|  |
| --- |
| 中国中医科学院中医药信息研究所 |
| 《构建中医脾系证候知识体系研究》子项目 |
| 中医脾系证候知识推理与展示平台 |
|  |
| **于彤** |
| **2013/12/25** |

|  |
| --- |
|  |

**目录**

[1 项目概述 1](#_Toc375749777)

[1.1 研究目标 1](#_Toc375749778)

[1.2 研究内容 1](#_Toc375749779)

[1.2.1 搭建中医脾系证候知识体系框架 1](#_Toc375749780)

[1.2.2 探索中医知识体系模式 2](#_Toc375749781)

[1.2.3 中医脾系证候知识体系的复杂关系展示 4](#_Toc375749782)

[1.3 预期研究结果 4](#_Toc375749783)

[2 术语与定义 6](#_Toc375749784)

[3 文献综述：中医药知识库系统研究进展 7](#_Toc375749785)

[3.1 中医药知识库系统的构建方法 7](#_Toc375749786)

[3.1.1 基于关系型数据库的中医药知识库系统 7](#_Toc375749787)

[3.1.2 基于本体技术的中医药知识库系统 7](#_Toc375749788)

[3.1.3 基于神经网络的中医药知识库系统 8](#_Toc375749789)

[3.2 中医药知识库系统的覆盖范围 9](#_Toc375749790)

[3.2.1 中医人体知识库 9](#_Toc375749791)

[3.2.2 中医疾病和证候知识库 9](#_Toc375749792)

[3.2.3 中医医案知识库 9](#_Toc375749793)

[3.2.4 中药知识库 10](#_Toc375749794)

[3.2.5 中医养生知识库 10](#_Toc375749795)

[3.3 中医药知识库系统的主要应用 10](#_Toc375749796)

[3.3.1 面向中医临床诊疗的专家系统 10](#_Toc375749797)

[3.3.2 中医计算机辅助教学系统 11](#_Toc375749798)

[3.4 小结 11](#_Toc375749799)

[4 需求分析 12](#_Toc375749800)

[4.1 中医药本体知识库 13](#_Toc375749801)

[4.1.1 构建多个实验和演示用的“小本体” 13](#_Toc375749802)

[4.1.2 已有知识库的转换和处理 13](#_Toc375749803)

[4.1.3 基于Jena的本体联合 13](#_Toc375749804)

[4.2 中医药知识推理系统 13](#_Toc375749805)

[4.2.1 基于Protégé的推理 13](#_Toc375749806)

[4.2.2 基于Pellet的推理 13](#_Toc375749807)

[4.2.3 推理结果验证 13](#_Toc375749808)

[4.3 中医药知识展示平台 13](#_Toc375749809)

[4.3.1 知识问答 13](#_Toc375749810)

[4.3.2 语义查询处理 15](#_Toc375749811)

[4.3.3 综合性的知识检索与浏览 15](#_Toc375749812)

[4.3.4 知识的图形化展示 15](#_Toc375749813)

[4.4 非功能需求 16](#_Toc375749814)

[4.4.1 数据安全 16](#_Toc375749815)

[4.4.2 用户手册 16](#_Toc375749816)

[4.4.3 系统部署 16](#_Toc375749817)

[5 技术方案 17](#_Toc375749818)

[5.1 技术路线 17](#_Toc375749819)

[5.2 基于OWL的知识建模与推理方法 17](#_Toc375749820)

[5.2.1 准备工作：OWL本体的转换、装载与抽取 18](#_Toc375749821)

[5.2.2 中医证候层次结构的构建方法 31](#_Toc375749822)

[5.2.3 证候概念的定义及相关的推理方法 35](#_Toc375749823)

[5.2.4 证候特征的表达及相关的推理方法 37](#_Toc375749824)

[5.2.5 病证复合概念建模案例 38](#_Toc375749825)

[5.2.6 证候加减的建模方法 43](#_Toc375749826)

[5.3 综合性知识展示系统 45](#_Toc375749827)

[5.3.1 基于关键词的知识检索 46](#_Toc375749828)

[5.3.2 知识浏览与导航功能 47](#_Toc375749829)

[5.3.3 顶层语义网络的浏览功能 51](#_Toc375749830)

[5.3.4 简单的中医药知识问答功能 55](#_Toc375749831)

[6 项目计划 59](#_Toc375749832)

[7 讨论 60](#_Toc375749833)

[7.1 本项目的贡献 60](#_Toc375749834)

[7.2 存在的问题与未来的工作 61](#_Toc375749835)

[相关网站 63](#_Toc375749836)

[参考文献 64](#_Toc375749837)

**图目录**

图 1‑1搭建中医脾系证候知识体系框架示意图 2

图 1‑2探索中医知识体系模式示意图 4

图 5‑1 构建中医脾系证候知识体系的总体技术路线图 17

图 5‑2 基于Protégé对本体知识库进行浏览、展示、推理和模块导出 26

图 5‑3 用Protégé打开中医脾系证候的OWL本体 27

图 5‑4 Protégé界面截图（以“肝气犯胃证”为例展示） 28

图 5‑5 OntoGrap界面截图（以中医脾系证候本体中的“津亏热结证”为例展示） 29

图 5‑6 Protégé的本体模块抽取界面（以“气虚证”为例） 30

图 5‑7 Protégé 将实体相关陈述打包的界面（以“气虚证”为例） 30

图 5‑8 Protégé 对抽出的本体模块命名的界面（以“气虚证”为例） 31

图 5‑9证候分类结构 33

图 5‑10围绕“气虚证”的证候结构图 34

图 5‑11围绕“气虚发热证”的证候结构图 34

图 5‑12围绕“气虚阳微证”的证候结构图 35

图 5‑13 围绕“气虚发热证”的概念图 37

图 5‑14“气虚发热证”继承“气虚证”的特征 38

图 5‑15围绕“呕吐病，食滞内停证”的概念图 40

图 5‑16围绕“呕吐病，胃阴不足证”的概念图 40

图 5‑17围绕“噎膈病，痰气交阻证”的概念图 41

图 5‑18围绕“噎膈病，瘀血内结证”的概念图 42

图 5‑19 围绕“津亏热结证，烦渴咽燥，噎食不下，或食入即吐，吐物酸热”的概念图 45

图 5‑20知识展示系统的主要功能 46

图 5‑21 基于关键词的知识检索功能 47

图 5‑22 以“概念”为单位的知识浏览与导航功能（以“脾胃阴虚证”为例） 48

图 5‑23 简单的分类导航界面 49

图 5‑24 语义关系的导航界面（以“香砂六君子汤”的组成为例） 50

图 5‑25 语义关系的导航界面（以“哪些药物治疗肝火犯胃证”为例） 50

图 5‑26 语义关系的导航界面（以“肠道湿热证的现象表达”为例） 50

图 5‑27 语义关系的导航界面（以“胃热炽盛证的现象表达”为例） 51

图 5‑28 顶层语义网络浏览界面 52

图 5‑29 顶层语义网络浏览界面（以“证候 由…组成 症状”为例） 53

图 5‑30 语义关系搜索界面（以“证候”为例） 54

图 5‑31 语义关系搜索界面（以“证候 由…组成 症状”为例） 54

图 5‑32 中医问答原型系统之辨证场景 55

图 5‑33 中医问答原型系统之证候信息问答场景（以“肝气犯胃证”为例） 56

图 5‑34 中医问答原型系统之证候信息问答场景（以“脾胃虚寒证”为例） 57

图 5‑35 中医问答原型系统之疾病信息问答场景（以“痰饮”为例） 58

图 7‑1 中医证候领域的概念图 60

图 7‑2 中医脾系证候知识展示平台界面截图 61

图 7‑3 中医药知识工程和服务模式 61

**表目录**

表 4.1 需求要点和完成情况 12

表 6.1 项目计划 59

# 项目概述

研究目的是建立中医脾系证候知识体系，以中医脏腑部分为示范研究。包括：一是基础研究：明确中医脏腑证侯的知识分类，实现准确描述中医脾系证候概念的含义，理顺概念间逻辑关系（包括语义关系），**同时支持语义关系推理，以期获得概念间的隐性语义关系**，为中医药信息后续研究搭建科研平台，为建立规范化数据库和统一检索平台提供可能；在此部分研究的基础上，探索中医知识体系框架的构建模式，二是应用研究：利用语义审核工具筛查语义关系，把分析对象之间错综复杂的共引网状关系简化并表示出来，实现概念的多维度、多层次关系的复杂展示，为中医基础理论知识的再利用提供服务。三是培养一支中医药信息理论研究的科研团队。

## 研究目标

①以建构中医脏腑证候概念及关联关系网络图为示范研究，探索建立中医知识体系模式。②建立一支中医药信息学理论研究的科研创新团队。

## 研究内容

### 搭建中医脾系证候知识体系框架

**确定概念及其关系，准确描述以脏腑的辨证定位的中医脾系证候概念的含义，理顺概念间语义关系，同时实现语义关系推理，并以期通过逻辑推理获取概念间的隐性语义关系。**

**收集相关概念**：证候是疾病过程中机体整体的动态、病态反应，包含了疾病的病因、病位、病性及邪正双方力量的对比。而证候均能以脏腑定位。本研究依托中医药学语言系统、中医文献、病案、中医药网站、中医古籍建立中医基础理论学科语料数据库，以五脏为纲，界定每一脏的证候范围，包括脏病证候、腑病证候、脏腑兼病证候，并对中医脏腑证候的古今概念进行准确描述。以中医药学语言系统现有58种语义关系为模板，对概念间语义关系进行系统梳理与关联规范；对古今术语进行提炼、总结，还原为概念及概念间语义关系，进行准确描述、系统梳理与关联规范；收集概念名称（同名、异名、古代名、英文名、拉丁名）、来源、年代、定义、释义、关联词等；根据中医基础理论学科特点，实现概念同异名规范、标准概念与通用概念关联规范、古代概念与现代中医药概念内涵合并。

**理顺关联关系：**找出证候概念之间、证候之间的所有语义类型、语义关联关系，并实现概念术语与概念术语之间语义关系的半自动构建。建立以脏腑证候概念及关联关系为主体的中医脏腑证候知识体系。

**实现逻辑推理: 由中医基础理论知识体系中多角度、小范围选取相关概念，依托protégé 4.0软件，将概念及概念间的语义关系完整表达出来，建立示范性小本体。运用pellet推理工具对示范性小本体进行逻辑推理，以期发现概念间的隐性语义关系。对经由推理产生的隐性语义关系进行验证，将符合中医基础理论学科原理及学科逻辑的语义关系补充入中医基础理论知识体系中。**

**补充知识关系：运用本体开发工具Jena软件对多个小型示范本体进行联合，形成联合本体，运用pellet推理工具对联合本体进行逻辑推理，对经由推理产生的概念间隐性语义关系进行验证，将符合中医基础理论学科原理及学科逻辑的语义关系补充入中医基础理论知识体系中。**

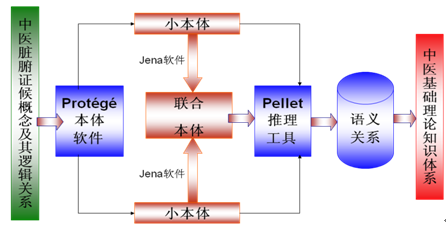


图 1‑1搭建中医脾系证候知识体系框架示意图

### 探索中医知识体系模式

**通过建构中医知识体系中医脏腑证候部分，摸索出一套较为成熟的知识体系构建方法。**

知识体系，综合来说，是指将学科相关所有知识加以科学分类，并按照一定的关系组成一个知识结构，统称为知识体系，具有开放、动态、通用和多层次的特点。知识体系的建立是为了使知识在接触→获取→应用→传播→保留→再应用的过程中，呈现出一种动态的发展。除了科学的组织显性知识，更是为了深度挖掘隐性知识，从而使知识创新成为可能。本研究在中医脏腑证候部分的示范研究基础上，以概念及概念间语义的关联关系这个角度，来探索中医学知识体系的构建原则和框架模式。

**构建原则：**

* 开放性、动态性：中医知识体系应包含中医内古今相关的知识，并具有可随时动态扩充的功能。
* 通用性、规范性：知识体系的各个环节必须规范化，才能适应各种科研需求，包括建立对知识概念间的组织与连接进行规范，建立知识库的制作规范，建立知识体系的组织构建与维护应用规范以及交互操作规范。

**探索中医知识体系框架模式：**

（1）调研知识体系构建现状：调研国内外相关行业中知识体系的构建模式，并思考本研究的应用前景。

（2）建立中医脾系证候知识体系框架：包括知识库、知识模块、推理模块、评估模块等四个模块。其中知识库主要把古今中医药相关的专门数据从知识源中提炼出来，并转化为明确的概念表示，完成知识的存储，重在确立提取相关数据的原则，人机结合验证原则；知识模块重在梳理并集合中医概念之间的逻辑关系集（包括语义关系），并设计出可随时扩充概念层次和结构，随时根据需求形成同类知识的多个大小不等的本体；推理模块主要是根据计算机方法和技术，在平台基础上设计出推理机，对本体的语义关系进行推理分析和信息比较，并作出推理结果，同时及时更新知识用于下一次的推理：评估模块根据推理结果完成人机结合验证，并对验证后不恰当的结果进行修订和评价，最终的评价结果将补充入知识模块。

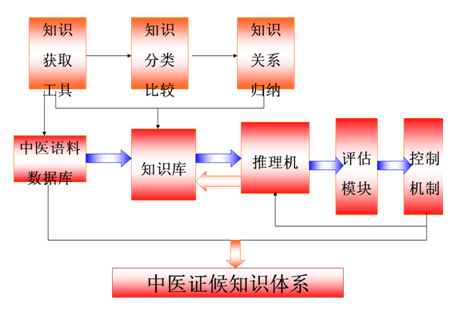


图 1‑2探索中医知识体系模式示意图

（3）探索中医知识体系模式：以证候部分研究为基础，探讨中医知识体系构建的模式。

### 中医脾系证候知识体系的复杂关系展示

**基于中医药语言系统，明晰知识元之间的多种关系。可用知识网络或概念地图的方式展示知识元之间的横向关系、树形关系以及网络关系。**

**中医药证候知识体系展示平台的搭建**

以我所服务器为运行平台，**采用Browser–Server 方式的Web 应用程序开发设计中医脾系证候知识体系展示平台。**

**中医脾系证候知识体系图形展示**

中医理论历史悠久，其概念间的关系错综复杂。本研究即要根据中医脾系证候理论概念间的关系，以概念为中心，展示与其相关的概念，并能逐级展开，以显示概念间的相互关系。在展示功能方面，平台不仅提供单一概念间的相互关系，还可提供多概念间、多层级的可选择性的知识元关系的展示；在展示形式方面，平台提供树形结构点击展开方式和网状展开方式两种形式；平台还提供点击概念展示和查询概念展示两种方式。平台将以多角度多元组合的形式展示中医脾系证候知识体系。

