



搜索

☐ Web ☒ www.fmddlmyy.cn

使用doxygen为C/C++程序生成中文文档（上）

按照约定的格式注释源代码，用工具处理注释过的源代码产生文档。通过这种方式产生文档至少有以下好处：

- 便于代码和文档保持同步。
- 可以对文档做版本管理。

很多编程语言都有类似的文档工具，例如：Java有javadoc，Ruby有rdoc。对于C/C++程序，我们可以用Doxygen生成文档。本文通过为一个C++程序“谁养鱼”建立文档，介绍了怎样在Windows平台使用Doxygen。

Doxygen比较适合制作API的接口文档，CHM是这类文档的常见格式。最新版本的Doxygen（目前是1.5.2）统一采用UTF-8作为输出文件的编码格式，但微软的CHM编译工具不支持UTF-8，这就为制作中文CHM文档带来麻烦。本文提出了解决这个问题的方法。

1 Doxygen简介

1.1 要做什么

使用Doxygen生成文档，主要是两件事：

1. 写一个配置文件（Doxyfile）。一般用Doxywizard生成后，再手工修改。
2. 按照Doxygen的约定，将代码“文档化”。

然后只要执行命令：

```
doxygen Doxyfile
```

就可以了。输入文件、输出目录、参数等都是在Doxyfile中配置的。

1.2 得到什么

Doxygen的输出格式主要有HTML、LATEX、RTF等：

- Doxygen在输出HTML文档时，可以自动准备用于制作CHM的项目文件（.hhp）、目录文件（.hhc）和索引文件（.hhk）。用HTML Help Workshop中的CHM编译器（hhc.exe）编译后生成CHM文件。
- Doxygen在输出LATEX文档的同时准备了转换到pdf格式的makefile。只要系统安装了合适的TEX工具，就可以从LATEX文档生成pdf文档。
- Doxygen输出的RTF格式，已经针对Word作了优化，可以较好地转换到Word文档。

1.3 需要什么

完成本文的范例需要以下工具：

1. Doxygen的最新版本，可以从[Doxygen的网站](#)下载。
2. Graphviz是一个图形可视化软件。Doxygen使用Graphviz生成各种图形，例如类的继承关系图、合作图，头文件包含关系图等。可以从[Graphviz的网站](#)下载Graphviz的最新版本。Doxygen使用了Graphviz的布局引擎dot，所以在文档中将其称作dot。
3. 为解决中文问题，需要使用Cygwin的iconv程序作编码转换。
4. 为解决中文问题，需要一个命令行的查找替换工具。我选择了[白杨](#)创作的工具fr。

可以从我的主页<http://www.fmddlmyy.cn>下载这些工具：[Doxygen 1.5.2](#)、[Graphviz 2.12](#)、[iconv \(GNU libiconv 1.9\)](#)和[fr 2.1.1.120](#)。

2 CHM格式的中文问题

前面说过：目前，Doxygen统一采用UTF-8作为输出文件的编码格式，但微软的CHM编译工具（hhc.exe）不支持UTF-8。如果直接用hhc.exe编译，中文不能正确显示。解决这个问题的思路很简单：

- 将Doxygen输出的html文件以及CHM的项目文件（.hhp）、目录文件（.hhc）和索引文件（.hhk）全部转换到GBK编码后，再用hhc.exe编译即可。

可以用iconv对文件作编码转换。对于html文件，除了文件内容的编码转换外，还要将

```
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
```

中的“UTF-8”替换成“gb2312”。

2.1 用批处理简化操作

我写了一些批处理文件（.bat）用于简化处理过程，包括：

2.1.1 clean.bat —— 清空以前输出

```
@echo off
echo 清空以前输出
if exist refman.chm del /f /q refman.chm
if exist output\html del /f /q output\html\*. *
if exist output\latex del /f /q output\latex\*. *
if exist output\rtf del /f /q output\rtf\*. *
if exist output del /f /q output\*. *
```

2.1.2 build.bat —— 调用doxygen生成文档

```
@echo off
echo 生成文档
doxygen Doxyfile
```

2.1.3 utf82gbk.bat —— 将指定文件（支持通配符）从utf-8编码转换到gbk编码

```
@echo off
echo 将%1从utf-8编码转换到gbk编码
for %%f in (%1) do copy %%f %%f.utf8
for %%f in (%1) do iconv -c -f utf-8 -t gbk %%f.utf8 > %%f
```

