

# 基于 Web 的可拓专家系统研究

张海涛, 董春游

(黑龙江科技学院 计算机与信息工程学院, 黑龙江 哈尔滨 150027)

**摘 要:**传统的专家系统的局限性使得构造基于 Web 的专家系统具有重要意义. 可拓学的发展为之提供了新的方法与手段. 首先明确提出可拓专家系统的概念, 然后结合当今流行的 Web 应用技术, 构建基于 B/S 程序架构、Web 的可拓专家系统, 使之具有高移植性、高扩展性, 且可以发挥可拓规则的扩展性和动态性的特点, 对实现系统的关键技术进行了说明, 系统运行良好, 证明了所阐述理论之可行性.

**关键词:**专家系统; 可拓学; 动态 Web; Java; 可拓规则

**中图分类号:** TP31 **文献标识码:** A **文章编号:** 1673-4785 (2009) 02-0175-05

## An extendable expert system based on the Web

ZHANG Hai-tao, DONG Chun-you

(College of Computer and Engineering, Heilongjiang Institute of Science and Technology, Harbin 150027, China)

**Abstract:** The limitations of traditional expert systems inspired us to study expert systems based on the WEB. The development of Extenics provides new methods and tools for this. Our concepts for the extendable expert system are presented in this paper. Using the latest Web application methodology, we constructed an extendable expert system based on the B/S program framework and the Web. It has high transplantability and extendability. The key technological requirements for system implementation were presented. The smooth operation of the system proved the feasibility of the proposed method.

**Keywords:** expert system; Extenics; dynamic Web system; Java; extendable rules

近年来可拓学的发展产生了很多研究成果, 将这些研究成果应用于专家系统中, 是专家系统发展的一个方向. 但如何将专家系统进一步与可拓学结合, 并移植到网上, 却是一个庞大而复杂的问题. 本文在此背景下依托前人的一些研究成果, 研究基于网络的可拓专家系统, 可以使数据库、知识库得到无限的扩展, 从而解决了传统专家系统知识库难以扩展且不能解决变化知识的问题.

### 1 可拓专家系统概念及研究现状

可拓学用可拓模型来描述客观世界的物和事, 既简洁、统一、使用方便, 又为拓展、变换、收敛创造了有利条件, 以此为基础的研究形成了以解决不相容问题为己任的定性、定量相结合的可拓工程方

法<sup>[1]</sup>. 应用可拓学的研究成果来解决 Web 专家系统中的实际问题是一个非常具有应用价值的领域.

可拓知识中的变换蕴含式是典型的变化知识, 它是目前可拓学发展过程中最适合将之应用于专家系统研究的一个创新点, 从而将专家系统研究向前推进了一步. 在专家系统的核心部分, 如知识库和推理机中运用可拓学相关理论研究成果的专家系统称为可拓专家系统. 近年来, 有不少学者尝试将可拓学与专家系统相结合, 取得了初步的成果, 但基于 Web 的可拓专家系统还少有人研究.

### 2 基于 Web 可拓 ES 的结构框架

#### 2.1 Web 技术应用

对于开发一个基于 Web 的专家系统, 合适的动态 Web 开发技术的选择关系到整个系统的开发、管理、与专家系统开发工具的集成以及可扩展性. 本文

研究的是一种使用 Java 开发的基于 Web 的专家系统,那么就规定了 Web 开发工具必须与 Java 具有良好的集成并使 Web 应用具有 Servlet 的一切优点.而用户界面的开发则可以采用 JSP 技术,使整个 Web 应用具有跨平台的特点,可以方便地移植.并使用 Struts 来构建 Web 应用,可以方便地管理业务逻辑的流程,便于系统的更新与扩展<sup>[2]</sup>.

使用 JSP/Servlet/JavaBean 等 Web 开发技术的可拓专家系统的框架如图 1 所示.

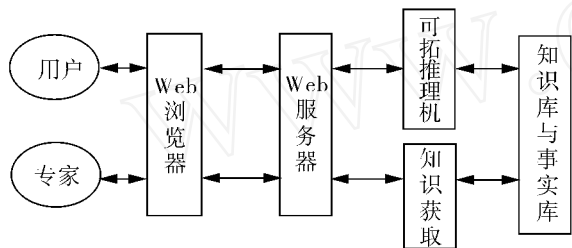


图 1 基于 Web 可拓专家系统框架

Fig 1 Framework of expert system with extension based on the Web

## 2.2 知识库设计

### 2.2.1 知识的可拓表示形式

可拓学的基元理论提供了定性与定量相结合的完整的表述信息的一种有效途径.专家知识可用基元来表示<sup>[3]</sup>.例如,有 2 条知识:1)“计算器是用来计算的”;2)“雪是白的”.可以用物元来表示: $R_1 = (\text{计算器}, \text{功能}, \text{计算})$ ;  $R_2 = (\text{雪}, \text{颜色}, \text{白})$ , 这样的知识称为可拓知识.

