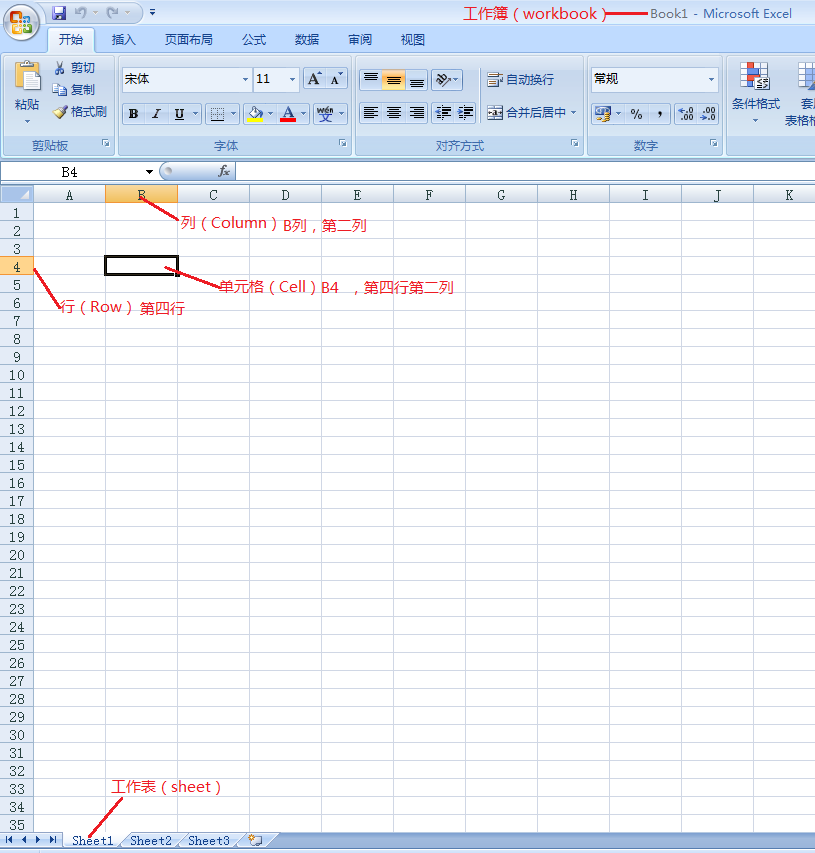
# POI操作Excel

## Excel简介

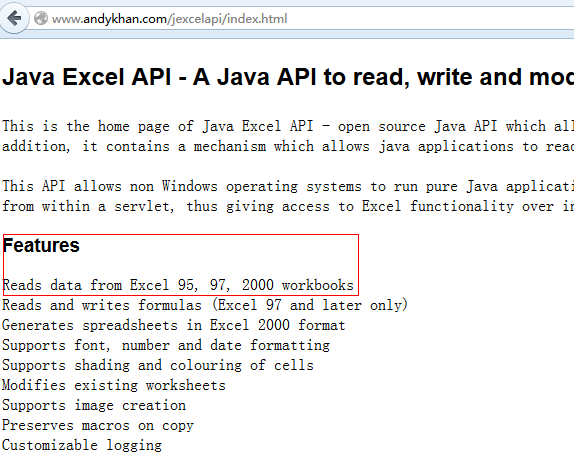
一个excel文件就是一个工作簿workbook，一个工作簿中可以创建多张工作表sheet，而一个工作表中包含多个单元格Cell，这些单元格都是由列（Column）行(Row)组成，列用大写英文字母表示，从A开始到Z共26列，然后再从AA到AZ又26列，再从BA到BZ再26列以此类推。行则使用数字表示，例如；A3 表示第三行第一列，E5表示第五行第五列。



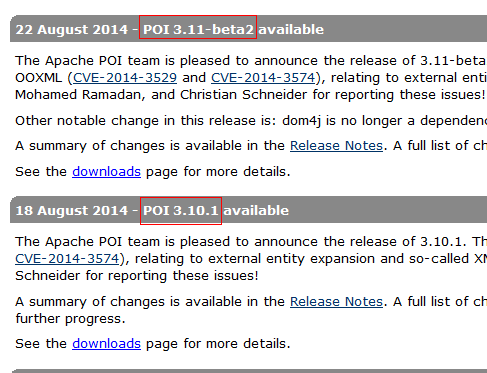
## POI工具包

JAVA中操作Excel的有两种比较主流的工具包： JXL 和 POI 。jxl 只能操作Excel 95, 97, 2000也即以.xls为后缀的excel。而poi可以操作Excel 95及以后的版本，即可操作后缀为 .xls 和 .xlsx两种格式的excel。