这个批处理文件要求系统当前路径上有iconv.exe。执行iconv时，使用-c参数忽略无法转换的字符。否则如果输入文件包含无法转换的字符，转换会失败。输入文件被备份到加过.utf8后缀的文件。

2.1.4 html-utf82gbk.bat —— 将指定html文件（支持通配符）从utf-8编码转换到gbk编码

```
@echo off
call utf82gbk %1
echo 将%1中的charset从UTF-8改为gb2312
fr %1 -f:charset=UTF-8 -t:charset=gb2312
```

这个批处理文件要求系统当前路径上有iconv.exe和[白杨](#)的fr.exe。

2.1.5 makechm.bat —— 用Doxygen的输出制作chm文件

```
@echo off
echo 将Doxygen输出文件编码从utf-8转换到gbk
set path=%path%;%cd%
```

```
cd output\html
```

```
echo 处理chm输入文件
call utf82gbk.bat index.hhp
call utf82gbk.bat index.hhc
call utf82gbk.bat index.hhk
call html-utf82gbk *.html
```

```
echo 生成chm文件
"C:\Program Files\HTML Help Workshop\hhc.exe" index.hhp
```

```
if exist index.chm copy index.chm ..\..\refman.chm
del /f /q *.chm
cd ..\..
```

这个批处理文件假设系统在目录“C:\Program Files\HTML Help Workshop\”安装了“HTML Help Workshop”。并假设输出目录是Doxyfile所在目录的子目录output。

2.1.6 rebuild.bat —— 重新生成chm文件

```
@echo off
call clean.bat
call build.bat
call makechm.bat
```

2.2 小结

了解DOS命令的朋友应该很容易看懂这些批处理吧。将这些批处理文件放在工作目录（即Doxyfile所在目录）后，每次只要键入rebuild就可以重新生成chm文件。必要时可以单独使用clean、build、makechm命令。utf82gbk和html-utf82gbk命令也可以单独使用。读者可以从我的主页www.fmddlmyy.cn下载这些[批处理文件](#)。