**中医脾系证候知识体系列表**

理清中医脾系证候基本概念，梳理以脏腑为病位的中医药证候知识体系，是本研究的根本目的。本研究不但能为中医基础理论提供图形化形象化的网状中医概念图形，更能清楚地以列表形式表现中医脾系证候概念，更加条理化的显示中医脾系证候知识体系。

中医脾系证候知识体系列表能具体罗列出概念的所有相关概念及其关系。这些关系不但能为研究人员在探究中医基础理论知识的关系上提供清晰的思路；还能为中医脾系证候知识的再利用提供服务。

## 预期研究结果

中医脾系证候知识体系及其网状关系展示平台；

在核心期刊发表论文4篇 ；

培养一支中医药信息理论研究人才队伍。

# 术语与定义

**Y**

**语义Web**

一个机器可以理解的Web。

未完成。

# 文献综述：中医药知识库系统研究进展

近年来，知识库系统成为中医药知识遗产保护和利用的创新手段。本文从构建方法、覆盖范围、应用情况等3方面对中医药知识库系统的研究进展进行了综述，以供中医药信息化领域的研究人员参考。

构建中医知识库系统，是指用人工智能技术把中医药理论和专家的经验按规范化、标准化的格式，组建成知识库，达到资源查询和共享。本文从构建方法、覆盖范围、应用情况等3方面对中医药知识库系统的研究进展进行了综述，以供中医药信息化领域的研究人员参考。

## 中医药知识库系统的构建方法

### 基于关系型数据库的中医药知识库系统

王连心等提出以关系型数据库为基础构建中药知识库的方案。据此方案，中药知识库分为3个层次：第一个层次是事实数据层，包含存在于文献资料中的原始数据；第二个层次是知识本体层，其内容是领域概念和领域规则；第三个层次是策略层，其内容是策略性的结论和规则。在中药知识库的构建中，可将关系型数据库作为事实数据层，在其基础上建立准则层和决策层，使中药数据上升为中药知识。陈国宁在构建中医咳感症诊断专家系统的过程中，采用关系数据库作为知识库的存储组织形式，并实现了知识的增添、删除、修改等操作，以支持推理机制和咨询模块。

### 基于本体技术的中医药知识库系统

本体（Ontology）是一个源自哲学领域的术语，探讨世间万物如何存在。在计算机科学领域，本体是针对一个概念体系的正式而明确的规范。本体一般针对特定领域和议题而建，它使交互各方对特定领域内共用的概念、词汇以及概念分类等达成一致，从而支持知识的共享和重用。近年来，本体技术逐渐成为构建中医药知识库系统的主流手段。

徐彬锋等将本体技术引入到医学知识组织当中，构建出包括中医本体和西医本体两大分支的医学本体框架。该本体定义了医学领域的类、类间关系和类及类间关系必须满足的约束。其中，中医本体包括中医疾病类、中医证类、中医舌诊类、中医脉诊类、中医病案类；西医本体包括疾病属性类、化学物质类、疾病类、诊断类和器官类。Protégé工具被用于构建和展示这一本体。这项工作展示了本体知识库的可重用性和可扩展性等优点，但构建全面、完整的医学本体仍是一个长期而复杂的过程。

谷建军基于本体技术构建了一个中医古籍知识库。中医古籍所蕴含的知识相当丰富，知识之间的关联关系错综复杂，每一部中医古籍都作为一个本体存在，将所有中医古籍构成一个综合本体的难度可想而知。谷建军提出基于中医古籍文献叙词表构建领域本体的方法，该叙词表以一个树状结构表达了中医古籍文献的概念体系，其中的概念来自于中医古籍内容。谷建军使用Protégé 2000工具将该叙词表转换为一个OWL本体，该本体继承了叙词表的概念结构，但对概念之间的语义关系进行了更为深入细致的描述，为建立文献之间的知识关联提供了可行的途径。

此外，还有很多学者开展了构建中医药本体知识库的探索工作。例如，车立娟等以Protégé为工具，构建了用于表达“肺阴虚证”病机规则的本体知识库；李新霞以Jena为工具，建立了中医脾胃病领域本体知识库；孙海舒等从规范控制、构建原则、本体构建工具、系统架构、术语规范化、构建本体模型等诸多方面，对基于本体的中医古籍知识库的构建进行了探讨；易钢等基于本体构建了一个中医知识库系统，其中包括中医概念、关系、推理及中医诊断等大量知识。这些工作初步验证了将本体技术用于构建中医药知识库的可行性，为中医领域内的知识组织提供一种新的方法。

### 基于神经网络的中医药知识库系统

中医诊断学知识具有不确定性、复杂性、模糊性等特点，需要能够充分模拟症状与证候之间的非线性映射关系的数据模型。人工神经网络（Artificial Neural Network, ANN）是由大量处理单元（神经元，Neurons）互连而成的网络，是对人脑的抽象、简化和模拟。ANN具有非线性映射功能，其通过对输入输出样本的自动学习，能够以任意精度逼近任意复杂的非线性映射，因此很适合模拟中医诊断的推理过程。

吴芸等使用神经网络方法，构建了面向中医舌诊八纲辨证的知识库。该项研究基于“舌象”与中医“八纲”（包括表、里、寒、热、虚、实等）之间的对应关系，建立了“概率神经网络”、“广义回归神经网络”和“学习矢量量化神经网络”等3种类型的神经网络，并尝试将其用于中医诊断推理。另外，周金海等也利用ANN算法构建了中医舌诊知识库，它所预测的舌像与八纲辨证的关系与临床诊断结果基本一致（个别属性存在误差）。这些工作说明采用神经网络技术构建中医诊断知识库是可行的。

## 中医药知识库系统的覆盖范围

### 中医人体知识库

赵静等初步建立了一个“三维可视人中医知识库”，将经络腧穴学等中医知识融入三维可视人模型中，构成了具有中医特色的人体三维医学影像图谱。该知识库使用语义网络方法，描述腧穴与经络之间的归属关系，以及腧穴与解剖结构（包括肌肉、骨骼、内脏、血管、神经等）之间的对应关系；它还包含腧穴标准定位、功能主治、针刺方法、进针层次、毗邻结构以及针刺意外等针灸临床知识。这项工作为今后开发基于“中国虚拟人”的腧穴和经络三维可视化模型提供了支持。另外，郑雷在三维人体可视化模型上对中医腧穴进行了系统定位，并开发了能够显示中医针灸针刺过程的三维影像浏览器。

### 中医疾病和证候知识库

在医学知识中，最重要、最核心的知识是关于疾病的知识。徐彬锋等以冠心病知识为例，建立了中西医知识相结合的本体知识库。该系统定义了“冠心病”及相关疾病的中英文名称、定义、发病率、病因和主要治疗方法等信息，并进一步定义了诊断学公理。这些公理是医学元知识（即关于医学知识的知识），比一般的医学知识具有更高的抽象层次。例如，发病率大于０小于１；疾病的并发症是一种疾病；疾病的人群死亡率小于发病率等。该系统再经完善、扩展后可用于中医辅助诊疗。

谷建军建立以“病证”为中心的本体模型，将中医古籍中的相关概念和知识加入本体之中。该本体包括证候、病位、病因、病机、鉴别诊断、治则、治法、用药、处方、治疗禁忌等类型，明确了类型的属性和相互关系，能较为准确地反映中医对“病症”的理解和认识。

李新霞以老中医关于脾胃病的病案为基础，建立了中医脾胃病领域本体知识库。该项目对中医症状术语进行了规范化处理，将零散的中医脾胃病知识组织起来，实现了疾病诊断、证候诊断、治法、方药等方面的推理功能，以支持知识共享和疾病诊疗。

### 中医医案知识库

中医医案是中医临床思维活动和辨证论治过程的记录，是中医理法方药综合应用的具体反映形式。特别是名老中医的医案，对于中医理论和方法的传承具有重要意义。近年来，在名老中医经验传承的背景下，中医医案知识库的构建成为学术界的一个热点。彭笑艳采用本体、Agent等技术，构建了中医医案知识库，其中主要包含中医肝病、胃病等领域的医学知识、材料和医案，为用户提供在线诊疗服务。曹宇峰基于本体技术设计了中医病案知识模型，构建了中医病案的知识库，以辅助学生通过大量病案实践来巩固所学的理论知识。

### 中药知识库

中药知识是提高中医临床与学术水平的重要环节和文献宝库之一，值得当代医家学习与研究。王连心等提出了一套中药知识库的设计方案，该知识库的内容包括中药的名称、出处、剂量、性味、归经、功效、主治、用药宜忌等信息，以及方剂的处方来源、剂型、药物组成、功效主治、制备方法、用法用量与禁忌等。这些知识通过对古典医籍、学术专著、学术期刊和报纸、各种学术会议论文和报告以及相关指南等进行普查与筛选、分析与整理而获得。该方案还对中药命名、药效和剂量的规范化问题进行了讨论。

### 中医养生知识库

《黄帝内经》是中医药理论体系的奠基之作，其养生理论和方法对中医养生学有着深远的影响[3]。臧知明采用Linux、MySQL、PHP等技术，初步构建了面向网络应用的《黄帝内经》养生知识库。臧知明对《黄帝内经》的语言特点和思维方式进行了分析，提出了基于词句的《黄帝内经》养生知识的表达模式，并提出了《黄帝内经》养生知识库的框架。

## 中医药知识库系统的主要应用

### 面向中医临床诊疗的专家系统

中医专家系统是指用计算机人工智能技术来模拟著名老中医诊疗病人的临床经验，从而使该软件具有专家水平的诊治病人的能力。知识库是专家系统的核心部分，它决定系统设计的正确与否，也影响系统其他模块设计的难易以及系统运行的效率。关幼波、郭荣江、马斌荣等于1979年研制了“中医关幼波诊疗肝病的计算机程序”。该系统开临床诊疗专家系统之先河，并启动了中医信息化的进程。陈国宁等采用产生式规则来表示专家知识，以知识库的形式对这些知识进行组织，并实现了混合推理的机制；在此基础上，实现了一个能对中医咳感症进行诊治的专家系统。在中医诊疗技术现代化的背景下，中医专家系统、智能化辅助中医诊疗系统以及中医诊疗仪器和设备的研发工作成为中医药信息化建设的热点，中医药知识库系统作为其中的支撑技术也得到了学术界的重视。

### 中医计算机辅助教学系统

中医药知识库能有效支撑领域知识的检索、浏览和形象化展示，因此对于中医教学工作具有良好的辅助作用。徐彬锋等将其开发的知识库系统用于《医学基础》课程的教学工作，取得了良好的效果。学习者反映，使用这个系统可以迅速定位到系统所包含的概念，从而加速对相关概念的理解。但该系统的内容仅限于知识库构建者所掌握的医学知识，而未覆盖所有的相关知识。为提升该系统在教学中的实际价值，尚需扩展其范围并完善其功能。

## 小结

中医药知识库系统在中医药信息化和中医药知识工程等领域中都扮演着核心的角色。在中医药领域，关系型数据库目前是实现知识库最为实用的手段；领域本体则在复杂知识建模和推理等方面显示了技术优势，且其实现技术和工具日趋成熟；学者们也在探索使用人工神经网络等机器学习和人工智能手段来实现知识库，体现出了较大的潜力。知识库系统已被用于中医疾病、证候、医案、中药、养生和古籍等诸多领域，且在辅助临床诊疗、辅助教学等方面取得了一些实际应用。但从现有的文献来看，中医药知识库系统的建设尚不成熟，还有很长的路要走。我们尚需在术语规范化、知识建模方法、本体工程方法、知识获取方法和推理机制等方面开展深入研究，并进一步开展中医药知识库系统在中医药信息化中的实际应用。

# 需求分析

本章给出平台系统的需求分析，包括功能需求和非功能需求。首先，根据项目任务书，分析整个项目完成情况，得出我们这个子项目的需求要点。

表 4.1 需求要点和完成情况

|  |  |  |
| --- | --- | --- |
| 需求要点（来自任务书） | 完成情况 | 讨论 |
| 理顺中医概念之间的关系，构建一个知识库来表达中医知识体系。 | 基本完成 | 已构建了一个知识库，它存储于关系型数据库中。  但可能缺乏有关中医治则、功效、病因等方面的知识，拟从中医病案库、TCMLS等库中导入，或补充进行一些知识编辑工作。 |
| 构建知识展示平台。 | 部分完成 | VSP开发了知识展示平台，包括知识推理、图形化展示等元素。  主要问题：友好性较差，展示效果一般；未能充分展示知识库内容和推理结果。  拟构建一个基于Web的中医药知识展示系统，以展示知识库内容和推理结果，与VSP开发的知识展示平台相互补充，满足任务书的要求。 |
| 构建多个小本体。 | 尚未开展 | 使用protégé 4.x软件，将概念及概念间的语义关系完整表达出来，建立示范性小本体。这项工作尚未开展。 |
| 将知识库转换为OWL本体，并导入Protégé软件。 | 尚未开展 | 使用Python或Java语言编写脚本，将存储与关系型数据库中的知识转化为OWL本体，并导入Protégé软件，从而基于Protégé进行推理、可视化等操作。 |
| 基于Jena的本体联合 | 尚未开展 | 运用本体开发工具Jena软件对多个小型示范本体进行联合，构成一个整合的知识库，支持综合性知识检索。 |
| 基于Pellet进行推理，推理结果需要人工验证。 | 尚未开展 | 运用pellet推理工具对示范性小本体进行逻辑推理，以期发现概念间的隐性语义关系。对经由推理产生的隐性语义关系进行验证，将符合中医基础理论学科原理及学科逻辑的语义关系补充入中医基础理论知识体系中。 |

