owl:disjointWith to specify that they are different

While defining Ontologies for any application domain, make sure that reusability has been ensured or not. If already existing Ontologies can be used for those applications, try to utilize those of.

In the same way, it is possible to specify that two instances are same and different. This can be done with the help of owl:sameIndividualAs property and owl:differentFrom property. While defining Ontologies for any application domain make sure that reusability has been ensured or not. If already existing Ontologies can be used for those applications, try to utilize those of.

D. Perform Ontology operations to obtain more meaningful knowledge base

Web resources are going to be accessed by content rather than keywords. As stated by Anderson and Whitelock [13] Semantic Web based Educational Systems design and working is based on three fundamental affordances. The first is the capacity for effective information storage and retrieval. The second is the capacity for nonhuman autonomous agents to augment the learning and information retrieval and processing power of human beings. And the third is the capacity of the Internet to support, extend and expand communications capabilities of humans in multiple formats across the bounds of time and space. It is important to note that non human agents are playing an essential role in the functionality of Internet. The Ontology based software agent is the one which perform one or combination of more than one Ontology based operations to fulfill the queries raised by the Learner. The below mentioned is the list of Ontology operations which can be performed on atomic or dependent Ontologies. Educational Web Servers are distributed with various e-learning related contents, documents and Ontologies which are available for access. Web Service interfaces help the learner to access the service of his/her choice. Semantically enhanced Web Services can be obtained by implementing different operations on Ontologies such as

- Merging of Ontologies: means creation of a new Ontology by linking up the existing ones.
- Refinement is mapping from Ontology A to another Ontology B so that every concept of Ontology A has equivalent in Ontology B, however primitive concepts from Ontology A may
correspond to non-primitive (defined) concepts of Ontology B.

- Unification is aligning all of the concepts and relations in Ontologies so that inference in Ontology can be mapped to inference in other Ontology and vice versa.

- Integration is a process of looking for the same parts of two different Ontologies A and B while developing new Ontology C that allows to translate between Ontologies A and B and so allows interoperability between two systems where one uses Ontology A and the other uses Ontology B.

- Inheritance means that Ontology A inherits everything from Ontology B. It inherits all concepts, relations and restrictions or axioms and there is no inconsistency introduced by additional knowledge contained in Ontology.

The above mentioned Ontology operations can be performed on atomic or non atomic dependent Ontologies. In OWL, different additional vocabulary can be used to implement these operations among Ontologies. These vocabularies are called class expressions. They provide us many basic set operations to be performed on Ontologies. For Example, consider merging of Ontologies in e-Learning domain:

```xml
<owl:Class rdf:ID="Course">
    <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Pedagogy" />
        <owl:Class rdf:about="#Assessment" />
        <owl:Class rdf:about="#Feedback" />
    </owl:unionOf>
</owl:Class>
```

Here owl:unionOf and rdf:parseType together informs the parser that Course Class is a union of Pedagogy, Learning Object, Assessment and Feedback classes. This way containership can be specified in OWL.

For example, creating subclasses can be done with the help of owl:subClassOf vocabulary. In e-Learning domain, consider that Professional is a Person. Learner is a professional.

```xml
<owl:Class rdf:ID="Person">
    <owl:Class>
        <owl:Class rdf:ID="Professional" />
        <rdfs:subClassOf rdf:resource="#Person" />
    </owl:Class>
</owl:Class>
```

From this OWL listing, a software agent infers Learner is a person. rdf:subPropertyOf is useful vocabulary construct to enforce is a relationship between properties. For example, taught_by is a subPropertyOf facilitated_by property can be specified in OWL as follows:

```xml
<owl:ObjectProperty rdf:ID="facilitated_by">
    <rdfs:domain rdf:resource="#Course" />
    <rdfs:range rdf:resource="#Facilitator" />
</owl:ObjectProperty>
```

The following Table I presents which Ontology operation can be achieved with what owl and/or RDF vocabulary.

<table>
<thead>
<tr>
<th>Ontology Operation</th>
<th>OWL/RDFS Vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merge</td>
<td>owl:unionOf</td>
</tr>
<tr>
<td>Refinement</td>
<td>owl:equivalentClass,</td>
</tr>
<tr>
<td></td>
<td>owl:disjointWith and</td>
</tr>
<tr>
<td></td>
<td>owl:complementOf</td>
</tr>
<tr>
<td>Integration</td>
<td>owl:unionOf, and</td>
</tr>
<tr>
<td></td>
<td>owl:intersectionOf</td>
</tr>
<tr>
<td>Inheritance</td>
<td>rdfs:subClassOf</td>
</tr>
<tr>
<td></td>
<td>rdfs:subPropertyOf</td>
</tr>
</tbody>
</table>

The steps mentioned above, elaborates the construction of Ontologies using OWL. This procedure is simple and helps to construct Semantics for any domain.

V. CONCLUSION

OWL has greater strength of providing complete expressiveness. FOAF provides additional constructs which helps to include more relationships between classes and resources. In this paper, we established a different approach for developing Ontologies for e-Learning knowledge base. This gives insight for developing knowledge base for any particular domain. The approach used in this paper introduces two important aspects 1) easiness in developing knowledge representation based on UML 2) OWL maximum expressiveness achieved using knowledge engineering approach. Instead of beginning construction of semantic entities in haphazard manner, well defined approach can facilitate semantic web based service development to get better results.

REFERENCES


