

新一代网络语言 Wiki

李学俊, 李龙澍, 徐 怡

(安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039;

安徽大学 计算机学院, 安徽 合肥 230039)

摘 要:在国外,新一代网络语言 Wiki 已经广泛应用于建立网络知识库,但国内的研究和应用刚刚起步,出现了如维基百科、天下维客等网络知识库。首先介绍了 Wiki 语言的来源、发展历史;然后详细叙述了 Wiki 语言的实现机制,并与 HTML 语言作了简单的比较;最后列举了主流的 Wiki 语言的实现平台——Wiki 引擎,并给出了基于现有 Wiki 引擎的 Oracle 网络知识库示例。

关键词:Wiki 语言; Wiki 引擎; 超文本链接

中图分类号:TP393

文献标识码:A

文章编号:1673-629X(2007)01-0085-03

A New Generation Network Language Wiki

LI Xue-jun, LI Long-shu, XU Yi

(Ministry of Education Key Lab. of IC & SP, Anhui University, Hefei 230039, China;

School of Computer Science and Engineering, Anhui University, Hefei 230039, China)

Abstract: Abroad the new generation network language Wiki has widely applied into building network knowledge base; however, domestic research and application on Wiki arrives just such as Wikipedia, allWiki. Firstly the original and history of Wiki is introduced. Then Wiki language, which is implementation mechanism, is presented in details, and is contrasted with HTML. Finally several popular Wiki engines which are implementation platforms of Wiki language are introduced, and an example of Oracle network knowledge base based on existing Wiki engine is given.

Key words: Wiki language; Wiki engine; HTML

0 引 言

Wiki 一词来源于夏威夷语的“wee kee wee kee”,原本是“快点快点”的意思,被译为“维基”或“维客”。Wiki 指一种超文本系统实现技术,用以实现超文本系统。这种超文本系统支持面向社群的协作式写作,同时也包括一组支持这种写作的辅助工具。可以在万维网/Web 的基础上对 Wiki 文本进行浏览、创建、更改,而且创建、更改、发布的代价远比 HTML 文本小;同时 Wiki 系统还支持面向社群的协作式写作,为协作式写作提供必要帮助;最后,Wiki 的协作者自然构成了一个社群,Wiki 系统为这个社群提供简单的交流工具。与其它超文本系统,如博客(Blog)、BBS 相比,Wiki 有使用方便及开放的特点,所以 Wiki 系统可以帮助人们在一个社群内共享某领域的知识,建立共享知识社区平台^[1]。

1 Wiki 语言的历史

1995 年沃德·坎宁安(Ward Cunningham)为了方便模式社群的交流建立了一个工具——波特兰模式知识库(Portland Pattern Repository)。在建立这个系统的过程中,Ward Cunningham 创造了 Wiki 的概念和名称,并且实现了支持这些概念的服务系统。这个系统是最早的 Wiki 系统。从 1996 年至 2000 年间,波特兰模式知识库围绕着面向社群的协作式写作,不断发展出一些支持这种写作的辅助工具,从而使 Wiki 的概念不断得到丰富。同时 Wiki 的概念也得到了传播,出现了许多类似的网站,如维基百科全书(Wikipedia)、维基文库(Wikisource)等,也开发出了很多支持 Wiki 技术的软件系统,如 MediaWiki^[2]。

2 Wiki 语言的特点

Wiki 语言具有如下几个特点:

1) 使用方便。

(1) 维护快捷:快速创建、存取、更改超文本页面。

(2) 格式简单:用简单的格式标记来取代 HTML 的复杂格式标记,类似所见即所得的风格。

(3) 链接方便:通过简单标记,直接以关键字名来建立链接(页面、外部连接、图像等)。

收稿日期:2006-04-25

基金项目:国家自然科学基金(60273043);安徽省自然科学基金(050420204);安徽省教育厅自然科学研究项目(2006KJ098B)

