

# 详解crifan的C#库：crifanLib.cs

版本：v1.0

Crifan Li

## 摘要

本文主要介绍了我，crifan，的C#库：crifanLib.cs的功能和用法。



## 本文提供多种格式供：

在线阅读	<a href="#">HTML</a> <sup>1</sup>	<a href="#">HTMLs</a> <sup>2</sup>	<a href="#">PDF</a> <sup>3</sup>	<a href="#">CHM</a> <sup>4</sup>	<a href="#">TXT</a> <sup>5</sup>	<a href="#">RTF</a> <sup>6</sup>	<a href="#">WEBHELP</a> <sup>7</sup>
下载（7zip压缩包）	<a href="#">HTML</a> <sup>8</sup>	<a href="#">HTMLs</a> <sup>9</sup>	<a href="#">PDF</a> <sup>10</sup>	<a href="#">CHM</a> <sup>11</sup>	<a href="#">TXT</a> <sup>12</sup>	<a href="#">RTF</a> <sup>13</sup>	<a href="#">WEBHELP</a> <sup>14</sup>

HTML版本的在线地址为：

[http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/html/crifanlib\\_csharp.html](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/html/crifanlib_csharp.html)

有任何意见，建议，提交bug等，都欢迎去讨论组发帖讨论：

[http://www.crifan.com/bbs/categories/crifanlib\\_csharp/](http://www.crifan.com/bbs/categories/crifanlib_csharp/)

## 修订历史

修订 1.0	2013-08-20	crf
--------	------------	-----

1. 从C#学习心得提取出来成立独立的book

2. 更新了N多函数的代码和用法

<sup>1</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/html/crifanlib\\_csharp.html](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/html/crifanlib_csharp.html)

<sup>2</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/htmls/index.html](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/htmls/index.html)

<sup>3</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/pdf/crifanlib\\_csharp.pdf](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/pdf/crifanlib_csharp.pdf)

<sup>4</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/chm/crifanlib\\_csharp.chm](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/chm/crifanlib_csharp.chm)

<sup>5</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/txt/crifanlib\\_csharp.txt](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/txt/crifanlib_csharp.txt)

<sup>6</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/rtf/crifanlib\\_csharp.rtf](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/rtf/crifanlib_csharp.rtf)

<sup>7</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/webhelp/index.html](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/webhelp/index.html)

<sup>8</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/html/crifanlib\\_csharp.html.7z](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/html/crifanlib_csharp.html.7z)

<sup>9</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/htmls/index.html.7z](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/htmls/index.html.7z)

<sup>10</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/pdf/crifanlib\\_csharp.pdf.7z](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/pdf/crifanlib_csharp.pdf.7z)

<sup>11</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/chm/crifanlib\\_csharp.chm.7z](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/chm/crifanlib_csharp.chm.7z)

<sup>12</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/txt/crifanlib\\_csharp.txt.7z](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/txt/crifanlib_csharp.txt.7z)

<sup>13</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/rtf/crifanlib\\_csharp.rtf.7z](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/rtf/crifanlib_csharp.rtf.7z)

<sup>14</sup> [http://www.crifan.com/files/doc/docbook/crifanlib\\_csharp/release/webhelp/crifanlib\\_csharp.webhelp.7z](http://www.crifan.com/files/doc/docbook/crifanlib_csharp/release/webhelp/crifanlib_csharp.webhelp.7z)

---

## 详解crifan的C#库 : crifanLib.cs:

Crifan Li

版本 : v1.0

出版日期 2013-08-20

版权 © 2013 Crifan, <http://crifan.com>

本文章遵从 : [署名-非商业性使用 2.5 中国大陆\(CC BY-NC 2.5\)](http://creativecommons.org/licenses/by-nc/2.5/)<sup>15</sup>

---

<sup>15</sup> [http://www.crifan.com/files/doc/docbook/soft\\_dev\\_basic/release/html/soft\\_dev\\_basic.html#cc\\_by\\_nc](http://www.crifan.com/files/doc/docbook/soft_dev_basic/release/html/soft_dev_basic.html#cc_by_nc)

---

# 目录

前言	viii
1. 本文目的	viii
2. crifanLib.cs的由来	viii
3. 最新最全的crifanLib.cs完整代码下载	viii
4. crifanLib.cs所包含的引用 ( using )	viii
4.1. crifanLib.cs中的宏定义	viii
4.2. crifanLib.cs中引用的所有的库	ix
4.3. crifanLib.cs中各个宏的解释	x
4.3.1. USE_GETURLRESPONSE_BW	x
4.3.2. USE_HTML_PARSER_SGML和USE_HTML_PARSER_HTMLAGILITYPACK	x
4.3.3. USE_DATAGRIDVIEW	xi
4.3.4. USE_JSON	xii
5. crifanLib.cs中的全局变量, 初始化代码, 私有函数	xii
1. crifanLib.cs之TreeView/TreeNode	1
1.1. 查找TreeNode的根节点:findRootTreeNode	1
1.2. 取消节点的高亮:unHighlightNode	2
1.3. 高亮TreeNode:highlightNode	2
2. crifanLib.cs之Unit Conversion	4
2.1. 盎司转千克:ounceToKiloGram	4
2.2. 千克转盎司:kiloGramToOunce	4
2.3. 英镑转千克:poundToKiloGram	5
2.4. 千克转英镑:kiloGramToPound	5
2.5. 英尺转厘米:inchToCm	6
2.6. 厘米转英尺:cmToInch	6
3. crifanLib.cs之Values	7
3.1. 和Javascript中Math.Random()等价的函数:mathRandom	7
4. crifanLib.cs之Time	8
4.1. 计算 ( 代码执行 ) 时间消耗 ( 的时间段/时长 ):elapsedTimeSpanInit,getElapsedTimeSpan	8
4.2. 获得 ( 从epoch时间纪元以来的 ) ( 以毫秒为单位的 ) 当前时间:getCurTimeInMillisec	9
4.3. 将毫秒转换为 ( 自1970年1月1日以来的 ) 本地时间:milliSecToDateTime	9
4.4. 将Javascript中的"new Date(xxx)"转换为C#中的DateTime变量:parseJsNewDate	10
5. crifanLib.cs之String	11
5.1. 格式化字符串中间对齐左右填充:formatstring	11
5.2. 初始化null的字符串位空字符串":emptyStringArray	11
5.3. 将感叹号"!"强制编码为"%21":encodeExclamationMark	12
5.4. 将"%21"解码为感叹号"!:decodeExclamationMark	12
5.5. 从字符串中提取单个的子字符串:extractSingleStr	13
5.6. 组合参数列表 ( 变成&xxx=yyy ):quoteParas	14
5.7. 去除文件名或路径中非法字符:removeInvChrInPath	15
5.8. 把\xXX转换为对应的字符:filterEscapeSequence	16
5.9. 从文件的URL地址中提取文件名:extractFilenameFromUrl	16
6. crifanLib.cs之Array	18
6.1. 从给定字符串中, 从指定位置, 提取指定长度的子字符串:getSubStrArr	18
7. crifanLib.cs之Cookie	19
7.1. 从Url中提取主机Host:extractHost	19
7.2. 从Url中提取域Domain:extractDomain	19
7.3. 从Url中提取域Domain的URL:getDomainUrl	20
7.4. 将Cookie的某一项的值, 添加到Cookie中:addFieldToCookie	20
7.5. 判断字符串是否是有效的cookie的某一项:isValidCookieField	22
7.6. 校验Cookie的名字是否有效/合法:isValidCookieName	23
7.7. 解析Cookie的名字和值:parseCookieNameValue	24
7.8. 解析Cookie的项和域值:parseCookieField	25

7.9. 解析 ( SetCookie的 ) 字符串为单个Cookie值:parseSingleCookie .....	27
7.10. 解析 ( Http访问所返回的 ) Set-Cookie的字符串为Cookie数组:parseSetCookie .....	28
7.11. 解析Javascript中的setCookie为Cookie变量:parseJsSetCookie .....	32
7.12. 判断Cookie是否已经过期/失效/无效:isCookieExpired .....	34
7.13. 将单个Cookie添加到Cookie数组变量中:addCookieToCookies .....	35
7.14. 判断Cookies中是否包含某个Cookie:isContainCookie .....	37
7.15. 更新本地Cookie:updateLocalCookies .....	39
7.16. 从一个CookieCollection获得一个Cookie的值:getCookieVal .....	40
8. crifanLib.cs之Serialize/Deserialize .....	41
8.1. 将一个对象序列化成字符串:serializeObjToStr .....	41
8.2. 将字符串反序列化为对象:deserializeStrToObj .....	42
9. crifanLib.cs之Http .....	44
9.1. 设置代理:setProxy .....	44
9.2. 清除当前cookie:clearCurCookies .....	44
9.3. 获得当前cookie:getCurCookies .....	45
9.4. 设置当前cookie:setCurCookies .....	46
9.5. 获得Url地址的响应:getUrlResponse .....	47
9.5.1. getUrlResponse的参数详解 .....	47
9.5.1.1. getUrlResponse的参数:url .....	47
9.5.1.2. getUrlResponse的参数:headerDict .....	47
9.5.1.3. getUrlResponse的参数:postDict .....	48
9.5.1.4. getUrlResponse的参数:timeout .....	48
9.5.1.5. getUrlResponse的参数:postDataStr .....	48
9.5.1.6. getUrlResponse的参数:readWriteTimeout .....	49
9.5.2. getUrlResponse 的用法详解 .....	49
9.5.2.1. 被getUrlRespHtml调用 .....	49
9.5.2.2. 只传入url而获得对应的url的response .....	50
9.6. 获得Url地址返回的网页内容:getUrlRespHtml .....	50
9.6.1. getUrlRespHtml的参数详解 .....	52
9.6.2. getUrlRespHtml 的功能详解 .....	52
9.6.2.1. 内部已默认指定了IE8的User-Agent .....	52
9.6.2.2. 默认是允许自动跳转的 .....	53
9.6.2.3. 默认已支持解压缩html .....	53
9.6.2.4. 已支持设置 ( 单个 ) 代理 .....	53
9.6.2.5. 支持网络超时设置 .....	54
9.6.2.6. 支持读写超时设置 .....	54
9.6.2.7. 支持自动处理cookie .....	54
9.6.3. getUrlRespHtml 的用法详解 .....	55
9.6.3.1. getUrlRespHtml用法示例：只传入url而获得html .....	55
9.6.3.2. getUrlRespHtml用法示例：传入各种header信息 .....	55
9.6.3.2.1. getUrlRespHtml用法示例：指定Referer .....	55
9.6.3.2.2. getUrlRespHtml用法示例：禁止自动跳转 .....	56
9.6.3.2.3. getUrlRespHtml用法示例：手动设置Accept .....	57
9.6.3.2.4. getUrlRespHtml用法示例：不保持连接 .....	57
9.6.3.2.5. getUrlRespHtml用法示例：设置Accept-Language .....	58
9.6.3.2.6. getUrlRespHtml用法示例：添加特定的User-Agent的header .....	58
9.6.3.2.7. getUrlRespHtml用法示例：设置ContentType .....	60
9.6.3.2.8. getUrlRespHtml用法示例：设置其他的特定的header .....	60
9.6.3.3. getUrlRespHtml用法示例：设置网页字符编码charset .....	61
9.6.3.4. getUrlRespHtml用法示例：设置网络超时timeout时间 .....	61
9.6.3.5. getUrlRespHtml用法示例：设置Stream的读写超时readWriteTimeout .....	61
9.6.3.6. getUrlRespHtml用法示例：POST操作 .....	62
9.6.3.6.1. postDict示例：getDomainPageRank .....	62
9.6.3.6.2. postDict示例：downloadSongtasteMusic .....	62
9.6.3.6.3. postDataStr示例：百度API上传文件 .....	63
9.6.3.6.4. postDataStr示例：网易的心情随笔 .....	64

9.7. 多次尝试版本的getUrlRespHtml:getUrlRespHtml_multiTry .....	64
9.7.1. getUrlRespHtml_multiTry 的参数详解 .....	65
9.8. 获得Url地址所返回的二进制数据流:getUrlRespStreamBytes .....	66
9.9. (谷歌)翻译一段话:translateString .....	67
9.10. 将中文翻译为英文:transzhcntoen .....	69
9.11. 查找获得域名的Page Rank:getDomainPageRank .....	69
9.12. 查找获得域名的Alexa Rank:getDomainAlexaRank .....	72
10. crifanLib.cs之File/Folder .....	76
10.1. 获得当前保存路径:getSaveFolder .....	76
10.2. 二进制(字节)数据存为文件:saveBytesToFile .....	76
10.3. (从网络上)下载文件(到本地):downloadFile .....	78
10.4. 调用资源管理器打开文件夹并选中文件:openFolderAndSelectFile .....	79
10.5. (调用系统默认程序直接)打开文件:openFileDirectly .....	80
11. crifanLib.cs之Screen .....	81
11.1. 获得当前任务栏的尺寸大小:getCurTaskbarSize .....	81
11.2. 获得当前任务栏的坐标位置:getCurTaskbarLocation .....	82
11.3. 获得当前屏幕的角落的坐标位置:getCornerLocation .....	83
12. crifanLib.cs之Runtime .....	85
12.1. 获得当前软件的版本:getCurVerStr .....	85
13. crifanLib.cs之Html Parse .....	86
13.1. 将HTML转换为XmlDocument:htmlToXmlDoc .....	86
13.2. 将HTML转换为HtmlAgilityPack的HtmlDocument:htmlToHtmlDoc .....	88
13.3. 去除HtmlNode中的子节点:removeSubHtmlNode .....	89
13.4. 去除HTML的标签tag:htmlRemoveTag .....	90
14. crifanLib.cs之集成DLL到exe中 .....	92
14.1. 集成DLL到exe中 .....	92
15. crifanLib.cs之DataGridView .....	94
15.1. 清楚DataGridView的内容:dgvClearContent .....	94
15.2. 让DataGridView显示行号:dgvDrawHeaderNum .....	94
15.3. 释放对象(变量):releaseObject .....	94
15.4. 导出DataGridView内容到Excel文件:dgvExportToExcel .....	95
15.5. 导出DataGridView内容到CSV文件:dgvExportToCsv .....	98
16. crifanLib.cs之JSON .....	100
16.1. JSON字符串转换为字典变量:jsonToDict .....	100
参考书目 .....	102

---

## 范例清单

1.1. findRootTreeNode的使用范例	1
1.2. unHighlightNode的使用范例	2
1.3. highlightNode的使用范例	3
2.1. ounceToKiloGram的使用范例	4
2.2. kiloGramToOunce 的使用范例	5
2.3. poundToKiloGram 的使用范例	5
2.4. kiloGramToPound 的使用范例	5
2.5. inchToCm 的使用范例	6
2.6. kiloGramToPound 的使用范例	6
3.1. mathRandom 的使用范例	7
4.1. elapsedTimeSpanInit,getElapsedTimeSpan 的使用范例	8
4.2. getCurTimeInMillisec 的使用范例	9
4.3. milliSecToDateTime 的使用范例	9
4.4. parseJsNewDate 的使用范例	10
5.1. formatstring 的使用范例	11
5.2. emptyStringArray 的使用范例	12
5.3. encodeExclamationMark 的使用范例	12
5.4. decodeExclamationMark 的使用范例	13
5.5. extractSingleStr 的使用范例	13
5.6. quoteParas 的使用范例	15
5.7. removeInvChrInPath 的使用范例	15
5.8. filterEscapeSequence 的使用范例	16
5.9. extractFilenameFromUrl 的使用范例	17
6.1. getSubStrArr 的使用范例	18
7.1. extractHost 的使用范例	19
7.2. extractDomain 的使用范例	20
7.3. getDomainUrl 的使用范例	20
7.4. addFieldToCookie 的使用范例	22
7.5. isValidCookieField 的使用范例	22
7.6. isValidCookieName 的使用范例	24
7.7. parseCookieNameValue 的使用范例	25
7.8. parseCookieField 的使用范例	26
7.9. parseSingleCookie 的使用范例	28
7.10. parseSetCookie 的使用范例	31
7.11. parseJsSetCookie 的使用范例	34
7.12. isCookieExpired 的使用范例	35
7.13. addCookieToCookies 的使用范例	37
7.14. isContainCookie 的使用范例	38
7.15. updateLocalCookies 的使用范例	40
7.16. getCookieVal 的使用范例	40
8.1. serializeObjToStr 的使用范例	41
8.2. deserializeStrToObj 的使用范例	42
9.1. setProxy 的使用范例	44
9.2. clearCurCookies 的使用范例	45
9.3. getCurCookies 的使用范例	45
9.4. setCurCookies 的使用范例	46
9.5. getUrlResponse 的使用范例：被getUrlRespHtml调用	49
9.6. getUrlResponse 的使用范例：只传入url	50
9.7. getUrlRespHtml用法示例：只传入url而获得html	55
9.8. getUrlRespHtml_multiTry 的使用范例	66
9.9. getUrlRespStreamBytes 的使用范例	67
9.10. translateString 的使用范例	69
9.11. transzhcntoen 的使用范例	69
9.12. getDomainPageRank 的使用范例	71
9.13. getDomainAlexaRank 的使用范例	75

10.1. getSaveFolder 的使用范例 .....	76
10.2. saveBytesToFile 的使用范例 .....	77
10.3. downloadFile 的使用范例 .....	79
10.4. openFolderAndSelectFile 的使用范例 .....	80
10.5. openFileDirectly 的使用范例 .....	80
11.1. getCurTaskbarSize 的使用范例 .....	82
11.2. getCurTaskbarLocation 的使用范例 .....	83
11.3. getCornerLocation 的使用范例 .....	84
12.1. getCurVerStr 的使用范例 .....	85
13.1. htmlToXmlDoc 的使用范例 .....	86
13.2. htmlToHtmlDoc 的使用范例 .....	89
13.3. removeSubHtmlNode 的使用范例 .....	90
13.4. htmlRemoveTag 的使用范例 .....	91
14.1. 集成DLL到exe中 的使用范例 .....	92
15.1. dgvClearContent 的使用范例 .....	94
15.2. dgvDrawHeaderNum 的使用范例 .....	94
15.3. releaseObject 的使用范例 .....	95
15.4. dgvExportToExcel 的使用范例 .....	97
15.5. dgvExportToCsv 的使用范例 .....	99
16.1. jsonToDict 的使用范例 .....	100

---

# 前言

## 1. 本文目的

本文目的在于，将自己的C#库crifanLib.cs中的函数都详细解释一遍  
以方便，看了我的库函数，知道如何使用。

## 2. crifanLib.cs的由来

之前在折腾[WLW \(Windows Live Writer\) Plugin-InsertSkydriveFiles](#)<sup>1</sup>的过程中，先后遇到很多问题，然后基本上也都自己解决了。对应的也写了相应的代码和函数。

后来又折腾了很多其他C#方面的东西，比如：

[downloadSonstasteMusic\(下载Songtaste歌曲\)](#)<sup>2</sup>

前前后后，就把其中比较常用或通用的功能，整理提取出来，放到一个单独的文件中，即crifanLib.cs  
此文就是专门针对每个函数，进行详细的解释其用法和给出示例。

## 3. 最新最全的crifanLib.cs完整代码下载

该文件，之前以帖子的方式发布到这里的：[crifan的C#函数库：crifanLib.cs](#)<sup>3</sup>

后来，就放到Google Code上去了，即：

所有的，完整的crifanLib.cs的内容，都是：

- 会不定期更新
- 最新版本始终都放在[google code](#)中的[crifanLib](#)<sup>4</sup>中的[crifanLib.cs](#)<sup>5</sup>了，需要的，自己去下载即可。

其中，当前，截止到2013-08-20，crifanLib.cs的最新版本是：

## 4. crifanLib.cs所包含的引用 ( using )

如果你在使用这些函数的遇到说某某函数，类等找不到，那很可能是没有包含对应的此处的引用。  
那么则请自行参考crifanLib.cs中的using部分，添加对应的引用。

### 4.1. crifanLib.cs中的宏定义

经过后来的版本升级，此时的crifanLib.cs中，已经包含了很多宏定义。

这些宏定义，主要用于，打开，关闭，某些库函数的，以便实现：

当你不想要使用某些函数，以及其会依赖到相关的库，的时候，则可以直接注释掉对应的宏，以实现此目的。

举例，比如，你此处，不想用.NET是3.5或更高的版本，也不想要使用JSON相关的函数，则可以在crifanLib.cs中，把JSON的宏注释掉，即：

---

<sup>1</sup> [http://www.crifan.com/crifan\\_released\\_all/website/dotnet/wlw\\_plugin\\_insertskydrivefiles/](http://www.crifan.com/crifan_released_all/website/dotnet/wlw_plugin_insertskydrivefiles/)

<sup>2</sup> [http://www.crifan.com/crifan\\_released\\_all/website/dotnet/downloadsongtastemusic/](http://www.crifan.com/crifan_released_all/website/dotnet/downloadsongtastemusic/)

<sup>3</sup> [http://www.crifan.com/crifan\\_csharp\\_lib\\_crifanlib\\_cs/](http://www.crifan.com/crifan_csharp_lib_crifanlib_cs/)

<sup>4</sup> <http://code.google.com/p/crifanlib/>

<sup>5</sup> <http://code.google.com/p/crifanlib/source/browse/trunk/csharp/crifanLib.cs>



```
//#define USE_JSON
```

如此，就不会使用到JSON相关的函数了：此刻的效果，主要是：

- 相关的函数jsonToDict等被注释掉
- 不需要用到 ( json所依赖的 ) .NET 3.5+才有的库：System.Web.Script.Serialization了

```
#if USE_JSON  
using System.Web.Script.Serialization; // json lib, need: .NET 3.5+  
#endif
```

## 4.2. crifanLib.cs中引用的所有的库

此处，就把crifanLib.cs目前所有依赖的库，即所有的using，都贴出来，供需要的人，自己添加自己所需要的：

```
//comment out following macros if not use them  
#define USE_GETURLRESPONSE_BW //for getUrIResponse use backgroundworker version  
//#define USE_HTML_PARSER_SGML //need SgmlReaderDll.dll  
//#define USE_HTML_PARSER_HTMLAGILITYPACK //need HtmlAgilityPack.dll  
//#define USE_DATAGRIDVIEW  
//#define USE_JSON
```

```
using System;  
using System.Collections.Generic;  
using System.Text.RegularExpressions;  
using System.Web; // for server  
using System.Net; // for client  
using System.IO;  
using System.Runtime.Serialization.Formatters.Binary;  
using System.Text;  
using System.Drawing;  
using System.Windows.Forms;  
using System.Reflection;  
using System.Diagnostics;  
using System.ComponentModel;  
using System.Globalization;
```

```
#if USE_JSON  
using System.Web.Script.Serialization; // json lib, need: .NET 3.5+  
#endif
```

```
#if USE_HTML_PARSER_SGML  
using Sgml;  
using System.Xml;  
#endif
```

```
#if USE_HTML_PARSER_HTMLAGILITYPACK  
using HtmlAgilityPack;  
#endif
```

```
#if USE_DATAGRIDVIEW
```

```
using Excel = Microsoft.Office.Interop.Excel;
using Microsoft.Office.Interop.Excel;
#endif
```

## 4.3. crifanLib.cs中各个宏的解释

如上所述，crifanLib.cs中包含了一些宏，用于控制一些相关的功能，是否使用。

此处，就对于这些宏，进行详细的解释：

### 4.3.1. USE\_GETURLRESPONSE\_BW

默认关闭此宏。

其背景是：

原先的getUrlResponse，是用于获得URL的响应，属于耗时操作，其在C#中使用时，一般都是出于默认的UI进程中。

导致结果是：当调用到getUrlResponse（以及相关的getUrlRespHtml等）函数时，UI失去响应，导致用户体验很不好。

所以后来又实现了一个BackgroundWorker版本的getUrlResponse

使得，当调用getUrlResponse，UI也可以得到响应了。

所以，如果你想要用BackgroundWorker版本的getUrlResponse，就可以打开此宏：

```
#define USE_GETURLRESPONSE_BW //for getUrlResponse use backgroundworker version
```

如果没此需求，就关闭此宏：

```
//#define USE_GETURLRESPONSE_BW //for getUrlResponse use backgroundworker version
```

### 4.3.2. USE\_HTML\_PARSER\_SGML和 USE\_HTML\_PARSER\_HTMLAGILITYPACK

之前解析HTML，最初用的就是这个sgml库：SgmlReaderDll.dll

但是明显不是很好用。

后来发现了另外一个库：HtmlAgilityPack.dll，后，发现比较好用，就更多的时候，都用HtmlAgilityPack.dll这个库了。

所以，推荐做法：

当涉及到HTML解析的时候，推荐用HtmlAgilityPack，不太推荐用Sgml

所以，一般设置为：

```
//#define USE_HTML_PARSER_SGML //need SgmlReaderDll.dll
#define USE_HTML_PARSER_HTMLAGILITYPACK //need HtmlAgilityPack.dll
```

即可。

当然，如果你两个库都使用，也是可以的。



### 使用sgml或HtmlAgilityPack时要有对应的dll库

此处很明显，当使用对应的库时，则必须有对应的dll库文件，即

- SgmlReaderDll.dll  
详见：[【记录】C#中的HTML解析](#)<sup>6</sup>
- HtmlAgilityPack.dll  
详见：[【记录】折腾C#中的HTML解析库：HtmlAglityPack](#)<sup>7</sup>

### 4.3.3. USE\_DATAGRIDVIEW

DataGridView是表格控件。

之前的很多折腾：

- [【整理】如何使用C#中的DataGridView控件](#)<sup>8</sup>
- [【已解决】C#中DataGridView中的数据导出为CSV](#)<sup>9</sup>
- [【已解决】C#中，清除DataGridView中已有的数据](#)<sup>10</sup>
- [【已解决】给C#的DataGridView中的DataGridViewButtonCell添加事件](#)<sup>11</sup>
- [【已解决】C#的DataGridView中，如何选中新添加的行](#)<sup>12</sup>
- [【已解决】C#的DataGridView中的单元格内添加按钮（整列都是按钮）](#)<sup>13</sup>
- [【已解决】C#的DataGridView中自动在行首添加行号](#)<sup>14</sup>
- [【已解决】将C#中的DataGridView中的数据，导出为Excel](#)<sup>15</sup>

整理出了这些函数：

- dgvClearContent
- dgvDrawHeaderNum
- releaseObject
- dgvExportToExcel
- dgvExportToCsv

