

一种可视化电力领域本体管理软件^①

杨永超 房立芳 (中国科学技术大学 自动化系 安徽 合肥 230027)

摘要: 在分析现有电力系统公用信息模型(Common Information Model, CIM)不足的基础上,设计了一种能将 CIM 描述语言由 UML 自动转换为基于描述逻辑的网络本体语言(Web Ontology Language based on Description Logics, OWL DL)的可视化本体编辑工具。实验结果表明:通过该工具可以完成对 CIM 由 UML 到 OWL DL 的语言转换,使得转换后的电力本体具有良好的计算机可推理性,并可以直观的对转换后的本体进行扩展。

关键词: 公用信息模型; 网络本体语言; 可视化; 转换算法; 电力领域本体

Visual Ontology Manager of Power Domain

YANG Yong-Chao, XIE Xing-Sheng, FANG Li-Fang

(Department of Automation, University of Science and Technology of China, Hefei 230027, China)

Abstract: Based on the analysis of the shortage of current CIM(Common Information Model)in power system, this paper designed a visual ontology editor of power domain. It could translate the description language of CIM from UML to OWL DL(Web Ontology Language based on Description Logics). Experimental results indicate that the editor can complete the CIM language translation from UML to OWL DL and the translated- ontology can be well inferred by machine and be extended intuitively.

Keywords: CIM; OWL; visual; translation algorithm; power domain ontology

1 引言

自本体在人工智能界提出以来,就成为国内外科研人员研究的热点,各个专业学科都在研究本领域的本体,以期获得本领域的共享概念模型,达到知识共享和知识复用的目标^[1,2]。目前许多研究领域都建立了自己的标准本体。电力系统是关系国计民生的重要领域,建立一个电力领域的本体就显得尤为重要。现有电力领域的本体 CIM^[3],其描述语言表达能力太弱并且不满足逻辑可推理性。而在实际工程应用中,随时有可能对电力概念类进行扩展,比如对类进行属性和关系的约束形成新类,现有的 CIM 描述语言由于不支持计算机推理,有可能会忽略这些扩展。因此需要使用一种描述能力更强并且支持有效推理的本体语言可

以对电力本体知识充分描述,以实现系统对扩展后 CIM 类的识别。OWL 结合一阶谓词逻辑的可判定子集——描述逻辑(Description Logics, DL)极大的丰富了本体语义,并为计算机推理本体创造了条件。从领域本体的目标——最终要求实现对领域知识的推理出发^[2],需要对现有 CIM 使用 OWL DL 语言进行修改重建。但是建设、修改领域本体是一项庞大的系统工程,需要经历一个复杂的反复过程^[4]。目前,本体的构建过程,大多仍由人工方式完成,本体编辑工具在一定程度上减少了本体重建的难度。因此设计一个界面简洁、操作直观的本体编辑工具将为本体的构建节约大量的时间。

本文从直观的编辑角度出发,设计了基于 CIM 模

^① 收稿时间:2010-03-25;收到修改稿时间:2010-04-27

型的电力领域可视化本体管理软件(Visual Ontology Manager of Power Domain, VOMPD), 它针对 CIM 模型特点, 可以自动完成其描述语言由 UML 到 OWL DL 的转换, 获得一个计算机可推理的电力领域本体, 并设计了直观的视图本体编辑界面, 由此领域专家可以对转换后的 CIM 进行扩展, 最终建立一个完备的电力领域本体。

