

文章编号:1006-8309(2004)02-0020-03

超文本可用性评价:迷失程度的测量

周荣刚¹,李怀龙²,张侃¹

(1. 中国科学院心理研究所,北京 100101;2. 淮北煤炭师范学院,安徽 淮北 235000)

摘要:迷失成为限制超文本系统使用的一个严重障碍,以此为指标评价超文本系统的可用性越来越受到重视。文章从路径测量、心理模型测量和主观评价3个方面回顾了测量(评价)迷失程度的工具,并对各种方法进行了简要的评述,提出了测量工具本身研究和使用方面的建议。

关键词:超文本;迷失;路径测量;心理模型测量;主观评价

中图分类号:TB18 **文献标识码:**A

1 引言

在影响超文本可用性的因素中,迷失(lostness or disorientation)最为突出:由于对文本系统没有一个很清晰的认识,用户在超文本使用过程中往往不知道身处何处、不知道下一步应该浏览什么、很难找回浏览过的信息或者由于过多的跳转而偏离了主题。这不仅影响了信息的搜索,而且使用户备受挫折、失去兴趣,最终放弃。有数字表明:在商业网站上购物者找到所要信息的几率只有42%,而最终放弃寻找欲购商品的几率为62%^[1]。正因为如此,超文本可用性的研究中有60%是关于迷失问题的^[2]。超文本可用性如何,一个很重要的指标就是考察用户在超文本使用过程中出现了多大程度的迷失,因此关于超文本迷失现象的研究则主要集中在测量的方法上。本文就测量迷失程度的方法进行回顾,并予以简要评述。

2 路径测量法(Path Measurements)

2.1 Smith的路径测量法^[3]

Smith认为传统的HCI评价方法并不适用超文本系统的可用性评价,对以任务完成时间和错误率作为评价超文本可用性的指标提出了质疑,并认为衡量迷路程度应该是从等级化操作绩效入手而不是依据用户的主观感觉。在此基础上,Smith从用户是否能迅速搜索到所要信息角度第一次提出了一系列测量迷失程度的数量化方法。通过记录用户在完成搜索任务的过程中所浏览的各类节点的数目,把搜索路径和最优(最少)路径之间的路径差作为迷失程度的参数,其所记录的

参数如下:

T——节点总数目

D——不同节点的数目

R——完成某一任务所需访问的节点数目

S——搜索时,访问的节点数目

N——搜索时,访问的不同节点数目

V——检查或修改时,访问的节点数目

T和D提供搜索过程中完成任务的一般信息;R在这里被认为是最优路径;N、S和V主要是用来区别搜索和修改时访问的节点数目。当搜索任务完成得十分完美的时候, $T=D=R, S=T$ 。据此,Smith以点(N/S, R/N)和点(1,1)之间的距离提出了测量迷失程度的公式:

$$L = \sqrt{\left(\frac{N}{S} - 1\right)^2 + \left(\frac{R}{N} - 1\right)^2}$$

当L=0时,搜索任务完成的最好。结合被试的口语报告和视频记录分析,Smith认为当L<0.4时用户被试没有出现迷失现象;而当L>0.5时,用户表现出明显的迷失现象;至于L [0.4, 0.5], Smith没有发现相关资料,不能肯定,但认为L=0.42可能是一个敏感点,要进一步的研究。

2.2 链接-权重迷失测量(LWLM, link-weighted lostness metric)^[2]

链表示不同节点间信息的联系,链的功能的强弱,直接影响节点的表现力,也影响到信息网络的结构和导航的能力。Otter等人认为不同的链所起的启动效应(Priming Effect)是不同的,一个设计完美的超文本中的链应该能够最大限度启动用户对其他信息链的知觉或记忆。而且不同类型的

基金项目:教育部“十五”规划课题(FBB011067);中国科学院院长基金资助项目(JHJ02013)

作者简介:周荣刚(1978-),男,安徽当涂人,博士研究生,主要从事人因学和工程心理学方面的研究工作。

链提供给用户能预测链接终点的程度是不一样的。所以从搜索路径角度发展迷失程度的测量工具,就应该考虑链接的类型。通过文献回顾,Utter

等人考虑了4种与路径测量迷失程度有关的链接,并给予了4种链接造成迷失的权重。见表1。

表1 外延型链接的迷失程度权重

	关系链接(1对1)		专门链接(1对多)	
	结合型链接	注解型链接	连续型链接	分类型链接
诱发迷失的可能性	高	低	相对低	中度
迷失程度的权重	1	4	3	2