所以，当你需要的时候，可以打开此宏：

```
#define USE_DATAGRIDVIEW
```

<sup>6</sup> [http://www.crifan.com/record\\_csharp\\_html\\_parser/](http://www.crifan.com/record_csharp_html_parser/)

<sup>7</sup> [http://www.crifan.com/csharp\\_play\\_with\\_html\\_parse\\_htmlagilitypack/](http://www.crifan.com/csharp_play_with_html_parse_htmlagilitypack/)

<sup>8</sup> [http://www.crifan.com/summary\\_how\\_to\\_use\\_csharp\\_datagridview/](http://www.crifan.com/summary_how_to_use_csharp_datagridview/)

<sup>9</sup> [http://www.crifan.com/csharp\\_export\\_datagridview\\_to\\_csv/](http://www.crifan.com/csharp_export_datagridview_to_csv/)

<sup>10</sup> [http://www.crifan.com/csharp\\_clear\\_all\\_datagridview\\_data/](http://www.crifan.com/csharp_clear_all_datagridview_data/)

<sup>11</sup> [http://www.crifan.com/csharp\\_datagridview\\_add\\_event\\_for\\_datagridviewbuttoncell/](http://www.crifan.com/csharp_datagridview_add_event_for_datagridviewbuttoncell/)

<sup>12</sup> [http://www.crifan.com/csharp\\_datagridview\\_select\\_newly\\_added\\_row/](http://www.crifan.com/csharp_datagridview_select_newly_added_row/)

<sup>13</sup> [http://www.crifan.com/csharp\\_datagridview\\_add\\_whole\\_button\\_column\\_in\\_cell/](http://www.crifan.com/csharp_datagridview_add_whole_button_column_in_cell/)

<sup>14</sup> [http://www.crifan.com/csharp\\_datagridview\\_auto\\_add\\_line\\_number\\_on\\_row\\_head/](http://www.crifan.com/csharp_datagridview_auto_add_line_number_on_row_head/)

<sup>15</sup> [http://www.crifan.com/csharp\\_datagridview\\_save\\_export\\_data\\_to\\_csv\\_or\\_excel/](http://www.crifan.com/csharp_datagridview_save_export_data_to_csv_or_excel/)

去使用相关函数。

#### 4.3.4. USE\_JSON

可以去开启JSON的宏：

```
#define USE_JSON
```

以去使用对应的函数：

- jsonToDict



#### json需要.NET 3.5+ 版本

json依赖的库是：System.Web.Script.Serialization，是需要.NET 3.5或更高的版本才可以的。

换句话说，如果你当前C#项目是2.0的，那么需要转为3.5或更高版本的，才可以用此JSON函数。

## 5. crifanLib.cs中的全局变量，初始化代码，私有函数

此处，顺便也把对应的，全局变量，初始化代码，私有函数等等，贴出来，供参考：

```
public struct pairItem
{
    public string key;
    public string value;
};

private Dictionary<string, DateTime> calcTimeList;

const char replacedChar = '_';

string[] cookieFieldArr = { "expires", "domain", "secure", "path", "httponly", "version" };

//IE7
const string constUserAgent_IE7_x64 = "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 6.1; WOW64; Trident/5.0; SLCC2; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET CLR 3.0.30729; Media Center PC 6.0; InfoPath.3; .NET4.0C; .NET4.0E)";
//IE8
const string constUserAgent_IE8_x64 = "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.1; WOW64; Trident/4.0; SLCC2; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET CLR 3.0.30729; Media Center PC 6.0; InfoPath.3; .NET4.0C; .NET4.0E)";
//IE9
const string constUserAgent_IE9_x64 = "Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; WOW64; Trident/5.0)"; // x64
const string constUserAgent_IE9_x86 = "Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; Trident/5.0)"; // x86
//Chrome
const string constUserAgent_Chrome = "Mozilla/5.0 (Windows; U; Windows NT 6.1; en-US) AppleWebKit/533.4 (KHTML, like Gecko) Chrome/5.0.375.99 Safari/533.4";
//Mozilla Firefox
```

```
const string constUserAgent_Firefox = "Mozilla/5.0 (Windows; U; Windows NT 6.1;
rv:1.9.2.6) Gecko/20100625 Firefox/3.6.6";
private string gUserAgent;

private WebProxy gProxy = null;

//default values:
//getUrlResponse
private const Dictionary<string, string> defHeaderDict = null;
private const Dictionary<string, string> defPostDict = null;
private const int defTimeout = 30 * 1000;
private const string defPostDataStr = null;
private const int defReadWriteTimeout = 30 * 1000;
//getUrlRespHtml
private const string defCharset = null;
//getUrlRespHtml_multiTry
private const int defMaxTryNum = 5;
private const int defRetryFailSleepTime = 100; //sleep time in ms when retry fail for
getUrlRespHtml

List<string> cookieFieldList = new List<string>();

CookieCollection curCookies = null;

//private long totalLength = 0;
//private long currentLength = 0;
#if USE_GETURLRESPONSE_BW
//indicate background worker complete or not
bool bNotCompleted_resp = true;
//store response of http request
private HttpWebResponse gCurResp = null;
#endif

private BackgroundWorker gBgwDownload;
//indicate download complete or not
bool bNotCompleted_download = true;
//store current read out data len
private int gRealReadoutLen = 0;
Action<int> gFuncUpdateProgress = null;

public crifanLib()
{
    //!!! for load embedded dll: (1) register resolve handler
    AppDomain.CurrentDomain.AssemblyResolve += new
ResolveEventHandler(CurrentDomain_AssemblyResolve);

    //http related
    gUserAgent = constUserAgent_IE8_x64;
    //set max enough to avoid http request is used out -> avoid dead while get response
    System.Net.ServicePointManager.DefaultConnectionLimit = 200;

    curCookies = new CookieCollection();
    // init const cookie keys
    foreach (string key in cookieFieldArr)
    {
        cookieFieldList.Add(key);
    }
}
```

```
//init for calc time
calcTimeList = new Dictionary<string, DateTime>();
#if USE_GETURLRESPONSE_BW
    gBgwDownload = new BackgroundWorker();
#endif

    //debug
    //gProxy = new WebProxy("127.0.0.1", 8087);
}

/*-----Private Functions-----*/

//!!! for load embedded dll: (2) implement this handler
System.Reflection.Assembly CurrentDomain_AssemblyResolve(object sender,
ResolveEventArgs args)
{
    string dllName = args.Name.Contains(",") ? args.Name.Substring(0,
args.Name.IndexOf(',')) : args.Name.Replace(".dll", "");

    dllName = dllName.Replace(".", "_");

    if (dllName.EndsWith("_resources")) return null;

    System.Resources.ResourceManager rm = new
System.Resources.ResourceManager(GetType().Namespace + ".Properties.Resources",
System.Reflection.Assembly.GetExecutingAssembly());

    byte[] bytes = (byte[])rm.GetObject(dllName);

    return System.Reflection.Assembly.Load(bytes);
}

// replace the replacedChar back to original ','
private string _recoverExpireField(Match foundPprocessedExpire)
{
    string recovedStr = "";
    recovedStr = foundPprocessedExpire.Value.Replace(replacedChar, ',');
    return recovedStr;
}

//replace ',' with replacedChar
private string _processExpireField(Match foundExpire)
{
    string replacedComma = "";
    replacedComma = foundExpire.Value.ToString().Replace(',', replacedChar);
    return replacedComma;
}

//replace "0A" (in \x0A) into '\n'
private string _replaceEscapeSequenceToChar(Match foundEscapeSequence)
{
    char[] hexValues = new char[2];
    //string hexChars = foundEscapeSequence.Value.ToString();
    string matchedEscape = foundEscapeSequence.ToString();
    hexValues[0] = matchedEscape[2];
    hexValues[1] = matchedEscape[3];
    string hexValueString = new string(hexValues);
```

```
int convertedInt = int.Parse(hexValueString, NumberStyles.HexNumber,
NumberFormatInfo.InvariantInfo);
char hexChar = Convert.ToChar(convertedInt);
string hexStr = hexChar.ToString();
return hexStr;
}

//check whether need add/retain this cookie
// not add for:
// ck is null or ck name is null
// domain is null and curDomain is not set
// expired and retainExpiredCookie==false
private bool needAddThisCookie(Cookie ck, string curDomain)
{
    bool needAdd = false;

    if ((ck == null) || (ck.Name == ""))
    {
        needAdd = false;
    }
    else
    {
        if (ck.Domain != "")
        {
            needAdd = true;
        }
        else// ck.Domain == ""
        {
            if (curDomain != "")
            {
                ck.Domain = curDomain;
                needAdd = true;
            }
            else // curDomain == ""
            {
                // not set current domain, omit this
                // should not add empty domain cookie, for this will lead execute
                CookieContainer.Add() fail !!!
                needAdd = false;
            }
        }
    }
}

return needAdd;
}

//quote the input dict values
//note: the return result for first para no '&'
private string _quoteParas(Dictionary<string, string> paras, bool spaceToPercent20 = true)
{
    string quotedParas = "";
    bool isFirst = true;
    string val = "";
    foreach (string para in paras.Keys)
    {
        if (paras.TryGetValue(para, out val))
        {
            string encodedVal = "";
```

```
    if (spaceToPercent20)
    {
        //encodedVal = HttpUtility.UrlPathEncode(val);
        //encodedVal = Uri.EscapeDataString(val);
        //encodedVal = Uri.EscapeUriString(val);
        encodedVal = HttpUtility.UrlEncode(val).Replace("+", "%20");
    }
    else
    {
        encodedVal = HttpUtility.UrlEncode(val); //space to +
    }

    if (isFirst)
    {
        isFirst = false;
        quotedParas += para + "=" + encodedVal;
    }
    else
    {
        quotedParas += "&" + para + "=" + encodedVal;
    }
}
else
{
    break;
}
}

return quotedParas;
}

/* get url's response
**/
private HttpResponseMessage _getUrlResponse(string url,
    Dictionary<string, string> headerDict = defHeaderDict,
    Dictionary<string, string> postDict = defPostDict,
    int timeout = defTimeout,
    string postDataStr = defPostDataStr,
    int readWriteTimeout = defReadWriteTimeout)
{
    //CookieCollection parsedCookies;

    HttpResponseMessage resp = null;

    HttpRequest req = (HttpRequest)WebRequest.Create(url);

    req.AllowAutoRedirect = true;
    req.Accept = "*/*";

    //req.ContentType = "text/plain";

    //const string gAcceptLanguage = "en-US"; // zh-CN/en-US
    //req.Headers["Accept-Language"] = gAcceptLanguage;

    req.KeepAlive = true;

    req.UserAgent = gUserAgent;
```



```
req.Headers["Accept-Encoding"] = "gzip, deflate";
//req.AutomaticDecompression = DecompressionMethods.GZip;
req.AutomaticDecompression = DecompressionMethods.GZip |
DecompressionMethods.Deflate;

req.Proxy = gProxy;

if (timeout > 0)
{
    req.Timeout = timeout;
}

if (readWriteTimeout > 0)
{
    //default ReadWriteTimeout is 300000=300 seconds = 5 minutes !!!
    //too long, so here change to 300000 = 30 seconds
    //for support TimeOut for later StreamReader's ReadToEnd
    req.ReadWriteTimeout = readWriteTimeout;
}

if (curCookies != null)
{
    req.CookieContainer = new CookieContainer();
    req.CookieContainer.PerDomainCapacity = 40; // following will exceed max default 20
cookie per domain
    req.CookieContainer.Add(curCookies);
}

if ((headerDict != null) && (headerDict.Count > 0))
{
    foreach (string header in headerDict.Keys)
    {
        string headerValue = "";
        if (headerDict.TryGetValue(header, out headerValue))
        {
            string lowercaseHeader = header.ToLower();
            // following are allow the caller overwrite the default header setting
            if (lowercaseHeader == "referer")
            {
                req.Referer = headerValue;
            }
            else if (
                (lowercaseHeader == "allow-autoredirect") ||
                (lowercaseHeader == "allowautoredirect") ||
                (lowercaseHeader == "allow autoredirect")
            )
            {
                bool isAllow = false;
                if (bool.TryParse(headerValue, out isAllow))
                {
                    req.AllowAutoRedirect = isAllow;
                }
            }
            else if (lowercaseHeader == "accept")
            {
                req.Accept = headerValue;
            }
            else if (
```

```
(lowercaseHeader == "keep-alive") ||
(lowercaseHeader == "keepalive") ||
(lowercaseHeader == "keep alive")
)
{
    bool isKeepAlive = false;
    if (bool.TryParse(headerValue, out isKeepAlive))
    {
        req.KeepAlive = isKeepAlive;
    }
}
else if (
    (lowercaseHeader == "accept-language") ||
    (lowercaseHeader == "acceptlanguage") ||
    (lowercaseHeader == "accept language")
)

{
    req.Headers["Accept-Language"] = headerValue;
}
else if (
    (lowercaseHeader == "user-agent") ||
    (lowercaseHeader == "useragent") ||
    (lowercaseHeader == "user agent")
)
{
    req.UserAgent = headerValue;
}
else if (
    (lowercaseHeader == "content-type") ||
    (lowercaseHeader == "contenttype") ||
    (lowercaseHeader == "content type")
)
{
    req.ContentType = headerValue;
}
else
{
    req.Headers[header] = headerValue;
}
}
else
{
    break;
}
}
}

if (((postDict != null) && (postDict.Count > 0)) || (!string.IsNullOrEmpty(postDataStr)))
{
    req.Method = "POST";
    if (req.ContentType == null)
    {
        req.ContentType = "application/x-www-form-urlencoded";
    }

    if ((postDict != null) && (postDict.Count > 0))
    {
```

```
    postDataStr = _quoteParas(postDict);
}

//byte[] postBytes = Encoding.GetEncoding("utf-8").GetBytes(postData);
byte[] postBytes = Encoding.UTF8.GetBytes(postDataStr);
req.ContentLength = postBytes.Length;

try
{
    Stream postDataStream = req.GetRequestStream();
    postDataStream.Write(postBytes, 0, postBytes.Length);
    postDataStream.Close();
}
catch (WebException webEx)
{
    //for prev has set ReadWriteTimeout
    //so here also may timeout
    if (webEx.Status == WebExceptionStatus.Timeout)
    {
        req = null;
    }
}
else
{
    req.Method = "GET";
}

if (req != null)
{
    //may timeout, has fixed in:
    //http://www.crifan.com/
fixed_problem_sometime_httpwebrequest_getresponse_timeout/
    try
    {
        resp = (HttpWebResponse)req.GetResponse();
        updateLocalCookies(resp.Cookies, ref curCookies);
    }
    catch (WebException webEx)
    {
        if (webEx.Status == WebExceptionStatus.Timeout)
        {
            resp = null;
        }
    }
}

return resp;
}

#if USE_GETURLRESPONSE_BW
private void getUrlResponse_bw(string url,
    Dictionary<string, string> headerDict = defHeaderDict,
    Dictionary<string, string> postDict = defPostDict,
    int timeout = defTimeout,
    string postDataStr = defPostDataStr,
    int readWriteTimeout = defReadWriteTimeout)
{
```

```
// Create a background thread
BackgroundWorker bgwGetUrlResp = new BackgroundWorker();
bgwGetUrlResp.DoWork += new DoWorkEventHandler(bgwGetUrlResp_DoWork);
bgwGetUrlResp.RunWorkerCompleted += new
RunWorkerCompletedEventHandler( bgwGetUrlResp_RunWorkerCompleted );

//init
bNotCompleted_resp = true;

// run in another thread
object paraObj = new object[] { url, headerDict, postDict, timeout, postDataStr,
readWriteTimeout };
bgwGetUrlResp.RunWorkerAsync(paraObj);
}

private void bgwGetUrlResp_DoWork(object sender, DoWorkEventArgs e)
{
    object[] paraObj = (object[])e.Argument;
    string url = (string)paraObj[0];
    Dictionary<string, string> headerDict = (Dictionary<string, string>)paraObj[1];
    Dictionary<string, string> postDict = (Dictionary<string, string>)paraObj[2];
    int timeout = (int)paraObj[3];
    string postDataStr = (string)paraObj[4];
    int readWriteTimeout = (int)paraObj[5];

    e.Result = _getUrlResponse(url, headerDict, postDict, timeout, postDataStr,
readWriteTimeout);
}

//void m_bgWorker_ProgressChanged(object sender, ProgressChangedEventArgs e)
//{
//    bRespNotCompleted = true;
//}

private void bgwGetUrlResp_RunWorkerCompleted(object sender,
RunWorkerCompletedEventArgs e)
{
    // The background process is complete. We need to inspect
    // our response to see if an error occurred, a cancel was
    // requested or if we completed successfully.

    // Check to see if an error occurred in the
    // background process.
    if (e.Error != null)
    {
        //MessageBox.Show(e.Error.Message);
        return;
    }

    // Check to see if the background process was cancelled.
    if (e.Cancelled)
    {
        //MessageBox.Show("Cancelled ...");
    }
    else
    {
        bNotCompleted_resp = false;
    }
}
```

```
// Everything completed normally.
// process the response using e.Result
//MessageBox.Show("Completed...");
gCurResp = (HttpWebResponse)e.Result;
}
}
#endif

private void getUrlRespStreamBytes_bw(ref Byte[] respBytesBuf,
    string url,
    Dictionary<string, string> headerDict,
    Dictionary<string, string> postDict,
    int timeout,
    Action<int> funcUpdateProgress)
{
    // Create a background thread
    gBgwDownload = new BackgroundWorker();
    gBgwDownload.DoWork += bgwDownload_DoWork;
    gBgwDownload.RunWorkerCompleted += bgwDownload_RunWorkerCompleted;
    gBgwDownload.WorkerReportsProgress = true;
    gBgwDownload.ProgressChanged += bgwDownload_ProgressChanged;

    //init
    bNotCompleted_download = true;
    gFuncUpdateProgress = funcUpdateProgress;

    // run in another thread
    object paraObj = new object[] {respBytesBuf, url, headerDict, postDict, timeout};
    gBgwDownload.RunWorkerAsync(paraObj);
}

private void bgwDownload_ProgressChanged(object sender, ProgressChangedEventArgs e)
{
    if (gFuncUpdateProgress != null)
    {
        // This function fires on the UI thread so it's safe to edit
        // the UI control directly, no funny business with Control.Invoke.
        // Update the progressBar with the integer supplied to us from the
        // ReportProgress() function. Note, e.UserState is a "tag" property
        // that can be used to send other information from the
        // BackgroundThread to the UI thread.

        gFuncUpdateProgress(e.ProgressPercentage);
    }
}

private void bgwDownload_DoWork(object sender, DoWorkEventArgs e)
{
    // // The sender is the BackgroundWorker object we need it to
    // // report progress and check for cancellation.
    // BackgroundWorker gBgwDownload = sender as BackgroundWorker;

    object[] paraObj = (object[])e.Argument;
    Byte[] respBytesBuf = (Byte[])paraObj[0];
    string url = (string)paraObj[1];
    Dictionary<string, string> headerDict = (Dictionary<string, string>)paraObj[2];
    Dictionary<string, string> postDict = (Dictionary<string, string>)paraObj[3];
}
```

```
int timeout = (int)paraObj[4];

//e.Result = _getUrlRespStreamBytes(ref respBytesBuf, url, headerDict, postDict, timeout);

int curReadoutLen;
int realReadoutLen = 0;
int curBufPos = 0;

long totalLength = 0;
long currentLength = 0;

try
{
    //HttpWebResponse resp = getUrlResponse(url, headerDict, postDict, timeout);
    HttpWebResponse resp = getUrlResponse(url, headerDict, postDict);
    long expectReadoutLen = resp.ContentLength;

    totalLength = expectReadoutLen;
    currentLength = 0;

    Stream binStream = resp.GetResponseStream();
    //int streamDataLen = (int)binStream.Length; // erro: not support seek operation

    do
    {
        //let up layer update its UI, otherwise up layer UI will no response during this func
exec time
        //now has make this function to call by backgroundworker, so not need this to
update UI
        //System.Windows.Forms.Application.DoEvents();

        // here download logic is:
        // once request, return some data
        // request multiple time, until no more data
        curReadoutLen = binStream.Read(respBytesBuf, curBufPos, (int)expectReadoutLen);
        if (curReadoutLen > 0)
        {
            curBufPos += curReadoutLen;

            currentLength = curBufPos;

            expectReadoutLen = expectReadoutLen - curReadoutLen;

            realReadoutLen += curReadoutLen;

            int currentPercent = (int)((currentLength * 100) / totalLength);

            if (currentPercent < 0)
            {
                currentPercent = 0;
            }

            if (currentPercent > 100)
            {
                currentPercent = 100;
            }
        }
    }
}
```

```
        gBgwDownload.ReportProgress(currentPercent);
    }
} while (curReadoutLen > 0);
}
catch (Exception ex)
{
    string errorMessage = ex.Message;
    realReadoutLen = -1;
}

//return realReadoutLen;

e.Result = realReadoutLen;
//gBgwDownload.ReportProgress(100);
}

private void bgwDownload_RunWorkerCompleted(object sender,
RunWorkerCompletedEventArgs e)
{
    // The background process is complete. We need to inspect
    // our response to see if an error occurred, a cancel was
    // requested or if we completed successfully.

    // Check to see if an error occurred in the
    // background process.
    if (e.Error != null)
    {
        //MessageBox.Show(e.Error.Message);
        return;
    }

    // Check to see if the background process was cancelled.
    if (e.Cancelled)
    {
        //MessageBox.Show("Cancelled ...");
    }
    else
    {
        bNotCompleted_download = false;

        // Everything completed normally.
        // process the response using e.Result
        //MessageBox.Show("Completed...");
        gRealReadoutLen = (int)e.Result;
    }
}
}
```

---

# 第 1 章 crifanLib.cs之TreeView/ TreeNode

## 1.1. 查找TreeNode的根节点:findRootTreeNode

```
/*
 * [Function]
 * find root TreeNode of current TreeNode
 * [Input]
 * some TreeNode
 *
 * [Output]
 * root TreeNode of input TreeNode
 *
 * [Note]
 */
public TreeNode findRootTreeNode(TreeNode curTreeNode)
{
    TreeNode rootTreeNode = curTreeNode.Parent;

    if (rootTreeNode == null)
    {
        //root parent is null
        rootTreeNode = curTreeNode;
    }
    else
    {
        //child parent is not null
        while (rootTreeNode.Parent != null)
        {
            rootTreeNode = rootTreeNode.Parent;
        }
    }

    return rootTreeNode;
}
```

### 例 1.1. findRootTreeNode的使用范例

```
//get input TreeNode's BrowseNode's SearchIndex
private string getSearchIndex(TreeNode curTreeNode)
{
    string strSearchIndex = "";

    //find the root node
    TreeNode rootTreeNode = crl.findRootTreeNode(curTreeNode);
```



## 1.2. 取消节点的高亮:unHighlightNode

```
/*
 * [Function]
 * un highlight tree node
 * [Input]
 * some TreeNode
 *
 * [Output]
 * restore color to background color
 *
 * [Note]
 */
public Color unHighlightNode(TreeView trvValue, TreeNode treeNode)
{
    Color oldColor = trvValue.BackColor;
    if (treeNode != null)
    {
        oldColor = treeNode.BackColor;
        treeNode.BackColor = trvValue.BackColor;
        treeNode.ForeColor = Color.Black;
    }

    return oldColor;
}
```

### 例 1.2. unHighlightNode的使用范例

```
else if (e.ClickedItem == tsmiRemoveFromSelection)
{
    if (curSelTreeNodeList.Contains(curSelTreeNode))
    {
        //remove selection
        curSelTreeNodeList.Remove(curSelTreeNode);

        //unhighlight node
        cri.unHighlightNode(trvCategoryTree, curSelTreeNode);
    }
}
```

## 1.3. 高亮TreeNode:highlightNode

```
/*
 * [Function]
 * highlight tree node
 * [Input]
 * some TreeNode
 *
 * [Output]
```

```
* set color to highlighted color
*
* [Note]
*/
public Color highlightNode(TreeView trvValue, TreeNode someNode)
{
    Color oldColor = trvValue.BackColor; //{Name=Window, ARGB=(255, 255, 255, 255)}"
    if (someNode != null)
    {
        oldColor = someNode.BackColor; //{Name=0, ARGB=(0, 0, 0, 0)}"

        // HTML #3399FF -> RGB(51,153,255)
        //{Name=MenuHighlight, ARGB=(255, 51, 153, 255)}"
        someNode.BackColor = SystemColors.MenuHighlight;

        //node.BackColor = nodeHIBackColor;

        //node.ForeColor = Color.FromArgb(255, 255, 255);
        someNode.ForeColor = Color.White;
    }

    return oldColor;
}
```

### 例 1.3. highlightNode的使用范例

```
if (e.ClickedItem == tsmiAddToSelection)
{
    if (!curSelTreeNodeList.Contains(curSelTreeNode))
    {
        // add to selection
        curSelTreeNodeList.Add(curSelTreeNode);

        //highlight node
        crl.highlightNode(trvCategoryTree, curSelTreeNode);
    }
}
```

---

## 第 2 章 crifanLib.cs之Unit Conversion

### 2.1. 盎司转千克:ounceToKiloGram

```
public float ounceToKiloGram(float ounce)
{
    float kiloGram = ounce * 0.028349523125F;

    return kiloGram;
}
```

#### 例 2.1. ounceToKiloGram的使用范例

```
float kiloGram = -1.0F;
string weightNumberStr = "";

//type1:
//http://www.amazon.com/Kindle-Fire-HD/dp/B0083PWAPW/ref=lp_1055398_1_1?
ie=UTF8&qid=1369487181&sr=1-1
//<td style="font-weight: bold;text-align:left; font-size: 12px; border-bottom: 1px solid
#e2e2e2;" align="right">Weight</td><td style="font-size:12px;">13.9 ounces (395 grams)</
td>
//http://www.amazon.com/Kindle-Paperwhite-Touch-light/dp/B007OZNZG0/
ref=lp_1055398_1_2?ie=UTF8&qid=1369487181&sr=1-2
//<td style="font-weight: bold;text-align:left; font-size: 12px; border-bottom: 1px solid
#e2e2e2;" align="right">Weight</td><td style="font-size:12px;">7.5 ounces (213 grams)</
td>
if (!calculatedKiloGram)
{
    if (crl.extractSingleStr(@"Weight</td><td style=""[^<>]+?">([\.\d]+) ounces",
productHtml, out weightNumberStr))
    {
        float onces = float.Parse(weightNumberStr);
        kiloGram = crl.ounceToKiloGram(ounces);
    }
}
```