2 VOMPD功能需求分析

VOMPD 的功能包括自动语言转换功能、可视化本体编辑功能和本体编辑的一般功能。图 1 简要描述了 VOMPD 工作时数据流处理过程。

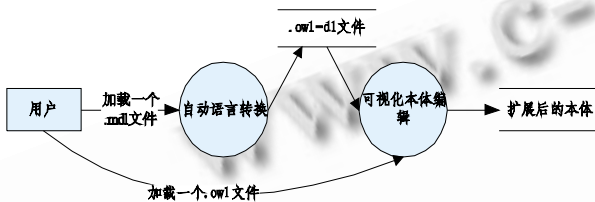


图 1 VOMPD 的数据流图

2.1 自动语言转换功能分析

国际电工技术委员会(IEC)制定了 IEC61970 标准, 其中作为描述电力领域中各元器件以及它们之间关系的基本模型——CIM 在一定程度上可以看作是电力领域本体^[3]。目前国内外已正式发表的 CIM, 只有基于 UML 的文档及能被应用程序正确读取的基于资源描述框架(Resource Description Framework, RDF)的 CIM/XML 型文档^[5]。其中 UML 只是一种建模语言, 不能直接用于实际工程。而 RDF 作为一种描述领域模型的知识表示语言, 在表达能力和语义表达精确程度上都存在很大缺陷。OWL DL 则可以在保证推理完备性和可判定性的前提下提供最强的表达能力^[6]。使用 OWL DL 描述 CIM 可以极大的丰富电力本体的语义。

VOMPD 集成了 Xpetal^[7]和 CIMTool^[8]的功能, 可以自动完成 CIM 文档由 UML 型到 OWL DL 型的转换。其中 Xpetal 可以实现 CIM 描述语言从 UML(仅限于 Rational Rose 编辑的 .mdl 扩展名文件)到 CIM/XML 的转换; CIMTool 是可以将 CIM 由 CIM/XML 型转换为 OWL Full 型的工具。

2.2 可视化的本体编辑功能分析

到目前为止, 世界上已经开发出超过 90 种的本体编辑工具^[4]。大多数编辑工具虽然功能丰富, 但要求用户有比较专业的计算机知识。一般而言, 专注于某一领域的专家并非一定是计算机专业人员, 因此设计一个界面简洁、操作直观的本体编辑工具将为本体的创建节约大量的时间。

VOMPD 具有可视化本体编辑及推理机推理功能, 丰富并检验转换后的本体。对本体的处理主要使用 Jena^[9], 它是一个由 HP 实验室开发的表示和处理半结构化数据的 Java 开源工具, 它将 RDF 看作 <subject, predicate, object> 三元组的集合, 提供了一系列的 API 函数对 RDF 进行管理、查询; 而推理功能的实现主要通过集成 Pellet 推理机^[10], 它是基于 Tableau 算法的描述逻辑推理机。具体实现如图 2 所示。

2.3 其他功能分析

除上述功能外, VOMPD 还具有基本的本体编辑功能, 例如载入本体、使用 OWL DL 语言非持久化存储本体、列表显示本体以及该本体 import 的本体、增减本体元素等。

3 VOMPD主要功能模块设计

如图 2 所示, 根据 VOMPD 所能处理的文档类型, 整个系统可以看作由两个功能模块组成: UML 文档处理模块和 OWL 文档处理模块。前者研究如何实现 CIM 描述语言由 UML 向 OWL DL 的自动转换; 后者则研究如何实现本体的编辑和推理。

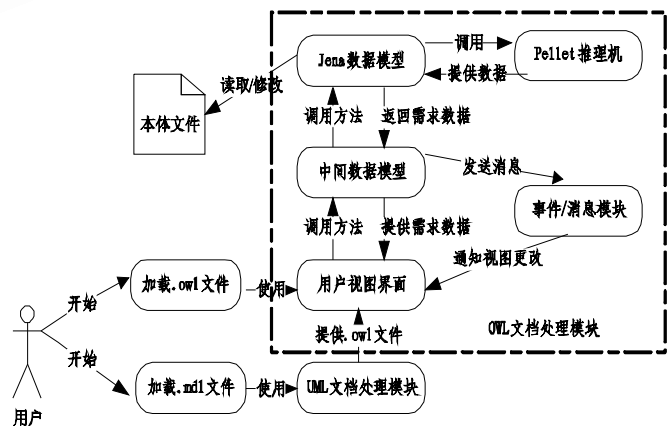


图 2 VOEPD 系统设计模块图

3.1 UML 文档处理模块

模块的界面设计如图 3 所示。该模块中,由 UML 型文档到 OWL Full 型文档的转换,可以直接通过 Java 语言调用可执行程序 Xpetal 和 CIMTool 实现。因此研究的重点是 CIM 从 OWL Full 型到 OWL DL 型的转换算法设计。该转换算法使用 Jena 提供的 RDF API 函数设计。根据 OWL DL 的语法约束,通过以下两个函数来对 OWL Full 形式的 CIM 增减语法规则,使其满足计算机推理计算。

