

·技术与方法·

一种基于本体集成的生物医学数据库中语义映射维护方法研究

肖清滔¹ 姚莉^{2△}

(1 第三军医大学 重庆 400038 2 国防科技大学信息系统与管理学院 湖南 长沙 410008)

摘要 生物医学数据库到生物医学本体的语义映射是基于本体集成生物医学数据库系统的一个重要环节。而生物医学本体随着学科的发展不断演化,造成了集成系统不稳定。针对这个问题,本文在本体演化条件下,分析并发现了语义映射的变化规律,设计了对应的维护流程和维护方法,并通过计算维护收益率证明了该方法对映射的维护是有效性的,从而增强了集成系统在本体演化条件下的稳定性。

关键词 本体;生物医学数据库;数据库集成;语义映射

中图分类号 Q819 **文献标识码** A **文章编号** 1673-6273(2011)14-2770-05

A Maintenance Method of Semantic Mapping of Biomedical Database Integration Based on Ontology

XIAO Qing-tao¹, YAO Li^{2△}

(1 Third Military Medical University, Chongqing 400038, China;

2 College of Information System and Management, National University of Defence Technology, Changsha 410008, China)

ABSTRACT: Semantic mapping from biomedical databases to biomedical ontology is a significant component of biomedical database integration system based on biomedical ontology. Biomedical ontology is evolving with development of the discipline, which is causing system instability. To address this issue, this paper analysed and found the changes of semantic mappings in the law under conditions in the biomedical ontology evolution, designed the maintenance process and the corresponding maintenance methods, and proved effectiveness of the method by calculating the benefit of the maintenance.

Key words: Ontology; Biomedical Database; Database Integration; Semantic Mapping

Classification(CLC): Q819 **Document code:** A

Article ID: 1673-6273(2011)14-2770-05

前言

生物医学的发展,伴随着互联网的广泛应用,尤其是基因研究的不断深入,生物医学数据急剧爆炸,呈现海量规模且数据量还在加速增长。截止2010年1月,权威的“核算研究在线分子生物学数据库集合(NARMBDC)”共收录了1230个分子生物数据库^[1]。然而这些数据库存放在不同的服务器上,使用不同的存储方式且相互隔离,形成了生物医学中的“信息孤岛”,实现数据共享和提高数据利用率成为生物医学的一个挑战。使用计算机和程序,将互相独立的数据库有机地组织在一起,对于提高学科领域的认知,以及对实验和临床诊断预测的指导,将起到巨大的意义。

生物医学数据库集成中的一个重要的方法是利用本体进行数据库语义集成,将不同的生物数据库按照语义关系映射到本体上,使用查询转换器将本体格式的用户的查询请求转换为底层数据库的查询操作,并将查询结果处理后返回给用户,在语义层面上实现数据共享。目前已出现多个基于本体的生物

数据库集成系统。K(o)hler等人利用本体,建立了一个生物学数据库语义集成系统SEMEDA^[2]。该系统利用本体的继承关系,实现了“is-a”的层次结构。SEMEDA系统将数据库中的表和属性映射到一个全局本体上,本体向外提供统一的接口,提供语义查询。Alonso-Calvo等人基于本体和主体(agent)建立了基因、蛋白质和疾病相关的异构数据库集成系统OntoFusion^[3]。该系统包含用户界面、词表服务、中介器和数据访问等4个模块,通过本体实现对各异构数据库的一致性访问。在该系统中,包含了三种映射:数据库模式间映射、数据库模式和本体间映射、本体间映射。Noy等人在本体仓库(REPOSITORY)的基础上建立了生物医学数据集成系统BioPortal系统^[4],利用134个本体集成了20个不同领域内著名的异构生物医学数据源。2009年,Min等人建立了前列腺癌本体(Prostate Cancer Ontology, PCO)并以此为基础开发了前列腺癌数据集成系统PCIS,为美国著名的福克斯·蔡斯癌症中心(Fox Chase Cancer Center, FC-CC),成功地集成了与两个与前列腺癌相关的数据库系统^[5]。PCIS是基于本体的数据集成方法在临床医学领域内的成功应用之一,为诊断提供了有效的技术保障。虽然基于本体的生物数据库集成系统正在被广泛应用于生物医学和生物信息学方面,但生物医学在不断发展,对领域的认知也在不断修正,这就造成了领域本体演化的持续性。而目前基于本体的生物数据库集成系统都未考虑到本体演化带来的映射维护问题。Min等人在对PCIS系统的评价中,也承认系统没有充分考虑到演化问

