

# 超文本技术

## Introduction to Hypertext

刘连芳

Liu Lianfang

(广西计算中心广西软件新技术实验室 南宁市星湖路 32 号 530022)

(Guangxi New Software Technology Lab., Guangxi Computing Center,

32 Xinghu Road, Nanning, Guangxi, 530022

email: lfl@Sun.ibep.ac.cn)

**摘要** 概要地介绍了超文本技术的产生与发展, 阐述什么是超文本, 超文本系统的结构、超文本与其它计算机技术的差别与联系, 概要地说明一下超文本的标准化问题。

**关键词** 超文本 超媒体 非线性 关系

**Abstract** Some scientists have been dreaming for ages of a kind of technology which can manage information as human mind. It finally became true in 90s. Ted Nelson coined "hypertext" for the new technology, and the word has been well accepted.

Hypertext is one of the most popular technologies. This paper introduces its emergence and development, explains what is hypertext, illustrates its structure, the differences and relations between hypertext and other computer technologies, then discusses the standardization briefly.

**Key words** hypertext, hypermedia, nonlinear, relation

### 1 为什么会产生超文本技术

当今时代是信息爆炸的时代, 信息量增长之快, 出乎了人们的预料。例如, 每一架波音飞机离开产地时, 都有 4 辆满载它的文档资料的牵引拖车随行<sup>[1]</sup>。信息量的高速增长促进了信息存贮技术的发展。目前数百兆、数千兆容量的磁盘、光盘、磁带已不罕见。信息量愈大, 查找所需要的信息就愈困难, 原有的信息管理技术已经不能适应新的形势。

一些新的计算机应用领域的出现, 也召唤新的信息管理与检索技术。例如, 在 CAD 领域, 要处理的信息类型多, 并且复杂, 不仅有字符、数字, 还有相关的各种形式的文档和图纸。CAD 这样新领域提出的问题, 很难用传统的技术解决。

近年来, 多媒体技术迅猛发展, 缩短了计算机与用户之间的距离, 同时, 多媒体信息管理与单纯的文字信息管理有极大的差别, 必须有新的管理技术。

所有这些, 都对传统的信息管理技术提出了挑战。

传统的计算机技术使计算机成了用户的主宰, 用户必须掌握相当多的计算机知识才能使用计算机。使用时, 还必须严格遵循计算机程序的“指示”。而程序的“逻辑思维”方式与一般用户的思维和工作方式大相径庭, 文化水平不高的或者从事的学科与计算机相距较远的, 一时难以适应, 使用户对计算机望而生畏。用户——计算机的主人, 变成了事实上的“奴隶”。

另一个严重的事实是, 信息之间的关系在人的头脑中以及现实世界的所有事物中都占有重要的地位。但是, 传统的计算机技术或者忽略了这些关系, 或者只是以隐式表示这些关系, 与现实生活相去甚远。

长期以来, 科学家们一直梦想有一种能按人的思维习惯组织和管理信息的方法, 能让用户主宰计算机。经过几十年的历程, 在 90 年代, 这种想法变成了现实, 即, 出现了新型的信息管理技术——超文本 (hypertext) 技术。

### 2 什么是超文本

为了解释什么是超文本, 我们先把超文本和传统文本 (text) 做个比较。

所有传统的文本,包括印在纸上的和存在计算机文件里的,都是顺序存放和阅读的。例如书,是从第一页顺序编排到最后一页,一篇文章(或一章、一节)接着另一篇文章(或一章、一节)。所有的都有头有尾。

事实上,人们的思维方式及顺序与传统文本的存放方式不同。人的思维是联想型的,人习惯于按事物之间的联系思考问题,继而采取行动。例如,在阅读书籍,特别是学习知识或做研究时,并不完全按照书的写作顺序,而是常常产生“跳跃”,从一页“跳”到相关的另一页,或从当前的书“跳”到相关的参考书或参考文献等等。

由于背景、知识水平、阅读目的不同,不同的人读同一本书就会产生不同的思路,因而导致不同的活动。例如,一个未接触过超文本技术的计算机工作者阅读“Hypertext & Hypermedia”一书时,肯定想知道超文本的定义和其它任何细节;一个熟悉超文本技术的计算机工作者阅读这本书时,可能会跳过一般性介绍,只阅读他以前并不了解的内容;一个研究超文本技术的学者,除了读这本书外,还会边读边查参考资料,在几本书之间或这本书与若干参考资料之间“跳”来“跳”去,或者在这本书的各章节之间来回“跳”来“跳”去。……