算法 1. 删除不必要的三元组

```
function Del_Unwanted_Elements (Model
CIMmodel):Model
//基本数据结构
List [1..n] wanted_subject_list;//需要保留的三元组的 subject 列表
```

Step 1. 三元组的 object 中含标识串 Package 的, 其对应 subject 添加至列表 wanted_subject_list, 然后遍历整个 CIM 文档, 将 subject \in wanted_subject_list 的所有三元组保留。

Step 2. 重新遍历文档, 删除 subject 或 object 中包含标识串 UML 或 Package 的三元组。

end function

算法 1 说明: 在 CIM/XML 文档中, 三元组的 object 中含 Package, 则 subject 可能是电力系统有关的概念术语, 而 subject 非电力系统术语的三元组语句可以删去。另外, 以命名空间 xmlns:j.1=http://langdale.com.au/2005/UML# 为前缀的是 UML 模型遗留组件描述, 也可以直接删除。再者, 三元组的 subject 或 object 中含 Package 字串的, 都是用来描述 Package 隶属关系或划分的语句, OWL DL 无法处理它们。在已经完成电力系统概念术语识别之后, 要删除所有直接包含 Package 的语句。

算法 2. 分离数据属性和对象属性, 并修改自定义的数据类型

```
function Norm_Property_Defs (Model
CIMmodel):Model
//基本数据结构
//CIMmodel 中 object 是数据属性的 subject 列表
DatatypeProps: List[1..n] of subject;
//CIMmodel 中 object 是对象属性的 subject 列表
```

ObjectProps: List[1..n] of subject;

//自定义数据类型列表

datatypeObjects: List[1..n] of resource;

Step 1. 遍历 cim.owl 的所有三元组, 如果某个三元组的 object 是 "rdfs#Property", 并且 subject 不属于 ObjectProps 或 DatatypeProps, 简单将 object 替换为 "owl#ObjectProperty", 否则删除此三元组。

Step 2. 如果三元组的 subject 是数据属性, 并且值域这个 predicate 对应的 object 不是 XML Schema 定义的简单类型(我们称之为自定义型), 那么将 object 保存至列表 datatypeObjects。

Step 3. 遍历所有三元组, 若模型中有该自定义型的等价数据类型(通过分析 OWL Full 文档我们知道, 该等价类型必为 XML Schema 的简单类型)则用等价的简单类型替换原数据属性的值域。

Step 4. 删除所有以自定义数据类型为 subject 的三元组。

end function

算法 2 说明: 在基于 OWL Full 描述的 CIM 中, 有三种描述属性的构词: rdfs#Property, owl#DatatypeProperty 和 owl#ObjectProperty。如果一个属性既被定义为 rdfs#Property 又被定义为 owl#ObjectProperty (或 owl#DatatypeProperty), 则为避免冗余删除前者定义; 如果一个属性只被定义为 rdfs#Property, 则简单将其替换为 owl#ObjectProperty。另外, 并非所有的 XML Schema 自定义数据类型都可以在 OWL 中使用[6], 但在 CIM 的 Domain 包中却定义了两类自定义数据类型: 基本型和枚举型。在转换中, 需要将基本型中所包含的自定义数据类型用等价的 XML Schema 简单类型替换, 并把枚举型的类处理为 OWL 的类。

通过以上的转换, 虽然不能完备的将模型转换为 OWL DL 型, 但是可以结合 Altova SemanticWorks 语法检验工具, 根据语法错误提示, 对转换后的模型进行手动修改。以 CIM 的版本 CIM10r7.mdl 为例, 要修改两处即可完成到 cim_dl.owl 的转换: 一是 Core2 包中的概念 Route, 既被定义为类又被定义为对象属性; 二是发现了无等价简单类型的自定义数据类型 Complex(复数型)。通过手工修改, 先删除 Route 的属性定义并增加一个对象属性 Routy, 然后将

