

基于自适应学习的小学英语本体构建方法研究

景 然¹ 骆力明^{1,2*} 刘 杰^{1,2} 周建设²

¹(首都师范大学信息工程学院 北京 100048)

²(首都师范大学高精尖创新中心 北京 100048)

摘 要 目前中国的小学英语学习资源非常丰富,学习者想要找到适合自身的学习资源的难度也明显提升。自适应学习系统可以依据不同的学习者特征和学习路径,推送出符合学习者的学习资源和测试题目,因此自适应学习系统的数据组织结构是非常重要的。而关系数据库无法体现出知识之间的内在关联,所以本研究提出使用本体技术将小学英语教材的相关知识有机关联的方法,为项目提供数据支持,以实现学习者的自适应学习,且希望对其他自适应系统的知识库构建起到借鉴作用。

关键词 自适应学习 本体构建 小学英语

中图分类号 TP311 文献标识码 A DOI:10.3969/j.issn.1000-386x.2017.08.004

RESEARCH ON THE CONSTRUCTIONAL METHOD OF ONTOLOGY FOR
PRIMARY STUDENTS' EFL ADAPTIVE LEARNING

Jing Ran¹ Luo Liming^{1,2*} Liu Jie^{1,2} Zhou Jianshe²

¹(College of Information Engineering, Capital Normal University, Beijing 100048, China)

²(Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing 100048, China)

Abstract The learning resources of primary English of China are currently very abundant. It leads to difficulty for learners to find appropriate learning resources which is significantly increased as well. Adaptive learning systems can match the recommended learning resources and evaluation questions based on learning characteristics and learning paths of different learners, thus, the data organization structure of adaptive learning system plays a key role. Relational databases hardly reflect the intrinsic correlation between knowledge of specific subject. Therefore, this study proposed a constructional method by using ontology to connect relatively knowledge, and developed adaptive learning system for the fourth-grade students of China to learn English from multiple perspectives. The method can be used as a reference for other adaptive systems.

Keywords Adaptive learning Ontology construction Primary english

0 引 言

传统课堂中的学习者只能看到同一种教材中的知识以及相同的资源,但实际上同一知识点可能会在不同年级和版本教材中的不同语境下出现。因此纸质教材的知识点相对独立,知识点之间的关联较少,一些关

联往往是由教师根据以往的教学经验进行搭建的。而在层出不穷的电子教材中,学习资源丰富类型多样,但是大多数个性化课程设计主要是以教师为中心,针对其需求把纸质教材电子化,并没有真正做到适应学习者的课程制定^[1-2]。因此,为学习者解决资源自适应推送问题尤为重要,需要对不同教材中相同知识点进行逻辑聚合,对不同知识点进行关联建立,这些可以通

收稿日期:2016-10-14。国家自然科学基金项目(61371194,61672361);北京市自然科学基金项目(4152012);本成果得到北京成像技术高精尖创新中心资助。景然,硕士生,主研领域:智能教育软件与知识工程。骆力明,教授。刘杰,副教授。周建设,教授。

过构建具有概念化和共享性特征的本体知识库来解决^[3],使学习者能够进行多维度学习。

目前本体技术已经在教育领域得到了一定的应用。如郑艺等以“计算机科学与技术”学科为例提出并实现了面向交叉融合学科知识组织的本体构建方法^[4],魏顺平提出了一种基于术语部件的领域本体自动构建方法,构建了教育技术学领域本体^[5]等。但在中小学教育领域的研究依然较少^[6],其中一些学者在教育领域本体构建方面进行了一定的探讨和研究。赵呈领等参考《义务教育物理课程标准(2011 年版)》、人教版《初中物理》教材、《初中物理知识清单》,以及从权威网站上获取的有关初中物理学科的教学资源,构建了初中物理学科的领域本体库^[7],王冰洁构建了小学英语学科本体^[8],丁国柱等构建了学习元语文本体架构^[9]等。

上述研究中,大部分中小学学科领域本体库都是根据课标,从教学角度出发,根据某一版本教材,按照一定的教学环节顺序构建本体库,其优势在于可以为教师或学习者推荐最适合这节课教学环节的所需要的资源,但其缺点在于忽略了学习者自身的能力等级。其中王冰洁构建的本体同样涉及到了小学英语学科,其研究站在教师角度,解决支持教师快速准确定位所需资源的知识本体构建问题,但无法根据学习者特征推送资源。本研究站在学习者角度,解决支持小学英语自适应学习的知识本体库构建问题。