由于时间、环境、目的的变化,同一个人在不同时间阅读同一本书也会有不同的想法和不同的阅读行为。最常见的例子是学生阅读教科书。学生的思维和阅读活动一次与一次不同,随学生掌握知识的程度及学习的需要不断变化。

由此可见,传统文本的存放和管理方式完全与人的联想思维方式及相应行动不相适应。

超文本,简而言之,是一种按人脑的联想思维,非线性地存储、组织、管理和浏览信息的技术。

如果用超文本技术组织一本书,就和传统的正文文件或印刷的书籍不同。所有正文(文章、段落、一句话、一个词)都按相互间的联系被组织成正文网。网中的信息不仅有传统书籍的全部内容,还加入了许多相关参考资料。例如作者简介,对书中关键词的解释或定义,与书中某些内容相关的其它书籍、文章、段落等等。内容远远超过了传统的书籍。由于是信息网,而不是顺序存放的书,所以这本“书”无所谓第一页和最后一页。从哪一段正文开始阅读,以及接下来读什么都由读者根据需要临时决定。选择“下一段”正文的依据既不是顺序,也不是索引,而是正文之间的联系。所以,超文本技术采用的组织和浏览方式更符合人的思维方式及工作习惯。

至今为止,超文本技术还没有确切的统一的定义,但是我们可以说:超文本是收集、存贮和浏览离散信息,建立和表示信息之间关系的技术。

超文本以信息和信息之间的关系全面地表示各种知识。它能模仿人的大脑,补充人的记忆,并辅助人的思维与交流。

超文本技术不同于传统的计算机技术,它不仅注重信息本身,更注重信息之间的关系。也就是它的重点并不是研究信息本身,而是除管理信息外,更重要的是超越信息空间,存贮、管理信息之间的非线性关系。这也就是 hypertext 的中文翻译应该为“超文本”,而不是“超级文本”的缘故。<sup>[2]</sup>

超文本技术让计算机用户做主宰。用户可以随意组织和浏览信息,使用户感到更自由、更友好。可以说,超文本技术给非计算机专业的用户带来了新的希望。从这种意义上来讲,超文本是某种程度上的接口技术。<sup>[3]</sup>超文本技术有它的特有的组织、管理信息的方式,有特有的结构,所以它又不仅仅是接口技术。

此外,超文本技术不仅涉及计算机科学,还涉及到认知科学和各个应用学科;不仅涉及信息管理领域;还会影响计算机的其它方面,如网络、多媒体、CAD、CASE、专家系统等等。所以,它引起了心理学家、各应用学科的专家及计算机各领域的专家的兴趣,并且倍受关注。

### 3 超文本的组成及有关概念

#### 3.1 基本组成

至今为止,超文本并没有完全统一的结构。但是,所有超文本系统的最基本组成可以说是相同的,那就是,任何超文本系统都是由两大部分:节点(node)和链(link)组成的。

节点在超文本中用于存贮离散信息,链用来表示信息之间的关系,并决定超文本的结构。

乍一看超文本是非常简单的,只由节点和链组成。但是,再深入研究就发现超文本技术实际上很复杂。首先,节点和链的类型复杂。其次,各个超文本系统所采用的结构和管理方法不同,就形成了形态各异的超文本系统。

• 节点 节点是超文本系统的基本信息单元。节点按表现形式分为基于框架(frame)和基于窗口(window)的两种。基于框架的节点的体积是固定的,它所能存储的信息一般相当于一个显示屏或半个显示屏的容量。这种节点简单,但信息量受限,不灵活。相对而言,基于窗口的节点对信息没有限制。信息量大,超过窗口可视范围,可以使用滚动带浏览,灵活、

方便。

根据内容区分,节点可分为正文节点、图像节点、声音节点、图形节点、动画节点、混合节点、正文索引节点、对象节点、规则节点等。不同类型节点有不同用途。对象节点用于表示某些知识,可以链接超文本以外的对象系统;规则节点存放推理用的规则;等等。其中的各种节点(例如,正文、图像、动画、声音节点)还可以互相补充说明。

