

基于 Web 的决策支持系统的研究与设计

Research and Design of Decision Support System Based on Web

肖春艳¹ 吴 斌² 史忠植²

(中国科技大学研究生院计算机学部 北京100039)¹

(中科院计算技术研究所智能信息处理开放实验室 北京100080)²

Abstract This paper discusses a decision support system based on Web. It first briefly introduces the basic theory, components and functions of this system. The paper also details the key technology used in this system, including the structure and advantage of browser/server based system, expert system, the technology of case-based reasoning and data mining over Web.

Keywords Decision support, Expert system, Case-based reasoning, Data mining, Web

1. 引言

自70年代中期 Keen 和 Scott Morton 创造“决策支持系统(DSS)”一词至今,作为用于管理的一种新型计算机信息系统-DSS 已经得到了迅速的发展,成为系统工程、管理科学、人工智能等领域十分活跃的研究课题。

随着 Internet 和网络技术的飞速发展,特别是 Web 技术的出现,为决策支持系统提供了一个新的应用空间。通过网络可以实现信息的快速收集、发布、存储、处理和交流,因此可以为决策过程提供大量有效而低成本的信息资源。从技术上讲,广泛可用的 Internet 连接,个人多媒体计算机的普及,先进的 Web 技术,优良的 Java 等程序设计语言,已经提供了开发基于 Web 的智能决策支持系统(Web Based Intelligent Decision Support System--WIDSS)的可能。用户可以在全球任何地方方便地使用基于 Web 的决策支持系统,为决策者提供远程决策的支持和服务,具有较好应用前景。

2. WIDSS 的系统模型

2.1 系统概述

DSS 可以为决策提供支持,充分运用可供利用的、合适的计算机技术,针对半结构化或非结构化问题,通过人机交互的方式帮助和改善管理决策制定的有效性。而将 DSS 与专家系统 ES 相结合,分别发挥 DSS 数值分析与 ES 符号处理的特点,形成的智能决策支持系统(IDSS)具有定量分析和定性分析相结合支持决策的能力,能够更加有效地解决半结构化或非结构化问题。

在基于 Web 的决策支持系统中,网络用户可以在客户端使用浏览器,通过网络访问集中在服务器端的应用,方便快捷地取得决策支持。系统具有良好的知识获取和解释子系统;运用知识发现工具进行数据存取和查询、识别数据的模式、得出规则并利用其它的数据来精炼、检验这些规则;同时该系统能够对外部数据(如 Internet 中的信息)进行处理并发掘出有价值的信息;在外部信息或内部规则、知识不完备的情况下,系统仍能进行一定的推理,并给出一个“满意”的结果;提供用户对以往决策案例的记忆。系统数据库、知识库、推理子系统、信

息发掘子系统之间的结构能适应在动态环境中不断进行更新,随外界环境的改变而相应地变化,提供较长时期的服务。

2.2 系统框架结构

基于 Web 的决策支持系统是一种开放式的体系结构,采用“三库”一体化结构,即数据库、模型库和知识库,利用专家系统技术,构造一种智能调度方式,获取相关业务方面的专家知识,形成决策知识库;利用推理机为业务管理人员提供决策支持;在系统地整理相关历史案例的基础上,形成范例库,利用范例推理系统完成案例的检索、复用、修正及保存,辅助业务处理;系统应用 Web 技术提供浏览器/服务器模式的决策支持;知识发现与数据挖掘为系统提供更深层次的知识发现和管理功能。为了更好地进行业务咨询和案例处理,系统还提供业务法规和合同范本检索功能,用户可以独立地查询或在决策过程中浏览相关规则和实例。具体的系统组成框图如图1所示。

