

代理型数字参考咨询系统设计

吴云标¹ 粟慧² 葛慧莉³ 周群芳⁴ 徐庆宁⁵

(1. 浙江工商大学计算机和信息工程学院, 杭州 310035; 2. 浙江大学信息资源管理系, 杭州 310028;

3. 浙江省科技信息研究院, 杭州 310007; 4. 浙江工业大学之江学院, 杭州 310024; 5. 浙江工业大学图书馆, 杭州 310032)

摘要 新千年以来, 网上数字参考咨询系统是图书情报领域研究热点之一, 但侧重于助理型。本文介绍一个基于咨询经验的启发式网上代理型系统, 着重探讨了与提高系统实用性及咨询知识库知识的权威性有关的问题, 如开放构架, 知识评审机制。

关键词 数字参考咨询 咨询知识库 代理型 开放构架 评审机制

A Design of DRS Surrogate System

Wu Yunbiao¹, Su Hui², Ge Huili³, Zhou Qunfang⁴ and Xu Qingning⁵

(1. School of Computer & Information Engineering, Zhejiang Gongshang University, Hangzhou 310035;

2. Department of Information Resource Management, Zhejiang University, Hangzhou 310028;

3. Zhejiang Institute of Science & Technology Information, Hangzhou 310007;

4. College of Zhijiang, Zhejiang University of Technology, Hangzhou 310024;

5. Library of Zhejiang University of Technology, Hangzhou 310032)

Abstract Since 2000, online digital reference system is one of hot topics in library and information science, but focuses on assistant system. This paper introduces a online heuristic surrogate system which are reference service experience-based. The discussion focuses on improving the practicability of the system and the validity of reference knowledge base, e.g. open structure, and the appraisalment mechanism of the reference knowledge.

Keywords digital reference services, reference knowledge base, surrogate system, open structure, appraisalment mechanism

1 系统设计方案的提出

数字参考咨询系统是一个利用计算机技术和通信技术帮助解答或直接解答图书情报机构参考咨询问题的系统。从不同的角度, 可以对其作出不同的分类。

根据系统对咨询事务的帮助程度, 可将其分为助理型与代理型。助理型数字参考咨询系统对图书

情报人员的咨询工作起辅助作用, 如帮助用户与咨询员之间建立起必要的沟通条件, 使咨询事务能顺利进行, 或者是帮助咨询人员提高咨询效率等, 其特征是不能离开咨询员独立工作, 只是在咨询员答疑过程中起到辅助咨询员的作用。这类系统也可称之为“有人值守型”数字参考咨询系统。与助理型相对应, 代理型的特征是系统独立工作, 不需要咨询员的参与。任何一个用户咨询时, 只与系统发生关系和进行交互, 对任意一次咨询事务, 没有咨询员在旁,

收稿日期: 2006年3月31日

作者简介: 吴云标, 男, 1958年生, 浙江工商大学计算机和信息工程学院副教授, 研究方向: 信息管理。粟慧, 女, 1960年生, 浙江大学信息资源管理系副教授, 研究方向: 信息管理。葛慧莉, 女, 1969年生, 浙江省科技信息研究院副研究馆员, 研究方向: 信息资源开发、用户信息服务。周群芳, 女, 1975年生, 浙江工业大学之江学院图书馆馆员, 研究方向: 信息管理。徐庆宁, 女, 1958年生, 浙江工业大学图书馆副研究馆员, 研究方向: 信息咨询与信息利用服务。

也不需要咨询员操纵或干预系统,咨询事务的开始、进行、结束完全由用户自己决定。这类系统也可称之为“无人值守型”数字参考咨询系统。

如果根据系统的工作机制来划分,可以分为启发式系统和非启发式系统,两者的区别在于其工作机制是否建立在启发式原则之上。而启发式系统根据启发式的表现方式和目标又可以进一步划分,启发是针对系统还是针对用户的。如果针对系统,其启发式原则是完全隐含的,用户在使用过程中不会感觉到,也不需要感觉到。而针对用户的启发是用户通过与系统交互,获得帮助,因而也称之为交互型。启发式系统的另一种划分标准是启发所依赖的基础,是基于咨询员的经验还是基于咨询提问的语言学特征(如根据词语特征、句法特征、语言学规律等)。

