

基于语义的XBRL财务报告共享与集成

张颖敏^{1,2}, 潘 定²(博士生导师)

【摘要】 本文针对XLink标记技术在语义表达和推理能力的缺陷,以及XBRL财务元素定义的多义性,提出多个组织之间的XBRL财务报告元数据的共享与集成模型,引入分布式扩展的时态描述逻辑DTD LBR对XBRL元数据进行语义形式化表达。通过元数据库管理多个组织的XBRL财务元数据,开发相应的接口程序实现XBRL本地本体的自动构建,最后编写推理规则并通过Jena推理引擎进行XBRL分类标准本体元数据的一致性检验,结果表明推理规则有效。

【关键词】 XBRL; 描述逻辑; 语义; 本体; XBRL集成

【中图分类号】 F275 **【文献标识码】** A **【文章编号】** 1004-0994(2016)09-0113-4

一、引言

XBRL是基于XML的可扩展的企业报告标记语言,目前主要用于企业对外财务报告的信息处理。经过十多年的推广应用,XBRL财务报告已经成为最有效的网络财务报告形式,被广泛应用于银行、证券、金融等领域。近年来,中、日、韩、欧美等国家纷纷强制或鼓励上市公司和金融机构发布XBRL财务报告。为了达到XBRL财务报告的可比性,XBRL国际组织和各国监管机构发布了一系列地区、行业的技术和管理规范,力图既能满足企业的个性化信息披露需求,又能保证XBRL财务元数据的一致性。但实践证明,目前尚无有效技术手段保证多个组织之间的财务元数据一致性。如何保障XBRL财务信息的可比性与操作性,实现多个组织之间的XBRL的共享与集成对于加快XBRL的推广具有重要意义。

本文将从XBRL元数据语义的角度,结合XBRL财务元素概念定义的多义性特点,引入分布式描述逻辑DTD LBR,试图研究基于元数据语义一致性的多个组织之间的XBRL财务元数据的集成管理方法。

二、多个组织之间的XBRL的集成模型

XBRL是一个主要针对财务报告信息化处理的标准化体系,其体系框架由XBRL技术规范、分类标准和实

例文档三个部分组成。其中,技术规范规定了分类标准和实例文档的主要结构以及所涵盖的财务元素;分类标准则阐明了财务元素的语义和语法,相当于帮助理解实例文档的词典;实例文档是经过转换可以最终呈现给用户的财务报告。

要想使分布在全球各地的XBRL财务信息能够实现集成、共享、自动组合、分析等高层次的智能化应用,首先需要解决的是多个组织XBRL系统(尤其是各个分类标准)之间的映射问题。本文借鉴Borgida(2003)提出的分布式描述逻辑中的包含桥规则概念,建立如图1所示的基于描述逻辑包含桥规则的多个组织之间的XBRL元数据集成模型。

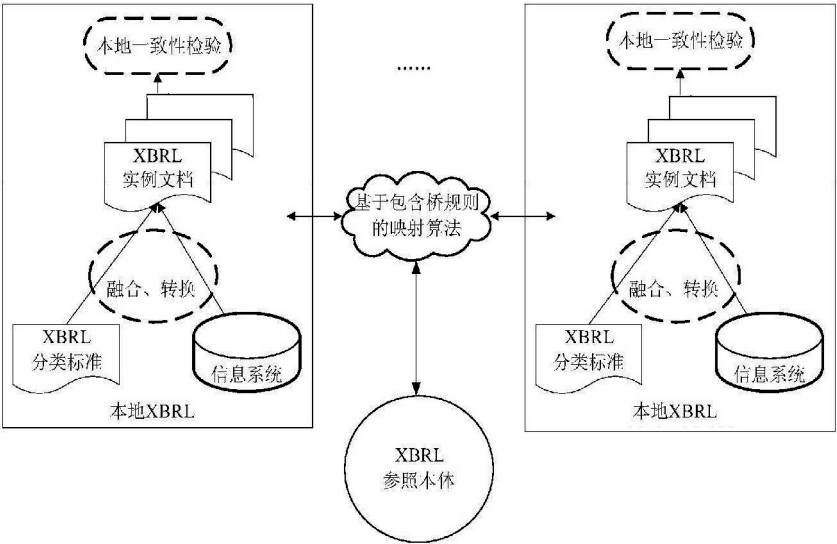


图1 多个组织之间的XBRL集成模型

【基金项目】 国家自然科学基金项目(项目编号:71171097;70771044)