资源本体库需满足小学英语学习自适应系统中各个模块的需求进行开发,提取多个纸质教材的知识点,整理并标注关联关系,并创建批量导入方式进行实例填充,最终形成满足自适应学习的小学英语资源本体库。

本文是根据学习者的角度来解决如何构建本体,它支持在小学英语自适应学习的一个问题。知识之间的关系可以通过本体比关系型数据库中反映更清晰,使学习者可以形成自己的知识网络。不同系统需要构建不同结构的本体知识库,本文构建的小学英语的本体知识库跟其他英语知识库相比,面向的自适应英语学习系统的结构更合理、功能更全面,需要本体知识库完美支持该系统的结构和功能,进而满足小学英语自适应学习系统的各个模块的具体要求。这需要我们从多个教科书中提取知识,整理和注释的关系,使用批处理模式导入实例,最终形成符合小学英语自适应学习的本体。

1 小学英语学科知识本体构建框架

本研究涉及的小学英语自适应学习系统需要根据不同学习者的学习情况进行个性化推荐,动态地调整学习内容和测试内容,这就要求系统中的各个模块在相互配合和调用时满足自适应学习要求。该系统中的数据层由资源本体库模块和学习模块组成,服务层由用户本体模块、测试模块以及机器人模块构成,如图 1 所示。

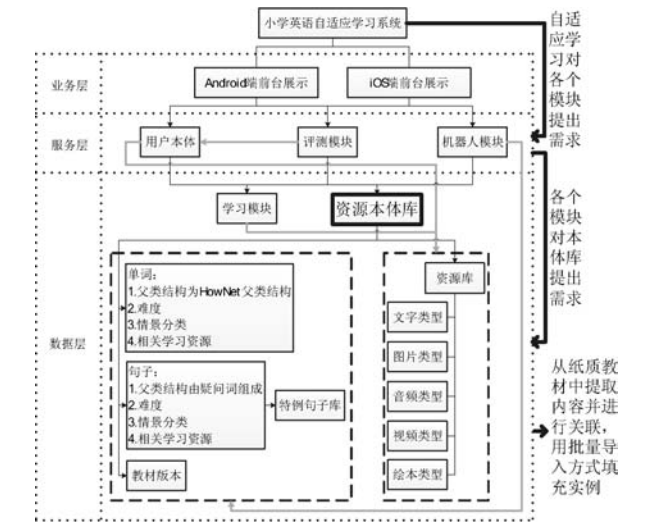


图 1 需求与本体结构对应关系图

评测模块会根据用户本体库中的信息找出用户最擅长与最不擅长的题型来得出测试题类型,测试内容来自学习模块中的学习内容,而学习内容调用的是资源本体库中的具体学习资源,机器人模块页调用资源本体库来应对用户提出的问题。

资源本体库需要满足与其交互的所有功能模块的调用以及查询需求,具体模块功能描述如下:

1.1 学习模块

需要根据用户本体中不同学习者的学习特征展示出内容偏爱的且符合其自身程度与资源类型的知识点。知识点包括单词和句子以及相关资源,即需要知识点对应的每一个学习资源有不同的类型和不同的难度。

1.2 机器人模块

机器人模块在与学习者对话时需在本体库中查找与提问句子相似度最高的问题并返回回答结果。在查找某单词及其相应资源时,若该单词不在本体库中,需要在 HowNet^[10]语义词典中查找该单词的父类,根据

父类查找到其下层的单词实例,对比出在本体库中相似度最高的单词,并把该单词及其资源推送给学习者。考虑到小学生会在提问的过程中输入中文,我们利用 HowNet 语义词典可以把中文翻译成英文,之后再行上述查找操作。在进行资源推荐时,需要根据学习者自身能力选择推荐适合难度的相关知识点,以及其他版本出现的相同知识点。

1.3 评测模块

需调用用户本体得到用户学习特征,再从知识本体库中调取适合此学习偏好的资源,处理成题目和选项,组成测试题。

根据以上对其他模块的功能和需求分析,需把资源本体库分为单词本体、句子本体、教材版本本体以及资源库四个部分。其中单词本体中的父类结构为 HowNet 的父类结构,在满足查找需求的同时保证了单词结构的稳定性。由于机器人模块的语句相似度算法是按照疑问词查找的,所以句子本体的父类则由疑问词组成,但这可能导致遗漏掉意思相近但疑问词不同的句子。例如 How about you? 和 What about you? 的疑问词分别是 How 和 What,但意思均为“你呢?”,这就需要人工找出所有特例句子,建立一个特例句子库,以满足问答的需求。资源自适应推荐时需要根据学习者程度推送,因此单词本体和句子本体中要有对难度控制和教材版本的描述。要构造支持自适应的资源本体库,必须符合小学生的思维特点和认知规律,并且以隐形教学方法为主,通过情景和联系展开学习,因此每个知识点都有对应的情景分类。系统在自适应推荐资源时针对该学习者偏好进行资源展示,因此资源库中的资源需区分为不同类型,类型分为文字类型、图片类型、音频类型、视频类型、绘本类型等。

