

基于本体技术构建的协同式网络课件编辑系统^{*})

程 静 邱玉辉 杨 明

(西南大学计算机与信息科学学院 重庆 400715)

摘 要 领域本体是解决异构系统语义互操作的关键技术。我们提出了基于协同式课件编辑环境的领域本体模型,该模型的特点在于建立了领域知识概念与学习资源和课件文档之间的关联。该模型采用 TRIPLE 语言实现,它能够方便地与 RDF 进行转换,并具有较强的推理能力。另外,本文还提出了基于领域本体的“语义冲突消除模型”,该模型是解决课件协同编辑过程中语义冲突的基础。

关键词 语义互操作,领域本体模型,TRIPLE,语义冲突消除模型

Construction of Cooperative Web Courseware Authoring System Based on Ontology

CHENG Jing QIU Yu-Hui YANG Ming

(Faculty of Computer and Information Science, Southwest University, Chongqing 400715)

Abstract Domain ontology is the key technology to resolve the semantic interoperability in heterogeneous systems. In the paper, we advance a model of domain ontology based on cooperative courseware authoring system, in which knowledge concepts can associate with learning resources and courseware documents. And we use TRIPLE as an ontology language to set up the domain ontology model, which can be transformed to RDF easily, and has the powerful reasoning capability. Based on the domain ontology, a semantic conflict elimination model is put forward, which is the foundation to resolve semantic confliction in authoring courseware cooperatively.

Keywords Semantic interoperability, Model of domain ontology, TRIPLE, Model of semantic conflict elimination

1 引言

20 世纪 90 年代中后期远程教育得到了高速发展,产生了大量 Web 学习平台。由于这些学习平台在开发语言、通讯机制、数据存储和体系结构等方面的不同,使得异构平台之间的互操作成为一个非常棘手问题。当前最为突出的表现是,学习资源在异构平台上共享、重用、交换、检索方面的困难。由此,国内外众多知名的组织开发了大量的元数据标准,期望采用元数据标注的方式解决资源交换的格式问题,其中包括著名的 IEEE LOM 标准、DC 元数据标准,以及 ADL SCORM 标准等。

元数据标准只能解决异构系统间信息交换的语法问题,关于信息交换中更深层的语义问题却是力所不及的。然而,在异构系统通信的过程中,针对本质相同而描述不同的事物形成一致的理解,即实现语义的互操作,却是异构系统互操作上亟待解决的另一问题。本体被定义为共享概念模型明确的形式化规范说明^[1],是解决系统语义互操作问题的重要工具。

当前,将本体技术应用于网络学习环境的构建,已经成为本体技术重要的应用领域之一。大量研究指出,对学习系统中领域知识进行建模,并利用领域模型控制知识点的表现,有助于实现系统的自适应功能^[2,3]。Lora Aroyo 领导的课题小组,长期致力于 e-Learning 系统中领域知识建模方面的研究,并利用本体技术先后开发了 AIMS 系统^[4]和 OntoAIMS 系统^[5],领域建模的工具也从概念图过渡到 DAML-S,以加强对领域概念的语义约束,并能够进行知识的推理和一致性的校验。Peter Dolog 等人致力于开发个性化学习环境,为了实现分布式学习环境下异构系统中学习者模型的交换和重用,

他们采用本体技术构建系统领域模型和学习者模型^[6],并利用基于一阶逻辑的推理规则,实现系统的自适应行为^[7]。

本文将就网络环境下课件协同编辑工具中对本体技术的使用展开研究。在协同编辑过程中,由于编辑者在地域、文化背景、知识背景之间的差异,使得“语义冲突”不断。另外,大多数的协同课件编辑系统往往难以处理跨学科的知识体系,而主要原因在于系统对于多学科的知识体系缺乏统一的知识描述机制。基于此,本文采用本体技术作为领域知识的描述手段,以期解决课件协同编辑过程中的语义冲突和跨学科知识描述等问题。文章结构安排如下:第 2 部分提出系统的体系结构模型;第 3 部分具体介绍利用本体技术所建立的领域模型;第 4 部分介绍了用于处理语义冲突的“语义冲突消除模型”;最后是小结。