图1所示的模型中,企业根据本地指定的分类标准(可在扩展的基础上),从企业的经营运作系统中进行数据的抽取、融合,转化为同一组织的XBRL实例文档,经过一致性检验后,通过对XBRL参照本体中的分类标准以及实例文档中的元数据及其关系的包含桥规则映射,实现多个组织之间的XBRL报告信息的共享与集成。本文将重点关注多个组织之间的XBRL财务数据映射。

1. 参照本体的构建。集成模型中的参照本体,主要是指包含财务元素丰富语义说明的分类标准。经过十多年的推广应用,XBRL分类标准的国际化与标准化的工作已得到业界的高度认可。会计准则作为会计核算工作必须遵守的规范和准则,是分类标准制定的主要参照,笔者认为可以选择公认度较高的国际会计准则作为构建XBRL分类标准参照本体的依据。

美国财务会计准则委员会(FASB)发布的公认会计准则(GAAP)和国际会计准则理事会(IASB)发布的会计准则(IAS)是业界公认最具影响力的会计准则。其中,IAS定位为全球通用的会计准则,而GAAP则主要针对美国国内的经济活动而制定,其通用性较IAS差。2011年IASB的监督机构IFRS(国际财务报告准则)基金会制定完成了IFRS的XBRL分类标准及其扩展。鉴于IFRS的国际通用性,本文将选择该套分类标准及其扩展作为XBRL集成模型的参照本体。

2. 多个组织之间的XBRL映射。XBRL共享与集成模型在保证同一组织发布的XBRL分类标准财务元数据一致性的基础上,通过包含桥规则实现本地XBRL本体与参照本体之间的映射,达到多个组织之间的XBRL财务数据的可比与共享集成。针对XBRL分类标准的多义性问题,李吉梅等(2012)研究了我国财政部、上交所和深交所各自发布的XBRL分类标准的语法和语义特征,将上述三套分类标准的异构类型分为结构(语法)异构和语义异构。本文研究XBRL财务元数据的语义一致性,将重点关注多个组织之间的XBRL分类标准的语义歧义与映射问题。

语义异构是指元素概念所表达的语义冲突,即语义歧义的问题。XBRL财务报告中,元素以及元素之间的关系主要通过分类标准中的模式文件以及链接库文件来获得。对元素概念的语义通过元素属性以及参考链接来描述。其中,元素属性用于同一组织的XBRL财务元素的区分,主要通过id、name等属性来标识;参考链接用于表明元素的外部权威参考来源,以判断元素的引用是否一致,如果一致则说明概念来

源于同一权威参考,可认为是同一概念。

下面将分析比较我国财政部和国际会计准则委员会发布的分类标准(CAS和IFRS)对财务元素概念的描述(如表1、表2所示),以进一步研究XBRL财务元素与参照本体的映射规则。

表1 CAS与IFRS对营业收入的属性描述比较

元素属性 分类标准	id	name	Xbrli:balance	Xbrli:periodType	type
CAS	cas_OperatingRevenue	OperatingRevenue	credit	duration	brli:monetaryItemType
IFRS	ifrs_Revenue	Revenue	credit	duration	brli:monetaryItemType

表2 CAS与IFRS对营业收入的参考链接描述比较

元素属性 分类标准	xlink:type	xlink:title	xlink:href
CAS	locator	OperatingRevenue	http://xbrl.iasb.org/taxonomy/2010-04-30/ifrs-cor_2010-04-30.xsd#ifrs_Revenue
IFRS	locator	loc_65	../.../ifrs-cor_2010-02-15_ed.xsd#ifrs_Revenue