作者简介 肖清滔(1984-),男,硕士,主要研究方向:生物信息学,生物信息计算。电话:023-68772164 E-mail: xqt@tmmu.edu.cn

△通讯作者 姚莉(1963-),女,博士生导师,教授,主要研究方向:人工智能

(收稿日期:2010-11-28 接受日期:2011-01-15)

题^[8]。本文就是针对这个问题,在本体演化条件下,分析并发现语义映射的变化规律,通过失效检测和一致性修正维护,处理本体演化条件造成的语义映射失效问题。

1 数据库模式和本体之间的语义映射

映射是一种特殊的元素转换机制^[9],在基于本体的数据库集成系统中,主要有三种生成方式^[10]:①从本体映射到数据库模式上;②将数据库模式和本体分别映射到某一中间件上;③从数据库模式映射到已有本体上。但在集成系统中,数据库模式生成的本体不能完全符合应用的需求^[8],因此要求集成系统中的本体语义应当大于被集成数据库的语义。目前的研究主要是采用第三种方法,这种方法能够最大限度地保留数据库模式中的语义,同时也能满足应用的需求。

语义映射是集成系统中介于数据库和本体之间的中间件,需要处理本体层和数据库之间的数据交换操作,如语义查询的转换和查询结果的处理。为更好的讨论本体演化条件下的语义映射维护问题,首先给出语义映射(本文的语义映射来自于前期工作 MVDSR 项目)的形式化表述以及其中隐含的概念定义。

定义 1 数据库模式可以用一个二元组表示: $\mathcal{T}=\langle R,U \rangle$, 其中:

SA 表述数据库模式;
 $U=(pk_1,pk_2,\cdots,pk_i,k_1,k_2,\cdots,k_j) \ i \in \mathbb{N}^0, j \in \mathbb{N}^0$
其中 $pk_1,pk_2,\cdots,pk_i$ 表示模式 R 中的主键;
 $k_1,k_2,\cdots,k_j$ 表示模式 R 中的非主键。

定义 2 数据库模式 $\mathcal{T}$ 中字段和本体 $\mathcal{O}$ 中元素的语义对应 $\mathcal{S}$ 可以分成两类(使用符号 $\mathcal{T}$ 和 $\mathcal{O}$ 区分数据库模式和本体):

$\mathcal{S}_1: \mathcal{T}: SA: pk_i \rightsquigarrow dp: c$: 表示模式 $\mathcal{T}$ 中主键 pk_i 与本体 $\mathcal{O}$ 中属性 dp 存在语义等价关系,其中 $dp \in DataProperty$,表示本体 $\mathcal{O}$ 中的数据类型属性;
 $\mathcal{S}_2: \mathcal{T}: SA: k_j \rightsquigarrow dp: c$: 表示模式 $\mathcal{T}$ 中非主键 k_j 与本体 $\mathcal{O}$ 中属性 dp 存在语义等价关系,其中 $dp \in DataProperty$,表示本体 $\mathcal{O}$ 中的数据类型属性。

定义 3 设数据库模式 $\mathcal{T}$ 和 OWL DL 本体 $\mathcal{O}$ 之间的语义对应关系集为 $\mathcal{S}$, 那么使用如下公式形式化定义 $\mathcal{T}$ 和 $\mathcal{O}$ 之间的一个语义映射:

$$\mathcal{T}(X) \rightarrow \phi(X,Y)$$

其中: $\mathcal{T}(X)$ 表示数据库中的一个模式, X 表示模式 $\mathcal{T}$ 中的字段,即 $X \in PK \cup K$;

