

超文本与人工智能 Hypertext and Artificial Intelligence

罗春晔

Luo Chunye

(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)

摘要 以专家超文本为例介绍了与超文本技术、人工智能技术相关的技术问题,并从面向用户的角度侧重讨论了在任务实施过程中用户与计算机系统间的智能的分布。

关键词 超文本 人工智能 专家系统 专家超文本

Abstract The combination of hypertext and artificial intelligence is a new research subject. Taking experttext as an example, this paper introduces some of the problems associated with artificial intelligence and hypertext. The two technologies are discussed from a user-oriented perspective which focuses on the distribution of intelligence between users and these computer systems during task performance.

Key words hypertext, artificial intelligence, expert system, experttext

1 超文本

1.1 什么是超文本?

超文本是一种非线性的信息管理技术,它强调按照人的联想思维方式进行管理,给读者以最大限度的自由。超文本的基本组成单元是节点(node)和链(link),节点存储信息的内容,链描述节点之间的关系。

传统线性文本和超文本的差别主要体现在不同文本间的链接上。线性文本的作者希望读者能够按从头到尾的顺序阅读,尽管有一些参考文章指向前面或者后面的章节,或者本文并没有包括的其它文本。这种参考关系有的是明确的,如作者可以这样表述:“这将在后面讨论”,“已经在前面讨论了”。但是,大部分参考关系是不明确的,需要由读者自己来确立,并且保存在记忆里,也许一个好的线性文本作者的标志就是能使读者不费力地寻找这种关系。

与线性文本相反,超文本试图使这种参考关系变得明确。超文本用链表达信息间的关系,而且希望超文本系统的用户能够沿着链浏览。因此,展现在不同

的读者面前的可以是文章的不同章节或者是其它的相关资料,而且可以以不同的顺序出现,这都取决于读者在链接关系网中是如何浏览的。

1.2 超文本的技术问题

任何事物都有其正反两个方面。对超文本来讲,它结构灵活,对用户几乎没有约束,用户可以在超文本提供的时空领域自由任意地浏览。这种灵活性的另一面又是什么呢?

• 链的混乱

超文本作者根据信息的使用方法、用途来建立节点之间的关系,也就是链。我们一定有这样的经验:不同的读者在阅读同一本小说时,由于他们的兴趣爱好不同、读书的目的不同往往会选择不同的阅读顺序。这也就是说,不同的读者会采用不同的使用信息的方法,其实,同一个读者在不同的时候也可能采用不同的方法使用同一份信息。这无疑给超文本作者带来了极大的困难,就好比要用同样的素材写出好几本小说。结果是读者常常抱怨超文本链反映的是超文本作者认为重要的,而不是读者所关心的。所以,超文本作者经常就把所有相关的东西都链到一起,很显然,链一下子多了许多,这就给读者带来了新的甚至更大的混乱和麻烦,也使迷航问题更加严重。

解决问题的关键是要建立完备的用户模型,而且超文本系统应该能根据用户的实际情况提供相应的信息,隐藏没有用到的信息。

· 迷航

迷航是当前超文本存在的一个普遍的问题,它指的是读者沿着超文本链浏览时经历的迷失方向的感觉。虽然浏览路径中的每一条链的关系明确,但当读者处于若干层时,他所在的当前位置与浏览开始处的关系变得模糊。这好比玩同义词游戏,它要求前后紧挨的两个词必须是同义词,但当游戏结束时,经常发现最后一个词和第一个词几乎沾不着边。随着信息量的扩充,系统的不断扩大,迷航的问题也会越来越严重。前面所提到的链的混乱更使迷航问题雪上加霜。

迷航产生的原因主要是链的数目很多,但是对链的类型的语义描述却很缺乏。当前主要是靠提供附加功能来辅助超文本导航,方法很多,但最常用的是采用标注过的节点和链组成的图。但一个超文本系统使用这样一种图所能达到的水平不过相当于传统滚动式编辑器所能够达到的水平,这显然是不够的。

· 链型的理解与描述

超文本的一个主要技术问题一直是理解和描述超文本中与文本节点相连的链的类型。经验表明,链的类型是很丰富的,而且对超文本读者来说通常是不明确和难以描述的,此外,超文本用户对他们所要浏览的链的类型的细节(即链的语义)并没有足够的了解。

2 人工智能^[1]

2.1 什么是人工智能?

