DOI:10.3969/j.issn.1671-3982.2017.09.002

· 研究与探讨 ·

糖尿病领域本体构建与语义推理实现

刘智锋,夏晨曦,黄　梨,莫雪盈,马国庆,马敬东

[摘要]目的:构建糖尿病领域本体库与糖尿病诊疗规则库,在此基础上实现语义推理,糖尿病领域知识的重用、潜在知识的揭示及共享,为后续糖尿病管理系统研究奠定基础,进而为基层全科医生诊疗糖尿病提供决策支持。方法:以国内糖尿病领域相关临床指南和领域专家知识为依据,抽取其中的概念以及概念之间的关系,借鉴七步法和骨架法,在斯坦福大学 Protégé 平台构建糖尿病领域本体,编写 SWRL 诊疗规则库,继而使用 JESS 推理机实现语义推理。结果:构建了较为完整的糖尿病领域本体库与糖尿病诊疗规则库,包含概念 233 条、实例 205 条、实例间关系 16 条、数值属性 18 条、SWRL 规则 28 条,并在此基础上实现了语义推理。结论:构建的糖尿病领域本体能够实现语义推理任务,是将本体技术应用于慢病诊疗领域的有益探索。

[关键词]领域本体;糖尿病;本体构建;基层医疗服务;语义推理

[中图分类号]R587.1　　[文献标志码]A　　[文章编号]1671-3982(2017)09-0007-05

Development of diabetes mellitus ontology and realization of semantic inference

LIU Zhi-feng, XIA Chen-xi, HUANG Li, MO Xue-ying, MA Guo-qing, MA Jing-dong

(Medical and Health Management School, Tongji Medical College, Central China University of Science and Technology, Wuhan 430030, Hubei Province, China)

Corresponding author: MA Jing-dong

[**Abstract**] **Objective** To lay the foundation for the successive research on diabetes mellitus (DM) management system and provide support for the general doctors at grass root level to make decisions by developing DM ontology database and DM diagnosis and treatment database for semantic inference, reuse of DM knowledge, revealing and sharing potential DM knowledge. **Methods** The DM ontology was established on the Stanford University Protégé Platform according to the 7-step method and skeletal method by extracting the concepts of DM ontology database and DM diagnosis and treatment database, and their relationship from domestic DM-related clinical guidelines and knowledge of DM experts. The SWRL diagnosis and treatment rules were then composed and the semantic inference was realized using the JESS inference engine. **Results** The developed DM ontology database and DM diagnosis and treatment database included 233 concepts, 205 examples, 16 relationships between examples, 18 data value properties, 28 SWRL rules, which could thus realize the semantic inference. **Conclusion** The developed DM ontology can realize semantic inference and is thus beneficial for the application of ontology technology in diagnosis and treatment of chronic disease.

[**Key words**] Domain ontology; Diabetes mellitus; Ontology establishment; Medical service at grass-root level; Semantic inference

糖尿病是一种严重的慢性非传染性疾病,2010 年我国 18 岁以上的糖尿病患病率达到 9.7%,并且患者人数有不断增加的趋势,给个人和社会带来极大的经济负担[1]。由于国内基层设备落后,医务人员水平有限,糖尿病的漏诊率与误诊率也较高。

本体作为语义网知识表示的核心,在国内外成为研究的热点。随着本体技术不断成熟,逐渐被用

[基金项目]中央高校基本科研业务费资助项目;"区域医疗机构知识网络形成机制研究"(2015AE017);华中科技大学大学生自主创新基金(16A324)支持

[作者单位]华中科技大学同济医学院医药卫生管理学院,湖北 武汉 430030

[作者简介]刘智锋(1995-),男,福建漳州人,研究方向为卫生信息管理、数据挖掘。

[通讯作者]马敬东(1978-),男,副教授,博士生导师,研究方向为卫生信息管理、卫生政策与管理。E-mail:jdma@mail.hust.edu.cn

于医学领域的知识表示。构建相应的知识库,以此搭建医疗专家系统,实现辅助医疗诊断等功能。如牟冬梅与范轶等构造了心电图领域本体,并在此基础上实现了本体的语义推理,用于对心脏病的诊断[2];张巍和张绚等为使用计算机辅助诊断和治疗高血压,提出了基于本体和案例推理的高血压诊疗系统模型[3]等。