### 2.2. 千克转盎司:kiloGramToOunce

```
public float kiloGramToOunce(float kiloGram)
{
    float ounce = kiloGram * 35.27396194958F;

    return ounce;
}
```

### 例 2.2. kiloGramToOunce 的使用范例

## 2.3. 英镑转千克:poundToKiloGram

```
public float poundToKiloGram(float pound)
{
    float kiloGram = pound * 0.45359237F;
    return kiloGram;
}
```

### 例 2.3. poundToKiloGram 的使用范例

```
else if (unitType.Equals("pounds"))
{
    float pound = float.Parse(weightNumberStr);
    kiloGram = crl.poundToKiloGram(pound);
}
```

## 2.4. 千克转英镑:kiloGramToPound

```
public float kiloGramToPound(float kiloGram)
{
    float pound = kiloGram * 0.45359237F;
    return pound;
}
```

### 例 2.4. kiloGramToPound 的使用范例

## 2.5. 英尺转厘米:inchToCm

```
public float inchToCm(float inch)
{
    float cm = inch * 2.54F;

    return cm;
}
```

### 例 2.5. inchToCm 的使用范例

```
dimensionInch.length = float.Parse(lengthInchStr);
dimensionInch.width = float.Parse(widthInchStr);
dimensionInch.height = float.Parse(heightInchStr);

dimensionCm.length = crl.inchToCm(dimensionInch.length);
dimensionCm.width = crl.inchToCm(dimensionInch.width);
dimensionCm.height = crl.inchToCm(dimensionInch.height);
```

## 2.6. 厘米转英尺:cmToInch

```
public float cmToInch(float cm)
{
    float inch = cm * 0.39370078740157F;

    return inch;
}
```

### 例 2.6. kiloGramToPound 的使用范例

---

## 第 3 章 crifanLib.cs之Values

### 3.1. 和Javascript中Math.Random()等价的函数:mathRandom

```
//equivalent of Math.Random() in Javascript
//get a 17 bit double value x, 0 < x < 1, eg:0.68637410117610087
public double mathRandom()
{
    Random rdm = new Random();
    double betweenZeroToOne17Bit = rdm.NextDouble();
    return betweenZeroToOne17Bit;
}
```

#### 例 3.1. mathRandom 的使用范例

---

# 第 4 章 crifanLib.cs之Time

此处是和时间 ( Time , DateTime等 ) 有关的函数

## 4.1. 计算 ( 代码执行 ) 时间消耗 ( 的时间段/时长 ) :elapsedTimeSpanInit,getElapsedTimeSpan

使用前, 先做最开始的初始化 :

```
private Dictionary<string, DateTime> calcTimeList;

//init for calc time
calcTimeList = new Dictionary<string, DateTime>();
```

每次使用之前, 使用 :

```
// init for calculate time span
public void elapsedTimeSpanInit(string keyName)
{
    calcTimeList.Add(keyName, DateTime.Now);
}
```

然后就可以获得对应的时间消耗了 :

```
// got calculated time span
public double getElapsedTimeSpan(string keyName)
{
    double milliSec = 0.0;
    if (calcTimeList.ContainsKey(keyName))
    {
        DateTime startTime = calcTimeList[keyName];
        DateTime endTime = DateTime.Now;
        milliSec = (endTime - startTime).TotalMilliseconds;
    }
    return milliSec;
}
```

### 例 4.1. elapsedTimeSpanInit,getElapsedTimeSpan 的使用范例

## 4.2. 获得（从epoch时间纪元以来的）（以毫秒为单位的）当前时间: `getCurTimeInMillisec`

```
//refer: http://bytes.com/topic/c-sharp/answers/713458-c-function-equivalent-javascript-gettime-function
//get current time in milli-second-since-epoch(1970/01/01)
public double getCurTimeInMillisec()
{
    DateTime st = new DateTime(1970, 1, 1);
    TimeSpan t = (DateTime.Now - st);
    return t.TotalMilliseconds; // milli seconds since epoch
}
```

### 例 4.2. `getCurTimeInMillisec` 的使用范例

```
double curMilliSecDouble = crl.getCurTimeInMillisec(); //1343392590725.6758
```

## 4.3. 将毫秒转换为（自1970年1月1日以来的）本地时间: `milliSecToDateTime`

```
// parse the milli second to local DateTime value
public DateTime milliSecToDateTime(double milliSecSinceEpoch)
{
    DateTime st = new DateTime(1970, 1, 1, 0, 0, 0);
    st = st.AddMilliseconds(milliSecSinceEpoch);
    return st;
}
```

### 例 4.3. `milliSecToDateTime` 的使用范例

```
double doubleVal = 0.0;
if (Double.TryParse(dateValue, out doubleVal))
{
    // try whether is double/int64 milliSecSinceEpoch
    parsedDatetime = milliSecToDateTime(doubleVal);
    parseOK = true;
}
```



## 4.4. 将Javascript中的"new Date(xxx)"转换为C#中的DateTime变量:parseJsNewDate

```
//parse xxx in "new Date(xxx)" of javascript to C# DateTime
//input example:
//new Date(1329198041411.84) / new Date(1329440307389.9) / new Date(1329440307483)
public bool parseJsNewDate(string newDateStr, out DateTime parsedDatetime)
{
    bool parseOK = false;
    parsedDatetime = new DateTime();

    if ((newDateStr != "") && (newDateStr.Trim() != ""))
    {
        string dateValue = "";
        if (extractSingleStr("@.*new\sDate\(((.+?)\).*", newDateStr, out dateValue))
        {
            double doubleVal = 0.0;
            if (Double.TryParse(dateValue, out doubleVal))
            {
                // try whether is double/int64 milliSecSinceEpoch
                parsedDatetime = milliSecToDateTime(doubleVal);
                parseOK = true;
            }
            else if (DateTime.TryParse(dateValue, out parsedDatetime))
            {
                // try normal DateTime string
                //refer: http://www.w3schools.com/js/js_obj_date.asp
                //October 13, 1975 11:13:00
                //79,5,24 / 79,5,24,11,33,0
                //1329198041411.3344 / 1329198041411.84 / 1329198041411
                parseOK = true;
            }
        }
    }

    return parseOK;
}
```

### 例 4.4. parseJsNewDate 的使用范例

```
DateTime expireTime;
if (parseJsNewDate(expire, out expireTime))
{
    parsedCk.Expires = expireTime;
}
```

---

# 第 5 章 crifanLib.cs之String

此处是和字符串 ( string等 ) 有关的函数

## 5.1. 格式化字符串中间对齐左右填充:formatstring

```
//input: [4] Valid: B0009IQZFM
//output: ===== [4] Valid: B0009IQZFM
=====
public string formatString(string strToFormat, char cPaddingChar = '*', int iTotWidth = 80)
{
    //auto added space
    strToFormat = " " + strToFormat + " "; //" [4] Valid: B0009IQZFM "

    //1. padding left
    int iPaddingLen = (iTotWidth - strToFormat.Length)/2;
    int iLefTotalLen = iPaddingLen + strToFormat.Length;
    string strLefPadded = strToFormat.PadLeft(iLefTotalLen,
cPaddingChar); //"===== [4] Valid: B0009IQZFM "
    //2. padding right
    string strFormatted = strLefPadded.PadRight(iTotWidth,
cPaddingChar); //"===== [4] Valid: B0009IQZFM
===== "

    return strFormatted;
}
```

### 例 5.1. formatstring 的使用范例

```
string strFullCategoryName = String.Format("FullCategoryName={0}",
curFullCategoryName);
string strFormattedFullCategoryName = crl.formatString(strFullCategoryName, '=');
```

## 5.2. 初始化null的字符串位空字符串"":emptyStringArray

```
//init the string array to empty
public string[] emptyStringArray(string[] strArr)
{
```

```
if (strArr != null)
{
    for (int idx = 0; idx < strArr.Length; idx++)
    {
        strArr[idx] = String.Empty;
        //strArr[idx] = "";
    }
}

return strArr;
}
```

### 例 5.2. emptyStringArray 的使用范例

```
//5 bullet
//public string[] bulletArr; // total 5 (or more, but only record 5)

productInfo.bulletArr = new string[5];
crl.emptyStringArray(productInfo.bulletArr);
```

## 5.3. 将感叹号"!"强制编码为"%21":encodeExclamationMark

```
// encode "!" to "%21"
public string encodeExclamationMark(string inputStr)
{
    return inputStr.Replace("!", "%21");
}
```

### 例 5.3. encodeExclamationMark 的使用范例

```
getItemsUrl += "id=" + encodeExclamationMark(folderId).ToLower();
```

## 5.4. 将"%21"解码为感叹号"!:decodeExclamationMark

```
// encode "%21" to "!"  
public string decodeExclamationMark(string inputStr)  
{  
    return inputStr.Replace("%21", "!");  
}
```

#### 例 5.4. decodeExclamationMark 的使用范例

```
folderId = decodeExclamationMark(folderId);
```

## 5.5. 从字符串中提取单个的子字符串:extractSingleStr

```
//using Regex to extract single string value  
// caller should make sure the string to extract is Groups[1] == include single () !!!  
public bool extractSingleStr(string pattern, string extractFrom, out string extractedStr)  
{  
    bool extractOK = false;  
    Regex rx = new Regex(pattern);  
    Match found = rx.Match(extractFrom);  
    if (found.Success)  
    {  
        extractOK = true;  
        extractedStr = found.Groups[1].ToString();  
    }  
    else  
    {  
        extractOK = false;  
        extractedStr = "";  
    }  
  
    return extractOK;  
}
```

#### 例 5.5. extractSingleStr 的使用范例

```
string resPreloadUrl = "";  
//var srf_uPreload = 'https://skydrive.live.com/handlers/resourcespreload.mvc?  
view=Folders.All&id;=250206&mkt;=EN-US';
```

```
string resPreloadP = @"var\ssrf_uPreload\s=\s'(.+?)";  
extractSingleStr(resPreloadP, html, out resPreloadUrl);
```



## 传入extractSingleStr的正则pattern中必须包含括号，即group

从代码中可见，传入extractSingleStr中的pattern，必须有一个括号，即一个group  
然后查找出来的内容，才能得以提取出来

## 5.6. 组合参数列表（变成 &xxx=yyy）:quoteParas

```
//quote the input dict values  
//note: the return result for first para no '&'  
public string quoteParas(Dictionary<string, string> paras, bool spaceToPercent20 = true)  
{  
    string quotedParas = "";  
    bool isFirst = true;  
    string val = "";  
    foreach (string para in paras.Keys)  
    {  
        if (paras.TryGetValue(para, out val))  
        {  
            string encodedVal = "";  
            if (spaceToPercent20)  
            {  
                //encodedVal = HttpUtility.UrlPathEncode(val);  
                //encodedVal = Uri.EscapeDataString(val);  
                //encodedVal = Uri.EscapeUriString(val);  
                encodedVal = HttpUtility.UrlEncode(val).Replace("+", "%20");  
            }  
            else  
            {  
                encodedVal = HttpUtility.UrlEncode(val); //space to +  
            }  
  
            if (isFirst)  
            {  
                isFirst = false;  
                quotedParas += para + "=" + encodedVal;  
            }  
            else  
            {  
                quotedParas += "&" + para + "=" + encodedVal;  
            }  
        }  
        else  
        {  
            break;  
        }  
    }  
}
```

```
}  
    return quotedParas;  
}
```

### 例 5.6. quoteParas 的使用范例

```
Dictionary<string, string> postDataDict = genPostsrfPostDict(html, login, passwd,  
isKeepLogin);  
postData += quoteParas(postDataDict);
```

## 5.7. 去除文件名或路径中非法字符:removeInvChrInPath

```
//remove invalid char in path and filename  
public string removeInvChrInPath(string origFileOrPathStr)  
{  
    string validFileOrPathStr = origFileOrPathStr;  
  
    //filter out invalid title and artist char  
    //char[] invalidChars = { '\\', '/', ':', '*', '?', '<', '>', '|', '\b' };  
    char[] invalidChars = Path.GetInvalidPathChars();  
    char[] invalidCharsInName = Path.GetInvalidFileNameChars();  
  
    foreach (char chr in invalidChars)  
    {  
        validFileOrPathStr = validFileOrPathStr.Replace(chr.ToString(), "");  
    }  
  
    foreach (char chr in invalidCharsInName)  
    {  
        validFileOrPathStr = validFileOrPathStr.Replace(chr.ToString(), "");  
    }  
  
    return validFileOrPathStr;  
}
```

### 例 5.7. removeInvChrInPath 的使用范例

```
string mid_tit;
```

```
if (crl.extractSingleStr(@"<p\s+?class=""mid_tit"">(?!<mid_tit>.+)<p>", respHtml, out mid_tit))
{
    albumInfo.name = crl.removeInvChrInPath(mid_tit);
}

string h1user;
if (crl.extractSingleStr(@"<h1\s+?class=""h1user"">(?!<h1user>.+)</h1>", respHtml, out h1user))
{
    albumInfo.author = crl.removeInvChrInPath(h1user);
}
```

## 5.8. 把\xXX转换为对应的字符:filterEscapeSequence

```
//convert \xXX into corresponding char
//eg: \x0A -> '\n'
public string filterEscapeSequence(string escapeSequenceStr)
{
    string filteredStr = Regex.Replace(escapeSequenceStr, @"\x\w{2}", new MatchEvaluator(_replaceEscapeSequenceToChar));

    return filteredStr;
}
```

### 例 5.8. filterEscapeSequence 的使用范例

## 5.9. 从文件的URL地址中提取文件名:extractFilenameFromUrl

```
//extract filename from url
//eg:
//http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KC/KC-slate-01-
lg_V401028090_jpg
//KC-slate-01-lg_V401028090_jpg
```

```
//file:///C:/Users/CLi/AppData/Local/Temp/WindowsLiveWriter-1737927945/  
supfilesC19F10/now-the-service-status-is-active_thu%5B1%5D.png  
//now-the-service-status-is-active_thu%5B1%5D.png  
public string extractFilenameFromUrl(string fullUrl)  
{  
    string filename = "";  
    string[] slashList = fullUrl.Split('/');  
    filename = slashList[slashList.Length - 1];  
    return filename;  
}
```

### 例 5.9. extractFilenameFromUrl 的使用范例

```
string imageUrl = imageUrlList[idx];  
gLogger.Info(String.Format("[{0}]={1}", idx, imageUrl));  
  
string picFilename = crl.extractFilenameFromUrl(imageUrl);
```



---

## 第 6 章 crifanLib.cs之Array

此处是和数组 ( Array ) 有关的函数

### 6.1. 从给定字符串中，从指定位置，提取指定长度的子字符串:getSubStrArr

```
//given a string array 'origStrArr', get a sub string array from 'startIdx', length is 'len'  
public string[] getSubStrArr(string[] origStrArr, int startIdx, int len)  
{  
    string[] subStrArr = new string[] { };  
    if ((origStrArr != null) && (origStrArr.Length > 0) && (len > 0))  
    {  
        List<string> strList = new List<string>();  
        int endPos = startIdx + len;  
        if (endPos > origStrArr.Length)  
        {  
            endPos = origStrArr.Length;  
        }  
  
        for (int i = startIdx; i < endPos; i++)  
        {  
            //refer: http://zhidao.baidu.com/question/296384408.html  
            strList.Add(origStrArr[i]);  
        }  
  
        subStrArr = new string[len];  
        strList.CopyTo(subStrArr);  
    }  
  
    return subStrArr;  
}
```

#### 例 6.1. getSubStrArr 的使用范例

```
string[] fieldExpressions = getSubStrArr(expressions, 1, expressions.Length - 1);
```

---

## 第 7 章 crifanLib.cs之Cookie

### 7.1. 从Url中提取主机Host:extractHost

```
//extrat the Host from input url
//example: from https://skydrive.live.com/, extracted Host is "skydrive.live.com"
public string extractHost(string url)
{
    string domain = "";
    if ((url != "") && (url.Contains("/")))
    {
        string[] splited = url.Split('/');
        domain = splited[2];
    }
    return domain;
}
```

#### 例 7.1. extractHost 的使用范例

```
string host = "";
host = extractHost(url);
```

### 7.2. 从Url中提取域Domain:extractDomain

```
//extrat the domain from input url
//example: from https://skydrive.live.com/, extracted domain is ".live.com"
public string extractDomain(string url)
{
    string host = "";
    string domain = "";
    host = extractHost(url);
    if (host.Contains("."))
    {
        domain = host.Substring(host.IndexOf('.'));
    }
    return domain;
}
```

### 例 7.2. extractDomain 的使用范例

```
private string gCurDomain;  
//update latest cookies  
gCurDomain = commLib.extractDomain(getItemsUrl);
```

## 7.3. 从Url中提取域Domain的URL:getDomainUrl

```
//extrat the domain url from original url  
//from  
//http://answers.yahoo.com/question/index?qid=20130323071141AA8PffP  
//get  
//http://answers.yahoo.com  
public string getDomainUrl(string url)  
{  
    string domainUrl = "";  
  
    Regex urlRx = new Regex(@"((https)|(http)|(ftp))://[\w\-\.]+" );  
    Match foundUrl = urlRx.Match(url);  
    if (foundUrl.Success)  
    {  
        //int slashIndex = foundUrl.Index + foundUrl.Length;  
        domainUrl = url.Substring(0, foundUrl.Length);  
    }  
    else  
    {  
        domainUrl = "";  
    }  
  
    return domainUrl;  
}
```

### 例 7.3. getDomainUrl 的使用范例

## 7.4. 将Cookie的某一项的值，添加到Cookie中:addFieldToCookie

```
//add recognized cookie field: expires/domain/path/secure/httponly/version, into cookie
public bool addFieldToCookie(ref Cookie ck, pairInfo pairInfo)
{
    bool added = false;
    if (pairInfo.key != "")
    {
        string lowerKey = pairInfo.key.ToLower();
        switch (lowerKey)
        {
            case "expires":
                DateTime expireDatetime;
                if (DateTime.TryParse(pairInfo.value, out expireDatetime))
                {
                    // note: here converted to local time: GMT +8
                    ck.Expires = expireDatetime;

                    //update expired filed
                    if (DateTime.Now.Ticks > ck.Expires.Ticks)
                    {
                        ck.Expired = true;
                    }

                    added = true;
                }
                break;
            case "domain":
                ck.Domain = pairInfo.value;
                added = true;
                break;
            case "secure":
                ck.Secure = true;
                added = true;
                break;
            case "path":
                ck.Path = pairInfo.value;
                added = true;
                break;
            case "httponly":
                ck.HttpOnly = true;
                added = true;
                break;
            case "version":
                int versionValue;
                if (int.TryParse(pairInfo.value, out versionValue))
                {
                    ck.Version = versionValue;
                    added = true;
                }
                break;
            default:
                break;
        }
    }

    return added;
}
//addFieldToCookie
```

### 例 7.4. addFieldToCookie 的使用范例

```
public bool parseSingleCookie(string cookieStr, ref Cookie ck)
{
    bool parsedOk = true;
    //Cookie ck = new Cookie();
    //string[] expressions =
cookieStr.Split(";".ToCharArray(),StringSplitOptions.RemoveEmptyEntries);
    //refer: http://msdn.microsoft.com/en-us/library/b873y76a.aspx
    string[] expressions = cookieStr.Split(new char[] { ';' },
StringSplitOptions.RemoveEmptyEntries);
    //get cookie name and value
    pairItem pair = new pairItem();
    if (parseCookieNameValue(expressions[0], out pair))
    {
        ck.Name = pair.key;
        ck.Value = pair.value;

        string[] fieldExpressions = getSubStrArr(expressions, 1, expressions.Length - 1);
        foreach (string eachExpression in fieldExpressions)
        {
            //parse key and value
            if (parseCookieField(eachExpression, out pair))
            {
                // add to cookie field if possible
                addFieldToCookie(ref ck, pair);
            }
        }
    }
}
```

## 7.5. 判断字符串是否是有效的cookie的某一项:isValidCookieField

```
public bool isValidCookieField(string cookieKey)
{
    return cookieFieldList.Contains(cookieKey.ToLower());
}
```

### 例 7.5. isValidCookieField 的使用范例

```
pair.key = ckFieldExpr.Substring(0, equalPos);
pair.key = pair.key.Trim();
if (isValidCookieField(pair.key))
```

```
{
    // only process while is valid cookie field
    pair.value = ckFieldExpr.Substring(equalPos + 1);
    pair.value = pair.value.Trim();
    parsedOK = true;
}
```

## 7.6. 校验Cookie的名字是否有效/合法:isValidCookieName

```
//cookie field example:
//
WLSRDAuth=FAAaARQL3KgEDBNbW84gMYrDN0fBab7xkQNmAAAEgAAACN7OQIVEO14E2ADnX8vEiz8fTuV7
DI6wTk5vXZbi2SEOHjt%2BbfDJMZGybHQm4NADcA9Qj/tBZOJ/
ASo5d9w3c1bTIU1jKzcm2wecJ5JMjvdmTCj4J0oy1oyxbMPzTc0iVhmDoyCIU1dgaaVQ15oF6LTQZBrA0EXdBxq6/
%2BUgYYB9DJdkSM/yFBXb2bXRTRgNJ1lruDtyWe%2Bm21bzKWS/
zFtTQEE56bIvn5ITesFu4U8XaFkCP/FYLiHj6gpHW2j0t
%2BvvxWUKt3jAnWY1Tt6sXhuSx6CFVDH4EYEEUALuqyxbQo2ugNwDkP9V5O%2B5FAyCf;
path=/; domain=.livefilestore.com; HttpOnly;;
//
WLSRDSecAuth=FAAaARQL3KgEDBNbW84gMYrDN0fBab7xkQNmAAAEgAAACJFcaqD2IuX42ACdJP23wgEz1q
%2BgXPFhyTzOCgQNBVfvpfCptSccxJHDixy47Hq8Cr6RGUeXSpipLSIFHumjX5%2BvcJWkqxDEczrmBsdGnUcbz4
%2BK310KIzq9J35DgfNdkl12oYX5LMIBzdiTLwN1%2Bx9DgsYmmgxPbcuZPe/7y7dlb00jNNd8p/
rKtG4KLLT4w3EZkUAOcUwGF746qfzngDIOvXWVvZjGzA; path=/; domain=.livefilestore.com;
HttpOnly; secure;;
//RPSShare=1; path=/;
//ANON=A=DE389D4D076BF47BCAE4DC05FFFFFFF&E=c44&W=1; path=/;
domain=.livefilestore.com;;
//NAP=V=1.9&E=bea&C=VTwb1vAsVjCeLWrDuow-
jCNgP5eS75JWWvYVe3tRppviqKixCvjggw&W=1; path=/; domain=.livefilestore.com;;
//RPSMaybe=; path=/; domain=.livefilestore.com; expires=Thu, 30-Oct-1980 16:00:00 GMT;

//check whether the cookie name is valid or not
public bool isValidCookieName(string ckName)
{
    bool isValid = true;
    if (ckName == null)
    {
        isValid = false;
    }
    else
    {
        string invalidP = @"\W+";
        Regex rx = new Regex(invalidP);
        Match foundInvalid = rx.Match(ckName);
        if (foundInvalid.Success)
        {
            isValid = false;
        }
    }
}
```

```
return isValid;
}
```

### 例 7.6. isValidCookieName 的使用范例

```
name = foundSetck.Groups[1].ToString();
value = foundSetck.Groups[2].ToString();
domain = foundSetck.Groups[3].ToString();
path = foundSetck.Groups[4].ToString();
expire = foundSetck.Groups[5].ToString();
secure = foundSetck.Groups[6].ToString();

// must: name valid and domain is not null
if (isValidCookieName(name) && (domain != ""))
{
    parseOK = true;

    parsedCk.Name = name;
    parsedCk.Value = value;
    parsedCk.Domain = domain;
    parsedCk.Path = path;
```

## 7.7. 解析Cookie的名字和值:parseCookieNameValue

```
// parse the cookie name and value
public bool parseCookieNameValue(string ckNameValueExpr, out pairItem pair)
{
    bool parsedOK = false;
    if (ckNameValueExpr == "")
    {
        pair.key = "";
        pair.value = "";
        parsedOK = false;
    }
    else
    {
        ckNameValueExpr = ckNameValueExpr.Trim();

        int equalPos = ckNameValueExpr.IndexOf('=');
        if (equalPos > 0) // is valid expression
        {
            pair.key = ckNameValueExpr.Substring(0, equalPos);
            pair.key = pair.key.Trim();
            if (isValidCookieName(pair.key))
            {
```

```
// only process while is valid cookie field
pair.value = ckNameValueExpr.Substring(equalPos + 1);
pair.value = pair.value.Trim();
parsedOK = true;
}
else
{
    pair.key = "";
    pair.value = "";
    parsedOK = false;
}
}
else
{
    pair.key = "";
    pair.value = "";
    parsedOK = false;
}
}
return parsedOK;
}
```

### 例 7.7. parseCookieNameValue 的使用范例

```
//string[] expressions =
cookieStr.Split(";",StringSplitOptions.RemoveEmptyEntries);
//refer: http://msdn.microsoft.com/en-us/library/b873y76a.aspx
string[] expressions = cookieStr.Split(new char[] { ';' },
StringSplitOptions.RemoveEmptyEntries);
//get cookie name and value
pairItem pair = new pairItem();
if (parseCookieNameValue(expressions[0], out pair))
{
```