作者简介:李学俊(1976-),男,安徽六安人,博士研究生,研究方向为机器学习、智能软件、Web 挖掘;李龙澍,博导,教授,研究方向为智能软件、知识工程、软件体系结构。

(4)命名平易:关键字名就是页面名称,并且被置于一个单层、平直的名空间中。

2)有组织。

(1)自组织的:同页面的内容一样,整个超文本的组织结构也是可以修改、演化的。

(2)可汇聚的:系统内多个内容重复的页面可以被汇聚于其中的某个,相应的链接结构也随之改变。

3)可增长。

(1)可增长:页面的链接目标可以尚未存在,通过点击链接;可以创建这些页面,从而使系统得到增长,知识得到累积。

(2)修订历史:记录页面的修订历史,页面的各个版本都可以被获取。

4)开放性。

(1)开放的:社群成员可任意创建、修改、删除页面。

(2)可观察:系统内页面的变动可以被访问者观察到。

3 Wiki 语言规则

可以在万维网/Web 的基础上对 Wiki 文本进行浏览、创建、更改。这样需要 Wiki 的实现机制——Wiki 语言。下面就排版、格式控制、图像、表格、超级链接等几个方面介绍 Wiki 语言^[3,4]。

3.1 排版

(1)标题。

= 文档内容 =, 表示一级标题; == 文档内容 ==, 表示二级标题; === 文档内容 ===, 表示三级标题等等, 依次类推。

(2)段落。

用空行开始一个新段落。可以使用 HTML 语言中的
, 但是最好不要这样。

(3)符号列表。

行首以符号 * 开始, 符号越多, 表示级别越低。例如, 代码如下:

* 一级列表 * * 二级列表 * * * 三级列表, 符号越多, 表示级别越低。

显示结果:

①一级列表

②二级列表

③三级列表, 符号越多, 表示级别越低

(4)数字列表。

行首以符号 # 开始, 符号越多, 表示级别越低。

例如, 代码为: # 一级数字列表; ## 二级数字列表; ### 三级数字列表。符号越多, 表示级别越低。

显示结果:

①一级数字列表

②二级数字列表

③三级数字列表, 符号越多, 表示级别越低

(5)水平线。

符号标记为 - - - -。

3.2 格式控制

和 HTML 标记相同, 粗体: 文档内容; 斜体: <i>文档内容</i>; 下划线: <u>文档内容</u>; 删除线: <strike>文档内容</strike>; 文字居中: <center>文档内容</center>。

3.3 超级链接

语法规则 1: [[名字空间: 链接目标(文档)|链接名称]], 名字空间可以为: Mozilla, Firefox。页面显示时, 自动隐藏名字空间和圆括号。

例如, 代码如下:

链接到: [[Firefox: http://wiki.donews.com|DoNews Wiki 国内最大的 IT 知识库]]

显示结果为: 链接到: DoNews Wiki 国内最大的 IT 知识库

语法规则 2: [URL 地址], 左右中括号可以省略。

3.4 图 像

语法规则: [[Image: 图像文件|参数 1|参数 2|说明文字]]

示例:

(1)在页面中显示一幅图像, 将宽度设为 100px:

[[Image: WikiLogo.jpg|100px]]

(2)在页面中显示图像的缩略图, 将宽度设为 40px, 并添加说明文字“WikiLogo”:

[[Image: WikiLogo.jpg|thumb|100px| WikiLogo]]

3.5 表 格

(1)Table 标记。

语法规则: || 表格参数 ||

(2)TD 标记。

语法规则: | “单元格参数” | 单元格内容

(3)TH 标记。

使用方法和 TD 类似。

(4)TR 标记。

语法规则: | - “参数”

示例, 代码为: || border = 1 | 火狐 | Firefox | - | 雷鸟 | Thunderbird ||

页面显示表格为:

火狐	Firefox
雷鸟	Thunderbird

3.6 其 它

Wiki 还支持环境变量, 商业符号、数学符号、希腊字符等特殊符号。

例如: || CURRENTYEAR || 表示当前年份, || SERVER || 表示当前服务器。

3.7 与 HTML 语言的比较

Wiki 和 HTML 语言都是一种生成超文本文档的标记语言, 但在功能和语法实现上, 都存在着很大区别。

HTML 用于开发设计网页,存放于服务器端,用户学习 HTML 中复杂的标记语言后才能设计出网页,设计网页的目的是让用户浏览,而不是编辑。但是支持 Wiki 技术的 Wiki 语言,要实现让用户自由地、平等地编辑网页,以实现协作共享的知识平台。HTML 语言是面向专业的开发人员,而 Wiki 语言是面向所有用户,由于两者功能目标不同,所以 Wiki 语言要比 HTML 语言更加简单、灵活。

4 实现平台——Wiki 引擎

Wiki 语言应用于建立网络知识社区,用户可以在社区内自由、平等地交流协作。随着 Wiki 语言的应用,已经出现了很多种支持该语言的引擎软件。按照引擎软件运行环境分为 3 类:运行于 PHP + MySQL 环境的 MediaWiki, CooCooWakka 等;运行于 JSP 环境的 JSPWiki 等;运行于 ASP 环境的 JsWiki, OperatorWiki 等^[5,6]。现将主流的引擎软件介绍如下。

4.1 MediaWiki

MediaWiki 全球最著名的开源 Wiki 引擎,运行于 PHP + MySQL 环境。MediaWiki 的原作者为德国的 Magnus Manske。从 2002 年 2 月 25 日被作为维基百科全书的系统软件,并有大量其他应用实例。目前 MediaWiki 的开发得到维基媒体基金会的支持。MediaWiki 一直保持着持续更新,当前最新版本为 1.4.13(2006-01-05)和 1.5.6(2006-01-19)稳定版本。

4.2 CooCooWakka

CooCooWakka 是一个中文 Wiki 引擎程序,运行于 PHP + MySQL 环境下。其最新版本为 0.0.9rc2 (2005. 12. 19)。CooCooWakka 是 2004 年最流行的中文 Wiki 引擎程序之一。因为是较早中文(简体、繁体)和多语种的 Wiki 引擎程序之一,至少上百个中文网站采用了这套引擎程序。其具有十分小巧、安装快速简单、运行速度较好的特点。

4.3 TWiki

TWiki 是一个开源的 Wiki 引擎程序。软件定位为灵活、强大、易于使用的企业协作平台,运行于 Perl 环境。TWiki 从 2001 年开始开发,大约每年发行一个重要版本。最新稳定版本为 2004-09-04 版本,最新 beta 版本为

2005-12-17 版本。Twiki 被很多大型商业公司采用,如 Yahoo, SAP, Motorola, Wind River 等。

4.4 PmWiki

PmWiki 是一个基于 PHP 的 Wiki 引擎,其主要特点是小巧易用,各项功能比较完善。PmWiki 在 2.0.6 版本之后已经较好地支持了中文。PmWiki 由 Patrick Michaud 开发,目前最新版本是 2.0.13(2005-11-10)和 2.1-beta17(2005-12-29)。

4.5 JSPWiki

JSPWiki 是一个纯 JSP/Servlet 编写的基于 J2EE 平台的 Wiki 引擎。具有支持中文,版本比较、权限管理等功能。

5 总结

随着 Wiki 语言的应用和发展, Wiki 语言将成为网络知识社区的主流交流协作语言,基于 Wiki 语言的共享知识平台将会应用于各个知识领域,形成分布式的网络知识库,促进知识的交流与发展。笔者采用现有的 Wiki 引擎 OperatorWiki,实现了基于 Wiki 的 Oracle 网络知识库,如图 1 所示。