很难给人工智能下一个准确的定义,它可以不确切地定义成一系列模拟人类决策的编程技术。人工智能程序不必要求得出一个问题的准确答案,因为很多问题可能就没有一个准确答案。实践证明,人工智能技术能最有效地处理那些非常大而且难以准确定义的问题,因为这些问题非常大,所以处理每个可能是不现实的;因为它们不能被准确定义,所以在一个相当长的时间里我们很难写出一个可用计算机处理的算法。所以传统程序常常讨论过程和变量,而人工智能程序常常讨论经验、概率和模式。

在人工智能技术当中,最常用的是基于规则的专家系统。专家系统是模仿人类专家在特定领域中决策的工具设备,它试图象人类专家一样提供建议、帮助、指导。基于规则指的是一种专家系统的建立方法,这种专家系统由两个部分组成:规则集和推理机。规则集是提供描述一个专家系统所拥有知识的规则的集

合,规则通常包括两个部分:IF部分(条件部分)和THEN部分(结论部分)。推理机负责推理分析获得结论,实现推理的方法一般有两种:正向推理,它从应用于当前情况的规则集中选择满足条件的规则,得出相应的结论部分,再把新得出的结论作为条件重复前面的过程,直到得到问题的答案;逆向推理,它和正向推理相反,先寻找满足结论部分的规则,再去证实相应的条件部分。

2.2 专家系统的技术问题

通过早期专家系统的研制人们发现专家系统除了提供建议以外还必须具备自我解释的功能,这已经得到了大多数人的认可。当然也许也有一些专家系统的应用并不需要解释功能,比如系统的使用者是一个专家或是经常的用户,他完全熟悉专家系统所涉及的应用领域,但这样的情况毕竟是少数。

显然,要提供解释功能要求一个专家系统必须能够自我参考或者自我检查,这已超出对一个计算机系统的一般要求。解释专家系统最通常的形式是用同一个推理机提供建议,进行解释,这样,一个采用正向和逆向推理的专家系统通常提供这样一种类型的解释:因为1号和2号规则成立,所以提出X问题。这种方法需要将精炼的自然语言文本的各个小部分与这样一个解释的不同部分相联系,所以有充分的理由认为这种解释功能是不够完善的,在该领域中还有许多东西需要进一步深入研究。

专家系统解释功能存在的一个主要问题是上下文相关性,用户需要解释的目的不同,采用的方式不同,而这些不同只能当专家系统有一个操作人员的完备的模型时才能得以实现,但是用计算机来实现这样的用户模型是非常困难的。一个典型的失败例子就是智能教学系统的应用,它的问题在于无法给系统提供一个不同学生用户的完备的模型。

2.3 系统智能

计算系统包括一个计算机系统,一个或者多个操作者,和一个执行任务的环境。当考虑一个人机子系统时,采用的一个有效角度是识别智能在系统何处。对于传统的数据库技术,这种智能主要取决于用户,用户必须描述从数据中提取相关信息的逻辑条件;超文本系统要求操作者能够给出沿链浏览的顺序,从这个意义上讲,超文本与数据库是相似的,相反,专家系统的智能主要取决于计算机系统,用户通常只是回答专家系统提出的问题。实际上,在专家系统中,用户唯一可以控制的人机对话领域就是这个系统的解释部分。

3 超文本与人工智能

3.1 规范的专家系统与超文本^[2]

专家系统和超文本都可以用图论来描述,“专家系统中支持知识表达的语义网和作为超文本文档模型的语义网的不同在于节点内容的组成成分的本质和标注关系(即链)的数学特性”。

