

超媒体系统 语义网络 信息语义 多媒体

计算机科学1998 Vol. 25 No. 3

一种基于语义网络的开放式超媒体系统结构

48-51

A Semantic Network Based Open Hypermedia System Architecture

李光亚 周学海 龚育昌 赵振西 TP391

(中国科学技术大学计算机系 合肥230027)

摘要 In this paper, an open hypermedia system architecture used semantic networks as the underlying link storage mechanisms. The construct of content register table was proposed to support semantic-based hyperlinking to various kinds of information units. Semantic constraints were classified to describe the domain knowledge effectively. At last future work was suggested.

关键词 Semantic network, Content register table, Semantic constraint, Linking protocol, Open hypermedia system

现在的开放式超媒体系统,绝大多数只能有效地利用文本及图象两种媒体,未能充分利用文本、图形、声音、动画及视频等媒体来展示信息,因此需要提出一种有效的系统结构来支持各种媒体信息的集成。语义网络是由概念(结点)以及概念之间的语义关系(链)所组成的一种有向图形式的知识表示模型^[1],它与超文本之间的类同关系在文[10]中已被提出。虽然现在大多数超媒体系统都有一潜在的语义网络,但未能充分利用语义网络本身所固有的语义推理能力、应用领域的描述能力及其灵活的查

询能力。对于开放式超媒体系统^[2],语义网络更有其独特的优点:开放系统中潜在的应用领域多种多样,所集成的各应用系统也各不相同,这就需要一种机制来支持面向信息语义的链接,而不是某一特定应用的特定文档,该机制应该透明地支持不同应用文档的集成,而不必将文档数据从一种格式转化为另一种格式。真正的开放式超媒体系统它应该满足以下需求:

R1:它支持面向信息语义的链接,即可以链接到文档内容的任意元素,而不管具体文档的格式,只要

李光亚 博士生,主要研究超媒体系统和多媒体数据库,周学海 博士生,主要研究多媒体数据库和面向对象程序设计,龚育昌 教授,主要研究数据库系统,CAI,赵振西 博士导师,主要研究数据库系统,面向对象和软件开发环境。

- [3] C. J. Glass and L. M. Ni, Fault-tolerant wormhole routing in meshes without virtual channels, IEEE Trans. parallel and Distributed Systems, 7(6)1996
- [4] J. H. Upadhyay et al., Efficient and balanced adaptive routing in two-dimensional meshes, prasant@iastate.edu
- [5] A. A. Chien and J. H. Kim, Planar-adaptive routing; low cost adaptive networks for multiprocessors, in Proc. 19th Int. Symp. Comput. Arch, 1992
- [6] P. E. Berman et al., Adaptive deadlock-and livelock-free routing with all minimal paths in torus networks, in Proc. 4th Symp. on Parallel Algorithms and Architectures, 1992
- [7] R. V. Boppana and S. Chalassai, A framework for designing deadlock-free wormhole routing algorithms, Same to [3], 7(2)1996
- [8] L. Schwiebert, D. N. Jayasimha, Optimal fully adaptive minimal wormhole routing for meshes, J. of parallel and distributed computing, (27)1995
- [9] P. T. Gaughan and S. Yalamanchili, A family of fault-tolerant routing protocols for direct multiprocessor networks, Same to [3], 6(5)1995
- [10] P. T. Gaughan, Adaptive routing protocols for hypercube interconnection networks, IEEE Comput. Mag., (26)1993
- [11] Jose Duato, A necessary and sufficient condition for deadlock-free adaptive routing in wormhole networks, Same to [3], May 1993
- [12] Jose Duato and Pedro Lopez, Highly adaptive wormhole routing algorithms for n-dimensional torus, DIMACS serials in Discrete Mathematics and Theoretical Computer Science, (21), 1995
- [13] W. J. Dally, Virtual-channel flow control, same to [3], 3(2)1992

所集成的文档管理系统能够对之加以识别和处理。

R2:它支持链的自动生成,传统的超媒体系统只有当链的两个或多个端点均生成之后,才可手工生成链,真正的开放式超媒体系统应根据特定应用领域的知识自动建立链。

R3:它应该将结点的内容管理和链服务分开,当前的系统虽然将结点内容存储和链存储分离,但大多数系统中具体文档管理的细节被外部应用强加于了链服务系统,如 Microcosm^[1]。