## 7.8. 解析Cookie的项和域值:parseCookieField

```
// parse cookie field expression
public bool parseCookieField(string ckFieldExpr, out pairItem pair)
{
    bool parsedOK = false;

    if (ckFieldExpr == "")
    {
        pair.key = "";
        pair.value = "";
        parsedOK = false;
    }
    else
```



```
{
    ckFieldExpr = ckFieldExpr.Trim();

    //some specials: secure/httponly
    if (ckFieldExpr.ToLower() == "httponly")
    {
        pair.key = "httponly";
        //pair.value = "";
        pair.value = "true";
        parsedOK = true;
    }
    else if (ckFieldExpr.ToLower() == "secure")
    {
        pair.key = "secure";
        //pair.value = "";
        pair.value = "true";
        parsedOK = true;
    }
    else // normal cookie field
    {
        int equalPos = ckFieldExpr.IndexOf('=');
        if (equalPos > 0) // is valid expression
        {
            pair.key = ckFieldExpr.Substring(0, equalPos);
            pair.key = pair.key.Trim();
            if (isValidCookieField(pair.key))
            {
                // only process while is valid cookie field
                pair.value = ckFieldExpr.Substring(equalPos + 1);
                pair.value = pair.value.Trim();
                parsedOK = true;
            }
            else
            {
                pair.key = "";
                pair.value = "";
                parsedOK = false;
            }
        }
        else
        {
            pair.key = "";
            pair.value = "";
            parsedOK = false;
        }
    }
}

return parsedOK;
} //parseCookieField
```

### 例 7.8. parseCookieField 的使用范例

```

foreach (string eachExpression in fieldExpressions)
{
    //parse key and value
    if (parseCookieField(eachExpression, out pair))
    {
        // add to cookie field if possible
        addFieldToCookie(ref ck, pair);
    }
    else
    {
        // if any field fail, consider it is a abnormal cookie string, so quit with false
        parsedOk = false;
        break;
    }
}
}

```

## 7.9. 解析 ( SetCookie的 ) 字符串为单个Cookie值:parseSingleCookie

```

//parse single cookie string to a cookie
//example:
//MSPShared=1; expires=Wed, 30-Dec-2037 16:00:00
GMT;domain=login.live.com;path=/;HTTPOnly= ;version=1
//PPAuth=CkLXJYvPpNs3w!fIwMOFcraoSIAVYX3K!
CdvZwQNwg3Y7gv74iqm9MqReX8XkJqtCFeMA6GYCWmb9m7CoIw!
ID5gx3pOt8sOx1U5qQPv6ceuyiJYwmS86IW*I3BEaiyVCqFvju9BMll7!FHQeQholDsi0xqzCHuW!
Qm2mrEtQPCv!qF3Sh9tZDjKcDZDI9iMByXc6R*J!JG4eCEUHIvEaxTQtftb4oc5uGpM!
YyWT!r5jXIRyxqzsCULtWz4lsWHKzwrNIBRbF!A7ZXqXygCT8ek6luk7rarwLLJ!qaq2BvS;
domain=login.live.com;secure= ;path=/;HTTPOnly= ;version=1
public bool parseSingleCookie(string cookieStr, ref Cookie ck)
{
    bool parsedOk = true;
    //Cookie ck = new Cookie();
    //string[] expressions =
    cookieStr.Split(";",".ToCharArray(),StringSplitOptions.RemoveEmptyEntries);
    //refer: http://msdn.microsoft.com/en-us/library/b873y76a.aspx
    string[] expressions = cookieStr.Split(new char[] { ';' },
    StringSplitOptions.RemoveEmptyEntries);
    //get cookie name and value
    pairItem pair = new pairItem();
    if (parseCookieNameValue(expressions[0], out pair))
    {
        ck.Name = pair.key;
        ck.Value = pair.value;

        string[] fieldExpressions = getSubStrArr(expressions, 1, expressions.Length - 1);
        foreach (string eachExpression in fieldExpressions)
        {
            //parse key and value
            if (parseCookieField(eachExpression, out pair))
            {

```

```

        // add to cookie field if possible
        addFieldToCookie(ref ck, pair);
    }
    else
    {
        // if any field fail, consider it is a abnormal cookie string, so quit with false
        parsedOk = false;
        break;
    }
}
}
else
{
    parsedOk = false;
}

return parsedOk;
} // parseSingleCookie

```

### 例 7.9. parseSingleCookie 的使用范例

```

Cookie ck = new Cookie();
// recover it back
string recoveredCookieStr = Regex.Replace(cookieStr, @"xpires=\w{3}" + replacedChar
+ @"\s\d{2}-\w{3}-\d{4}", new MatchEvaluator(_recoverExpireField));
if (parseSingleCookie(recoveredCookieStr, ref ck))
{
    if (needAddThisCookie(ck, curDomain))
    {
        parsedCookies.Add(ck);
    }
}
}

```

## 7.10. 解析 ( Http访问所返回的 ) Set-Cookie的字符串为Cookie数组:parseSetCookie

```

// parse the Set-Cookie string (in http response header) to cookies
// Note: auto omit to parse the abnormal cookie string
// normal example for 'setCookieStr':
// MSPOK= ; expires=Thu, 30-Oct-1980 16:00:00
GMT;domain=login.live.com;path=/;HTTPOnly= ;version=1,PPAuth=Cuyf3Vp2wolkjba!
TOR*0v22UMYz36ReuiwxZZBc8umHJYPIRe4qupywVFFcIpbJyvYZ5ZDLBwV4zRM1UCjXC4tUwNuKvh21iz6gQb0
TYGfowB9VQpGA8esZ7iCRucC7d5LiP3ZAv*j4Z3MOecaJwmPHx7!
wDFdAMuQUZURhHuZwJiLzHP1j8ppchB2LExniHO6IGAdZo1f0qzSWsZ2hq*yYP6sdy*FdTTKo336Q1B0i5q8jUg
wNc6eSnFdeouX; domain=login.live.com;secure= ;path=/;HTTPOnly= ;version=1,PPLState=1;
domain=.live.com;path=/;version=1,MSPShared=1; expires=Wed, 30-Dec-2037 16:00:00
GMT;domain=login.live.com;path=/;HTTPOnly= ;version=1,MSPPre= ;domain=login.live.com;path=/;Expires=

```

```
30-Oct-1980 16:00:00 GMT,MSPCID= ; HTTPOnly= ;
domain=login.live.com;path=/;Expires=Thu, 30-Oct-1980 16:00:00
GMT,RPSTAuth=EwDoARAnAAAUWkziSC7RbDJKS1VkhugDegv7L0eAAOfCAY2+pKwbV5zUlu3XmBbgrQ8Edak
+FKtdGSuz/FpCYutqiHWdfdt0YF21US7+1bPxuLJ0MO+wVXB8GtjLKZaA0xCXIU5u01r
+DOsxSVM777DmplaUc0Q4O1+Pi9gX9cyzQLAgRKmC/
QtIbVnKDA2YAAAhIwqiXOVR/DDgBocoO/n0u48RFgh79X2Q
+gO4FI5GMc9Vtpa7SUJjZCCfoaitOmcxhEjlVmR/2ppdfJx3Ykek9OFzFd
+ijtn7K629yrVFt3O9q5L0IWoxfDh5/
daLK7lqJGKxn1KvOew0SHIOqxuuhYRW57ezFyicxkxSI3aLxYFiqHSu9pq
+TIITqiflyfcAcw4MWpvHxm9on8Y1dM2R4X3sxuwrLQBpvNsG4oIaldTYIhMEnKhmxrP6ZswxzteNqIRvMEKsxiksB
Cnm6QYBZNSPawc6aAedZioeYwaV3Z/i3tNrAUwYTqLXve8oG6ZNXL6WLT/
irKq1EMiilK6Cw8IT3G13WYdk/U9a6YZPJC8LdqR0vAHYpsu/xRF39/
On+xDNPE4keITHJBptweOeWQfsMDwvgrYnMBKAMjplZwE=;
domain=.live.com;path=/;HTTPOnly= ;version=1,RPSTAuthTime=1328679636;
domain=login.live.com;path=/;HTTPOnly= ;version=1,MSPAAuth=2OIAAMHXtDIF0tpaK1afG2n*AAXdfCnCBIF
nH5G285k7AswJOrsr*gY29AVAbsiz2UscjIGHkXiKrTvIzkV2M;
domain=.live.com;path=/;HTTPOnly= ;version=1,MSPProf=23ci9sti6DZRrkDXftt1b3IhHmdheWicTZU2zdJS9!
zCloHzMKwX30MfEAcCyOjVt*5WeFSK3I2ZahtEaK7HPFMm3INMs3r!
JxI8odP9PYRHivop5ryohTMYzWZzj3gVVurcEr5Bg6eJJws7rXOggo3cR4FuKltXwz*FVX0VWuB5*aJhRkCT1GZn*L
domain=.live.com;path=/;HTTPOnly= ;version=1,MSNPPAuth=CiGSMoUOx4gej8yQkdfBvN!
gvffvAhCPeWydcRAbcg!
O2lRhVb4gruWSX5NZCBPsyrtZKmhLhRLTUUIxxPA7LIhqW5TCV*YcInlG2f5hBzwzHt!
PORYbg79nCkw65LKG399gRGtJ4wvXdNIhHNldkBK1jVXD4PoqO1Xzdcpv4sj68U6!
oGrNK5KgRSMXXpLJmCeuhUcsRW1NmInqQXpyanjykpYocZy0vq!6PIxkj3gMaAvm!
1vO58gXM9HX9dA0GloNmCDnRv4qWDV2XKqEKp!A7jiIMWTmHup1DZ!
*YcTDX3nUVQ1zAYSMjHmmbMDxRJECz!1XEwm070w16Y40TzuKAJVugo!pyF!
V2OaCsLjZ9tdGxGwEQRyi0oWc*Z7M0FBn8Fz0Dh4DhCzl1NnGun9kOYjK5itrF1Wh17sT!
62ipv1vI8omeu0cVRww2Kv!
qM*LFgwGIPOnNHj3*VulQOuaoliN4MUUxTA4owDubYZoKAwF*yp7Mg3zq5Ds2!
I9Q$$; domain=.live.com;path=/;HTTPOnly= ;version=1,MH=MSFT;
domain=.live.com;path=/;version=1,MHW=; expires=Thu, 30-Oct-1980 16:00:00
GMT;domain=.live.com;path=/;version=1,MHList=; expires=Thu, 30-Oct-1980 16:00:00
GMT;domain=.live.com;path=/;version=1,NAP=V=1.9&E=bea&C=zfjCKKBD0TqjZIWGgRTp__NiK08Lme_0XFai
$9;domain=login.live.com;path=/,pres=; expires=Thu, 30-Oct-1980
16:00:00 GMT;domain=.live.com;path=/;version=1,LOpt=0;
domain=login.live.com;path=/;version=1,WLSSC=EgBnAQMAAAAEgAAACoAASfCD
+8dUptvK4kvFO0gS3mVG28SPT3Jo9Pz2k65r9c9KrN4ISvidiEhxXaPLCSpkfa6fxH3FbdP9UmWAA9KnzKFJu/
IQnkZC3rzzMcVUMjblUpSVVyscJHcfSXmpGGgZK4ZCxPqXaI9EZ0xWackE4k5zWugX7GR5m/
RzakyVizWAFwA1gd9vwYA7VazI9QKmk/
UCjJPECCAAAoQoAAAFwBjcmlmYW4yMDAzQGhvdG1haWwY29tAE8AABZjcmImYW4yMDAzQGhvdG1haWw
UAANQKmk/
Uf0RPAAAAAAAAAAAAAAAAAADgA1OC4yNDAuMjM2LjE5AAUAAAAAAAAAAAAAAAAABBAABAAABAAABAAAA
domain=.live.com;secure= ;path=/;HTTPOnly= ;version=1,MSPSoftVis=@72198325083833620@:~;
domain=login.live.com;path=/;version=1
// here now support parse the un-correct Set-Cookie:
// MSPReq=;/Version=1;version&lt=1328770452&id=250915&co=1;
path=/;version=1,MSPVis=$9;
Version=1;version=1$250915;domain=login.live.com;path=/,MSPSoftVis=@72198325083833620@:~;
domain=login.live.com;path=/;version=1,MSPBack=1328770312;
domain=login.live.com;path=/;version=1
public CookieCollection parseSetCookie(string setCookieStr, string curDomain)
{
    CookieCollection parsedCookies = new CookieCollection();

    // process for expires and Expires field, for it contains ','
    //refer: http://www.yaosansi.com/post/682.html
    // may contains expires or Expires, so following use xpires
```

```

string commaReplaced = Regex.Replace(setCookieStr, @"xpires=\w{3},\s\d{2}-\w{3}-\d{4}", new MatchEvaluator(_processExpireField));
string[] cookieStrArr = commaReplaced.Split(',');
foreach (string cookieStr in cookieStrArr)
{
    Cookie ck = new Cookie();
    // recover it back
    string recoveredCookieStr = Regex.Replace(cookieStr, @"xpires=\w{3}" + replacedChar + @"\s\d{2}-\w{3}-\d{4}", new MatchEvaluator(_recoverExpireField));
    if (parseSingleCookie(recoveredCookieStr, ref ck))
    {
        if (needAddThisCookie(ck, curDomain))
        {
            parsedCookies.Add(ck);
        }
    }
}

return parsedCookies;
}

```

函数所输入的setCookieStr的值，是类似这种的：

```

MSPOK= ; expires=Thu, 30-Oct-1980 16:00:00
GMT;domain=login.live.com;path=/;HTTPOnly= ;version=1,PPAuth=Cuyf3Vp2wolkjba!
TOR*0v22UMYz36ReuiwxZZBc8umHJYPIRe4qupywVFFcIpbJyvYZ5ZDLBwV4zRM1UCjXC4tUwNuKvh21iz6gQb0
TYGfowB9VQpGA8esZ7iCRucC7d5LiP3ZAv*j4Z3MOecaJwmPHx7!
wDFdAMuQUZURhHuZWJiLzHP1j8ppchB2LExnlHO6IGAdZo1f0qzSWsZ2hq*yYP6sdy*FdTTKo336Q1B0i5q8jUg
wNc6eSnFDeouX; domain=login.live.com;secure= ;path=/;HTTPOnly= ;version=1,PPLState=1;
domain=.live.com;path=/;version=1,MSPShared=1; expires=Wed, 30-Dec-2037 16:00:00
GMT;domain=login.live.com;path=/;HTTPOnly= ;version=1,MSPPre= ;domain=login.live.com;path=/;Expires=
30-Oct-1980 16:00:00 GMT,MSPCID= ; HTTPOnly= ;
domain=login.live.com;path=/;Expires=Thu, 30-Oct-1980 16:00:00
GMT,RPSTAuth=EwDoARAnAAAUWkziSC7RbDJKS1VkhugDegv7L0eAAOfCAY2+pKwbV5zUlu3XmBbgrQ8Edak
+FKtdGSuz/FpCYutqiHwdfdt0YF21US7+1bPxuLJ0MO+wVXB8GtjLKZaA0xCXIU5u01r
+DOsxSVM777DmplaUc0Q4O1+Pi9gX9cyzQLAgRkMc/
QtIbVnkDA2YAAAhIwqiXOVR/DDgBocoO/n0u48RFGH79X2Q
+gO4FI5GMc9Vtpa7SUJjZCCfoaitOmcxhEjlVmR/2ppdfJx3Ykek9OFzFd
+ijtn7K629yrVFt3O9q5L0IWoxfDh5/
daLK7lqjGKxn1KvOew0SHIOqxuuhYRW57ezFyicxkxSI3aLxYFiqHSu9ppq
+TIITqiflyfcAcw4MWpvHxm9on8Y1dM2R4X3sxuwrLQBpvNsG4oIaldTYIhMEnKhmxrP6ZswxzteNqIRvMEKsxiksB
Cnm6QYBZNSPawc6aAedZioeYwaV3Z/i3tNrAUwYTqLXve8oG6ZNXL6WLT/
irKq1EMilK6Cw8IT3G13WYdk/U9a6YZPJC8LdqR0vAHYpsu/xRF39/
On+xDNPE4keITHJBptweOeWQfsMDwvgrYnMBKAMjplZwE=;
domain=.live.com;path=/;HTTPOnly= ;version=1,RPSTAuthTime=1328679636;
domain=login.live.com;path=/;HTTPOnly= ;version=1,MSPAuth=2OIAAMHXtDIF0tpaK1afG2n*AAXdfCnCBIF
nH5G285k7AswJOrsr*gY29AVAbsiz2UscjIGHkXiKrTvIzkV2M;
domain=.live.com;path=/;HTTPOnly= ;version=1,MSPPProf=23ci9sti6DZRrkDXftt1b3lHhMdheWicTZU2zdJS9!
zCloHzMKwX30MfEAcCyOjVt*5WeFSK3l2ZahtEaK7HPFMm3INMs3r!
JxI8odP9PYRHivop5ryohTMYzWZj3gVVurcEr5Bg6eJjws7rXOggo3cR4FuKLTxwz*FVX0VWuB5*aJhRkCT1GZn*L
domain=.live.com;path=/;HTTPOnly= ;version=1,MSNPPAuth=CiGSMoUOx4gej8yQkdfBvN!
gvffvAhCPeWydcRAbcg!
O2lRhVb4gruWSX5NZCBPsyrZKMHlHRLTUUIxxPA7LIhqW5TCV*YcInlG2f5hBzwzHt!
PORybg79nCkvw65LKG399gRGtJ4wvXdnIhHNIldkBK1jVXD4PoqO1Xzdcpv4sj68U6!

```

```
oGrNK5KgRSMXXpLJmCeehUcsRW1NmInqQXpyanjykpYOcZy0vq!6PIxkj3gMaAvm!
1vO58gXM9HX9dA0GloNmCDnRv4qWDV2XKqEKp!A7jiIMWTmHup1DZ!
*YcTDX3nUVQ1zAYSMjHmmbMDxRJECz!1XEwm070w16Y40TzuKAJVugo!pyF!
V2OaCsLjZ9tdGxGwEQRyi0oWc*Z7M0FBn8Fz0Dh4DhCzl1NnGun9kOYjK5itrF1Wh17sT!
62ipv1vI8omeu0cVRww2Kv!
qM*LFgwGIPOnNHj3*VulQOuaoliN4MUUxTA4owDubYZoKAwF*yp7Mg3zq5Ds2!
l9Q$$; domain=.live.com;path=/;HTTPOnly= ;version=1,MH=MSFT;
domain=.live.com;path=/;version=1,MHW=; expires=Thu, 30-Oct-1980 16:00:00
GMT;domain=.live.com;path=/;version=1,MHList=; expires=Thu, 30-Oct-1980 16:00:00
GMT;domain=.live.com;path=/;version=1,NAP=V=1.9&E=bea&C=zfjCKKBD0TqjZIWGgRTp__NiK08Lme_0XFai
$9;domain=login.live.com;path=/,pres=; expires=Thu, 30-Oct-1980
16:00:00 GMT;domain=.live.com;path=/;version=1,LOpt=0;
domain=login.live.com;path=/;version=1,WLSSC=EgBnAQMAAAAEgAAACoAASfCD
+8dUptvK4kvFO0gS3mVG28SPT3Jo9Pz2k65r9c9KrN4ISvidiEhxXaPLCSpkfa6fxH3FbdP9UmWAa9KnzKFJu/
lQNkZC3rzzMcVUMjblUpSVVyscJHcfSXmpGGgZK4ZCxPqXaIl9EZ0xWackE4k5zWugX7GR5m/
RzakyVizWAFwA1gD9vwYA7Vazl9QKMk/
UCjJPECCAAAoQoAAAFwBjcmlmYW4yMDAzQGhvdG1haWwUy29tAE8AABZjcmImYW4yMDAzQGhvdG1haWwU
UAANQKMk/
Uf0RPAAAAAAAAAAAAAAAAAADgA1OC4yNDAuMjM2LjE5AAUAAAAAAAAAAAAAAAAABBAABAAABAAABAAAAA
domain=.live.com;secure= ;path=/;HTTPOnly= ;version=1,MSPSoftVis=@72198325083833620@:@;
domain=login.live.com;path=/;version=1
```

此处同时支持解析那些“非正常”的Set-Cookie：

```
MSPRequ=;/;Version=1;version&lt;=1328770452&id=250915&co=1;
path=;/;version=1,MSPVis=$9;
Version=1;version=1$250915;domain=login.live.com;path=/,MSPSoftVis=@72198325083833620@:@;
domain=login.live.com;path=/;version=1,MSPBack=1328770312;
domain=login.live.com;path=/;version=1
```

### 例 7.10. parseSetCookie 的使用范例

```
resp = (HttpWebResponse)req.GetResponse();
//update latest cookies
gCurDomain = commLib.extractDomain(getItemsUrl);
CookieCollection parsedCookies = commLib.parseSetCookie(resp.Headers["Set-Cookie"],
gCurDomain);
commLib.updateLocalCookies(parsedCookies, ref skydriveCookies);
```

另外一个例子：

```
resp = (HttpWebResponse)req.GetResponse();
// here resp.Cookies may be uncorrect, so parse the returned Set-Cookie to get real cookies
parsedCookies = commLib.parseSetCookie(resp.Headers["Set-Cookie"], gCurDomain);
commLib.updateLocalCookies(parsedCookies, ref skydriveCookies);
```

[【已解决】又发现一个C#中解析Set-Cookie的一个bug：无故地添加cookie的path域<sup>1</sup>中的例子：](#)

```
HttpWebResponse addNk1Response = crl.getUrIResponse(addNk1Url, headerDict:
headerDict, postDict: postDict);//<script>location.href='/add/'</script>
String curDomain = crl.extractHost(addPhpUrl);//new.guguyu.com
CookieCollection parsedCookies = crl.parseSetCookie(addNk1Response.Headers["Set-
Cookie"], curDomain);
CookieCollection curCookies = crl.getCurCookies();
crl.updateLocalCookies(parsedCookies, ref curCookies);
crl.setCurCookies(curCookies);
```

为了更加方便使用，又添加了一个重载函数：

```
// parse Set-Cookie string part into cookies
// leave current domain to empty, means omit the parsed cookie, which is not set its
domain value
public CookieCollection parseSetCookie(string setCookieStr)
{
    return parseSetCookie(setCookieStr, "");
}
```

所以上述调用此函数时，也可以不指定对应的domain：

```
resp = (HttpWebResponse)req.GetResponse();
//update latest cookies
CookieCollection parsedCookies = commLib.parseSetCookie(resp.Headers["Set-Cookie"]);
commLib.updateLocalCookies(parsedCookies, ref skydriveCookies);
```

当然，此时要注意，domain为空的cookie，一般来说，在后续的http的请求中，往往都是由于domain不匹配，而变成无效的cookie。

所以此处你需要知道自己在干什么，搞清楚了，再去使用此不指定domain的版本的parseSetCookie。

## 7.11. 解析Javascript中的setCookie为Cookie变量:parseJsSetCookie

```
//parse Javascript string "$Cookie.setCookie(XXX);" to a cookie
// input example:
//
$Cookie.setCookie('wla42','cHJveHktYmF5LnB2dC1jb250YWN0cy5tc24uY29tfGJ5MioxLDIBOEI4QkY1MDFBMzI=
Date(1328842189083.44),1);
```

<sup>1</sup> [http://www.crifan.com/csharp\\_another\\_set\\_cookie\\_parse\\_bug\\_add\\_wrong\\_path/](http://www.crifan.com/csharp_another_set_cookie_parse_bug_add_wrong_path/)

```

//
$Cookie.setCookie('wla42','YnkyKjEsOUE4QjhCRjUwMUEzOEEzNiwwLCww','live.com','/',new
Date(1329198041411.84),1);
//$Cookie.setCookie('wla42', 'YnkyKjEsOUE4QjhCRjUwMUEzOEEzNiwwLCww', 'live.com', '/',
new Date(1329440307389.9), 1);
//$Cookie.setCookie('wla42',
'cHJveHktYmF5LnB2dC1jb250YWN0cy5tc24uY29tfGJ5MioxLDIBOEI4QkY1MDFBMzhBMzYsMSwwLDA=',
'live.com', '/', new Date(1329440307483.5), 1);
//$Cookie.setCookie('wls', 'A|eyJV-t:a*nS', '.live.com', '/', null, 1);
//$Cookie.setCookie('MSNPPAuth',''.live.com','/',new Date(1327971507311.9),1);
public bool parseJsSetCookie(string singleSetCookieStr, out Cookie parsedCk)
{
    bool parseOK = false;
    parsedCk = new Cookie();

    string name = "";
    string value = "";
    string domain = "";
    string path = "";
    string expire = "";
    string secure = "";

    //                1=name    2=value    3=domain    4=path    5=expire
6=secure
    string setckP = @"\"$Cookie\\.setCookie\\('(\\w+)',\\s*(.*?)',\\s*('[\\w\\.]+)',\\s*(.+)',\\s*(.+?),
\\s*(\\d?)\\)";
    Regex setckRx = new Regex(setckP);
    Match foundSetck = setckRx.Match(singleSetCookieStr);
    if (foundSetck.Success)
    {
        name = foundSetck.Groups[1].ToString();
        value = foundSetck.Groups[2].ToString();
        domain = foundSetck.Groups[3].ToString();
        path = foundSetck.Groups[4].ToString();
        expire = foundSetck.Groups[5].ToString();
        secure = foundSetck.Groups[6].ToString();

        // must: name valid and domain is not null
        if (isValidCookieName(name) && (domain != ""))
        {
            parseOK = true;

            parsedCk.Name = name;
            parsedCk.Value = value;
            parsedCk.Domain = domain;
            parsedCk.Path = path;

            // note, here even parse expire field fail
            //do not consider it must fail to parse the whole cookie
            if (expire.Trim() == "null")
            {
                // do nothing
            }
            else
            {
                DateTime expireTime;
                if (parseJsNewDate(expire, out expireTime))
                {

```



```
        parsedCk.Expires = expireTime;
    }
}

if (secure == "1")
{
    parsedCk.Secure = true;
}
else
{
    parsedCk.Secure = false;
}
} //if (isValidCookieName(name) && (domain != ""))
} //foundSetck.Success

return parseOK;
}
```

### 例 7.11. parseJsSetCookie 的使用范例

## 7.12. 判断Cookie是否已经过期/失效/无效:isCookieExpired

```
//check whether a cookie is expired
//if expired property is set, then just return it value
//if not set, check whether is a session cookie, if is, then not expired
//if expires is set, check its real time is expired or not
public bool isCookieExpired(Cookie ck)
{
    bool isExpired = false;

    if ((ck != null) && (ck.Name != ""))
    {
        if (ck.Expired)
        {
            isExpired = true;
        }
        else
        {
            DateTime initExpiresValue = (new Cookie()).Expires;
            DateTime expires = ck.Expires;

            if (expires.Equals(initExpiresValue))
            {
```

```

        // expires is not set, means this is session cookie, so here no expire
    }
    else
    {
        // has set expire value
        if (DateTime.Now.Ticks > expires.Ticks)
        {
            isExpired = true;
        }
    }
}
else
{
    isExpired = true;
}

return isExpired;
}