实际上,在这两种类型间有某种程度上的互补性,当用图论模型来表达它们时,这种互补性十分明显。超文本中的节点语义非常丰富,这是因为它们基本上自然语言文本,而专家系统中的节点语义显得相对贫乏,是因为节点包括一个规范的知识描述,例如规则。相反,专家系统中的链是准确描述的,超文本中的链则几乎没有语义上的描述。实际上,超文本链的问题在于它们有非常丰富的含义,以至超文本读者或作者无法理解,这当然与上面提及的“系统中的智能”这一特点是一致的:专家系统使用的是一种可计算的表达方式,但同时以省略知识的丰富涵义为代价(就是说所有的专家系统所涉及的知识领域都是非常局限的),与此相对,超文本却有着十分丰富的内容,以至超出了常人智力所能理解的范围,使人迷失方向。

3.2 专家超文本系统

专家超文本系统(Expertext)是专家系统(EXPERT systems)和超文本(hyperTEXT)的一种结合,这种结合方式带来的明显的积极效果是:该新系统既包含超文本中内涵丰富的节点又包含专家系统中准确定义的可计算的链,这是我们所期望的;然而相反的作用却是:节点间贫乏的知识表达和一系列丰富但难以计算的链的结合,这对使用这个假设系统的用户而言,是可怕的,他们将在理解节点和链上遇到很多困难。

一个十分有效的专家超文本系统,应当具有易为用户所理解而且丰富的节点,因为它们将构成自然语言文本,这也许是目前所有表达形式中最丰富的一种,当然不排除图解、数字、表格、图片等表达形式。类似地,通过使用推理机能够操作的准确描述和可计算的链,专家超文本系统用户可在节点表现的顺序上获得建议和指导,这样,超文本的导航问题就可以在很大程度上得以缓解或解决。

从“系统智能”这个角度而言,这样一个专家超文本系统的智能是分布在操作者与专家超文本计算机系统之间。人的智能体现在对节点内容的理解上,机器的智能体现在对网状链的浏览上,这种人机智能的分配显然是很明智的。人类擅长理解将事件、事实

同时罗列的复杂问题(即陈述性知识),但在处理冗长的关系链时却力不从心(即理解过程性知识)。相反,人工智能注重解决诸如棋类、装配游戏等问题和通用问题解决方法,即主要处理人类难以对付的冗长的因果链关系。Johnson-Laird的工作成果^[3]支持了这一观点:通过进一步丰富三段论的形式而不改变其逻辑内容,可提高人类解决三段论问题的能力。而机器长于解决用符号等表达的精练的逻辑问题,但在处理现实问题(如开汽车、火车)时却无能为力。

3.3 专家超文本的应用

专家超文本系统极有潜力替代或扩展当前的专家系统和超文本系统。对于专家系统来说,一个专家超文本系统能够提供一个附加的人机交互方式。过去,操作者仅能要求系统给予对某个问题的解释,但专家超文本系统可为操作者提供一个规则的文字叙述,同时操作者在操作过程中能够影响基本语义网的浏览。这样,操作者就可以指导专家系统,因为他能够理解专家超文本系统的意图。这样一个专家超文本系统就比传统专家系统有着更广泛的应用领域,因为操作者能够在操作过程中检测出不合适的规则调用,并能提供浏览网络的更有效的方法,而不必考虑这种浏览的复杂层次。

专家超文本系统可以补充和完善超文本系统,一般来说,专家超文本系统将提供给操作者或读者文字信息,并且指导他们浏览整个专家超文本文档中与任务相关的部分。指导有几种形式:系统有时自动提供文本的下一章节或列出建议读者下一步应当阅读的内容清单和次序。在专家超文本中,返回原来浏览过的节点并选择另一条路径是很容易的,因为可计算的链容许逆向推理,因此导航中的这类问题可以得到解决。显然,这样一个专家超文本系统可以防止与节点相关的正文的无效重显,除非这是读者的要求。

这种超文本类型的专家超文本系统有潜在的许多应用,Narula^[4]建立了一个超文本系统模型,作为训练第三世界医生治疗、诊断贫血的工具。一般,这种应用应当采用一个专家系统,但是这种方法的长处在于容许对超文本进行结构性的浏览,从而使Narula的系统不要求计算机有读者模型,这个系统存在的问题是:当读者在系统中迷失方向时,尽管计算机没有自动提供不恰当的信息,读者仍可能被不恰当或多余的信息所包围。一个专家超文本系统的实施应该能够减少这种导航问题且能让读者继续控制提供信息的内容。