R4:它应该支持知识的可重用性,同一应用领域的不同应用之间应能共享领域知识。

R5:它应该支持链的推导,即允许作者根据原有的链类型动态产生新的链类型。

R6:它应该保证所集成的各个文档管理系统的自治性。

R7:链的完整性应该由超媒体系统来管理,既降低了作者的认知负载,又减少了用户迷路问题的发生。

本文正是针对上述要求,提出了一种基于语义网络的开放式超媒体系统结构。

1 一种开放式超媒体系统结构

基于以上目标,本文提出了一种基于语义网络的开放式超媒体系统结构 OHSA (Open Hypermedia System Architecture),如图1所示。

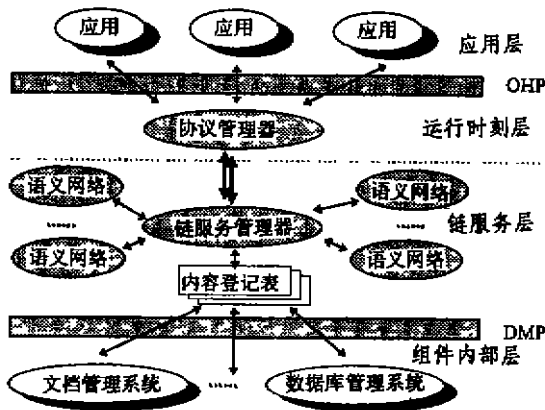


图1 开放式超媒体系统结构 OHSA

图1中的分层结构沿用了 Dexter 参考模型^[5]中的三个层次的概念,但每层内部的具体结构以及层与层之间的接口与 Dexter 模型都不尽相同,因为 Dexter 参考模型不能很好地支持开放式超媒体系统的设计,OHSA 各层的结构说明如下。

应用层由不同应用领域的多个应用所组成,它

们通过 OHP (Open Hypermedia Protocol) 与运行时刻层进行通信,运行时刻层的协议管理器负责根据不同的应用领域选择不同的语义网络进行存取,此外协议管理器还负责对各个应用之间的并发控制,具体 OHP 的内容见第2节。链服务层的链服务管理器具体完成对语义网络和内容登记表的查询,每一个语义网络用来描述某一应用领域的知识,这样同一应用领域的不同应用对应着相同的语义网络,满足了 R4 的需求。为了支持 R3,我们将结点的内容管理(由外部数据管理系统如文档管理系统、数据库系统维护),同链服务(由链服务管理器对语义网络和内容登记表进行存取来实现)分开,而内容登记表是为了记录所有外部数据管理系统所能处理的概念标识而提出的一种新的构件,一种可能的表记录形式为(管理系统名,文档名,概念标识)。各外部数据管理系统通过 DMP (Document Management Protocol) 向链服务层登记其所能处理的概念标识。概念标识的具体形式是不确定的,如字段名、图象中某一人物、URL 或文本中的某一单词等,只要外部数据管理系统能够对该标识进行解释并存取,但就目前的技术而言,全文检索、多媒体按内容查询等技术的发展使得关键字是一个比较好的选择,如关键字 John 可以代表数据库中的一条记录、某个 Video 中 John 的出现或通常的文本中单词 John 的出现,这些概念标识所代表的具体内容数据是由链服务层通过 DMP 向外部数据管理系统而取得的。

2 链接协议:OHP 和 DMP

标准 OHP 的概念最初是由 Davis 在文[3]中提出的,它负责外部应用与开放式超媒体系统的通信,对于大多数外部应用来说,链接协议是它们获得超媒体功能的基础。一个标准 OHP 的出现将允许一个支持 OHP 的应用与任何支持 OHP 的开放式超媒体系统进行通信,很显然这增强了应用的可重用性。典型的 OHP 的内容可以分为四个部分:①链的激发,当一个链被应用激发后,应用以某种形式将链的信息告知超媒体系统,而超媒体系统根据此信息对链及链的端点进行定位;②浏览信息的初始化,一个应用文档被展示前,超媒体系统必须将有关此文档的链网络信息告知应用,以便让文档对其外观进行改变等;③结构信息的产生和维护,这需要应用和超媒体系统双方交换信息以便维护超媒体网络,例如当用户删除或增加一条链时,相应的链服务层中链的信息被删除或增加;④合作消息的通知,这牵涉到同一超媒体语义网络的各个应用之间的数据一致性问