综上，本项目的主要任务分为3部分：

（1）构建本体知识库：现有的知识库存储于关系型数据库中，与任务书要求不符。将已有知识库转换为OWL本体，构建多个小本体，将它们联合起来，导入Protégé软件。

（2）构建知识推理系统：运用pellet推理工具对本体知识库进行逻辑推理，以期发现概念间的隐性语义关系。对经由推理产生的隐性语义关系进行验证，将符合中医基础理论学科原理及学科逻辑的语义关系补充入中医基础理论知识体系中。

（3）构建知识展示系统：构建基于Web的知识展示系统，以问答、检索、浏览等方式展示知识库内容，与VSP开发的知识展示平台相互补充，满足任务书的要求。

## 中医药本体知识库

### 构建多个实验和演示用的“小本体”

使用protégé 4.x软件，将概念及概念间的语义关系完整表达出来，建立示范性小本体。

### 已有知识库的转换和处理

将中医药知识库（也可能包括中医个案库、TCMLS等补充数据）转换为OWL本体，并导入Protégé软件。使用Python或Java语言编写脚本，将存储与关系型数据库中的知识转化为OWL本体，，并导入Protégé软件，从而基于Protégé进行推理、可视化等操作。

### 基于Jena的本体联合

运用本体开发工具Jena软件对多个小型示范本体进行联合，构成一个整合的知识库，支持综合性知识检索。

## 中医药知识推理系统

### 基于Protégé的推理

将中医药知识库并导入Protégé软件，从而基于Protégé进行推理、可视化等操作。

### 基于Pellet的推理

运用Pellet推理工具对示范性小本体进行逻辑推理，以期发现概念间的隐性语义关系。

### 推理结果验证

对经由推理产生的隐性语义关系进行验证，将符合中医基础理论学科原理及学科逻辑的语义关系补充入中医基础理论知识体系中。

## 中医药知识展示平台

### 知识问答

系统以知识问答的方式，向用户提供知识库系统中的内容。症状是医生可以观察到的，病因是医生可能从对用户的问询和对时令节气的判断中得到的；因此在中医推理系统里，系统输入应该既包括症状，又包括病因（比如六邪）。系统可通过提供选项来不断获取用户信息，并展示推理过程和依据。下面是一些知识问答的案例：

1. 输入症状： 怒

输出：

怒->愤怒->肝火炽盛->心火亢盛/肝火犯肺；

怒->郁怒->肝气郁结->脾气不运->肝郁脾虚。

2. 输入症状：耳鸣

输出：

耳鸣->肾精亏虚(肾开窍于耳)

耳鸣->肝胆火热(足少阳胆经入耳中)

3. 输入病因：风邪+寒邪

输出：

风邪+寒邪->风寒表实证

4. 输入病因：风邪+热邪

输出：

风邪+热邪->风热表证

5. 输入症状：协痛，咳嗽阵作，甚则咳血，急躁易怒，头胀，头晕，口苦，口渴

输出：

以上『』->肝火犯肺

### 综合性的知识检索与浏览

系统能够根据用户的请求，向用户提供综合性的知识检索与浏览服务。用户输入检索词，系统将知识库中的知识，按照主题进行综合性展示。例如，平台能够在同一张页面中对“人参”的知识在一张综合性页面中展示出来。另外，主题页面之间可通过超链接关联起来。例如，可由“四君子汤”的页面跳转到的“人参”的页面（方剂组成关系），由“人参”的页面跳转到“人参皂苷”等等。

### 知识的图形化展示

中医领域概念之间的关系错综复杂。例如，五脏中的肾按阴阳学说划分为阴，按五行划分为水，而肾可以进一步划分为肾阴和肾阳，而肾阴和肾阳又有多种变化关系。知识图谱是展示这些复杂关系的有效手段。针对中医药这样一个复杂领域，静态图往往各种线条纷繁杂乱，用户分不清楚，故而不能满足需求。因此在可视化方面，拟采用动态图谱的方式，以动态可拖动的元素的形式来展现中医药知识模型。

## 非功能需求

### 数据安全

保持数据的一致性，不会丢失数据，能够安全、可靠地运行。

### 用户手册

提供详尽、易懂的用户手册。

### 系统部署

协助用户在用户指定的服务器上安装部署。

# 技术方案

本章给出中医脾系证候知识展示平台的技术方案。构建中医脾系证候知识体系的总体技术路线如所示。下面将讨论拟解决的关键问题，阐述平台的设计原则、体系结构、核心方法和核心代码等内容。

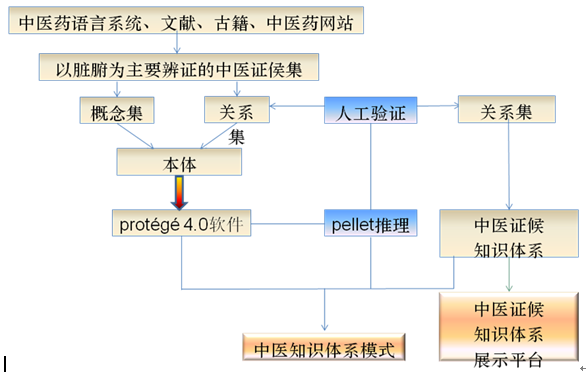


图 5‑1 构建中医脾系证候知识体系的总体技术路线图

## 基于OWL的知识建模与推理方法

“中医脾系证候”本体的核心内容在于表达中医脾系证候的知识体系。该本体以八纲辨证、精气血津液辨证、脏腑辨证和经络辨证等多种方式对证候概念进行了系统组织，并对证候的相关信息进行了精确表达。为描述证候的临床表现、治疗方法等外延知识，该本体还定义了相关的症状、疾病、中药和方剂等概念。

中医脾系证候本体分类体系的根节点是“事物”。本体首先将“事物”分为“实体”和“事件”两类。在“事件”下面定义了“证候”、“疾病”、“症状”，在“实体”下面则定义了“方剂”和“中药”等。其中，“证候”是该本体的核心部分。下面将从证候层次结构、证候定义、证候特征、病证复合概念和证候加减等方面，介绍该本体所使用的知识建模与推理方法。

### 准备工作：OWL本体的转换、装载与抽取

在此，介绍将基于Oracle的知识库转换为OWL本体并载入Protégé，以及从OWL中抽取本体模块的方法。中医脾系证候加工平台的后台为一个关系型数据库（设为d，代表database），用于存储知识库的内容。为将中医脾系证候载入Protégé中进行浏览、编辑和推理，需要将关系型数据库的内容转换为OWL格式。在本节中，将讨论转换方案并介绍转换脚本的代码。

转换脚本的功能是将关系型数据库d转换为一个OWL文件（设为f，代表file），分为如下部分：

1. 建立数据库d的连接：

|  |
| --- |
| import mysql.connector  #设置数据库配置参数  config = {  'user': '…', #设置数据库用户名  'password': '…', #设置数据库用户密码  'host': '127.0.0.1', #设置数据库主机IP地址  'database': 'spleen', #设置数据库名  'raise\_on\_warnings': True,  }  #采用配置参数config，建立数据库连接  cnx = mysql.connector.connect(\*\*config)  #获得访问数据库的游标  cursor = cnx.cursor(**buffered=True**) #若buffered为ture，在cursor中会获得SQL语句执行后的所有结果 |

1. 打开OWL文件f，用于写入知识库内容：

|  |
| --- |
| #以“可写（‘w’）”的方式打开'spleen-1.0.owl'，编码方式为‘utf-8’  f = open('spleen-1.0.owl', 'w', encoding='utf-8') |

1. 在f中写入命名空间的前缀

|  |
| --- |
| #在f中写入XML Schema的前缀  f.write('Prefix: xsd: <http://www.w3.org/2001/XMLSchema#>')  #在f中写入换行符（’\n’）  f.write('\n')  #在f中写入OWL的前缀  f.write('Prefix: owl: <http://www.w3.org/2002/07/owl#>')  f.write('\n')  #在f中写入XML的前缀  f.write('Prefix: xml: <http://www.w3.org/XML/1998/namespace>')  f.write('\n')  #在f中写入RDF Schema的前缀  f.write('Prefix: rdfs: <http://www.w3.org/2000/01/rdf-schema#>')  f.write('\n')  #在f中写入RDF的前缀  f.write('Prefix: rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>')  f.write('\n')  #在f中写入默认的命名空间  f.write('Prefix: : <http://www.example.com/>')  f.write('\n') |

1. 为本体命名：

|  |
| --- |
| #在本示例程序中，将本体命名为“http://www.example.com/”，可根据部署情况修改为其他名称  f.write('Ontology: <http://www.example.com/>\n')  f.write('\n') |

1. 为本体添加注释：

|  |
| --- |
| #在RDF中，用“rdfs:comment”为本体添加注释  f.write('Annotations:')  f.write('\n')  f.write(' rdfs:comment "中医脾系证候知识库。"@zh,')  f.write('\n')  f.write(' rdfs:comment "An OWL-based knowledge base for syndromes  in TCM spleen system."@en')  f.write('\n') |

1. 在本体中添加注释属性：

|  |
| --- |
| #在RDF中，用“rdfs:seeAlso”表达“另请参见”的意思  f.write('AnnotationProperty: rdfs:seeAlso')  f.write('\n')  #在RDF中，用“rdfs:seeAlso”表达“定义来源”的意思  f.write('AnnotationProperty: rdfs:isDefinedBy')  f.write('\n')  #在RDF中，用“rdfs:seeAlso”表达“标签”的意思  f.write('AnnotationProperty: rdfs:label')  f.write('\n')  #在RDF中，用“rdfs:seeAlso”表达“注释”的意思  f.write('AnnotationProperty: rdfs:comment')  f.write('\n') |

1. 从数据库中读取语义网络。如下表所示，语义网络在数据库中被存储为“Subject-Property-Value”的表格：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Subject | Property | Value | Subject | Property | Value |
| 人参 | 功效 | 补脾 | 四君子汤 | 组成 | 人参 |
| 人参 | 功效 | 益肺 | 四君子汤 | 组成 | 白术 |
| 人参 | 功效 | 生津 | 四君子汤 | 组成 | 茯苓 |
| 人参 | 功效 | 止渴 | 四君子汤 | 组成 | 甘草 |
| 人参 | 主治 | 劳伤 | 四君子汤 | 主治 | 脾胃气虚 |
| 人参 | 主治 | 虚损 | 四君子汤 | 功效 | 益气 |
| 人参 | 主治 | 食少 | 四君子汤 | 功效 | 健脾 |
| 人参 | 主治 | 倦怠 | … | … | … |

转换脚本将该表读出，存入一个Python Dict对象中，以供进一步处理：

|  |
| --- |
| graph={人参： {功效：[补脾，益肺，生津，止渴],  主治：[劳伤，虚损，食少，倦怠]}，  四君子汤： {组成：[人参，白术，茯苓，甘草]，  主治：[脾胃气虚]，  功效：[益气，健脾]},  …… } |

数据读取的代码如下：

|  |
| --- |
| cursor.execute("SELECT subject, property, value FROM graph")  graph = {}  for (subject, property, value) in cursor:  if (subject in graph) :  if property in graph[subject]:  graph[subject][property].append(value)  else:  graph[subject][property] = [value]  else:  graph[subject] = {property:[value]} |

1. 定义对象属性（ObjectProperty）。将知识库中的关系名称（如"由...组成"、"来源于"、"被...治疗"、"治疗"等）定义为OWL中的对象属性。

|  |
| --- |
| #列举知识库中的关系名称：  rel\_props = ["由...组成", "来源于", "被...治疗", "治疗", "加", "有基础方",  "组成", "现象表达", "…的现象表达", "由……组成", "症状加（调整）",  "中药加（调整）", "证候（调整）", "减", "方剂（调整）", "由……导致",  "导致", "方剂加（调整）", "中药减（调整）", "加（反向）", "症状减（调整）"]  #将知识库中的关系名称定义为OWL中的“对象属性（ObjectProperty）”：  for (name) in rel\_props:  f.write('ObjectProperty: ' + name + '\n')  f.write(' Annotations:\n')  f.write(' rdfs:label "' + name + '"@zh\n')  f.write('\n') |

1. 定义标注属性（ObjectProperty）。将知识库中的标注属性（如"病位"、"归经"、"功效"、"性味"等）定义为OWL中的标注属性。

|  |
| --- |
| #列举知识库中的关系名称：  anno\_props = ["病因病机", "病位", "归经", "功效", "性味", "作者", "出版社",  "出版时间", "出处", "病性"]  #将知识库中的关系名称定义为OWL中的“标注属性（AnnotationProperty）”：  for (name) in anno\_props:  f.write('AnnotationProperty: ' + name + '\n')  f.write(' Annotations:\n')  f.write(' rdfs:label "' + name + '"@zh\n')  f.write('\n') |

1. 定义OWL类型（OWL Class）。将知识库中的语义类型（如"中药"、"方剂"、"证候"、"症候"等）定义为OWL Class。

|  |
| --- |
| #从知识库的cls表中查出语义类型名称：  cursor.execute("SELECT value FROM cls")  for (col) in cursor:  if (col[0] != None) & (col[0] != ''):  owl\_classes.append(col[0])  #将知识库中的语义类型定义为OWL Class：  for (name) in owl\_classes:  f.write('Class: ' + name + '\n')  f.write(' Annotations:\n')  f.write(' rdfs:label "' + name + '"@zh\n')  f.write('\n') |

1. 读取知识库中所有概念的标识和名称。

|  |
| --- |
| #从知识库的def表中查出所有概念的标识和名称：  cursor.execute("SELECT id, name FROM def")  names = {}  for (id, name) in cursor:  names[id] = name |

1. 对每个概念进行处理。

|  |
| --- |
| #遍历知识库中每个概念，id为概念的id号，name为概念名称，subject为概念在库中的完整标识。  for (id, name) in names.items():  subject = 'spleen:o' + str(id)  #以下是对当前概念进行处理的代码 |

下面介绍对每个概念的处理方法。

1. 读取当前概念的类型，存入types变量。

|  |
| --- |
| types = set()  if (subject in graph) :  if ('类型' in graph[subject]) :  for obj in graph[subject]['类型']:  types.add(obj) |