3.4 人工智能与超文本的现状^[1,2]

把人工智能和超文本技术结合起来的商业产品

已经开始出现,它们用不同的方法实现,应用于不同的领域,但是毫无疑问,各种技术的融合已经开始。

Neuron Data 的 Nexpert Object 是一个适用于许多不同计算机的专家系统外壳,它把专家系统和 HyperCard 的卡堆结合起来;

德克萨斯仪器公司的 HyperTrans 系统能应用人工智能技术实现文档分析,并能建议建立最合适的链;

Cogent 软件公司的 HyperBase 系统在超媒体功能里融进了 Prolog 语言的专家系统功能,它可以开发智能的自适应的文档;

1st Class 专家系统中的 1st Class HT 是一个与超文本系统紧密结合的专家系统外壳,它试图产生能解释已知数据的规则,它被用来开发能维护在线文档的诊断系统;

Knowledge Garden 的 KnowledgePro Windows 是一个在微软窗口环境下的应用系统建造工具,它支持决策树和超文本;

贝尔通讯实验室的 Thoth I 是一个在超文本中嵌入语义的实验系统;

RUBRIC (基于规则的计算机信息检索系统)和 I3R (信息检索的智能接口)都提供了一个更智能和复杂的方法来对文档进行检索和分类;

CODER (复合文档专家检索系统)既能自动对文档进行语法分析,把文档转换为信息检索系统的标准形式,又能处理用户的询问;

JANUS 系统把 CRACK 专家系统 (厨房设计系统)与 PHI 超文本系统 (树状基于观点的信息系统)结合起来;

在 CASE (计算机辅助软件工程)领域,KBRA (基于知识的软件助手)允许多种类型的信息以任何次序输入,根据和操作者的对话能够自动生成所需的文档;类似,APRIS (应用软件原型实施系统)试图把需求自动转换为计算机设计的描述。

3.5 计算机辅助文档生成

专家超文本系统的一个成功应用就是从一个专家超文本文档生成线性文档,这主要是为了商业上和

工业上的文档重用。比如,几乎所有的制造产品,特别是计算机系统产品需要生成大量的不同的线性文档。即使是一个简单的产品也需要用户手册、服务手册和维修手册,还有许多与产品的制造和市场有关的手册和文档。文档重用的真实用意在于能用一个专家超文本文档产生出许多不同的线性文档。这种方法的好处是这些不同的文档可以缩短制作时间,而且能保持协调一致,有些是一个大的线性文档的一部分,有些仅公用文本的某些部分。比如,整个服务手册可能是更大而且更详细的维修手册的一部分,而用户手册则包括服务手册的某些部分和服务所不需要的附加的章节。另外,该系统还有一个优点就是支持多用户写作。作为一个 CASE 工具,专家超文本允许许多程序员、系统设计人员、管理人员、质量保证的评估人员、市场销售人员等等在软件的整个开发周期中共同制作一个单独的专家超文本文档,这样在软件交付使用的同时全套的必需文档也自动生成了。

4 小结

超文本和人工智能技术的结合,成功解决了超文本和人工智能技术各自遇到的问题,并且已经形成了一门新技术,它已在计算机辅助文档生成等许多领域得到广泛应用。

参考文献

- 1 Littleford Alan. Artificial Intelligence and Hypermedia. Hypertext/Hypermedia Handbook. USA: McGraw-Hill, 1991, 242~258.
- 2 Diaper D, Rada R. EXPERTTEXT: Hyperizing Expert System and Expertizing Hypertext. Hypermedia/Hypertext and Object-Oriented Databases. 1989.
- 3 Johnson-Laird. Mental Models: Towards a Cognitive Science of Language. Inference and Consciousness. UK: Cambridge University Press, 1983.
- 4 Narual. A prototype hyperdocument on anaemia using the GUIDE professional hypertext system. School of Tropical Medicine. UK: University of Liverpool, 1989.