$\phi(X,Y)$ 是一个谓词合取公式,这里的谓词表示本体 $\mathcal{O}$ 中的概念 $\mathcal{C}$ 、关系 $\mathcal{R}$ (或 ObjectProperty) 和属性 $\mathcal{A}$ (或 DataProperty),用于表示模式 $\mathcal{T}$ 中的语义, Y 是合取公式中的变量。

与通常的表示方法类似,公式中蕴涵符号 (" $\rightarrow$ ") 左部的变量是全域量化;未出现在左部仅出现在右部的变量 Y 是存在量化。这种描述方法在一些研究模式到描述逻辑表示的研究文献(如文献^[11,12])中已经使用。 $\mathcal{T}(X) \rightarrow \phi(X,Y)$ 公式也用于表示一个模式可能的语义。

定义 4 数据库模式 $\mathcal{T}$ 和 OWL DL 本体 $\mathcal{O}$ 之间的语义对应关系集为 $\mathcal{S}$, 那么语义映射表示为一个三元组:

$$SM=\langle M, \mathcal{T}, \mathcal{O} \rangle$$

其中 M 可定义为两个集合,其中一个集合是语义对应关系集,包含多个语义对应关系,另一个集合是模式 $\mathcal{T}$ 到本体 $\mathcal{O}$ 的一组一阶逻辑公式 $\mathcal{O}(X) \rightarrow \phi(X,Y)$ 。

2 本体演化下的语义映射维护过程

本体作为语义 Web 中的一个重要组成部分,它通过定义精确的共享术语,以提供某一特定领域可重用的知识。在生物医学领域,随时间的推移,领域知识在不断变化,而本体演化就是这一过程的最终反映,这也是生物医学本体在初始建立时不完备的因素之一。语义映射作为基于本体的数据库集成系统中的核心部件之一,需要根据生物医学本体的不同变化,做出相应的变化以抵消本体演化对系统带来的影响。

2.1 本体演化对语义映射的影响

本体演化不是本文研究重点,但在语义映射的形式化 $\mathcal{T}(X) \rightarrow \phi(X,Y)$ 中, $\phi(X,Y)$ 使用本体中的概念、关系和属性来表示模式 $\mathcal{T}$ 中的语义信息。因此重点关注本体演化中的概念、关系和属性的变化。

为了具体分析本体变化对语义映射产生的影响,本文将从 "变化类型轴" 和 "变化元素轴" 两个维度来分析本体不同变化对语义映射产生的影响。其中 "变化类型轴" 包含 "增加、更新、删除。" 变化元素轴 " 包括 "概念、关系和属性。本体中概念、关系和属性有其特有的性质,如关系和属性会随着概念的继承而向下传递。根据这些元素在本体中的作用和关系,可以得到本体变化条件下的语义映射影响表,见表 1。

表 1 本体变化下的语义映射影响表
Tab 1 Semantic mapping influence table under ontology changing

			增加 Increment	更新 Update	删除 Deletion
概念 Notion	无继承关系 No succession relations	参与对应	不失效 Non-lapse	用户确认 User affirm	失效、无法维护 Lapse, can't attendance
		不参与对应	不失效	用户确认	自动维护 / 阻抗失配
	有继承关系 Succession relations	参与对应	不失效	用户确认	自动维护 / 阻抗失配
		不参与对应	不失效	用户确认	自动维护
关系 Relationship	参与对应	参与对应	不失效	用户确认	失效、无法维护
		不参与对应	不失效	用户确认	自动维护 / 阻抗失配
	不参与对应	参与对应	不失效	用户确认	自动维护 / 阻抗失配
		不参与对应	不失效	用户确认	自动维护

