

●李 敏 张玉峰

基于多知识库协作的信息导航模型研究*

摘 要 信息导航系统需要类型多样、操作灵活的知识库体系提供知识支持。这些知识库有:用户知识库、专家知识库、信息特征知识库、领域知识库、常识库、元知识库和语言知识库等。利用外部指针和黑板式协作模式,可建立一个多知识库协作的信息导航模型。图2。参考文献7。

关键词 信息导航 知识库体系 知识库协作 信息导航模型

分类号 G350

ABSTRACT An information navigation system requires the knowledge supports of diversified and operable knowledge base systems. In this paper, the authors introduce the types of knowledge bases and an information navigation model. 2 figs. 7 refs.

KEY WORDS Information navigation. Knowledge base system. Knowledge base cooperation. Information navigation model.

CLASS NUMBER G350

理想的信息导航涵盖两方面的内容:一是对信息资源所蕴含的知识内容进行剖析、组织,使其能够提供符合用户多样化需求的导航模式;二是在已有的导航知识的指导下进行信息查询,在整个导航过程中,所有导航行为离不开知识的支持。其中尤为重要的是各类知识的导航指导作用。未来的信息导航系统将借助于储量巨大、类型多样、便于灵活操作的知识,使整个系统的运作呈现出高度的智能性特征。因此,如何合理地存储、组织、管理、应用和共享知识,以及进行知识库的合理构建就成了信息导航系统深入发展所面临的重要问题。

1 信息导航系统的知识库体系

1.1 信息导航系统中的知识组织问题

信息导航系统需要大量知识的支持。然而知识库体系中的知识单元之间具有无法想象的关联复杂度,从而形成一个庞大、多维、交织的知识空间,人们难以找到一个合理的框架以便有效地描述其中的知识(包括内容、属性、关联等各方面)。

类是知识的一种自然属性。哪怕某一知识单元不易明确其确切类属,但其必定具有某类或某些类的特征属性,也就是有归类的趋向。因此,知识的分类存储一直以来都是一种高效的组织方法。类最大的优势就在于等级分明、一目了然,可以对知识迅速定位。

知识领域的交叉融合所产生的海量而复杂的知

识关联虽然给我们的科学研究带来了创新契机,但却打破了人类知识体系固有的类从属层次关系,影响了知识组织的有序化。若要维护知识的类组织优势,就需要将所有知识按照一定的主题、功能分别存储。而且如果我们将所有的知识关联都加以表达描述的话,如此庞大的知识体系必定会被复杂、交错的关联链接网络所困扰,同时众多纠缠的关联链在管理、维护上也将成为导航系统难以承受的负担。另一方面,如果所有知识都组织在单一知识库中,虽然对库中的知识遍历迅速、调用便利,然而这样不仅要求所有知识必须具有一致的知识表示形式和推理策略,而且对问题的理解、建模都要困难得多^[1]。总之,由于知识类目、关联繁多且不易控制,过于集中的知识组织方式必然会影响知识库的存取效率。

因此,只有提高知识库的模块化程度,将知识库中的知识分成一组相互兼容、共享,但又相互独立的模块,从而构造出分布式的知识库体系,才能使知识的存储结构化、简单化。针对任何一个知识领域,或是从任何一种知识属性出发来进行知识的分别组织,使得其知识属性单一、知识关联清晰,从而易于知识库的管理、维护。

1.2 信息导航系统中的知识库类型

庞大而有序的知识体系能够辅助信息导航系统的正常工作。从知识的学科特性以及应用范围来看,可以有以下几种类型:①一般性知识(非专业知识):包括语言学知识和常识,可以被用于所有的学科领

* 本文系教育部人文社会科学博士点基金项目“基于信息构建的智能导航机理研究”(03JB870002)的成果之一。

域;②学科领域知识:同问题领域相关的实体、属性、事件和关系,以及有关的词典等;③知识的知识(元知识):包括专家的专门知识和系统的控制策略知识等^[2]。

根据需要可以将知识按照领域、功能、属性等的不同,进行分别存储,构建结构层次丰富、运作呈一体化模式的知识库体系来辅助系统的导航任务,从而使导航任务得以顺利高效地完成,并可以对系统运作进行整体智能性协调和控制。信息导航系统中的知识库体系一般可以包括如下类型的知识库。