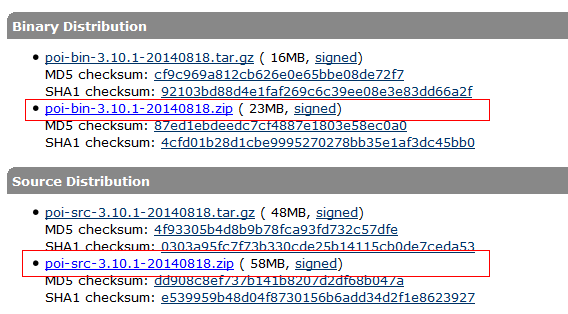
JXL的官网为：<http://www.andykhan.com/jexcelapi>



POI全称 Poor Obfuscation Implementation,直译为“可怜的模糊实现”，利用POI接口可以通过JAVA操作Microsoft office 套件工具的读写功能。官网：<http://poi.apache.org> ，POI支持office的所有版本，并且在接下来的演示中需要从前端页面导入用户上传的版本不确定的excel文件，所以选择POI来讲解。在官网，下载POI ：



在官网中下载jar包以及源码包：



下载完后，打开“poi-bin-3.10.1-20140818”获取操作excel需要的jar包，并将这些jar包复制到项目中。对于只操作2003 及以前版本的excel，只需要poi-3.10.1-20140818.jar ，如果需要同时对2007及以后版本进行操作则需要复制

poi-ooxml-3.10.1-20140818.jar，

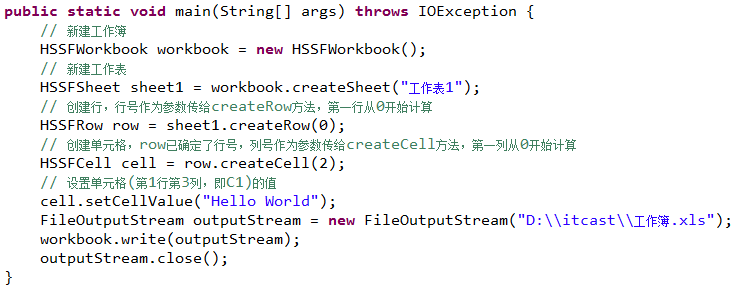
poi-ooxml-schemas-3.10.1-20140818.jar，以及复制在ooxml-lib目录下的xmlbeans-2.6.0.jar，dom4j-1.6.1.jar。

## POI Hello World

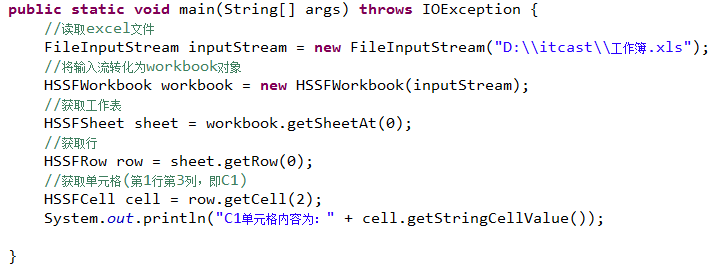
在POI包中有如下几个主要对象和excel的几个对象对应：

|  |  |
| --- | --- |
| HSSFWorkbook | Excel 工作簿workbook |
| HSSFSheet | Excel 工作表 sheet |
| HSSFRow | Excel 行 |
| HSSFCell | Excel 单元格 |