1. 读取每个概念的属性值类型，存入equivs、super\_classes、some\_restriction、annotations等变量。

|  |
| --- |
| super\_props = ["上位词","有基础方"]  label\_props = ["病名", "名称", "别名", "方名"]  equivs = set()  super\_classes = set()  some\_restriction = set()  annotations = {'rdfs:label "' + name + '"@zh'}    if (subject in graph) :  for (property, values) in graph[subject].items():  #将“上位词”、“有基础方”等关系直接翻译为OWL中的父子类关系  if (property in super\_props) :  for obj in values :  super\_classes.add(obj.replace('spleen:', '', 1))  #将“描述”翻译为rdfs:comment  if (property == '描述') :  for obj in values :  annotations.add('rdfs:comment "' + obj + '"@zh')  #将概念关系翻译为some restriction  if (property in rel\_props):  for obj in values :  some\_restriction.add(property + " some " +  obj.replace('spleen:', '', 1))  #将概念注释翻译为owl中的AnotationProperty  if (property in anno\_props):  for obj in values :  annotations.add(property + ' "' + obj + '"@zh')  #将概念标签翻译为rdfs:label  if (property in label\_props):  for obj in values :  annotations.add('rdfs:label "' + obj + '"@zh') |

1. 对“证候加减”的处理。

|  |
| --- |
| #若概念类型为“证候加减”，则执行下段代码  if ('证候加减' in types):  clauses = set()  #对症状加（调整）的处理  if ('症状加（调整）' in graph[subject]):  for obj in graph[subject]['症状加（调整）']:  clauses.add(obj.replace('spleen:', '', 1))  #对症状减（调整）的处理  if ('症状减（调整）' in graph[subject]):  for obj in graph[subject]['症状减（调整）'] :  clauses.add('(not ' + obj.replace('spleen:', '', 1) + ')')  #对证候（调整）的处理  if ('证候（调整）' in graph[subject]):  for obj in graph[subject]['证候（调整）'] :  clauses.add(obj.replace('spleen:', '', 1))  if len(clauses) > 1:  equivs.add(' and '.join(clauses))  #对方剂（调整）的处理  if ('方剂（调整）' in graph[subject]):  for obj in graph[subject]['方剂（调整）'] :  some\_restriction.add("被...治疗 some " + obj.replace('spleen:', '', 1))  #对中药加（调整）的处理  if ('中药加（调整）' in graph[subject]):  for obj in graph[subject]['中药加（调整）'] :  some\_restriction.add("被...治疗 some " + obj.replace('spleen:', '', 1)) |

1. 对“方剂加减”的处理。

|  |
| --- |
| #若该概念“有基础方”，则处理其加减情况：  if ('有基础方' in graph[subject]):  clauses = set()  #对方剂“加药物“的处理  if ('加' in graph[subject]):  for obj in graph[subject]['加'] :  clauses.add("(加 some " + obj.replace('spleen:', '', 1) + ")")  #对方剂“减药物“的处理  if ('减' in graph[subject]):  for obj in graph[subject]['减'] :  clauses.add("(减 some " + obj.replace('spleen:', '', 1) + ')')  for obj in graph[subject]['有基础方'] :  clauses.add(obj.replace('spleen:', '', 1))  if len(clauses) > 1:  equivs.add(' and '.join(clauses)) |

1. 将该概念的OWL定义写入f。

|  |
| --- |
| #将概念声明为Class  f.write('Class: o' + str(id) + '\n')  f.write('\n')  #将概念注释写入f  f.write(' Annotations:\n')  f.write(' ' + ',\n '.join(annotations) + '\n')  f.write('\n')  #将概念的等价关系写入f  if len(equivs) != 0 :  f.write(' EquivalentTo:\n')  f.write(' ' + ',\n '.join(equivs) + '\n')  f.write('\n')  #将概念的父子类关系写入f  super\_classes |= some\_restriction  if len(super\_classes) != 0 :  f.write(' SubClassOf:\n')  f.write(' ' + ',\n '.join(super\_classes) + '\n')  f.write('\n') |

1. 将缓存在f中的内容输出到文件，并关闭f：

|  |
| --- |
| f.flush()  f.close() |

（程序结束）

运行上述转换脚本，即可生成一个面向中医脾系证候知识的OWL本体文件。下面介绍使用Protégé对该本体文件进行处理的方法。Protégé是一个免费、开源的本体编辑器，其支持的本体文件格式包括OWL，RDF（S）和XML Schema等。Protégé主要包括2个版本：桌面客户端和Web应用程序。在本项目中，主要使用Protégé客户端两个版本或支持建模本体。Protégé是基于Java的，具有可扩展性，并提供了一个“即插即用（plug-and-play）”的编程环境，支持快速原型设计和应用开发。

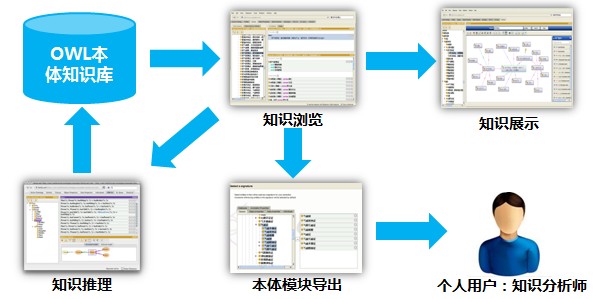


图 5‑2 基于Protégé对本体知识库进行浏览、展示、推理和模块导出

如 所示，下面介绍通过Protégé装载这一本体，进行本体的浏览、编辑、图形化展示、推理、抽取等工作的方法：

1. 下载和安装：从Protégé的官方网站（http://protege.stanford.edu/）上下载Protégé的安装程序（本项目使用的是Protégé 4.1版本），并按照Protégé的用户文档[[1]](#footnote-1)完成Protégé的安装。
2. 如 所示，打开Protégé，选择“Open OWL Ontology”，在弹出的文件选择器中选择中医脾系证候的OWL本体文件，即可将该本体载入Protégé中。

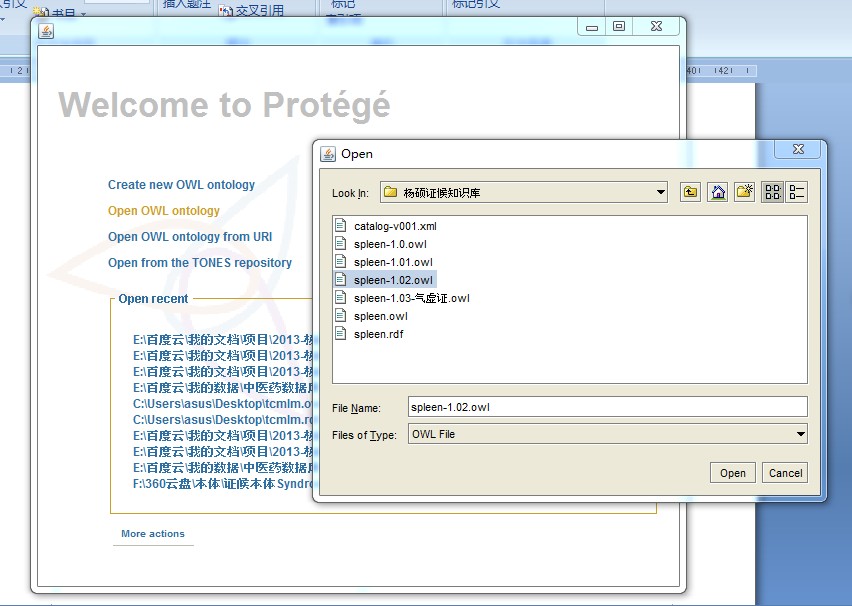


图 5‑3 用Protégé打开中医脾系证候的OWL本体

1. 如 所示，可使用Protégé对中医脾系证候的OWL本体进行浏览和编辑，具体方法详见Protégé的技术文档（http://protege.stanford.edu/doc/users.html）。

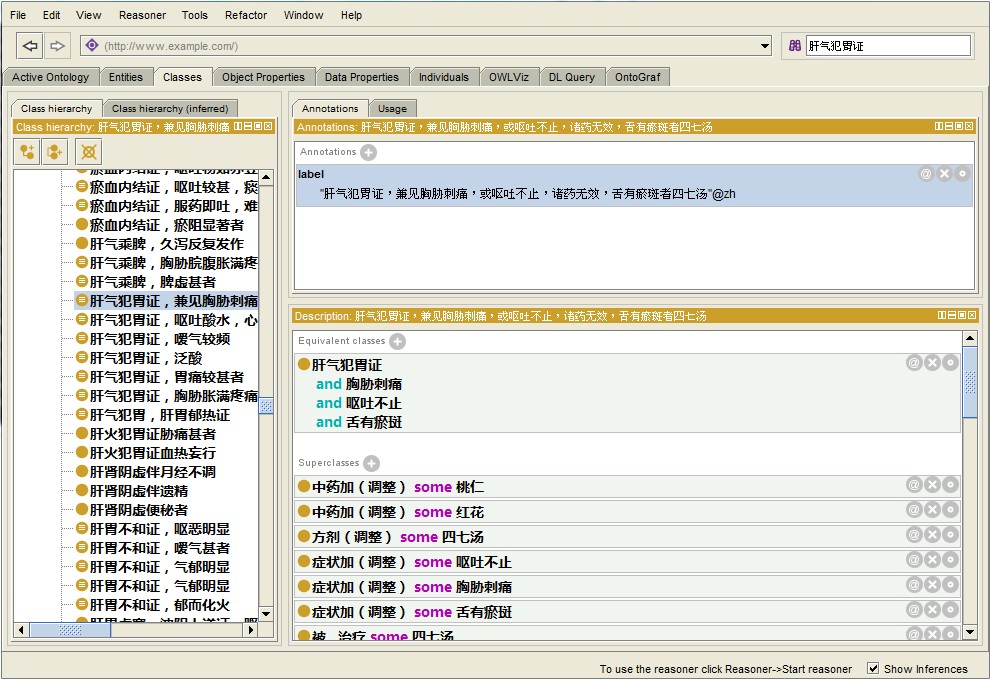


图 5‑4 Protégé界面截图（以“肝气犯胃证”为例展示）

1. 如 所示，使用OntoGraf插件对OWL本体进行图形化展示。OntoGraf是一款OWL本体的图形化展示工具，支持用户以交互的方式对OWL本体中的语义关系进行浏览。它提供多种自动布局方法，以通过合理的方式展示本体的结构。可展示的语义关系包括：父子类关系、等价关系、隶属关系等等。用户可手动调节图形布局，并对关系和节点类型进行过滤，以创建自己想要的本体视图。

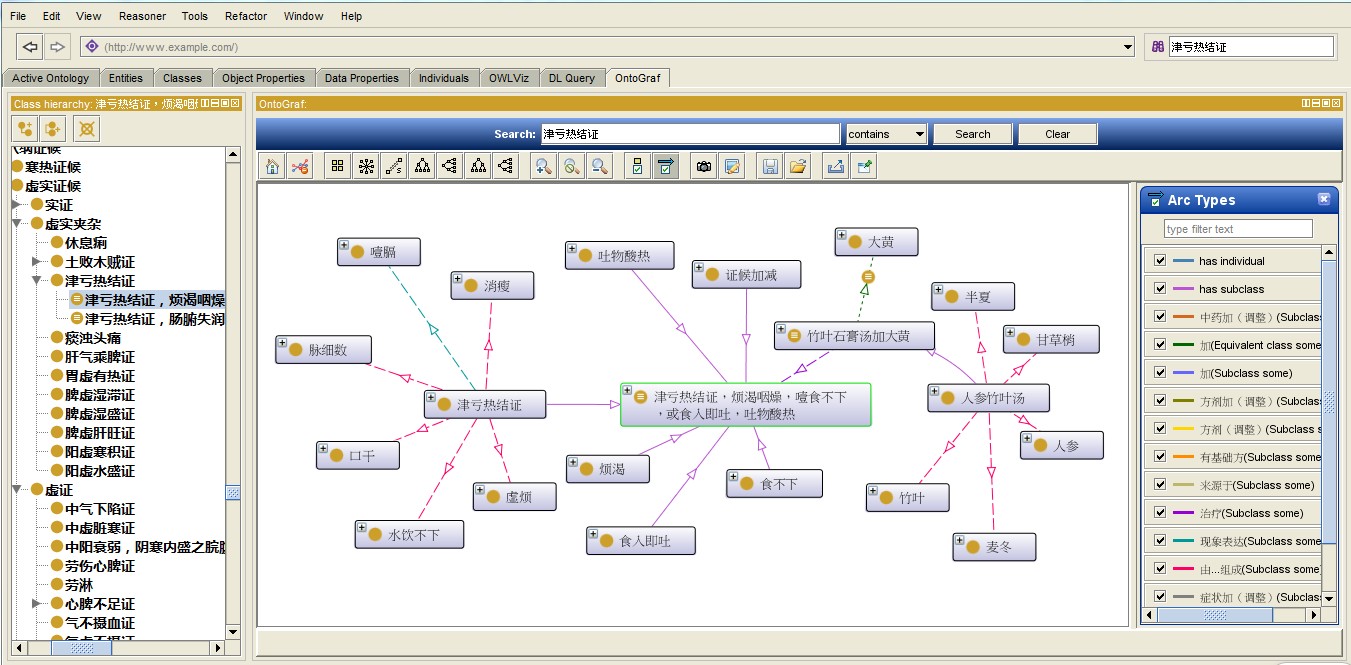


图 5‑5 OntoGrap界面截图（以中医脾系证候本体中的“津亏热结证”为例展示）

1. 可使用Protégé从本体中抽取出某个模块，以满足特定应用的需求。用户可选择一组实体（如气虚证及其直接的子类）（如图 5‑6 所示），与这些实体相关的本体内容会被自动筛选出来，打包为一个本体模块（如图 5‑7 所示），并导出为一个本体文件（如图 5‑8所示）。