·链 在超文本中可以定义的链的类型很多,最常见的是导航链。例如名称链(节点间的关系通过名字链接)、聚焦链(节点B放大节点A)、全景链(聚焦链的反)等属于导航链。此外,还有一大类是组织推理链,如索引链、属性链、组成链、隐含链。隐含链用于链接推理树中的事实,索引链链接正文索引节点和相应的索引项。

各种类型的节点已经使超文本成为一种异常灵活的知识表示工具,再加上节点之间各种类型的链接关系,为超文本提供了浏览和推理能力,更进一步提高了超文本的表达能力和灵活性。节点和链的类型的复杂性和灵活性决定了超文本的复杂性和灵活性。

### 3.2 有关概念

除“节点”和“链”以外,超文本还有些专用术语。

·浏览(browse) 有了节点和链,就有了信息网。在超文本系统中,用户查看信息的操作是根据信息之间的关系,也就是沿着链进行的。比如,“超文本”和“多媒体”之间有链,用户从“超文本”这个节点出发,沿着链就可以找到“多媒体”。这种沿链查看相关信息的操作在超文本中叫作“浏览”。

·导航(navigation) 超文本利用链引导用户浏览信息的功能叫导航。每个超文本系统都要提供导航工具,用户才能在信息空间中遨游。

·锚(anchor) 链是表现信息之间关系的重要工具。锚是链的端点,可以在任何节点的任何位置上。链通过锚连接节点。源节点上的锚叫源锚,目标节点上的锚叫目标锚。一般,源锚是可见的,可供用户作为浏览起点来选择。

·按钮(button) 按通常的理解,按钮是一个可以按动的物体。超文本系统使用一些类似物理按钮的图形供用户用鼠标选择,从而激活相应的操作,例如跳转。在超文本中,也称这种图符为按钮。

如果“按动”按钮激活的操作是转向相连的目标节点,那么,它的功用就与锚一样。但一般来说,按钮不会出现在有信息的地方,并且它所激活的操作不限于转向相连目标节点。例如,可以激活某个外部程

序。

·超媒体库(hyperbase) 超媒体的概念本文后面将予以介绍。超媒体库是相对数据库而言。顾名思义超媒体库就是超文本中存贮节点和链的机构。

## 4 超文本的产生与发展

### 4.1 发展历史

超文本从思想的诞生到技术上的实现及发展,经历了几十年的时间。其中值得一提的有几个事件<sup>[4]</sup>。

·1945年,Vinnevar Bush 提出以非线性方式组织正文的方法。Bush 提出了一种叫做 Memex 的机器。Memex 包含有一个庞大的文档库,Bush 的原始思想就是把广为散布的信息聚集在一起形成一本新书。任何人都可以根据自己的意愿在其中的信息之间建立联系,形成某个特定的浏览路径。但是,由于40年代硬件条件的限制,Bush 的想法没能变为现实。

·1963年,斯坦福研究所的 Engelbert 在 Memex 思想的影响下进行研究,于1968年产生了一个叫 NLS (ON Line System) 的 hypertext 系统。NLS 由文件组成,文件又由少于3000个词的段落组成,它们之间可以按层次结构或非层次结构链接。

·1964年,Ted Nelson 为这种技术杜撰了“hypertext”这个词。Ted Nelson 所参加的课题 Xanadu 开始于1960年。Xanadu 的目的是把世界范围的文学著作放进计算机,并且用超文本技术连接它们,使它们易于使用。1988年一个大软件公司接手了这个课题,想把它商品化。后来,“hypertext”这个词得到公认,成了这种非线性信息管理技术的专用词汇。

·1987年11月,在美国北卡罗来纳州召开了第一次国际超文本学术研讨会。

除此之外,值得一提的是继1987年的超文本研讨会之后,自1989年起,每年都有一次国际超文本学术交流会,单年在美国,双年在欧洲举行。美国的计算机协会 ACM 下专门成立了“超文本”兴趣小组 SIGLink,为从事超文本研究的人们提供交流的环境。ACM 下的另一个小组 SIGhyper 则致力于超文本标准的研究。

