

文章编号: 2096-1472(2017)-06-55-05

使用本体标注构建SCORM本体模型的方法

张波¹, 申发海², 韦永林¹

(1.贺州学院数学与计算机学院, 广西 贺州 542899;

2.重庆邮电大学计算机科学与技术学院, 重庆 400065)

摘要: 针对SCORM学习对象模型语义缺失的问题, 结合语义Web本体标注的技术, 提出一种使用语义标注实现SCORM模型本体化的方法。该方法从SCORM规范中提取语义信息构建SCORM本体模型, 以公理的方式形式化表示了SCORM的语义信息和语义约束。并以此为基础在现有SCORM学习资源中扩展语义标注信息, 使SCORM学习资源中的元素和属性映射到SCORM本体的类和属性, 使用SCORM本体模型为学习资源提供语义约束和知识推理。该方法遵照SCORM语法、语义, 以及扩展思想, 将SCORM信息模型映射为SCORM知识模型, 把SCORM学习对象的使用提高到知识重用和知识互操作的层面。

关键词: 语义标注; SCORM; 描述逻辑; 本体模型; E-Learning技术标准

中图分类号: TP318 **文献标识码:** A

A Modeling Method to Structure SCORM Ontology by Semantic Annotation

ZHANG Bo¹, SHEN Fahai², WEI Yonglin¹

(1.College of Mathematics and Computer Technology, Hezhou University, Hezhou 542899, China;

2.College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Focusing on the semantic loss of the SCORM learning object model and combining with the semantic Web ontology annotation technology, the paper proposes a method applying semantic annotations technology to implement semantic SCORM information model. This method extracts semantic information from the SCORM specification to structure SCORM ontology, in which the SCORM semantic information and semantic constraints are formalized through axioms. Then on this basis, the existing SCORM resources can be extended with some semantic annotation, to map elements and attributes in the SCORM model to the SCORM ontology. Based on the syntax and semantics of SCORM, this research transforms the SCORM information model into the SCORM knowledge model, and promotes utility of SCORM learning objects to knowledge reusability and knowledge interoperability.

Keywords: semantic annotation; SCORM; description logic; ontology model; E-Learning technology standards

1 引言(Introduction)

在远程教育技术标准领域, 由于ADL(Advanced Distributed Learning)的SCORM(Sharable Content Object Reference Model)规范^[1]具有可访问性、协作性、持久性和重用性等优势, 因而得到了广泛的应用和研究^[2]。但是由于目前SCORM信息模型依然采用XML Schema作为模型定义语言, XML和XML Schema技术并不具备语义定义的能力, 造成现有SCORM信息模型的语义信息和语义约束的缺失。为此, 国内外SCORM主要研究机构也正在积极进行SCORM语义化研究。关于这项研究, 目前国内外研究思路全部集中于使用语义网(Semantic Web)思想和技术方面。其中, 卡内基梅隆的学习系统结构实验室(LSAL: Learning Systems Architecture Lab of Carnegie Mellon University)是众多SCORM本体化研究机构之一, LSAL明确提出将语义网技术纳入SCORM 2.x框架之中以实现SCORM的语义建模, 以构

建下一代更复杂的和更具自适应能力的智能学习系统。

鉴于语义网技术在知识表示和知识推理方面的强大生命力, 在远程教育技术标准领域, 使用语义网技术和思想研究各种教育技术标准已经成为一个趋势。其中对IEEE LOM(IEEE Learning Object Metadata)标准的研究最为充分, 其研究的方向主要是通过对IEEE LOM标准的本体化提高学习对象检索和分类的效率。IMS LD标准(IMS Learning Design specification)的本体化研究主要集中于IMS LD知识表示和语义约束方面^[3]。在目前对SCORM标准的本体模型研究中, 国内外学者和研究机构已经取得了很多经验和探索。主要的研究思路是(1)使用描述逻辑对SCORM信息模型进行本体建模^[4,5]; (2)SCORM元数据语义建模研究^[6]; (3)SCORM XML Schema信息模型转换为本体表示语言研究^[7,8]。在目前SCORM标准本体模型研究中, 所有的研究思路都集中在推翻现有的SCORM信息模型表示方法, 改用OWL、DAML-S等