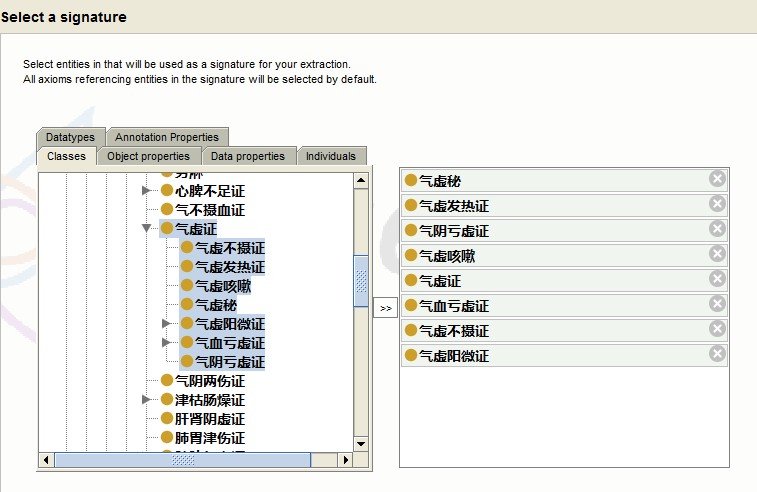


图 5‑6 Protégé的本体模块抽取界面（以“气虚证”为例）

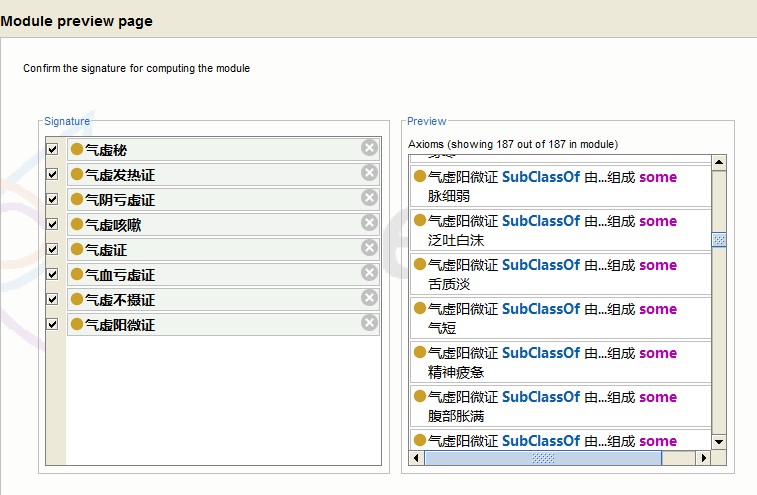


图 5‑7 Protégé 将实体相关陈述打包的界面（以“气虚证”为例）

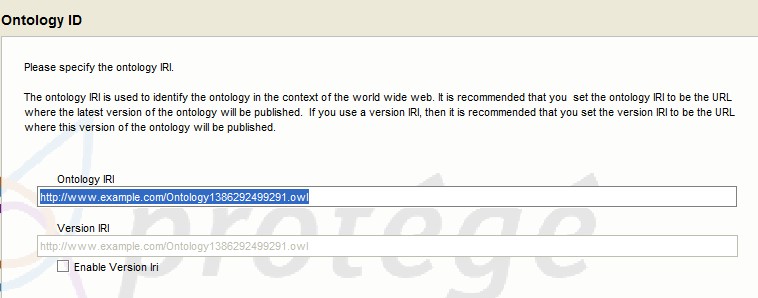


图 5‑8 Protégé 对抽出的本体模块命名的界面（以“气虚证”为例）

### 中医证候层次结构的构建方法

证候层次结构，也就是证型之间通过层次关系构成的层次结构。使用证候层次结构，可对纷繁复杂的病理变化及临床表现进行统一的分门别类。证候分类的基本原则被称为八纲。各种证候按照病变位置的深浅分为表证和里证；按照病性分为寒证和热证；按照邪正斗争的力量对比分为虚证和实证。可以将证候总归为阴阳两大类，其中里证、寒证和虚证属于阴证；表证、热证和实证属于阳证。除八纲之外，证候的分类方法还包括“精气血津液证候（如气、血、精、津液等）”、“脏腑证候（如胃、脾脏、肠道等）”和“经络证候（如足太阴脾经证、足阳明胃经证等）”等。

中医证候本体的核心内容，在于表达中医证候的层次结构。在OWL语言中，“类（owl:Class）”定义了一组具有共同特征的个体，这些个体被称为这个类的实例。可将每个中医证候定义为一个“类（owl:Class）”，并通过“rdfs:subClassOf”来定义证候之间的层次关系。例如，下面的OWL公式将“脏腑证候”定义为“证候”的子类：

|  |
| --- |
| **Class：脏腑证候**  **SubClassOf:**  **证候** |

可通过该方法进一步定义“精气血津液证候”、“脏腑证候”和“经络证候”等等。用OWL可声明一个类的多个父类。例如，下面的OWL公式将“脾虚气陷证”声明为“脾脏证候”、“气证”和“虚证”的共同之类：

|  |
| --- |
| **Class：脾虚气陷证**  **SubClassOf:**  **脾脏证候, 气证, 虚证** |

可以通过owl:disjointWith表达类之间的互斥关系：“A owl:disjointWith B”表示任意一个A的实例不可能同时是B的实例。例如：{寒证 owl:disjointWith 热证.}表明“寒证”和“热证”是一对互斥的范畴。OWL还可以声明一组类之间是两两互斥的：“DisjointClasses: 方剂 证候 脏腑”。

如 所示，该本体将“证候”分为如下4类：

（1）“八纲证候”，按寒热、虚实、表里、阴阳对证候分类；

（2）“精气血津液证候”，按气、血、精、津液对证候分类；

（3）“脏腑证候”，按胃、脾脏、肠道等对证候分类；

（4）经络证候，如足太阴脾经证、足阳明胃经证等。

接下来，可对每类证候进行进一步细分，如“八纲证候”可被细分为“阴证”、“阳证”、“虚证”、“实证”、“寒证”、“热证”、“表证”和“里证”；“足阳明胃经证”又可被细分为“协热下利证”和“阳明腑实证”等；以此类推。下面举几个对证候分类进行表达的例子：



图 5‑9证候分类结构

（1）气虚证及其子类

“气虚证”有多种分类的维度：从“八纲辨证”的角度分析，它属于“虚证”；从“精气血津液证候”的角度分析，它属于“气证”；从“脏腑证候”的角度分析，它属于“脾脏证候”。鉴于此，我们同时将“气虚证”定义为“气证”、“虚证”和“脾脏证候”这3个类的子类（围绕“气虚证”的证候结构图如 所示）：

|  |
| --- |
| **Class: 气虚证**  **SubClassOf:**  **气证，虚证，脾脏证候** |

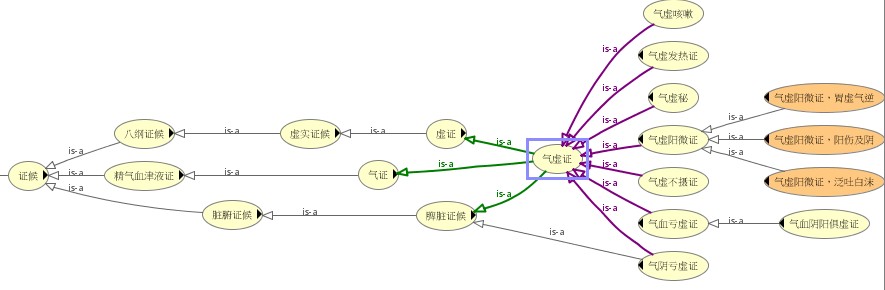


图 5‑10围绕“气虚证”的证候结构图

本体进一步定义了“气虚证”的子类，包括“气虚下陷证”、“气虚不摄证”、“气虚发热证”、“气虚咳嗽证”等等。例如，“气虚发热证”在八纲辨证体系中属于“热证”和“虚证”，在精气血精液辨证体系中属于“气证”，其定义如下（围绕“气虚证”的证候结构图如 所示）：

|  |
| --- |
| **Class: 气虚发热证**  **SubClassOf:**  **气证，热证，虚证** |

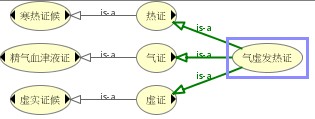


图 5‑11围绕“气虚发热证”的证候结构图

又如“气虚阳微证”在八纲辨证体系中属于“寒证”和“虚证”，在精气血精液辨证体系中属于“气证”；在脏腑辨证体系中属于“胃证候”（围绕“气虚阳微证”的证候结构图如 所示）：

|  |
| --- |
| **Class: 气虚阳微证**  **SubClassOf:**  **寒证，气证，虚证，胃证候** |

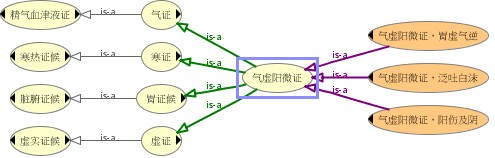


图 5‑12围绕“气虚阳微证”的证候结构图

本体在“气虚阳微证”之下，又列举了“气虚阳微证，胃虚气逆”，“气虚阳微证，泛吐白沫”和“气虚阳微证，阳伤及阴”等复杂的证候加减变化情况。我们将在第5.2.6节中对“证候加减”的建模方法进行详细讨论。

### 证候概念的定义及相关的推理方法

可以通过OWL对证候进行定义。例如，我们可将“阴证”理解为“与‘阴’有关的证候”，据此给出“阴证”的OWL定义：

|  |
| --- |
| **Class: 阴证**  **EquivalentTo:**  **证候 and （概念上相关 some 阴）** |

“阴虚证”是一种抽象的证型，指的是机体内阴的亏虚。它可被定义为“阴证”和“虚证”的交集：

|  |
| --- |
| **Class: 阴虚证**  **EquivalentTo:**  **阴证 and虚证** |

同理，可将“阳虚证”定义为“**阳证 and 虚证**”。

肾虚是位于肾的证候的统称。证候可以指定疾病的病位，即疾病的影响所发生的位置。例如，肾虚是影响肾中的正气（可以是气、血、阴和阳）的证候，这一定义可以表示如下：

|  |
| --- |
| **Class: 肾虚**  **EquivalentTo:**  **虚证 and （位于 some 肾） and （影响 some （位于 some 肾））** |

肾阳虚指出肾阳是亏虚的。其中，术语肾阳是肾中之阳的意思，它是机体中阳气的根源。我们可以将肾阳定义为“**阳 and （位于 some 肾）**”，并将肾阳虚表示如下：

|  |
| --- |
| **Class: 肾阳虚证**  **EquivalentTo:**  **虚证 and （影响 some 肾阳）** |

同理，可以将脾阳虚定义为“**虚证 and （位于 some 脾） and （影响 some 脾阳）**”。最后，可以将脾肾阳虚定义为“脾和肾中阳气的不足所造成的疾病”：

|  |
| --- |
| **Class: 脾肾阳虚证**  **EquivalentTo:**  **虚证 and （位于 some 脾） and （位于 some 肾）**  **and （影响 some （阳 and （位于 some 脾）））**  **and （影响 some （阳 and （位于 some 肾）））** |

另外，可以将多个证型组合起来，构成更加复杂的证型。例如，脾肾阳虚也可以被定义为脾阳虚和肾阳虚的交：“**肾阳虚 and 脾阳虚**”。

根据上述本体定义，一个推理机可以自动推理出脾肾阳虚、肾阳虚、脾阳虚、阳虚，肾虚和虚之间的隶属关系。推理也可根据OWL定义来判断证候概念之间的等价关系。例如，假设有多个本体设计师在一起协作开发本体。设计师甲录入了“肺脾气虚证”这一概念，并将其定义为：

|  |
| --- |
| **Class: 肺脾气虚证**  **EquivalentTo:**  **气虚证 and (位于 some 肺) and (位于 some 脾)** |

设计师乙则录入了“脾肺气虚证”这一概念，并将其定义为：

|  |
| --- |
| **Class: 脾肺气虚证**  **EquivalentTo:**  **气虚证 and (位于 some 脾) and (位于 some 肺)** |

因这两个定义实质上是等价的，推理机会据此发现“肺脾气虚证”和“脾肺气虚证”之间的等价性，并自动生成等价关系。

### 证候特征的表达及相关的推理方法

OWL可被用于表达证候的各种特征。例如，根据《中医内科学》，气虚证的症状包括“少气”、“懒言”、“无力”等。我们用“由…组成”来表达证候与症状之间的关系，并可通过**SubClassOf**和**some**表达证候的特征：

|  |
| --- |
| **Class: 气虚证**  **SubClassOf:**  **由…组成 some 乏力，**  **由…组成 some 少气，**  **由…组成 some 懒言，…** |

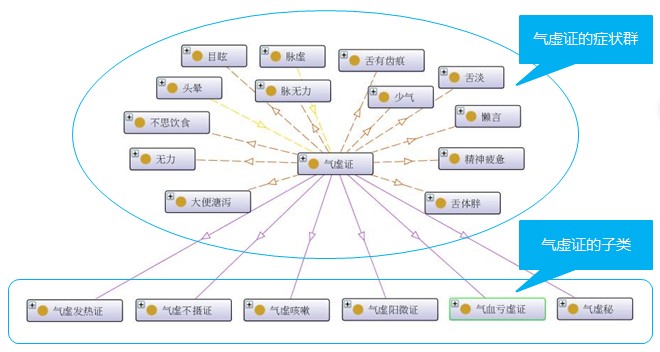


图 5‑13 围绕“气虚发热证”的概念图

如 所示，基于证候层次结构，可以实现证候特征的自动继承：若在本体中定义“气虚证”的症状为“少气”、“懒言”、“无力”等，则“气虚证”的子类（如“气虚发热证”）将自动继承这些症状，而无需重复声明。如 所示，在Protégé工具中启动推理机（如Fact++）即可实现“证候特征继承”推理。