2 系统的体系结构

在课件的编辑过程中,主要涉及三类不同的资源,它们分别是:1)用于课件创作的基本素材资源;2)所有学习内容形成的知识结构;3)最终形成的课件资源。由于每类资源的处理任务大相径庭,为此我们构建了不同的编辑工具并形成了相应的资源库,分别是:与基本素材资源相对应的学习资源编辑工具和学习资源库、与知识结构相对应的领域本体编辑工具和领域知识本体库,以及课件编辑工具和课件库,如图 1 所示。

学习资源编辑工具的主要功能有:(1)搜索原始素材;(2)编辑素材;(3)对于需要入库的素材,加上 LOM 元数据转换为学习对象之后入库;(4)利用 SCORM 规范将简单素材打包成复杂的学习对象入库;(5)实现学习资源库中基于元数据的

^{*} 基金项目:西南大学校青年基金(SWNUQ2004027)。程 静 硕士,讲师,主要研究领域为异构学习系统互操作;邱玉辉 教授,博士生导师,主要研究领域为语义网络;杨 明 博士研究生,副教授,主要研究领域为语义网络。

检索,以方便提取恰当的学习对象进行课件的编辑。

领域本体编辑工具和领域知识本体库是系统的核心组件,它们的主要功能在于,将知识体系中的知识点和知识点间的关系,转化为本体库中的概念和概念间的关系。由于本体在领域建模方面的能力,使得知识点概念和概念间的关系不仅能够获得丰富的语义,更能够获得本体语言所提供的语义推理能力,为协同课件编辑过程中“语义冲突”的解决奠定基础。

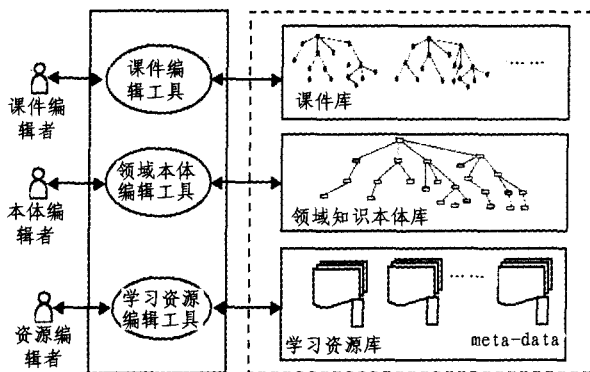


图1 系统的体系结构

课件编辑工具的主要功能在于,按照课件编辑者的意图,提取学习资源库中的学习对象,根据领域知识本体库所表现

的知识结构,进行课件的编排。常见的课件单位有:课程、主题、活动。“活动”是带有一定学习任务的学习单元,是教学过程展开的基本单位。一个学习活动中往往由若干学习对象构成教学过程中所需资源,它们最终的执行顺序由特定的排序信息控制。“主题”是带有一定教学目标的、基于一次课堂教学所需要的课件内容,它由若干个活动按照一定的顺序组织起来。相关主题构成了一门课程,课程是学习者在一段时间内学习的课件的集合。

3 构建领域知识本体

对领域知识建模已经成为众多网络学习平台在建设过程中必须要考虑的问题,原因在于领域知识模型能够更好地控制网页对知识点的表现粒度,从而能够获得系统的自适应能力;另外,个性化学习系统往往将领域知识模型作为学习者知识模型的基本框架,同时还可以作为构建个性化学习目标的基础^[2,3]。能够进行领域建模的工具很多,如向量、概念图^[8]、本体^[5-7]等。本系统采用本体语言 TRIPLE 建立领域知识模型,它能够方便地与 RDF 进行转换,且具有强大的推理能力。