语义网语言作为下一代SCORM信息模型定义语言,而忽视了XML技术在定义结构化信息模型方面的巨大优势,同时也忽视了目前存在大量符合现有版本SCORM的规范学习对象资源的现实。虽然XML Schema的设计是为了提供XML数据模型的语法约束,对信息模型语义表示能力的缺乏是根本性的不足,但是XML技术在表示数据的语法结构化方面,是现有的语义网本体语言无法匹及的。若SCORM模型完全替换为OWL、DAML、RDF等语义网语言表示的本体模型,固然具有了语义定义能力,为下一代智能学习系统的知识重用和知识互操作提供了基础,但是也将导致信息模型原有的清晰的结构化优势的丧失。而且,由于SCORM标准在E-Learning领域的广泛应用,现有的SCORM学习对象需要重新转换为本体表示语言,也将限制了新版本SCORM模型的推广和使用。

本文借鉴现有网页语义标注研究的思想,提出使用本体标注的方式实现SCORM学习对象模型的本体化。该方法在现有SCORM学习对象模型的基础上,增加语义标注信息,实现SCORM对象模型的元素和属性映射到一个SCORM本体模型,而符合SCORM规范的学习对象,作为本体模型的实例元素。在这个SCORM本体模型中,对SCORM学习对象模型进行语义定义和语义约束。该方法分为建立SCORM语义概念模型和对现有SCORM学习对象进行语义标注三个步骤。第一,建立一个用于定义SCORM信息模型的本体模型。把现有SCORM XML Schema信息模型中的元素映射为本体模型中类(Class)的概念,把SCORM中属性映射为本体模型中数据类型属性(Data Type Property),以此为基础建立SCORM概念模型。该本体模型用于对SCORM元素和属性的语义表示。在该本体模型中,原标准中元素和属性的名称空间、元素名、属性名完全保持一致。第二,通过OWL DL公理对SCORM本体模型进行语义扩充,使之增加SCORM语义信息和语义约束。语义信息和语义约束提取自SCORM规范文档,并使用描述逻辑对这些语义约束进行表示,并转换为OWL DL公理增加到SCORM本体模型中。第三,依据XML本身具有的可扩展性的特点,根据现有网页语义标注的思想,对现行的SCORM XML Schema模型定义,以及SCORM学习对象进行语义标注,在不改变现有模型语法结构的基础上,增加语义本体映射信息,以实现SCORM标准的本体化。

2 SCORM信息模型的本体映射(Ontology mapping of SCORM information model)

在本文提出了SCORM语义化方案中,首先需要建立一个与SCORM信息模型一致的SCORM本体模型。该本体模型语义化表示SCORM信息模型,作为SCORM信息模型的本体映射。该本体模型完全遵从SCORM规范,所有的语义信息均完全遵照SCORM规范,包括类名、属性名、名称空间、类约束、属性约束、公理约束等。该本体模型把现有SCORM XML Schema信息模型中的元素映射为本体模型中类(Class)的概念,把SCORM中属性映射为本体模型中数据类型属性(Data Type Property),把SCORM学习对象的数据映射为本体模型中的个体。

SCORM信息模型中定义了用于描述学习对象的XML元素和属性。在SCORM本体概念模型中,SCORM模型中的XML元素映射为类,类名和类名称空间将保持一致。SCORM模型中的元素所使用的命名空间有: <http://ltsc.ieee.org/xsd/LOM>(前缀为lom)、http://www.imsglobal.org/xsd/imscp_v1p1(前缀为imscp)、<http://www.imsglobal.org/xsd/imsss>(前缀为imsss)、http://www.adlnet.org/xsd/adlcp_v1p3(前缀为adlcp)、http://www.adlnet.org/xsd/adlseq_v1p3(前缀为adlseq)、http://www.adlnet.org/xsd/adlnav_v1p3(前缀为adlnav)^[1]。具有代表性元素的映射如表1所示。