2 小学英语学科本体构建

2.1 本体构建思路框架

本体结构确定后,需要人工提取本体术语和本体术语关系。其中本体术语即本体中的内容,本体术语关系即本体术语之间的关联关系,将这两个部分链接在一起即本体术语关系标注。由于本体术语以及本体术语关系数量较大,所以选择使用代码代替人工进行关系标注,从而形成完整的本体。本体构建思路框架如图 2 所示。

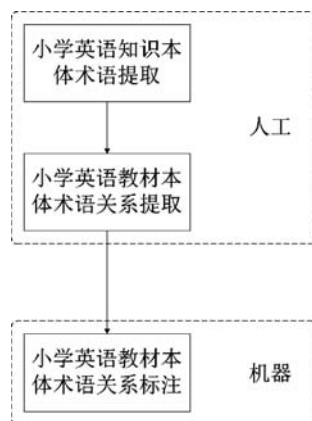


图2 本体构建思路框架图

2.2 本体术语提取

2.2.1 小学英语学科领域本体构建基础

1) 确认本体构建的目的

形式化地描述小学英语知识点及其相关资源和关系,以根据不同学习者各自的学习需求,为其推荐相应不同类型和内容的学习资源,即支持学习者自适应学习。

2) 资源的抽取与整理

为了符合一线教学要求,并且可以更加全面地覆盖小学英语的知识点,本研究参考了以下内容:《义务教育英语课程标准(2011)版》,北师大版、北京版、外研社新标准、外研社(一年级起点)教育部审定 2013 版、人教版、朗文版这 6 本纸质教材,以及课外读本《全脑思维图解英汉词典》。

3) 本体术语的分类

小学英语学科领域的本体结构确定为单词本体、句子本体、教材版本本体以及资源库四个部分之后,即可抽取本体术语,需要抽取本体术语的本体包括单词本体、句子本体和教材版本本体。本体术语可以从上述 8 个部分中抽取出来,并且提取出的术语可以分成两个类型:知识本体术语和教材组织本体术语。知识本体术语是指描述小学英语领域相关知识点及其资源的本体术语;教材组织本体术语是指描述小学英语学科教材结构的本体术语。

2.2.2 义务教育英语课程标准(2011)版中的本体术语

要对资源进行整理与抽取,就需要对课标和教材进行分析。由于《义务教育英语课程标准(2011)版》(以下简称课标)是我国国家教育部统一制定的针对小学英语学科基本规范和质的基本纲领性文件,所以本资源库的构建首先应该符合课标标准。课标课程总目标中的语言知识^[1]是英语知识内容的直接体现,它包括语音知识、词汇知识、语法知识、功能意念知识、话题知识这 5 个部分。对应这 5 个方面,需要抽取单

词发音音频、拼写与相关图片、词性与词性属性(对应不同词性各自的属性)、主题-功能意念以及主题-话题这 5 种本体术语。课标规定的九级目标体系中,二级目标是 6 年级结束时应达到的基本要求,但是不同版本的纸质教材也会选择 200~300 个的五年级单词,因此也要抽取出“难度”标准。

2.2.3 教材中的本体术语

教材是一线教学以及学习者获得知识的重要参考内容,目前我国小学英语学科教材有很多版本,本研究选择北京市小学常用的上述 6 个版本的教材,从中抽取本体术语。教材中的本体术语包括教材内容与教材结构两个方面。

1) 教材内容中的本体术语

小学英语的知识主要包括单词和句子两个方面。本研究以四年级为例,针对单词本体提取出 6 本教材中所有的单词,以及与单词相关的课文原句和情景段落;针对句子本体提取出问句与其回答,以及相应情景对话、疑问词、句型以及相关单词。这样单词本体与句子本体的内容是对应的,在学习单词的同时可以进一步学习如何在对话中使用该单词,反之亦然。教材中还包含一些静态资源,如图片、音频、视频等,在丰富了学习资源的同时,也保证了资源来源的权威性。