新千年以来,数字参考咨询服务的研究和实践不断深入和发展,特别是网上参考咨询系统的研究开发呈现热潮,从相关研究文献可以看出,网上参考咨询系统研究基本集中在基于国际互联网的助理型系统上,其主要问题是提问处理的延迟性和服务成本较高^[1]。

代理型数字参考咨询系统采用“无人值守”的咨询服务方式,突破了现有数字参考咨询服务的形式^[2]或形态^[3]。系统一旦提供网上服务,服务的主导权就交于了用户,什么时候接受服务,什么环境下接受服务,乃至接受服务的程度都可由用户自由决定。在此意义上可以说充分体现了用户至上的服务理念,极大程度地提高了参考咨询服务的灵活性。其全天候服务特点不仅与网络环境下用户的学习生活习惯相适应,也可使图书情报人员摆脱简单或常规问题的咨询服务,而致力于深层的服务。因此图书情报机构如能将网上无人值守咨询服务与有人值守咨询服务结合起来,有效地节约人力财力,有利于提升咨询服务竞争力。

本文介绍一个基于咨询经验的启发式网上代理型系统原型,用户通过与系统的交互获得启发式引导和帮助从而完成所需信息或答案的查找。

2 系统简介

系统由管理子系统、交流子系统、服务子系统组成。

管理子系统实现对系统的全面管理,包括人员和知识管理,以及相关监测统计分析,用于保障系统

正常运行,提高服务效能。

交流子系统是会员咨询知识的交流平台,主要为咨询员提供讨论咨询知识及其制作的平台,为会员开辟了“后台通道”,同时为咨询知识块进入咨询知识总库前提供公示场所。

服务子系统实现图书情报机构的参考咨询服务,其启发式答疑服务功能模拟图书情报人员解答咨询,帮助用户解决信息查询中的问题。包含三个服务子模块:启发式服务子模块、FAQ 服务子模块、有人值守咨询服务子模块,见图1。需要指出,本系统的有人值守咨询服务功能,只是一个简单的服务原型,目的是实现无人值守服务系统与有人值守服务系统的结合,限于篇幅,本文不作详细讨论。

图1中FAQ库表示常见问题解答知识库,RR库表示启发式答疑知识库,PDRS库表示有人值守咨询知识库。其中RR库为通用型、可扩展,当前存有1100多个页面,相当于上千条规则(即因果关系。规则数超过页面数,见第4节介绍)。

对于WEB服务,如何提高用户体验效果至关重要。本系统在设计过程中认真考虑了这一问题,采用了以下方法:(1)减少硬件环境限制,对客户端硬件配置无特殊要求,也不需要下载安装任何软件、插件或进行特别配置。(2)方便用户,不需注册、认证便可使用服务。(3)多用户响应,服务基本上不受同时进入系统的用户数量的影响。(4)提供多咨询答疑入口,有57个入口作为交互起点(图1中的信息源类型只列了大类),方便用户根据不同情况选择。(5)提供搜索方式作为补充。搜索时,通过下拉列表,可以对“信息需求”、“已知条件”、“资源类型”进行限定,缩小范围,以便快速获得所需帮助。搜索界面采用了AJAX技术^[4],有类似“Google suggest”^[5]即搜索提示的效果。(6)结合有人值守咨询服务方式,使咨询服务更完善。用户在启发式咨询服务下不能解决的问题,则将被直接指引到有人值守的咨询服务。

3 启发式答疑系统知识库

系统知识库涉及常见问题解答知识库、启发式答疑知识库与有人值守咨询知识库,启发式咨询服务系统与前两者有直接关系。这里重点介绍启发式咨询知识库。启发式答疑咨询系统的知识库系统逻辑结构见图2。用户通过人机交互模块获得咨询知识库的帮助,咨询知识库中的知识指导用户如何获