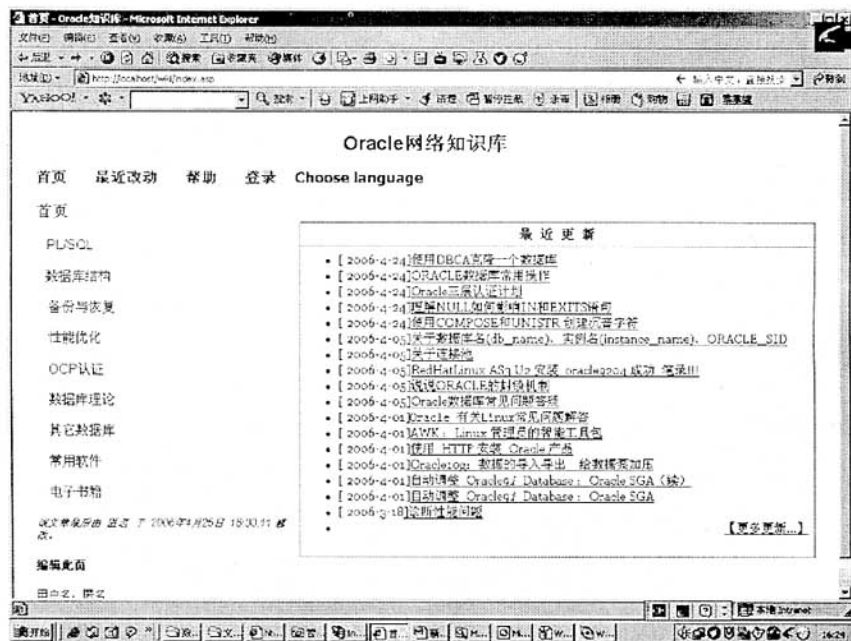


图 1 基于 Wiki 的 Oracle 网络知识库

参考文献:

- [1] 钟翠玲. Wiki 是一种社会运动[EB/OL]. 2006-12-06. <http://tech.sina.com.cn/1/2006-04-11/1159898896.shtml>.
- [2] 龙 楼. WIKI 扫盲手册[EB/OL]. 2006-01-23. <http://homepage. yesky.com/489/2283489.shtml>.

铅球决赛、跳高决赛^[5]。

并作如下规定：

(1)进行男子(女子)项目径赛的同时交叉进行女子(男子)项目的田赛。

(2)比赛分学女、学男、教工男、教工女四组。

(3)同一比赛项目尽量女子项目先比赛。

为了简化问题,排除了在多个比赛场所同时进行比较的情况。将每个比赛项目看成不同的事件,共得 26 个事件。

1. 学女 100 米预赛;2. 学女 100 米复赛;3. 学女 100 米决赛;4. 学男 100 米预赛;5. 学男 100 米复赛;6. 学男 100 米决赛;7. 学男 200 米预赛;8. 学男 200 米决赛;9. 学女 4×100 米决赛;10. 学男 4×100 米决赛;11. 教工男 4×100 米决赛;12. 学女 800 米决赛;13. 教工女 800 米决赛;14. 教工男 800 米决赛;15. 学男 1500 米决赛;16. 学女跳远;17. 教工女跳远;18. 教工男跳远;19. 学女铅球;20. 学男铅球;21. 教工女铅球;22. 教工男铅球;23. 学女跳高;24. 学男跳高;25. 学男标枪;26. 教工男标枪。

由问题描述,为求得符合要求的 R-时刻表,首先分析出各个项目之间的约束条件。要考虑到运动员兼项,例如许多运动员既报短跑也报跳高跳远,所以尽量不要让这些项目同时比赛。在这里仅给出一些项目之间的约束条件:

$R(1,4) = \{m\}; R(1,18) = R(1,21) = \{=\}; R(4,13) = \{m\}; R(4,16) = \{s\}; R(4,17) = \{=\}; R(13,2) = \{m\}; R(13,6) = \{f\}; R(2,5) = \{m\}; R(2,19) = \{s\}.$