```

### 例 7.12. isCookieExpired 的使用范例

```

//extract cookies for upload file
cookiesForUploadFile = new CookieCollection();

foreach (Cookie ck in skydriveCookies)
{
    if ((ck.Domain == constDomainLiveCom) && (!commLib.isCookieExpired(ck)))
    {
        Cookie ckToAdd = new Cookie(ck.Name, ck.Value, ck.Path, ck.Domain);
        ckToAdd.HttpOnly = ck.HttpOnly;
        ckToAdd.Expires = ck.Expires;
        ckToAdd.Secure = ck.Secure;
        ckToAdd.Version = ck.Version;
        cookiesForUploadFile.Add(ckToAdd);
    }
}

//!!! if not seperatly set new domain value, then will overwrtten the original domain of
cookie in skydriveCookies
foreach (Cookie ckNew in cookiesForUploadFile)
{
    ckNew.Domain = constDomainUsersStorageLive;
}

```

## 7.13. 将单个Cookie添加到Cookie数组变量中:addCookieToCookies



```
}
```

### 例 7.13. addCookieToCookies 的使用范例

```
//ref CookieCollection localCookies
foreach (Cookie newCookie in cookiesToUpdate)
{
    if (isContainCookie(newCookie, omitUpdateCookies))
    {
        // need omit process this
    }
    else
    {
        addCookieToCookies(newCookie, ref localCookies);
    }
}
```

## 7.14. 判断Cookies中是否包含某个Cookie:isContainCookie

```
//check whether the cookies contains the ckToCheck cookie
//support:
//ckTocheck is Cookie/string
//cookies is Cookie/string/CookieCollection/string[]
public bool isContainCookie(object ckToCheck, object cookies)
{
    bool isContain = false;

    if ((ckToCheck != null) && (cookies != null))
    {
        string ckName = "";
        Type type = ckToCheck.GetType();

        //string typeStr = ckType.ToString();

        //if (ckType.FullName == "System.string")
        if (type.Name.ToLower() == "string")
        {
            ckName = (string)ckToCheck;
        }
        else if (type.Name == "Cookie")
        {
            ckName = ((Cookie)ckToCheck).Name;
        }
    }
}
```

```
if (ckName != "")
{
    type = cookies.GetType();

    // is single Cookie
    if (type.Name == "Cookie")
    {
        if (ckName == ((Cookie)cookies).Name)
        {
            isContain = true;
        }
    }
    // is CookieCollection
    else if (type.Name == "CookieCollection")
    {
        foreach (Cookie ck in (CookieCollection)cookies)
        {
            if (ckName == ck.Name)
            {
                isContain = true;
                break;
            }
        }
    }
    // is single cookie name string
    else if (type.Name.ToLower() == "string")
    {
        if (ckName == (string)cookies)
        {
            isContain = true;
        }
    }
    // is cookie name string[]
    else if (type.Name.ToLower() == "string[]")
    {
        foreach (string name in ((string[])cookies))
        {
            if (ckName == name)
            {
                isContain = true;
                break;
            }
        }
    }
}

return isContain;
} //isContainCookie
```

### 例 7.14. isContainCookie 的使用范例

```
foreach (Cookie newCookie in cookiesToUpdate)
```

```

{
    if (isContainCookie(newCookie, omitUpdateCookies))
    {
        // need omit process this
    }
    else
    {
        addCookieToCookies(newCookie, ref localCookies);
    }
}

```

## 7.15. 更新本地Cookie:updateLocalCookies

主要用于管理本地Cookie。

比如提交某http请求后，返回一些cookie，然后加入到本地Cookies数组变量中，用于后续使用。

```

// update cookiesToUpdate to localCookies
// if omitUpdateCookies designated, then omit cookies of omitUpdateCookies in
cookiesToUpdate
public void updateLocalCookies(CookieCollection cookiesToUpdate, ref CookieCollection
localCookies, object omitUpdateCookies)
{
    if (cookiesToUpdate.Count > 0)
    {
        if (localCookies == null)
        {
            localCookies = cookiesToUpdate;
        }
        else
        {
            foreach (Cookie newCookie in cookiesToUpdate)
            {
                if (isContainCookie(newCookie, omitUpdateCookies))
                {
                    // need omit process this
                }
                else
                {
                    addCookieToCookies(newCookie, ref localCookies);
                }
            }
        }
    }
}
}

//updateLocalCookies

//update cookiesToUpdate to localCookies
public void updateLocalCookies(CookieCollection cookiesToUpdate, ref CookieCollection
localCookies)
{
    updateLocalCookies(cookiesToUpdate, ref localCookies, null);
}
}

```

**例 7.15. updateLocalCookies 的使用范例**

```
resp = (HttpWebResponse)req.GetResponse();
updateLocalCookies(resp.Cookies, ref curCookies);
```

## 7.16. 从一个CookieCollection获得一个Cookie的值:getCookieVal

```
// given a cookie name ckName, get its value from CookieCollection cookies
public bool getCookieVal(string ckName, ref CookieCollection cookies, out string ckVal)
{
    //string ckVal = "";
    ckVal = "";
    bool gotValue = false;

    foreach (Cookie ck in cookies)
    {
        if (ck.Name == ckName)
        {
            gotValue = true;
            ckVal = ck.Value;
            break;
        }
    }

    return gotValue;
}
```

**例 7.16. getCookieVal 的使用范例**

---

# 第 8 章 crifanLib.cs之Serialize/Deserialize

## 8.1. 将一个对象序列化成字符串:serializeObjToStr

```
// serialize an object to string
public bool serializeObjToStr(Object obj, out string serializedStr)
{
    bool serializeOk = false;
    serializedStr = "";
    try
    {
        MemoryStream memoryStream = new MemoryStream();
        BinaryFormatter binaryFormatter = new BinaryFormatter();
        binaryFormatter.Serialize(memoryStream, obj);
        serializedStr = System.Convert.ToBase64String(memoryStream.ToArray());

        serializeOk = true;
    }
    catch
    {
        serializeOk = false;
    }

    return serializeOk;
}
```

### 例 8.1. serializeObjToStr 的使用范例

```
[Serializable]
public struct loginInfo_t
{
    public bool valid;
    public string username;
    public string cid;
    public string appid;
    public string bitProtocol;
    public string canary;
    public CookieCollection cookies;
    public DateTime createTime; // record the login info(cookie) create time
    public DateTime lastUpdateTime;// last update the login info(cookie)'s time
};

private bool updateLoginInfo(skydrive.loginInfo_t loginInfo)
{
```



```
bool updateOk = false;

string serializedStr = "";

loginInfo.lastUpdateTime = DateTime.Now;

if (skydrive.comLib.serializeObjToStr(loginInfo, out serializedStr))
{
    Settings.Default.loginInfoStr = serializedStr;
    Settings.Default.Save();

    updateOk = true;
}
```

## 8.2. 将字符串反序列化为对象:deserializeStrToObj

```
// deserialize the string to an object
public bool deserializeStrToObj(string serializedStr, out object deserializedObj)
{
    bool deserializeOk = false;
    deserializedObj = null;

    try
    {
        byte[] restoredBytes = System.Convert.FromBase64String(serializedStr);
        MemoryStream restoredMemoryStream = new MemoryStream(restoredBytes);
        BinaryFormatter binaryFormatter = new BinaryFormatter();
        deserializedObj = binaryFormatter.Deserialize(restoredMemoryStream);

        deserializeOk = true;
    }
    catch
    {
        deserializeOk = false;
    }

    return deserializeOk;
}
```

### 例 8.2. deserializeStrToObj 的使用范例

```
//restore login info
object deserializedObj = null;
if (skydrive.comLib.deserializeStrToObj(Settings.Default.loginInfoStr, out deserializedObj))
```

```
{  
    loginInfo = (skydrive.loginInfo_t)deserializedObj;
```

---

# 第 9 章 crifanLib.cs之Http

此处是和网络 ( Http等 ) 有关的函数

## 9.1. 设置代理:setProxy

```
/* set proxy
 * Note:
 * 1. current only support http proxy
 * 2. current only support single proxy
 */
public void setProxy(string proxyIp, int proxyPort)
{
    gProxy = new WebProxy(proxyIp, proxyPort);
}
```

### 例 9.1. setProxy 的使用范例

```
public crifanLib cri;
cri = new crifanLib();
cri.setProxy("127.0.0.1", 8087);
```

然后后续的 ( 去用getUrlRespHtml等等 ) 去访问网络, 就会自动使用该代理了。

## 9.2. 清除当前cookie:clearCurCookies

```
/*
 * Note: currently support auto handle cookies
 * currently only support single caller -> multiple caller of these functions will cause cookies
accumulated
 * you can clear previous cookies to avoid unexpected result by call clearCurCookies
 */
public void clearCurCookies()
{
    if (curCookies != null)
    {
        curCookies = null;
        curCookies = new CookieCollection();
    }
}
```

### 例 9.2. clearCurCookies 的使用范例

```
//http://www.crifan.com/
example_of_how_to_use_ie9_f12_to_capture_the_real_music_mp3_address_of_some_songtaste_musc/
// here must clear previous cookies
// otherwise access html with previous cookies will get fault html:
//提示信息: 对不起, 该用户不存在! 3 秒钟以后系统将自动跳转!
crl.clearCurCookies();

string respHtml = "";
respHtml = crl.getUrlRespHtml(songInfo.url, stHtmlCharset);
```

另外[InsertSkydriveFiles](#)<sup>1</sup>中的一个例子：

```
private void clearGolobalValues()
{
    //gCurDomain = "";
    skydriveCookies = null;
    commLib.clearCurCookies();
}
```

## 9.3. 获得当前cookie:getCurCookies

```
/* get current cookies */
public CookieCollection getCurCookies()
{
    return curCookies;
}
```

### 例 9.3. getCurCookies 的使用范例

```
string primeRespHtml = getSkydriveRespHtmlLogin(ref resp);
skydriveCookies = getCurCookies();
```

另外【已解决】又发现[一个C#中解析Set-Cookie的一个bug：无故地添加cookie的path域](#)<sup>2</sup>中的一个例子：

<sup>1</sup> <http://code.google.com/p/insertskydrivefiles/>

<sup>2</sup> [http://www.crifan.com/csharp\\_another\\_set\\_cookie\\_parse\\_bug\\_add\\_wrong\\_path/](http://www.crifan.com/csharp_another_set_cookie_parse_bug_add_wrong_path/)

```
crl = new crifanLib();

HttpWebResponse addNk1Response = crl.getUrlResponse(addNk1Url, headerDict:
headerDict, postDict: postDict);//<script>location.href='/add/'</script>
String curDomain = crl.extractHost(addPhpUrl);//new.guguyu.com
CookieCollection parsedCookies = crl.parseSetCookie(addNk1Response.Headers["Set-
Cookie"], curDomain);
CookieCollection curCookies = crl.getCurCookies();
crl.updateLocalCookies(parsedCookies, ref curCookies);
crl.setCurCookies(curCookies);
```

## 9.4. 设置当前cookie:setCurCookies

主要用于，重置当前的cookie，设置为所需的状态。

```
/* set current cookies */
public void setCurCookies(CookieCollection cookies)
{
    curCookies = cookies;
}
```

### 例 9.4. setCurCookies 的使用范例

```
skydriveCookies = new CookieCollection();
skydriveCookies = loginInfo.cookies;
setCurCookies(skydriveCookies);
```

另外【已解决】又发现一个C#中解析Set-Cookie的一个bug：无故地添加cookie的path域<sup>3</sup>中的一个例子：

```
crl = new crifanLib();

HttpWebResponse addNk1Response = crl.getUrlResponse(addNk1Url, headerDict:
headerDict, postDict: postDict);//<script>location.href='/add/'</script>
String curDomain = crl.extractHost(addPhpUrl);//new.guguyu.com
CookieCollection parsedCookies = crl.parseSetCookie(addNk1Response.Headers["Set-
Cookie"], curDomain);
CookieCollection curCookies = crl.getCurCookies();
crl.updateLocalCookies(parsedCookies, ref curCookies);
crl.setCurCookies(curCookies);
```

<sup>3</sup> [http://www.crifan.com/csharp\\_another\\_set\\_cookie\\_parse\\_bug\\_add\\_wrong\\_path/](http://www.crifan.com/csharp_another_set_cookie_parse_bug_add_wrong_path/)

## 9.5. 获得Url地址的响应:getUrlResponse

```
/* get url's response
**/
public HttpWebResponse getUrlResponse(string url,
    Dictionary<string, string> headerDict = defHeaderDict,
    Dictionary<string, string> postDict = defPostDict,
    int timeout = defTimeout,
    string postDataStr = defPostDataStr,
    int readWriteTimeout = defReadWriteTimeout)
{
#if USE_GETURLRESPONSE_BW
    //BackgroundWorker Version getUrlResponse
    HttpWebResponse localCurResp = null;
    getUrlResponse_bw(url, headerDict, postDict, timeout, postDataStr, readWriteTimeout);
    while (bNotCompleted_resp)
    {
        System.Windows.Forms.Application.DoEvents();
    }
    localCurResp = gCurResp;

    //clear
    gCurResp = null;

    return localCurResp;
#else
    //non-BackgroundWorker Version getUrlResponse
    return _getUrlResponse(url, headerDict, postDict, timeout, postDataStr);
#endif
}
```

从上面的代码中可以看出，此处的getUrlResponse内部的实现，是依赖于是否设置宏USE\_GETURLRESPONSE\_BW，而去调用对应的BackgroundWorker版本的，还是非BackgroundWorker版本的\_getUrlResponse

此处，getUrlResponse，是用来返回HttpWebResponse的，且支持N多参数。

### 9.5.1. getUrlResponse的参数详解

下面就对于getUrlResponse的各个参数，进行详细解释一下：

#### 9.5.1.1. getUrlResponse的参数:url

要访问的url地址

必填参数，无默认值。

支持http，也支持https类型的地址。

#### 9.5.1.2. getUrlResponse的参数:headerDict

headerDict的意思是，header的dict，即用于存放对应的header信息

默认的headerDict的值为defHeaderDict

defHeaderDict值是null :

```
private const Dictionary<string, string> defHeaderDict = null;
```

作用是, 当不指定对应的header信息时, 默认为空

常见用法中, 一般也不需要指定此headerDict

当然, 有时候, 需要用到一些header, 比如其中最最常见的referer等等。

### 9.5.1.3. getUrlResponse的参数:postDict

postDict即POST的dict, 用于存放post数据。

默认的postDict的值为defPostDict

defPostDict值是null :

```
private const Dictionary<string, string> defPostDict = null;
```

一般的GET时, 无需指定此参数。

只有当是POST时, 才可能会用到此postDict。

### 9.5.1.4. getUrlResponse的参数:timeout

timeout用于指定网络超时的最大允许时间, 单位是毫秒ms。

默认的timeout的值为defTimeout

defTimeout值是30000毫秒==30秒 :

```
private const int defTimeout = 30 * 1000;
```

注意, 此timeout, 是针对于http网络发送请求后, 得到服务器的响应之前, 这段时间, 是否超时, 即和GetResponse和GetRequestStream有关。

一般来说, 也不需要设置此timeout, 即无需改变对应的默认超时时间。

当然, 如果有需要, 可以根据你自己的情况修改为更合适的值。

### 9.5.1.5. getUrlResponse的参数:postDataStr

postDataStr是用来传递, 特殊的POST的数据是以回车为分隔符的那些POST数据的。

postDataStr的默认值为defPostDataStr

defPostDataStr值也是null :

```
private const string defPostDataStr = null;
```

需要注意的是，如果是GET，很明显无需关系此参数，而如果是POST，正常情况下，也只需要去设置对应的postDict参数即可，对应的内部处理POST数据，都是以'&'为分隔符的。

但是，有些特殊的POST，POST的数据是以回车为分隔符的，比如之前折腾【[记录](#)】给BlogsToWordPress添加支持导出网易的心情随笔<sup>4</sup>时遇到这种特殊情况，此时，才需要你用到此去设置postDataStr

### 9.5.1.6. getUrlResponse的参数:readWriteTimeout

readWriteTimeout指的是，针对于获得了response后，用StreamReader去read或write时，对应的超时时间。单位是毫秒ms。

readWriteTimeout的默认值是defReadWriteTimeout

defReadWriteTimeout值是30000毫秒==30秒：

```
private const int defReadWriteTimeout = 30 * 1000;
```

注意，参考微软官网的解释：[HttpWebRequest.ReadWriteTimeout 属性](#)<sup>5</sup>其默认的ReadWriteTimeout是300秒=5分钟，太长了。

所以，此处才把默认时间改短一些的，否则，5分钟的超时时间，太长了。

此参数，是经过多次折腾后，才搞明白的，详见：[【已解决】C#中在GetResponseStream得到的Stream后,通过StreamReader去ReadLine或ReadToEnd会无限期挂掉 + 给StreamReader添加Timeout支持](#)<sup>6</sup>

## 9.5.2. getUrlResponse 的用法详解

getUrlResponse参数太多，但是其实也是自己一点点，从无到有，加进去的，以适应各种应用需求。

此处，就来通过例子来说明，如何使用此getUrlResponse函数。

### 9.5.2.1. 被getUrlRespHtml调用

其实，此处的getUrlResponse，在绝大多数的时候，都是被，我的另外一个函数：getUrlRespHtml，所调用的。

即，getUrlRespHtml，调用，getUrlResponse，获得对应的HttpWebResponse，然后后续再处理，得到返回的html的。

所以，用起来，一般都是这样的：

#### 例 9.5. getUrlResponse 的使用范例：被getUrlRespHtml调用

```
// valid charset:"GB18030"/"UTF-8", invliad:"UTF8"
public string getUrlRespHtml(string url,
    Dictionary<string, string> headerDict = defHeaderDict,
    string charset = defCharset,
```

<sup>4</sup> [http://www.crifan.com/blogstowordpress\\_add\\_feeling\\_card\\_for\\_163\\_netease\\_blog/](http://www.crifan.com/blogstowordpress_add_feeling_card_for_163_netease_blog/)

<sup>5</sup> <http://msdn.microsoft.com/zh-cn/library/system.net.httpwebrequest.readwritetimeout.aspx>

<sup>6</sup> [http://www.crifan.com/csharp\\_httpwebresponse\\_streamreader\\_readline\\_dead\\_again/](http://www.crifan.com/csharp_httpwebresponse_streamreader_readline_dead_again/)



```

        Dictionary<string, string> postDict = defPostDict,
        int timeout = defTimeout,
        string postDataStr = defPostDataStr,
        int readWriteTimeout = defReadWriteTimeout)
    {
        string respHtml = "";

        HttpResponseMessage resp = getUrlResponse(url, headerDict, postDict, timeout, postDataStr,
readWriteTimeout);

```

关于此种用法，更详细的代码和解释，参见下面要介绍的：[第 9.6 节 “获得Url地址返回的网页内容:getUrlRespHtml”](#)

### 9.5.2.2. 只传入url而获得对应的url的response

getUrlResponse的相对次要的用法是：当有时候，不仅仅需要html，而且也要关心和处理HttpResponseMessage时，此时，才会考虑直接调用getUrlResponse（而不是去调用getUrlRespHtml）而直接使用getUrlResponse的话，相对简单的用法就是，只传入对应的url即可：

#### 例 9.6. getUrlResponse 的使用范例：只传入url

```

const string constSkydriveUrl = "https://skydrive.live.com/";
HttpResponseMessage resp = getUrlResponse(constSkydriveUrl);

```

## 9.6. 获得Url地址返回的网页内容:getUrlRespHtml

```

// valid charset:"GB18030"/"UTF-8", invliad:"UTF8"
public string getUrlRespHtml(string url,
        Dictionary<string, string> headerDict = defHeaderDict,
        string charset = defCharset,
        Dictionary<string, string> postDict = defPostDict,
        int timeout = defTimeout,
        string postDataStr = defPostDataStr,
        int readWriteTimeout = defReadWriteTimeout)
    {
        string respHtml = "";

        HttpResponseMessage resp = getUrlResponse(url, headerDict, postDict, timeout, postDataStr,
readWriteTimeout);

        //long realRespLen = resp.ContentLength;
        if (resp != null)
        {
            StreamReader sr;
            Stream respStream = resp.GetResponseStream();

```

```
if (!string.IsNullOrEmpty(charset))
{
    Encoding htmlEncoding = Encoding.GetEncoding(charset);
    sr = new StreamReader(respStream, htmlEncoding);
}
else
{
    sr = new StreamReader(respStream);
}

try
{
    respHtml = sr.ReadToEnd();

    //while (!sr.EndOfStream)
    //{
    //    respHtml = respHtml + sr.ReadLine();
    //}

    //string curLine = "";
    //while ((curLine = sr.ReadLine()) != null)
    //{
    //    respHtml = respHtml + curLine;
    //}

    ////http://msdn.microsoft.com/zh-cn/library/system.io.streamreader.peek.aspx
    //while (sr.Peek() > -1) //while not error or not reach end of stream
    //{
    //    respHtml = respHtml + sr.ReadLine();
    //}

    //respStream.Close();
    //sr.Close();
    //resp.Close();
}
catch (Exception ex)
{
    // 【未解决】C#中StreamReader中遇到异常：未处理ObjectDisposedException,无法访问
    // 已关闭的流
    //http://www.crifan.com/
    //csharp_streamreader_unhandled_exception_objectdisposedexception_cannot_access_closed_stream
    //System.ObjectDisposedException
    respHtml = "";
}
finally
{
    if (respStream != null)
    {
        respStream.Close();
    }
    if (sr != null)
    {
        sr.Close();
    }
    if (resp != null)
    {
        resp.Close();
    }
}
```

```

    }
}

return respHtml;
}

```

## 9.6.1. getUrlRespHtml的参数详解

很明显可以看出，此处的getUrlRespHtml的很多参数，和前面介绍的[第 9.5 节 “获得Url地址的响应:getUrlResponse”](#) 非常类似。

此处，针对于getUrlRespHtml的参数，也要再解释一下：

其他参数，包括url,headerDict,postDict,timeout,postDataStr,readWriteTimeout，都和getUrlResponse的参数含义相同。所以不再赘述。

另外还有参数，需要解释一下：

- charset  
charset表示返回的网页内容，用何种字符编码去解码。

charset默认值是defCharset

defCharset的值是：

```
private const string defCharset = null;
```

此处，之所以defCharset的值，不是我们所常见的GBK，UTF-8等等，是因为此处是为了支持，当不设置charset时，就不去尝试用某种编码去解码通过StreamReader所读取出来的内容。

这样的就可以获得，原始的，返回的html，可以供有需要的人，后期进行自己的处理，比如自己去解码等等。

## 9.6.2. getUrlRespHtml 的功能详解

getUrlRespHtml内部，已经实现了足够多的，相对比较复杂的功能，对此，需要详细解释一下：

### 9.6.2.1. 内部已默认指定了IE8的User-Agent

getUrlRespHtml内部调用getUrlResponse，内部已经加上了对应的User-Agent了。

当然默认使用的是IE8的User-Agent，相关部分的代码是：

```

//IE7
const string constUserAgent_IE7_x64 = "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT
6.1; WOW64; Trident/5.0; SLCC2; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET CLR 3.0.30729;
Media Center PC 6.0; InfoPath.3; .NET4.0C; .NET4.0E)";
//IE8
const string constUserAgent_IE8_x64 = "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT
6.1; WOW64; Trident/4.0; SLCC2; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET CLR 3.0.30729;
Media Center PC 6.0; InfoPath.3; .NET4.0C; .NET4.0E)";

```

```
//IE9
const string constUserAgent_IE9_x64 = "Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; WOW64; Trident/5.0)"; // x64
const string constUserAgent_IE9_x86 = "Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; Trident/5.0)"; // x86
//Chrome
const string constUserAgent_Chrome = "Mozilla/5.0 (Windows; U; Windows NT 6.1; en-US) AppleWebKit/533.4 (KHTML, like Gecko) Chrome/5.0.375.99 Safari/533.4";
//Mozilla Firefox
const string constUserAgent_Firefox = "Mozilla/5.0 (Windows; U; Windows NT 6.1; rv:1.9.2.6) Gecko/20100625 Firefox/3.6.6";
private string gUserAgent;

gUserAgent = constUserAgent_IE8_x64;

req.UserAgent = gUserAgent;
```

所以，不会出现，被服务器当做普通的机器人或蜘蛛爬虫的情况。

### 9.6.2.2. 默认是允许自动跳转的

内部相关代码：

```
req.AllowAutoRedirect = true;
```

默认是启用了自动跳转的。

如果想要禁止自动跳转，可以去给headerDict中加上对应的"AllowAutoRedirect"为"false"的参数

更多使用实例，详见后续的例子。

### 9.6.2.3. 默认已支持解压缩html

内部相关代码：

```
req.Headers["Accept-Encoding"] = "gzip, deflate";
//req.AutomaticDecompression = DecompressionMethods.GZip;
req.AutomaticDecompression = DecompressionMethods.GZip |
DecompressionMethods.Deflate;
```

相关帖子：[【已解决】C#中HttpRequest使用Proxy后异常<sup>7</sup>](#)

### 9.6.2.4. 已支持设置（单个）代理

内部相关代码：

<sup>7</sup> [http://www.crifan.com/csharp\\_httpwebrequest\\_proxy\\_abnormal\\_not\\_working/](http://www.crifan.com/csharp_httpwebrequest_proxy_abnormal_not_working/)

```
private WebProxy gProxy = null;

req.Proxy = gProxy;
```

关于如何设置代理，详见：[第 9.1 节 “设置代理:setProxy”](#)

### 9.6.2.5. 支持网络超时设置

即前面所解释的参数：[第 9.5.1.4 节 “getUrlResponse的参数:timeout”](#)，指的是网络方面的超时，和 GetResponse和GetRequestStream有关

内部相关部分的代码是：

```
if (timeout > 0)
{
    req.Timeout = timeout;
}
```

### 9.6.2.6. 支持读写超时设置

即前面所解释的参数：[第 9.5.1.6 节 “getUrlResponse的参数:readWriteTimeout”](#)，指的是 StreamReader或StreamWriter的读写超时，和readLine之类的有关。

内部相关部分的代码是：

```
if (readWriteTimeout > 0)
{
    //default ReadWriteTimeout is 300000=300 seconds = 5 minutes !!!
    //too long, so here change to 300000 = 30 seconds
    //for support TimeOut for later StreamReader's ReadToEnd
    req.ReadWriteTimeout = readWriteTimeout;
}
```