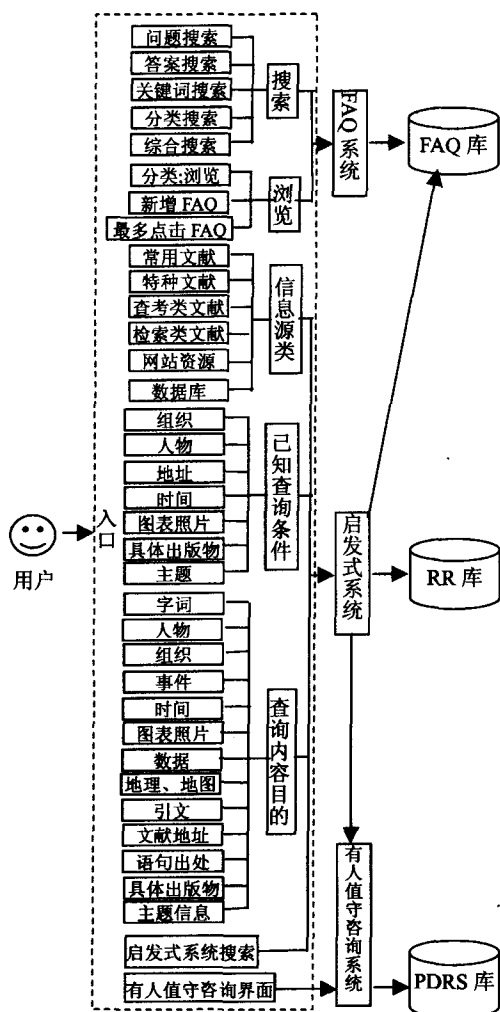


图1 服务子系统的主体结构

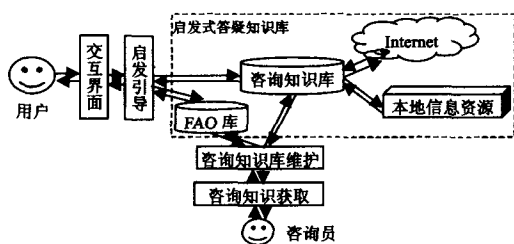


图2 启发式答疑知识库系统整体结构

得问题的答案或相关文献。在提供用户交互过程中,系统具有演绎功能,用户通过查找和推理选择(由系统给出提示)获得帮助。系统的知识接受、更新则由咨询专家通过知识获取模块实现。

图2所示启发式咨询系统的知识库结构与通常的代理型参考咨询系统的知识库结构^[6]有一个明显的不同,这里的咨询知识库中并不包含任何咨询问题的答案,也不包含任何文献信息资源,它是一个帮

助查询的方法与经验的集成库。系统的知识库(虚线部分)是一个包含咨询知识库、FAQ库、本地信息资源库、国际互联网的动态库。图2中的FAQ库是常见问题解答服务子系统的FAQ库中的一部分。

对于代理型启发式参考咨询系统的知识库,知识更新与权威可靠是至关重要的。进入互联网时代,人们面临着信息资源急剧增加,载体形式多样、交流向传输速度快密度高的通道聚集、资源地址多变,用户使用习惯个性化等诸多变化,所有这一切对咨询服务提出了更高的要求,因此咨询知识库就必须具备动态适应性。开放构架是加强和提高系统动态适应性的有力措施,而且在开放条件下,启发式答疑咨询知识库的大小可以不受限制,可以容纳来源于任何人的查询方法与经验(在通过评审之后)。

具有更新能力的咨询知识库并不意味着对进入的咨询知识可以不作要求,为了保证在被咨询知识库接受以前其知识准确可靠的,并真实地再现咨询专家的咨询过程,本系统特别建立了知识评审机制,以提高入库咨询知识的权威性和可信度。

