

用 XML 建造专家系统知识库

詹子鹏, 李龙澍

(安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039)

摘 要:专家系统是一种基于知识的智能系统, 主要包括知识库、数据库、推理机制、解释机制、人机接口和知识获取等功能模块。它的主要特征是, 系统中有一个巨大的知识库, 存贮某个领域的专门知识, 有很多种技术管理知识库。随着网络技术的发展, XML 已经成为重要的网络计算语言。文中就如何用 XML 建造专家系统的知识库以及其相关操作做了深入的探讨。

关键词:知识库; 专家系统; XML 数据处理

中图分类号: TP182

文献标识码: A

文章编号: 1673-629X(2007)07-0031-04

Construct the Expert System Knowledge Base with XML

ZHAN Zi-peng, LI Long-shu

(Ministry of Education Key Lab. of Intelligent Computing & Signal Processing, Anhui Univ., Hefei 230039, China)

Abstract: Expert system is procedure design method that a procedure for having specialized knowledge in large quantity system. An expert system is constituted from knowledge base, database, reasoning mechanism, explaining mechanism, man-machine interface and knowledge obtaining. Its main characteristic is, has an enormous knowledge base in the system, saves the specialized knowledge of a certain realm in it, has a lot of technocracies knowledge base. Along with technical development in network, XML has become the important network calculation language. This text discussed that how to use knowledge base and its related operations that XML construct expert systems.

Key words: knowledge base; expert system; XML data processing

0 引言

在网络上, XML 作为一种交换商务数据的标准越来越流行^[1]。在 XML 发布之前, 国际互联网的发展受到 HTML 如下几个问题的束缚:

- 1) HTML 无法描述数据内容, 而这一点恰恰是数据检索、电子商务所必需的。
- 2) HTML 对数据表现的描述能力是十分不够的, 如 HTML 还不能描述矢量图形、科学符号等对象, 目前只能通过图像来表现这些对象。
- 3) HTML 实例置标语言的地位, 完全不能适应对新标记需求的发展需要。

XML 的出现, 使上述问题都得到很好的解决。试想在互联网世界, 如果大家都使用不同的语言, 互相交

换信息时都要进行翻译, 那将是一个多么难以沟通和交流的世界。在互联网世界, XML 之所以重要, 恰恰是由于它扮演了“国际语言”角色的缘故。此外, XML 更为互联网世界提供了定义各行各业的“专业术语”的工具。因为可以用 XML 设计自己的文件类型。信息的内容将更加丰富, 更加方便使用。在 XML 中超文本链接的能力比 HTML 更强。XML 能使得浏览器的表现形式更加丰富。

专家系统使用了基于知识的程序设计方法, 它的主要特征是, 系统中有一个巨大的知识库, 存贮某个领域的专门知识, 要存储这个知识库, 可供选择的方法可以有多种, 笔者就如何用 XML 存储专家系统的知识库以及其相关操作做探讨。

文中所用计算机编程语言为 ASP, XML, HTML, 其中 Microsoft Active Server Pages(ASP)是服务器端脚本编写环境, 使用它可以创建和运行动态、交互的 Web 服务器应用程序。使用 ASP 可以组合 HTML 页、脚本命令和 ActiveX 组件以创建交互的 Web 页和基于 Web 的功能强大的应用程序。ASP 应用程序很容易开发和修改。

收稿日期: 2006-09-20

基金项目: 安徽省自然科学基金项目(050420204); 安徽省教育厅自然科学基金项目(2006KJ098B); 安徽省高校拔尖人才基金项目(05025102)