相关折腾见：[【已解决】C#中在GetResponseStream得到的Stream后,通过StreamReader去ReadLine或ReadToEnd会无限期挂掉 + 给StreamReader添加Timeout支持](#)<sup>8</sup>

### 9.6.2.7. 支持自动处理cookie

此处已经支持，getUrlRespHtml内部，自动处理cookie。

内部相关部分的代码是：

```
CookieCollection curCookies = null;
```

<sup>8</sup> [http://www.crifan.com/csharp\\_httpwebresponse\\_streamreader\\_readline\\_dead\\_again/](http://www.crifan.com/csharp_httpwebresponse_streamreader_readline_dead_again/)

```

curCookies = new CookieCollection();

if (curCookies != null)
{
    req.CookieContainer = new CookieContainer();
    req.CookieContainer.PerDomainCapacity = 40; // following will exceed max default 20
    cookie per domain
    req.CookieContainer.Add(curCookies);
}

resp = (HttpWebResponse)req.GetResponse();
updateLocalCookies(resp.Cookies, ref curCookies);

```

其中，注意到，设置了最大支持40个cookie，是因为，之前折腾[InsertSkydriveFiles](#)<sup>9</sup>期间，遇到相对极端的情况：cookie超过默认的20多个，一个CookieContainer都装不下了，所以才改为40个，以便支持如此多的cookie。

### 9.6.3. getUrlRespHtml 的用法详解

getUrlRespHtml的参数够多，用法，也有很多种。

此处，就来通过例子来说明，如何使用此getUrlResponse函数。

#### 9.6.3.1. getUrlRespHtml用法示例：只传入url而获得html

getUrlRespHtml最常用，也是最简单的用法，就是：直接传入url，然后获得返回的html

代码如下：

##### 例 9.7. getUrlRespHtml用法示例：只传入url而获得html

```

string mainJsUrl = "http://image.songtaste.com/inc/main.js";
string respHtmlMainJs = getUrlRespHtml(mainJsUrl);

```

其中，getUrlRespHtml内部，会自动帮你处理各种细节，比如cookie，header中的User-Agent等等内容，而你就直接可以得到对应返回的html了。

#### 9.6.3.2. getUrlRespHtml用法示例：传入各种header信息

很多时候，在折腾抓取网页和模拟登陆时，都会遇到，需要额外再指定一些header，用于实现一些特定的目的。

##### 9.6.3.2.1. getUrlRespHtml用法示例：指定Referer

比如，添加对应的Referer，以便成功模拟网页逻辑，获得所需返回的内容的：

```

string tmpRespHtml = "";
Dictionary<string, string> headerDict;
//(1)to get cookies

```

<sup>9</sup> <http://code.google.com/p/insertskydrivefiles/>

```
string pageRankMainUrl = "http://pagerank.webmasterhome.cn/";
tmpRespHtml = getUrlRespHtml(pageRankMainUrl);
//(2)ask page rank
string firstBaseUrl = "http://pagerank.webmasterhome.cn/?domain=";
//http://pagerank.webmasterhome.cn/?domain=answers.yahoo.com
string firstWholeUrl = firstBaseUrl + noHttpPreDomainUrl;
headerDict = new Dictionary<string, string>();
headerDict.Add("referer", pageRankMainUrl);
tmpRespHtml = getUrlRespHtml(firstWholeUrl, headerDict: headerDict);
```



### header中的Referer支持大小写任意

由具体的实现代码：

```
string lowercaseHeader = header.ToLower();
// following are allow the caller overwrite the default header setting
if (lowercaseHeader == "referer")
{
    req.Referer = headerValue;
}
```

可以看出，此处的"referer"，写成常见的首字母大写"Referer"也是可以的。

#### 9.6.3.2.2. getUrlRespHtml用法示例：禁止自动跳转

如[第 9.6.2.2 节 “默认是允许自动跳转的”](#)所述，默认是启用了自动跳转的，想要禁止自动跳转，可以通过header去设置：

```
Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict.Add("AllowAutoRedirect", "false");
string respHtml = getUrlRespHtml(yourUrl, headerDict: headerDict);
```



### header中的AutoRedirect支持多种写法

由具体的实现代码：

```
else if (
    (lowercaseHeader == "allow-autoredirect") ||
    (lowercaseHeader == "allowautoredirect") ||
    (lowercaseHeader == "allow autoredirect")
)
{
    bool isAllow = false;
    if (bool.TryParse(headerValue, out isAllow))
    {
        req.AllowAutoRedirect = isAllow;
    }
}
```

```
    }
}
```

可以看出，此处的"AllowAutoRedirect"，写成别的形式，也是支持的，比如："allowautoredirect"，"allow-autoredirect"，"Allow-Autoredirect"，"allowautoredirect"，"Allow Autoredirect"

### 9.6.3.2.3. getUrlRespHtml用法示例：手动设置Accept

此处默认的Accept是"\*/\*"，如果想要指定不同的类型，可以手动通过header去设置：

```
Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict.Add("Accept", "text/html");
string respHtml = getUrlRespHtml(yourUrl, headerDict: headerDict);
```

关于Accept更多可能的取值，自己参考官网的解释：[14.1 Accept](#)<sup>10</sup>



#### header中的Accept支持大小写任意

由具体的实现代码：

```
else if (lowecaseHeader == "accept")
{
    req.Accept = headerValue;
}
```

可以看出，此处的"Accept"，写成别的形式，也是支持的，比如："accept"

### 9.6.3.2.4. getUrlRespHtml用法示例：不保持连接

此处默认的KeepAlive是true的，如果不想继续保持连接，则可以通过header去禁止：

```
Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict.Add("Keep-Alive", "false");
string respHtml = getUrlRespHtml(yourUrl, headerDict: headerDict);
```



#### header中的KeepAlive支持多种写法

由具体的实现代码：

```
else if (
```

<sup>10</sup> <http://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html#sec14.1>



```

        (lowercaseHeader == "keep-alive") ||
        (lowercaseHeader == "keepalive") ||
        (lowercaseHeader == "keep alive")
    )
}
bool isKeepAlive = false;
if (bool.TryParse(headerValue, out isKeepAlive))
{
    req.KeepAlive = isKeepAlive;
}
}

```

可以看出，此处的"Keep-Alive"，写成别的形式，也是支持的，比如："keep-alive"，"keepalive"，"KeepAlive"，"keep alive"，"Keep Alive"

### 9.6.3.2.5. getUrlRespHtml用法示例：设置Accept-Language

此处默认没有指定Accept-Language，有需要的话，可以去通过header设置：

```

Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict.Add("Accept-Language", "en-US"); //"zh-CN"
string respHtml = getUrlRespHtml(yourUrl, headerDict: headerDict);

```

关于Accept-Language更多可能的取值，自己参考官网的解释：[14.4 Accept-Language](#)<sup>11</sup>



### header中的Accept-Language支持多种写法

由具体的实现代码：

```

else if (
    (lowercaseHeader == "accept-language") ||
    (lowercaseHeader == "acceptlanguage") ||
    (lowercaseHeader == "accept language")
)
{
    req.Headers["Accept-Language"] = headerValue;
}

```

可以看出，此处的"Accept-Language"，写成别的形式，也是支持的，比如："accept-language"，"acceptlanguage"，"AcceptLanguage"，"accept language"，"Accept Language"

### 9.6.3.2.6. getUrlRespHtml用法示例：添加特定的User-Agent的header

如[第9.6.2.1节“内部已默认指定了IE8的User-Agent”](#)所述，我此处的getUrlRespHtml，默认添加的User-Agent是IE8的。

<sup>11</sup> <http://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html#sec14.4>

如果有需要，你可以自己换成别的，比如Firefox的User-Agent：

```
//Mozilla Firefox
const string constUserAgent_Firefox = "Mozilla/5.0 (Windows; U; Windows NT 6.1; rv:1.9.2.6)
Gecko/20100625 Firefox/3.6.6";
Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict.Add("User-Agent", constUserAgent_Firefox);
string respHtml = getUrlRespHtml(yourUrl, headerDict: headerDict);
```

其中，关于各种浏览器的User-Agent，你可以自己去网络上找到。也可以参考我代码中的值：

```
//IE7
const string constUserAgent_IE7_x64 = "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT
6.1; WOW64; Trident/5.0; SLCC2; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET CLR 3.0.30729;
Media Center PC 6.0; InfoPath.3; .NET4.0C; .NET4.0E)";
//IE8
const string constUserAgent_IE8_x64 = "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT
6.1; WOW64; Trident/4.0; SLCC2; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET CLR 3.0.30729;
Media Center PC 6.0; InfoPath.3; .NET4.0C; .NET4.0E";
//IE9
const string constUserAgent_IE9_x64 = "Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1;
WOW64; Trident/5.0)"; // x64
const string constUserAgent_IE9_x86 = "Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1;
Trident/5.0)"; // x86
//Chrome
const string constUserAgent_Chrome = "Mozilla/5.0 (Windows; U; Windows NT 6.1; en-US)
AppleWebKit/533.4 (KHTML, like Gecko) Chrome/5.0.375.99 Safari/533.4";
//Mozilla Firefox
const string constUserAgent_Firefox = "Mozilla/5.0 (Windows; U; Windows NT 6.1;
rv:1.9.2.6) Gecko/20100625 Firefox/3.6.6";
```



## header中的User-Agent支持多种写法

由具体的实现代码：

```
else if (
    (lowercaseHeader == "user-agent") ||
    (lowercaseHeader == "useragent") ||
    (lowercaseHeader == "user agent")
)
{
    req.UserAgent = headerValue;
}
```

可以看出，此处的"User-Agent"，写成别的形式，也是支持的，比如："user-agent"，"user agent"，"User Agent"，"UserAgent"，"useragent"

### 9.6.3.2.7. getUrlRespHtml用法示例：设置ContentType

此处默认情况下，对于GET，没有指定ContentType，对于POST，已经指定了"application/x-www-form-urlencoded"。

如果你有别的特殊需求，需要设置ContentType的话，可以去通过header设置：

```
Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict.Add("Content-Type", "text/plain");
string respHtml = getUrlRespHtml(yourUrl, headerDict: headerDict);
```

关于Content-Type更多可能的取值，自己参考官网的解释：[14.17 Content-Type](#)<sup>12</sup>



### header中的Content-Type支持多种写法

由具体的实现代码：

```
else if (
    (lowercaseHeader == "content-type") ||
    (lowercaseHeader == "contenttype") ||
    (lowercaseHeader == "content type")
)
{
    req.ContentType = headerValue;
}
```

可以看出，此处的"Content-Type"，写成别的形式，也是支持的，比如："content-type"，"contenttype"，"ContentType"，"content type"，"Content Type"

### 9.6.3.2.8. getUrlRespHtml用法示例：设置其他的特定的header

在很多时候，都需要设置，某些其他的，非标准的，header信息，则也可以去通过header设置。

比如，之前折腾[InsertSkydriveFiles](#)<sup>13</sup>时所用到的：

```
string createFolerUrl = "https://skydrive.live.com/API/2/AddFolder?lct=1";

Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict.Add("Accept", "application/json");
headerDict.Add("Referer", constSkydriveUrl);
headerDict.Add("Canary", gCanary);
headerDict.Add("Appid", gAppid);
headerDict.Add("X-Requested-With", "XMLHttpRequest");
headerDict.Add("Cache-Control", "no-cache");

string postDataStr = genCreateFolderPostData(folderName, parentId, cid);
```

<sup>12</sup> <http://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html#sec14.17>

<sup>13</sup> <http://code.google.com/p/insertskydrivefiles/>

```
respJson = getUrlRespHtml(createFolerUrl, headerDict:headerDict,  
postDataStr:postDataStr);
```



### 指定某些特定的header

由具体的实现代码：

```
else  
{  
    req.Headers[header] = headerValue;  
}
```

可以看出，此处，不限制你所指定的，其他某些特殊的header，但是你自己要清楚，你设置的什么header，是用来干什么用的。

### 9.6.3.3. getUrlRespHtml用法示例：设置网页字符编码charset

有时候，已经网页是某种编码的，所以为了正确解析返回的html，需要指定对应的字符编码charset：

```
string songtasteUserUrl = "http://www.songtaste.com/user/351979/";  
string songtasteHtmlCharset = "GB18030";  
string respHtmlUnicode = getUrlRespHtml(songtasteUserUrl,  
charset:songtasteHtmlCharset);
```

即可返回对应的，已经解码后的，Unicode字符串了。

### 9.6.3.4. getUrlRespHtml用法示例：设置网络超时timeout时间

如果你觉得默认的网络超时时间30秒不合适，可以自己另外指定，比如：

```
int timeoutInMilliSec = 10 * 1000;  
string respHtml = getUrlRespHtml(someUrl, timeout:timeoutInMilliSec);
```

### 9.6.3.5. getUrlRespHtml用法示例：设置Stream的读写超时readWriteTimeout时间

如果你觉得默认的Stream的读写超时时间30秒不合适，可以自己另外指定，比如：

```
int streamRdWrTimeout = 20 * 1000;  
string respHtml = getUrlRespHtml(someUrl, readWriteTimeout:streamRdWrTimeout);
```

### 9.6.3.6. getUrlRespHtml用法示例：POST操作

在模拟登陆时，往往会用到POST，会传递对应的POST数据

此处，主要有两种方式传递POST数据：

- postDict  
一般都是通过postDict传递数据进去

然后内部通过quoteParas转换为对应的post data，是以"&"为分隔符的。

- postDataStr  
个别情况下，特殊的情况下，会用到此postDataStr

其传递的post数据，是以换行为分隔符的。此时需要，不设置postDict（默认为null），然后设置对应的postDataStr即可。

下面，针对两种情况，都给出对应的多个示例来说明如何使用：

#### 9.6.3.6.1. postDict示例：getDomainPageRank

比如，之前折腾：[第 9.11 节 “查找获得域名的Page Rank:getDomainPageRank”](#) 时所用到的：

```
//Method 1: use http://www.pagerankme.com/
queryUrl = "http://www.pagerankme.com/";
postDict = new Dictionary<string, string>();
postDict.Add("url", domainUrl);
respHtml = getUrlRespHtml(queryUrl, postDict: postDict);
```

#### 9.6.3.6.2. postDict示例：downloadSongtasteMusic

比如，之前折腾：[DownloadSongtasteMusic](#)<sup>14</sup>时所用到的：

```
const string stHtmlCharset = "GB18030";

Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict.Add("x-requested-with", "XMLHttpRequest");
// when click play
// access http://songtaste.com/time.php, post data:
//
str=5bf271ccad05f95186be764f725e9aaf07e0c7791a89123a9addb2a239179e64c91834c698a9c5d82f1ced3fe
Dictionary<string, string> postDict = new Dictionary<string, string>();
postDict.Add("str", str);
postDict.Add("sid", sid);
postDict.Add("t", "0");
string getRealAddrUrl = "http://songtaste.com/time.php";
songInfo.realAddr = crl.getUrlRespHtml(getRealAddrUrl, headerDict:headerDict,
postDict:postDict, charset:stHtmlCharset);
```

<sup>14</sup> [http://www.crifan.com/crifan\\_released\\_all/website/dotnet/downloadsongtastemusic/](http://www.crifan.com/crifan_released_all/website/dotnet/downloadsongtastemusic/)

### 9.6.3.6.3. postDataStr示例：百度API上传文件

比如，之前折腾：[【未解决】通过百度API上传单个文件出现403的错误](#)<sup>15</sup>时所遇到的就是，post数据是以换行符非分隔符的，所以就要去直接设置对应的postDataStr：

```
string[] token = respTokenJson.Split(',');

string tokenStr = token[2].Split(':')[1].Trim("");

byte[] fileBytes = null;
string filename = "fileForUpload2.txt";
string fullPath = @"d:\\" + filename;
using (FileStream fs = new FileStream(fullFilePath, FileMode.Open))
{
    fileBytes = new byte[fs.Length];
    fs.Read(fileBytes, 0, fileBytes.Length);
}

StringBuilder buffer = new StringBuilder();
char[] fileCh = new char[fileBytes.Length];
for (int i = 0; i < fileBytes.Length; i++)
    fileCh[i] = (char)fileBytes[i];

buffer.Append(fileCh);
//postDict = new Dictionary<string, string>();
//postDict.Add("file", buffer.ToString());

string postDataStr = buffer.ToString();

string uploadSingleFileUrl = "https://pcs.baidu.com/rest/2.0/pcs/file?";
Dictionary<string, string> queryParaDict = new Dictionary<string, string>();
queryParaDict.Add("method", "upload");
queryParaDict.Add("access_token", tokenStr);
queryParaDict.Add("path", "/apps/测试应用/" + filename);
uploadSingleFileUrl += crifanLib.quoteParas(queryParaDict);

curCookies = crifanLib.getCurCookies();
newCookies = new CookieCollection();
foreach (Cookie ck in curCookies)
{
    if (ck.Name == "BAIDUID" || ck.Name == "BDUSS")
    {
        ck.Domain = "pcs.baidu.com";
    }

    newCookies.Add(ck);
}
crifanLib.setCurCookies(newCookies);

string boundaryValue = "----WebKitFormBoundaryS0JIa4uHF7yHd8xJ";
string boundaryExpression = "boundary=" + boundaryValue;
```

<sup>15</sup> [http://www.crifan.com/baidu\\_api\\_single\\_file\\_upload\\_403\\_forbidden/](http://www.crifan.com/baidu_api_single_file_upload_403_forbidden/)

```

headerDict = new Dictionary<string, string>();
headerDict.Add("Pragma", "no-cache");
headerDict.Add("Content-Type", "multipart/form-data;" + " " + boundaryExpression);

postDataStr = boundaryValue + "\r\n"
    + "Content-Disposition: form-data; name=\"file\"" + "\r\n"
    + postDataStr + "\r\n"
    + boundaryValue;

//string str = crifanLib.getUrlRespHtml(
//    string.Format(@"https://pcs.baidu.com/rest/2.0/pcs/file?method=upload&path=%2Fapps%2F%E6%B5%8B%E8%AF%95%E5%BA%94%E7%94%A8%2F78.jpg&access_token={0}",
//        tokenStr),
//    headerDict, postDict);
string respJson = crifanLib.getUrlRespHtml(uploadSingleFileUrl, headerDict:headerDict,
    postDataStr: postDataStr);

```

#### 9.6.3.6.4. postDataStr示例：网易的心情随笔

比如，之前折腾：[【记录】给BlogsToWordPress添加支持导出网易的心情随笔](#)<sup>16</sup>时所遇到的就是，post数据是以换行符非分隔符的，所以就要去直接设置对应的postDataStr：

```

string postDataStr =
    "callCount=1" + "\r\n" +
    "scriptSessionId=${scriptSessionId}187" + "\r\n" +
    "c0-scriptName=BlogBeanNew" + "\r\n" +
    "c0-methodName=getBlogs" + "\r\n" +
    "c0-id=0" + "\r\n" +
    "c0-param0=" + "number:" + userId + "\r\n" +
    "c0-param1=" + "number:" + startBlogIdx + "\r\n" +
    "c0-param2=" + "number:" + onceGetNum;

//http://api.blog.163.com/ni_chen/dwr/call/plaincall/BlogBeanNew.getBlogs.dwr
string getBlogsDwrMainUrl = blogApi163 + "/" + blogUser + "/" + "dwr/call/plaincall/"
    + "BlogBeanNew.getBlogs.dwr";

Dictionary<string, string> headerDict = new Dictionary<string, string>();
headerDict = new Dictionary<string, string>();
//Referer http://api.blog.163.com/crossdomain.html?t=20100205
headerDict.Add("Referer", "http://api.blog.163.com/crossdomain.html?t=20100205");
headerDict.Add("Content-Type", "text/plain");

string blogsRespHtml = getUrlRespHtml(getBlogsDwrMainUrl, headerDict:headerDict,
    postDataStr:postDataStr);

```

## 9.7. 多次尝试版本的 getUrlRespHtml:getUrlRespHtml\_multiTry

默认的getUrlRespHtml只允许一次，即当出错时，就返回空字符串了，就不再继续了。

<sup>16</sup> [http://www.crifan.com/blogstowordpress\\_add\\_feeling\\_card\\_for\\_163\\_netease\\_blog/](http://www.crifan.com/blogstowordpress_add_feeling_card_for_163_netease_blog/)

此处的getUrlRespHtml\_multiTry，是带多次尝试的版本。

其完整代码是：

```
public string getUrlRespHtml_multiTry
    (string url,
     Dictionary<string, string> headerDict = defHeaderDict,
     string charset = defCharset,
     Dictionary<string, string> postDict = defPostDict,
     int timeout = defTimeout,
     string postDataStr = defPostDataStr,
     int readWriteTimeout = defReadWriteTimeout,
     int maxTryNum = defMaxTryNum,
     int retryFailSleepTime = defRetryFailSleepTime)
{
    string respHtml = "";

    for (int tryIdx = 0; tryIdx < maxTryNum; tryIdx++)
    {
        respHtml = getUrlRespHtml(url, headerDict, charset, postDict, timeout, postDataStr,
readWriteTimeout);
        if (!string.IsNullOrEmpty(respHtml))
        {
            break;
        }
        else
        {
            //something wrong
            //maybe network is not stable
            //so wait some time, then re-do it
            System.Threading.Thread.Sleep(retryFailSleepTime);
        }
    }

    return respHtml;
}
```

### 9.7.1. getUrlRespHtml\_multiTry 的参数详解

很明显可以看出，此处的getUrlRespHtml\_multiTry的很多参数，和前面介绍的[第 9.6 节 “获得Url地址返回的网页内容:getUrlRespHtml”](#) 非常类似。

此处，还有另外两个参数，需要解释一下：

- maxTryNum  
maxTryNum表示最大（当出错时）重试次数。

maxTryNum默认值是defMaxTryNum

defMaxTryNum的值是5：

```
private const int defMaxTryNum = 5;
```



当你需要，在出错时，重试更多次，则可以修改此参数。

- retryFailSleepTime  
retryFailSleepTime表示在每次出错之后，sleep的时间。

retryFailSleepTime默认值是defRetryFailSleepTime

defRetryFailSleepTime的值是100毫秒：

```
private const int defRetryFailSleepTime = 100; //sleep time in ms when retry fail for
getUrlRespHtml
```

此处，是为了，尽量适应网络不稳定等异常情况，在出错后，sleep一段时间重试，以希望实现，网络不稳定的时候，经过多次尝试，且每次错误后会sleep，达到增大网络访问成功的机会。

### 例 9.8. getUrlRespHtml\_multiTry 的使用范例

```
//respHtml = curl.getUrlRespHtml(viewHtmlUrl);
respHtml = curl.getUrlRespHtml_multiTry(viewHtmlUrl);
```

## 9.8. 获得Url地址所返回的二进制数据流:getUrlRespStreamBytes

```
public int getUrlRespStreamBytes(ref Byte[] respBytesBuf,
    string url,
    Dictionary<string, string> headerDict,
    Dictionary<string, string> postDict,
    int timeout,
    Action<int> funcUpdateProgress)
{
    int realReadoutLen = 0;
    getUrlRespStreamBytes_bw(ref respBytesBuf, url, headerDict, postDict, timeout,
funcUpdateProgress);
    while (bNotCompleted_download)
    {
        System.Windows.Forms.Application.DoEvents();
    }
    realReadoutLen = gRealReadoutLen;

    //clear
    gRealReadoutLen = 0;

    return realReadoutLen;
}
```

}

**例 9.9. getUrlRespStreamBytes 的使用范例**

```

public bool downloadStMusicFile(string musicRealAddr, string fullnameToStore, out string
errStr, Action<int> funcUpdateProgress)
{
    bool downloadOk = false;
    errStr = "未知错误！";

    if (musicRealAddr == null ||
        musicRealAddr == "" ||
        fullnameToStore == null ||
        fullnameToStore == "")
    {
        errStr = "Songtaste歌曲真实的地址无效！";
        return downloadOk;
    }

    Dictionary<string, string> headerDict = new Dictionary<string, string>();
    //headerDict.Add("Referer", "http://songtaste.com/music/");
    headerDict.Add("Referer", "http://songtaste.com/");

    //const int maxMusicFileLen = 100 * 1024 * 1024; // 100M
    const int maxMusicFileLen = 300 * 1024 * 1024; // 300M
    Byte[] binDataBuf = new Byte[maxMusicFileLen];

    int respDataLen = crl.getUrlRespStreamBytes(ref binDataBuf, musicRealAddr,
headerDict, null, 0, funcUpdateProgress);
    if (respDataLen < 0)
    {
        errStr = "无法读取歌曲数据！";
        return downloadOk;
    }
}

```

**9.9. (谷歌) 翻译一段话:translateString**

```

//-----
//translate strToTranslate from fromLanguage to toLanguage
//return the translated string
//return empty string if error
//some frequently used language abbrv:
//Chinese Simplified: zh-CN
//Chinese Traditional: zh-TW
//English:          en
//German:           de
//Japanese:         ja
//Korean:           ko