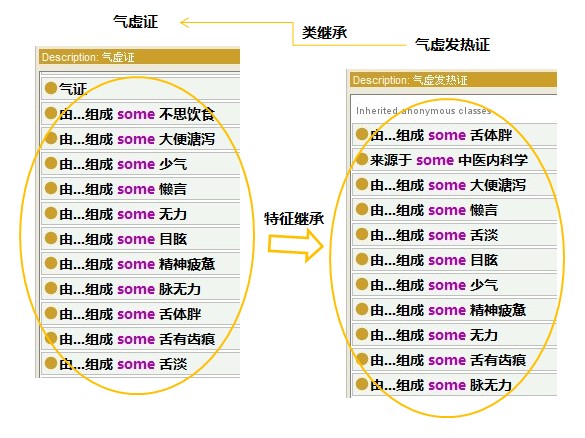


图 5‑14“气虚发热证”继承“气虚证”的特征

### 病证复合概念建模案例

从《中医内科学》的体例来看，中医内科目前采用的是一种基于“病证”复合概念的知识体系。《中医内科学》的各论即是按“病证”分章的，即分为肺系病证、心系病证、脾胃系病证、肝胆病证、肾系病证、气血津液病证、肢体经络病证等章。在每章中，各节基本上是围绕一种“疾病”论述，在“辨证论治”中的“证治分类”中，则引入各种“证候”的概念进行具体论述。例如，在“脾胃系病证”这一章的第一节讲的就是“胃痛”，在“证治分类”中论述“胃痛”可能分成“寒邪客胃证”、“饮食伤胃证”、“肝气犯胃证”等多种证候，并讨论每种“证候”的临床表现、证机概要、治法、代表方、常用药等。因此，中医内科学实质上是围绕“胃痛——寒邪客胃证”、“胃痛——肝气犯胃证”这类病证复合体进行阐述的。为准确表达中医内科学知识，需要对病证复合概念的内涵与外延进行准确表达。

（1）呕吐病、食滞内停证

根据教材《中医内科学》，在“呕吐病”之下的“食滞内停证”，其症状表述为“呕吐酸腐，脘腹胀满，嗳气厌食，大便或溏或结，舌苔厚腻，脉滑实”，代表方为“保和丸加减（该方以消食和胃为主，兼有理气降逆之功效，适用于饮食停滞，浊气上逆的呕吐）。

|  |
| --- |
| **Class: 呕吐病，食滞内停证**  **EquivalentTo:**  **呕吐病 and 食滞内停证,**  **厌食 and 呕吐酸腐 and 嗳气 and 脘腹胀满 and 苔厚腻 and (脉实 and 脉滑)**  **and (大便干结 or 大便溏)**  **SubClassOf:**  **被治疗 some 保和丸加减** |

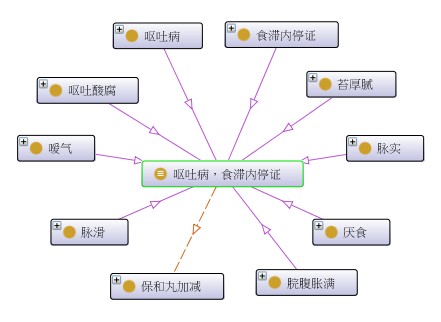


图 5‑15围绕“呕吐病，食滞内停证”的概念图

（2）呕吐病、胃阴不足证

根据《中医内科学》，“呕吐病”之下的“胃阴不足证”，其症状表述为：“呕吐反复发作，或时作干呕，似饥而不欲食，口燥咽干，舌红少津，脉象细数，其代表方为“麦门冬汤加减（该方滋阴养胃，适用于呕吐反复，或时作干呕的阴虚证）”。

|  |
| --- |
| **Class: 呕吐病，胃阴不足证**  **EquivalentTo:**  **呕吐病 and 胃阴不足证,**  **口燥 and 咽干 and 少津 and 时作干呕 and 舌红 and 饥不欲食 and**  **(脉数 and 脉细)**  **SubClassOf:**  **被治疗 some 麦门冬汤** |

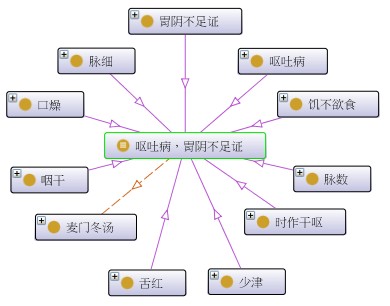


图 5‑16围绕“呕吐病，胃阴不足证”的概念图

（3）噎膈病、痰气交阻证

根据《中医内科学》，“噎膈病”之下的“痰气交阻证”，其症状表述为“吞咽梗阻，胸膈痞满，甚则疼痛，情志舒畅时稍可减轻，情志抑郁时则加重，嗳气呃逆，呕吐痰涎， 口干咽燥，大便艰涩，舌质红，苔薄腻，脉弦滑；其代表方为“启膈散加减”。

|  |
| --- |
| **Class: 噎膈病，痰气交阻证**  **EquivalentTo:**  **噎膈病 and 痰气交阻证,**  **口干 and 吞咽梗阻 and 呃逆 and 呕吐痰涎 and 咽燥 and 嗳气 and 大便艰涩 and**  **舌质红 and 苔薄腻 and (情志抑郁时加重 and 情志舒畅时可减轻) and (脉弦 and 脉滑)**  **and (胸膈疼痛 or 胸膈痞满)**  **SubClassOf:**  **被治疗 some 启膈散加减** |

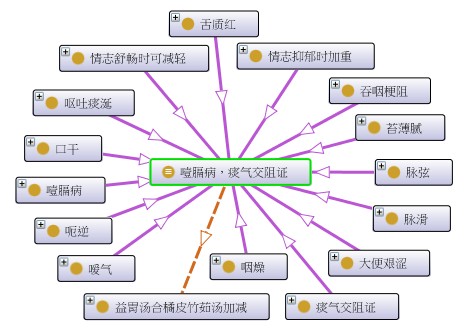


图 5‑17围绕“噎膈病，痰气交阻证”的概念图

（4）噎膈病、瘀血内结证

根据《中医内科学》，“噎膈病”之下的“瘀血内结证”，其症状表述为“饮食难下，或虽下而复吐出，甚或呕出物如赤豆汁，胸膈疼痛，固着不移，肌肤枯燥，形体消瘦，舌质紫暗，脉细涩”，代表方为“通幽汤加减”。

|  |
| --- |
| **Class: 噎膈病，瘀血内结证**  **EquivalentTo:**  **噎膈 and 瘀血内结证,**  **形体消瘦 and 肌肤枯燥 and 舌质紫暗 and (固着不移 and 胸膈疼痛) and**  **(脉涩 and 脉细) and (呕出物如赤豆汁 or 食下而复吐出 or 饮食难下)**  **SubClassOf:**  **被治疗 some 通幽汤加减** |

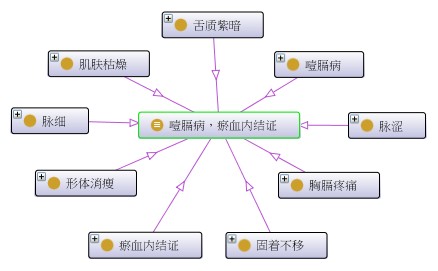


图 5‑18围绕“噎膈病，瘀血内结证”的概念图

（5）呃逆病、胃阴不足证

根据《中医内科学》，“呃逆病”之下的“胃阴不足证”呃声短促而不得续，口干咽燥，烦躁不安，不思饮食，或食后饱胀，大便干结，舌质红，苔少而干，脉细数。

代表方：益胃汤合橘皮竹茹汤加减。

### 证候加减的建模方法

在中医临床领域，同一个证候会有多种变化，所使用的方药治疗方案也要酌情加减变化。例如，根据《中医内科学》（pp. 190），“胃痛”之中的“肝气犯胃证”有多种加减变化：“如胃痛较甚者，可加川楝子、延胡索以加强理气止痛；嗳气较频者，可加沉香、旋覆花以顺气降逆；泛酸者加乌贼骨、煅瓦楞子中和胃酸……”。又如，根据《中医内科学》（pp. 216），“噎嗝”之中的“津亏热结证”也有多种加减变化：“胃火偏盛者，加栀子、黄连清胃中之火；烦渴咽燥，噎食不下，或食入即吐，吐物酸热者，改用竹叶石膏汤加大黄泻热存阴”。在上述的知识表述中，涉及到“肝气犯胃证，胃痛较甚”、“肝气犯胃证，嗳气较频”、“津亏热结证，胃火偏盛”、“津亏热结证，烦渴咽燥，噎食不下，或食入即吐，吐物酸热”等复合概念，我们在此称之为“证候加减”。另外，其中还涉及到“方剂加减”的概念，如“竹叶石膏汤加大黄”。

在本节中，我们结合具体案例探讨如何使用OWL语言来定义“证候加减”和“方剂加减”的概念，并表达证候加减变化的知识。

（1）“津亏热结证：烦渴咽燥，噎食不下，或食入即吐，吐物酸热者，改用竹叶石膏汤加大黄”的形式化表达

知识表达的基本思路是：首先定义“津亏热结证”、“咽干”、“烦渴”、“吐物酸热”、“食入即吐”、“竹叶石膏汤”、“大黄”等概念（相关的概念图如所示），再定义“津亏热结证，烦渴咽燥，噎食不下，或食入即吐，吐物酸热”这一证候加减概念以及“竹叶石膏汤加大黄”这一方剂加减概念，最后建立这两者之间的“治疗关系”。

为定义“证候加减”这一符合概念，首先要定义，再用“**and**”、“**or**”和“**not**”等连接词将它们逐步连接起来：（1）“烦渴咽燥”可被表达为“**咽干 and 烦渴**”；（2）“噎食不下”可被表达为“**not 进食**”；（3）“食入即吐，吐物酸热”可被表达为“**吐物酸热 and 食入即吐**”；（4）“烦渴咽燥，噎食不下，或食入即吐，吐物酸热”可被表达为“**(咽干 and 烦渴) and ((not 进食) or (吐物酸热 and 食入即吐))**”；（5）将“津亏热结证”与“烦渴咽燥，噎食不下，或食入即吐，吐物酸热”连接起来。该“证候加减”的最终定义如下：

|  |
| --- |
| **Class: 津亏热结证，烦渴咽燥，噎食不下，或食入即吐，吐物酸热**  **EquivalentTo:**  **津亏热结证 and ((咽干 and 烦渴) and ((not 进食) or (吐物酸热 and 食入即吐)))** |

为定义“竹叶石膏汤加大黄”这一方剂加减概念，首先要定义方剂“竹叶石膏汤”和中药“大黄”，其次定义“加”这一对象属性，再用“**and**”和“**some**”等连接词将方剂和中药连接起来：

|  |
| --- |
| **Class: 竹叶石膏汤加大黄**  **EquivalentTo:**  **(竹叶石膏汤 and (加 some 大黄))** |

最后，我们在“津亏热结证，烦渴咽燥，噎食不下，或食入即吐，吐物酸热”和“竹叶石膏汤加大黄”之间建立“被治疗”，将“津亏热结证：烦渴咽燥，噎食不下，或食入即吐，吐物酸热者，改用竹叶石膏汤加大黄”定义为：

|  |
| --- |
| **Class: 津亏热结证，烦渴咽燥，噎食不下，或食入即吐，吐物酸热**  **EquivalentTo:**  **津亏热结证 and ((咽干 and 烦渴) and ((not 进食)**  **or (吐物酸热 and 食入即吐)))**  **SubClassOf:**  **被治疗 some 竹叶石膏汤加大黄** |

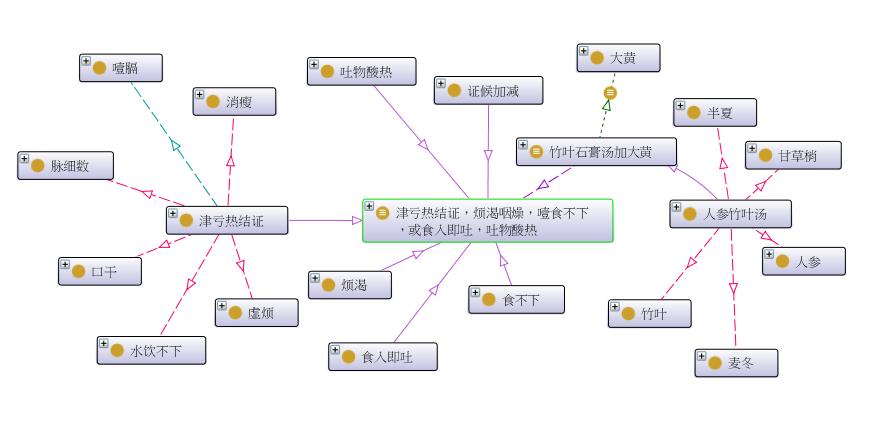


图 5‑19 围绕“津亏热结证，烦渴咽燥，噎食不下，或食入即吐，吐物酸热”的概念图

## 综合性知识展示系统

在本项目中，我们采用国际主流的Web技术，在中医脾系知识库的基础上，开发了知识展示系统。如 所示，该系统实现了四项主要功能：（1）基于关键词的知识检索；（2）知识浏览与导航功能；（3）语义网络的检索和浏览功能；（4）简单的知识问答功能。



图 5‑20知识展示系统的主要功能

### 基于关键词的知识检索

系统提供了基于关键词的知识检索功能。如图 5‑21 所示，用户可输入关键词——“肝气犯胃证”，系统将会提供“肝气犯胃证”等概念的基本信息，并推荐“肝气犯胃证”的相关链接。系统提供了“肝气犯胃证”、“胃痛”等概念的链接：当用户点击这些概念的链接时，系统将引导到这些概念的相关页面。