在糖尿病领域,国外有控制糖尿病食物领域本体的构建并通过推理对糖尿病患者饮食方案提供建议[4],有基于本体构建糖尿病临床支持系统为医务工作者提醒患者的相关病情[5],还有结合本体与语义网规则语言(SWRL)以及根据不同患者的不同身体状况,为糖尿病患者提供个性化服务如提供合适的药物[6]等。

国外也有应用糖尿病的本体技术为糖尿病患者或医务工作者提供个性化服务,但是基本没有为基层全科医生提供转诊服务。目前国内糖尿病专家系统研究主要在糖尿病诊断和数据挖掘方面,而基于本体,利用糖尿病相关指南以及专家知识对糖尿病进行诊断和治疗的研究却并不多[7]。

本文构建了糖尿病领域本体与诊疗规则模型并实现了语义推理,为后续基于本体的糖尿病管理系统研究奠定了基础。以此为基础开发基于本体的糖尿病管理系统,为基层全科医生糖尿病患者诊疗提供科学规范的决策支持,能够有效提高糖尿病的知晓率和降低各种并发症的致残率,对于控制血糖、减轻糖尿病带来的经济与社会负担具有重要意义。

1 糖尿病领域本体构建

1.1 知识库来源

糖尿病领域本体相关的概念、属性、关系以及诊疗规则主要来自国内糖尿病领域相关的临床指南及糖尿病领域专家。其中主要抽取有关糖尿病风险评估、患者症状、糖尿病诊断方法、糖尿病诊疗路径、相关治疗药物如降糖药、调脂药与降压药以及糖尿病的生活干预方式等知识,经过反复论证,确保概念以及概念关系的准确性及统一性。并按照格鲁伯提出的清晰性、完整性、一致性、最大单调可扩展性以及最小承诺原则构建本体五项原则,保证本体构建的科学性。在语义推理阶段,则根据患者健康档案数据,创建相应的患者实例,实现语义推理。

1.2 本体构建方法

目前,国内外本体构建方法主要有两种:一是直接从知识工程的角度出发构建本体,称为本体工程;二是在现有词表资源的基础上直接向本体转化[8]。

本体工程强调构造本体时要按照一定的规范和标准。目前为止,其主要的方法有 TOVE 法、七步法、骨架法、IDEF5 法以及 SENSUS 法等,还没有统一的构造本体的标准。以上方法都是在构建具体项目的本体时总结出来的。相较于其他的本体构建方法,斯坦福大学开发的七步法构建领域本体最为成熟[9],但缺少“检查和评估以及用户的反馈”环节[10]。因此本文借鉴骨架法领域本体构建方法,引进本体评价环节,创建了新的糖尿病领域本体构建方法(图 1)。

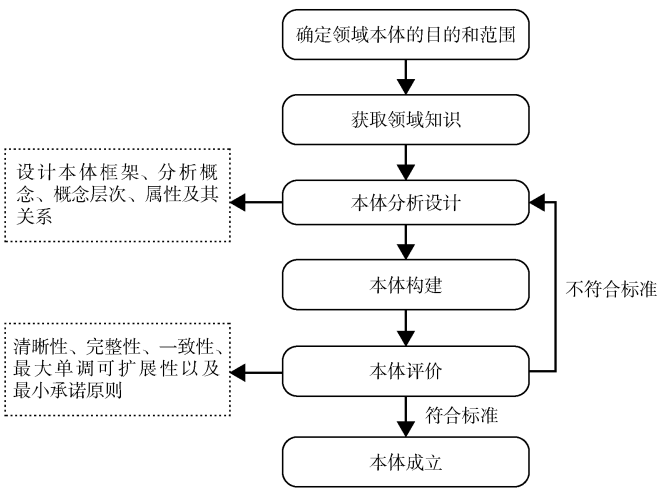


图 1 糖尿病领域本体构建方法

Maja Hadzic 和 Elizabeth Chang 等学者提出了人类疾病本体模型即疾病类型、表型、治疗手段与病因四维模型[11]。本文在此基础上将其拓展为一种含有疾病类型、临床表现、治疗手段、检查项目、并发症以及患者 6 个维度的糖尿病本体构建模型，使用新的本体构建方法在斯坦福大学 Protégé 3.5 平台构建糖尿病领域本体。