产生式规则作为一种推理机制广泛地应用于控制和决策领域,它是计算机专家系统和决策支持系统中形成控制规则和决策规则的有效工具.产生式规则具有“if<条件> then <结论>”的形式,产生式规则包括左边(A)和右边(B)2 部分,分别称为产生式规则的前件和后件.前件决定规则的可用条件;后件表示条件成立时应采取的行动,或条件成立时产生的结果,它是产生式被调用以后所做的事情.如果把产生式规则中前件的结点和后件的结点换成基元,就得到可拓规则.也就是说,在可拓规则中,知识结点是用基元来表示的.用基元来表示专家知识,形成可拓知识元,用可拓知识元来表示规则的前件和后件,这是可拓规则的基本特征.

在当前盛行的专家系统中,经常以(对象,属性,值)三元组的形式作为产生式系统的规则,例如(细菌,染色斑,革兰氏阳性)(细菌,形态,球状)

(细菌,生长结构,链状) $\Rightarrow$ (细菌,类别,链球菌).就是这样一个产生式,这一产生式很容易用基元表示: $R_1 = (\text{细菌}, \text{染色斑}, \text{革兰氏阳性})$ ,  $R_2 = (\text{细菌}, \text{形态}, \text{球状})$ ,  $R_3 = (\text{细菌}, \text{生长结构}, \text{链状})$ ,  $R_4 = (\text{细菌}, \text{类别}, \text{链球菌})$ .则这个产生式规则可表示为:if  $R_1 \quad R_2 \quad R_3$  then  $R_4$ <sup>[4]</sup>.

利用可拓推理的拓展性推理(包括发散推理规则、蕴含推理规则、相关推理规则和可扩推理规则)、传导性推理、共轭推理等多种推理规则,使可拓规则在求解问题时功能更为强大,使知识规则体现可拓思维方式.

对于基元的产生式,可以利用基元的拓展性而开拓出新的基元,从而利用基元变换,产生动态变化的知识.对基元的变换会产生传导变换,传导变换的过程也可用推理规则来描述,即传导推理规则.传导推理是指前件为可拓变换,后件为它所引起的传导变换的推理<sup>[5]</sup>.如:对于两类规则  $A \Rightarrow B$ ,  $B \Rightarrow N$ .一般情况  $A = a_i$ ,  $B = b_j$ .若存在条件的可拓变换  $T_c$ ,  $T_c(B) = A$ ,并存在结论的可拓变换  $T_r$ (它为  $T_c$  的传导变换),  $T_r(N) = P$ ,则成立可拓变换规则知识(变化知识):

$$(T_c(B) = A) \Rightarrow (T_r(N) = P). \text{即 } \text{if } T_c(B) = A \text{ then } T_r(N) = P.$$

对于 2 条同类规则  $A \Rightarrow P$ ,  $C \Rightarrow P$ .若存在可拓变换  $T_c(B) = A$ ,则成立可拓变换规则知识  $(T_c(B) = A) \Rightarrow P$ .即  $\text{if } T_c(B) = A \text{ then } P$ .

类似地,根据物的共轭性质,也可以得到一系列的推理规则,称为共轭推理.在一定条件下,物的某一共轭部的变化,会导致其相应的共轭部发生传导变换<sup>[6]</sup>.

### 2.2.2 可拓知识库的建立

对于不同的知识表示方法,知识库的构建也具有不同的方法和形式.在大多数的专家系统尤其是诊断专家系统中,知识往往最容易通过产生式规则来表示.通常一条产生式规则由规则号、规则前提和规则结果组成,前提和结果之间是 if-Then 的关系<sup>[7]</sup>.可以在关系数据库中将 3 部分分别用 3 个字段来表示可拓规则,或者前提组成事实数据库,结果组成结果数据库,另外的一个可拓规则库则将两者之间的关系对应.在关系数据库中,存在 3 种关系:一对一、一对多和多对多.一对一是最简单的一种关系,如果可拓知识可以表示成一对一的关系,那么只需一个表就可以表示可拓知识.如果可拓知识比较

复杂,只能表示成一对多和多对多的关系,那么事实与结果独立存储,另外由规则库来表达两者联系的方式,这无疑更适合可拓知识的表示以及维护.在构建可拓知识库以及日后的知识维护过程中,知识的冗余与冲突是值得关注的,可拓知识的冗余可以通过关键字的设置来缓解,而适当的冲突消解策略如基于可信度排序则可以适当地缓解冲突.