Complex 直接修改为简单数据类型 float, 并删除有关 Complex 定义的语句。

3.2 OWL 文档处理模块

模块的界面设计如图 4 所示。该模块包含了基本的本体列表编辑、查看、连接推理机等功能, 其重点是可视化设计本体编辑界面。通过以上转换算法得到的 OWL DL 文档, 是最基本的电力领域本体。用户可以根据自己的需求对本体增加类或属性等作进一步扩展。VOEPD 构造了一个基于概念层次树结构的图形化编辑模块, 用户可以通过增减图形节点来编辑概念, 形象直观的编辑方式可以更加快捷有效的扩展本体。其中图形显示算法如算法 3 所示。另外, 通过对概念图形类增加事件响应, 可以实现本体概念的增减。

算法 3. 图形显示算法

function repaint()

//基本数据结构

Classtuxing classtuxing//概念图形。

Suboftuxing suboftuxing//上下位连接图形。

Step 1. 读取概念图形集合中的表示 OWLthing 的图形。按照图形参数在界面的左边中心位置绘制一个圆。

Step 2. 获得 OWLthing 的直接子节点(假设为 n 个)对应图形集合中的对象集合 nodes。设置迭代绘制方法, 变更绘制对象的圆心。

Step 3. 绘制父子关系之间的连线。

Step 4. 如果 nodes 中每个对象对应的概念还有子概念, 返回 Step2, 修改 OWLthing 参数为当前选择概念。

end function

算法 3 说明: Classtuxing 和 Suboftuxing 是用来图形化显示概念以及它们之间关系的类, 分别用圆形节点和直线表示。该算法主要是将所有的概念从 OWL thing 开始搜索, 根据概念所在层次和该层次中位置不断生成其对应的概念图形对象, 实现了通过节点及之间的连线直观的显示概念层次图。

4 实验测试

4.1 UML 文档处理测试

以 CIM10r7.mdl 为例, 当打开此类文件时, 程序将调用 Xpetal 和 CIMTool, 将 CIM 自动由 UML

型转换为 OWL Full 型, 如下图“对应 OWL Full 文档”所示, 当单击按钮“从 Full 到 DL 转换”时, 则文本区显示对应的 OWL DL 文档。然后将经过转换算法得到的 OWL DL 型文档保存, 最后使用 OWL 语法检验工具——Altova Semanticworks[11]进一步检验。经过少量修改可以得到完全满足 OWL DL 语法的文件 cim_dl.owl。转换结果如图 3 所示。

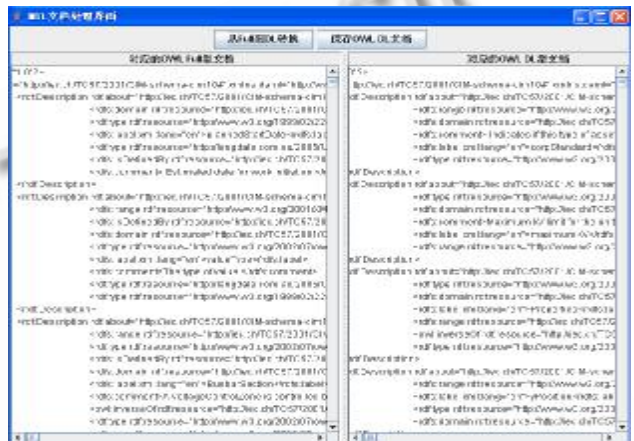


图 3 UML 到 OWL DL 转换测试

4.2 OWL 文档处理测试

我们将得到的文件 cim_dl.owl, 使用 OWL 文档模块进行查看编辑。通过观察其概念层次结构树, 与 Rational Rose 编辑工具中显示的 CIM10r7.mdl 概念类图进行仔细分析对比, 证实了转换算法结果的正确性和概念无损性。

