

IACA—智能化农业知识获取平台的设计与实现

余锋林^{1,2} 王儒敬¹

¹(中国科学院合肥智能机械研究所 安徽 合肥 230031)

²(中国科学技术大学自动化系 安徽 合肥 230026)

摘要 IACA(the Intelligent Agricultural Knowledge Acquisition Platform)是一种农业领域的智能化知识获取平台,为了实现农业知识获取及可视化,阐述了 IACA 的思想和基于面向对象技术的系统设计和实现方案,并采用基于“以用户为中心的设计”方法设计,利用拟人知识化、可视化的技术实现了 IACA 平台。

关键词 知识获取 知识地图 知识表示 XML

DESIGN AND REALIZATION OF THE INTELLIGENT AGRICULTURAL KNOWLEDGE ACQUISITION PLATFORM

Yu Fenglin^{1,2} Wang Rujing¹

¹(Hefei Institute of Intelligent Machines, CAS, Hefei 230031, Anhui, China)

²(Department of Automation, University of Science and Technology of China, Hefei 230026, Anhui, China)

Abstract IACA is an intelligent agricultural knowledge acquisition platform. The key points of IACA are described, and the object-oriented system is designed and implemented based on the User-Centered Design method and the techniques of personification and visualization.

Keywords Knowledge acquisition Knowledge map Knowledge representation XML

0 引言

知识获取技术已经发展多年了,在理论和实践上已经逐步形成了一些方法和理论。2005 年国际知识和信息可视化大会,力图把在各自领域独自发展的信息可视化和知识可视化技术融合在一起。会上多位学者使用概念图相关技术作为连接它们的桥梁,从某种意义上为可视化知识获取奠定了基础。概念图自问世以来,经过了从纸笔到计算机工具的发展历程。在这个过程中,许多学术机构研究开发了概念图设计软件工具,如 CMap^[1]、Webster^[2]等。其中 IBM 沃森研究中心设计的 Webster^[2]是基于网络的概念图工具,允许用户灵活地使用多种名词术语来将知识和信息表示成结构化的概念图。在这种结构下,Webster 简化了知识构建的过程和抽象知识的多层次展示,便于用户实现可视化的知识认知表示结构,有助于用户的理解和学习。这些知识获取工具都是在概念图基础上发展而来的,大多集中在概念性的知识,以语言形式的文本格式来表示知识。从知识共享复用的角度来说,知识库应当尽量以结构化的形式来表示,其接口也应该很容易获取。另外,大多数知识获取工具缺乏实时的知识检验求精活动;如果知识工作者编辑知识时,在实时检验系统帮助下,能及时地发现一些比较隐蔽的错误,这样就会大大降低知识库的维护成本。为了实现知识获取及可视化,本文以科技部农业成果转化基金项目“智能化农业信息处理开发平台中试与应用”为背景,采用拟人化、知识可视化的技术设计实现了一个智能化农业知识获取平台——IACA。

1 工具体系结构设计

1.1 设计原则目标

计算机系统中的人机界面是研究人、计算机以及它们之间相互影响的一门综合学科,其涉及到计算机硬件、软件、网络、认知心理学等多门学科,是人与计算机之间传递和交换信息的媒介,是计算机系统向用户提供的综合操作环境。近年来,国际上已广泛采用“以用户为中心的设计”方法^[3]。该方法已被国际标准化组织(ISO)作为正式标准——以人为中心的交互系统设计过程而发布。这种方法的主要特征是用户的积极参与,对用户及其任务要求清楚了解;在用户和技术之间适当分配功能;反复设计解决方案;其主要设计活动是了解并确定使用背景;确定用户和组织要求;提出设计方案;根据要求评价设计。

其中按照用户行为特性,可以把用户分为 3 类,定义如下:

- 1) 低级初学型 对新的系统不熟悉,但对计算机还有些经验的用户。
- 2) 中级熟练型 对系统有相当多的经验,但不具备扩展系统的能力。
- 3) 高级专家型 具备对系统如何操作的良好知识以及维护、修改基本系统的能力。