利用以上几个对象，我们简单创建一个Excel工作表，往里面的C1单元格写入和读出“Hello World”：



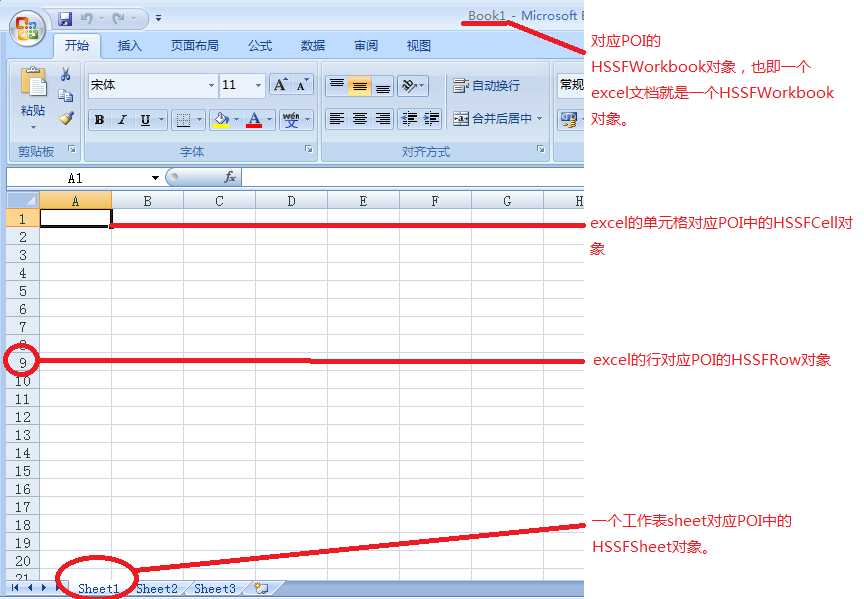
再利用上述几个对象将“Hello World”读取出来：



## POI与Excel

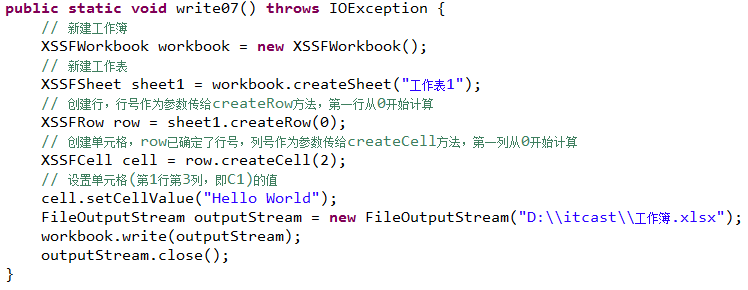
通过上门的例子我们知道如下信息：

1. Excel 的工作簿对应POI的HSSFWorkbook对象；
2. Excel 的工作表对应POI的HSSFSheet对象；
3. Excel 的行对应POI的HSSFRow对象；
4. Excel 的单元格对应POI的HSSFCell对象。

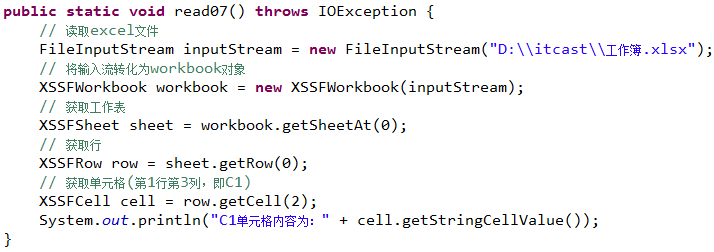


POI 也能对07以后的excel版本进行读写，读写方法和读写03版是一样的，只是对象名称变了；原来各对象的开头字母H变为X，操作方式不变。

1. Excel 的工作簿对应POI的XSSFWorkbook对象；
2. Excel 的工作表对应POI的XSSFSheet对象；
3. Excel 的行对应POI的XSSFRow对象；
4. Excel 的单元格对应POI的XSSFCell对象。