2　领域本体模型及推理规则

2.1　本体框架

糖尿病领域本体的框架主要由疾病类型、临床表现、治疗手段、检查项目、并发症以及患者 6 个核心类组成。在这 6 个核心类的基础上，逐层创建子类，增添其属性与关系以并创建实例，进而逐步完善糖尿病领域本体。

2.2　概念与概念层次

根据临床指南定义糖尿病领域的概念及它们的关系，确定概念与概念的层次。国内糖尿病领域相关的临床指南包含了大量术语以及关系，通过较完整的抽取，尽可能保证糖尿病领域概念的完整性，而概念的抽取确保了类属性的定义以及实例的创建。该研究使用自顶向下与自底向上的方法相结合定义概念，并使用本体描述语言 OWL DL 描述糖尿病领域本体，对其进行编码与形式化。

2.3　概念的属性与关系

本体中的关系代表了在领域中概念之间的交互作用，主要包含基本关系和自定义关系两部分。除了系统预设的 4 种基本关系（即 part-of 概念之间部分与整体的关系，kind-of 概念之间的继承关系、父子类关系，instance-of 概念的实例和概念之间的关系、类对象和类之间的关系），attribute-of 某个概念是另一概念的属性外，该研究还定义了 12 种关系，丰富了糖尿病领域本体库。其中，属性分为对象属性和数据类型属性，对象属性表示两个实例之间的关系，数据类型属性描述类实例与数据类型值的关系。该研究构建了 7 个数据类型属性，部分属性与关系见表 1。

表 1　糖尿病领域本体部分属性

序号	属性名称	属性类型	属性描述
1	妊娠情况	对象属性	测试者是否妊娠
2	并发症情况	对象属性	测试者是否有并发症
3	FPG 值	数据属性	测试者空腹血糖值
4	2HPG 值	数据属性	测试者的 OGTT 试验（糖负荷试验）后 2 小时的血糖值
5	HbA1c 值	数据属性	测试者糖化血红蛋白值
6	收缩压值	数据属性	测试者收缩压的值
7	诊断结果	对象属性	测试者是否有糖尿病
8	推荐药物	对象属性	给测试者推荐使用的抗糖药
9	是否转诊	对象属性	测试者是否需要转诊到上级医院

2.4　实例的创建

实例是概念的具体化，继承了概念的属性，当构建大量的实例后，就逐渐形成相关领域的知识库。实例数据经过处理形成符合格式的实例数据，方便事实库的形成，从而与规则库结合，在 JESS 推理机中实现推理。经过概念、属性以及实例的构建，最终形成部分糖尿病领域本体库。

2.5　SWRL 规则库

语义网规则语言（SWRL）是万维网联盟（W3C）于 2004 年提出的，用于描述推理规则。它是以 OWL 子语言 OWL DL/OWL Lite 和 RuleML 为基础的规则描述语言，其目的是为了让规则可与 OWL 产生结合，从而提高本体的推理能力[12]。糖尿病诊疗规则库使用 SWRL 进行构建。

结合国内糖尿病相关的权威指南、政策文件以及糖尿病专家的建议，形成规范的糖尿病诊疗路径（图 2），继而生成糖尿病诊疗规则库，主要包含早期糖尿病筛查规则、糖尿病的诊断规则、糖尿病的分型规则、糖尿病的药物推荐规则以及转诊规则，共 28 条规则（图 3）。如早期糖尿病筛查规则：人（？P）∧