设计的目标是希望能够满足不同层次的用户快捷地进行知

收稿日期:2006-07-31。国家 863 高科技发展研究计划资助项目(2003AA118070)。余锋林,硕士生,主研领域:知识获取, WAP 综合推理机。

识获取;减少用户记忆负担、学习操作的时间,简化操作方法等,使所设计的人机界面能尽可能地减少人的认知负担。

1.2 系统结构设计

IAKA 是以一种以多种工具为核心,为高效率开发知识库而设计和实现的大型智能计算机软件系统,加上与之配套的各种辅助工具,形成一种集成化知识获取工具包。其支持的功能包括:推理过程、数据存取、自然语言对话的界面、过程的界面、解释功能等。按设计目标要求,工具应该支持不同层次的三类用户。为达到预期的设计目标,IAKA 必须构建成为一个开放的易操作的系统,为提高用户的设计和开发效率,极大简化知识获取工作,缩短研制周期提供系统设计和开发的必要环境。

笔者参与了科技部农业成果转化基金项目“智能化农业信息处理开发平台中试与应用”整体架构设计,完成单机各个子系统的设计和开发。结合本论文的特点,从功能模块来划分,工具的内部结构如图1所示。

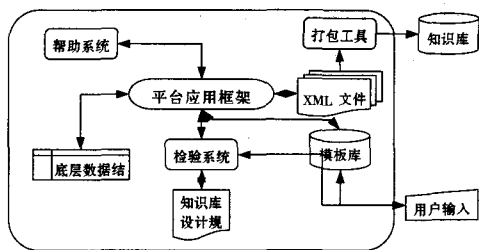


图1 IAKA 整体框架

用户进行的知识获取并绘制成知识地图的工作,主要是将外界的信息进行可视化编码,大致过程可以分为三部分^[4]。首先,将初始数据通过分析计算转成抽象的概念;其次,将抽象的概念转成可视化的概念;最后,将得到的可视化概念转成对应的几何数据,并映射到显示屏上。在知识库生成的过程中,用户的每一步操作都要经过检验系统的确认;在知识库生成的最后阶段,检验系统还会完成知识库求精任务,以保证最终所得知识库的一致性与正确性。校验系统主要的工作如下:

- 1) 用户所定义的变量是否符合平台命名法则的校验、修正;
- 2) 用户所开发的知识库是否符合平台 XML Schema 以及产生式规范;
- 3) 用户设计的框架是否符合平台框架设计的要求;
- 4) 在维护过程中友好地提示用户进行正确的操作,如果操作错误则返回重新操作;
- 5) 最后对整个知识库进行最终求精检验。

2 使用工具进行可视化知识获取

2.1 开发模式选择

如图2所示,IAKA 为用户提供了三种开发模式,而且三种方式可以相互切换操作。

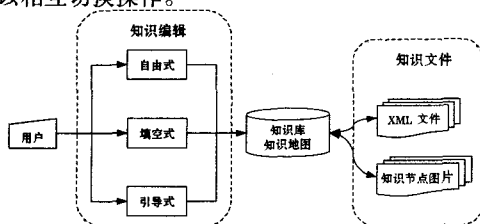


图2 知识获取编辑模式

1) 自由模式

在这种模式下,用户可以自由地建立各级知识单元,而不必非要按照某种次序进行设计;在设计过程中,单元之间的关系可以随时调整。对规则知识单元进行产生式规则编辑时,用户可以直接按照 EBNF 范式的规定直接编写、修改产生式规则脚本语言。

2) 填空模式

使用填空模式进行知识获取时,用户在智能引导的方式下通过与系统问答式的会话过程进行系统构建。在设计过程中,用户的每一步操作都会在系统的提示下进行。与自由模式不同,各级知识单元的建立有一定的先后次序,且它们之间的关系在建立的过程中也随之确立。对规则知识单元编写产生式规则时,用户可以利用事先设计好的编辑框和下拉列表进行编写、修改产生式规则脚本,即进行“填空”操作。