表1 SCORM部分元素的类映射

Tab.1 Class map of partial SCORM elements

SCORM模型XML元素名	SCORM本体模型的类
imscp:manifest	&imscp;manifest
imscp:metadata	&imscp;metadata
lom:ifeCycle	&lom;ifeCycle
imscp:organizations	&imscp;organizations
imscp:organization	&imscp;organization
imscp:item	&imscp;item
adlcp:timeLimitAction	&adlcp;timeLimitAction
imsss:sequencing	&imsss;sequencing
adlseq:objectives	&adlseq;objectives
adlnav:presentation	&adlnav;presentation

在SCORM本体概念模型中,SCORM模型中的XML属性映射为Datatype Property,属性名和属性名称空间将保持一致。具有代表性属性的映射如下表所示。

表2 SCORM部分属性的映射

Tab.2 Class map of partial SCORM datatype property

SCORM模型属性	SCORM本体模型的Datatype Property	值域	定义域
imscp:default	&imscp;default	&xsd;IDREF	&imscp;organizations
adlseq:objectives	&adlseq;objectives	&xsd;boolean	&imscp;organizations
GlobalToSystem	GlobalToSystem	&xsd;boolean	&imscp;organizations
imscp:isvisible	&imscp;isvisible	&xsd;boolean	&imscp;item
imscp:parameters	&imscp;parameters	&xsd,string	&imscp;item
adlcp:readSharedData	&adlcp;readSharedData	&xsd;boolean	&adlcp;map
adlcp:writeSharedData	&adlcp;writeSharedData	&xsd;boolean	&adlcp;map

鉴于论文篇幅的原因,本文只给出了具有代表性的元素和属性的本体映射,其他元素和属性映射方法与表中所列元素的映射无异。

3 SCORM本体模型公理表示(Axiomatic representation of SCORM ontology model)

SCORM本体模型不仅包括SCORM规范的元素、属性的语义定义,还包括SCORM规范的语义约束。对比SCORM XML Schema信息模型,SCORM本体的最大的优势之一在于形式化表示了学习对象模型中的概念之间的约束。前面介绍了SCORM本体模型的类和属性在SCORM信息模型中的映射,本节是在前面建立的SCORM本体概念模型的基础上

通过公理的形式，增加SCORM规范的语义信息。主要包括SCORM元素从属关系的语义建模、SCORM属性与元素之间的从属关系建模，以及从SCORM规范中提取语义信息和语义约束的建模。

为了表示SCORM规范中XML元素之间的从属关系，在本体模型中增加两个ObjectProperty属性：“subElementOf”和“subElement”，两个属性互为反函数。这两个属性来源于XML语言的元素语义，目的是便于对SCORM规范的语义约束公理表示。其中“subElementOf”表示子元素与父元素之间的从属关系，“subElement”表示父元素与子元素之间的从属关系。

在使用subElementOf和subElement对SCORM本体概念模型中的类进行约束时，还需要考虑SCORM规范中元素之间的值域约束、必要性约束、值域约束，以及规范文本上下文约束^[9]，这些语义信息包含在SCORM XML Schema信息模型定义XSD文件、Content Aggregation Model规范文档、Sequencing and Navigation规范文档和Run-Time Environment规范文档中。下文以organizations和organization元素为例，说明该形式化表示思想。

在SCORM规范中，organization元素是organizations的子元素，且要求organizations至少需由一个organization子元素。使用描述逻辑对该从属关系的公理表示为：

$$\text{organizations} \sqsubseteq (\exists \text{subElementOf.organization}) \cap (\geq 1 \text{subElementOf}(1))$$

描述逻辑的语义形式化表示为：

$$\begin{aligned} \text{organizations}' \sqsubseteq & (\exists \text{subElementOf.organization})' \cap (\geq 1 \text{subElementOf})' \\ (\exists \text{subElementOf.organization})' = & \{x \mid \exists y.(x,y) \in \text{subElementOf}' \text{ and } y \in \text{organization}'\} \\ (\geq 1 \text{subElementOf})' = & \{x \mid \#\{y \mid (x,y) \in \text{subElementOf}'\} \geq 1\} \end{aligned} \quad (2)$$

OWL抽象语法表示为：

```
Class(
  organizations partial
  ( restriction ( subElementOf
    someValueForm(organization) ) )
  ( restriction ( subElementOf
    minCardinality(1) ) )
)
```