2.3 数据库连接

Java数据库连接 (Java data base connectivity, JDBC)是一种用于执行 SQL 语句的 Java API 它由一组用 Java 编程语言编写的类和接口组成. JDBC 为工具/数据库开发人员提供了一个标准的 API,使他们能够用纯 Java API 来编写数据库应用程序. JDBC 扩展了 Java 语言的功能.使用 JDBC,向各种关系数据库发送 SQL 语句成为很容易的一件事.将 Java 和 JDBC 结合起来,将会使程序经过一次编译即可在其他任何平台上运行. JDBC API 用于连接 Java 应用程序与各种关系数据库.这使得人们建立客户服务器应用程序时,通常把 Java 作为编程语言,把任何一种浏览器作为应用程序的友好界面,把 Internet 或 Intranet 作为网络主干,把有关的数据库作为数据库后端.

2.4 Web 程序与可拓 ES 事实库及知识库的交互

在基于 Web 的专家系统中,Web 页面的表单首先要从事实库中读取事实.供用户选择后提交到可拓推理机,可拓推理机需要访问可拓知识库,将用户表单提交的事实与可拓知识库的规则前提进行匹配,进而进行可拓推理,最后将推理结果返回给 Web 页面.

2.5 可拓推理机设计

2.5.1 可拓规则的匹配

可拓规则的匹配根据基元的特性,对可拓规则的匹配方法有欧式距离匹配方法和可拓距匹配方法.对于基元属性取精确数量值的规则,要求证据与规则前件绝对相符,否则认为不匹配,因此采用欧式距离匹配方法.判断证据与规则前件的最简单的精确匹配办法是计算证据与条件之间的欧式距离,如果距离为 0,则认为精确匹配,如果距离大于 0,则认为不匹配.若有多维物元:

$$\begin{bmatrix} R, & c_1, V_1, \\ & c_2, V_2 \\ & \dots \\ & c_n, V_n \end{bmatrix}$$

证据为

$$\begin{bmatrix} R, & c_1, V_1, \\ & c_2, V_2 \\ & \dots \\ & c_n, V_n \end{bmatrix},$$

则匹配度:

$$P = \frac{\sqrt{(V_1 - V_1)^2 + (V_2 - V_2)^2 + \dots + (V_n - V_n)^2}}{\sqrt{\sum_{i=1}^n (V_i - V_i)^2}} =$$

而对于基元属性取特定区间值的规则,要求证据的属性值落在特定区间内,因此采用可拓学中的距函数来计算匹配度.用 Java 语言设计实现可拓推理机,使推理机可以在任何平台上使用.

可拓学中的距匹配方法:设  $x$  为实域  $(-, +)$  上的任一点,  $X_0 = (a, b)$  为实域上的任一区间,称  $(x, x_0) = |x - (a + b) / 2| - (a - b) / 2$  为点  $x$  与区间  $X_0$  之间的距.

若有多维物元  $\begin{bmatrix} R, & c_1, V_1, \\ & c_2, V_2 \\ & \dots \\ & c_n, V_n \end{bmatrix}$  证据为

$$\begin{bmatrix} R, & c_1, V_1, \\ & c_2, V_2 \\ & \dots \\ & c_n, V_n \end{bmatrix}, \text{ 则匹配度: } P = 1/n \left( (v_1, V_1) + (v_2, V_2) + \dots + (v_n, V_n) \right) = 1/n \sum_{i=1}^n (v_i, V_i).$$

如果  $P < 0$ , 证据与条件匹配;如果  $P = 0$ , 证据点落在属性区间边界上,勉强匹配;如果  $P > 0$ , 证据与条件不匹配.对于其他特殊情况,如出现不同特征的距相互抵消,则不能说明匹配,应试具体情况分析.

2.5.2 可拓推理机算法及实现

1) 可拓推理机算法的推理思想是:首先,建立专家给定的可拓规则集和用户给出的证据集;其次,计算证据和规则前件是否匹配成功并激活规则;然后,将得到的推理结论加入证据集结论;最后,循环执行匹配,直到整个系统不再产生新的推理结论为止,退出推理.

算法步骤:

初值:设有  $N$  条可拓推理规则,  $M$  条证据.令  $r = 1$ .

判断:如果  $1 < r < M$ , 从证据集中取出第  $r$  条证据 ProofSet( $r$ );如果  $r > M$ , 转 .

令  $i = 1$ .