3.1 领域本体模型

我们建立领域模型的思想在于,利用知识点概念建立学习资源的索引,并实现与课件库的关联。具体模型如图2所示。

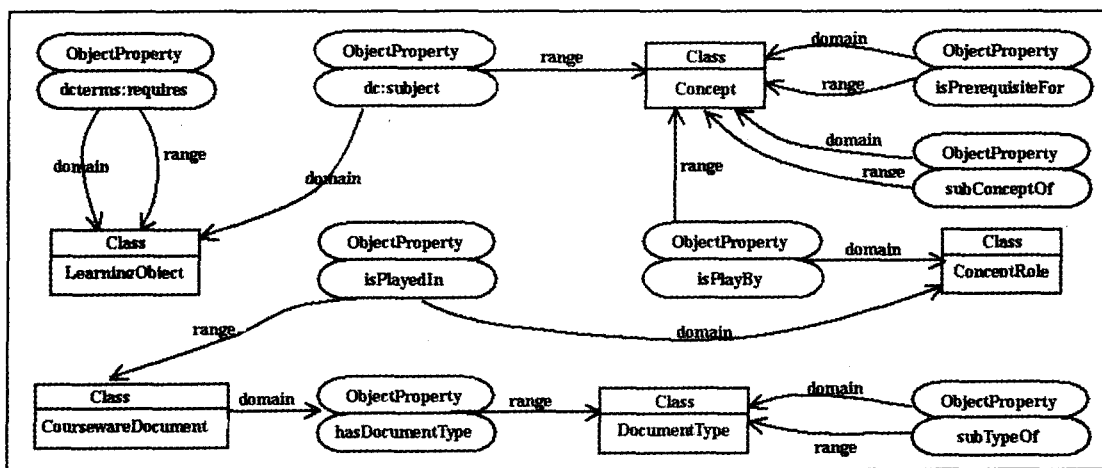


图2 领域本体模型

从图2中可以看到,领域本体模型中有三个主要类,即知识点概念类(Concept)、学习对象类(LearningObject)和课件文档类(CoursewareDocument)。知识点概念之间的关系有两种:知识点间的先修关系(isPrerequisiteFor)和知识点之间的包含关系(subConceptOf),并由此建立了 Concept 连接自身的属性 isPrerequisiteFor 和 subConceptOf。在学习对象之间,由于复杂的学习对象可以由简单学习对象组合而成,故建立属性 dc:terms:requires,表示学习对象之间的组合关系。由于采用知识点概念对学习资源进行索引,因此需要建立知识点概念和学习对象之间的关联,即学习对象的主题为某个知识点概念,我们采用属性“dc:subject”将 Concept 和 LearningObject 关联起来。知识点概念的详细内容和相应素材是集成在课件文档中进行表现的。不同的知识点概念在同一课件文档中往往会起到不同的作用,比如有的可以作为课程的引子,有的则是课程讲述的核心。由此,我们建立了 Concep-

tRole类,用来描述 Concept 在课件文档中的不同作用。而 Concept 类和 CoursewareDocument 类之间通过 ConceptRole 类和两个属性:isPlayedIn 和 isPlayedBy 关联起来。

3.2 领域本体的实例

我们采用 TRIPLE 语言来描述领域本体模型。TRIPLE^[9]是一个基于 Horn 逻辑的分层规则语言,支持 RDF 三元组的表示、推理和转换。对于其它能够进行推理的本体语言,如 DAML+OIL、RDF Schema 等,可以直接用 TRIPLE 实现,也可以编写为外部推理模块与 TRIPLE 语言程序进行交互。由于 TRIPLE 中有命名空间(NameSpaces)、资源(Resources)和陈述(Statements)等基本成分,因此可以方便将 RDF 陈述转换为 TRIPLE 程序。

以下为 TRIPLE 语言书写的一个实例:

```
sun.java: ' java/concepts/class.html ' [
rdf:type -> doc:CoursewareDocument;
```

```

dc:subject -> doc:OO Class ] .
doc:OO Class [
rdf:type -> doc:Concept;
doc:isPrerequisiteFor -> doc:OO Inheritance;
doc:subConceptOf -> doc:Object-Oriented ] .
doc:ClassesIntroduction [
rdf:type -> doc:ConceptRole;
doc:isPlayedBy -> doc:OO Class;
doc:isPlayedIn ->
sun:java; 'java/concepts/class.html'; ] .

