**字体**

1.字体类型分类？

字体类型有如下表：

|  |  |
| --- | --- |
| 类型 | SubType赋值（描述） |
| Type0 | Type0(一个复合字体从派生的CIDFont字形组合) |
| Type | Type1（使用1型字体定义字体的类型） |
| MMType1（一种多主字体，1型字体的扩展，允许从单个字体生成多种字体样式） |
| Type3 | Type3（使用 PDF 图形操作符的字符串来定义字形的字体） |
| TrueType | TrueType基于 TrueType 字体格式的字体 |
| CIDFont | CIDFontType0（一个 CIDFont，字形基于 1 型 字 体 技 术） |
| CIDDontType2（一个 CIDFont，字形基于 TrueType 型字体技术） |

1. 编码

除了Type3的字体，每个字体的程序有一个内置的编码。在某些情况下，一个

PDF 字体字典可以更改字体的内置编码来匹配生成的文本应用程序。编码字符的

灵活性有以下两个原因：

• 它允许根据现有的各种约定来显示文本编码。例如，微软 Windows 和苹果Mac 操作系统使用不同编码标准的拉丁文字，而且许多应用程序使用自己的专用编码。

• 它允许应用程序指定字符怎样从大的字符集里选中进行编码。一些字符集包括超过 256 个字符，包括连字，重音符号，以及高质量排版或非拉丁书写系统所需的其他符号。在相同的字符设置中不同的编码可以选择不同的子集。

3.简单字体与复杂字体

**简单字体：**

有以下的特征：

 1. 字体的字形由一个从某字符串获得的单字节字符代码确定，该字符串由文本显示运算符给出。可以推出，这些代码索引形成 256 个字形的表格；从代码到字形的映射就称为字体的译码过程。

 2.每个字形有一个单一的指标，包括水平位移或设置，宽度。也就是说，简单的字体只支持横向书写模式。

 3.除了Type0字体，在非标记的 PDF 文件的Type3字体，以及某些标准 Type1字体，每个字体列表中包含一个子列表，字体描述符，包含字体宽度指标以及字体的其他属性。在那些属性中的是一个可选的字体文件流，包含着字体构成程序。

**复杂字体：**

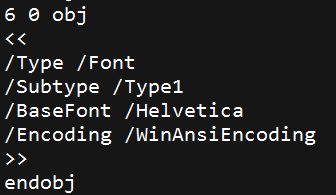
一种复合字体也称为Type0的字体,是一个叫 CIDFont 的一个字体对象中得到字形的字体。一种复合字体由字体字典的子类型值是Type0表示。Type0字体被称为根字体，和其关联的 CIDFont 被称为及其子体。

**区别：**

当前字体是复合时，文本显示运算符与简单的字体是不同的。对于简单字体而言，显示的字符串的每个字节选择一种字形，而为复合字体时，可以解码的一个或多个字节序列从子代 CIDFont 选择一种字形。这项设施支持大量的字符集的使用，如中国、日语以及韩语。复合字体还简化了有扩音器的编码要求的字体的系统。

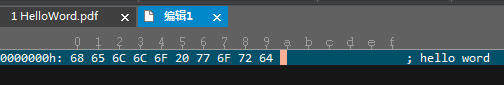
4.pdf中hello world 如何显示出来？

如下图所示：



这里定义了编码规范为WinAnsiEncoding，这种编码表将常用的256个字符和其他特殊字符按照16进制按照字典的顺序进行了重新编码，所以只需要在文本中嵌入相应的16进制数据，使用定义的字典来映射在pdf中显示的数据。

5.为什么pdf中的文档能够复制到其他编译器进行显示？



我们可以看到，将hello word这个单词进行复制的时候，实际上只是复制了相应的二进制数据到相应的txt文件中，其编码格式并没有被一同被复制过去，只需要按照相应的二进制编码格式规范反向生成相应的文字即可。

样例解析：HelloWord.pdf