OWL语句表示为：

```
<owl:Class rdf:ID=" &imscp;organizations" >
  <rdfs:subClassOf>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
        <owl:Restriction>
          <owl:onProperty rdf:resource="subElementOf"/>
          <owl:someValuesFrom rdf:resource=" &imscp;or
ganization"/>
        </owl:Restriction>
        <owl:Restriction>
          <owl:onProperty rdf:resource="subElementOf"/>
          <owl:minCardinality rdf:datatype=" &xsd:int">1</
owl:minCardinality>
```

```
</owl:Restriction>
</owl:intersectionOf>
</owl:Class>
</rdfs:subClassOf>
</owl:Class>
```

SCORM属性与元素之间的从属关系可以通过增加元素的约束性公理实现。在具体的属性进行形式化时，也需要考虑属性的值域约束、必要性约束、值域约束，以及规范文本上下文约束。下文以organizations元素的structure属性为例，说明该形式化思想。

在SCORM规范中，organizations元素可以具有一个structure属性，structure属性属于可选属性，且数据类型为字符串类型。使用描述逻辑对该从属关系的公理表示为：

$$\text{organizations} \sqsubseteq (\forall \text{structure.xsd:string}) \cap (\leq 1 \text{structure}) \quad (3)$$

描述逻辑的语义形式化表示为：

$$\begin{aligned} \text{organizations}' \sqsubseteq & (\forall \text{structure.xs:string})' \cap (\leq 1 \text{structure})' \\ (\forall \text{structure.xs:string})' = & \{x \mid \forall y.(x,y) \in \text{structure}' \rightarrow y \in (\text{xsd:string})'\} \\ (\leq 1 \text{structure})' = & \{x \mid \#\{y \mid (x,y) \in \text{structure}'\} \leq 1\} \end{aligned} \quad (4)$$

OWL抽象语法表示为：

```
Class(
  organizations partial
  ( restriction ( structure
    allValueForm(xsd:string) ) )
  ( restriction ( structure
    maxCardinality(1) ) )
)
```

OWL语句表示为：

```
<owl:Class rdf:ID=" &imscp;organizations" >
  <rdfs:subClassOf>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
        <owl:Restriction>
          <owl:onProperty rdf:resource=" &imscp;st
ructure"/>
          <owl:allValuesFrom rdf:resource=" &xsd:string"/>
        </owl:Restriction>
        <owl:Restriction>
          <owl:onProperty rdf:resource=" &imscp;st
ructure"/>
          <owl:maxCardinality rdf:datatype=" &xsd:int">1</
owl:minCardinality>
        </owl:Restriction>
      </owl:intersectionOf>
    </owl:Class>
  </rdfs:subClassOf>
</owl:Class>
```

除了上述两种基本约束公理以外，还需要对从SCORM规范中提取语义信息和语义约束，并通过描述逻辑对这些语义约束进行建模。下文以SCORM Content Aggregation技术文档第CAM-3-11页引用IMS Content Packaging规范对于

Item元素引用resource元素的约束公理为例,说明SCORM语义信息的提取和形式化表示方法。

规范文档的原句为:“Only leaf Items (Item elements that do not have children) may reference a learning resource.”在SCORM概念模型语境中,对该句的语义解释为:在Item的个体中,具有imscp:hasPart属性和具有Item子元素存在异或关系。一阶逻辑表示为:

$$\forall a | a \in \text{Item} \wedge (\exists b | b \in \text{Item} \wedge \text{subElement}(a,b) \oplus \exists c | c \in \text{owl:Thing} \wedge \text{identifierref}(a,c)) \quad (5)$$

描述逻辑表示为:

$$\text{Item} \subseteq (\exists \text{subElement.Item} \cap \neg \exists \text{identifierref.T}) \cup (\neg \exists \text{subElement.Item} \cap \exists \text{identifierref.T}) \quad (6)$$

OWL抽象语法表示为:

```
Class(Item, partial unionOf(
  intersectionOf(restriction(identifierref
someValueFrom(owl:Thing)) ,
  complementOf(restriction(subElement
someValueFrom(Item))) ),
  intersectionOf(complementOf(restriction(identifierref
someValueFrom(owl:Thing))) ,
  restriction(subElement someValueFrom(Item)))
))
```