由此 Otter 等人结合 Smith 的研究提出了新的测量迷失程度的路径公式:

$$LWLM = L / (LW / 4)$$

这里 LWLM 是指链接 - 权重迷失测量值, L 是 Smith 公式里的迷失度, LW 是诱发迷失程度的权重, LWLM = 0 时表示没有迷失。

总的来说,路径测量法的核心就是通过客观记录构建数学公式来计算最优路径的偏离量,以此衡量迷失程度,作为超文本可用性的指标。与任务完成的时间相比,最优路径偏离可以敏感地反映出学习者超文本使用中网络空间的迷失程度^[2],但是路径测量法依然是以搜索任务或信息获取(information retrieve)为基础的,完成任务时所要求的最优路径并不鼓励用户浏览和探索,而且超文本系统各色各样,所以最优路径的判断是没有固定标准的。尽管如此,路径测量法作为一种评价工具还是很具有应用价值和良好发展前景的,毕竟在研究中可以使一些变量得到控制。

3 心理模型测量法(The mental models metric)

心理模型测量法是基于这样的假设:如果用户对超文本的组织结构能形成一个很好的心理模型,那么将有助于超文本的阅读,较为容易找到所要找的信息,出现迷失现象的可能性就小。心理模型实际上就是用户对超文本系统的认知表征(cognitive presentation),这种表征是以图式(schema)的形式产生的。在定位自身所处位置或作方位判断时,图式提供了空间参照体系或框架,帮助导航,这样的体系或框架同样存在于超文本信息空间中。如果超文本结构设置合理、符合人信息加工的特点、可用性强,那么用户在使用过程中就容易建立起文本结构的清晰的心理模型或图式,从而使参照框架更为完备、提供的知觉线索更为丰富,这样在操作的时候便可以形成对下一操作的预期,使迷失成为不可能。要以基于文本结构的心理模型来测量用户使用超文本过程中的迷失程度,并以此作为超文本可用性的评价指标,首

先要解决的问题就是找到能够测量或量化用户心理模型的方法。

3.1 Utter 等人的绘图(drawing)测量法^[2]

结合前人^[3]的研究,Utter 等人要求被试以盒子图(boxes)代替页面(节点)、以线条(line)代替链接,画出从起始页面(节点)到其所认为的特定信息所在页面之间的路径图,而这样的路径在实际超文本系统中也是存在的,因此计算二者之间的差异便可数量化测定用户的迷失程度:

- AD ——实际所画的节点数目
- RD ——要求所画的节点数目
- C ——正确画出的节点数目
- CCP ——正确画出且正确放对位置的节点数目
- LBE ——在出现第一个错误之前,所画出的层次结构的深度。

据此 Utter 等人提出了测量心理模型的公式:

$$AMMH = 1 / 3 (C / AD + CCP / RD + LBE / RD)$$

其中,AMMH 表示的是超文本心理模型正确率(The accuracy of mental models for hypertext), C/AD 测量的是所画节点的正确率, CCP/RD 测量的是节点位置放置的正确率, LBE/RD 表示的是第一个错误之前正确的深度。

AMMH 的取值范围是[0, 1],心理模型建构的质量随之渐进:0 表示心理模型建构的很差,该超文本系统容易导致用户的迷失;1 表示心理模型建构的很好,该超文本系统可用性很强。

3.2 自由回忆组串(clustering in free recall)测量法^[4~6]

自由回忆是让被试尽可能多地回忆在实验网站上看到的内容。如果被试对信息组织得好,那么信息较容易被提取。如果信息是没有组织的,帮助回忆的提取线索就较少,造成迷失的可能性就较大。组串是指记忆中目标信息的组织结果,反映了用户是如何组织文本信息的,如果超文本结构体系设计合理,用户以组串形式回忆项目(节

点)就会变得较为容易。通过临近回忆内容的分析,组串测量的校正比率(the adjusted ratio of clustering measure, ARC)可以被用来估计被试记忆组织的组织量,可以反映被试的记忆组织,以此作为超文本系统的可用性指标。计算 ARC 的公式是:

$$ARC = \frac{R - E(R)}{\max R - E(R)}$$

$$E(R) = \frac{\sum n_i^2}{N} - 1$$

其中 R 是指观察到的类别重复的总数;maxR 是类别重复的可能的最大数 (maxR = N - k;N 是所有回忆项目的总数,k 是回忆内容类别的总数);E(R)是期待类别重复的数目, n_i 是类别 i 中回忆出的项目数。ARC 的取值范围是[0,1],为 1 时表示信息组织得非常好,迷失的可能性很小。