2) 教材结构中的本体术语

教材结构中的本体术语分为三层。第一层为教材版本,包含北师大版、北京版、外研社新标准、外研社(一年级起点)教育部审定 2013、人教版、朗文版这 6 本纸质教材。第二层为教材册数,描述教材所涉及的学期,小学阶段一共 6 个年级,12 个学期,所以使用 1 至 12 来表示教材册数。第三层为知识点所属单元,同样用整数表示。

2.2.4 课外读本中的本体术语

本研究参考外研社出版的《全脑思维图解英汉词典(小学点读笔版)》^[12],它使用思维导图展示了单词之间清晰的关系,非常符合本体的知识组织形式,可以作为本体构建的结构参考依据^[13]。该书的内容丰富了每个单词的资源,从而从中抽取出现有单词的联想、同义词、反义词、拓展、百科、用法、延伸例句、常用用法这 8 类内容,让学习者可以从多方位学习同一个知识点,加深印象,建立自己的知识网络,从而提高学习绩效。

2.2.5 小学英语学科各类本体术语数量汇总

通过上述三方面对本体术语的抽取,小学英语学科的本体术语数量汇总如表 1 所示。本研究的本体术语从课标、北师大版、北京版、外研社新标准、外研社(一年级起点)教育部审定 2013、人教版、朗文版纸质

教材,以及课外读本《全脑思维图解英汉词典》这 8 个角度进行抽取。

表 1 小学英语学科各类本体术语数量汇总表

类别		本体术语	数量
知识本体术语	单词本体术语	单词	2 118
		主题-功能意念	418
		主题-话题	633
		HowNet 中的 Taxonomy	1 105
		词性	1 089
		词性属性	800
		中文含义	1 105
		难度	2
		课文原句	1 063
		情景段落	963
		联想	864
		同义词	48
		反义词	101
		拓展	368
		百科	111
		用法	112
		延伸例句	668
		常用	403
	句子本体术语	问题疑问词	22
		问题	261
		回答	261
		情景对话	259
		句型	147
教材本体术语	教材版本		6
	册数		72
	单元		37

2.3 本体术语关系提取与标注

本体术语提取过后还需要提取本体术语之间的关系,本体术语与本体关系的建立即搭建了一个网状关系知识库,以便计算机可以进行推理推导出本体之间的语义关系。

2.3.1 小学英语学科领域本体构建基础

针对小学英语学科的特征,本研究涉及到了 23 种本体关系。综合这些本体关系的特点,本研究将这 23 种本体关系分为三大类,分别为层级关系、属性关系以及关联关系,如表 2 所示。

表 2 小学英语学科本体关系类别表

层级关系	上下位关系	有向关系
	成员关系	
	前后继关系	
属性关系	属性关系	无向关系
关联关系	同义词关系	
	反义词关系	
	相关关系	

1) 层级关系

层级关系即两个本体术语之间明显的层次结构关系,是有向关系,本体中使用 `rdfs:subClassOf` 和 `rdf:type` 表示。本研究中的层级关系包括上下位关系、成员关系以及前后继关系。上下位关系是指本体术语中父节点与子节点之间的关系,主要存在于知识本体中,本研究参考 HowNet 中 Taxonomy 的层级关系进行层次构建。成员关系表示本体术语 A 是本体术语 B 的一个组成部分,主要存在于资源库中,如图片、音频、视频、绘本分别为资源库的成员。前后继关系表示学期和单元之前的前驱与后继关系,一般存在于教材组织本体中。例如第二学期为第一学期的后继节点,反之为前驱节点,单元关系与学期关系类似。

2) 属性关系

属性关系是指本体术语 A 为本体术语 B 的一个特征性质,是有向关系,在本体中用数据类型属性(Data Property)表示,本研究所涉及的属性如表 3 所示。每个知识本体术语都具有表 3 中所呈现的属性,这些属性值共同描述了一个具体的本体术语。

表 3 小学英语学科本体属性表

类别	属性	注解
单词本体中的属性关系	主题-功能意念	对应该单词在课标中的功能意念目标
	主题-话题	对应该单词在课标中的话目标
	词性	包含名词、动词、冠词、代词、形容词、副词、数词、介词、连词、感叹词等
	词性属性	若词性为名词,词性属性则为可数名词或不可数名词,如 boat 的词性属性为 boats;若词性为动词,则词性属性为该单词的原形/第三人称/现在进行式/过去式/现在完成式,如 do 的词性属性为 do/does/doing/did/done
	中文含义	中文解释
	难度	按照课标规定,若为二级词汇则难度为 2;若为五级词汇泽难度为 5
	课文原句	该单词所在教材中的课文原句
	情景段落	该单词所在教材中的情景段落
	拓展	该单词同其他词语联合构成新的拓展词
	百科	与该单词所表达事物相关的人文、自然科学知识
	用法	该单词的同义词、反义词用法辨析、英美不同用法、常见错误说明
	延伸例句	课外参考书目中配合该单词的常用对话和单句
	常用	英语中该单词的固定搭配或常用搭配