因此,开放式构架是本系统设计的基本要求,评审制度是本系统咨询知识库的知识可靠性、权威性的保障。下面分别介绍其实现要点。

4 咨询知识库的结构

咨询知识库的宏观结构由四大分支构成:咨询问题类型、已知条件类型、信息资源类型、检索技巧类型,这是咨询知识库的基本层次。每一分支又分成若干类型,称为从属层次。“咨询问题”按用户要查找的内容或目的区分,分成13个类,“已知条件”当前列了7种类型,信息资源类型分为6大类31小类,“检索技巧”按其应用范围划分为6大类。详细内容参见文献^[7]。

这里重点讨论咨询知识库的微观结构。

基于规则的知识表示是在产生式的基础上发展起来的,由于它与人类专家表示知识的方法十分相似,所以是应用最广泛的知识表示方法之一^[8]。

当用户想要知道某个缩略语的含义时,咨询员会向用户推荐查询缩略语词典;当用户要想知道某个地名的来龙去脉,咨询员会向用户推荐查询地名词典等。所以从微观上看,咨询专家的参考咨询知识或经验,可以方便地被描述成IF-THEN格式。

将IF-THEN格式的规则存储在计算机中,通常可使用以下方式:

RULE(No, FatherNode, SonNode, Condition, ...)

其中, No——规则编号, 它是推理机进行推理的索引号; FatherNode——父节点, 即应用该规则前的状态; SonNode——子节点, 即应用该规则后的状态; Condition——规则成立条件; 省略号表示其他可能的项。

但是, 这种表示方式虽然能被咨询专家所理解, 却不符合人们的习惯。所以在经典的专家系统构造时, 需由知识工程师对专家知识进行转化并输入。这种模式不适合于系统的开放构架。

本启发式答疑系统的设计目标并不采用自动推理机, 其启发的对象是用户。如果在一个节点下有多个规则可以选择, 系统就提供了推荐选择, 但决定权在用户, 用户可以根据该节点的说明文字, 以及自身的情况来判断并作出选择。所以向用户提供清楚的应用规则前后的状态, 对启发效果来说显得更为重要。每一个状态节点对应于一个面向用户展示的页面, 因此本系统以页面作为储存的基本单位, 而不是以推理规则作为储存的基本单位。

所用的页面储存方式为:

PAGE(pageNo, content, ...)

其中, pageNo——页面编号, 它是页面转换的索引号, 对应于上面的节点; content——页面内容; 省略号表示其他可能的项。

页面内容其实包含了该页面节点派生出来的所有规则, 即到其他子节点的链接。它将前面的规则转换成页面转移, 所以不需要咨询专家将“条件——结果”, 或者“现象——措施”等作“如果——那么”解, 而只需要理解成“当前页面(状态)经某个链接转到下一个相应页面(状态)”, 这是很自然的。如图3所示页面:

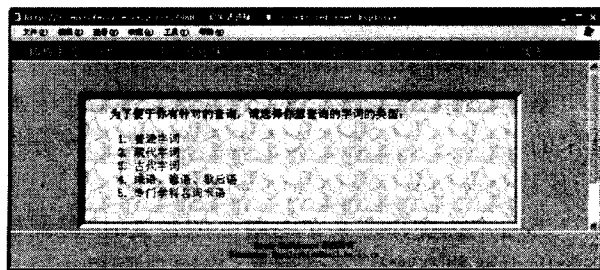


图3 查字词

图3是“查字词”的页面, 按查找目的的不同, 可以将字词的类型分为五类: (1)普通生词; (2)现代生词; (3)古代生词; (4)成语、谚语、歇后语; (5)专门

学科名称术语。这五种查找的类型相应地链接到下层不同的页面。编写这样的页面对于参考咨询员来说是不太难的, 因为它并不要求咨询员为了系统的内在表示机制而改变其思维习惯。