表中“不失效”表示本体变化不会对映射产生影响,语义映射依然能够被用于集成系统中;“用户确认”表示结果在语法上符合应用,但需要用户对语义进行确认;“自动维护/阻抗失配”表示映射的自动维护过程中可能产生两种情况:一是利用本体的特性自动对映射进行维护修改,得到满足要求的结果;另一种是映射失效且无法进行维护,这种结果称之为阻抗失配;“无法维护”表示出现了阻抗失配情况,这时只能由用户重新确定语义对应关系集,完全产生新的语义映射。

2.2 本体演化下的语义映射维护过程

对语义映射维护的本质是对数据库模式和本体之间语义关联的维护,这种维护可以有效提高集成系统的稳定性。当本体发生演化后,需要及时对系统内部进行适应性修改,以保障

系统的稳定性,为此系统必须能够及时发现并捕获引发系统中映射变动的本体变化。其次,需要分析捕获到的本体变化,根据变化类型对语义映射采取不同的维护措施。维护结果的可用性验证是维护工作中最重要的一环,也是保证维护组件有效运转的一个关键环节。在数据库集成系统中,语义映射主要用于连接异构数据库和本体,对外提供统一的访问接口,转换查询请求和查询结果,因此语义映射维护后需要在系统中进行试运行,确认是否满足需求,确保数据库和本体之间的语义关联与映射在语义上保持一致。参考本体演化过程^[13],将语义映射的基本维护过程归纳为:“捕获-分析-修改-验证-传播”,如图1所示。

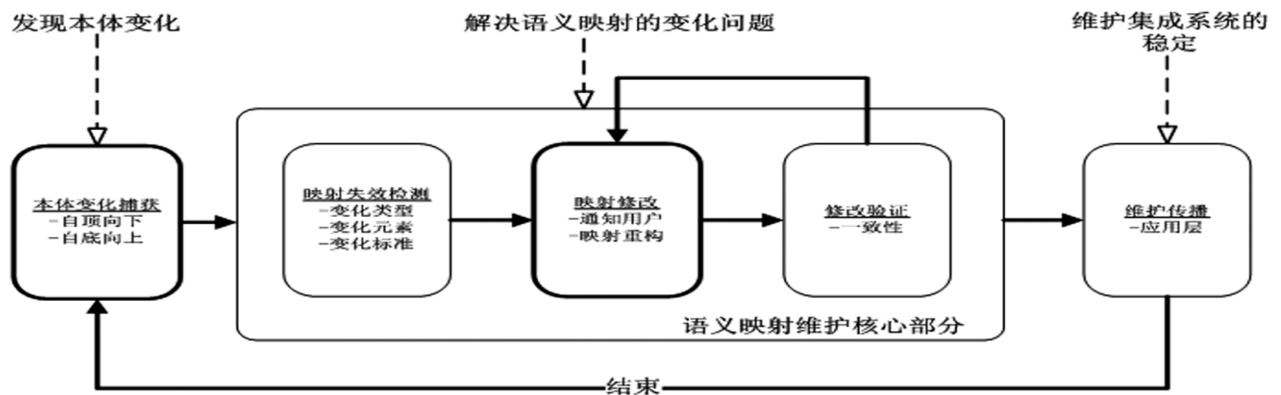


图1 本体演化条件下的语义映射维护基本过程

Figure 1 Semantic mapping attendance unit process under ontology changing

本体变化捕获:本体演化是一个过程,本体变化的捕获需要从整体上考察演化前后的本体在结构和语义上的变化,而不是局限于某一个时间片段,以便根据变化的本体元素集发现失效的映射。

映射失效检测:需要根据捕获到的本体变化,进而得到变化元素集合和变化类型集合,之后与语义映射进行交叉比较,根据表1所列出的可能性,进行语义映射失效检测。

映射修改:映射的修改遵循最小化修改的原则,尽可能地重用映射,或是映射的某一部分,减少计算量,降低计算复杂度。

修改验证:维护后的语义映射需要满足两个基本条件:一是一致性。即维护得到的语义映射需要在语义和结构上分别与本体和 $T(X) \rightarrow \phi(X,Y)$ 保持一致;二是可用性。维护后的语义映射至少能够满足应用需求。