FPG 值(？ p,？ x) ∧ swrlb:lessThan(？ x,7.0) ∧ swrlb:greaterThanOrEqual(？ x,6.1) ∧ _2HPG 值(？ p,？ y) ∧ swrlb:lessThan(？ y,7.8)→ 风险评估结果(？ p,有空腹血糖受损)。表示的是如果一个人,空腹血糖大于等于 6.1,且小于 7.0,葡萄糖负荷后 2 小时血糖值小于 7.8,那么这个人有空腹血糖受损。

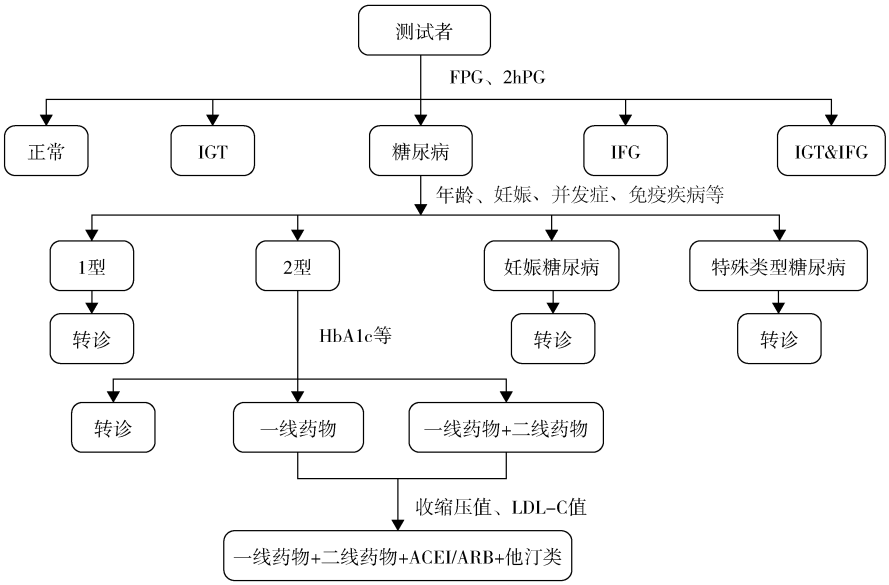


图 2 部分糖尿病诊疗流程

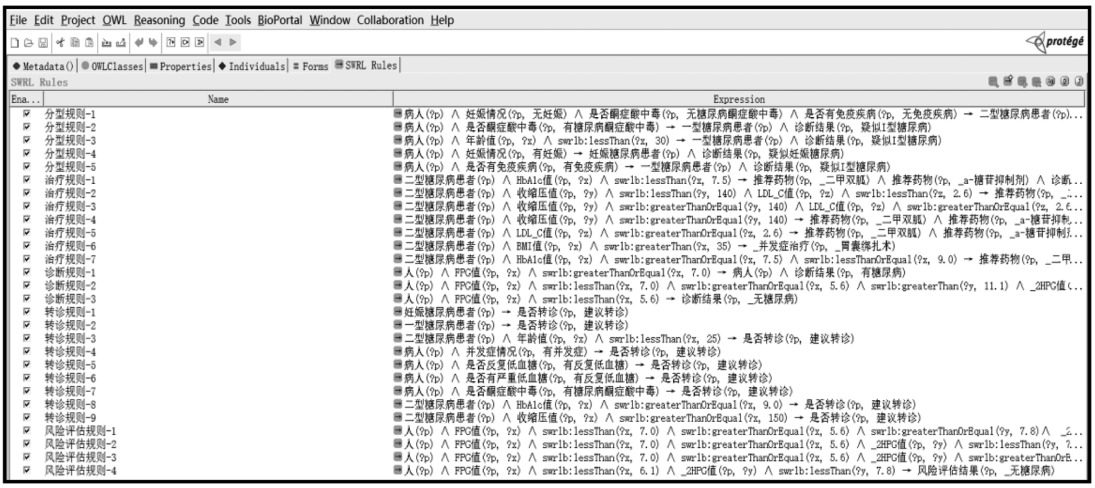


图 3 糖尿病诊疗规则库

2.6 语义推理实现

Jess 推理引擎主要由事实库、规则库和推理机 3 部分组成。先根据本体,创建相应的实例。Jess 推理机将患者领域本体转化为 Jess 事实库,再将 SWRL 规则转化为 Jess 规则库,推理机结合 Jess 事实库与 Jess 规则库,实现本体语义推理,揭示糖尿病领域潜在的知识,实现早期糖尿病的筛查、诊断、糖尿病药物推荐以及转诊服务。如图 4 所示,实例中患者李明(化名),最终诊断为糖尿病,疑似为 2 型糖尿病,并且合并有高血压和高血脂,最终推荐使用抗糖药物有 a-糖苷抑制剂、二甲双胍,降压药有血管紧张素受体拮抗剂、血管紧张素转换酶抑制剂以及降脂药他汀类药物。由于收缩压大于 150mmHg,专家建议转诊到上级医院确诊并制定诊疗方案。



图 4　语义推理实现

3　讨论

本文从 Maja Hadzic 和 Elizabeth Chang 提出的四维人类疾病本体模型出发，在此模型基础上综合应用七步法与骨架法本体构建方法，构造了较为完整的糖尿病领域本体，并且实现了基于糖尿病领域本体的语义推理，是本体技术在糖尿病领域的有意义的尝试，为糖尿病管理系统的开发打下基础。

本研究还存在一定的局限性，主要在于以下几个方面。构建的糖尿病领域本体可能不够完整，需要更多的领域专家参与；本体全部由人工创建，工作量大且繁琐，后续我们将对本体的自动化构建进行有益的探索；采用基于产生规则式的推理，不能很好地将部分知识转化为结构化的规则存储于糖尿病知识库。健康饮食与运动对糖尿病的治疗非常重要，因此糖尿病诊疗规则库也需要不断地调整和完善，使之不仅能实现糖尿病高危人群的筛查、糖尿病诊断、抗糖尿病药物的推荐以及转诊建议，还将实现糖尿病饮食与运动方面的指导功能。

【参考文献】

[1] 中华医学会糖尿病分会. 中国 2 型糖尿病防治指南(2013 年版)[J]. 中华糖尿病杂志, 2014, 6(7): 447-495.

[2] 牟冬梅, 范　轶. 基于 SWRL 推理机制的心电图本体设计与实现[J]. 情报学报, 2010, 29(5): 954-960.