作者简介: 詹子鹏(1982-), 男, 安徽阜阳人, 硕士研究生, 研究方向为智能软件; 李龙澍, 教授, 博士生导师, 研究方向为智能软件。

1 专家系统的组成

专家系统(Expert System)是一个具有大量专门知识的程序系统。它应用人工智能[Artificial Intelligence, AI]技术,根据一个或多个人类专家提供的特殊领域知识进行推理、模拟人类专家作决定的过程来解决那些需要专家才能解决的复杂问题^[2]。目前,专家系统不仅限于解决科学问题,而且已经开始用于工业、企业界并已渗透到社会的方方面面,如医学、地质、故障诊断、计算机软硬件、化学分析和生物工程等领域。专家系统按知识表示技术可分为:基于逻辑的专家系统,基于规则的专家系统,基于语义网络的专家系统,基于框架的专家系统。而按任务类型则可分为:解释型:用语分析符号数据,进而阐明这些数据的实际意义;预测型:根据对象的过去和现在情况来推断对象的未来演变结果;诊断型:根据输入信息来找出对象的故障和缺陷;调试型:给出已确定的故障的排除方案;维修型:指定并实施纠正某类故障的规划;规划型:根据给定目标拟订行动计划;设计型:根据给定要求形成所需方案和图样;监护型:完成实时监测任务;控制型:完成实时控制任务;教育型:诊断型和调试型的组合,用于教学和培训。

专家系统第一个重要组成部分是知识库,它存储从专家那里得到的、关于某个领域的专门知识^[3]。专家系统的第二个组成部分是推理机,它具有依据一定的策略进行推理的能力,即能够根据知识推导出结论,而不是简单地搜索现成的答案。

为了建造知识库,需要解剖如何存储知识的问题,这就是知识表示,确切地说就是如何以计算机能够存储的形式表达知识。另一个与此相关的问题是如何从专家那里获得知识,即所谓的知识获取问题,在建立专家系统的过程中,这是个重要的问题。

专家系统的很多能力来自于对所存储的大量专家知识和恰当地应用有效的推理技术。推理方法不可能完全独立于所要解决问题的种类,但同时,推理模型的研究又不能过于依赖某一特定问题,以致在其他场合不能适用。所要寻求的推理方法应该是很有效的,同时又可普遍地应用于描述和解决一类问题。

推理机除了具有推理能力以外,还包括和使用者的能力及解释的能力,这被称为问题咨询。在某些系统里,使用者可以很方便地使用自然语言进行人机对话。

2 基于 XML 的网络专家系统特色

XML 的优势之一是开放性,它允许各个组织、个人建立适合自己需要的置标集合,并且这些置标可以

迅速地投入使用^[4]。这一特征使得 XML 可以在电子商务、政府文档、司法、出版、CAD/CAM、保险机构、厂商和中介组织信息交换等领域中一展身手,针对不同的系统、厂商提供各具特色的独立解决方案。

XML 的优势之二是分离性,它的数据存储格式不受显示格式的制约。一般来说,一篇文档包括三个要素:数据、结构以及显示方式。对于 HTML 来说,显示方式内嵌在数据中,这样在创建文本时,要时时考虑输出格式,如果因为需求不同而需要对同样的内容进行不同风格的显示时,要从头创建一个全新的文档,重复工作量很大。此外 HTML 缺乏对数据结构的描述,对于应用程序理解文档内容、抽取语义信息都有诸多不便。而与之不同,XML 把文档的三要素独立开来,分别处理。首先把显示格式从数据内容中独立出来,保存在样式单文件(StyleSheet)中,这样如果需要改变文档的显示方式,只要修改样式单文件就行了。XML 的自我描述性质能够很好地表现许多复杂的数据关系,使得基于 XML 的应用程序可以在 XML 文件中准确高效地搜索相关的数据内容,忽略其他不相关部分^[5]。

XML 还有其他许多优点,比如它有利于不同系统之间的信息交流,完全可以充当网际语言。大家知道,网络专家系统是利用电子手段尤其是互联网进行和用户进行交互活动。从技术上说,网络专家系统是通过互联网传输和交换知识库中的数据,并能根据知识库中的数据进行人工或自动处理,为访问该系统的用户提供方便、快捷、准确的服务。XML 的可扩展性和自相容性等特点,使它成为数据交换的有力工具^[6]。

随着 XML 标准体系的成熟和技术的发展,已经出现了相当多的客户化工具,尤其是可视化工具的出现,使得人们可以无须了解 XML 的细节就能够编写出需要的 XML 文档,而浏览器对 XML 越来越强的支持能力,对 XML 应用起到了巨大的促进作用。