如果  $1 < i < N$ , 从规则集取出第  $i$  条规则 RuleSet( $i$ ); 如果  $i > N$ , 转 ;

比对证据名和规则前件名, 若相同, 则转 , 否则转 .

计算匹配度: 根据规则前件的不同, 用不同的公式计算匹配度.

所有规则前件的子前提都比对后, 则得到推理结论或新的证据. 若得出新的证据, 则将证据添加到证据集中,  $M++$ . 若得出结论, 则将结论添加到结论集中.

$i++$ , 转 .

$r++$ , 转 .

输出推理结论集.

## 2) 可拓推理机的设计实现.

考虑到专家系统的跨平台应用, 可拓推理机用 Java 语言开发, 在 Windows 平台上测试通过, 打包成一个 ExI.jar 文件. 可拓推理机主要有如下几个类: ExEngine(引擎类)、ExRuleSet(可拓规则集类)、ExEvidenceSet(可拓证据集类)、ExConclusionSet(结论集类)、ExRule(可拓规则类)、ExEvidence(可拓证据类)、ExConclusion(结论类)、ExRuleIF(可拓规则前件类)、ExRuleTHEN(可拓规则后件类)、ExVariable(基元类).

基元类是推理机的最小运算单元, 它可以是一个物元、关系元或事元的表示, 它是构成可拓规则前件、结论的基本元素. 可拓规则类由可拓规则前件类和可拓规则后件类组成. 可拓规则前件类由若干个基元类组成, 这些基元类之间可以是任何逻辑关系, 可拓规则后件类由可拓证据类或结论类组成. 若干个可拓规则类构成可拓规则集类, 若干个可拓证据类构成可拓证据集, 若干个结论构成结论集. 整个引擎类由可拓规则集、可拓证据集、结论集构成. 引擎类是可拓推理机的核心类, 它为专家系统的实现提供了一个运行平台. 引擎类的 add() 方法是一个多态方法, 可以用来向推理引擎添加用来推理的规则集和证据集. 引擎类的 run() 方法用来实现可拓推理算法, 即根据规则集和证据集推出结论集, 并输出最后的推理结论.

## 3 人机接口

系统的实现主要是引入了目前主流设计技术: 基于 Web 的 B/S 结构程序框架. 即人机接口处以网络浏览器为实现依托, 用户只要安装有 Web 浏览器, 通过网络联接即可使用本系统, 能做到客户端零

维护, 数据库及知识库使用 SQL Server 集中管理, 真正意义上实现数据库及知识库的“无限”扩充.

## 4 应用实例: 基于可拓学的口腔疾病诊断网络专家系统

本系统是基于可拓方法对口腔溃疡疾病信息进行描述, 制定可拓规则, 建立可拓知识库和事实库. 利用可拓推理和可拓变换去解决问题, 相比常规专家系统能体现出知识的动态变化性的特点.

### 4.1 系统开发环境及结构

由于整个系统使用的 JSP、Servlet、JDBC 都基于 Java 语言, 因此基于 Web 的 (网络故障诊断) 专家系统实例分析具有良好的移植性与平台无关性. 因此在开发环境的选择上基本不受任何限制.

实例选择的系统平台 Windows XP Home Edition, 数据库选择的是 Microsoft SQL Server 2000, 应用服务器选择的是 Tomcat

### 4.2 可拓知识表示与可拓知识库设计

从口腔学专家那里, 得到以下规则及传导推理规则:

- 1) (粘膜, 状态, 充血) (粘膜, 状态, 水肿) = > (溃疡, 程度, 初期);
- 2) (溃疡面, 形态, 规则) (溃疡, 愈合时间, 数天至数周) = > (溃疡, 类型, 良性);
- 3) (溃疡面, 形态, 不规则) (溃疡, 愈合时间, 数月数年) = > (溃疡, 类型, 恶性);
- 4) (溃疡面, 底部, 平滑) (溃疡, 病程规律, 反复发生) (溃疡, 全身症状, 淋巴结正常) = > (溃疡, 类型, 良性);
- 5) (溃疡面, 药物的敏感程度, 差) (溃疡面, 底部, 不平) (溃疡, 全身症状, 淋巴结肿大) = > (溃疡, 类型, 恶性);
- 6) (粘膜, 表面状况, 齐平) (粘膜, 状态, 充血和糜烂) (上皮, 连续性, 不中断) = > (口腔, 疾病类型, 糜烂);
- 7) (溃疡, 病程规律, 无复发史) (上皮, 连续性, 中断) (溃疡面, 底部, 坚韧) = > (溃疡, 类型, 恶性).