维护传播:在集成系统中,语义映射往往还与系统的部分接口关联,因此当语义映射发生变化时,就需要对相应的外部接口进行修改。但接口的维护问题不是比本文研究内容,在此不做讨论。

3 映射维护

语义映射维护的前提是得到本体演化中的变化集合,在生物医学领域,已经有许多著名且公开发行的本体,如基因组学领域的基因本体(Gene Ontology)、氨基酸本体、疾病本体(Disease Ontology)等等,这些本体随着领域内知识的变化不断演

化,通过日志等多种形式记录其演化过程。本文主要关注在已知演化过程后,对语义映射的维护。

3.1 映射失效检测机制

数据库和本体间的语义映射作为语义集成中间环节,使用本体解决数据库间的异构性。在动态环境中,整个应用系统的语义是随着本体变化而不断变化,当本体发生变化后,需要验证语义映射是否还能够继续语义连接生物医学本体和各数据库。

根据定义3和定义4的定义,语义映射SM包含了语义对应关系集 S 和模式 T 到本体 O 的一组逻辑公式 $T(X) \rightarrow \phi(X,Y)$,因此当本体 O 变为 O' ,需要验证 $SM = \langle M, T, O' \rangle$ 是否成立。在定义3中,谓词公式 $\phi(X,Y)$ 是由本体 O 中的概念 C 、关系 R 和属性 A 组成,即 $\phi(X,Y) = \bigwedge A(Y,X) \wedge \bigwedge R(Y_i) \wedge \bigwedge R(Y_m, Y_n) \mid X \in PK \cup K$,那么使用语义相容性来进行语义映射失效检测。

定义5 若本体 O 演化为 O' 后,语义映射SM在本体 O' 中依然成立,那么称语义映射SM与本体 O' 语义相容。

语义映射SM与本体语义相容,表明了演化后本体依然能够通过映射SM描述数据库中的语义。而这种语义相容性不仅需要验证SM中语义对应关系集 S 是否有效,同时也需要验证本体 O' 是否包含谓词公式 $\phi(X,Y)$ 中所使用的概念 C 、属性 A 以及概念之间关系 R 。若语义映射在本体演化后不失效,当且仅当上述两个条件都成立。3.2中将详细阐述不同条件下语义映射的维护。

3.2 语义映射维护

语义映射维护的一个原则是映射的重用性,即对语义映射进行最小化修改。根据语义映射检测的原理以及映射的定义,可从映射的组成方式上将映射的维护分为两个步骤:语义对应维护(Semantic-Correspondence Maintenance, SCM)和关系维护(Relationship Maintenance, RM)。

(1) 语义对应维护 (Semantic-Correspondence Maintenance): 映射中的语义对应如定义 2,即模式 T 中某一字段的语义可描述为概念 c 有属性 dp 。若映射失效是由于语义对应引起的,则存在如下三种情况:

① $dp \notin O' \wedge c \notin O'$: 此时语义对应关系已不存在,即演化后的本体不能充分描述当前模式的语义。此时不论对映射进行如何操作,都无法完成映射的修复,只有在人工参与的条件,重新确定新的语义对应关系。

② $dp \notin O' \wedge c \in O'$: 演化后的本体依然包含模式的概念,但模式属性语义却发生了改变,即原有对应属性 dp 发生了变化,此时可以利用函数 $\text{Domain}^{-1}(c)$ 和模式属性的数据类型,得到属性集合 Δdp 。使用 Δdp 替换失效的 dp ,这其中可能出现非目的映射,这就需要用户从 SM' 中选择选择最恰当的结果。

③ $dp \in O' \wedge c \in O'$: 由于本体中元素的依赖关系,利用函数 $\text{Domain}^{-1}(dp)$ 可以得到一个概念集 Δc 。然而属性会随着概念的变化而表示不同的语义,如属性 ID 在概念 Gene 和概念 Protein 中就表示了两种不同的语义,因此需要将 Δc 引入关系维护中,经过处理后得到 Δc 的一个真子集或空集。