综上所述,与用其他语言实现的专家系统相比较,基于 XML 的网络专家系统有如下特点:

- 1) 在线查询,使用方便。用户无需安装客户端,直接访问相关网站,就可以使用该网站下设的专家系统进行咨询活动。
- 2) 信息量大,指导性强,信息及时准确而赢得广大用户的青睐。
- 3) 知识库内容更新快捷方便,而且内容实用、新颖。系统管理员可以直接修改网络专家系统的知识库,所以用户不用在线下载升级包去升级本地客户端,从而为用户省下了大量的时间、精力。
- 4) 界面更友好,便于用户查阅,更贴近用户,服务于用户。ASP 语言结合 CSS(样式表文件)可以使用户

界面更为清爽、整洁。

3 知识库管理模块的实现

知识库管理主要由 index.asp, PostNew.asp, List.xml 和 Del.asp 四个模块组成。

1) index.asp 是浏览知识库中内容的主界面(图 1 是此模拟程序在本地机上的演示图(系统:winXP, IE 浏览器 6.0)),这是 index.asp 的运行效果图。

```
具体代码如下:
<% strSourceFile = Server.MapPath("List.xml")
Set objXML = Server.CreateObject("Microsoft.
FreeThreaddXMLDOM")
objXML.load(strSourceFile)
'定义一个目标 XML 文件,从 XML 文件中<
Knowledgebase></Knowledgebase>里存储或提
取知识
Set objRootsite = objXML.documentElement. se-
lectSingleNode("Knowledgebase")
%>
<%
    PageSize=2
    '获取子节点数据
    AllNodesNum=objRootsite.childNodes.length-1
    PageNum=AllNodesNum\PageSize+1
    '获得起始节点
    StarNodes=PageNo*PageSize-1
    '获得结束节点
    EndNodes=(PageNo-1)*PageSize
    '判断起始节点数是否超过总的节点数
    if StarNodes>AllNodesNum then
        '如果超过则结束节点要减去(StarNodes-AllNodesNum)的差
        值,否则下标会超界出错
        EndNodes=EndNodes-(StarNodes-AllNodesNum)
        StarNodes=AllNodesNum
    end if
    while StarNodes>=EndNodes
        name=objRootsite.childNodes.item(StarNodes).childNodes.item
        (1).text
        .....
    %>
```

代码说明:这段代码将数据从 List.xml 中读入,并显示到一个与用户交互的界面中。如语句:name=objRootsite.childNodes.item(StarNodes).childNodes.item(1).text 即是从 List.xml 中读出用户自定义的标签<name></name>中的数据内容,并把它赋给变量 name,再通过 ASP 语句将变量 name 的数值显示到 IE

浏览器中。

2)PostNew.asp 是向知识库中添加内容的主界面,图 2 是 PostNew.asp 的运行效果图。



图 1 index.asp 运行效果图



图 2 PostNew.asp 运行效果图

具体代码如下:

```
name = Request.Form("name")
'要求输入待添加记录的名称
text = Request.Form("text")
'要求输入待添加记录的内容
Posttime = now()
'记录添加记录的时间,便于更新或其他操作
Set oListNode = objXML.documentElement.selectSingleNode
(" Knowledgebase "). AppendChild ( objXML.createElement
("item"))
'在节点<Knowledgebase>下创建一个节点<item>
Set oDetailsNode = oListNode.appendChild(objXML.createElement
("id"))
oDetailsNode.Text = id
.....
```

代码说明:这段代码将用户在交互界面输入的信息存储到 List.xml 中。如语句: Set oDetailsNode = oListNode.appendChild(objXML.createElement("name"))

oDetailsNode.Text = name

即是在 List.xml 文件中创建一个节点 <name>, 并将变量 name 接收的值赋给标签 <name> </name> 中的数据内容。

3) List.xml 负责存储(输出)由用户界面输入(调用)的数据, 其代码如下:

```
<? xml version="1.0" encoding="gb2312"? >
<xml><Knowledgebase>
<item>
<id>1</id>
<name>感冒</name>
<Posttime>2004-5-14 10:28:05</Posttime>
<causeofill>受凉</causeofill>
<disease>头疼,发热</disease>
<medicine>感冒灵,一天三次,一次一粒。</medicine>
<remarks>注意休息,祝早日康复!</remarks>
</item>
<item>
<id>2</id>
.....
</item>
</Knowledgebase></xml>
```