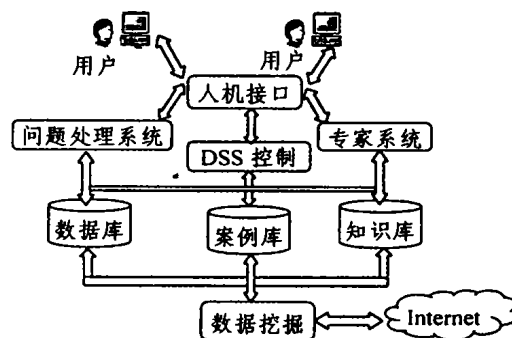


图1

其中:1)人机接口是决策支持系统与用户交互的窗口,它提供友好的人机交互界面,通过对自然语言和所陈述的决策问题的理解,更好地实现决策支持所需要的功能,如人机对话、数据存取、机器推理等;2)问题处理系统由问题分析和问题求解组成,包括问题库的管理、调度和求解控制;3)DSS 控制是对库进行管理、协调和维护并在内外部环境起到消息传递的作用;4)专家系统包括知识处理和推理控制,可以将它看成是一个智能部件,与 DSS 数据处理相结合构成 IDSS;5)数

肖春艳 硕士生,主要研究方向为人工智能,知识处理和专家系统。吴 斌 博士生,主要研究方向为人工智能,数据挖掘和群体智能。史忠植 研究员,博士生导师,主要研究方向为人工智能、机器学习和分布式人工智能。

据挖掘是从各库中找出一些特征,并更好地组织各库,另外还可以从 Internet 上查找相关资料。

2.3 系统功能描述

为了给需要决策支持的用户提供远程访问的服务,系统应用了 Internet 网络环境以及先进的 Web 技术,采用一种 Browser/Server 的体系结构。智能决策支持系统的应用程序安装在 Web 服务器上并一直连接 Internet,随时准备为用户提供系统服务;在客户端接入 Internet,并使用浏览器连入 Web 服务器,网络用户可以通过访问本系统迅速获得参考决策,并能将实际结果及时反馈给系统。

决策系统主要有知识管理模块、业务处理模块、个案处理模块和知识挖掘四个模块构成:

(1)知识管理模块:对知识库、范例库、业务规则库等进行管理,提供友好的各库建立工具,制订方便的索引,实现快捷的检索功能;

(2)交互式业务处理模块:推理机根据知识库进行推理,再通过黑板、人机交互生成业务处理建议;

(3)基于范例推理的个案处理模块:通过范例检索、范例复用、范例修正、范例保存等过程实现个案处理;

(4)知识挖掘:从各库中找出一些特征,并更好地组织各库,另外还提供从 Internet 上查找相关的业务管理案例和资料。

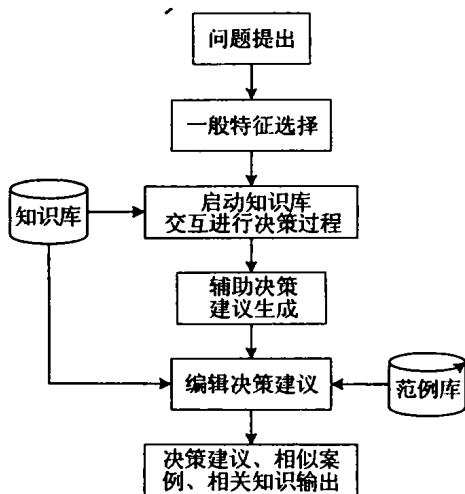


图2

提供决策支持的工作流程如图2所示。其中:1)问题的提出:提出具体的决策问题,选择一定的类别;2)一般特征选择:选择决策问题的一般特征;3)启动知识库、交互进行决策过程:启动待处理的知识库调入到系统中,成为决策支持的前提,与基于案例的推理进行交互访问进行决策;4)辅助决策建议生成:是一个提供决策支持的过程,通过推理机对新建的知识库进行智能判断,并且提出相应的一般性质的决策建议支持;5)编辑决策建议:不同的问题具有其自身不同的性质和特点,用户可以通过查看业务规则库、案例库对相应的业务管理和案例分析做出更为详细的决策建议支持;6)决策建议、相关案例、相关法规输出:输出所处理的问题的决策建议。