目前可供阅读的有关超文本技术的书刊有每年的学术会议论文集、美国的“HyperlinkMagazine”、英国的“Hypermedia”以及一些专著。

### 4.2 一些早期的系统<sup>[4]</sup>

超文本技术近年来发展比较快,应用也在不断增加。众所周知的 Internet 软件 3W (World Wide Web) 就是采用超文本技术的一个例子。

以下介绍几个在80年代末期推出的比较著名的

系统。

• Intermedia 是美国布朗大学于 1987 年为研究及教学开发的系统。布朗大学研究 hypertext 已有 20 多年的历史。他们的第一个系统是 1968 年开发的,叫超文本编辑系统,用于产生有关阿波罗空间程序的文档。

Intermedia 是在 4.2 BSD UNIX 下开发的,可以在 IBM RT、PC 和 Sun 工作站上运行,也可以在 Machintosh II 上运行。它由正文编辑器、图形编辑器、扫描图像浏览和三维对象浏览系统等几部分组成。Intermedia 基本上只限于在布朗大学内使用。

• NoteCards 是美国智能系统实验室 1987 年开发的一个 hypertext 系统。NoteCards 有良好的浏览工具,包括一个组织复杂的 NoteCards 网络的层次结构系统。

NoteCards 的作者把 NoteCards 看作思想处理环境,以便与用于教育的 Intermedia 等系统相区别。它的主要目的与 Engelbert 的 Augment 相似,但是结构不同。NoteCards 是网络式的,而 Augment 是轮廓式的。

基本的 NoteCards 结构是一个语义网。语义网由一组 NoteCards 组成,NoteCards 之间由类型链连接。NoteCards 提供了一组工具去组织、显示和管理语义网,还提供了建立、管理语义网的程序的方法和协议。

NoteCards 的 4 个基本实体为:卡片,链,浏览器,文件盒。

一张卡片是一个 3×5 的电子索引卡片。每张卡片有一个标题,卡片内可以有正文或图形,用户可以对正文或图形进行编辑。不同类型的 NoteCards 在卡片类型继承层次结构中定义。

链把相关卡片连接成网。链是有向的。

浏览器是一个包含有 NoteCards 网的结构图的卡片。浏览器是支持编辑功能的工具。用户可以通过对浏览图上节点和边进行操作来编辑语义网的基本结构。

文件盒用于建立层次范畴结构,以便有效地存贮和检索卡片。

• Guide 是第一个用于 Machintosh 的 hypertext 系统。它是 Kent 大学 1982 年开始的课题。Guide 是一个建立电子文档的工具。它比 NoteCards 更面向正文。Guide 使用了 4 种链类型:替换、注释、参考和查询。

Guide 最初是在 UNIX 环境下开发的,后来移植到 Machintosh。

• HyperCard 是 Apple 公司 1987 年在 Machintosh 上开发的。它可能是应用最广的 hypertext 系统,一个重要原因就是 Apple 随 Machintosh 机奉送 HyperCard。

HyperCard 的节点是与屏幕一样大小的卡片。卡片之间用“按钮”连接。HyperCard 提供一些画图工具,这样,就很容易设计包含有复杂图形和正文的卡片。HyperCard 还包括 HyperTalk 语言。在 HyperCard 的卡片上,用户可以设计命令按钮。HyperTalk 语言则用于编写与命令按钮相连的命令文件,例如:

```
on mouse up
  go to next card
end mouse up
```

当用户在命令按钮的图符上按动鼠标器按钮时,系统就显示下一张卡片。

HyperTalk 易于使用,所以许多非程序员用 HyperCard 开发了很多简单的系统。HyperCard 特别适用于作一个接口设计工具。它鼓励这样一种程序设计风格:先设计接口,然后,再通过链接命令按钮和相关的按钮命令文件来添加所需要的功能。

HyperCard 不是一种完全的面向对象程序设计环境,但是,它具有许多面向对象程序设计语言的特性。在面向对象程序设计中的 5 个重要组成部分是:对象、方法、消息、类和继承性。在 HyperTalk 里有 5 种对象:卡片叠、背景、卡片、命令按钮和字段。每个对象都可以发送和接收消息。每个对象都可以和一个包含一个方法的命令文件相连。