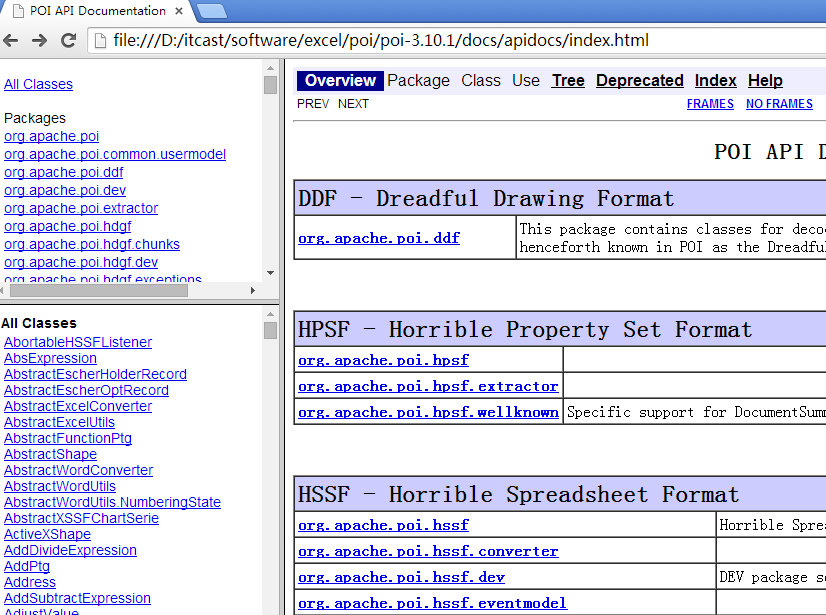


创建07以后版本excel



读取07以后版本excel

在“poi-bin-3.10.1-20140818\poi-3.10.1\docs\apidocs”目录中，点击“index.html”查看POI api文档，我们可以查询POI中所有这些对象的属性和方法。



从api文档中我们了解到：

HSSFWorkbook 和 XSSFWorkbook 都实现了Workbook接口；

HSSFSheet 和 XSSFSheet 实现了Sheet接口；

HSSFRow 和 XSSFRow 实现了Row接口；

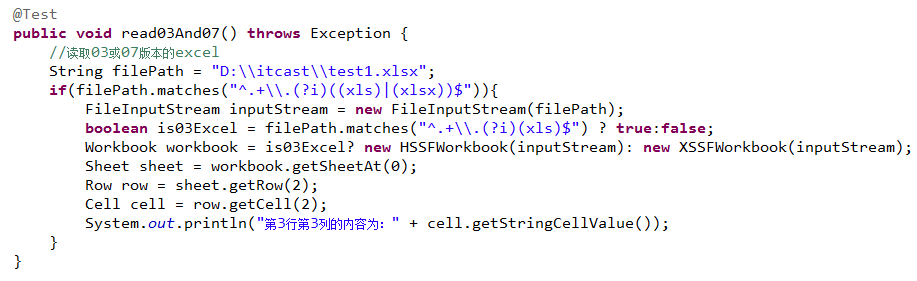
HSSFCell 和 XSSFCell 实现了Cell接口；

因为这两类处理对象共同实现了对应的同一接口，届时将大大方便和简化了同时处理不同格式的excel文件的编码工作。如；在处理03和07版本的excel文件时利用统一的接口就可以做到分析两个版本的excel数据。

POI同时读入03和07版本的excel。

方法一：判断文件的名称后调用对应版本的读入方法。

方法二：根据不同版本用Workbook接口来读入文件并统一处理。

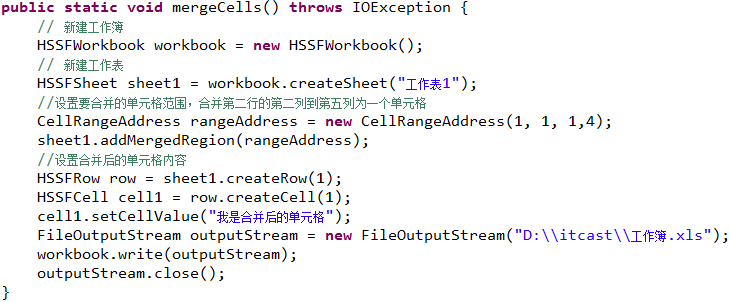


## 格式化Excel

在POI中可以利用格式化对象来格式化excel文档；也即设置excel内容的样式。POI中主要的格式化对象常用的有合并单元格、设置单元格字体、边框，背景颜色等。

**合并单元格**

在POI中有一个CellRangeAddress对象,中文直译是 单元格范围地址，主要用于在单元格的合并上，这个对象的构造方法[**CellRangeAddress**](file:///D:\itcast\software\excel\poi\poi-3.10.1\docs\apidocs\org\apache\poi\hssf\util\CellRangeAddress.html#CellRangeAddress(int, int, int, int))(int firstRow, int lastRow, int firstCol, int lastCol) 有4个参数，分别表示（起始行号，终止行号， 起始列号，终止列号）, 设置这个对象中要合并的单元格范围后，工作表对象sheet调用方法[**addMergedRegion**](file:///D:\itcast\software\excel\poi\poi-3.10.1\docs\apidocs\org\apache\poi\hssf\usermodel\HSSFSheet.html#addMergedRegion(org.apache.poi.ss.util.CellRangeAddress))([CellRangeAddress](file:///D:\itcast\software\excel\poi\poi-3.10.1\docs\apidocs\org\apache\poi\ss\util\CellRangeAddress.html) region) ，将上述设置的CellRangeAddress对象作为参数传入即可合并单元格。