咨询知识库的宏观结构与微观结构使咨询知识库从咨询人员或专家直接获取知识而不必受知识工程师的干预或帮助成为可能。

5 咨询答疑知识制作

咨询库的微观结构为咨询员或专家向库中输入知识提供了必要条件, 但不是充分条件。为此, 本系统利用一个良好的人机界面, 通过人机对话, 在键盘和屏幕的输入/输出交互方式下, 对知识库进行增删、修改、扩充和更新。这种方法较为简便, 容易被理解和掌握^[9]。

本系统为每一个有权创建知识块的咨询人员提供了一个独立的建设平台, 其功能主要有答疑知识制作、闭合性检验与知识提交。所谓知识块是指一组相关页面, 存在且只存在一个首页, 其余页面均可以看成是首页的派生页面。页面由知识编辑器编辑完成。知识编辑器包括页面添加、修改、删除以及页面浏览查找。其中的添加功能界面见图4。

图4中一大片空白区域为页面内容输入编辑框。咨询答疑知识块制作将页面类型分成三种: 检索技巧类、信息资源类、条件内容类。这三种页面的制作方式略有不同。

向咨询库添加知识被处理成在树的某个节点上生长出新的分支, 分支用知识块实现。任何一个咨询员, 如果认为知识库的某个地方需要扩展, 只要记下这生长点的页号, 自行编写分支知识块就行。

在提交知识块以前必须先通过该知识块的闭合性检验, 其中知识块的块内链接检查由机器自动完成。知识块的块内链接检查, 是指检查当前知识块中的块内链接情况, 即知识块中的页面相互之间的链接。此外, 块内检查还包括检查该知识块所有页面的页面类型是否一致等。这些都由机器自动完成, 不用人工干预。

构建知识块的重点是确定核心路径图, 以勾画出该知识块包含的主要内容和链接顺序。比如, 已知人物姓名查与该人物的有关的信息, 其核心路径图, 见图5。

图5中的“技巧”是指相应环境下的检索技巧。该核心路径图表示, 已知人物名称查与该人物有关

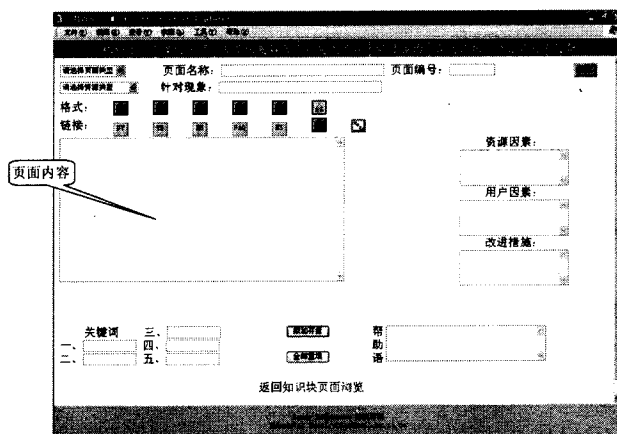


图4 知识编辑器的添加功能界面

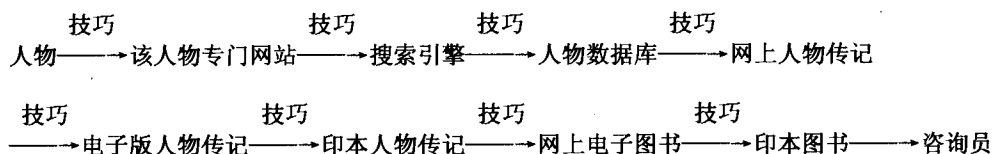


图5 已知姓名查该人物的核心路径

的信息,首先推荐查询该人物的人物网站,查人物网站涉及网站搜索技巧;以此类推,如果用户的问题还是得不到解决,最后将引导用户转向有人值守咨询服务。

核心路径图的顺序是咨询知识和经验的反映,但是没有一条路径能适合于所有用户的信息需求,即对所有用户的信息需求都是最佳的路径是不存在的。确定顺序时应考虑如下几个原则:(1)资源越权威越靠前;(2)页面的利用率越高越靠前;(3)页面越简洁的越靠前,牵涉的页面数越多越靠后;(4)用户检索知识越多,所走的路径应越短等。当然这些原则有时是相互矛盾的,落实起来有相当难度,这里不作讨论。由此不难明白,获取启发式咨询知识是一项工作量很大、工序繁杂而又细致的专业劳动。有时需要通过反复的应用和不断的修正,才能逐步趋于完善。