在 HyperTalk 中类和继承的概念没有那么明显。

HyperCard 提供 5 个级别的用户:浏览、键入、绘图、写作和编写命令文件。这样,不同的用户可以为不同的目的使用 HyperCard。

• KMS KMS (Knowledge Management System 知识管理系统)的前身是 ZOG。ZOG 是美国卡内基-梅隆大学自 1972 年至 1985 年开发的用于结构式浏览的 hypertext 系统,曾于 1982 年作为计算机辅助信息管理系统安装在一艘原子能航空母舰——USS Carl Vinsen 上。它的应用系统包括一个联机的策略手册和一个联机的维护手册。

ZOG 的商业版 KMS,是由知识系统公司开发的。ZOG、KMS 的基本要素不是节点,而是框架(frame)。框架的尺寸或是和屏幕一样,或者是屏幕的二分之一。框架内容可以是正文、图形和图象。每个正文项可以和另一个框架相链接。还可以激活一个程序。

框架之间有两种类型的链:层次链和交互参考

链。为了帮助用户在信息网内“航行”，框架都是格式化的，具有名字、标题、框架体，连接低层链的树项以及特殊项和命令项。为了简化系统和提高速度，ZOG 和 KMS 既不使用相互覆盖的窗口，也不使用图形浏览器。它的开发重点是放在快速检索正文的能力、支持多个用户以及尽量减少系统响应时间方面。在 Sun 和 Apollo 工作站上运行，KMS 访问和显示框架所花时间平均不足 1 秒。KMS 数据库可以分布在无限的文件服务器上。只要磁盘空间允许，数据库可以任意大。

KMS 的应用范围很广，它可以用于电子排版、联机技术手册、项目管理、问题分析、ES 用户接口、软件工程、计算机辅助外语翻译等等。

## 5 超文本的特长与不足

### 5.1 特长

超文本技术在存储、查询、浏览信息和知识表达方面相当灵活，它的思想更符合人的思维和习惯。超文本技术将计算机的存贮、表现能力和人脑筛选信息的能力有机地结合起来，更有效地发挥计算机和人脑各自的潜力，提高查找信息的效率。采用超文本技术开发的系统比传统的线性技术给用户更好的感觉和视觉，给用户以充分自由，大大缩短了人与计算机之间的距离。

超文本技术尽可能保留管理对象的本来面目，对管理对象的类型、体积没有规则化的要求。这就使超文本可以适用于任何信息管理。

超文本可以存贮、管理、表现信息之间的任何关系，可以如实描述现实世界的系统，因而可以解决线性技术解决不了的问题，例如工程数据库的建立。

### 5.2 不足

由于信息量大、类型繁多、关系复杂，而超文本又太灵活，所以，超文本系统的最普遍问题是用户容易迷失方向，不知道自己在信息网的什么地方，已经浏览过哪些信息。超文本的设计者们在努力解决这个问题，并且这个问题在某些系统里已达到某种程度的解决。

第二个问题是查找信息的效率问题。如果仅仅凭借关系（也就是沿链）浏览信息，如果一个系统有上万、数十万，甚至上百、上千万个节点，如果起始节点找的不好，那么，可想而知，要“顺藤摸瓜”从中摸到所要的瓜会是一个相当艰巨的任务，效率将是很低的。因此，提高超文本系统的效率也是超文本工作者的一个重要任务。

第三个问题是实际应用问题。从理论上分析，超

文本是相当美好的技术，但是，国外的专家、学者对应用做过一些课堂实验。实验结果表明超文本系统的实际应用还存在一些问题。有些学生认为使用超文本系统的感觉不如传统的文本，因此，对超文本的应用尚有争论。

## 6 超文本及相关技术

超文本有它的长处，也有其不足。每一种技术都包含这两方面。如果根据应用领域的需求，将相关技术结合起来，各取所长，可以更好地发挥它们的作用。

### 6.1 超文本与多媒体

近年来，由于计算机硬件的高速发展，高速、大容量微机的出现，使计算机象电视那样，能够以人类喜闻乐见的图、文、声、像并茂的形式向人们提供信息，从而，将使计算机渗透进包括家庭、娱乐、新闻、艺术、商业在内的人类生活的各个领域，开创计算机科学的新时代。