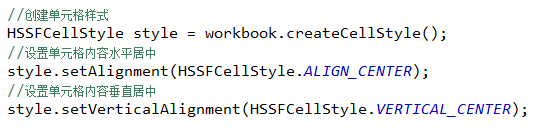


【注意：上图中合并单元格后，单元格的名称是第一个单元格；即上面中合并了第二行的第二列到第五列，合并后的单元格叫B2，而其它被合并的单元格已经无效了，不能对无效单元格设置值。如果进行了设置将不显示。】

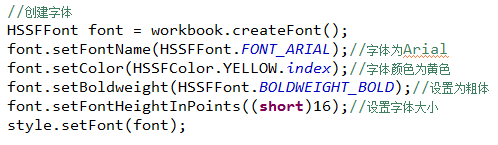
**设置单元格样式**

首先要设置单元格样式则要先初始化POI中的单元格样式对象HSSFCellStyle，然后在样式对象中设置不同的样式（内容位置、字体、背景、颜色、边框等）。单元格样式是由工作簿workbook创建的，一个工作簿可以创建多个样式。

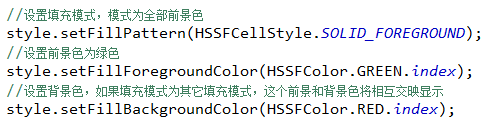
1. 设置单元格内容位置；设置水平位置 [**setAlignment**](file:///D:\itcast\software\excel\poi\poi-bin-3.10.1-20140818\poi-3.10.1\docs\apidocs\org\apache\poi\hssf\usermodel\HSSFCellStyle.html#setAlignment(short))(short align) ，设置垂直位置[**setVerticalAlignment**](file:///D:\itcast\software\excel\poi\poi-bin-3.10.1-20140818\poi-3.10.1\docs\apidocs\org\apache\poi\hssf\usermodel\HSSFCellStyle.html#setVerticalAlignment(short))(short align)



1. 设置单元格字体；POI中的字体对象为HSSFFont，字体是由工作簿创建，可以用于多个单元格上。



1. 设置单元格背景色。





## POI应用案例

### 项目环境

Web框架：SSH(struts2+spring+hibernate)

数据库：mysql-5.6.20

JDK版本：jdk1.6

MyEclipse版本：MyEclipse 2013

应用服务器：apache-tomcat-6.x

### 导出数据到Excel

【需求】在项目数据库中，有一张用户表user，里面有系统的用户数据；现在需要将这个用户表的数据导出到excel中提供用户下载。要求excel要有个一个主标题，各个列名称。

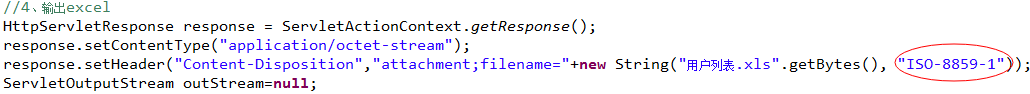


【分析】① 从数据库中读取用户数据；② 创建excel并写入数据； ③ 将excel输出到浏览器提供用户下载。

【编码】利用POI中的HSSFWorkbook创建一个excel，再利用 HSSFSheet 创建一个sheet，HSSFRow/HSSFCell创建行/列。创建的excel第一行标题（CellRangeAddress合并单元格），第二行列名，第三行开始是具体用户数据。



【注意】在输出excel时，要设定要好respone的头部ContentType为application/octet-stream或者application/x-excel，Header的Content-Disposition要正确设置下载文件名称编码，以便可以在各大浏览器中都能够正确输出中文的下载文件名称。



### 导入Excel数据

【需求】将一张有用户数据的Excel导入到数据库中并展示在页面。

【分析】① 获取excel数据；② 解析excel 数据；③ 插入数据到数据库。

【编码】使用正则表达式【^.+\\.(?i)((xls)|(xlsx))$】判断用户上传的excel是什么版本，使用HSSFWorkbook来读取2003以前版本，使用XSSFWorkbook来读取2007及以后版本。逐个遍历sheet，row，cell获取每个单元格对应数据并将数据转换为用户对象插入数据库中。当导入完毕后返回列表页面。