```

```

//French:          fr
//more can be found at:
//http://code.google.com/intl/ru/apis/language/translate/v2/using_rest.html#language-params
public string translateString(string strToTranslate, string fromLanguage, string toLanguage)
{
    string translatedStr = "";
    string transRetHtml = "";

    ////following refer: http://python.u85.us/viewnews-335.html
    //string googleTranslateUrl = "http://translate.google.cn/translate_t";
    //Dictionary<string, string> postDict = new Dictionary<string, string>();
    //postDict.Add("hl", "zh-CN");
    //postDict.Add("ie", "UTF-8");
    //postDict.Add("text", strToTranslate);
    //postDict.Add("langpair", fromLanguage + "|" + toLanguage);
    //const string googleTransHtmlCharset = "UTF-8";
    //string transRetHtml = getUrlRespHtml(googleTranslateUrl,
    charset:googleTransHtmlCharset, postDict:postDict);

    ////http://translate.google.cn/#zh-CN/en/%E4%BB%96%E4%BB%AC%E6%98%AF
    %E8%BF%99%E6%A0%B7%E8%AF%B4%E7%9A%84
    //string googleTransBaseUrl = "http://translate.google.cn/#";
    //strToTranslate = "他们是这样说的";
    //string encodedStr = HttpUtility.UrlEncode(strToTranslate);
    //string googleTransUrl = googleTransBaseUrl + fromLanguage + "/" + toLanguage + "/"
    + encodedStr;
    //string transRetHtml = getUrlRespHtml(googleTransUrl);

    //http://translate.google.cn/translate_a/t?client=t&text=%E4%BB%96%E4%BB%AC
    %E6%98%AF%E8%BF%99%E6%A0%B7%E8%AF%B4%E7%9A%84&hl=zh-CN&sl=zh-
    CN&tl=en&ie=UTF-8&oe=UTF-8&multires=1&ssel=0&tssel=0&sc=1
    //strToTranslate = "他们是这样说的";
    string encodedStr = HttpUtility.UrlEncode(strToTranslate);
    string googleTransBaseUrl = "http://translate.google.cn/translate_a/t?";
    string googleTransUrl = googleTransBaseUrl;
    googleTransUrl += "&client=" + "t";
    googleTransUrl += "&text=" + encodedStr;
    googleTransUrl += "&hl=" + "zh-CN";
    googleTransUrl += "&sl=" + fromLanguage; // source language
    googleTransUrl += "&tl=" + toLanguage; // to language
    googleTransUrl += "&ie=" + "UTF-8"; // input encode
    googleTransUrl += "&oe=" + "UTF-8"; // output encode

    try
    {
        transRetHtml = getUrlRespHtml_multiTry(googleTransUrl);
        //[[["They say", "他们是这样说的", "", "Tāmen shì zhèyàng shuō de"]], "zh-CN", [[["They",
    [5],0,0,1000,0,1,0],["say",[6],1,0,1000,1,2,0]],["他们是",5,[[["They",1000,0,0],["they are",0,0,0],
    ["they were",0,0,0],["that they are",0,0,0],["they are the",0,0,0]],[[0,3]],"他们是这样说的"],["这样
    说",6,[[["say",1000,1,0],["said",0,1,0],["say so",0,1,0],["says",0,1,0],["say this",0,1,0]],[[3,6]],""],,
    [["zh-CN"]],1]

        if (extractSingleStr(@"\[\\["(.+?)", ".+?", "", transRetHtml, out translatedStr))
        {
            //extrac out:They say

```

```
    }  
  }  
  catch  
  {  
    // if pass some special string, such as "彭德怀", then will occur 500 error  
    // here tmp not process the error, just omit it here  
  }  
  
  return translatedStr;  
}
```

### 例 9.10. translateString 的使用范例

```
string strToTranslate = "他们是这样说的";  
string translatedStr = translateString(strToTranslate, "zh-CN", "en");
```

## 9.10. 将中文翻译为英文:transzhcntoen

```
public string transZhcnToEn(string strToTranslate)  
{  
  return translateString(strToTranslate, "zh-CN", "en");  
}
```

### 例 9.11. transzhcntoen 的使用范例

```
string strToTranslate = "他们是这样说的";  
string translatedEnglishStr = transZhcnToEn(strToTranslate);
```

## 9.11. 查找获得域名的Page Rank:getDomainPageRank

```
//get page rank for some domain url  
//para: http://answers.yahoo.com  
//return: 7  
public int getDomainPageRank(string domainUrl)  
{  
  int pageRank = 0;
```

```

string queryUrl = "";
string respHtml = "";
Dictionary<string, string> postDict = new Dictionary<string, string>();
string rankStr = "";
bool prevMethodFail = true;

if ((pageRank == 0) && prevMethodFail)
{
    //Method 1: use http://www.pagerankme.com/
    queryUrl = "http://www.pagerankme.com/";
    postDict = new Dictionary<string, string>();
    postDict.Add("url", domainUrl);
    respHtml = getUrlRespHtml_multiTry(queryUrl, postDict: postDict);
    //<a href="http://www.pagerankme.com" target="_blank" style="text-decoration:none;color:#000000;">PageRank 7</a>
    rankStr = "";
    if (extractSingleStr(@"<a href=""http://www\pagerankme\.com"" target=""_blank"" style=""\.\.+?"">PageRank (\d+)</a>", respHtml, out rankStr))
    {
        pageRank = Int32.Parse(rankStr);
        prevMethodFail = false;
    }
    else
    {
        prevMethodFail = true;
    }
}

if ((pageRank == 0) && prevMethodFail)
{
    //Method 2: use http://moonsy.com/pagerank_checker/
    //(1) http://moonsy.com/pagerank_checker/
    queryUrl = "http://moonsy.com/pagerank_checker/";
    postDict = new Dictionary<string, string>();
    postDict.Add("domain", domainUrl);
    postDict.Add("Submit", "CHECK");

    respHtml = getUrlRespHtml_multiTry(queryUrl, postDict: postDict);

    //<h3>Your Page Rank: 7/10
    rankStr = "";
    if (extractSingleStr(@"<h3>Your Page Rank.+?(\d+)/10", respHtml, out rankStr))
    {
        pageRank = Int32.Parse(rankStr);
        prevMethodFail = false;
    }
    else
    {
        prevMethodFail = true;
    }
}

if ((pageRank == 0) && prevMethodFail)
{
    //Method 3: use http://pagerank.webmasterhome.cn/
    string noHttpPreDomainUrl = Regex.Replace(domainUrl, "((https)|(http)|(ftp))://", "");

    //http://pagerank.webmasterhome.cn/prLoading.asp?domain=answers.yahoo.com

```

```

string tmpRespHtml = "";
Dictionary<string, string> headerDict;
//(1)to get cookies
string pageRankMainUrl = "http://pagerank.webmasterhome.cn/";
tmpRespHtml = getUrlRespHtml_multiTry(pageRankMainUrl);
//(2)ask page rank
string firstBaseUrl = "http://pagerank.webmasterhome.cn/?domain=";
//http://pagerank.webmasterhome.cn/?domain=answers.yahoo.com
string firstWholeUrl = firstBaseUrl + noHttpPreDomainUrl;
headerDict = new Dictionary<string, string>();
headerDict.Add("referer", pageRankMainUrl);
tmpRespHtml = getUrlRespHtml_multiTry(firstWholeUrl, headerDict: headerDict);

string baseUrl = "http://pagerank.webmasterhome.cn/prLoading.asp?domain=";
//http://pagerank.webmasterhome.cn/prLoading.asp?domain=answers.yahoo.com
queryUrl = baseUrl + noHttpPreDomainUrl;
headerDict = new Dictionary<string, string>();
headerDict.Add("referer", firstWholeUrl);
respHtml = getUrlRespHtml_multiTry(queryUrl, headerDict: headerDict);

//'  (7/10)'
rankStr = "";
if (extractSingleStr(@"\((\d+)/10\)", respHtml, out rankStr))
{
    pageRank = Int32.Parse(rankStr);
    prevMethodFail = false;
}
else
{
    prevMethodFail = true;
}
}

//TODO:
//Google PR (PageRank) Checker
//http://www.searchbliss.com/seo-tools/google-pagerank-checker.php
//tmp is "We're sorry, the Google PR check is currently being repaired."
//future: if Ok, mayby can use it

return pageRank;
}

```

### 例 9.12. getDomainPageRank 的使用范例

```

public struct searchItemInfo
{
    public string title;
    public string googleUrl; // with google appendix
    public string originalUrl;
    public string description;
    //add domain url and rank
    public string domainUrl;
}

```

```

public int pageRank;
public int alexaRank;
};

singleItemInfo.domainUrl = crifanLib.getDomainUrl(singleItemInfo.originalUrl);
singleItemInfo.pageRank = crifanLib.getDomainPageRank(singleItemInfo.domainUrl);
singleItemInfo.alexRank = crifanLib.getDomainAlexaRank(singleItemInfo.domainUrl);

```

## 9.12. 查找获得域名的Alexa Rank:getDomainAlexaRank

```

//get alexa rank for some domain url
//para: http://answers.yahoo.com
//return: 4
public int getDomainAlexaRank(string domainUrl)
{
    int alexaRank = 0;
    string queryUrl = "";
    string respHtml = "";
    Dictionary<string, string> postDict = new Dictionary<string, string>();
    string alexaRankStr = "";
    bool prevMethodFail = true;

    //string noHttpPreDomainUrl = Regex.Replace(domainUrl, "((https)|(http)|(ftp))://", "");

    if ((alexRank == 0) && prevMethodFail)
    {
        //Method 1: use http://www.searchbliss.com/rank.asp
        string mainUrl = "http://www.searchbliss.com/rank.asp";
        respHtml = getUrlRespHtml_multiTry(mainUrl);
        //<input type="hidden" name="RAC" value="EIS">
        string accessCode = "";
        if (extractSingleStr(@"<input\s+type=""hidden""\s+name=""RAC""\s+value=""([A-Z]+)"">", respHtml, out accessCode))
        {
            queryUrl = "http://www.searchbliss.com/rank.asp";
            //AC EIS
            //RAC EIS
            //rank http://hubpages.com
            postDict = new Dictionary<string, string>();
            //postDict.Add("domain", noHttpPreDomainUrl);
            postDict.Add("AC", accessCode);
            postDict.Add("RAC", accessCode);
            postDict.Add("rank", domainUrl);
            respHtml = getUrlRespHtml_multiTry(queryUrl, postDict: postDict);
            //<a href="http://www.alex.com/data/details/main/http://hubpages.com"
            target="_blank">444</a>
            if (extractSingleStr(@"<a\s+href=""http://www\.alex\.com/data/details/main/\.+?"\s+target=""_blank"">(\d+)\s+</a>", respHtml, out alexaRankStr))
            {
                //alexRank = Int32.Parse(alexRankStr);
            }
        }
    }
}

```

```

        if (Int32.TryParse(alexaRankStr, out alexaRank))
        {
            prevMethodFail = false;
        }
        else
        {
            prevMethodFail = true;
        }

        prevMethodFail = false;
    }
    else
    {
        prevMethodFail = true;
    }
}
else
{
    prevMethodFail = true;
}
}

#if USE_HTML_PARSER_HTMLAGILITYPACK
if ((alexaRank == 0) && prevMethodFail)
{
    //Method 2: use http://www.alexa.com/
    string tmpUrl = "http://www.alexa.com";
    //to get cookies
    string tmpRespHtml = getUrlRespHtml_multiTry(tmpUrl);
    //then do work
    queryUrl = "http://www.alexa.com/search";
    //http://www.alexa.com/search?q=crifan.com&r=home_home&p=bigtop
    queryUrl += "?q=" + domainUrl;
    queryUrl += "&r=" + "home_home";
    queryUrl += "&p=" + "bigtop";
    respHtml = getUrlRespHtml_multiTry(queryUrl);

    HtmlAgilityPack.HtmlDocument htmlDoc = htmlToHtmlDoc(respHtml);
    HtmlNode rootHtmlNode = htmlDoc.DocumentNode;

    //<span>
    //
    //<span class="traffic-stat-label">Alexa Traffic Rank:</span>
    //<a href="/siteinfo/yahoo.com#trafficstats">
    //4</a>
    //</span>

    //<span class="traffic-stat-label">Alexa Traffic Rank:</span>
    //<a href="/siteinfo/crifan.com#trafficstats">
    //170,557</a>
    //</span>
    //HtmlNode trafficHtmlNode = rootHtmlNode.SelectSingleNode("//span/
span[@class='traffic-stat-label']/a[@href]");
    //HtmlNode trafficHtmlNode = rootHtmlNode.SelectSingleNode("//span/
span[@class='traffic-stat-label']/a");
    //HtmlNodeCollection trafficHtmlNodes = rootHtmlNode.SelectNodes("//span/
span[@class='traffic-stat-label']");

```



```

    XmlNode trafficXmlNode = rootXmlNode.SelectSingleNode("//span/
span[@class='traffic-stat-label']");
    if ((trafficXmlNode != null) && (trafficXmlNode.InnerText.StartsWith("Alexa Traffic
Rank:")))
    {
        XmlNode parentNode = trafficXmlNode.ParentNode;
        XmlNode aHrefNode = parentNode.SelectSingleNode("./a[@href]");
        string trafficNumberStr = aHrefNode.InnerText;
        alexaRankStr = trafficNumberStr.Trim().Replace(",", "");

        //speical:
        //"No Data"
        //alexaRank = Int32.Parse(alexaRankStr);
        if(Int32.TryParse(alexaRankStr, out alexaRank))
        {
            prevMethodFail = false;
        }
        else
        {
            prevMethodFail = true;
        }
    }
    else
    {
        prevMethodFail = true;
    }
}
#endif

if ((alexaRank == 0) && prevMethodFail)
{
    //Method 3: use http://moonsy.com/alexa_rank/

    //(1) http://moonsy.com/alexa_rank/
    queryUrl = "http://moonsy.com/alexa_rank/";
    postDict = new Dictionary<string, string>();
    //postDict.Add("domain", noHttpPreDomainUrl);
    postDict.Add("domain", domainUrl);
    postDict.Add("Submit", "CHECK");

    respHtml = getUrlRespHtml_multiTry(queryUrl, postDict: postDict);

    //<h2>Alexa Rank of <b>ANSWERS.YAHOO.COM</b> is : <b>4</b></h2>
    alexaRankStr = "";
    if (extractSingleStr(@"<h2>Alexa Rank of.+?is.+?(\d+).+?</h2>", respHtml, out
alexaRankStr))
    {
        //alexaRank = Int32.Parse(alexaRankStr);
        if (Int32.TryParse(alexaRankStr, out alexaRank))
        {
            prevMethodFail = false;
        }
        else
        {
            prevMethodFail = true;
        }
    }

    prevMethodFail = false;
}

```

```
    }
    else
    {
        prevMethodFail = true;
    }
}

//TODO:
//maybe future can use:
//http://www.dakola.com/tools/alexa/

return alexaRank;
}
```

### 例 9.13. getDomainAlexaRank 的使用范例

```
public struct searchItemInfo
{
    public string title;
    public string googleUrl; // with google appendix
    public string originalUrl;
    public string description;
    //add domain url and rank
    public string domainUrl;
    public int pageRank;
    public int alexaRank;
};

singleItemInfo.domainUrl = crifanLib.getDomainUrl(singleItemInfo.originalUrl);
singleItemInfo.pageRank = crifanLib.getDomainPageRank(singleItemInfo.domainUrl);
singleItemInfo.alexaRank = crifanLib.getDomainAlexaRank(singleItemInfo.domainUrl);
```

---

# 第 10 章 crifanLib.cs之File/Folder

## 10.1. 获得当前保存路径:getSaveFolder

调用对应的FolderBrowserDialog控件，得到用户所选的（保存文件的）路径

```
public string getSaveFolder(FolderBrowserDialog fbdSave)
{
    string saveFolderPath = "";
    //string saveFolderPath = System.Environment.CurrentDirectory;
    //fbdSave.SelectedPath = System.Environment.CurrentDirectory;
    DialogResult saveFolderResult = fbdSave.ShowDialog();
    if (saveFolderResult == System.Windows.Forms.DialogResult.OK)
    {
        saveFolderPath = fbdSave.SelectedPath;
    }
    else if (saveFolderResult == System.Windows.Forms.DialogResult.Cancel)
    {
        saveFolderPath = "";
    }
    return saveFolderPath;
}
```

### 例 10.1. getSaveFolder 的使用范例

```
//private System.Windows.Forms.FolderBrowserDialog fbdSaveFolder;
string saveFolderPath = getSaveFolder(fbdSaveFolder);
```

## 10.2. 二进制（字节）数据存为文件:saveBytesToFile

```
//save binary bytes into file
public bool saveBytesToFile(string fileToSave, ref Byte[] bytes, int dataLen, out string errStr)
{
    bool saveOk = false;
    errStr = "未知错误！";

    try
    {
        int bufStartPos = 0;
        int bytesToWrite = dataLen;
```

```
FileStream fs;
fs = File.Create(fileToSave, bytesToWrite);
fs.Write(bytes, bufStartPos, bytesToWrite);
fs.Close();

saveOk = true;
}
catch (Exception ex)
{
    errStr = ex.Message;
}

return saveOk;
}
```

### 例 10.2. saveBytesToFile 的使用范例

```
public bool downloadStMusicFile(string musicRealAddr, string fullnameToStore, out string
errStr, Action<int> funcUpdateProgress)
{
    bool downloadOk = false;
    errStr = "未知错误!";

    if (musicRealAddr == null ||
        musicRealAddr == "" ||
        fullnameToStore == null ||
        fullnameToStore == "")
    {
        errStr = "Songtaste歌曲真实的地址无效!";
        return downloadOk;
    }

    Dictionary<string, string> headerDict = new Dictionary<string, string>();
    //headerDict.Add("Referer", "http://songtaste.com/music/");
    headerDict.Add("Referer", "http://songtaste.com/");

    //const int maxMusicFileLen = 100 * 1024 * 1024; // 100M
    const int maxMusicFileLen = 300 * 1024 * 1024; // 300M
    Byte[] binDataBuf = new Byte[maxMusicFileLen];

    int respDataLen = crl.getUrlRespStreamBytes(ref binDataBuf, musicRealAddr,
headerDict, null, 0, funcUpdateProgress);
    if (respDataLen < 0)
    {
        errStr = "无法读取歌曲数据!";
        return downloadOk;
    }

    if (crl.saveBytesToFile(fullnameToStore, ref binDataBuf, respDataLen, out errStr))
    {
        downloadOk = true;
    }
}
```

## 10.3. (从网络上) 下载文件 (到本地) :downloadFile

```
//download file from url
//makesure destination folder exist before call this function
//input para example:
//http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KC/KC-slate-01-
lg_V401028090_jpg
//download\B007OZNZG0\KC-slate-01-lg_V401028090_jpg
public bool downloadFile(string fileUrl, string fullnameToStore, out string errStr,
Action<int> funcUpdateProgress)
{
    bool downloadOk = false;
    errStr = "未知错误!";

    if ((fileUrl == null) || (fileUrl == ""))
    {
        errStr = "URL地址为空!";
        return downloadOk;
    }

    if ((fullnameToStore == null) || (fullnameToStore == ""))
    {
        errStr = "文件保存路径为空!";
        return downloadOk;
    }

    //const int maxFileLen = 100 * 1024 * 1024; // 100M
    const int maxFileLen = 300 * 1024 * 1024; // 300M
    const int lessMaxFileLen = 100 * 1024 * 1024; // 100M
    Byte[] binDataBuf;
    try
    {
        binDataBuf = new Byte[maxFileLen];
    }
    catch (Exception ex)
    {
        //if no enough memory, then try alloc less
        binDataBuf = new Byte[lessMaxFileLen];
    }

    int respDataLen = getUrlRespStreamBytes(ref binDataBuf, fileUrl, null, null, 0,
funcUpdateProgress);
    if (respDataLen < 0)
    {
        errStr = "无法下载文件数据!";
        return downloadOk;
    }

    if (saveBytesToFile(fullnameToStore, ref binDataBuf, respDataLen, out errStr))
```

```
{
    downloadOk = true;
}

return downloadOk;
}
```

### 例 10.3. downloadFile 的使用范例

```
public void updateProgress(int percentage)
{
    //pgbDownload.Value = percentage;
}

public void downloadPictures(string productUrl, string respHtml, out string[]
picFullnameList)
{
    //.....

    string[] imageUrlList = amazonLib.extractProductImageList(respHtml);
    gLogger.Info("Extracted image url list:");
    if (imageUrlList != null)
    {
        picFullnameList = new string[imageUrlList.Length];
        for (int idx = 0; idx < imageUrlList.Length; idx++)
        {
            string imageUrl = imageUrlList[idx];
            gLogger.Info(String.Format("[{0}]={1}", idx, imageUrl));

            string picFilename = crl.extractFilenameFromUrl(imageUrl);

            string picFullFilename = Path.Combine(picFolderFullPath, picFilename);
            string errorStr = "";
            gLogger.Info(String.Format("Downloading {0} to {1}", imageUrl, picFullFilename));
            crl.downloadFile(imageUrl, picFullFilename, out errorStr, updateProgress);
        }
    }
}
```

## 10.4. 调用资源管理器打开文件夹并选中文件:openFolderAndSelectFile

```
//open folder and select file
public void openFolderAndSelectFile(string fullFilename)
{
    System.Diagnostics.Process.Start("Explorer.exe", "/select," + fullFilename);
}
```

**例 10.4. openFolderAndSelectFile 的使用范例**

```
string outputFilename = txbExpAlertFilename.Text + ".xls";
string fullFilename = Path.Combine(saveFolderPath, outputFilename);
//.....
crifanLib.openFolderAndSelectFile(fullFilename);
```

## 10.5. (调用系统默认程序直接) 打开文件:openFileDirectly

```
//open file/url/...
public void openFileDirectly(string fullFilename)
{
    System.Diagnostics.Process.Start(fullFilename);
}
```

**例 10.5. openFileDirectly 的使用范例**

```
private void btnOpenOutputFolder_Click(object sender, EventArgs e)
{
    if (Directory.Exists(txbOutputFolder.Text))
    {
        cri.openFileDirectly(txbOutputFolder.Text);
    }
}
```

---

# 第 11 章 crifanLib.cs之Screen

## 11.1. 获得当前任务栏的尺寸大小:getCurTaskbarSize

```
// get current taskbar size(width, height), support 4 mode: taskbar bottom/right/up/left
public Size getCurTaskbarSize()
{
    int width = 0, height = 0;

    if ((Screen.PrimaryScreen.Bounds.Width == Screen.PrimaryScreen.WorkingArea.Width)
    &&
        (Screen.PrimaryScreen.WorkingArea.Y == 0))
    {
        //taskbar bottom
        width = Screen.PrimaryScreen.WorkingArea.Width;
        height = Screen.PrimaryScreen.Bounds.Height -
Screen.PrimaryScreen.WorkingArea.Height;
    }
    else if ((Screen.PrimaryScreen.Bounds.Height ==
Screen.PrimaryScreen.WorkingArea.Height) &&
        (Screen.PrimaryScreen.WorkingArea.X == 0))
    {
        //taskbar right
        width = Screen.PrimaryScreen.Bounds.Width -
Screen.PrimaryScreen.WorkingArea.Width;
        height = Screen.PrimaryScreen.WorkingArea.Height;
    }
    else if ((Screen.PrimaryScreen.Bounds.Width ==
Screen.PrimaryScreen.WorkingArea.Width) &&
        (Screen.PrimaryScreen.WorkingArea.Y > 0))
    {
        //taskbar up
        width = Screen.PrimaryScreen.WorkingArea.Width;
        //height = Screen.PrimaryScreen.WorkingArea.Y;
        height = Screen.PrimaryScreen.Bounds.Height -
Screen.PrimaryScreen.WorkingArea.Height;
    }
    else if ((Screen.PrimaryScreen.Bounds.Height ==
Screen.PrimaryScreen.WorkingArea.Height) &&
        (Screen.PrimaryScreen.WorkingArea.X > 0))
    {
        //taskbar left
        width = Screen.PrimaryScreen.Bounds.Width -
Screen.PrimaryScreen.WorkingArea.Width;
        height = Screen.PrimaryScreen.WorkingArea.Height;
    }

    return new Size(width, height);
}
```



**例 11.1. getCurTaskbarSize 的使用范例**

```
Size curTaskbarSize = crl.getCurTaskbarSize();
```

## 11.2. 获得当前任务栏的坐标位置:getCurTaskbarLocation

```
// get current taskbar position(X, Y), support 4 mode: taskbar bottom/right/up/left
public System.Drawing.Point getCurTaskbarLocation()
{
    int xPos = 0, yPos = 0;

    if ((Screen.PrimaryScreen.Bounds.Width == Screen.PrimaryScreen.WorkingArea.Width)
    &&
        (Screen.PrimaryScreen.WorkingArea.Y == 0))
    {
        //taskbar bottom
        xPos = 0;
        yPos = Screen.PrimaryScreen.WorkingArea.Height;
    }
    else if ((Screen.PrimaryScreen.Bounds.Height ==
    Screen.PrimaryScreen.WorkingArea.Height) &&
        (Screen.PrimaryScreen.WorkingArea.X == 0))
    {
        //taskbar right
        xPos = Screen.PrimaryScreen.WorkingArea.Width;
        yPos = 0;
    }
    else if ((Screen.PrimaryScreen.Bounds.Width ==
    Screen.PrimaryScreen.WorkingArea.Width) &&
        (Screen.PrimaryScreen.WorkingArea.Y > 0))
    {
        //taskbar up
        xPos = 0;
        yPos = 0;
    }
    else if ((Screen.PrimaryScreen.Bounds.Height ==
    Screen.PrimaryScreen.WorkingArea.Height) &&
        (Screen.PrimaryScreen.WorkingArea.X > 0))
    {
        //taskbar left
        xPos = 0;
        yPos = 0;
    }

    return new System.Drawing.Point(xPos, yPos);
}
```

### 例 11.2. getCurTaskbarLocation 的使用范例

```
Point curTaskbarLocation = crl.getCurTaskbarLocation();
```

## 11.3. 获得当前屏幕的角落的坐标位置:getCornerLocation

```
// get current right bottom corner position(X, Y), support 4 mode: taskbar bottom/right/up/left
public System.Drawing.Point getCornerLocation(Size windowSize)
{
    int xPos = 0, yPos = 0;

    if ((Screen.PrimaryScreen.Bounds.Width == Screen.PrimaryScreen.WorkingArea.Width)
    &&
        (Screen.PrimaryScreen.WorkingArea.Y == 0))
    {
        //taskbar bottom
        xPos = Screen.PrimaryScreen.WorkingArea.Width - windowSize.Width;
        yPos = Screen.PrimaryScreen.WorkingArea.Height - windowSize.Height;
    }
    else if ((Screen.PrimaryScreen.Bounds.Height ==
    Screen.PrimaryScreen.WorkingArea.Height) &&
        (Screen.PrimaryScreen.WorkingArea.X == 0))
    {
        //taskbar right
        xPos = Screen.PrimaryScreen.WorkingArea.Width - windowSize.Width;
        yPos = Screen.PrimaryScreen.WorkingArea.Height - windowSize.Height;
    }
    else if ((Screen.PrimaryScreen.Bounds.Width ==
    Screen.PrimaryScreen.WorkingArea.Width) &&
        (Screen.PrimaryScreen.WorkingArea.Y > 0))
    {
        //taskbar up
        xPos = Screen.PrimaryScreen.WorkingArea.Width - windowSize.Width;
        yPos = Screen.PrimaryScreen.WorkingArea.Y;
    }
    else if ((Screen.PrimaryScreen.Bounds.Height ==
    Screen.PrimaryScreen.WorkingArea.Height) &&
        (Screen.PrimaryScreen.WorkingArea.X > 0))
    {
        //taskbar left
        xPos = Screen.PrimaryScreen.WorkingArea.X;
        yPos = Screen.PrimaryScreen.WorkingArea.Height - windowSize.Height;
    }
}
```

```
return new System.Drawing.Point(xPos, yPos);  
}
```