从以上两套分类标准对营业收入的描述比较可以看出,虽然两者对营业收入元素的标签和标识不同,但参考链接的URL路径xlink:href均指向ifrs的核心模式文件,并且参考同为#ifrs_Revenue元素,即其权威引用相同,可以认为两者具有语义一致性。

根据上述XBRL财务元素的语义一致性规则,建立不同组织发布的XBRL分类标准与IFRS分类标准之间的元素映射与语义检测规则库,可为多个组织之间的XBRL财务报告的共享与集成分析提供基础。

3. XBRL财务元数据的形式化。XBRL中含有大量的元数据,用于描述商业事实的背景参数,具有明确的语义。为了加强XBRL元数据的可比性,更好地表示XBRL技术规范的元数据语义,XBRL国际组织于2011年发布了XBRL抽象模型1.0,于2012年修订并发布其最新版的XBRL抽象元模型2.0,该元模型以概念及建模方式展示了XBRL技术规范中元模型的抽象表达,可以加深技术和非技术人员对于XBRL的认识,以便开发出相应的应用和接口程序。

本文将引入分布式扩展时态描述逻辑DTD LBR对多个组织之间的XBRL集成模型中的映射规则进行形式化表达,在此基础上构建多个组织之间的XBRL财务元素的语义映射规则库。根据元数据管理的三层结构模型(从抽象到具体,自上而下分别为:元模型→元数据→数据),XBRL元数据可以对应为XBRL元模型中元类的实例,财务元素之间的关系可以对应于元类之间关联关系的实例。如CAS中对利润表的营业收入与营业成本(Operating Cost)、营业税金及附加

(Business Tax And Surcharge)等元素之间的聚合关系可以形式化为:

Operating Revenue ⊃ Operating Cost ⊃ Business Tax And Surcharge ⊃ Distribution Costs ⊃ Administrative Expenses ⊃ Finance Costs ⊃ Impairment Loss ⊃ Profit Or Loss Arising From Changes In Fair Value ⊃ Investment Income ⊃ Operating Profits ⊃ Income Statement Line Items

上式表示营业成本、营业税金及附加等元素通过聚合关系组成利润表列报事项元素(Income Statement Line Items)。

4. XBRL 财务元数据的一致性检验。元数据一致性是保障文件系统可用性与可靠性的最基本要求,只有元数据描述一致才能保证数据编辑与使用的正确性。互联网之父 Tim Berners-Lee 认为语义 Web 的语义是形式符号的指称。因此,语义一致性在语义 Web 的层面来说,是指形式化的符号与其指向的对象是否一致的问题。

XBRL 分类标准是相关财务报告领域权威性术语的“词汇”,具有明显的本体特征。本文将以本体作为其语义检测的主要技术,运用表达力较强的时态扩展描述逻辑 DTDLBR,对 XBRL 分类标准中的财务元素概念以及概念间的关系进行形式化表示,并以本体自身的概念一致性检验功能以及知识库的推理机来检测其语义冲突,以此保证分类标准编制的正确性。

XBRL 分类标准中财务元数据的一致性包括财务元素概念以及概念间关系的一致性。XBRL 模式文件对财务元素

的定义主要通过元素名以及元素的 6 个基本属性(name、type、xbrli:periodType、nilable、abstract、substitutionGroup)、与该元素相关的标签以及参考链接的 href 属性来描述。由于是对同一组织的 XBRL 分类标准的检测,不允许存在同名元素定义,因此,如果模式文件的概念同名则认为是同一个概念。此外,对于同一个 XBRL 分类标准来说,如果财务元素不同名,但上述 6 个基本属性和参考链接均相同,则也认为是相同的元素概念。因此,对元素概念的一致性检测主要从上述两方面来展开。对于概念间关系的一致性,则需要引入相应的技术规范,如制定分类标准的会计准则,编制一致性检测规则以实现元素关系的冲突检测。