3. 系统应用的关键技术

3.1 Web 技术

随着 Internet 这种开放、分布的信息空间的快速发展,相应的 Web 技术也得到了飞速发展。Web,即世界范围的信息

网(World Wide Web),是基于超文本传输协议的客户端/服务器系统,它具有速度快、可维护性好、成本低和功能灵活等特点。从服务器的角度看,每一个 Web 站点是由一台主机、Web 服务器及许多 Web 页组成,形成一个树状结构,每一个 Web 页都是以超文本的格式编写的,它包含各种以文字、图形、声音、动画及超文本链接所组成的信息,借助于一个简单且通用的图形用户界面的 Web 浏览器,用户可以方便地浏览网上的各种不同的信息资源。

目前,Internet 应用正逐渐从以信息发布为主转向信息交流和共享,应用系统的体系结构也从传统的 C/S 结构发展到三层结构到多层结构的 Browser/Server 模式。B/S 模式是一种基于 Hyperlink、HTML、Java 的三级或多级 Client/Server,将应用程序模块放在服务器一侧,其中的复杂数据计算操作都由 Server 上的应用程序实现,客户只需将网页下载到客户端的浏览器来完成各种应用。这种模式由浏览器、Web 服务器、应用服务器、数据服务器四个层次组成,结构图如图3所示。

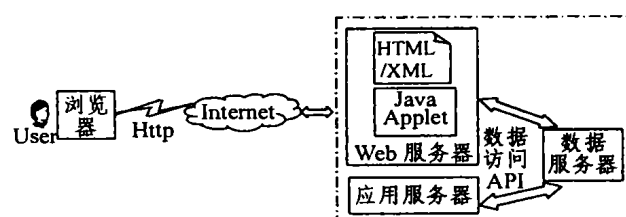


图3

在 Web 的开发和应用中,常用到 Java、PHP、ASP、XML 等语言,它们的特点是面向对象的、多线程的、动态的、跨平台和具有网络安全性。Java applets 可以通过 Internet 传送,并可以在任何可运行 Java 的 Web 浏览器上运行,克服了以往的 Web 仅能发布静态信息的不足,增添了动态特性和交互性,极大地扩展了 Web 的功能和应用领域。而 XML 语言被认为是很有潜力的描述语言,它可以表述数据的内容和结构,创建特殊的数据结构并且有利于结构化和半结构化数据的集成。在 Web 环境下,服务器提供数据库信息,在客户端用 Web 浏览器进行查询。将 Web 与数据库集成的方法可以使用 CGI 技术、专用的 Web 服务器 API 及 JDBC。JDBC 是一组由 Java 语言编写的类和接口组成、可以执行 SQL 语句的 Java 语言应用程序编程接口 API,并通过 Java 程序访问数据库。

采用 Web 技术具有如下优点:

(1)TCP/IP 通讯协议在广域网和局域网上得到广泛使用,已成为工业标准,为异构网络互联提供了平滑的技术支持。

(2)浏览器(Web Browser)作为各种系统通用的客户端软件,给用户提供了统一的图形化界面,减少了用户在学习新软件方面所遇到的困难。

(3)Web Server 上的 Servlet 可以与后端数据库相连,实现动态的网上信息发布。

(4)网页编程语言能够跨平台运行,易于维护。

(5)性能/价格比高。

通过 Web 技术的使用(包括 TCP/IP 协议、统一的浏览器、Java 语言等),可以统一组织和管理企业内部不同部门的不同格式的信息,进而解决企业间以及企业中各部门的信息共享问题,为基于 Web 的决策支持开发和应用奠定了计算机

技术基础。

3.2 专家系统技术

在决策过程中,通过人机交互作用发挥专家的经验、推测和判断,对决策的全过程提供有效的支持,提高了决策支持的及时性和正确性,也充分体现了决策系统的智能化程度。专家系统中,由知识库管理工具按照业务类别建立并维护专家知识库,通过面向对象的推理机启动各项业务处理流程,同时利用建议编辑黑板进行交互。

系统实现可视化知识管理,应用面向对象的知识表达方式将知识库建立在关系数据库基础之上,利用图形接口界面技术来帮助知识工程师构造可视化的知识库。知识工程师将知识对象有序地组织成一种树型结构,树的结构与实际应用中的具体问题密切相关。对知识库的操作接口选用了流行的标准接口:开放数据库连接(ODBC),并且通过 SQL(结构化查询语句)对数据库进行查询、选取、添加等。