### 例 11.3. getCornerLocation 的使用范例

```
this.Location = crl.getCornerLocation(this.Size);
```

---

# 第 12 章 crifanLib.cs之Runtime

## 12.1. 获得当前软件的版本:getCurVerStr

```
public string getCurVerStr()
{
    string curVerStr = "";
    Assembly asm = Assembly.GetExecutingAssembly();
    FileVersionInfo fvi = FileVersionInfo.GetVersionInfo(asm.Location);
    curVerStr = String.Format("{0}.{1}", fvi.ProductMajorPart, fvi.ProductMinorPart);
    return curVerStr;
}
```

### 例 12.1. getCurVerStr 的使用范例

```
//update version string
this.Text += " v" + getCurVerStr();
```

---

# 第 13 章 crifanLib.cs之Html Parse

## 13.1. 将HTML转换为 XmlDocument:htmlToXmlDoc

```
#if USE_HTML_PARSER_SGML
//convert html to XML document
public XmlDocument htmlToXmlDoc(string html)
{
    // setup SgmlReader
    SgmlReader sgmlReader = new SgmlReader();
    sgmlReader.DocType = "HTML";
    sgmlReader.WhitespaceHandling = WhitespaceHandling.All;
    sgmlReader.CaseFolding = Sgml.CaseFolding.ToLower;

    string decodedHtml = HttpUtility.HtmlDecode(html);
    sgmlReader.InputStream = new StringReader(decodedHtml);

    // create document
    XmlDocument xmlDoc = new XmlDocument();
    xmlDoc.PreserveWhitespace = true;
    xmlDoc.XmlResolver = null;
    xmlDoc.Load(sgmlReader);

    return xmlDoc;
}
#endif
```

### 例 13.1. htmlToXmlDoc 的使用范例

```
//(1) with xmlNs
string withXmlNsUrl = "http://fiverr.com/gigs/search?utf8=%E2%9C
%93&query=seo&x=15&y=13&page=2";
string withXmlNsHtml = getUrlRespHtml(withXmlNsUrl);
XmlDocument xmlDocWithNs = htmlToXmlDoc(withXmlNsHtml);
```

另外，贴出，完整的示例代码：

```
//example code for html parse
void _demoHtmlParse()
{
    #if USE_HTML_PARSER_SGML
    //Method 1: use htmlToXmlDoc
    //(1) with xmlNs
```

```

    string withXmlnsUrl = "http://fiverr.com/gigs/search?utf8=%E2%9C
%93&query=seo&x=15&y=13&page=2";
    string withXmlnsHtml = getUrlRespHtml(withXmlnsUrl);
    XmlDocument xmlDocWithNs = htmlToXmlDoc(withXmlnsHtml);
    //<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://
www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
    //<html xmlns:og="http://ogp.me/ns#" xmlns:fb="http://www.facebook.com/2008/fbml"
xmlns="http://www.w3.org/1999/xhtml" lang="en" xml:lang="en" >
    // <head>
    // ...
    XmlNamespaceManager m = new XmlNamespaceManager(xmlDocWithNs.NameTable);
    m.AddNamespace("w3org", "http://www.w3.org/1999/xhtml");
    XmlNode titleNode = xmlDocWithNs.SelectSingleNode("//
w3org:h1[@itemprop='name']", m);
    string title = titleNode.InnerText;

    //(2) without xmlns
    string withoutXmlnsUrl = "http://www.amazon.com/gp/new-releases/appliances/
ref=zg_bsnr_nav_0";
    //<!DOCTYPE html>
    //<html>
    //<head>
    //...
    string withoutXmlnsHtml = getUrlRespHtml(withoutXmlnsUrl);
    XmlDocument xmlDocNoNs = htmlToXmlDoc(withoutXmlnsHtml);
    XmlNodeList pageNodeList = xmlDocNoNs.SelectNodes("//ol[@class='zg_pagination']/
li[@class]");
    #endif

    //common part
    //how to use Attributes
    //XmlNodeList pageNodeList = xmlDoc.SelectNodes("//ol[@class='zg_pagination']/
li[@class]");
    //if (pageNodeList != null)
    //{
    //    for (int pageIndex = 1; pageIndex < pageNodeList.Count; pageIndex++)
    //    {
    //        XmlNode curPageNode = pageNodeList[pageIndex];
    //        //<li class="zg_page" id="zg_page2"><a page="2"
ajaxUrl="http://www.amazon.com/gp/new-releases/appliances/
ref=zg_bsnr_appliances_pg_2/191-0874592-3518518?ie=UTF8&pg=2&ajax=1"
href="http://www.amazon.com/gp/new-releases/appliances/
ref=zg_bsnr_appliances_pg_2/191-0874592-3518518?ie=UTF8&pg=2">21-40</a></li>
    //        XmlNode ajaxUrlNode = curPageNode.SelectSingleNode("./a[@href]");
    //        string pageUrl = ajaxUrlNode.Attributes["href"].Value;
    //    }
    //}

    #if USE_HTML_PARSER_HTMLAGILITYPACK
    //Method 2: use htmlToHtmlDoc
    string testUrlWithXmlns = "http://sd.csdn.net/";
    string respHtml = getUrlRespHtml(testUrlWithXmlns);

    //<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://
www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
    //<html xmlns="http://www.w3.org/1999/xhtml">
    //<head>

```

```

HtmlAgilityPack.HtmlDocument htmlDoc = htmlToHtmlDoc(respHtml);

//<div class="tabcontent" id="sc1">
// <ul>
// <li><a href="http://www.csdn.net/article/tag/%E4%BA%A7%E5%93%81"
target="_blank">产品</a></li>
// .....
// <li><a href="http://www.csdn.net/article/tag/%E8%AE%BE%E8%AE%A1"
target="_blank">设计</a></li>
// </ul>
//</div>
//...
//<div class="tabcontent" id="sc4">
// <ul>
// ...
// <li><a href="http://www.csdn.net/article/tag/%E6%95%B0%E6%8D%AE%E5%BA
%93" target="_blank">数据库</a></li>
// </ul>
//</div>

//here, no need to take care the html xmlns
//is better than SGMLReader
HtmlNode rootHtmlNode = htmlDoc.DocumentNode;
HtmlNodeCollection htmlNodes = rootHtmlNode.SelectNodes("//
div[@class='tabcontent']");
foreach (HtmlNode link in htmlNodes)
{
    HtmlAttribute att = link.Attributes["id"];
    string idHref = att.Value;
}

```

## 13.2. 将HTML转换为HtmlAgilityPack的HtmlDocument:htmlToHtmlDoc

```

public HtmlAgilityPack.HtmlDocument htmlToHtmlDoc(string html)
{
    HtmlAgilityPack.HtmlDocument htmlDoc = new HtmlAgilityPack.HtmlDocument();

    //http://www.crifan.com/
    htmlagilitypack_html_tag_form_option_no_child_via_sibling_get_innertext/
    //make some html tag: form/option, has child
    HtmlNode.ElementsFlags.Remove("form");
    HtmlNode.ElementsFlags.Remove("option");

    htmlDoc.LoadHtml(html);

    return htmlDoc;
}

```

**例 13.2. htmlToHtmlDoc 的使用范例**

```
//Method 2: use htmlToHtmlDoc
string testUrlWithXmlns = "http://sd.csdn.net/";
string respHtml = getUrlRespHtml(testUrlWithXmlns);

//<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://
www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
//<html xmlns="http://www.w3.org/1999/xhtml">
//<head>
HtmlAgilityPack.HtmlDocument htmlDoc = htmlToHtmlDoc(respHtml);
```

注意，使用此函数之前，需要开启对应的宏USE\_HTML\_PARSER\_HTMLAGILITYPACK，以及添加对应的dll库HtmlAgilityPack.dll的引用。

## 13.3. 去除HtmlNode中的子节点:removeSubHtmlNode

```
//remove sub node from current html node
//eg:
//"script"
//for
//<script type="text/javascript">
public HtmlNode removeSubHtmlNode(HtmlNode curHtmlNode, string
subNodeToRemove)
{
    HtmlNode afterRemoved = curHtmlNode;

    ///method 1: fail
    ///foreach (var subNode in afterRemoved.Descendants(subNodeToRemove))
    ///foreach (HtmlNode subNode in afterRemoved.Descendants(subNodeToRemove))
    ///{
    // //An unhandled exception of type 'System.InvalidOperationException' occurred in
mscorlib.dll
    // //Additional information: Collection was modified; enumeration operation may not
execute.

    // //afterRemoved.RemoveChild(subNode);
    // //curHtmlNode.RemoveChild(subNode);
    // subNode.Remove();
    //}

    //method 2: OK
    HtmlNodeCollection foundAllSub = curHtmlNode.SelectNodes(subNodeToRemove);
    if ((foundAllSub != null) && (foundAllSub.Count > 0))
    {
        foreach (HtmlNode subNode in foundAllSub)
        {
            curHtmlNode.RemoveChild(subNode);
        }
    }
}
```



```

    }
}

return afterRemoved;
}

```

### 例 13.3. removeSubHtmlNode 的使用范例

```

HtmlNode curBulletNode = allBulletNodeList[idx];

HtmlNode noJsNode = crl.removeSubHtmlNode(curBulletNode, "script");
HtmlNode noStyleNode = crl.removeSubHtmlNode(curBulletNode, "style");

string bulletStr = noStyleNode.InnerText;

```

## 13.4. 去除HTML的标签tag:htmlRemoveTag

```

/*
 * [Function]
 * remove html tag, retain html content
 * [Input]
 * html, with tag
 *
 * [Output]
 * pure content, no html tag
 *
 * [Note]
 */
public string htmlRemoveTag(string html)
{
    string filteredHtml = "";

    if (!string.IsNullOrEmpty(html))
    {
        HtmlAgilityPack.HtmlDocument htmlDoc = new HtmlAgilityPack.HtmlDocument();
        htmlDoc.LoadHtml(html);
        if (htmlDoc == null)
        {
            return "";
        }

        // 1. remove all comments
        //(1)get all comment nodes using XPath
        HtmlNodeCollection commentNodeList = htmlDoc.DocumentNode.SelectNodes("//comment()");
        if (commentNodeList != null)
        {

```

```
foreach (HtmlNode comment in commentNodeList)
{
    //(2) remove comment node itself
    comment.ParentNode.RemoveChild(comment);
}

//2. get all content
foreach (var node in htmlDoc.DocumentNode.ChildNodes)
{
    filteredHtml += node.InnerText;
}

return filteredHtml;
}
```

### 例 13.4. htmlRemoveTag 的使用范例

```
HtmlAgilityPack.HtmlDocument htmlDoc =
crl.htmlToHtmlDoc(googleSearchRespHtml);
HtmlNodeCollection liNodeList = htmlDoc.DocumentNode.SelectNodes("//
li[@class='g']");
foreach (HtmlNode liNode in liNodeList)
{
    HtmlNode h3ANode = liNode.SelectSingleNode("./h3[@class='r']/a");
    if (h3ANode != null)
    {
        googleSearchResultItem singleResultItem = new googleSearchResultItem();

        //string titleHtml = h3ANode.InnerHtml; //"Amritanandamayi Math to
<em>sponsor charity</em> events - Times Of India"
        string titleHtml = h3ANode.InnerText; //"Amritanandamayi Math to sponsor
charity events - Times Of India"
        string filteredTitle = crl.htmlRemoveTag(titleHtml);
```

# 第 14 章 crifanLib.cs之集成DLL到exe中

## 14.1. 集成DLL到exe中

```
public yourClassname()
{
    //!!! for load embedded dll: (1) register resolve handler
    AppDomain.CurrentDomain.AssemblyResolve += new
    ResolveEventHandler(CurrentDomain_AssemblyResolve);

    InitializeComponent();

    ...
}

//!!! for load embedded dll: (2) implement this handler
System.Reflection.Assembly CurrentDomain_AssemblyResolve(object sender,
ResolveEventArgs args)
{
    string dllName = args.Name.Contains(",") ? args.Name.Substring(0,
args.Name.IndexOf(',')) : args.Name.Replace(".dll", "");

    dllName = dllName.Replace(".", "_");

    if (dllName.EndsWith("_resources")) return null;

    System.Resources.ResourceManager rm = new
    System.Resources.ResourceManager(GetType().Namespace + ".Properties.Resources",
    System.Reflection.Assembly.GetExecutingAssembly());

    byte[] bytes = (byte[])rm.GetObject(dllName);

    return System.Reflection.Assembly.Load(bytes);
}
```

### 例 14.1. 集成DLL到exe中的使用范例

```
public crifanLib()
{
    //!!! for load embedded dll: (1) register resolve handler
    AppDomain.CurrentDomain.AssemblyResolve += new
    ResolveEventHandler(CurrentDomain_AssemblyResolve);

    //.....
}
```

```
//!!! for load embedded dll: (2) implement this handler
System.Reflection.Assembly CurrentDomain_AssemblyResolve(object sender,
ResolveEventArgs args)
{
    string dllName = args.Name.Contains(",") ? args.Name.Substring(0,
args.Name.IndexOf(',')) : args.Name.Replace(".dll", "");

    dllName = dllName.Replace(".", "_");

    if (dllName.EndsWith("_resources")) return null;

    System.Resources.ResourceManager rm = new
System.Resources.ResourceManager(GetType().Namespace + ".Properties.Resources",
System.Reflection.Assembly.GetExecutingAssembly());

    byte[] bytes = (byte[])rm.GetObject(dllName);

    return System.Reflection.Assembly.Load(bytes);
}
```

关于如何把DLL集成到exe中，详见：[【已解决】C#中集成DLL库到自己的exe程序中<sup>1</sup>](#)

<sup>1</sup> [http://www.crifan.com/csharp\\_integrate\\_dll\\_library\\_into\\_exe\\_application/](http://www.crifan.com/csharp_integrate_dll_library_into_exe_application/)

---

## 第 15 章 crifanLib.cs之DataGridView

### 15.1. 清楚DataGridView的内容:dgvClearContent

```
public void dgvClearContent(DataGridView dgvValue)
{
    dgvValue.Rows.Clear();
}
```

#### 例 15.1. dgvClearContent 的使用范例

```
dgvClearContent(dgvSearchedAlerts);
```

### 15.2. 让DataGridView显示行号:dgvDrawHeaderNum

```
//draw the row index
public void dgvDrawHeaderNum(DataGridView dgvValue)
{
    for (int index = 0; (index <= (dgvValue.Rows.Count - 1)); index++)
    {
        int number = index + 1;
        dgvValue.Rows[index].HeaderCell.Value = String.Format("{0}", number);
    }
}
```

#### 例 15.2. dgvDrawHeaderNum 的使用范例

```
dgvDrawHeaderNum(dgvSearchedAlerts);
```

### 15.3. 释放对象 ( 变量 ) :releaseObject

```
//release object
public void releaseObject(object obj)
{
    try
    {
        System.Runtime.InteropServices.Marshal.ReleaseComObject(obj);
        obj = null;
    }
    catch (Exception ex)
    {
        obj = null;
        //MessageBox.Show("Exception Occured while releasing object " + ex.ToString());
    }
    finally
    {
        GC.Collect();
    }
}
```

### 例 15.3. releaseObject 的使用范例

```
xlWorkbook.Close(true, misValue, misValue);
xlApp.Quit();

releaseObject(xlWorkSheet);
releaseObject(xlWorkbook);
releaseObject(xlApp);
```

## 15.4. 导出DataGridView内容到Excel文件:dgvExportToExcel

```
public void dgvExportToExcel( DataGridView dgvValue,
                             string excelFullFilename,
                             bool isAutoFit = true,
                             bool isHeaderBold = true,
                             List<int> omitRowIdxList = null,
                             List<int> omitColumnIdxList = null,
                             List<int> useTagColumnIdxList = null)
{
    Excel.Application xlApp = new Excel.Application();
    Excel.Workbook xlWorkbook;
    Excel.Worksheet xlWorkSheet;

    object misValue = System.Reflection.Missing.Value;
    xlApp = new Excel.ApplicationClass();
```

```
xlWorkBook = xlApp.Workbooks.Add(misValue);
xlWorkSheet = (Excel.Worksheet)xlWorkBook.Worksheets.get_Item(1);

int rowIdx = 0, realRowIdx = 0;
int columnIdx = 0, realColumnIdx = 0;
const int excelRowHeader = 1;
const int excelColumnHeader = 1;

//save header
for (columnIdx = 0, realColumnIdx = 0; columnIdx <= dgvValue.ColumnCount - 1;
columnIdx++)
{
    if ((omitColumnIdxList != null) && omitColumnIdxList.Contains(columnIdx))
    {
        //omit this column
    }
    else
    {
        //excelRowHeader and excelColumnHeader -> jump over the excel buildin row and
column
        xlWorkSheet.Cells[0 + excelRowHeader, realColumnIdx + excelColumnHeader] =
dgvValue.Columns[columnIdx].HeaderText;

        realColumnIdx++;
    }
}

const int excelTitleRow = 1;
//save cells
for (rowIdx = 0, realRowIdx = 0; rowIdx <= dgvValue.RowCount - 1; rowIdx++)
{
    if ((omitRowIdxList != null) && omitRowIdxList.Contains(rowIdx))
    {
        //omit this row
    }
    else
    {
        for (columnIdx = 0, realColumnIdx = 0; columnIdx <= dgvValue.ColumnCount - 1;
columnIdx++)
        {
            if ((omitColumnIdxList != null) && omitColumnIdxList.Contains(columnIdx))
            {
                //omit this column
            }
            else
            {
                //note here use [columnIdx, rowIdx], not [rowIdx, columnIdx]
DataGridViewCell curCell = dgvValue[columnIdx, rowIdx];
if ((useTagColumnIdxList != null) &&
useTagColumnIdxList.Contains(columnIdx))
            {
                xlWorkSheet.Cells[(realRowIdx + excelTitleRow) + excelRowHeader,
realColumnIdx + excelColumnHeader] = curCell.Tag;
            }
            else
            {
```

```

        xlWorkSheet.Cells[(realRowIndex + excelTitleRow) + excelRowHeader,
realColumnIdx + excelColumnHeader] = curCell.Value;
    }

    realColumnIdx++;
}
}

realRowIndex++;
}
}

//formatting
//(1) header to bold
if (isHeaderBold)
{
    Range headerRow = xlWorkSheet.get_Range("1:1", System.Type.Missing);
    headerRow.Font.Bold = true;
}
//(2) auto adjust column width (according to content)
if (isAutoFit)
{
    Range allColumn = xlWorkSheet.Columns;
    allColumn.AutoFit();
}

//output
xlWorkbook.SaveAs( excelFullFilename,
    XlFileFormat.xlWorkbookNormal,
    misValue,
    misValue,
    misValue,
    misValue,
    XlSaveAsAccessMode.xlExclusive,
    XlSaveConflictResolution.xlLocalSessionChanges,
    misValue,
    misValue,
    misValue,
    misValue);
xlWorkbook.Close(true, misValue, misValue);
xlApp.Quit();

releaseObject(xlWorkSheet);
releaseObject(xlWorkbook);
releaseObject(xlApp);
}

```

#### 例 15.4. dgvExportToExcel 的使用范例

```

string outputFilename = txbExpAlertFilename.Text + ".xls";
string fullFilename = Path.Combine(saveFolderPath, outputFilename);

List<int> omitColumnIdxList = new List<int>();
//omit the last column: View page

```



```
omitColumnIdxList.Add(dgvSearchedAlerts.ColumnCount - 1);

crifanLib.dgvExportToExcel(dgvSearchedAlerts, fullFilename, omitColumnIdxList:
omitColumnIdxList);
```

## 15.5. 导出DataGridView内容到CSV文件:dgvExportToCsv

```
public void dgvExportToCsv(DataGridView dgvValue,
    string csvFullFilename,
    string delimiter = ",",
    List<int> omitRowIdxList = null,
    List<int> omitColumnIdxList = null,
    List<int> useTagColumnIdxList = null)
{
    StreamWriter csvStreamWriter = new StreamWriter(csvFullFilename, false,
System.Text.Encoding.UTF8);

    int rowIdx = 0, realRowIdx = 0;
    int columnIdx = 0, realColumnIdx = 0;

    //output header data
    string headerRowStr = "";
    for (columnIdx = 0, realColumnIdx = 0; columnIdx <= dgvValue.ColumnCount - 1;
columnIdx++)
    {
        if ((omitColumnIdxList != null) && omitColumnIdxList.Contains(columnIdx))
        {
            //omit this column
        }
        else
        {
            headerRowStr += dgvValue.Columns[columnIdx].HeaderText + delimiter;

            realColumnIdx++;
        }
    }
    csvStreamWriter.WriteLine(headerRowStr);

    //output rows data
    for (rowIdx = 0, realRowIdx = 0; rowIdx <= dgvValue.RowCount - 1; rowIdx++)
    {
        if ((omitRowIdxList != null) && omitRowIdxList.Contains(rowIdx))
        {
            //omit this row
        }
        else
        {
            string eachRowStr = "";
            for (columnIdx = 0, realColumnIdx = 0; columnIdx <= dgvValue.ColumnCount - 1;
columnIdx++)
```

```
{
    if ((omitColumnIdxList != null) && omitColumnIdxList.Contains(columnIdx))
    {
        //omit this column
    }
    else
    {
        DataGridViewCell curCell = dgvValue[columnIdx, rowIdx];//
dgvValue.Rows[rowIdx].Cells[columnIdx]
        if ((useTagColumnIdxList != null) &&
useTagColumnIdxList.Contains(columnIdx))
        {
            eachRowStr += curCell.Tag + delimiter;
        }
        else
        {
            eachRowStr += curCell.Value + delimiter;
        }

        realColumnIdx++;
    }
}
csvStreamWriter.WriteLine(eachRowStr);

realRowIndex++;
}
}

csvStreamWriter.Close();
}
```

### 例 15.5. dgvExportToCsv 的使用范例

```
string outputFilename = txbExpAlertFilename.Text + ".csv";
string fullFilename = Path.Combine(saveFolderPath, outputFilename);

List<int> omitColumnIdxList = new List<int>();
//omit the last column: View page
omitColumnIdxList.Add(dgvSearchedAlerts.ColumnCount - 1);

crifanLib.dgvExportToCsv(dgvSearchedAlerts, fullFilename, omitColumnIdxList:
omitColumnIdxList);
```

# 第 16 章 crifanLib.cs之JSON

## 16.1. JSON字符串转换为字典变量:jsonToDict

```
#if USE_JSON
/*
 * [Function]
 * convert json string into dictionary object
 * [Input]
 * json string
 * [Output]
 * object, internally is dictionary
 * [Note]
 * 1.you should know the internal structure of the dictionary
 * then converted to specific type of yours
 */
public Object jsonToDict(string jsonStr)
{
    JavaScriptSerializer jsonSerializer = new JavaScriptSerializer() { MaxJsonLength =
int.MaxValue };
    Object dictObj = jsonSerializer.DeserializeObject(jsonStr);

    return dictObj;
}
#endif
```

### 例 16.1. jsonToDict 的使用范例

```
string kibMasJson = "";
string colorImagesJson = "";

if (crl.extractSingleStr(@"window\.kibMAS\s*=\s*(\[.+\?\])\s*;\s*window\.kibConfig\s*=",
productHtml, out kibMasJson, RegexOptions.Singleline))
{
    //2. json to dict
    Object[] dictList = (Object[])crl.jsonToDict(kibMasJson);

    //3. get ["preplayImages"]["L"]
    imageUrlList = new string[dictList.Length];
    crl.emptyStringArray(imageUrlList);

    for (int idx = 0; idx < dictList.Length; idx++)
    {
        Dictionary<string, Object> eachImgDict = (Dictionary<string, Object>)dictList[idx];
        Object imgUrlObj = null;
        if (eachImgDict.ContainsKey("preplayImages"))
        {
            eachImgDict.TryGetValue("preplayImages", out imgUrlObj);
        }
    }
}
```

```
else if (eachImgDict.ContainsKey("imageUrls"))
{
    eachImgDict.TryGetValue("imageUrls", out imgUrlObj);
}

if (imgUrlObj != null)
{
    //"L" : "http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KC/KC-
slate-01-lg_V401028090_jpg",
    //"S" : "http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KC/KC-
slate-01-sm_V401028090_jpg"

    //"L" : "http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KC/KC-
slate-03-lg_V400694812_jpg",
    //"S" : "http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KC/KC-
slate-03-sm_V400694812_jpg",
    //"rich": {
// src: "http://g-ecx.images-amazon.com/images/G/01/misc/untranslatable-
image-id.jpg",
// width: null,
// height: null
//}

//Type curType = imgUrlObj.GetType();
Dictionary<string, Object> imgUrlDict = (Dictionary<string, Object>)imgUrlObj;
Object largeImgUrObj = "";
if (imgUrlDict.TryGetValue("L", out largeImgUrObj))
{
    //[0] "http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KT/KT-
slate-01-lg_V395919237_jpg"
    //[1] "http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KT/KT-
slate-02-lg_V389394532_jpg"
    //[2] "http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KT/KT-
slate-03-lg_V389394535_jpg"
    //[3] "http://g-ecx.images-amazon.com/images/G/01//kindle/dp/2012/KT/KT-
slate-04-lg.jpg"
    //[4] "http://g-ecx.images-amazon.com/images/G/01/kindle/dp/2012/KT/KT-
slate-05-lg_V389394532_jpg"
    imageUrlList[idx] = largeImgUrObj.ToString();
}
else
{
    //something wrong
    //not get all pic
}
}
else
{
    //something wrong
}
}
```

---

# 参考书目

[1] [crifan的的库函数 : crifanLib](#)<sup>1</sup>

[2] [crifan的C#库 : crifanLib.cs](#)<sup>2</sup>

---

<sup>1</sup> [http://www.crifan.com/crifan\\_released\\_all/crifanlib/](http://www.crifan.com/crifan_released_all/crifanlib/)

<sup>2</sup> <http://code.google.com/p/crifanlib/source/browse/trunk/csharp/crifanLib.cs>