(1)用户知识库。用户知识库通过用户的专业背景、需求偏好和导航行为等知识,以及使用系统的方式经验与交互记录、对系统运行过程与结果的反馈知识等,帮助导航系统理解用户的真实需求,挖掘用户的潜在需求,同时协助用户明确自己的模糊需求^[3]。

(2)专家知识库。专家知识库包括专家特定领域知识、执行专门查询任务的经验知识、任务管理决策知识、启发式知识等。专家知识库能够辅助用户正确制定导航策略,指导用户路径选择,使其不偏离目标领域范围。

(3)信息特征知识库。信息特征知识库一般可以包括信息的组织模式知识(如分类知识、主题知识等)、存储结构知识、表现形式知识等。信息特征知识库起到指导导航行为,提高信息的导航效率以及控制导航结果的知识处理质量。

(4)领域知识库。领域知识库中的知识包括专业领域的概念与概念之间的各种关系,以及一定专业领域中被公认的常识性知识。领域知识可划分为三个层次,即具体事实知识、领域概念知识和通用概念知识。领域知识库能辅助系统优化导航策略,有效缩小搜索范围,并对导航结果进行知识获取与集成等^[4]。

(5)常识库。常识库是用来存放相对于专业知识的那部分一般性常识知识。常识知识在量上要远大于专业知识,而且在问题求解的过程中,离开常识知识的辅助,很多推理判断无法正常运行^[5]。常识知识可以帮助导航系统联想式推测用户意图,并且协助系统准确分析问题内容。

(6)元知识库。元知识主要包括各类数据库内容的元数据、关于领域知识的知识(即领域知识的性质、功能、特点、组成的知识,以及如何管理、控制和使用领域知识的知识)、系统控制知识、计划策略知识、

各种优化的分布式导航策略和协作调控协议等。系统需要利用元知识对知识及其应用进行系统调控。

(7)语言知识库。语言知识库存放导航任务操作所需要的语言学知识,主要是字典、词典、语法知识、多语言互转换知识以及各领域词汇的特定性与相关性知识等,能够支持用户提问与信息源的自然语言处理问题和人机交互中的自然语言会话,以及多语种信息源的导航互操作等。

另外,各知识库可以通过推理学习功能使自身得到不断扩展,发挥较强的重构能力,使得信息导航系统在解决问题时能够完成复杂的知识挖掘、获取、集成等操作。

2 知识库体系的协作模式

在信息导航系统中构建如此庞大的知识库体系,其知识组织难免会有协调上的问题。毕竟知识之间一般都是相互融通的,过度分割使得部分知识关联断裂,终将成为知识的孤立,影响知识创新与学科发展。从另一方面来说,不同领域知识库中的知识在概念上具有二义性,这会造成知识在语义上的不一致;同样,不同专家的知识之间,以及用户知识同专家知识之间极有可能产生冲突,容易在导航决策的指导与被指导上出现问题。因此需要在知识库之间建立合理而高效的协作关系。导航系统的整体优势不仅同各类知识库息息相关,还取决于各分布式知识库的协作共享状态的形成与维持。

从控制论观点看,系统的智能是要通过系统与环境、子系统与子系统、系统与部件之间的动态交互过程得以反映^[6]。因此导航系统的各种分布式知识库之间应当建立动态交互、协调合作等协作模式,达到信息导航智能化的目的。也就是说整个知识库体系的运作模式应该处于一种谐振状态,通过一定的协调控制,各类知识库在功能上相互扶持,产生共振效应,从而增加知识库体系在导航支持方面的有序度。

在信息导航系统中,相互独立存在的知识库之间的协作模式有很多选择,这里提出了一个知识库外部指针和黑板机制共同作用下的协作模式,如图1所示。也就是要通过外部指针将不同知识库中相关联的知识联系在一起,并通过黑板机制展示、分派导航任务,以及在知识库通信交互方面起到支持作用。

2.1 外部指针协作

同一知识库中的知识单元间的关联指针属于知识库的内部指针,而指向其他知识库的指针就称为外

部指针。外部指针的目的就是使每一个知识库不再是完全孤立的个体,当其执行任务的时候都会得到其他知识库的知识辅助。

外部指针有两种功用:一种是知识关联链接的延续。由于为了保证任一知识库的单纯简洁,知识库中的知识依照领域、属性、功能等的不同被分别存储。这样就将一些相关联的知识单元被人为地割裂开来。外部指针就是在导航的过程中,根据任务执行中的需要,通过相应的知识关联链接,在不同的知识库之间进行知识沟通。由于这种关联不像同一知识库中的知识之间来得紧密频繁,所以虽然知识库之间的通信相对于同一知识库中的通信要烦琐耗时,但从整体上看,导航任务的完成还是快捷省力的。外部指针的另外一种功用就是它可以在系统运行时,随时从某一知识库调用其他知识库中的知识,借以辅助指导导航任务。因此整个导航任务可以多向并行执行,从而提高了工作效率。