题。

与 OHP 定义的客户-服务器之间的通信不同, DMP(文档管理协议)定义的是服务器-服务器之间的通信协议,其实现主要基于概念描述,由前所述,这保证了各个外部数据管理系统的自治性,即由外部数据管理系统负责结点或锚的数据存取,而超媒体系统只是通过 DMP 向外部数据管理系统存取数据,这满足了 R6 的需求,典型的 DMP 的形式可作这样设计(部分):①Register(概念标识,文档名,属性名,值);②UnRegister(概念标识,文档名);③GetObjects(概念标识,文档名,属性名);④UpdateObject(概念标识,文档名,属性名)。其中①和②是外部数据管理系统向内容登记表进行概念登记时所用的, Register 用来登记概念标识条目,而 UnRegister 则是当外部数据管理系统中的数据被删除时,为了保证超媒体系统中数据的完整性而向链服务层发出的调用,它将概念标识所对应的实例删除,这部分解决了开放式超媒体系统中的数据完整性问题^[6];而③和④是链服务层向外部数据管理系统存取数据内容或概念标识属性时所用的。

OHP 和 DMP 一起实现了超媒体系统与外部系统或应用的接口,它支持面向信息语义的链接(满足了 R1),因为任一文档或文档内部的元素,如图象中的某一物体,只要外部数据管理系统能够对它加以识别,应用就可以把它作为结点或锚来处理,这需要外部数据管理系统具有一定的按内容查询的能力,这一领域已取得了相当的进展^[4,9]。而传统的由文本中的一字段所组成的锚,可以通过在内容登记表中将该字段值作为概念标识加以实现。

3 面向应用领域的语义约束

语义网络是一种由有向图所构成的知识表示模式,在有向图中,概念单元被表示成结点,而结点和结点之间的关系被表示成链,很显然,语义网络和超媒体链网络之间存在着一定的对应关系。一个自由的语义网络对应于允许任意链类型的超媒体网络,一个限定的语义网络允许有限的链类型,而一个面向特定应用领域的语义网络则是具有语义约束的、有限链类型的超媒体网络。因此,语义约束在用语义网络描述特定应用领域时具有十分重要的作用^[10],它维护了语义网络中所有概念标识实例之间的数据完整性,并通过对语义网络的创建和操作的限定,更加精确地描述了特定应用领域的知识。

在本文所提出的基于语义网络的 OHSA 中,我们将语义约束进行了分类:结构约束、关系约束和计

算约束。结构约束维护的是语义网络中某个子网所构成的图的类型,如线性、树和 DAG 等,这种类型的约束维护的往往是由同一类或某几类链类型所生成的子网图的类型,如利用语义网络描述书的章节结构时,所生成的语义网络类型只能是一树型结构,这就限制了用户对该语义网络的插入操作。而关系约束维护的是超媒体网络结点类型之间允许或不允许存在的链类型,不仅仅是简单的图的类型。一个简单的关系约束的例子用规则表示为:

```
LinkType1(Node1, Node2) :- NodeType1
(Node1), NodeType2(Node2)
```

它所表示的是链类型1只能建立在结点类型1和结点类型2之间。我们进一步可以将关系约束分为子图包含约束和子图排除约束,这两种约束结合了前面所讲到的结构约束,如图2所示:

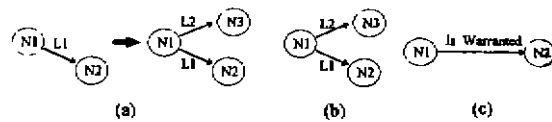


图2 面向应用领域的语义约束例子

图2(a)表示的是子图包含约束,即若 N1和 N2 之间有链 L1相连,则必存在结点 N3,而 N1和 N3 之间有链 L2相连。如学生 S1(N1)需要选修课程 C1(N2),则必须已选修过课程 C2(N3),这种语义约束可以通过子图包含约束来实现;而图2(b)代表了子图排除约束,即不允许存在象(b)那样的子图,它所表示的是语义网络中的析取关系,如“John is a Chinese”和“John is an American”这种析取关系可通过(b)来约束。

计算约束表示的则是对一超媒体网络结点、链或子图进行操作的条件,如文[12]中由 Petri 网中 Token 所限定的超媒体网络执行限制即是一种计算约束,这种约束一般都牵涉到实例化结点或链的属性值,并根据这些属性值来限定对这些实体的操作,如在文[1]中描述银行系统的语义网络中,为了批准某项贷款,要求贷款申请者的收入必须超过10000美元,则可以通过由如下规则所表示的计算约束来描述:

```
Is Warranted(Node1, Node2),
-Loan_Contractor(Node1),
warrant_Letter(Node2),
Node1.name = Node2.Loaner
Node1.income > 10000.
```

在上述规则中假定 Loan_Contractor 类型的结点有一 income 属性,则上述规则约束了在贷款合同结点与批准书结点之间进行链接时必须进行的语义约

束检查。

面向应用领域的语义约束满足了R2的需求,在基于上述OHSA的系统中,用户无需手工生成链,当某一文档或文档内部的信息元素被外部数据管理系统通过DMP中的Register命令向内容登记表登记为语义网络中某一概念标识的实例后,链可以由系统本身根据上面所提到的三种语义约束规则来自动生成,这当然需要外部数据管理系统通过DMP接口对概念标识实例的属性进行赋值。这既维护了语义网络数据的完整性,又极大地降低了用户的认知负载,满足了R7的需求。