```

该实例是对关于 Java 程序设计课程中一个课件文档的描述。该课件文档的名称是“sun java: 'java/concepts/class.html'”，其中 sun java 是一个自定义名称空间的缩写，该课件文档完整的 URI 应是“sun java # java/concepts/class.html”。课件文档的类型是 CoursewareDocument，所关联的知识点主题是 OO Class。OO Class 是一个 Concept，它是 Object-Oriented 概念之下的一个子概念，它的前驱知识点概念是 OO Inheritance，并且课件文档“sun java: 'java/concepts/class.html'”在表现知识点概念 OO Class 的作用是 ClassesIntroduction，而 ClassesIntroduction 的类型是 ConceptRole。

4 构建基于本体的“语义冲突消除”模型

在本系统提供的课件编辑环境中，课件的编辑者可以来自于不同的地区或不同的领域，他们通过网络连接到该平台。由于背景不同，常常导致他们对同一问题理解的角度也不一样。为了促进相互之间的协作和沟通，必须提供相应的机制保证他们对问题的内涵达成一致的理解。这种机制体现在两个方面，一是采用本体对领域知识进行建模，二是建立基于本体的“语义冲突消除”模型，从而解决协作编辑过程中经常发生的语义冲突问题。

基于本体的“语义冲突消除”模型是一个五元组 $M = \langle \text{Role}, k, \rho, \mu, \lambda \rangle$ ，其中 Role 表示从发生语义冲突直到进行消除过程中参与动作的相关角色； k 表示当前产生语义冲突的词汇或文字段； ρ 表示对产生语义冲突的文字段 k 的触发操作； μ 表示对文字段 k 的语义解释动作； λ 表示将文字段 k 的语义解释发送给相关接受者。具体说明如下：

定义 1 参与语义冲突消除模型的角色集 Role

$\text{Role} = \{\text{知识点触发者}, \text{知识点解释者}, \text{知识点接受者}\}$ 。
 知识点触发者表示在课件协作编辑过程中对当前共同关注材料中的某些文字段 k 提出疑问的人，知识点触发者必须对提出疑问的部分 k 作标记；知识点解释者则需要对知识触发者标记的部分 k 进行基于本体的解释，即检索到与 k 语义相似的本体概念集 C ；知识点接受者是那些对本体概念集 C 的相关素材正在编辑的人。

定义 2 知识点触发操作 ρ

$\rho(k)$ 表示对材料中某段文字 k 执行的触发操作，该触发操作会对文字段 k 进行标注。

定义 3 知识点解释操作 μ

$\mu: \mu(k) \rightarrow C, C \subseteq O$

对文字段 k 的解释操作 $\mu(k)$ ，实际上是在领域本体集 O 中寻找与文字段 k 在语义相似度^[10]上小于一定阈值的本体集 C 。

定义 4 语义解释发送操作 λ

$\lambda(k)$ 表示将通过解释操作 $\mu(k)$ 获得的本体概念集 C 发送给相应的接受者，即那些正在编辑相应本体概念所对应资源的人。

该模型具体的执行过程可以参考图 3。首先，某课件编辑者在课件创作过程中发现了让人困惑的文字段 k ，利用知识点触发操作 $\rho(k)$ 向知识点解释者提出解释申请；接下来，知识点解释者接受申请，并利用知识点解释操作 $\mu(k)$ 在本体库 O 中搜索语义相似的本体集 C ；第三步，将本体集 C 发送给相应的知识点接受者，以消除语义冲突。