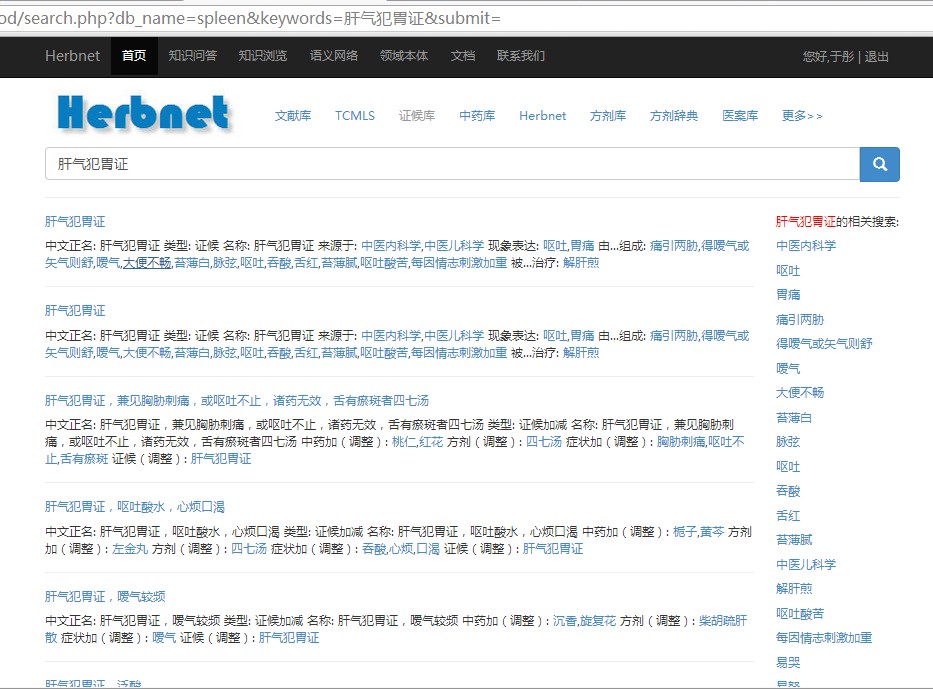


图 5‑21 基于关键词的知识检索功能

### 知识浏览与导航功能

系统以Web的方式向网络用户提供知识浏览与导航功能。系统以概念为单位对中医脾系证候知识库进行展示，它提供概念信息的展示功能，以及基于概念语义关系的知识导航功能。例如，系统对“脾胃阴虚证”信息的展示界面如 所示，该界面提供了“脾胃阴虚证”的中文正名、类型、来源等信息，基于“现象表达”关系导航到该证候相关的疾病，并基于“由…组成”关系导航到该证候相关的“症状”。



图 5‑22 以“概念”为单位的知识浏览与导航功能（以“脾胃阴虚证”为例）

如 所示，系统还提供了一个简单的分类导航界面。在该界面中，左侧为一个简单的类型列表，提供了语义类型的名称及其实例数量；当用户选择某一类型时，系统列出该类型的实例。当用户点击某一概念的链接时，系统会导航到该概念的展示界面。

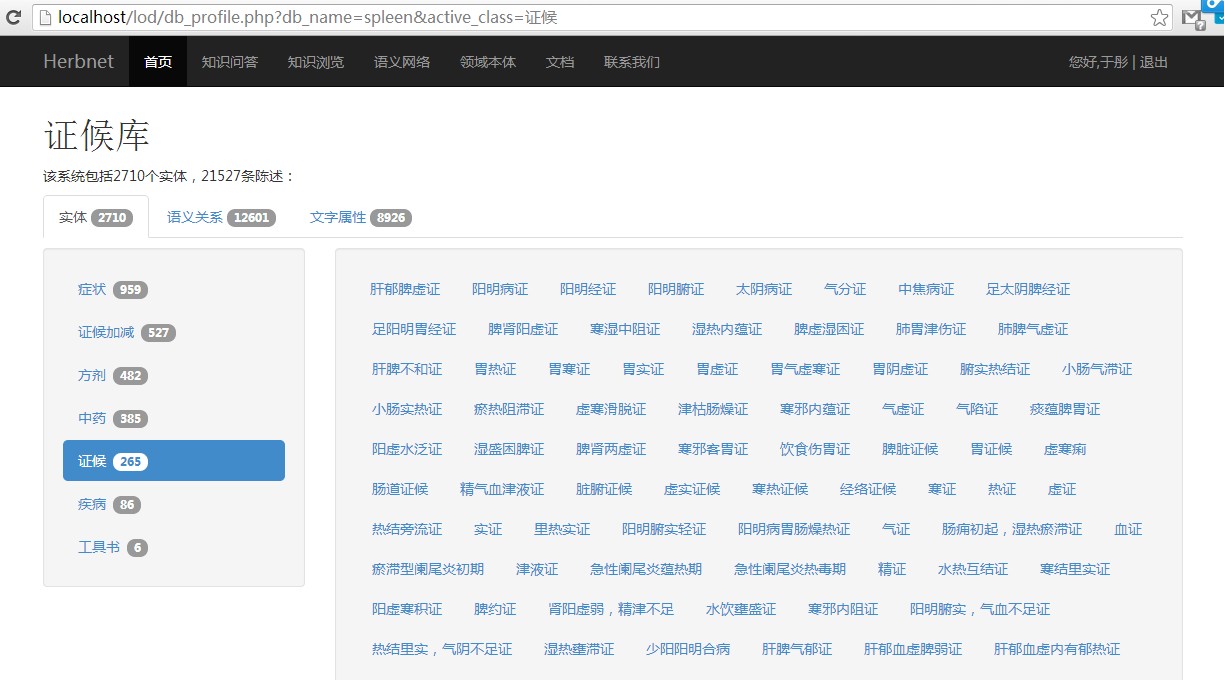


图 5‑23 简单的分类导航界面

如图 5‑24 所示，系统还提供了知识库中语义关系的导航界面。在界面的左侧，列出了知识库中的语义关系（如“由…组成”、“来源于”、“上位词”、“下位词”等）以供用户选择。比如，当用户选择“由…组成”时，系统则分页显示知识库中所有的“由…组成”关系。用户还可以输入关键词对语义关系进行筛选。例如，当用户输入“香砂六君子汤”时，系统会列出所有包含“香砂六君子汤”的“由…组成”关系，用户即可了解到“香砂六君子汤”是由哪些中药组成的。



图 5‑24 语义关系的导航界面（以“香砂六君子汤”的组成为例）

又如 所示，若用户想了解“哪些药物治疗肝火犯胃证”，可在语义关系列表中选择“治疗”，并在搜索框中输入“肝火犯胃证”，系统则列出与“肝火犯胃证”有“治疗”关系的中药或方剂。



图 5‑25 语义关系的导航界面（以“哪些药物治疗肝火犯胃证”为例）

再如 所示，若用户选择“现象表达”关系，并在搜索框中输入“肠道湿热证”，则可了解“与肠道湿热证”相关的疾病（如“便血”）。



图 5‑26 语义关系的导航界面（以“肠道湿热证的现象表达”为例）

中则显示出了与“胃热炽盛证”有“现象表达”关系的实体（如“消渴”等）。



图 5‑27 语义关系的导航界面（以“胃热炽盛证的现象表达”为例）

### 顶层语义网络的浏览功能

顶层语义网络是由语义类型所构成的关系网络。如 所示，我们开发了一个用于浏览顶层语义网络的界面，它列出“证候库”中的语义类型（包括症状、证候加减、方剂、中药、证候、疾病和工具书）供用户选择；当用户选择某一类型（如“证候”）时，系统会列出相关的语义关系（如“证候 由…组成 症状”）。



图 5‑28 顶层语义网络浏览界面

如 所示，用户可点击某一语义关系，进一步查看该关系的实例。例如，“证候 由…组成 症状”的实例包括“脾肺气虚证 由…组成 气短”、“脾肺气虚证 由…组成 腹胀”等。



图 5‑29 顶层语义网络浏览界面（以“证候 由…组成 症状”为例）

系统还提供了另一种语义网络展示工具，被称为“语义关系搜索”。用户可以任选主体、谓词和客体的类型，系统则将符合条件的语义关系分门别类地展示出来。如 所示，用户可选择“证候”作为主体，查看相关的语义关系。系统则显示出“证候 由…组成 症状”、“证候 来源于 工具书”、“证候 现象表达 疾病”等关系。对于每条顶层语义关系，系统都给出了一些实例供用户参考。例如，“证候 由…组成 症状”的例子包括“肺脾气虚证 由…组成 气短”,“肺脾气虚证 由…组成 倦怠”等。又如，若用户若想了解“哪些类型的事物可以治疗证候”，则可选择主体为“任意事物”，谓词为“治疗”，客体为“证候”，系统会显示“方剂 治疗 证候”等关系。

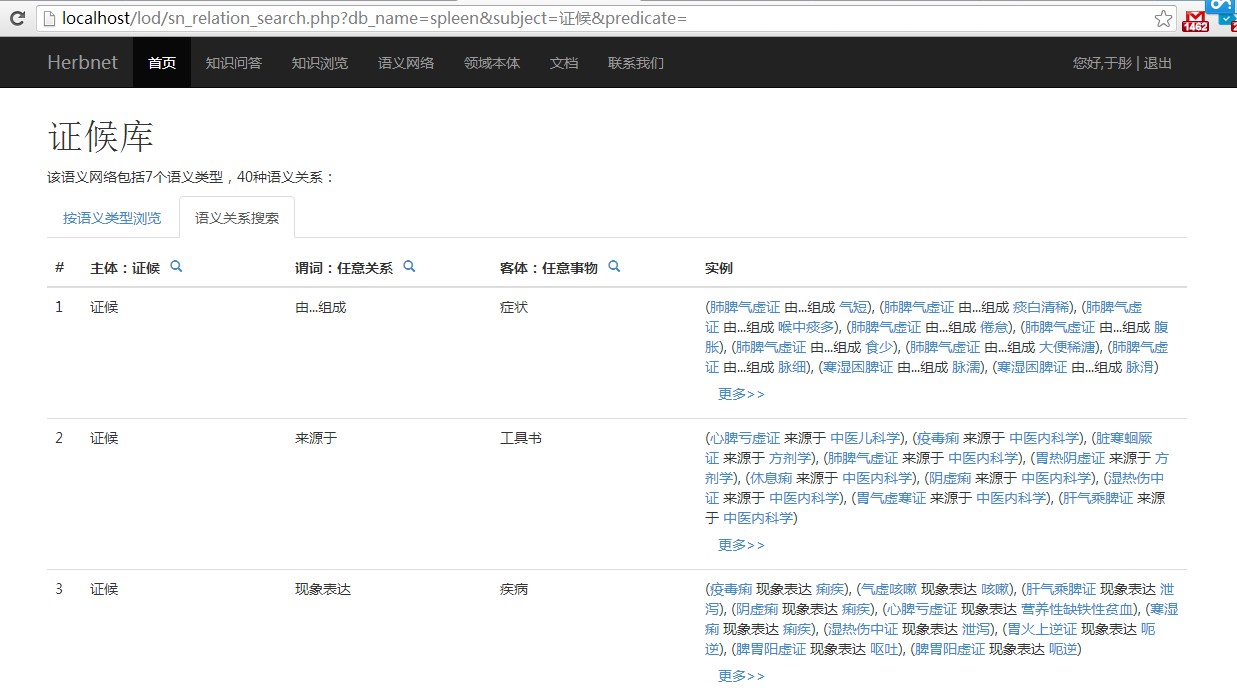


图 5‑30 语义关系搜索界面（以“证候”为例）

用户若想查看更多的实例，可点击“查看更多>>”，则系统将转到如图 5 3 所示的语义关系搜索界面（以“证候 由…组成 症状”为例展示）。

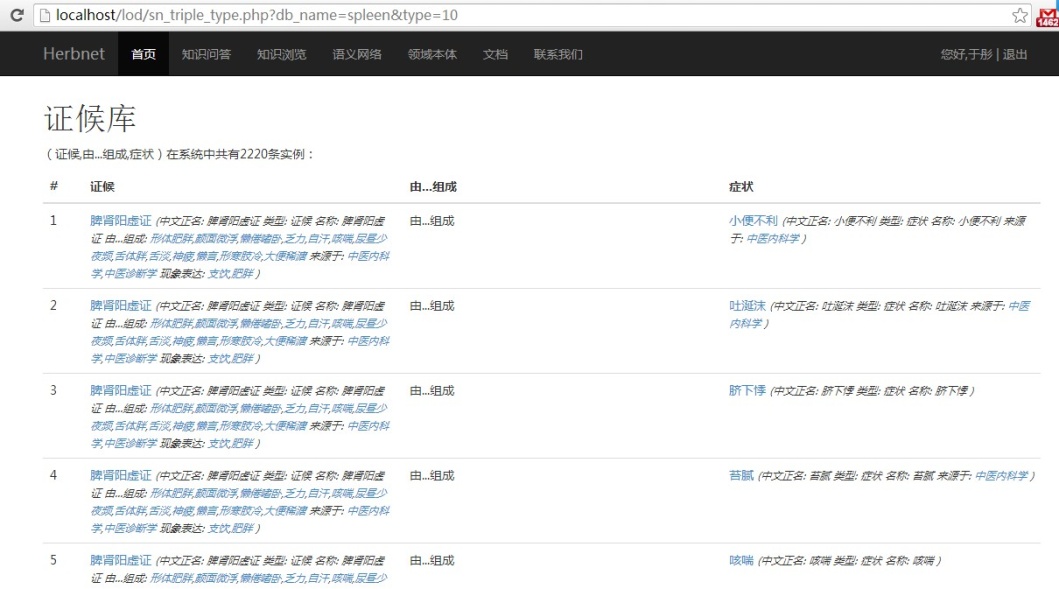


图 5‑31 语义关系搜索界面（以“证候 由…组成 症状”为例）

### 简单的中医药知识问答功能

问答系统是知识库系统的一般交互方式，也是展示推理能力的一种常用方式。问答系统的一个典型案例是Wolfram|Alpha，它是一部提供知识的引擎，能利用所存储的大量专业知识，完成问答、分析和报告生成等任务。Wolfram|Alpha与传统的搜索引擎不同，它提供的是准确的知识，而非大量的搜索结果。在本项目中，我们借鉴Wolfram|Alpha，构建了一个中医问答系统的雏形，提供根据症状辨别证候，以及推荐处方等常见的问答形式。

图 5‑1 中医问答系统示意图

如上图所示，中医问答提供类似于搜索的界面，但“更加智能”：它不是提供成千上万的搜索结果，而是提供问题的答案和准确的知识。用户可与系统进行交互性的问答，如：

|  |
| --- |
| Q：患者具有“郁怒”的症状。  A: 推荐意见：  辨证: 患者有证候肝气郁结证、肝气不运证、肝郁脾虚证、血瘀证。  立法: 患者的治法为疏肝补脾法。  处方: 患者的处方为柴胡疏肝汤。  Q: 你是如何推理出来的？  A: 我的推理过程和依据是…… |