续表 3

类别	属性	注解
句子本体中的属性关系	回答	课文原句中该问句对应的回答
	情景对话	该句子所在的情景段落
	句型	该句子所运用的句型
	相关单词	该句子中所用到的重点单词
句子本体中的属性关系	回答	课文原句中该问句对应的回答
	情景对话	该句子所在的情景段落
	句型	该句子所运用的句型
	相关单词	该句子中所用到的重点单词

3) 关联关系

关联关系是指两个本体术语之间具有语义相关或相似的关系,在本体中用数据类型属性表示,所有关联关系如表 4 所示。

表 4 关联关系表

关联关系	特征(Character)	注解
同义词关系	Transitive(传递性), Symmetric(对称性), Reflexive(自反性)	与该单词意义相近的词
反义词关系	Symmetric(对称性), Irreflexive(反自反性)	与该单词意义相反的词
联想关系	Symmetric(对称性)	与该单词所表达的内容相关的其它单词

这 3 种关联关系涉及 4 种特征,具体描述如下:

(1) Transitive(传递性)

若该关系具有传递属性,那么对于任意 x,y 和 z , 则 $P(x,y)$ 与 $P(y,z)$ 蕴含 $P(x,z)$, 例如 trip 是 journey 的同义词,journey 是 tour 的同义词,那么 trip 也是 tour 的同义词。

(2) Symmetric(对称性)

若该关系被声明为对称属性,那么对任意 x 和 y , 则 $P(x,y)$ 当且仅当 $P(y,x)$, 例如 white 是 black 的反义词,那么 black 也是 white 的反义词。

(3) Reflexive(自反性)

若该关系是自反的,则存在 $P(x,x)$, 例如 white 是自身的同义词。

(4) Irreflexive(反自反性)

若该关系是反自反的,则不存在 $P(x,x)$, 例如 white 不是自身的反义词。

2.3.2 本体术语关系标注

对本体术语和本体关系进行抽取之后,就要实现把本体关系搭建在本体术语之间,即本体术语关系的标注。目前本体术语关系标注的方法有自动标注法和

人工标注法,由于本研究的本体术语和本体属性数量较多,若使用人工标注所需的人力精力较大,所以本研究选择自动标注法完成本体术语关系标注。在使用 Protégé 软件对本体进行构建的过程中,其自身的树形结构层次已经自动标注生成了层级关系写入 RDF 文件中,使用 Jena 框架可以直接读取层级关系信息,实现本体术语关系的自动标注。因此,先使用 Protégé 软件构建具有树形结构的类结构、对象属性与数据属性结构的类。每个实例的类都是基于 HowNet 的 Taxonomy 结构的,计算机需从 HowNet 中查找到该单词的父类,之后用 Jena 框架自动标注每个实例的与类的上下位关系。这样就减少了大量人工工作量,提高了工作效率。

本研究中的实例填充使用批量导入方式,可以先用 Excel 整理好所有属性,避免使用 Protégé 软件人工填充的复杂操作。并且若本体大小超过 JDK 的内存限制,则导致内存溢出,无法继续在 Protégé 软件中运行。因此本研究使用更加易操作的批量填充实例法,程序流程图如图 3 所示。