(2) 关系维护 (Relationship Maintenance): 上一步操作中,已对概念和属性进行了相关维护操作,而关系维护的实质是验证逻辑公式 $\mathcal{R}(Y_m, Y_n)$ 是否对变化后的概念集 c' 成立。此时需要分两种情况进行讨论:

① 若 $\exists r \in \mathcal{R}, r \in \mathcal{R}'$ ($\mathcal{R}'$ 是本体 O' 的关系集),即本体演化导致关系 r 失效,则对于映射 SM 中有关系 $r(y_i, y_j)$ 的概念 c_i, c_j 进行如下操作: 设集合 $\mathcal{R}^{\Delta} = (\mathcal{R}r)$, 在 Δ 中寻找概念 c_i 和 c_j 之间的一个关系 $r\Delta$:

I 若存在这样的 $r\Delta$,则在映射 SM 中使用关系 $r\Delta$ 替换关系 r ;

II 若集合 $\mathcal{R}^{\Delta}$ 中不存在这样的关系 $r\Delta$,遵照概念互斥原则则将映射 SM 拆分。

定义 6 (概念互斥) 本体中概念 c_i 和 c_j ,若本体中存在关系 r ,使得 $c_i \in \text{Domain}(r)$ 和 $c_j \in \text{Range}(r)$ 同时成立,那么称概念 c_i 和 c_j 是互斥的。

② 若 $\mathcal{R} \subset \mathcal{R}'$, 设 $\mathcal{C}^{\Delta} = \text{Domain}(\mathcal{R}) \cup \text{Range}(\mathcal{R})$, 判断 $\mathcal{C}^{\Delta}$ 与 $\mathcal{C}$ 的关系,确定对关系 $\mathcal{R}$ 的操作:

I 若 $\mathcal{C}^{\Delta} \cap \mathcal{C} = \emptyset$,即本体演化导致概念 $\mathcal{C}$ 之间的原有关系 $\mathcal{R}$ 完全失效,此时映射 SM 应当利用语义对应重新寻找概念 $\mathcal{C}$ 之间的关系;

II 若 $\mathcal{C}^{\Delta} \cap \mathcal{C} = \nabla \mathcal{C} \subset \mathcal{C}$,则对映射 SM 按照概念互斥原则进行拆分,得到多个映射 SM' ;

III 若 $\mathcal{C}^{\Delta} \cap \mathcal{C} = \mathcal{C}$,则保留关系 $\mathcal{R}$ 。

分析: 在成熟的生物本体中,本体的演化主要集中在局部变化上,变化的幅度和范围较小,出现情况 I 的几率趋于 0,情况 II 则表明演化后的本体改变了映射中部分概念之间的关系,

即出现新的知识推翻了原有的概念间的关系,而由于映射是数据库模式语义到本体的满射,即一个映射能够完整地使用本体片段来描述数据库模式的语义,因此当出现 $\mathcal{C}^{\Delta} \cap \mathcal{C} = \nabla \mathcal{C} \subset \mathcal{C}$ 时,就需要根据 $\nabla \mathcal{C}$ 和 $(\mathcal{C} - \nabla \mathcal{C})$ 两个集合对映射进行拆分,以保证映射是语义上的满射;情况 III 说明本体演化并没有改变数据库模式对应的概念间的关系,因此不用对映射中的关系 $\mathcal{R}$ 进行修改。特别地对于 $\Delta \mathcal{C}$,可以利用 $(\Delta \mathcal{C} \cap \mathcal{C})$ 来确定当出现 $dp \in O' \wedge c \in O'$ 时,映射的可维护性以及维护过程。

4 结果评比

目前在映射维护的研究中,对于维护效果的评价主要是计算收益性(Benefits)。文献^[14-16]将收益性定义为: 对映射使用维护相比于对映射进行重新生成所节省的成本,并且都采用 blank-sheet 方法计算维护结果的收益率(benefit)。Benefit 反映了映射维护的价值,即 benefit 越大,维护的价值越高。