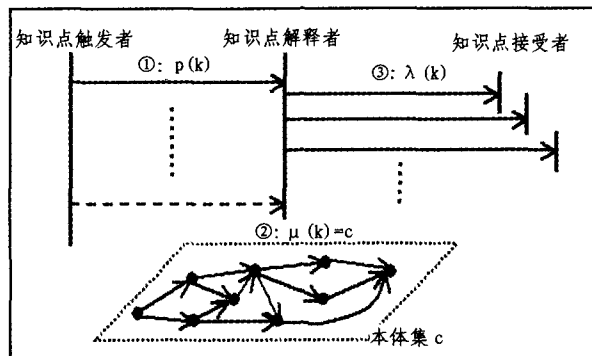


图 3 语义冲突消除过程

小结 本文介绍了我们在构建协同式课件编辑系统时所采取的关键技术，即采用本体技术构建了领域知识模型，并基于该模型提出了课件协作编辑过程中语义冲突消除的方法。我们所提出的领域本体模型的优点在于，能够将课件编辑所需的学习资源与最终形成的课件文档关联起来，且建立了基于知识点概念的索引；另外，该模型也是后期课件编辑过程中语义冲突消除的基础。

在本系统中，我们采用 TRIPLE 语言来描述领域本体模型。TRIPLE 语言具有与 RDF 相似的语言成分，如命名空间、基于三元组的陈述语句等，因此能够方便地实现 RDF 描述向 TRIPLE 程序的转换；另外，TRIPLE 是基于 Horn 逻辑的，具有较强的推理能力，且能够外挂 OWL、DAML 等本体语言，功能强大，且易于理解。在 3.2 节中给出了一个 TRIPLE 程序实例，具体说明领域本体的含义和 TRIPLE 语言的使用方法。

本系统当前还只是一个简单的原型系统，仅实现了其中较为核心的模块，还需进一步完善的工作包括：1) 实现基于语义的资源搜索工具，从而能够更准确描述编辑者当前的检索需求；2) 采用领域本体建立个性化编辑者模型，以便编辑者模型能够在网络上异构的系统中进行迁移；3) 本系统当前还只能基于异步的模式进行工作，对于同步编辑的工作方式还需要在通讯机制、存储机制等方面开展更多的工作。

参考文献

- 1 邓志鸿, 唐世渭, 张铭, 杨冬青, 陈捷. Ontology 综述. 北京大学学报(自然科学版), 2002, 38(5): 730~738
- 2 Brusilovsky P. Developing adaptive educational hypermedia systems: From design models to authoring tools. In: Murray T, Blessing S, Ainsworth S, eds. Authoring tools for advanced technology learning environment. Dordrecht: Kluwer Academic Publishers, 2003. 377~409
- 3 Aroyo L, Dolog P, Houben G-J, Kravcik M, Naeve A, Nilsson M, Wild F. Interoperability in Personalized Adaptive Learning.

Educational Technology & Society, 2006, 9 (2): 4~18

4 Aroyo L, Dicheva D, Cristea A. Ontological support for web courseware authoring. In: Proceedings of ITS' 02, Biarritz, France, 2002

5 Aroyo L, Pokraev S, Brussee R. Preparing SCORM for the Semantic Web. On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, LNCS 2888. 621~634

6 Dolog P, Henze N, Nejd W, Sintek M. Personalization in Distributed e-Learning Environments. In: Proceedings of the 13th International World Wide Web Conference, New York, USA, 2004

7 Dolog P, Henze N, Nejd W, Sintek M. Towards the Adaptive

Semantic Web. In: Principles and Practice of Semantic Web Reasoning (PPSWR'03), Mumbai, India, December 2003

8 周竹荣, 吴敬花, 邱玉辉. 基于概念图的网络课件与资源库集成. 计算机应用, 2005, 25(10): 2302~2305

9 Sintek M, Decker S. TRIPLE-A Query, Inference, and Transformation Language for the Semantic Web. International Semantic Web Conference (ISWC), Sardinia, June 2002

10 Resnik O. Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity and Natural Language. Journal of Artificial Intelligence Research, 1999, 11: 95~130