确定可拓变换规则知识:  $T_1$ : (溃疡面, 形态, 规则) = (溃疡面, 形态, 不规则);  $T_2$ : (溃疡, 愈合时间, 数天至数周) = (溃疡, 愈合时间, 数月至数年); 也存在可拓变换:  $T_3$ : (溃疡, 类型, 良性) = (溃疡, 类型, 恶性). 则有: if {  $T_1 T_2$ : (溃疡面, 形态, 规则) (溃疡, 愈合时间, 数天至数周) = (溃疡面, 形态, 不规则) (溃疡, 愈合时间, 数月至数

年) } then {  $T_3$ : (溃疡,类型,良性) = (溃疡,类型,恶性) }.

在数据库中基元表示见表 1 (部分).

表 1 基元知识表 (部分)

Table 1 Basic-element knowledge ( part)

$R_i$	$N$	$C$	$V$
1	粘膜	状态	充血
2	粘膜	状态	水肿
3	溃疡	程度	初期
4	溃疡	形态	规则
5	溃疡	愈合时间	数天至数周
6	溃疡	类型	良性
7	溃疡面	形态	不规则

可拓规则存储关系表内容见表 2 (部分):

表 2 可拓规则表 (部分)

Table 2 Extension rules ( part)

规则号	病 症		病 因
1001	$R_1$	$R_2$	$R_3$
1002	$R_4$	$R_5$	$R_6$
1003	$R_7$	$R_8$	$R_9$
1004	$R_{10}$	$R_{11}$	$R_{12}$
1005	$T_1$	$T_2$	$T_3$

4. 3 Web系统设计

基于 Web的专家系统是一个以数据库操作为主、相对逻辑复杂的系统,而且推理过程都具有一定的相似性,适合组件重用,因此非常适用于使用 MVC设计模式. 本文的实例便采用了基于 Struts框架的 MVC模式来开发 Web系统. 根据上文系统流程分析的步骤,整个系统的主要页面由首页、登录、最新消息、口腔病症介绍、口腔溃疡诊断专家系统组成. 图 2是专家系统首页 (粘膜部分):



图 2 口腔溃疡诊断专家系统首页 (部分)

Fig 2 Home page of oral ulcer diagnosis expert system (part)

专家系统的诊断过程是整个系统的核心部分,它包含了专家知识获取以及推理过程. JavaBean首先获取病人症状的列表,将症状申明成事实. 推理机

将用户表单提交的事实与可拓知识库的规则前提进行匹配,进而进行推理,最后将推理结果返回给 Web页面,将结果展示给用户.

本例中,当患者提交的症状中当溃疡面由规则变得不规则,愈合时间由原来的数天到后来的数月甚至数年时,专家系统可利用可拓推理规则得出患者已经由良性溃疡转变为恶性溃疡了.

5 结束语

可拓学用可拓模型来研究事物的拓展性,可拓推理规则是解决矛盾问题的推理形式. 而动态 Web技术是一种开发网络专家系统的适宜环境,本文探讨了将可拓学初步应用于网络专家系统中,取得了异于常规专家系统的论断效果. 随着可拓学研究的不断深入,应用可拓学的研究成果来解决基于 Web的专家系统问题将有着令人鼓舞的应用前景.

参考文献:

[1] 杨春燕,蔡 文. 可拓工程 [M]. 北京:科学出版社,2007: 3-5.  
[2] 飞思科技. JSP应用开发详解 [M]. 北京:电子工业出版社,2004: 45-48.  
[3] JOHNSON R, SONS J W. Professional Java development with the spring framework [M]. [ S 1 ]Wiley Publishing, 2005: 301-305.  
[4] CAIWen. Extension management engineering and applications[J]. International Journal of Operations and Quantitative Management, 1999 (1): 50-87.  
[5] 李立希,杨春燕. 可拓策略生成系统 [M]. 北京:科学出版社,2006: 76-79.  
[6] 杨春燕,蔡 文. 可拓信息—知识—智能形式化体系研究 [J]. 智能系统学报, 2007, 2 (3): 8-11.  
YANG Chunyan, CAIWen. A formalized system of extension information—knowledge—intelligence [ J ]. CAAI Transactions on Intelligent Systems, 2007, 2 (3): 8-11.  
[7] 蔡 文,杨春燕,何 斌. 可拓逻辑初步 [M]. 北京:科学出版社,2003: 56-59.

作者简介:



张海涛,男,1982年生,硕士研究生,主要研究方向为专家系统.



董春游,男,1962年生,教授,硕士生导师,主要研究方向为人工智能、决策支持系统. 曾获得煤炭科学技术三等奖和黑龙江省科技进步三等奖. 发表学术论文 20余篇,出版专著 1部.