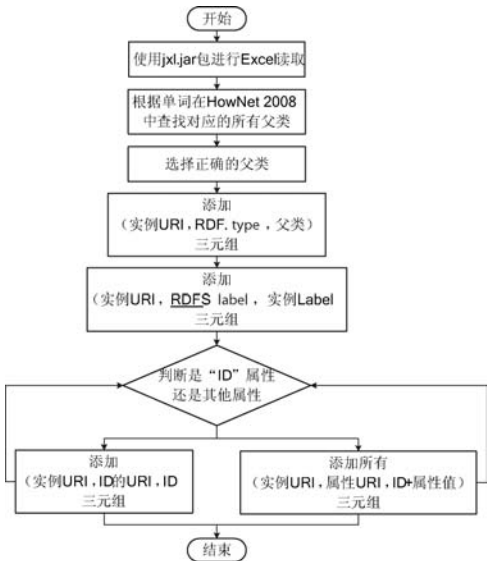


图 3 批量添加实例程序流程图

2.3.3 本体知识库构建

经过上述所有操作之后,本研究的知识库结构如图 4 所示。

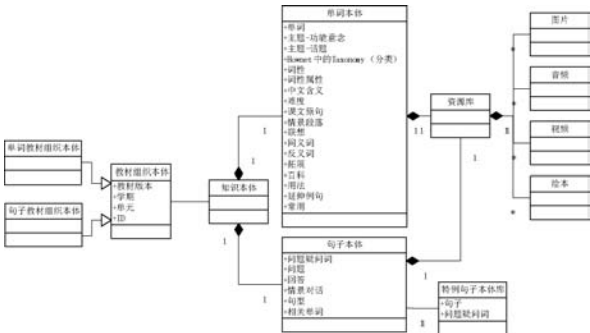


图 4 本研究知识本体结构图

本研究的知识本体分为 3 部分:知识本体、教材组织本体和资源库。知识本体由单词本体和句子本体构成,具有上述不同的属性。每个单词和句子都对应不同的资源,资源中包括图片、音频、视频、绘本等资源,以应对不同学习者的不同需求。由于本系统中计算句子相似度模块的算法是在相同的疑问词中查找相似度最高的句子,这就会遗漏不同疑问词中语义相似度高的句子,所以建立一个特例句子本体库,链接此类句子,以提高查找精度。单词教材组织本体和句子教材组织本体都继承教材组织本体,包含教材特征的所有属性。该本体作为知识本体中的一个属性呈现,即每一个单词和句子都有的那个独一无二的 ID,表现形式为教材版本/学期/单元/ID,例如:backpack 的教材本体为 1/10/7/1,代表该单词出现在北师大版教材中第 10 册第 7 单元中。

2.4 本体术语的形式化编码

Fuseki 数据库中的三元组部分截图如图 5 所示。

<http://www.semant ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_14f80e b6_a211_4520_a4 65_d3fb0057bee5>	<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_28c5fcc 3_920f_486d_beee _0f2abac08bce>	"(2/8/17/2)船是重要 的水上交通工具。 在石器时代,就出 现了最早的船—— 独木舟。"@zh
<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_36c776 e2_486c_4bb4_81 19_c40cb90c62b4 >	<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_4ca74c 2b_2ea9_43ed_a7 18_539813773729 >	"(2/8/17/2)17.旅游 与交通- (58) 交通 运输方式"@zh
<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_7537ef 9d_005a_4ca2_91 2c_ad76bf3cf8a2>	<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_806996 51_270d_4d1b_9d b2_46537509ff54>	"(2/8/17/2)北京 版"@zh
<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_806996 51_270d_4d1b_9d b2_46537509ff54>	<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_806996 51_270d_4d1b_9d b2_46537509ff54>	"(2/8/17/2)1.They also have dragon boat races.It sounds fun."@zh
<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_806996 51_270d_4d1b_9d b2_46537509ff54>	<http://www.seman ticweb.org/primary EnglishOntology/fo urthGrade#OWLDa taProperty_806996 51_270d_4d1b_9d b2_46537509ff54>	"(2/8/17/2)小船: 艇"@zh

图 5 Fuseki 数据库中的三元组部分截图

在对本体术语的抽取与本体关系的提取与标注之后,需要将这些存储在计算机中,方便计算机对本体进行计算和处理,以实现资源聚合。由于本研究的数据量较大,若选择文本存储,会导致低效的数据处理。因此本研究选择使用数据库存储方式。数据库存储方式包括关系数据库存储和非关系数据库存储,虽然关系数据库存储技术较为成熟,但其对语义的兼容性不是很好,所以本研究选择语义数据库——Fuseki 数据库^[14]。Fuseki 是一个提供访问 SPARQL1.1 标准的服务器,增强了计算机对数据的理解力和处理程度,可以通过 HTTP 使用 SPARQL 语句对数据库中的三元组进行增加、删除、修改、查找等操作。