3) 引导模式

IAKA 工具中存储有一些知识库的基本框架模型,用户从模型库中挑选一个比较合适的框架模型,然后按照实际需要原有框架基础之上对知识单元进行删减,或者修改知识单元的具体属性,改写产生式求解规则。在这种模式下,用户不必从零开始,而是直接利用已有的知识库模型,这样既节省了开发时间和工作量,又达到了知识共享与复用的理想效果。

2.2 知识可视化的实现

IAKA 的知识地图可视化结构如图3所示。

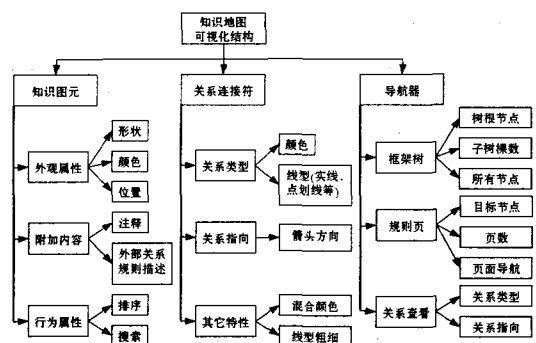


图3 知识地图可视化结构

1) 知识单元

知识图元是知识单元的图形化,IAKA 中的知识图元有两种:框架知识图元和规则知识图元。它们有一些共同的属性,如形状、外观、颜色等。同时,它们还有自己的独特属性,例如框架知识图元会有全局关系描述;而规则单元则可能有自己的规则描述,可以通过规则查看器来浏览规则。

2) 关系连接符

知识图元之间的连接线反映了知识单元之间的局部关系。有的开发工具(如 CMap)在关系连接线上带上关系标签,如果节点数太多的话,会给用户带来视觉上的混乱,造成节点空间布局困难度大大增加。鉴于这种情况,IAKA 采用特定的线型和特定的箭头符号的组合来表示不同的关系。这样,既在节点空间布局上有所缓和,更是在内存空间上有很大节省。

3) 导航器

导航器使得用户能够在不同的知识地图页面之间快速地切换。针对分页式知识地图的特性,IAKA 主要使用了三种导航器:框架树导航、规则页导航和关系查看器。框架树导航器中包

(下转第 28 页)

```

F: =FU{(vi,ci),(ci,vj)};
end
else if Idg(vj) > 1 then
/* Idg(vj)表示 G 中 vj 的入度 */
begin
C: =CU{cj};
F: =FU{(vi,cj),(cj,vj)};
end;
end
C: =CU{c_start,c_finish};
F: =FU{(c_start,Entry(G)),
(Exit(G),c_finish)};
M: ={c_start};
end.

```

7 结 论

软件工作模型的变化能够在大部分的软件演化实践中得到体现,尤其是在软件工作流程发生改变这样的演化中最为明显。建模者只需要构建出演化后理想的工作模型作为目标,使得软件的演化参照这个目标来进行。本文提出的通过模型对比来获得软件演化过程模型的方法,使建模者不必为繁琐的演化细节耗费过多精力,而把注意力集中到对需求的分析以及对系统功能的理解上,这样能够更好地确保演化之后的软件质量。分层的模型结构为模型的建立和转换带来了一定的复杂性,所以选择适当的粒度对模型进行分层是很重要的。

参 考 文 献

- [1] Lehman M M, Ramil J F. Software Evolution and Software Evolution Processes. *Annals of Software Engineering*, 2002, 14(1-4): 275-309.
- [2] Ramil J F, Lehman M M, Kahen G. The FEAST Approach to Quantitative Process Modelling of Software Evolution Processes. *Proc. PROFES 2000 2nd International Conference*.
- [3] Osterweil L J. Software Processes are Software Too. *Proceedings of the 9th International Conference on Software Engineering*, Monterey, CA, March, 1987; 2-13.
- [4] *Proceedings of the conference on The future of Software engineering*, May 2000; 25-34 Limerick Ireland.
- [5] Li Tong, Yang Hongji. A Language and Approach to Modelling Software Process for Software evolution.
- [6] Bernardinello L, De Cindio F. A survey of Basic Net Models and Modular Net Classes. LNCS vol. 609, Springer Verlag, 1992.
- [7] 周之英. 现代软件工程. 第二册. 科学出版社, 2000.