(上接第 256 页)

表 2 标准攻击下的检测结果对比

攻击典型	改进后方案 改进前方案	检测错误率 检测错误率
高斯滤波 3×3	0	0
高斯滤波 3×3	0	0
尖锐化	0	0
PEO	0	0
10%JPEG 压缩	0.20	0.26
15%JPEG 压缩	0.03	0.06
20%JPEG 压缩	0	0
缩放 0.75 倍	0	0
缩放 0.9 倍	0	0.27
缩放 1.5 倍	0	0.07
缩放 2.0 倍	0	0.06
朝 x 方向缩放 0.8 倍	0	0.15
朝 x 方向缩放 1.2 倍	0	0.09
朝 y 方向缩放 0.8 倍	0	0.14
朝 y 方向缩放 1.2 倍	0	0.08

(续表)

去除 17 行 5 列	0	0.22
去除 5 行 17 列	0	0.20
去除 5 行 1 列	0	0.10
去除 1 行 5 列	0	0.14
顺时针旋转 0.25°+剪切	0.14	0.31
逆时针旋转 0.25°+剪切	0.18	0.26
顺时针旋转 0.25°+剪切+缩放	0.08	0.26
逆时针旋转 0.25°+剪切+缩放	0.18	0.29
线性几何变换 (1, 0.13, 0.0008, 0.011, 1.008)	0.40	0.42
随机几何弯曲	0.52	0.56
剪切 1%	0.43	0.45
shearing-x-0%-y-1%	0.28	0.43
shearing-x-1%-y-0%	0.26	0.41
shearing-x-1%-y-1%	0.39	0.41

从表 1 和表 2 中可以看出:改进后的方案与改进前的方案相比,改进后的方案比改进前的方案在抵抗各种标准攻击和非标准攻击方面都有了明显的改善,尤其是在高斯白噪声攻击下对改进前方案鲁棒性的改善最大。这里所加的高斯白噪声较强,因为在这种较强的高斯噪声的干扰下,而且当嵌入容量较大时,会改进前水印方案的鲁棒性。除此之外,随机剪切对改进前的影响也较大,这也是因为这种剪切方式对 DCT 系数的影响较大的缘故。



(a)原始图像 256×256



(b)加强了标志位的图像

RJ
XY

(c)原始 LOGO

图 4 嵌入水印的图像

实验结果表明,特别是在高斯噪声影响下,改进后的方案明显优于原方案。

参考文献

1 张春田, 苏育挺. 信息产品的版权保护技术——数字水印. 电信科学, 1998, 14(12): 15~17

2 茹国宝, 杨锐, 夏双奎. 基于 Turbo 码的图像数字水印技术. 武汉大学学报(理学版), 2003, 49(5): 633~636

3 Dong P, Galatsanos N P. Affine Transformation Resistant Watermarking Based on Image Normalization[J]. In: Image Pro International Conference, 2003 (3): 24~28

4 杨东林, 叶梧. Turbo CDMA 多用户检测的研究. 电子科技大学学报, 2002, 29(3): 247~251

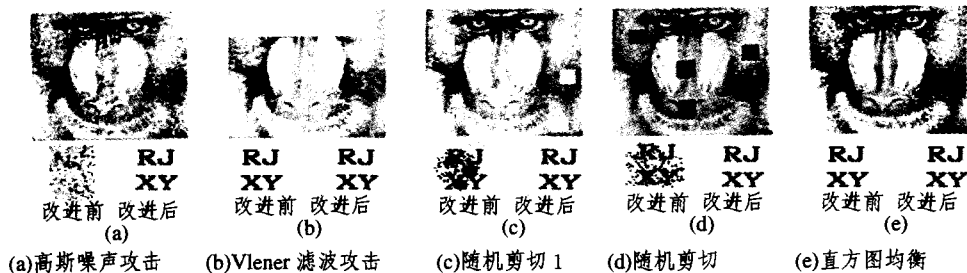


图 5 非标准攻击下的检测性能