2.2 黑板协作

一般来说,各知识库之间会自动通过外部指针进行关联。然而当该知识库本身不知道应向哪一个知识库请求支援的时候,就可以利用黑板机制的协调功能来达到目的。黑板其实是一个共享的问题求解工作空间,即一个虚拟动态的问题空间^[7]。每一个知识库意味着一个知识源,涵盖特定领域问题求解所需的专门知识,均可解决一种相对完整、独立的特定子问题。任一知识库都可以通过黑板了解到整个问题的内容,并且接受自己能处理的子问题,最后将问题结果返回给黑板。

这一过程主要需通过多个规则库来完成,包括问题分解规则库、子问题派发规则库、冲突协调规则库、结果合成规则库等。首先,问题分解规则库预先根据用户提问以及导航过程中所遇到的问题的复杂度,将问题分解成一系列较为独立的子问题,形成一个问题树,并将其发布在黑板上。黑板接受该系列子问题后,通过子问题派发规则库,为每一个子问题寻找相应的知识库,并逐一派发出,使各子问题得以并行处理,最后各知识库将子问题的处理结果提交给黑板。当结果全部提交给黑板,且没有任何子问题遗留下来,则黑板将所有结果转给结果合成规则库,将其有机地合并成一个最终解,返回给系统。

当某一知识库在问题解决的过程中发现存在自己无法解决的部分,同时又不存在有效的外部指针,只有进一步提取子问题,提交并发布在黑板上,请求

支援,同时停止问题的处理。黑板标记问题的来源知识库,并指示问题分派规则库再次根据该子问题所涉及的领域范围进行问题派发。当该子问题的结果返回后,黑板将结果直接转达给问题来源知识库,使得该知识库可以继续处理问题。因此黑板的首要作用就是知识库之间的通信和交互的中介。

当从不同的知识库中调用知识的时候,知识的冲突问题不可避免。黑板另一个作用就是将冲突的知识发布,并利用冲突协调规则库将知识冲突降到最低限度,找到一个最佳的解决方案。

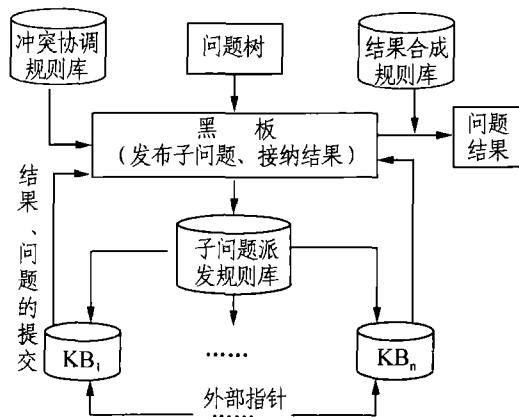


图1 知识库的协作模式

3 基于多知识库协作的信息导航

3.1 问题的分解与合成

在复杂的信息空间执行导航任务,势必会遇到导航路径选择方面的问题,很多时候需要提交给存储有丰富知识的知识库体系来辅助解决。待解决的问题首先由问题分解规则库根据其复杂度进行问题分解,得到相关度较低、独立性较高的一系列子问题。子问题根据其主次从属关系构成一个问题树。对问题树可以进行特定规则下的遍历,这样不仅不会遗漏任何问题,也不会在分派任务时偏离问题的主旨。同时在结果合并过程中,顺着问题树进行逆向遍历可以保证结果合并的正确性与合理性。

问题树中的叶子节点都是最终分解出来的子问题,且从左向右子问题的重要程度及相关程度均依次减弱。每个叶子节点都通过子问题派发规则库赋予一个任务链,任务链指向该子问题相关的知识库。任务链有主链和辅链之分,拥有主链的子问题节点说明是该问题的中心问题,拥有辅链的子问题节点说明是该问题的旁支问题,这在子问题派发优先序、问题解决深度,以及在结果合并上均起到了参照作用。