三、实证分析

对于 XBRL 财务元数据的一致性检验,本文主要利用本体技术来实现,需要数据转换、元数据存储以及本体推理机等技术支持。因此,综合上述对 XBRL 财务元数据的一致性检验任务,本文通过构建元数据存储库,实现本体元数据的实时自动获取。在此基础上,设计从元数据库到本体的自动映射与转化的原型系统,该系统能够支持从 Oracle 元数据库到 OWL 本体的自动转换,再通过本体编辑工具 Protégé 内置的 Pellet 推理机实现对本地本体的一致性检验。由于多个组织之间的 XBRL 涉及多个 XBRL 分类标准本体,目前还没有专门支持多本体的软件和工具,因此,本文将建立多本体的一致性推理规则,并通过 Jena 推理机来实现多本体的概念一致性检验。

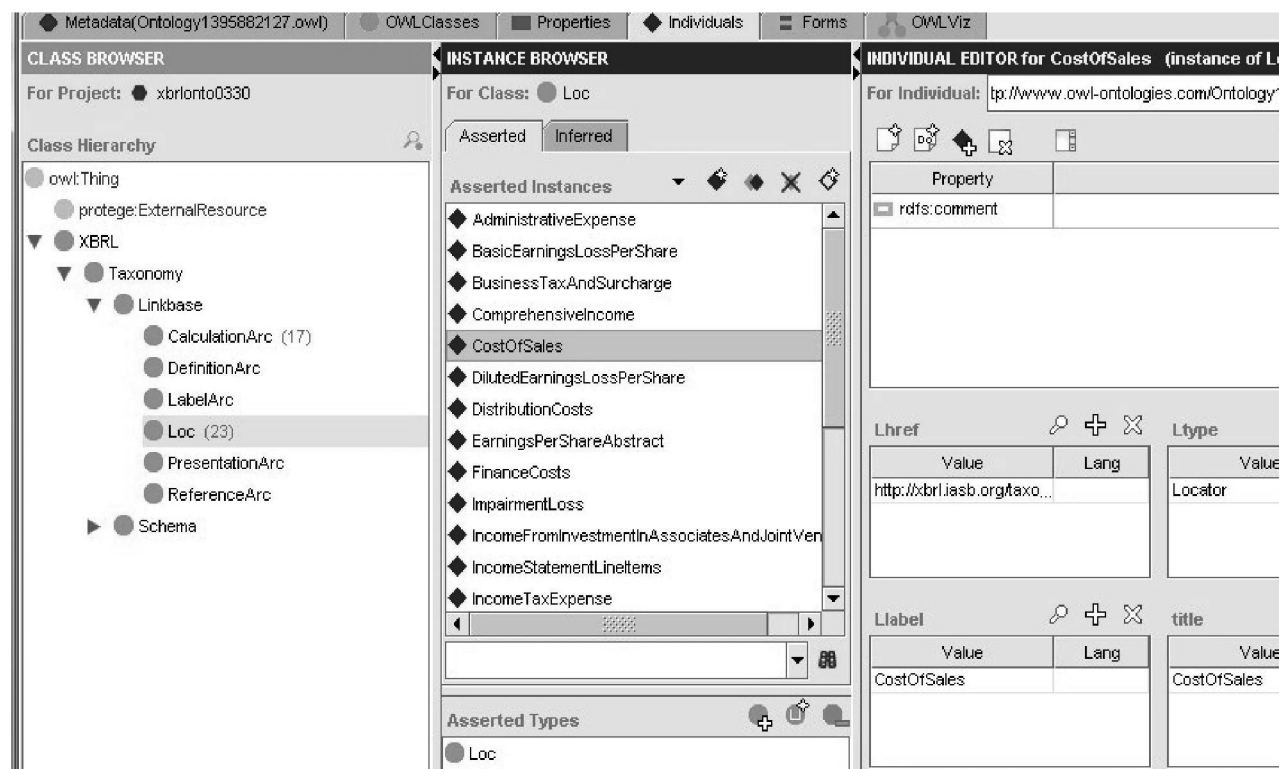


图 2 protégé构建的 XBRL CAS 分类标准本体片段

1. 本体的构建。目前还没有对分布式本体的编辑和管理工具,因此,本文考虑开发Jena的推理接口,实现分布式本体的概念一致性检测。

元数据存储库Metabase可用于存储和管理多个本体的元数据,因此,本文考虑建立多个本体的元数据存放于Metabase中,然后通过Jena的推理规则对多个分布式本体的元数据进行一致性检测。下面以财政部的CAS分类标准和深交所的FR分类标准为例,在Metabase中建立并存储分类标准元数据。通过开发元数据读取接口程序,从Metabase获取分类标准的元数据并创建本体,Protégé通过导入数据生成的本体结构如图2所示。XBRL本体的建立为后续的本体推理、一致性检验奠定了基础。

