



### 送9264元话费

两年不用交话费新老号均可办理月送1.6G流量、1250分钟话费



沃·3G

搜索

用户名：

密码：

飞诺网 加入收藏

飞诺网 软件编程 C C++ Java VB Delphi Foxpro 汇编语言 游戏开发 移动开发 软件工程师 软工与管理 VC shell编程 C#

C++系列教程 C++实例 C++技术文档 C++/C语言函数 Mangos

飞诺网开发频道

您当前的位置：飞诺网 >> c/c++ >> C++技术文档

# 用 C 语言编写一个网络蜘蛛

www.diybl.com 时间：2007-12-23 作者：佚名 编辑：andyoo 点击：5462 [评论]

[北京总院 三甲医院 间质性肺炎](#)  
国家津贴专家亲诊,间质性肺炎在线咨询 全国肺  
病中医诊疗中心,值得信赖!  
[www.feibing999.com](#)

[名品商标转让](#)  
专业权威商标转让机构。正规商标局备案机构 10  
多年专业商标转让经验!专业.诚信  
[www.tmtmw.com](#)

[女人喜欢又大又坚挺的秘密\(图\)](#)  
巩固爱情 享受性福生活 纯天然提取 见效快不反  
弹  
[www.168vi.com](#)

```

/*****
功能：从 src 中分析出网页链接，并加入到当前节点的子节点上
*****/

void GetLink(char * src)
{
    char * pa, * pb, * pc;
    char * myanchor = 0;
    int len = 0;

    pa = src;
    do {
        if((pb = strstr(pa, "href=")))    {
            pc = strchr(pb + 6, '"');
            len = strlen(pb + 6) - strlen(pc);
            GetMemory(&myanchor, len);
            memcpy(myanchor, pb + 6, len);
        }
        else if((pb = strstr(pa, "href=")))    {
            pc = strchr(pb + 6, '"');
            len = strlen(pb + 6) - strlen(pc);
            GetMemory(&myanchor, len);
            memcpy(myanchor, pb + 6, len);
        }
        else if((pb = strstr(pa, "href=")))    {
            GetAfterPosWithSlash(pb + 5, &pc);
            len = strlen(pb + 5) - strlen(pc);
            GetMemory(&myanchor, len);
            memcpy(myanchor, pb + 5, len);
        }
    }
    else {goto __returnLink ;}

    /*
    if(DEBUG)    {
        if(strcmp(NodeCurr->dir, "/"))    fprintf(stdout, "%s http://%/s/%s/%s
", myanchor, NodeCurr->host, NodeCurr->dir, strcmp(NodeCurr->page, "")?NodeCurr->page:"");
        else    fprintf(stdout, "%s http://%s%s%s ", myanchor, NodeCurr->host, NodeCurr-

```

频道地图

- C++
- C++系列教程
- C++实例
- C++技术文档
- C++/C语言函数
- Mangos
- C
- C++
- Java
- VB
- Delphi
- Foxpro
- 汇编语言
- 游戏开发
- 移动开发
- 软件工程师
- 软工与管理
- VC
- shell编程
- C#

- C++技术文档相关文章
- c++--SDI的启动和关闭的调试...
- Effective STL 学习（1）
- 一个用C++模拟.Net事件机制...
- using libxml2连载(2 使用简...
- CRuntimeClass的应用
- 如何学好C语言和其他计算机...
- 我所用过的 VC 技巧 之 WinS...
- getch()、getche()和getchar(...
- 在vista下,如何判断某个进程...
- 简单的Hash\_Table
- 学习vc的经验之谈
- c、c++程序的存储空间布局
- 在Vista操作系统中通过manif...
- algorithms websites
- [转载] 约瑟夫问题的数学方法(...
- 关于作用域的心得
- 为什么C++（感谢waterwalk...
- 在MFC类中各种类的指针的获...
- vmware下为linux挂载(mount...
- const、mutable、volatile意...
- 快速幂取模求a的b次方余c
- 对UTF8编码的初步认识！
- 皇冠足球.金沙投注网.泰山投...
- Turbo C 2中的常用快捷键