3.2 实验与分析

为验证算法的有效性,在 Windows 操作系统下,用 Delphi 实现了上述算法。

通过图 2 界面将所有事件之间的关系输入,利用上述算法得出事件发生的先后顺序。其中合并端点记录了每个事件端点在化简后的矩阵中所对应的端点号。R 时刻表给出了事件端点的排列顺序,即事件发生的先后等关系。最终可得运动会项目安排顺序如图 3 所示。

从事件的安排顺序来看,该算法能够较好地实现一般运动会的项目安排。但它只给出了一种安排顺序。

4 结束语

改进了求解时间规划问题 R 时刻表算法中相容子集

算法,使得相容子集算法时间复杂度减小一半。并通过实验验证了该算法的实用性。

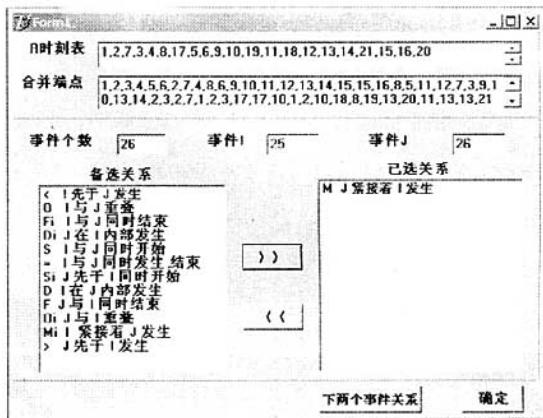


图 2 R-时刻表交互界面

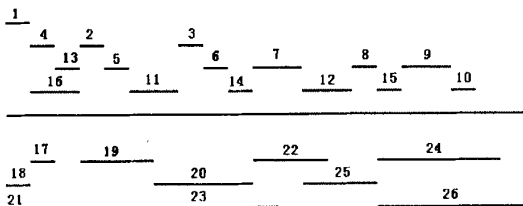


图 3 项目安排顺序

但该算法还具有一定的缺陷和不完善的地方。算法的主要不足是:很难设置事件之间的约束关系,只能排列出事件发生的先后顺序,不能给出事件发生的起始时间和结束时间,只适用于单成分。实际生活中许多问题是多成分的,关于多成分和涉及时间宽度约束的时间规划问题将在以后讨论。

参考文献:

- [1] Allen J F. Maintaining knowledge about temporal interval[J]. Communication of the ACM, 1983, 26(11): 43-52.
- [2] 张 钺, 张 铃. 时间规划的关系矩阵法[J]. 计算机学报, 1991, 14(6): 411-422.
- [3] 张 钺, 张 铃. 问题求解的理论及应用[M]. 北京: 清华大学出版社, 1990: 308-312.
- [4] 方思行. 一种有效的 R-时刻表综合算法[J]. 华南理工大学学报: 自然科学版, 1995, 23(9): 43-46.
- [5] 中国体育协会. 田径竞赛规则[M]. 北京: 人民体育出版社, 2002: 123-145.

(上接第 87 页)

- [3] 和 洋. 天下维客使用手册[EB/OL]. 2006-05-10. <http://www.allwiki.com/wiki/Help:%E7%BC%96%E8%BE%91>.
- [4] 飞 鸟. 10 大流行 Wiki 引擎[EB/OL]. 2005-06-02. [http://www.allwiki.com/wiki/Wiki%E5%BC%95%E6%](http://www.allwiki.com/wiki/Wiki%E5%BC%95%E6%93%8E)

93%8E.

- [5] Gustavo. JSPWiki 引擎[EB/OL]. 2006-04-07. <http://www.jspwiki.org/>.
- [6] Pagallo G. MediaWiki 引擎[EB/OL]. 2006-03-04. <http://www.Mediawiki.org/>.