6 咨询答疑知识评审

任何咨询员制作的知识块都是临时知识块,不直接进入系统的咨询知识库。凡完成的知识块都需作为临时知识块提交审定,评审合格后方可转入知识库。

评审过程主要为两段。由咨询答疑制作模块提交的知识块首先经过一个高级咨询员的初步审阅,由其决定提交的知识块是退回还是接受。不合基本要求的将被退回,如果接受,则让其进入公示状态,相当于让会员试运行。经过一定时间的会员试运行、评论,如果有修改意见,则由负责审定该知识块的高级咨询员进行修改,这一段过程可能反复,然后决定该知识块是否批准录用。

实时咨询时,咨询人员容易因时间或记忆的限制而使咨询结果的周密性受损,而本系统中咨询员在设计知识块时,对各种可能性可以经过周密的考虑,从而保障所作指引的准确、合适,并且通过评审修改,进一步提高其正确性、权威性。

咨询经验往往是隐性知识,很难转为显性知识。评审机制有利于图书情报参考咨询人员和专家的专有知识得以正确显化。保存在咨询知识库的知识和经验,不受时间和空间的限制,有效避免了因工作调动或退休而流失,使参考咨询技巧和经验得到总结和精练。有了评审机制,有利于防止咨询知识的重复生产。加入系统的热心咨询员(职业的或志愿的)越多,富有经验的高级专家越多,咨询知识库的实用性就越好,正确性、权威性越有保障。

7 未来的工作

本系统设计虽然遵循参考咨询范式,充分利用了各种相关研究成果,但作为真正应用咨询经验的网上代理型启发式咨询系统是一次全新的尝试,许多工作有待进一步深化:

(1)当前国内有人值守的参考咨询系统经过这几年来大量研究开发,已经取得了不少成绩。如国家科技图书文献中心参考咨询系统,直接应用到国际互联网,用户反映良好^[10]。CALIS 分布式联合虚拟参考咨询系统(CVRS)将正式进入市场化销售渠道^[11]。所以正如有专家指出的,如何将有人值守的参考咨询服务与无人值守的参考咨询服务结合起来,实现1+1大于2的目标;这是可能的。

(2)系统设计了用户点击路径跟踪等相关监测统计分析,这将有助于评价咨询知识的有用性。但可能需要有一定的积累,才能分析数据,修改系统,增加功能,逐步完善监测体系。

参 考 文 献

[1] Marc Meola, Sam Stormont. Starting and Operating Live

Virtual Reference Services. Neal-Schuman Publishers, 2002.

- [2] 黄敏, 杨宗英. 网上咨询服务的主要形式与发展趋势. 大学图书馆学报, 2003(1).
- [3] 王梅. 论数字参考服务. 中国图书馆学报, 2003(5).
- [4] Dave Crane, Eric Pascarello, Darren James. Ajax in Action. Manning Publications Co., 2005.
- [5] <http://www.google.com/webhp?complete=1&hl=en>.
- [6] Craig McDonald, John Weekert. Libraries and Expert Systems. Taylor Graham, 1991.
- [7] 吴云标, 栗慧, 葛慧莉, 周群芳, 徐庆宁. 无人值守参考咨询服务研究. 中国图书馆学报, 2006(3).
- [8] Janet F. An introduction to artificial intelligence. Janet Finlay and Alan Dix, London: UCL Press, 1996.
- [9] 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. 北京: 北京出版社, 1997.
- [10] 杨和东, 吕世灵, 张志平. 网上实时参考咨询系统的设计与实现. 情报学报, 2006(1).
- [11] <http://www.lib.sjtu.edu.cn/chinese/virtual-reference-desk/CVRStbirdmeeting.doc>.

(责任编辑 王建平)