构造可视化的知识库系统对于系统的维护十分有益。知识工程师可以方便地添加、删除和修改各类知识,浏览各种对象和属性,赋予初始值,调整对象的相对关系并能够及时更新数据库。在对于知识树中的各个结点的关系和每个结点的方法描述可以通过打印的方式一目了然,从而高效地建立、扩展和维护决策支持系统的知识库。

推理机是指基于知识的推理的计算机实现,它包括推理与控制两个方面。推理控制策略决定了对知识进行选择的方法,对推理效果与效率有重要影响。而面向对象知识表达的推理过程是一个面向知识库系统的、基于假设产生-假设证实机制的、由高层到底层逐步求精的递推过程。其结构如图4所示。

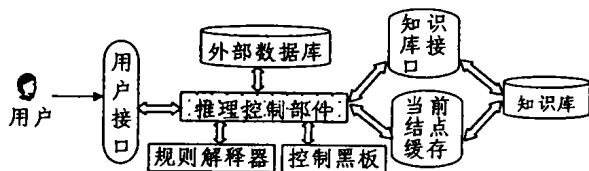


图4

推理机在问题求解过程中应用内存组织的优化方法读取知识库中的知识对象到当前结点的缓存中;利用控制黑板进行推理机和知识对象树中结点的通信;通过规则解释器将对象方法中的过程脚本进行翻译执行,在执行中可以调用各函数和模型,并且通过函数接口对外部数据库进行访问,最终给出业务处理结果的辅助建议。

3.3 基于范例推理技术

应用基于范例推理技术进行个案处理,用于辅助业务处理。在个案处理的过程中,源案例库由已有的案例形成,采用基于范例的推理技术,根据案例类别从对应的源案例库中找到与咨询案例相似的案例,然后把这些案例的处理结果复用到咨询案例中。

一般情况下,由目标范例的特性联想到记忆中的源范例,在最初检索后,进一步检索两个类似体的更多细节,证实它们之间的可类比性和差异,并将获得的源范例集按与目标范例的可类比程度进行优先级排序。选择最优源范例,建立它与目标范例之间的一一对应,利用一一对应关系转换源范例完整或部分的求解方案,从而获得目标范例的完整或部分的求解方案。若目标范例得到部分解答,把结果加到目标范例的初始描述中,从头开始整个类比过程;若求解方案未能给目标范例

以正确的解答,则解释方案失败的原因,且调用修补过程来修改所获得的方案,系统记录失败的原因,以避免以后再出现同样的错误。最后,评价类比求解的有效性,整个类比过程是递进地进行的。

系统中个案处理的具体流程如图5所示。主要模块说明如下:

(1)范例库生成:基于范例推理方法的效率和范例表示紧密相关。范例表示涉及到范例中存放信息选择、范例内容描述结构、范例库如何组织和索引。对历史案件进行分析和整理以后,设计范例库的组织 and 索引方案,生成初始范例库。

(2)范例检索:利用检索信息从源范例库中检索并选择潜在可用的源范例。碰到一个新问题时,首先是从记忆或范例库中回忆出与当前问题相关的最佳范例。由于范例匹配不是精确的,只能是部分匹配或近似匹配,因此要求有一个相似度的评价标准。该标准定义得好,会使得检索出的范例十分有用,否则将会严重影响后面的过程。

(3)范例修正:修正过程的输入是解方案、失败报告及一个解释,然后修改这个解以排除失败的因素。当复用阶段产生的求解结果不好时,需要对其进行修正。修正的第一步是对复用结果进行评估,如果成功,则不必修正,否则需对错误进行修正。

进行结果评估,可以依据它在实际环境中运行后的反馈,也可以通过咨询完成。等待反馈的过程可以通过模拟时间环境来实现。

修正错误一般涉及错误探测和寻找原因。寻找原因是为了对错误进行解释分析,以找出原因,修改造成错误的原因使其不再发生。此外,修改也可以使用领域知识模型进行自修补或由用户输入完成。

(4)范例保存:新问题得到了解决,要把它加入到范例库中,此过程涉及选取哪些信息保留、如何把新范例有机集成到范例库中、修改和精化源范例库,其中包括泛化和抽象等过程。这是学习也是知识获取。