2.5 确认与评估

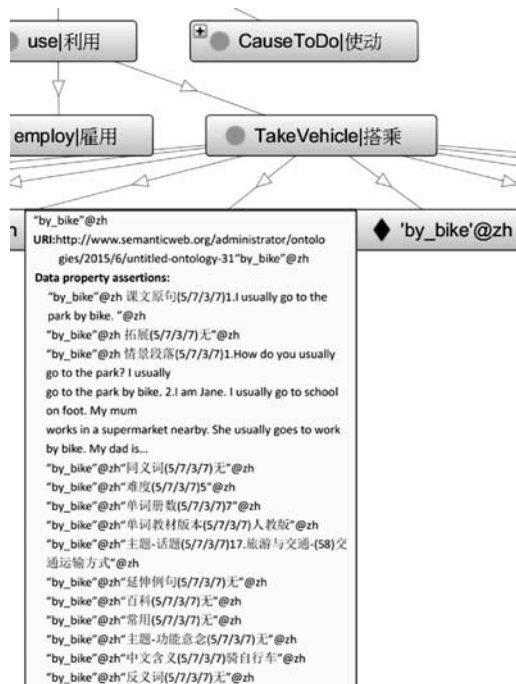


图6 用 Protégé 软件展示的本体部分截图

经过上述的本体构建步骤后,构建出了初步的小学英语学科领域本体,由于所构建的本体并非是静态的。它会根据学习者学习的过程不断修改和完善本体术语和关系的识别以及本体关系的标注,所以在这个过程中要保证本体的正确性、一致性和有效性。本研究中的知识本体术语和教材本体术语都是从课标以及一线教材中抽取出来的,课标是教育部颁发的针对课程质量的指导性文件,而一线教材则是根据课标编纂的;资源库中的内容一部分来源于一线教材中的静态资源,另一部分通过对内容分析得出,这部分由三位一线教师共同整理和抽取,因此这部分本体术语具有较强的权威性和正确性。本体关系标注的过程中,除了计算机自动标注出的本体层级关系以外,其余的属性关系和相似关系都是通过计算机读取教师整理好的

Excel 文件,使用 Jena 框架批量添加处理的。在未来本体库的不断修改和完善的过程中,依然是由教师把关进行,以确保本体的正确性、一致性。为了知识推理操作,还需把 Fuseki 中的所有三元组重新写回. owl 文件中,写回的. owl 文件用 Protégé 软件打开部分截图如图 6 所示。

本研究的小学英语学科本体库作为底层数据库为 iOS 端与 Android 端 APP 提供数据支持,根据学习者的认知水平和学习偏好,为知识学习模块提供知识点资源展示数据,为评测模块提供题目和选项数据,为机器人模块提供问题相应回答数据。各个模块满足小学英语自适应学习系统的需求,本研究本体库能满足各个模块提出的需求,保证本体的有效性。

3 结 语

本研究中将本体技术运用在小学英语领域中,作为自适应学习系统中的底层数据库可以加强系统自适应推荐的性能,促使学习者形成符合其自身的知识网络结构。本研究基于小学英语学科,在满足学生和广大一线教师需求的基础上,设计了针对该领域的本体库构建方案。其中重点描述了有机融合多版本教材的本体术语抽取与本体关系的提取和标注,对知识进行多维度描述,并使用计算机自动标注本体关系,将生成的 OWL 文件以三元组的形式存储于 Fuseki 数据库中,并使用计算机批量添加实例,最终完成知识本体库的构建。

本研究中仅以四年级为例进行本体知识库的构建,将在以后的研究中不断添加更新。在后续研究中,可以将一个知识点中某个类型资源标注出不同难度等级,以适应不同学习者能力需求。

目前中国的英语学习资源十分丰富,但是如何让学习者快速准确地查找到适合自身的资源是困难的。本文开发了支持小学英语自适应学习系统的本体库,使用本体可以加强系统自适应推荐的性能,促使学习者形成符合其自身的知识网络结构。旨在为小学英语学科自适应学习系统的资源库构建提供一个可借鉴的案例。目前该本体库已经应用于自适应系统中,并正在针对试用进行最后的调试。

参 考 文 献

- [1] 曹双双,王移芝. 泛在学习自适应学习系统模型研究[J]. 现代教育技术,2012,22(7):101-104.
- [2] 姜强,赵蔚. 自适应学习系统述评及其优化机制研究[J]. 现代远程教育,2011(6):57-63.

元组时,多级块折叠模型往往只需要更少的物理元组。在构造查询结果时,所需要进行的连接操作也更少。在实验场景中,与块折叠表数据模型相比,多级块折叠表数据模型的平均查询延时减少了 28.04%,吞吐率提高了 69.68%。

4 结 语

本文对块折叠表数据模型进行了拓展,提出了一种适用于多租户数据存储的多级块折叠表数据模型。该数据模型通过设置基础数据表存储租户同构的基础数据,通过设置多组多张块内结构各异的块表存储租户的自定义数据,本文通过一个多级块折叠数据模型、块折叠表数据模型和弹性拓展表数据模型的对比实验,验证了多级块折叠表数据模型具有较高的查询性能与空间效率。本文实验分析部分对多级块折叠表数据模型所做的有效性分析为一般性分析,如何结合生产环境中更多的实际使用的数据库引擎的缓存、索引等技术,对数据模型的有效性进行进一步分析与验证,是本文未来研究的方向。