2. 单个XBRL本体的一致性检验。对单个本体的一致性检验可以通过本体推理引擎来实现。考虑推理规则的可扩展性,本文将选择Jena推理机,并通过编写推理规则,实现对多个组织之间的XBRL分类标准的概念一致性推理。具体来说,就是把写好的规则保存到规则库文件中,通过Jena对规则库规则进行检测。

为验证推理规则的正确性,本文创建了一个测试类testArc,并定义该类与营业利润(Operating Profits)和营业收入(Revenue)的弧关系具有相同的个体Operating Profits to Revenue。Jena的推理结果表明,推理规则能发现冲突;该类同时属于相离的两个类,其个体也存在冲突。

3. 多个组织之间的XBRL本体一致性检验。由于参考链接是对元素的权威出处的描述,可以认为如果参考链接的URL地址相同,则认为是相同的概念。因此,对多个组织之间的XBRL分类标准的概念一致性检测可以通过参考链接的判断来实现。具体可以通过建立以下推理规则来实现:

Rule1:(?x http://www.owl-ontologies.com/
Ontology1395882127.owl#xlink:href?y)->(x Equal y)

运用Jena推理机对上述推理规则进行检测。研究结果表明,本文建立的推理规则能正确检测元素概念和参照关系的一致性。如果两个XBRL财务元素的参考出处不同,则需要进一步分析该元素与其他元素对应的计算关系,以及计算关系对应的元素的定义,以判断语义表达是否一致。

四、结论

本文首先根据多个组织之间XBRL财务元素概念的多

义性特点,构建了多个组织之间的XBRL的集成模型。该模型在保证单个XBRL元数据一致性的基础上,引入分布式时态描述逻辑DTD LBR对多个XBRL之间的元数据语义进行形式化表达,分别从同一组织和多个组织之间两个层面对XBRL的财务元数据一致性进行检测。

实证中应用元数据存储库Metabase存放了我国财政部的CAS分类标准和深交所的FR分类标准两套元数据。通过开发的数据转换程序模块,自动生成XBRL分类标准本体;利用Protégé内置的Pellet推理机进行本体的一致性检测,保证本地本体的正确性和有效性。

最后对CAS和FR两套分类标准财务元数据的参考链接关系设计相应的推理规则,并通过Jena推理机进行检测。结果表明,推理规则是有效的。

主要参考文献:

Ding Pan, Yingmin Zhang, Dong Wang. On formalization and reasoning algorithm in distributed XBRL system [J]. Information Technology, 2013(12).

A Borgida, L. Serafini. Distributed description logics: assimilating information from peer sources [J]. Journal of Data Semantics, 2003(1).

李吉梅,杜美杰.基于XBRL的异构财务信息集成算法研究[J].吉林大学学报(工学版),2012(9).

吴忠生,张天西,陈志德等.基于领域本体的XBRL财务报告转换研究[J].计算机应用研究,2013(12).

Franz Baader, Diego Calvanese, Deborah L. McGuinness, et al.. The description logic handbook: Theory, implementation, and applications[M]. London: Cambridge, 2003.

朱小栋,黄志球,沈国华.基于描述逻辑的数据挖掘元数据的一致性检验[J].小型微型计算机系统,2009(2).

高志强,潘越,马力等.语义Web原理及应用[M].北京:机械工业出版社,2009.

O'Riain S., Curry E., Harth A.. XBRL and open data for global financial ecosystems: a linked data approach[J]. International Journal of Accounting Information Systems, 2012(2).

作者单位:1.广州民航职业技术学院,广州510403; 2.暨南大学管理学院,广州510623