把新范例加入到范例库中,需要对它建立有效的索引,使能够对它作出有效的回忆。可以对范例库的索引内容甚至结构进行调整,如改变索引的强度或特征权值。

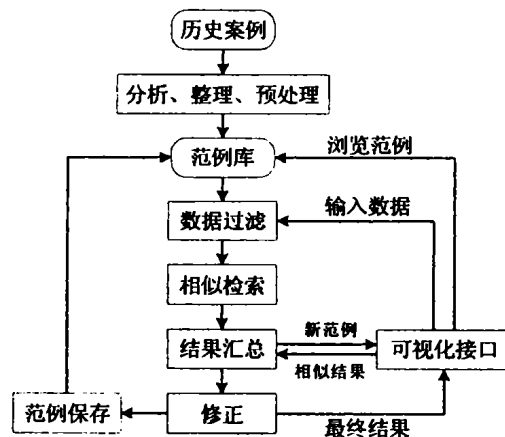


图5

3.4 数据挖掘及知识发现

目前,网上数据资源非常丰富,通过把数据挖掘技术在 Web 上应用,即 Web Mining 技术,对 Web 上蕴含的知识和信息进行充分的挖掘和利用,对决策过程会起到很好的辅助

作用。

数据挖掘是从大量的、不完全的、有噪声的、模糊的、随机的数据中抽取和精化新的模式,也可以称为知识发现(KDD)。1989年,Fayyad 定义“KDD 是从数据集中识别出有效的、新颖的、潜在有用的,以及最终可理解的模式的非平凡过程”。相对来讲,数据挖掘主要流行于统计界、数据分析、数据库和管理信息系统界;而知识发现则主要流行于人工智能和机器学习界。

数据挖掘过程可简单地理解为:数据准备(data preparation)、数据开采,以及结果的解释评估(interpretation and evaluation)。数据挖掘使用的主要方法有关联规则发现、分类规则发现、聚类规则发现、依赖关系发现、预测模式发现等。数据挖掘的范围非常广泛,可以是经济、工业、农业、军事、社会、商业、科学的数据或卫星观测得到的数据。数据的形态有数字、符号、图形、图像、声音等。数据组织方式也各不相同,可以有结构、半结构、非结构的。数据挖掘的结果可以表示成各种形式,包括规则、法则、科学规律、方程或概念网。应用于 Web 的数据挖掘流程如图6所示。

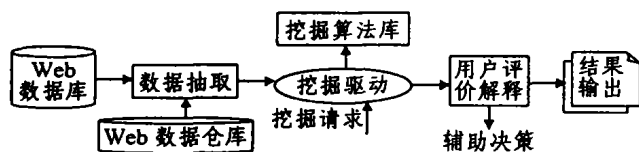


图6

Web 数据挖掘的方法是从 Web 数据库及数据仓库中提取相关数据,其中数据仓库是面向主题的、集成的、稳定的、不同时间的数据集,用以支持经营管理中的决策制订过程。对提取的数据进行二义性分析,消除不一致性,集成数据。挖掘驱动根据挖掘请求在挖掘算法库中选择合适的挖掘算法,并使用选定的算法执行数据挖掘任务。用户评价可以借助专家系统的方法,以一种直观的方式来表现数据挖掘的结果,提供

(上接第139页)

相分离,大大提高了模型的清晰度和模型的可复用性,从而减少在编译器构造过程中的重复劳动;(2)在 GCC 中,对机器描述处理程序处理方式要比 Zephyr 的层次性好。

结束语 GCC 和 Zephyr 编译基础设施提供了相应的编译器重组技术,将编译前端、后端和机器描述各部分相对分离又相互配合,使在编译系统的重组过程中,仅修改或替换与用户研究者有关的部分,大大缩短了编译器的开发周期,从而达到快速地建立高质量的编译器的目的。通过提供支持多源语言多目标编译系统所需解决的三个根本问题的关键技术,从而为快速实现多语言和多平台移植提供了有力的支持。而这些编译器输出可执行代码的质量并没有因为它支持多目标而降低。研究表明^[3],对于多数目标机而言,利用这两种设施产生的编译器输出代码的质量与产品编译器所输出代码的质量差不多。