(1) 全局 benefit 计算

若从生物医学数据库为出发点,整体考察生物医学本体和数据库之间的映射的可维护性,则 benefit 的计算应采用下式:

$$\text{benefit} = 1 - \frac{\#N_{\text{invalid-mappings}}}{\#N_{\text{mappings}}}$$

从上式中可以看出,当某一生物医学数据库与本体之间的失效语义映射数量 $N_{\text{invalid-mappings}}$ 增大时,映射的可维护性会变得很低,即当 $N_{\text{invalid-mappings}} \Rightarrow N_{\text{mappings}}$ 时, $\text{benefit} \Rightarrow 0$,此时应当重新生成该数据库和本体的语义映射。

(2) 局部 benefit 计算

若失效映射数量较少,那么以映射为考察点,计算维护的成本。设计算执行单步的时间是 Δt ,本体中概念、属性和关系的数目分别为 M_1, M_2, M_3 ,语义映射 SM 中语义对应、涉及概念和关系数目分别为 n_1, n_2, n_3 ,那么 benefit 计算如下:

$$\text{benefit} = 1 - \frac{T_{\text{maincane}}}{T_{\text{re-generation}}} \quad \text{其中:}$$

$$T_{\text{maincane}} = T_{\text{SCM}} + T_{\text{RM}} = (\Delta t \cdot n_1 + \Delta t \cdot n_2) + (\Delta t \cdot n_3 + \Delta t \cdot n_3 + \Delta t \cdot n_2) = \Delta t \cdot (n_1 + 2n_2 + 2n_3)$$

$$T_{\text{re-generation}} = \Delta t \cdot n_1 + \Delta t \cdot M_2 + \Delta t \cdot M_3 = \Delta t \cdot (n_1 + M_2 + M_3)$$

带入 benefit 计算式,得到:

$$\text{benefit} = 1 - \frac{n_1 + 2n_2 + 2n_3}{n_1 + M_2 + M_3}$$

在生物医学本体中,由于 $n_2 \ll M_1$ 且 $n_3 \ll M_3$,因此利用上式计算得到的 benefit 值较高。

从上述的分析中可以看出,在基于生物医学本体集成的医学数据库中,在生物医学本体演化后,应当根据失效的映射数量来决定对映射应当采取维护操作,还是重新生成整个数据库和本体之间的映射。

5 结论及下一步研究方向

基于生物医学本体进行数据库集成实现数据共享是目前生物医学研究中的一个热点,其中建立本体和数据库之间的语义映射是一种重要的集成方法。但生物医学的发展必然引发生物医学本体的演化。而本文分析了本体和数据库间映射在生物医学本体演化下的变化规律,设计了维护过程,以及相应的维

护方法,阐述了语义映射问题的解决方案,有效保证了动态环境下的生物医学集成系统的稳定性。

但是也要注意,语义的维护不能完全离开用户,但用户难以理解形式化表示的映射,也难于操作映射。因此在后续的工作中,我们将引入可视化的思想,从易懂性和易操作性解决动态环境下的映射维护问题,在语义层向用户展现多个生物数据库之间的语义关系。

参考文献(References)