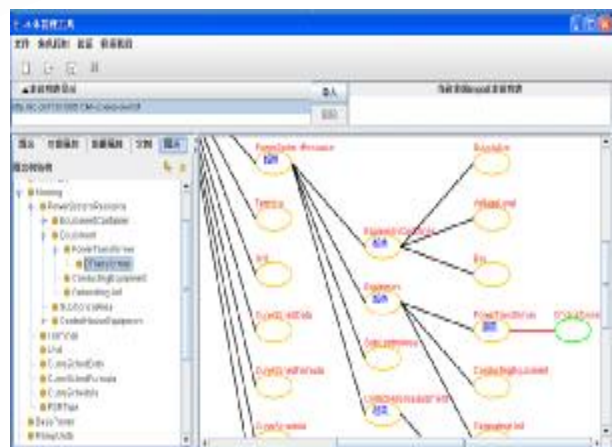


图 4 可视化界面推理测试

最后使用 OWL 文档处理模块对 `cim_dl.owl` 进行扩展, 在 `Wires` 包中变压器类的基础上增加对绕组数目的约束来构建双绕组变压器类: `DTransformer`, 并定义 `DTransformer ≡ PowerTransformer ∩ (=2Contains_Transformer Windings.TransformerWinding)`, 那么根据描述逻辑推理可知 `DTransformer` 是 `PowerTransformer` 的子类, 使用 `Pellet` 推理机推理后图形如图 4。

经 `Pellet` 推理机推理后, `DTransformer` 被识别为 `PowerTransformer` 的子类, 连线颜色改变。

5 结束语

本文在分析当前电力领域本体 `CIM` 不足的基础上, 设计了一种基于 `CIM` 的可视化电力领域本体管理软件(`VOMPD`)。并对 `VOMPD` 的两个功能模块的主要算法进行了描述, 最后分别对软件的主要功能进行了实验测试。结果表明: `VOMPD` 可以正确的分析 `.mdl` 和 `.owl` 两种扩展类型的文件。它可以自动将 `CIM` 的描述语言由 `UML` 转换为 `OWL DL`, 使得转换后的领域本体具有良好的计算机可推理性, 并可以通过视图化编辑界面直观的对 `OWL DL` 语言描述的 `CIM` 进行扩展, 最终电力领域专家可以在此基础上建立一个完备的电力本体。目前, 我们已将研究成果作为构建电力领域本体的工具, 应用到数据集成系统中, 取得了良好的效果^[12]。

参考文献

- 1 邓志鸿,唐世渭,张铭,杨冬青,陈捷. *Ontology 研究综述*. 北京大学学报(自然科学版), 2002,38(5):730 - 737.
- 2 陆建江,张亚非,苗壮,周波. *语义网原理与技术*. 北京: 科学出版社, 2007.
- 3 Scott Neumann,Arnold DeVos,et al. *Use Of the CIM Ontology*. Pacific Northwest National Laboratory.
- 4 Michael Denny. *Ontology Tools Survey, Revisited*[2010-04-22].<http://www.xml.com/pub/a/2004/07/14/onto.html>.
- 5 IEC 61970, *Energy Management System Application Program Interface(EMS-API)Part 301 - 503*.
- 6 Grigoris Antoniou, Frank van Harmelen, *A Semantic Web Primer (second edition)*. The MIT Press, 2008.
- 7 Langdale Consultants. *The Xpetal Tool*[2010-04-22]. <http://www.langdale.com.au/styler/xpetal/>.
- 8 Langdale Consultants. *CIMTool*[2010-04-22]. <http://www.cimtool.org/>.
- 9 Ian Dickinson. *Jena Ontology API[S/OL]*. 2010-04-22. <http://jena.sourceforge.net/ontology/>.
- 10 MindSwap. *Pellet*[2010-04-22]. <http://pellet.owldl.com/>.
- 11 Altova. *Semanticworks*[2010-04-22]. <http://www.altova.com/semanticworks.html>.
- 12 谢兴生,张一鸣,余银等. 一种支持智能匹配检索的数据集成系统. *模式识别与智能系统*, 2009,22(1):40 - 46.