3 创建配置文件

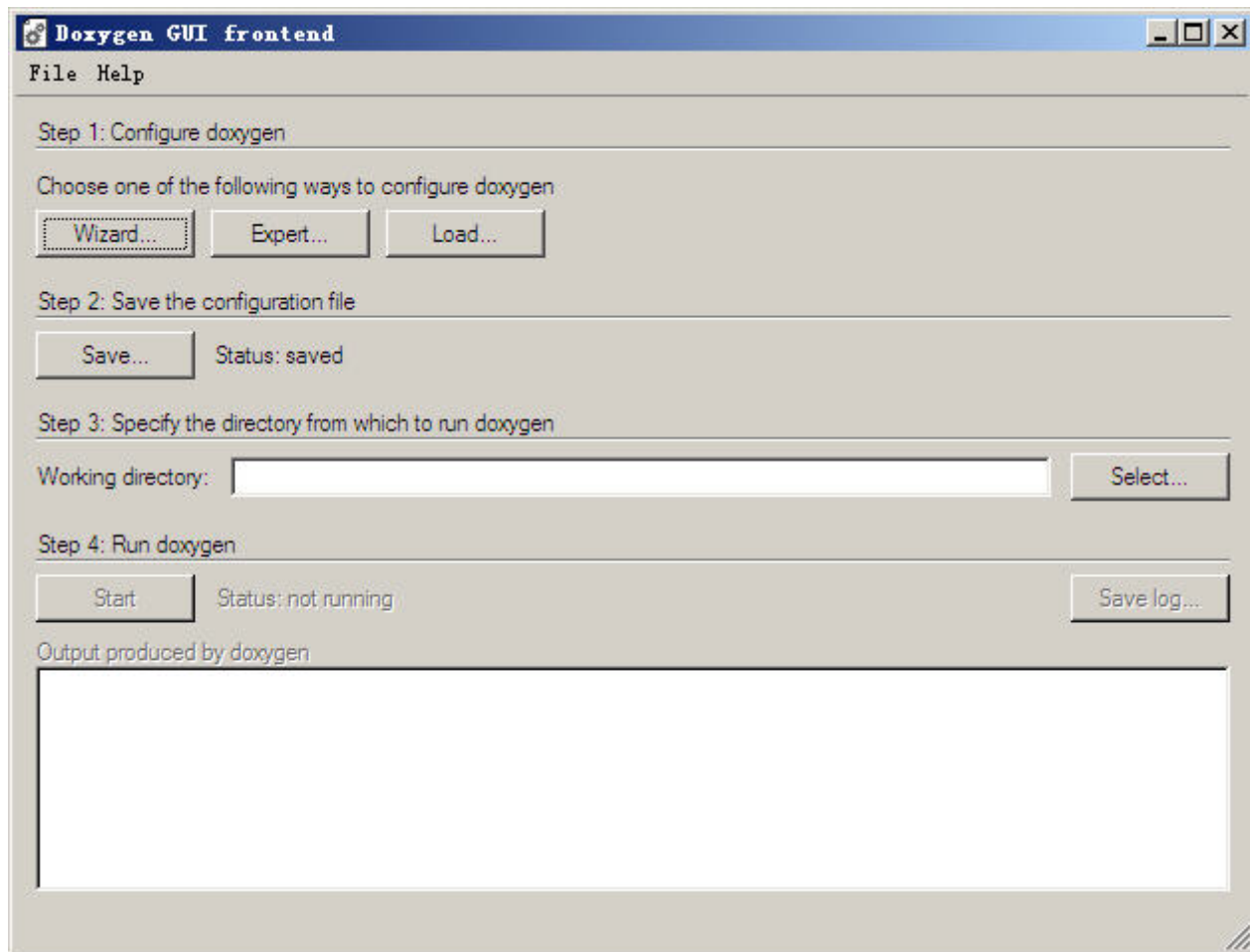
3.1 准备工作

“谁养鱼”是我最近写的一个小程序，它用推理法求解爱因斯坦谜题——“谁养鱼”。读者可从[《穷举和推理：用C++程序求解“谁养鱼”》](#)下载未文档化的程序。

制作文档前，我们要完成以下准备工作：

1. 安装Doxygen、Graphviz和“HTML Help Workshop”。我使用的HTML Help Workshop版本是4.74.8702.0，英文版。网上有汉化版本，但运行时会出错。
2. 将iconv和fr程序放到系统路径包含的目录，例如c:\windows\system32。
3. 建立一个空目录fish，放入要注释的程序（fish\src），建立制作文档的工作目录（fish\doc），将前面介绍的批处理文件放到doc目录。