本文通过从总体结构及其所采用的关键技术进行了分析比较,得出了 Zephyr 所采用的技术比 GCC 具有较强的优势,但 GCC 是自由软件,目前使用更广泛一些。GCC 和 Zephyr 编译基础设施所采用的技术代表着交叉编译技术的发展方向,如何进一步提高编译后端生成的自动化程度,是我们进一步探讨的课题。

一种人机交互的友好界面,将数据仓库、数据挖掘和专家系统结合,是进行综合决策最有效的技术途径。

目前已有一些著名的系统和应用实例,如加拿大 Simon Fraser 大学开发的一个多任务 KDD 系统 DBMiner(目的是把关系数据库和数据采集集成在一起,以面向属性的多级概念为基础发现各种知识),IBM 公司 Almaden 研究中心开发的 QUEST(为新一代决策支持系统的应用开发提供高效的数据挖掘基本构件),Acknosoft 公司开发的诊断和预测系统 CASSIOPEE, HNC 公司开发的信用卡欺诈估测系统 FALCO 等。

结束语 Internet 的出现和普及给企业提供了更新的挑战 and 机遇,这对管理者也提出了更高的要求。面对网络提供的共享信息、快速沟通、相互合作等优点,管理者应能快速、准确地对生产和经营的业务过程作出决策。通过研究和开发基于 Web 的决策支持系统,可以为管理者提供远程决策支持和服务,通过网络为企业的管理和经营活动获得丰富信息资源、案例咨询和决策建议,以更好地满足客户的需求,实现更高的自身价值。

参考文献

- 1 史忠植. 高级人工智能. 科学出版社, 1997
- 2 史忠植, 沈沧海, 李云峰. 智能决策支持系统开发平台 IDSDP. CIMS-CHINA'96, 1996
- 3 陆汝钤. 专家系统开发环境. 科学出版社, 1994
- 4 高洪深. 决策支持系统--理论方法案例. 清华大学出版社, 2000
- 5 Carmel E, Herniter B C. MEDIANSS: Conceptual design of a system for negotiation session. In: Proc. of 9th Conf. on DSS, DSS-89 Transactions, 1989. 239~253
- 6 Kersten G E, Noronha S J. Supporting international negotiations with a WWW-based system: [Interim Report (IR-97-49)]. International Institute for Applied Systems Analysis, Laxenburg, Austria, 1997

参考文献

- 1 Norman R, Jack W D. Machine Description to Build Tools for Embedded Systems. <http://www.cs.virginia.edu/zephyr/>
- 2 Andrew A, Jack D, Norman R. The Zephyr Compiler Infrastructure. <http://www.cs.virginia.edu/zephyr/>
- 3 Manuel E B, Jack W, D. Target-specific Global code improvement: Principles and Application: [Technical Report CS-94-42]. Department of Computer Science University of Virginia. <http://www.cs.virginia.edu/zephyr/>
- 4 <http://gcc.gnu.org/>
- 5 赵克佳, 杨灿群, 罗红兵. 多语种多平台编译系统剖析——GNU CC 的解剖与移植. 1997
- 6 Norman R, Mary F F. Specifying Representations of Machine Instructions. ACM Trans. on Programming Languages and Systems, 1997, 19(3): 492~524
- 7 Norman R, Mary F F. New Jersey Machine-Code Toolkit Architecture Specifications. <http://www.cs.virginia.edu/zephyr/>
- 8 Jack W D, Steve C L, Norman R. Zephyr Code-Generation Interfaces. <http://www.cs.virginia.edu/zephyr/>
- 9 Mark W B, Jack W D. A Formal Model and Specification Language for Procedure Calling Conventions. The 22nd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL '95, San Francisco, CA, January 1995
- 10 Nikil D, Alex N, Hiroyuki T, Ashok H. New Directions in Compiler Technology for Embedded Systems. In: Proc. of the Conf. on Asia South Pacific Design Automation Conf. Yokohama Japan, 2001