参 考 文 献

- [1] Chong F, Carraro G. Architecture Strategies for Catching the Long Tail [EB/OL]. (2006-06) [2016-03-21]. <http://msdn.microsoft.com/en-us/library/aa479069.aspx>.
 - [2] Chong F, Carraro G, Wolter R. Multi-Tenant Data Architecture [EB/OL]. (2006-06) [2016-03-21]. <https://msdn.microsoft.com/en-us/library/aa479086.aspx>.
 - [3] Aulbach S, Grust T, Jacobs D, et al. Multi-tenant databases for software as a service: schema-mapping techniques [C]// ACM SIGMOD International Conference on Management of Data. ACM, 2008:1195-1206.
 - [4] Saraswathi M, Bhuvaneswari T. Multitenant SaaS model of cloud computing: Issues and solutions [C]// International Conference on Communication and Network Technologies. IEEE, 2014:27-32.
 - [5] Fang S, Tong Q. A comparison of multi-tenant data storage solutions for Software-as-a-Service [C]// International Conference on Computer Science & Education. IEEE, 2011:95-98.
 - [6] Yaish H, Goyal M, Feuerlicht G. An Elastic Multi-tenant Database Schema for Software as a Service [C]// IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing. IEEE, 2011:737-743.
 - [7] Yaish H, Goyal M. A Multi-tenant Database Architecture Design for Software Applications [C]// IEEE, International Conference on Computational Science and Engineering. 2013:933-940.
 - [8] Gorti I, Shiri N, Radhakrishnan T. A Flexible Data Model for Multi-tenant Databases for Software as a Service [C]// IEEE, International Conference on Computational Science and Engineering. 2013:1059-1066.
 - [9] 唐圣潘,周肆清,丁长松. 基于 SAAS 模式的共享模型的数据扩展技术研究 [J]. 计算机技术与发展, 2011, 21(8):63-66.
 - [10] 周文琼,李庆忠,范路桥,等. SaaS 模式多租户数据存贮模型的研究与实现 [J]. 计算机科学, 2013, 40(10):194-197.
 - [11] Weissman C D, Bobrowski S. The design of the force.com multitenant internet application development platform [C]// ACM SIGMOD International Conference on Management of Data, SIGMOD 2009, Providence, Rhode Island, Usa, June 29-July. 2009:889-896.
 - [12] Ni J, Li G, Wang L, et al. Adaptive Database Schema Design for Multi-Tenant Data Management [J]. IEEE Transactions on Knowledge & Data Engineering, 2014, 26(9):2079-2093.
-
- (上接第 24 页)
- [3] 胡瑛,贾积有. 学习对象内容本体描述框架研究 [J]. 开放教育研究, 2009, 15(2):102-106.
 - [4] 郑艺,应时,陈旭. 面向交叉融合学科知识组织的本体构建方法 [J]. 情报杂志, 2014(3):143-149.
 - [5] 魏顺平. 基于术语部件的领域本体自动构建方法研究——以教育技术学领域本体构建为例 [J]. 电化教育研究, 2013(5):62-67.
 - [6] 马捷,刘小乐,黄岚,等. 教育领域本体构建研究 [J]. 情报理论与实践, 2012, 35(7):104-108.
 - [7] 赵呈领,黄志芳,万力勇,等. 基于初中物理课程的学科领域本体库构建研究 [J]. 电化教育研究, 2014(8):64-70.
 - [8] 王冰洁. 基于语义网的小学英语资源动态聚合系统设计与开发研究 [D]. 北京:北京师范大学, 2013.
 - [9] 丁国柱,余胜泉. 基于本体学习算法的学科本体辅助构建研究——以学习元平台语文学科知识本体的构建为例 [J]. 中国电化教育, 2015(3):81-89.
 - [10] 董振东,董强. HowNet [OL]. 2016-03-22. <http://www.keenage.com/>.
 - [11] 中华人民共和国教育部. 义务教育英语课程标准:2011 年版 [M]. 北京师范大学出版社, 2011.
 - [12] 章思英. 全脑思维图解英汉词典:小学点读笔版 = Mind Map Primary English-Chinese Dictionary: Primary E-pen Edition [M]. 外语教学与研究出版社, 2011.
 - [13] 邱聃. 基于思维导图的小学教育语义本体库构建 [J]. 电子设计工程, 2016, 24(3):53-56.
 - [14] Jena A. Apache Jena Fuseki [Z]. The Apache Software Foundation, 2014.