- C++技术文档热门文章
- MFC中edit控件的使用
- cout<<格式输出的使用
- C/C++ 文件读写操作总结
- C语言 给字符数组赋值的方法
- assert() 函数用法

```

>dir, strcmp(NodeCurr->page, "")?NodeCurr->page:"");
    }
*/
    if(strlen(myanchor) > 0)    AddChildNode(NodeCurr, myanchor);
    if(pc + 1)    pa = pc + 1;
}while(pa);
__returnLink:
    return;
}

/*****
功能：为当前节点增加子节点
*****/
void AddChildNode(WEBNODE * node, char * src)
{
    int WebPort, len;
    char * WebHost = 0, * PageAddress = 0, * WebDir = 0, * pC = 0;
    WEBNODE * NewNode;
    char filename[MAXFILENAME + 1] = "";
    char IsFromRoot = 0;

    if(!src)    return;
    if(!strncasecmp(src, "mailto:", strlen("mailto:")))    return ;
    if(strstr(src, ".css"))    return;
    if(strstr(src, ".xml"))    return;
    if(strstr(src, ".ico"))    return;
    if(strstr(src, ".jpg"))    return;
    if(strstr(src, ".gif"))    return;
    if(strstr(src, "javascript:"))    return;
    if(strstr(src, "+"))    return;

    ret = GetHost(src, &WebHost, &PageAddress, &WebPort, &WebDir);
    if(ret)    {
        len = strlen(node->host);
        GetMemory(&WebHost, len);
        strcpy(WebHost, node->host);

        WebPort = node->port;

        IsFromRoot = !strncmp(src, "/", 1);
        if(IsFromRoot && (src + 1))    Rstrchr(src + 1, '/', &pC);
        else if(!IsFromRoot)    Rstrchr(src, '/', &pC);
        else    pC = 0;

        if(pC)    {
            if(IsFromRoot)    len = strlen(src + 1) - strlen(pC);
            else    len = strlen(src) - strlen(pC) + strlen(node->dir) + 1;
            GetMemory(&WebDir, len);
            if(IsFromRoot)    memcpy(WebDir, src + 1, len);
            else    {memcpy(WebDir, node->dir, strlen(node->dir)); strcat(WebDir, "/"); memcpy(WebDir + strlen(node->dir) + 1, src, strlen(src) - strlen(pC));}

            if(pC + 1)    {
                len = strlen(pC + 1);
                GetMemory(&PageAddress, len);
                strcpy(PageAddress, pC + 1);
            }
            else    {

```

[C/C++ 笔试、面试题目大汇总](#)  
[DOS内部命令和批处理命令列表](#)  
[C++ string详解](#)  
[C++中STL中的map用法详解](#)  
[华为C/C++笔试题\(附答案, ...\)](#)

```
len = 1;
GetMemory(&PageAddress, len);
memcpy(PageAddress, e, len);
}
}
else {
```

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
							18	19	20	21	22					

[提供虚拟机（intel vt）技术支持](#)

如果图片或页面不能正常显示请[点击这里](#)

 发表评论

[【文章投稿】](#) [【收藏此页】](#) [【飞诺社区】](#) [【发表评论】](#) [【关闭】](#)

上一篇：解决：fatal error C1189: #error : include 'stdafx.h' before including this file for PCH      下一篇：linux的sh脚本编程

C++技术文档推荐文章

- 编程中遇到的一些头文件(c++)  
linux下产生core文件  
在ActiveX控件中引入窗体技术  
动态内存分配之11.1  
C++ strtok应用方式浅析  
一个典型网络引擎的设计与实现  
对于C++数据类型说明  
计划任务
- 关于accept函数需要注意的一点  
Qt MetaObject system 详解之三：QMetaObject接口...  
最近工作不太忙了，呵呵，写点东西润色一下枯燥的生活  
Effective C++连载  
C++中extern "C"含义深层探索  
vc字符串总结（转）  
关于STRUCT的几点认识  
用协变模板参数进行矢量对象操作

文章评论

暂无数据！

请您留言

昵称：

验证码：

[注册会员](#) [会员登录](#)

发表评论

也许你对以下文章感兴趣

- 首日封
- c++派生类转换为基类与public、p...
- 几分钟教会您使用Visual C++ 6.0设计程序
- Pascal简介
- 流操作符的重载
- NMHDR和NM\_TREEVIEW转换问题
- 返回引用类型的函数
- fstream seekg tellg && dos2unix && 对一个文件边读边写