目前，中医问答原型系统目前实现了3种用户场景：

（1）中医辨证场景

该场景如 所示：在用户可输入一组症状，如口干,唇干,干呕,呃逆,面色潮红,皮肤干燥,大便干结等。系统从知识库中找出最为匹配的证候（如脾胃阴虚证），并列出相关证候的简要信息。



图 5‑32 中医问答原型系统之辨证场景

（2）中医证候信息问答场景

该场景如 所示，用户输入证候（如“肝气犯胃证”），系统根据知识库向用户推荐合适的方剂。在方剂推荐中，系统可根据知识库处理“肝气犯胃证兼有胃痛”等各种证候加减变化的情况。系统还根据知识库向用户推荐与该证候相关的疾病。



图 5‑33 中医问答原型系统之证候信息问答场景（以“肝气犯胃证”为例）

“脾胃虚寒证”的信息问答场景如 所示。当用户输入证候“脾胃虚寒证”之后，系统会向用户推荐“小建中汤（来源于教材《方剂学》）”、“丁萸理中汤（来源于教材《中医儿科学》）”等方剂。系统还试图基于简单的间接关系推理来找出推荐方剂。例如，系统可根据“小建中汤加党参白术 治疗 腹痛”和“脾胃虚寒证 现象表达 腹痛”，推测出“[小建中汤加党参白术](http://localhost/lod/entity.php?id=1679&db_name=spleen)”治疗“[脾胃虚寒证](http://localhost/lod/entity.php?id=2108&db_name=spleen)”，并将“[小建中汤加党参白术](http://localhost/lod/entity.php?id=1679&db_name=spleen)”作为推荐方剂提供给用户。用户点击“[小建中汤加党参白术](http://localhost/lod/entity.php?id=1679&db_name=spleen)”信息右侧的“推理依据”，即可查看系统推荐意见背后的推理依据。

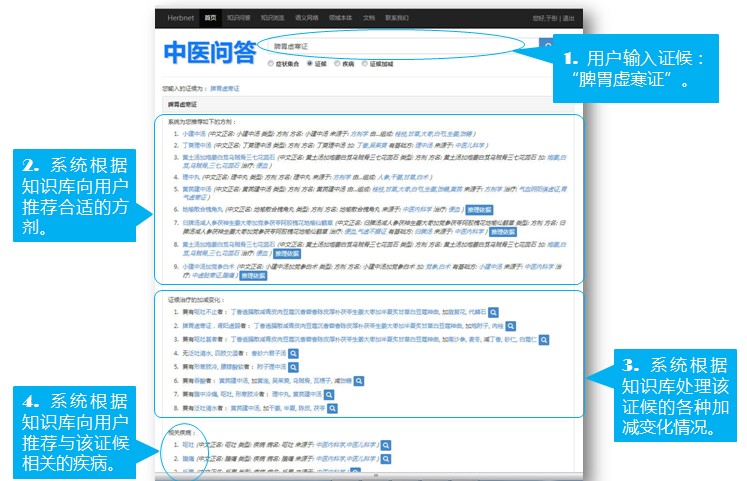


图 5‑34 中医问答原型系统之证候信息问答场景（以“脾胃虚寒证”为例）

（3）中医疾病信息问答场景

如 所示，用户可输入疾病名称（如痰饮），系统基于知识库向用户推荐方剂（如苓桂术甘汤），并列出相关证候的简要信息。



图 5‑35 中医问答原型系统之疾病信息问答场景（以“痰饮”为例）

# 项目计划

对项目开发工作作出如下安排：

表 6.1 项目计划

|  |  |  |
| --- | --- | --- |
| 时间 | 任务 | 说明 |
| 9月 | 需求分析&项目规划 | 1. 需求分析文档。 2. 制定项目规划。 |
| 10月 | 预研工作 | 1. 采用Protégé工具，构建多个“小本体”，设计并实现一些中医证候领域的推理案例 2. 为满足推理实验和知识展示的需要，进行必要的数据预处理和转换工作 3. 构建中医知识展示平台的原型系统 4. 中期总结会（报告推理实验结果，解释技术方案，演示原型系统，听取反馈意见） |
| 11月 | 系统研发 | 1. 将中医脾系证候知识库转换为OWL格式， 2. 将OWL本体载入Protégé中，开展本体建模与推理研究。 3. 研发中医药知识展示平台。 4. 系统部署 5. 成果汇报会（系统演示，听取反馈意见） |
| 12月 | 系统部署与修改，文档撰写 | 1. 完成系统部署、测评和改进工作 2. 撰写并提交技术报告和用户手册 3. 撰写论文 4. 项目验收会（待定） |

# 讨论

目前，在中医界尚未出现一个严格意义上的OWL本体知识库。一个严格意义上的本体知识库应至少：（1）对某个有限领域的知识体系进行系统表达；（2）支持推理，即从已有知识中推出隐含知识；（3）支持知识查询，用户可以通过知识查询功能来查询这个知识库，辅助个人进行决策。中医界已经发表了一些OWL本体知识库方面的工作，但所开发的知识库一般都不支持推理，或仅支持简单的推理案例。本项目将有望产生一个严格意义上的OWL中医本体知识库，起到示范作用。

## 本项目的贡献

本项目的贡献有如下3点：

（1）对中医脾系证候的知识体系进行了系统的梳理和表达。如 所示，项目中构建的知识库已录入概念2710个，其中证候有265个，证候加减527个，疾病86个，方剂482个，中药385个，症状959个；在这些概念之间建立了10471条语义关系，分为30种。中医脾系证候本体在表达能力和表达的丰富性方面都强于TCMLS。它采用了OWL 2的新特性，对于中医证候学中的一些难以用原有的语义网络表述的复杂知识进行了合理的建模；它还增加了对“证候加减”和“方剂加减”的处理，涉及到了更为复杂的知识内容；并且，它对于现代文献搜索检阅赋予了语义特征，提供发散性思维的关联搜索模式。

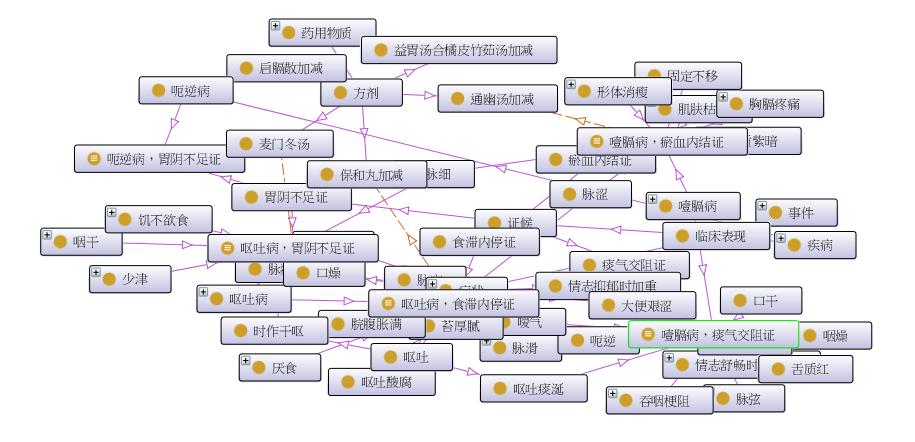


图 7‑1 中医证候领域的概念图

（2）开发了中医脾系证候知识展示系统，提供了丰富多样的知识展示方式。该系统支持如下4方面的功能：（a）基于关键词的知识检索，向用户提供与关键词相匹配的概念的简要信息和链接，以及相关概念列表；（b）以概念为单位的知识浏览，向用户提供概念信息以及相关概念的跳转功能；（c）顶层语义网络的展示和浏览功能，支持语义关系的搜索和浏览；（d）简单的知识问答功能，支持辨证问答、证候信息问答和疾病信息问答，为临床决策提供支持。



图 7‑2 中医脾系证候知识展示平台界面截图

（3）使用OWL和Protégé等国际先进技术，探索出了一种中医药知识工程和服务的新模式。本项目采用的模式有3个要点：第一、构建基于Web的知识加工系统，支持多名领域专家进行协作式的知识获取；第二、构建基于Web的知识展示系统，向领域专家提供随时、随地可访问的知识服务；第三、将知识库内容（或其中的部分内容）载入单机版的Protégé本体编辑系统，支持领域专家个人使用，进行知识的图形化展示和推理实验。这套模式对于TCMLS等其他系统的知识加工与服务具有借鉴意义。



图 7‑3 中医药知识工程和服务模式

## 存在的问题与未来的工作

目前中医脾系证候知识库的内容尚不充足。在未来的工作中，拟进一步开展知识加工工作，增加病因、病机、功效等概念的表达，并集成TCMLS和中医个案库等系统中的相关数据，以丰富和补充知识库内容，满足机器学习、临床决策支持和推理的需要。

拟进一步研究中医证候知识的表达方法。首先，拟进一步斟酌中医药领域语义关系的用法。如“方剂”和“中药”的关系是否为“由…组成”（如“保元汤 由…组成 肉桂”）？“证候”与“疾病”的关系是否为“现象表达”？其次，拟研究如何表达“中阳衰弱，阴寒内盛之脘腹剧痛证”这样的复杂概念（将其表达为证候吗？）最后，拟研究概念的归类问题，如“湿热痢疾”应归入“证候”还是“疾病”？

拟进一步深入研究推理方法，从而发现知识库中蕴含的潜在知识。拟研究如何通过推理实现中医领域的概念的自动关联；比如在症状的“呕吐”中，将“湿滞脾胃证”和“平胃散”关联起来。另外，基于OWL 2的推理属于确定性推理，其基础都在于领域知识和规则的确定性，因此它只能对中医知识体系中确定的部分进行建模和推理。但中医知识体系和推理模式也具有不确定性，这是OWL 2目前无法支持的。如何将不确定性推理方法与OWL 2相结合，是未来需要解决的一个问题。

拟采用Jena作为语义Web的编程框架，进一步研发面向中医脾系证候知识库的语义技术。Jena用于将RDF文档或数据流解析为Java程序能够处理的数据结构，用来支持语义Web。Jena框架包括如下的组件：（1）RDF API，支持RDF和三元组文件的读写；（2）OWL API；（3）SPARQL查询引擎；（4）推理机：包括传递性推理机（Transitive reasoner）、RDFs规则推理机（RDFS rule reasoner）、Web本体推理机（包括OWL，OWL Mini，OWL Micro，DAML等语言的推理机）、通用规则推理机（Generic rule reasoner），可支持多种推理需求。拟采用Jena，进一步研发专门面向中医药领域的推理方法，以及将中医脾系证候知识库与其它知识库相互融合的方法等。

拟进一步完善知识展示系统，实现别名检索，扩展检索等功能。实现基于知识库的语义查询处理。用户通过一个查询界面输入语义查询（即SPARQL语句），系统对查询进行解析和执行，对所有的查询结果进行语义封装，构成综合性的结果集，在用户界面上展示出来。

拟利用TCMLS中的中英文对应关系，或Google翻译等在线翻译系统，为证候库中的概念添加英文标签，使之升级为一个国际化的中英文双语本体；并将一些典型的案例翻译成英文，以供撰写英文论文。

# 相关网站

* [Wolfram|Alpha: Computational Knowledge Engine](http://www.baidu.com/link?url=9-zWOw7TwWJcSK4VEKlSF2du4a73-BdGP1vF2IShXTfqbtSfliqXwqil5E_FHdju)
* Virtuoso: http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/VOSUbuntuNotes

# 参考文献

1. 路耀华．思维模拟与知识工程[M]．清华大学出版社, 1997.
2. 马斌荣. 中医专家系统与中医知识库-中医领域计算机软件的开发与应用[M]. 北京出版社, 1998.
3. 徐彬锋,温志浩,罗小刚等.基于本体的医学知识库构建及应用[J].北京生物医学工程,2011,30(6):618-623.
4. 臧知明.《黄帝内经》养生知识库框架构建研究[D]. 广西中医学院,2005.
5. 谷建军.基于叙词表的中医古籍文献领域本体建模方法研究[D].中国中医科学院,2006.
6. 车立娟,王瑾德,周强,高成勉,高红艳,邢斌. 基于“肺阴虚证”本体的中医证候知识库构建方法研究[J]. 上海中医药大学学报,2009,23(4):18-20.
7. 李新霞.基于本体的中医学脾胃病知识库的构建[D].南京理工大学,2008.
8. 孙海舒,符永驰,张华敏等. 基于本体论构建中医古籍知识库的探索[J]. 医学信息学杂志,2011,32(3):64-68.
9. 彭笑艳.基于中医医案的知识库构建[D].北京科技大学,2009.
10. 曹宇峰.病案分析引导下的中医智能教学的研究[D].首都师范大学,2005.
11. 王连心,孟庆刚,王志国等.中药知识库设计浅析[J].世界中医药,2011,06(6):535-537.
12. 胡滨.中医药文献检索[M].上海:上海科技出版社, 2007:33-45.
13. 赵静,庄天戈,刘红菊等.基于语义网络方法的三维可视人中医知识库[J].上海交通大学学报,2005,39(4):517-521,526.
14. 郑雷.结合中医针灸的三维医学影像学研究[D].上海交通大学,2004.
15. 吴芸,周昌乐,张志枫等.中医舌诊八纲辨证神经网络知识库构建[J].计算机应用研究,2006,23(6):188-189,198.
16. 周金海,杨涛,沈大庆等.基于ANN的中医舌诊八纲辨证知识库构建与应用[J].计算机应用研究,2010,27(5):1771-1772,1790.
17. 陈国宁,陈秋莲,李陶深等.一个中医咳感症诊断专家系统的设计[J].广西大学学报(自然科学版),2001,26(2):101-104.
18. 易钢,罗尧岳.基于本体的中医知识库系统的研究[J].医学信息（上旬刊）,2010,23(10):3516-3518.
19. 易钢.应用Protégé构建中医药学本体方法研究[J].电脑知识与技术,2012,08(1):223-225.
20. 杨继红.中医古籍数字化资源建设概述[J].现代情报,2008,28(5):136-138.
21. Gruber TR. Ontology. 2008. Entry in the Encyclopedia of Database Systems, Ling Liu and M. Tamer Özsu (Eds.), Springer-Verlag.

1. <http://protegewiki.stanford.edu/wiki/Protege4GettingStarted> [↑](#footnote-ref-1)