[3] 张　巍, 张　绚, 陈俊杰. 基于本体的高血压诊疗系统推理模型研究[J]. 计算机工程与设计, 2013, 34(11): 4016-4020.

[4] Cantais J, Dominguez D, Gigante V, *et al*. An example of food ontology for diabetes control[C]//Proceedings of the International Semantic Web Conference 2005 workshop on Ontology Patterns for the Semantic Web. 2005: 1-9.

[5] Chalortham N, Buranarach M, Supnithi T. Ontology development for type ii diabetes mellitus clinical support system[C]//4th International Conference on Knowledge Information and Creativity Support Systems. 2009.

[6] Hempo B, Arch-int N, Arch-int S, *et al*. Personalized care recommendation approach for diabetes patients using ontology and SWRL[M]//Information Science and Applications. Springer Berlin Heidelberg, 2015: 959-966.

[7] 陈桂芬, 汪　江, 杨志刚. 基于本体的规则推理和案例推理结合的糖尿病诊疗专家系统研究[J]. 长春大学学报, 2016, 26(6): 19-25.

[8] 丁晟春, 李岳盟, 甘利人. 基于顶层本体的领域本体综合构建方法研究[J]. 信息系统, 2007, 30(2): 236-240.

[9] 张　旭, 朱勤东. 环保领域本体的构建研究[J]. 微型机与应用, 2015, 34(10): 92-95.

[10] 尚新丽. 国外本体构建方法比较分析[J]. 图书情报工作, 2012, 56(4): 116-119.

[11] Hadzic M, Chang E. Ontology-based support for human disease study[C]//Hawaii International Conference on System Sciences. IEEE, 2005: 143a-143a.

[12] 周　亮, 黄志球, 倪　川. 基于 SWRL 规则的本体推理研究[J]. 计算机技术与发展, 2015, 25(10): 67-70, 75.

[收稿日期: 2017-09-01]

[本文编辑: 王天津]