- [1] Cochrane G R, Galperin M Y. The 2010 Nucleic Acids Reserach Database Issue and online Database Collection: a community of data resources[J]. Oxford University Press,2010, 38.
- [2] 李金洋. 一种基于本体的关系数据库语义集成方法研究 [D]. 国防科技大学, 2008
LI Jin-yang. A relational data base semantic integration method study based on noumenon [D]. National University of Defense Technology, 2008
- [3] Köhler J, Philippi S, Lange M. SEMEDA:ontology based semantic integration of biological databases [J]. Bioinformatics of Oxford Univerity Press, 2003, 18
- [4] Antonella Poggi, Domenico Lembo, Diego Calvanese, et al. Linking Data to Ontologies [C]. In: Proceedings of the 2006 International Workshop on OWL: Experiences and Directions (OWLED2006), 2006:133-173
- [5] Alonso-Calvo R, Maojo V, Billhardt H, et al. An agent-and ontology-based system for integrating public gene,protein,and disease databases [J]. Journal of Biomedical Informatics, 2007
- [6] Noy N F, Shah N H, Whetzel P L, et al. BioPortal:ontologies and integrated data resources at the click of a mouse [J]. Nucleic Acids Research, 2009
- [7] Yu C, Popa L. Semantic Adaptation of Schema Mappings [C]. In: VLDB'05.2005
- [8] Min H, Maniona F J, Goralczyk E, et al. Integration of prostate cancer clinical data using an ontology[J]. Journal of Biomedical Informatics, 2009, 6
- [9] Object-Management-Group. MOF 2.0 Query/Views/Transformations RFP[Z]. <http://www.omg.org/cgi-bin/doc?ad/2002-4-10>
- [10] 瞿裕忠, 胡伟, 郑东栋, 等. 关系数据库模式和本体间映射的研究综述[J]. 计算机研究与发展,2008, 45(2)
Qu Yu-zhong, Hu Wei, Zheng Dong-dong, et al. Mapping Between Relational Database Schemas and Ontologies: The State of the Art[J]. Journal of Computer Research and Development, 2008, 45(2)
- [11] Yuan An A B, Mylopoulos J. Constructing Complex Semantic Mappings between XML Data and Ontologies[Z]. 2005
- [12] Yuan An A B J M. Inferring Complex Semantic Mappings between Relational Tables and Ontologies from Simple Correspondences[Z]. 2005
- [13] Stojanovic L. Methods and Tools for Ontology Evolution [D]. ph.D, University of Karlsruhe, 2004
- [14] Lerner B S. A model for compound type changes encountered in schema evolution [C]. In: ACM Transactions on Database Systems, 1996
- [15] Nash A, Bernstein P A, Melnik S. Composition of mappings given by embedded dependencies[C]. In: PODS '05.Baltimore, Maryland, USA: 2005
- [16] An Y, Hu X, Song I. Round-Trip Engineering for Maintaining Conceptual-Relational Mappings [J]. Journal of Database Management (JDM) 2010, 21(3): 36-68
- [17] Kwon Y.S., Kobayashi A., Kajiyam S.I., et al. Antimicrobial constituents of *Angelica dahurica* roots. *Phytochemistry*, 1997, 44: 887-889
- [18] Kimuar Y., Okuda H.. Histamine-release effectors from *Angelica dahurica* var.*dahurica* root. *J. Nat. Prod*, 1997, 60:249-251
- [19] Chen I.S., Chang C..T, Sheen W.S., et al. Coumarins and antiplatelet aggregation constituents from *Formosan peucedanum*. *Phytochemistry*, 1996, 41:525-527
- [20] Yang L.L, wang M.C., Chen L.G., et al. Cytotoxic ativity of coumarins from the furits of *Cnidium monniere* on leukemia cell lines. *Planta Medica*, 2003, 69:1091-1093

(上接第 2766 页)

- [14] Wang K.T., Liu H.T, Chen X.G, et al. Identification and determination of active components in *Angelica dahurica* Benth and its medicinal preparation by capillary electrophoresis. *Talanta*, 2001, 54:753-755
- [15] 郑虎占, 董泽宏, 余靖, 等. 中药现代研究与应用[M]. 北京: 学苑出版社. 1997: 1511-1522
Zheng Hu account, Dongze Hong, Yu Jing, et al. Chinese Modern Research and Application [M]. Beijing: Academy Press, 1997: 1511-1522
- [16] Dekermendjian K., Ai J.L., Nielsen M., et al. Characterisation of the furanocoumarin phellopterin as a rat brain benzodiazepine receptor partial agonist in vitro. *Neuroscience Letters*, 1996, 219:151-153