五彩缤纷的多媒体技术的确富有吸引力。目前用计算机存储、表现多媒体信息已经变为现实。但是，要让多媒体象人们所期望的那样投入实际应用，还必须要有相应的管理技术。

目前，可利用的多媒体信息管理技术有两类。一类是传统的数据库技术，另一类是本文介绍的超文本技术。

传统数据库技术管理多媒体信息，称之为多媒体数据库。数据库技术成熟，有较好的应用基础，这是它的优势，但是，由于 6.2 节阐述的原因，多媒体数据库有它的较大局限性。

应用超文本管理多媒体信息，这种技术在计算机界有一个专用名词：超媒体（hypermedia）。超媒体技术超越了媒体信息空间，它不仅管理各种媒体信息，还组织管理媒体信息之间的非线性关系。

由于超文本是本文所介绍的非线性技术的统称，所以，凡涉及超媒体技术之处，均可使用超文本一词替换，二者意义相同。如果一个系统所管理的对象仅仅是文本一种媒体，那么对它的确切称谓应该是超文本系统。在大多数情况下，超文本和超媒体二词可以混用。所以，尽管本文的前 5 节讨论的均是超文本技术，实际上它们对超媒体同样适用。

超媒体技术将多媒体技术的生动表现形式和超文本技术自由的、符合人脑思维习惯的交流方式融为一体，将为信息管理领域带来革命性的变化。

### 6.2 超文本与传统数据库

一般人熟知的传统数据库技术将信息规则化为文件、记录、字段。传统数据库技术在信息检索、报

表生成、数据统计方面具有强大的功能。它适合于管理大数据量、少类型的同质信息。但是，传统数据库对非同质信息显得无能为力。

超文本适宜于管理复杂类型的非同质信息，自由、灵活。但是，在检索、报表生成、统计方面明显不足。

如果将超文本技术和传统数据库技术，则可以发挥各自的优势。

### 6.3 超文本、多媒体、数据库三者的关系

超文本、多媒体、数据库三种技术可以说都是计算机科学的重要分支，而三种技术又可以如上所述互相结合。图1是作者对这三种技术之间关系理解的一个图示。

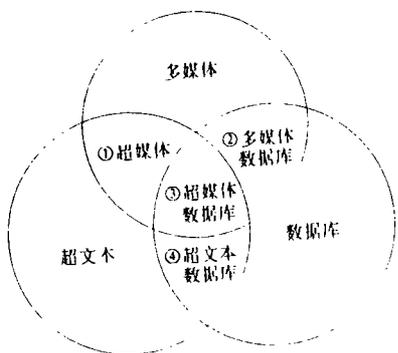


图1 超文本、多媒体、数据库技术的关系

Fig. 1 The relation among hypertext, multimedia, and database

如图所示，将超文本技术用于多媒体信息管理，即所谓的超媒体技术；用数据库管理多媒体信息，即多媒体数据库；若将三种技术结合在一起，则为超媒体数据库。

### 6.4 其它

在超文本中融进专家系统思想可以开发出某种程度的智能超文本系统，去除盲目性，减少无效的浏览，或者人机交互式地进行推理，最终导致正确的结论或到达正确的目的地。

超文本与软件工程方法结合，可以产生优秀的CASE工具；超文本与计算机辅助设计及数据库结合可以建成工程数据库……

总之，超文本技术与其它计算机技术相结合，最大限度地发挥各自的优势，将对计算机技术的发展起到无法估量的影响。

## 7 结论

计算机技术的发展目标是合谐的应用、良好的人机接口。超媒体正是符合这种发展需求的非线性技术。它可以为用户提供图、文、声、像并茂的信息，以及符合人的思维习惯的交流方式，是使计算机渗透进各学科、各领域（包括家庭）的理想选择。但是，超媒体技术还处于婴儿时期<sup>[5]</sup>，许多问题尚有待深入研究。

### 参考文献

- 1 Devlin J & Berk E. Why Hypertext, Hypertext/Hypermedia Handbook, Intertext Publications, 1991, 9.
- 2 刘连芳, 吴玉军, 罗春晔等. 智能计算机接口与应用进展, 清华大学出版社, 1995.
- 3 Jakob Nielsen. HYPERText & HYPERMedia, Academic Press, 1989, 4.
- 4 刘连芳. 超文本技术. 计算机世界, 1993.
- 5 洪德威. 多媒体——通信发展的新领域. 电子展望与决策, 1994, 总第54期: 27~29.
- 6 Jeff conklin. A Survey of Hypertext, Mcc Technical Report, 1987.