OWL语句表示为:

```
<owl:Class rdf:resource=" &imscp;Item" >
  </rdfs:subClassOf>
  <owl:Class>
    <owl:unionOf rdf:parseType="Collection">
      <owl:Class>
        <owl:intersectionOf rdf:parseType="Collection">
          <owl:Restriction>
            <owl:onProperty rdf:resource="&imscp;resourceref" />
            <owl:someValuesFrom
rdf:resource="&owl;Thing" />
          </owl:Restriction>
          <owl:Class>
            <owl:complementOf>
              <owl:Restriction>
                <owl:onProperty
rdf:resource="subElement" />
                <owl:someValuesFrom
rdf:resource="&imscp;Item" />
              </owl:Restriction>
            </owl:complementOf>
          </owl:Class>
        </owl:intersectionOf>
      </owl:Class>
    </owl:unionOf>
  </owl:Class>
```

```
<owl:intersectionOf
rdf:parseType="Collection">
  <owl:Restriction>
    <owl:onProperty
rdf:resource="subElement" />
    <owl:someValuesFrom
rdf:resource="&imscp;Item" />
  </owl:Restriction>
  <owl:Class>
    <owl:complementOf>
      <owl:Restriction>
        <owl:onProperty rdf:res
ource="&imscp;resourceref" />
        <owl:someValuesFrom
rdf:resource="&owl;Thing" />
      </owl:Restriction>
    </owl:complementOf>
  </owl:Class>
</owl:intersectionOf>
</owl:Class>
</owl:unionOf>
</owl:Class>
</rdfs:subClassOf>
</owl:Class>
```

4 SCORM信息模型的语义标注(Semantic annotation of SCORM information model)

SCORM信息模型语义标注的目的是在不改变目前SCORM信息模型语法结构的基础上,通过增加XML属性或标签,把SCORM数据作为本体实例数据映射到SCORM本体概念模型,以达到SCORM数据具有明确语义,为机器理解和智能处理提供基础^[10]。目前本体语义标注的方式主要包括两种,其一是将标注结果嵌入到原文档中,其二是将文档中的所有标注结果独立存放在一个文档中,并与原文档保持链接。由于SCORM规范本身具有的扩展机制,为通过增加扩展元素的方式对SCORM信息模型增加语义标注提供了便利。本文提出的语义标注方法借鉴了SAWSDL^[11](Semantic Annotations for WSDL and XML Schema)语义标注的思想,采用扩展modelReference属性,把标注信息直接写入SCORM信息模型定义xsd文件和SCORM学习对象清单xml文件(imsmmanifest.xml),完成对SCORM的语义标注。

本文为modelReference属性设定的名称空间为http://www.adlnet.org/SSCORM,其XML Schema的定义为:

```
<xs:schema targetNamespace="http://www.adlnet.org/SSCORM" xmlns="http://www.adlnet.org/SSCORM" xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:attribute name="modelReference" type="listOfAnyURI" />
  <xs:element name="attrExtensions">
```

```

<xs:complexType>
  <xs:anyAttribute namespace="##any"
processContents="lax" />
</xs:complexType>
</xs:element>
</xs:schema>

```

modelReference属性可以用于模式定义xsd文件中的<xsd:element>、<xsd:complexType>、<xsd:simpletypes>和<xsd:attribute>等元素中,表示该元素映射为对应的本体概念。该属性也可以直接用于SCORM学习对象清单xml文件(imsmanifest.xml),标注XML元素实例作为对应个体与某一本体概念相同。

下文以SCORM信息模型的根元素manifest为例,示例为manifest元素增加语义标注。

```

<xsd:element name="manifest" type="manifestType"
sscrom:modelReference="http://www.imsglobal.org/
xsd/imscp_v1p1#manifest"/>
<xsd:complexType name="manifestType">
  <xsd:sequence>
    <xsd:element ref="metadata" minOccurs="0"
sscrom:modelReference="http://www.imsglobal.org/
xsd/imscp_v1p1#metadata"/>
    <xsd:element ref="organizations"
sscrom:modelReference="http://www.imsglobal.org/
xsd/imscp_v1p1#organizations"/>
    <xsd:element ref="resources"
sscrom:modelReference="http://www.imsglobal.org/
xsd/imscp_v1p1#resources"/>
    .....
  </xsd:sequence>
  <xsd:attributeGroup ref="attr.version"/>
  .....
</xsd:complexType>
<xsd:attributeGroup name="attr.version">
  <xsd:attribute name="version" type="xsd:string"
sscrom:modelReference="http://www.imsglobal.org/
xsd/imscp_v1p1#version"/>
</xsd:attributeGroup>