好了，现在可以运行Doxywizard创建配置文件。

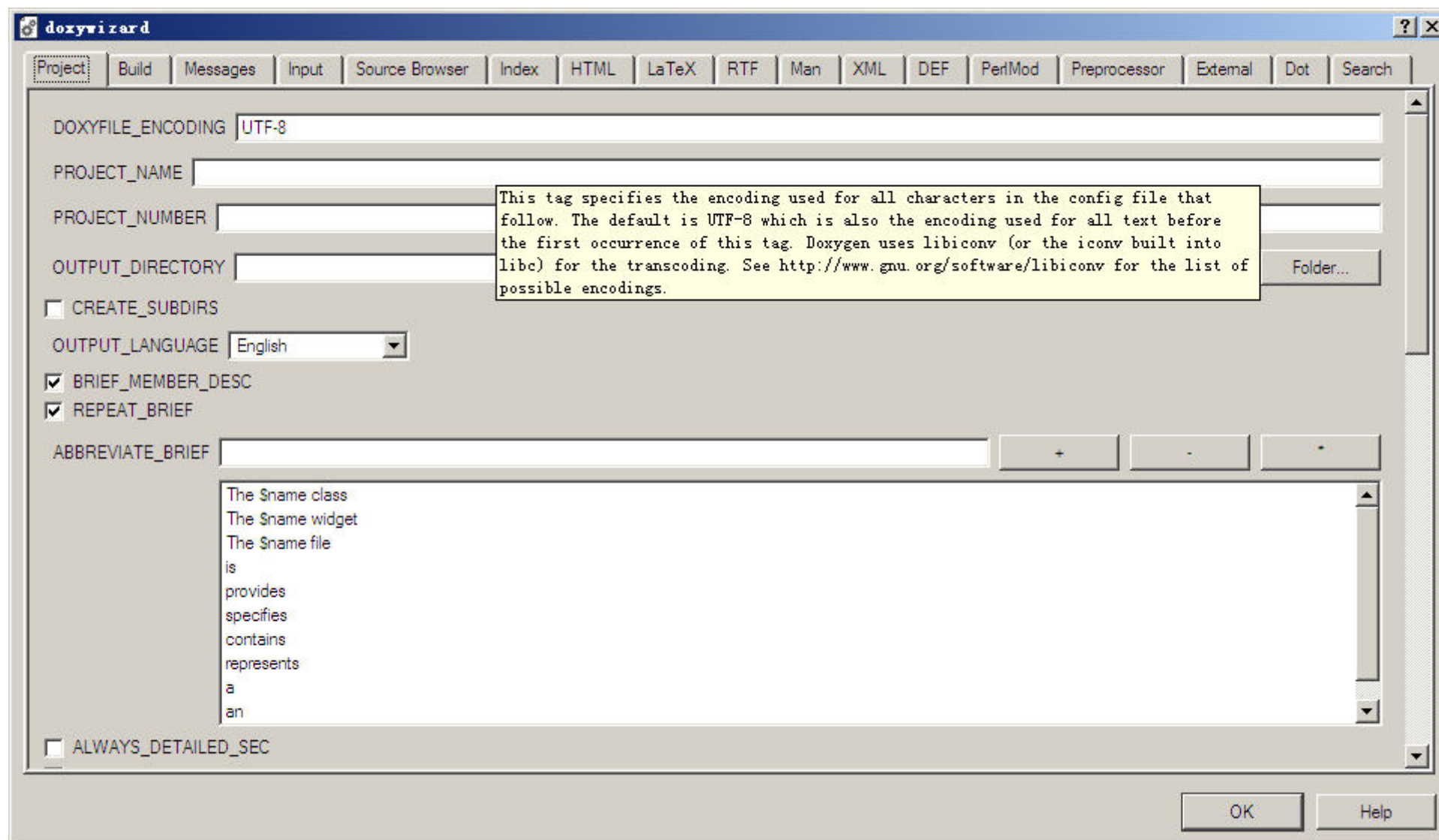


可以直接点“Save...”按钮，将配置保存在doc\Doxyfile。这时，Doxyfile的内容是Doxygen的默认设置。Doxyfile是普通文本文件，我们可以直接打开编辑。不过在Doxywizard的界面上填写也很方便，每个参数都有详细提示。建议用Doxywizard完成第一次设置。以后如果需要调整个别参数，可以

直接编辑Doxyfile。

3.2 填写参数

点“Expert...”按钮后，开始填写配置参数。



:), Doxygen是不是有很多参数？其实大多数参数都有缺省值，需要填写的不算多，下面分页介绍：

3.2.1 Project页

DOXYFILE_ENCODING是Doxyfile的文本编码。如果文件中有中文字符，可以填写GBK。

填写项目名（PROJECT_NAME）、项目版本（PROJECT_NUMBER）、输出目录（OUTPUT_DIRECTORY）和输出语言（OUTPUT_LANGUAGE）。输出目录可以按Doxyfile的相对目录填写。输出语言相当于程序资源，选择Chinese。

Doxywizard的中文支持不完善，中文字符会被存为乱码。我们直接编辑Doxyfile，填写：

PROJECT_NAME = 谁养鱼

取消FULL_PATH_NAMES。我们修改了以下参数：

DOXYFILE_ENCODING	GBK
PROJECT_NAME	谁养鱼
PROJECT_NUMBER	1.0
OUTPUT_DIRECTORY	output
OUTPUT_LANGUAGE	Chinese
FULL_PATH_NAMES	NO

3.2.2 Messages页

在Messages页将WARN_LOGFILE填写为build.log。这样，Doxygen会将编译时出现的警告和错误保存在build.log，我们可以对照修改。

WARN_LOGFILE	build.log
--------------	-----------

3.2.3 Input页

指定输入源文件目录（INPUT），将输入文件编码（INPUT_ENCODING）改为GBK。

INPUT	..\src
INPUT_ENCODING	GBK

FILE_PATTERNS参数是Doxygen要处理的文件类型，缺省值包括Doxygen支持的所有文件类型。不能用Doxygen文档化任意文件类型。例如Doxygen不支持汇编程序。

3.2.4 Source Browser页

选择SOURCE_BROWSER，在文档中包含源代码。

SOURCE_BROWSER	YES
----------------	-----

3.2.5 Html页

选择GENERATE_HTMLHELP后，Doxygen会准备生成chm文件需要的项目文件、目录文件和索引文件。可以通过参数HTML_HEADER和HTML_FOOTER定制页面，参数值是包含定制内容的文件名。例如，我们可以建立文件html_foot，内容为：

```
<p align="right"><A HREF="http://www.fmddlmyy.cn/text20.html" target="_top">穷举和推理：用C++程序求解“谁养鱼”</A></p>
</BODY>
</HTML>
```