3.2 基于多知识库协作的信息导航模型

如图2所示。

在基于多知识库协作的导航模型中,其运作流程

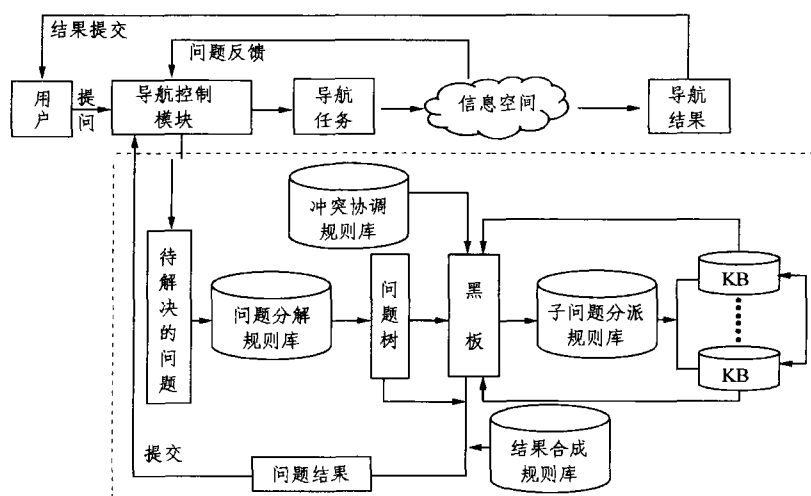


图2 基于多知识库协作的信息导航模型

用户提问输入到信息导航系统后,导航控制模块将用户提问转化成导航任务,并在信息空间中执行该任务。当系统遇到导航路径选择等问题,导航控制模块可以将问题转到导航系统的知识库体系解决。导航控制模块还具有查重、过滤等问题预处理功能。

导航中问题解决的过程是:

首先,将待解决的导航问题经由问题分解规则库进行问题分析与分解,并将各级子问题按照从属、主次关系,形成一个关系明晰、层次有序的问题树。其中问题树的叶子节点都是需要分派出去的待解子问题,每个叶子节点分别具有三个部分:存放子问题的值域、存放指向父节点层次链的指针域、存放指向对应知识库任务链的指针域。

其次,对问题树进行遍历,在子问题派发规则库的辅助下,依次向对应的知识库派发子问题,其中要根据子问题的重要程度或是子问题之间是否具有嵌套关系,决定任务派发的顺序。

然后,通过任务链将任务派发给各个知识库,在相应的知识库中搜索解决问题的知识。如果需要应用其他知识库的知识辅助问题解决,则可以通过已有的外部指针调用相应知识库中的知识。如果遇到没有预先设置好指针的调用问题,就将该问题提取出来,提交到黑板发布,再通过子问题派发规则库决定由哪些对应知识库处理,标记问题来源知识库后,派发子问题并将收到的结果重返问题来源知识库,保证其继续完成任务。如果某子问题提交频繁,就需要在相应知识库之

间设置一个固定的外部指针,以供以后直接调用。

最后当所有的结果都返回到黑板,利用结果合成规则库与问题树的关系结构,将所有结果过滤、合成,并将合成结果提交给导航系统,继续执行在信息空间的导航任务,并将最终结果提交给用户,完成导航。

知识库之间的协作问题是基于知识的信息导航系统研究的重要组成部分,协作带来的不仅是导航效率的提高,而且还增强了知识之间的融通与创新,使知识在信息导航系统中的作用和价值得以提升。

参考文献

- 1 曹泽文,张维明等.任务链:知识库组织的一种新方法.计算机工程,2002,28(1)
- 2 张玉峰.智能情报系统.武汉:武汉大学出版社,1991
- 3 张玉峰,李敏等.论信息检索与知识检索.中国图书馆学报,2003,29(5)
- 4 傅谦,张申生等.基于本体的领域知识共享研究.计算机集成制造系统—CIMS,2001,7(11)
- 5 Nils J. Nilsson. 人工智能.北京:机械工业出版社,2000
- 6 J. S. Albus. Outline for a Theory of Intelligence. IEEE Transactions on System, Man and Cybernetics, 1991, 21(3)
- 7 张维明.智能协作信息技术.北京:电子工业出版社,2002

李敏 武汉大学信息管理学院在读博士研究生。通信地址:武汉。邮编430072。

张玉峰 武汉大学信息管理学院教授。通信地址同上。

(来稿时间:2005-04-13)