总体来看,心理模型测量法试图描绘出人们知觉超文本信息的心理图式。如果说路径测量法反映了用户网络空间的迷失程度,那么心理模型则反映了用户概念空间的迷失程度,这为超文本系统的开发者和设计者更能提供有用信息。但是这种测量只是间接地反映了迷失程度,而且使用起来所受干扰因素多。相对路径测量法而言,这种方法用得较少。

4 主观测评法

尽管 Smith^[3]认为网络或者超文本中的迷失应该由操作的绩效来反映而不是根据用户的感觉加以衡量,Utter 等人^[2]也认为用户报告的主观迷失感觉并不可靠,但是在他们的研究中都用了用户的自我报告、视频分析或问卷。而且往往会出现这样的问题:由于个人认知风格的差异,有人就倾向于多点击页面或节点,同时也比较容易地找到自己所需要的信息;另外,用户使用网络或超文本不仅仅在于知道某种信息在哪儿,更重要的是理解信息。所以超文本可用性的研究不能忽视网络环境下的学习,一个网站的可用性如何也应该考虑到网站(超文本)的易学性,客观的记录是很难反映出学习效果好坏的。所以测量用户的主观迷路感觉是有必要的。

4.1 非线性媒体迷失程度评价工具 (Non-Linear Media Disorientation Assessment)^[7,8]

在 Smith 之前,Beasley 等人在一项关于认知地图结构的超媒体研究中,就曾以问卷的形式测量迷失程度。该主观问卷共有 10 个题目,迷失分值从 10 分到 50 分,10 分表示没有迷失,50 分表

示迷失的程度严重。后来 Baylor 也使用了该问卷,结果都表明主观迷失感和学习效果有很强的负相关。但是他们都没有报告该问卷本身的信度。

4.2 主观迷失感 (Perceived Disorientation) 评价工具

Ahuja 和 Webster^[9]发展了评价 Web 系统的主观迷失问卷,该问卷共有 7 个项目,每个项目都为陈述句,让被试从“非常不同意”到“非常同意”7 点中选择自己的评判分值。在周荣刚等人^[10]的研究中把这 7 个项目翻译成中文,Alpha 系数为 0.8579。结果^[9,10]表明:主观迷失感与导航绩效相比,更能预测超文本下的学习效果。该问卷的 7 个项目见表 2。

表 2 测量主观迷失感的 7 个项目(中英文)

我感觉迷路了。
I felt lost.
我觉得自己似乎在绕圈。
I felt like I was going around in circles.
再找到浏览过的网页比较困难。
It was difficult to find a page that I had previously viewed.
页面之间的导航对我来说是个问题。
Navigating between pages was a problem.
我不知道如何才能到达我想要找的位置。
I didn't know how to get to my desired location.
我觉得辨别不出方向了。
I felt disorientation.
浏览一会儿,我不知道下一步该去哪里。
After browsing for a while I had no idea where to go next.

与路径测量法和心理模型法相比,主观测评法使用更为便捷、费用低,更能预测超文本环境下用户的学习效果。虽然非线性媒体迷失程度评价工具并没有报告本身的可靠性,而主观迷失感问卷的检验也是基于特定信息的正式搜索任务,尽管如此,通过问卷测量用户的主观感受,以此进行超文本系统的可用性评价,是一种可行且便利的评价方法,最有可能适用于自然浏览行为或非正式搜索。

5 结语

迷失已成为限制超文本系统使用的一个严重障碍,从测量迷失程度入手进行超文本可用性评价越来越受到关注。最优路径测量、心理模型测量和主观评价是该领域研究的三个视角。就性质而言,主观评价和心理模型更能反映用户概念空间的迷失程度,而最优路径测量则反映了网络(物理)空间的迷失程度;就是否直接反映迷失程度

(下转第 25 页)