然后将HTML_FOOTER的值设为html_foot。

GENERATE_HTMLHELP	YES
HTML_FOOTER	html_foot

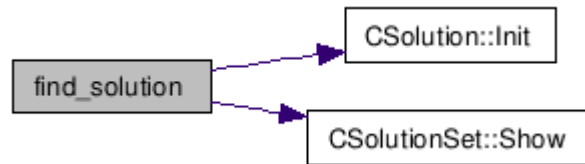
3.2.6 LaTeX页

取消GENERATE_LATEX，不产生LaTeX输出。

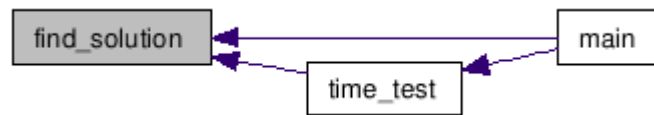
GENERATE_LATEX	NO
----------------	----

3.2.7 Dot页

在Dot页，可以选上UML_LOOK、CALL_GRAPH和CALLER_GRAPH。CALL_GRAPH是本函数调用其它函数的示意图，例如：



CALLER_GRAPH是本函数调用者的示意图，例如：



UML_LOOK	YES
CALL_GRAPH	YES
CALLER_GRAPH	YES

3.3 运行Doxygen

对于“谁养鱼”这个例子，其它参数都可以使用缺省值。从命令行进入doc目录后（参见附录1）运行rebuild.bat，就可以产生refman.chm。这时，我们还没有对程序作任何文档化，输出仅包含Doxygen通过Dot生成的示意图。

我们可以编辑Doxyfile，将EXTRACT_ALL设为YES，再rebuild。这时，Doxygen会自动提取程序的所有要素，包括文件、函数、变量、类型定义、枚举、枚举值、宏定义等。



只想看看结果的朋友可以下载[这个chm文件](#)。Doxygen配置文件中所有参数的详细参考可以查阅Doxygen手册中的“Configuration”页。上篇到此结束。下篇将介绍文档化程序的方法。

后记（2007/7/15）

不少朋友说按照我的说明不能产生期待的结果，这一定是我的文章表述不清了，不好意思。最近，我手头事情比较多，这几个月恐怕没时间写本文的下篇了。好在网上介绍Doxygen文档化的文章还是不少的，少我这篇应该也没什么。

其实，这篇文章的目的只是让不熟悉doxygen的朋友能够快速地了解这个工具。大家如果真的要使用doxygen，还是应该多花些时间看看它的文档。我把本文范例的工作目录打包放在我的主页www.fmddlmyy.cn上，需要的朋友可以[下载](#)。这是个未完成的版本，我去掉了EXTRACT_ALL，注释了几个函数。因为注释不完整，编译（在doc目录rebuild）时还会产生非空的build.log。log是善意的提示，可以帮助我们完善文档。

附录1 快速进入命令行并转到指定目录

如果经常用到命令行，可以在注册表的“HKEY_CLASSES_ROOT\Folder\shell”下建立“Command Prompt Here”项及其子项“command”，将“command”项的默认值设为字符串值“cmd.exe”。这时，只要在资源管理器的任意目录上点击右键并选择“Command Prompt Here”就可以快速进入命令行并转到指定目录。

我将这个注册表项保存成[reg文件](#)：

Windows Registry Editor Version 5.00

```
[HKEY_CLASSES_ROOT\Folder\shell\Command Prompt Here]
```


```
[HKEY_CLASSES_ROOT\Folder\shell\Command Prompt Here\command]  
@="cmd.exe"
```

需要的朋友可以下载后双击导入注册表即可。

在vista上，这样操作不能进入指定目录，也没有必要这样做。在vista中：只要在资源管理器中先按下shift键，再用右键点击文件夹，就会出现“在此处打开命令窗口”的菜单项，选择即可。在左侧的目录树上，这样操作无效。



☐ Web ☒ www.fmddlmyy.cn

[个人主页](#) | [留言本](#) | [我的空间](#) | [我的程序](#) |  fmdd@263.net