(上接第6页)

含了所有的主框架树和子框架树,位于屏幕的左侧。规则页面应对为一个具体的问题求解推理有向图,规则页导航器将所有的规则页以列表的方式显示在屏幕左侧。用户通过操作规则页列表即可转到相应的规则页面。关系查看器同规则页导航器类似,显示了节点之间的关系列表以及关系的起始和结束。

2.3 获取知识并绘制知识地图

根据知识获取过程中领域专家参与的程度,使用 IAKA 生成知识库可以分为三种方法:(1)知识工作者收集科技文献,整理相应应用领域的典型应用场景,进行结构、逻辑抽象,生成系

统骨架。(2)知识工作者与领域专家协作,由领域专家进行系统分析,由知识工作者来实现系统框架。(3)知识工作者利用现有的框架,或先对应用场景抽象分析,建构系统框架雏形,然后与领域专家合作进行反复的审核修正,最终完成知识库系统。其中方法(3)较好地综合了知识工作者和领域专家各自的优势,使用得比较多。采用这种方法进行知识获取的一般步骤为:

- 1) 收集数据信息,对数据进行系统抽象、提炼,得到一些基本的知识单元抽象概念;
- 2) 从应用系统模型库中选出一个最适合的构架模型,建立一个知识库结构的雏形;
- 3) 根据知识单元抽象概念绘制知识地图,确立框架结构,规则求解页等,形成可视化的知识库结构;
- 4) 反复地进行系统构架的评估修正,包括处理逻辑结构的完善性评估、修正,以及逻辑结构、计算处理的参数抽象数据的充分必要性分析、修正等;
- 5) 如果已经将收集和提炼的知识内容完全都展现在知识图上时,就将当前最小的知识单元作为叶节点;
- 6) 修剪知识地图,将那些看似不相关的内容去掉,并补充其它的必要知识;从整个知识图的角度,修正各个知识节点的状态属性、行为属性和描述属性,以及节点间的关系。

3 可视化知识获取工具的特点

IAKA 是一个面向对象体系结构的开放性系统开发平台,在几个农业示范区得到了很好的应用,它的主要技术特点包括:

- 集成化多级知识单元的知识表示方法;
- 可视化的知识获取方法;
- 结构化的知识库文档;
- 以用户为中心的拟人化开发环境。

4 小 结

为了实现知识获取及可视化,本文采取面向对象的程序设计技术,并利用拟人知识化、可视化的技术设计实现了一个智能型农业知识获取平台 IAKA。该工具是基于“以用户为中心的设计”方法设计的,主要特征是用户的积极参与,对用户及其任务要求清楚了解;在用户和技术之间适当分配功能;反复设计解决方案。在科技部农业成果转化基金项目“智能化农业信息处理开发平台中试与应用”中得到了很好的应用;利用该工具已经成功开发出了多个综合型知识库,在几个 863 农业示范区得到了很好的推广。

参 考 文 献

- [1] Canas A J, Hill G, Carff R. CmapTools: A knowledge modeling and sharing environment. *Proceedings of the first international conference on concept mapping*, 2004, 1: 125-133.
- [2] Sherman R Alpert. *Comprehensive Mapping of Knowledge and Information Resources: The Case of Webster*. *Knowledge and Information Visualization*, LNCS 3426, 2005: 220-237.
- [3] Vredenburg K, Isensee S, Righi C. *User-Centered Design: An Integrated Approach* [M]. New Jersey: Prentice Hall, 2001.
- [4] Ed H Chi. A Taxonomy of Visualization Techniques using the Data State Reference Model. *Proceedings of the IEEE Symposium on Information Visualization*, 2000.