- [4] Irwin DE. Memory for position and identity across eye movements[J]. Journal of Experimental Psychology: Learning, Memory, and Cognition, 1992, 18:307-317.
- [5] Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions[J]. Nature, 1997, 390: 279-281.
- [6] Rensink RA. The dynamic representation of scenes[J]. Visual Cognition, 2000, 7:117-142.
- [7] Prinzmetal W, Presti DE, Posner MI. Does attention affect visual feature integration? [J]. Journal of Experimental Psychology: Human Perception Psychophys, 1986, 47:497-505.
- [8] De Valois RL, Yund EW, Helper N. The orientation and direction selectivity of cells in macaque visual cortex[J]. Vision Research, 1982, 22:531-544.
- [9] Stefurak DL, Boynton RM. Independence of memory for categorically different colors and shapes[J]. Perception & Psychophysics, 1986, 39:164-174.
- [10] Treisman A. Focused attention in the perception and retrieval of multidimensional stimuli[J]. Perception & Psychophysics, 1977, 22:1-11.
- [11] Simons DJ. In sight, out of mind: When object representations fail[J]. Psychological Science, 1996, 7:301-305.
- [12] Wheeler ME, Treisman AM. Binding in short-term visual memory[J]. Journal of Experimental Psychology, 2002, 131(1):48-64.
- [13] Rensink RA. Visual search for change: A probe into the nature of attentional processing[J]. Visual Cognition, 2000, 7:345-376.
- [14] Rensink RA, O'Regan JK, Clark JJ. To see or not to see: The need for attention to perceive changes in scenes[J]. Psychological Science, 1997, 8:368-373.

[收稿日期]2003-05-16

[修回日期]2004-02-19

(上接第22页)

而言,主观评价最为直接,最优路径测量其次,而心理模型测量法则间接反映迷失程度;就使用成本和使用分析是否便捷而言,主观评价成本最低而且便捷,最优路径测量需要编制程序,心理模型测量进行大样本分析起来较为困难;从更为适用的范围而言,心理模型测量适合超文本组织结构方面的探讨,如信息的组织方式、网络导航系统的设计,更能为系统开发者提供建议;最优路径测量可能更适合导航帮助设置、链接类型和任务类型等方面的可用性研究;主观评价则为一般的评价,受限制少、适用范围较大,但提供给设计者具体建议相对较少。这些测量工具本身还不成熟,最好结合使用,尤其是主观评价和客观记录的结合使用。以迷失度作为超文本可用性评价的指标时,还应结合其他评价方法或指标。关于测量迷失工具研究本身,一方面要对各种方法进行比较,相互参考;另一方面要继续精化。

以上所述测量工具的开发和形成多是基于正式的或非正式的特定信息搜索任务之上的,如何在鼓励支持浏览等自然的超文本使用环境中测量迷失程度值得考虑研究。

参考文献:

- [1] 刘正捷,张丽萍,陈燕. 可用性工程对 IT 产品的影响 [N]. 计算机世界报, 2002-06-04(36:B1、B2).
- [2] Otter M, Johnson H. Lost in hyperspace: metrics and mental models[J]. Interacting with Computers, 2000, 13:1-40.
- [3] Smith PA. Towards a practical measure of hypertext usability[J]. Interacting with Computers, 1996, 8(4):365-381.
- [4] Roenker DL, Thomppson CP, Brown SC. Comparison of measures for the estimation of clustering in free recall[J]. Psychological Bulletin, 1971, 76(1):45-48.
- [5] Zhang H, Salvendy G. The implications of visualization ability and structure preview design for web information search tasks[J]. International Journal of Human-Computer Interaction, 2001, 13(1):75-95.
- [6] 赵晨. 超文本环境中影响信息结构设计的因素[D]. 北京:中国科学院心理研究所, 2002.
- [7] Beasley RE, Waugh ML. Cognitive mapping architectures and hypermedia disorientation: an empirical study [J]. Journal of Educational Multimedia and Hypermedia, 1995, 4(2/3):239-255.
- [8] Baylor A. Perceived disorientation and incidental learning in a web-based environment: internal and external factors [J]. Journal of Educational Multimedia and Hypermedia, 2001, 10(3):227-251.
- [9] Ahuja JS, Webster J. Perceived disorientation: an examination of a new measure to assess web design effectiveness [J]. Interacting with Computers, 2001, 14:15-29.
- [10] Zhou RG, Li HL, Zhang K. The effect of contextual navigation aids and cognitive style on usability of hypertext [A]. Dai GZ. Proceedings of APHCI2002 [C]. Beijing: Science Press, 2002. 98-107.

[收稿日期]2003-06-20

[修回日期]2003-12-12