```

manifest元素是SCORM学习对象imsmanifest.xml文件的根元素,其XML语法约束时在imscp_v1p1.xsd文件中定义,在上述示例代码中,在imscp_v1p1.xsd文件中,通过使用语义标注modelReference属性,使得XML元素和定义映射为本体模型的类和属性。manifest元素映射为SCORM本体模型的http://www.imsglobal.org/xsd/imscp_v1p1#metadata类,metadata元素映射为http://www.imsglobal.org/xsd/imscp_v1p1#metadata类,organizations元素映射为http://www.imsglobal.org/xsd/imscp_v1p1#organizations类,version属性映射为http://

www.imsglobal.org/xsd/imscp_v1p1#version属性。以此模式定义语言建立的SCORM学习对象清单文件,XML元素将映射为对应类的个体。

学习管理系统和学习对象编辑系统可以使用本体推理机对SCORM本体模型和SCORM学习对象清单文件进行本体查询、本体推理和验证。目前常用的本体推理机有Pellet、Racer、FaCT++、Jess、Jena等,可以根据学习管理系统的功能需求选择不同的推理机。

5 结论(Conclusion)

针对现有SCORM规范学习对象模型存在的语义缺失的问题,本文分析了现有远程教育技术标准本体模型研究的成果,结合语义Web本体标注的技术,提出了一种使用本体语义标注实现SCORM学习对象模型本体化的方案。该方案的思想是保持现有SCORM学习对象XML语法结构的基础上,运用SCORM模型的扩展机制,通过增加一个modelReference属性,实现SCORM学习对象模型的XML元素和属性与SCORM本体概念模型的映射。该SCORM本体概念模型的类和属性均来源于SCORM标准,包括类名、名称空间、属性名、元素之间的从属关系、值域约束、必要性约束、值域约束等,均与SCORM规范保持一致。本研究还对原SCORM规范文本中的语义信息进行提取,并通过描述逻辑实现形式化表示。这些语义信息以公理的形式扩充到SCORM本体概念模型中,解决SCORM模型的语义缺失的问题。

该方法既能结合XML Schema在描述结构化数据的优势,又可以使得XML元素映射到本体模型的类和属性,通过语义WEB技术保障SCORM学习对象的语义表示,也为目前已经存在的SCORM学习对象的本体化提供便捷可行的方式。

参考文献(References)

- [1] ADL.SCORM 2004 4th Edition[EB/OL].<https://www.adlnet.gov/adl-research/scorm/scorm-2004-4th-edition/>.
- [2] Supachanun Wanapu,Chun Che Fung,Nittaya Kerdprasop. An investigation on the correlation of learner styles and learning objects characteristics in a proposed Learning Objects Management Model(LOMM)[J].Education and Information Technologies,2016,21(5):1113-1134.
- [3] Ricardo R.Amorim,et al.A Learning Design Ontology based on the IMS Specification[J].Educational Technology and Society,2006,32(3):1436-4522.
- [4] 张金玲.基于描述逻辑与SCORM规范的学习内容参考本体[D].桂林:广西师范大学,2013.
- [5] 骆亮.SCORM语义化及其应用研究[D].重庆:西南大学,2009.
- [6] 孙玉娣,刘亮亮.SCORM语义化的教学资源库构建与实现[J].计算机应用与软件,2012,12:219-222.
- [7] Lora Aroyo,Stanislav Pokraev,Rogier Brussee.Preparing SCORM for the Semantic Web[J].On The Move to Meaningful Internet Systems,2003,2888:621-634.
- [8] Chrisa Tsinaraki,Stavros Christodoulakis.XS2OWL:A Formal