代码说明: 标签 <Knowledgebase>, <id>, <name>, <causeofill>, <disease>, <medicine>, <remarks>等都是由用户定义的, 从标签的命名上直观地反映了标签包含的数据内容的具体含义, 这正是 XML 的优点之一, 譬如: <id> </id> 中的数据内容是表示这是第几条记录, <name> </name> 中的数据内容是表示这条记录的名称是知识库中哪个部分的名称, <causeofill>, </causeofill>, <disease>, </disease>, <medicine>, </medicine>, <remarks>, </remarks> 中的内容表示的是记录具体内容。用户可以通过交互界面的查询功能搜索知识库中的知识规则, 如果能找到, 即匹配查找成功, 那么输出这条记录, 如果由多条记录同时满足, 则列表给出或给出一个相似度最高的数据, 如果未能找到, 即匹配查找不成功, 则显示没有这项记录, 并询问用户是否要添加所查记录^[7]。例如, 用户输入“发热 头疼”, 经过查找匹配, 系统可以找到相应的记录, 则输出:

“病因: 受凉

病症: 发热, 头疼

用药: 感冒灵, 一天三次, 一次一粒

备注: 注意休息, 祝早日康复!!”

4) Del.asp 是用户删除已经添加到知识库中的记录的交互界面, 其代码如下:

```
Set objRootsite = objXML. documentElement. selectSingleNode
("Knowledgebase")
AllNodesNum = objRootsite. childNodes. length - 1
response. write AllNodesNum
objXML. DocumentElement. RemoveChild(objRootsite)
objRootsite. parentNode. removechild(objRootsite)
response. write "ok"
objXML. save(strSourceFile)
```

4 结束语

XML 在实际应用中的重要作用越来越明显, 作为一种数据交换的标准, 它已经逐渐地流行起来^[1]。它允许各个领域自定义自己行业内部的通用标记, 并且易于阅读和编写, 为网络上交换数据提供了很好的标准。

如前面所说, 一个完整的专家系统主要包括知识库、数据库、推理机制、解释机制、人机接口和知识获取等功能模块。文中着重探讨了如何用 XML 建造专家系统的知识库以及其相关操作。

参考文献:

- [1] 黄理, 曹林有, 张勇, 等. ASP.NET/XML 深入编程技术[M]. 北京: 北京希望电子出版社, 2002.
- [2] 程慧霞, 李龙澍, 倪志伟, 等. 用 C++ 建造专家系统[M]. 北京: 电子工业出版社, 1996.
- [3] 张全寿, 周建峰. 专家系统建造原理及方法[M]. 北京: 中国铁道出版社, 1992.
- [4] Ennsen L. WEBSPIHERE 环境下 XML 与 XSL 编程[M]. 北京: 机械工业出版社, 2001.
- [5] Baartse M. ASP 与 XML 高级编程[M]. 康博译. 北京: 清华大学出版社, 2002.
- [6] Freeman A, Jones A. Microsoft. NET XML Web 服务程序设计[M]. 北京: 清华大学出版社, 2003.
- [7] 胡海璐. XML Web Services 高级编程范例[M]. 北京: 电子工业出版社, 2003.

(上接第 30 页)

- [2] Abrial J R. B 方法[M]. 裴宗燕译. 北京: 电子工业出版社, 2004.
- [3] 邹盛荣, 郑国梁. B 语言和方法与 Z、VDM 的比较[J]. 计算机科学, 2002(10): 136-138.
- [4] 肖美华, 薛锦云. 形式化方法 B 及其程序规约机理[J]. 计

算机工程 2004, 30(16): 16-18.

- [5] Szyperski C, Gruntz D, Murer S. 构件化软件: 超越面向对象编程[M]. 第 2 版. 王千祥, 曹东刚, 左继宏译. 北京: 电子工业出版社, 2004.