当前大多数超媒体系统中,链的类型是由作者事先固定好了的,这显然具有很大的缺陷:作者事先不可能考虑到所有的链类型,更何况每个浏览者所期望的超媒体网络链类型也是各不相同的。利用上面所提出的语义约束,我们可以支持虚拟链的产生,即用户可以根据自身的需求在结点之间定义新的链类型,它相当于数据库中视图的概念。

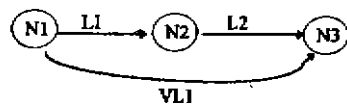


图3 虚拟链类型的定义

图3中若想在Node Type1型结点N1和Node Type3型结点N3之间直接建立一链,我们可以通过语义约束建立一虚拟链类型VLI来实现:

VLI(Node1, Node3) ← Node Type1(Node1), Node Type2(Node2), Node Type3(Node3), Link Type1(Node1, Node2), Link Type2(Node2, Node3).

一旦用户定义了这种虚拟的链类型,以后系统便会自动根据语义网络的实例化来动态地生成这种链类型,当然该链实例的生成是通过调用上述规则来完成的,而其中LinkType1(Node1, Node2)和LinkType2(Node2, Node3)两个子句的判定则要通过别的规则来进行判断,这满足了R5的需求,传统的超媒体系统中链的类型则是由作者事先定义好的,这显然缺乏必要的灵活性。

结论 本文针对当前开放式超媒体系统所存在的缺陷,提出了一种基于语义网络的开放式超媒体系统结构。然而,实现中如何将内容登记表中各记录进行有效的组织,使之不致成为系统的瓶颈,是一个值得深入研究的问题。此外,我们虽然提出了面向应用领域的语义约束问题,但如何在系统中记录各种约束(如以规则的形式)以及这些约束的执行机制,这在系统实现时值得进一步的考虑。语义路径^[2]是

一种有效的机制,如何将这种机制同超媒体所特有的浏览机制相结合,关系到系统的查询能力,一种可行的办法是如文[1]中所提出的一种知识查询语言(KQL),它既结合了传统的信息检索,又能让用户通过指定语义路径来对语义网络进行遍历。

随着多媒体数据的广泛使用,对于按内容进行链接变得十分必要。关于这方面的研究已经取得了很大的进展^[1-9]。一旦某个外部数据管理系统具备了某种程度的按内容查询能力,便可通过内容登记表将其所管理的文档输入给超媒体系统的应用,而这些对于用户来说是透明的。此外,随着软件构件技术(如Agent)的发展以及操作系统对复合文档的支持(如OLE、CORBA),使得OHP和DMP的实现变得更加高效和灵活。

参考文献

- [1] A. Celentano et al., Knowledge-Based Document Retrieval in Office Environments, The Kabiria Systems, ACM TOIS, 13(3)1995
- [2] D. Cunliffe et al., Query-Based Navigation in Semantically Indexed Hypermedia, Hypertext'97, April 1997, Southampton UK
- [3] H. Davis et al., OHP, A Draft Proposal for a Standard Open Hypermedia Protocol, Hypertext'96, March 1996, Washington DC
- [4] A. Gupta, Semantic Queries with Pictures, The VIMSYS Model, VLDB'97, Sept. 1991
- [5] F. Halasz, The Dexter Hypertext Reference Model, CACM, 37(2), 1994
- [6] 李光亚、周学海等,超媒体系统的开放性探析,计算机科学,1997
- [7] J. Mayfield et al., SNITCH, Augmenting Hypertext Documents with a Semantic Net, Intl. J. of Intelligent and Cooperative Information Systems, 2(3), 1993
- [8] J. Peckham, Semantic Data Models, ACM Comput. Surv., 20(3)1988
- [9] 周学海、李光亚等,按内容检索的图象数据库系统数据模型,软件学报,已录用
- [10] W. Weigang, et al., A Semantic Net for Reusable Hypertext, Proc. of Workshop on Intelligent Hypertext, 1993
- [11] H. Davis et al., Toward an Integrated Information Environment with Open Hypermedia System, ACM Hypertext'92, Dec. 1992, Milan Italy
- [12] P. D. Stotts, et al., Petri-net Based Hypertext: Document Structure with Browsing Semantics, ACM, TOIS, 7(1)1989