**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**

编码是信息从一种形式或格式转换为另一种形式的过程也称为计算机编程语言的代码简称编码。

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。

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。

在相同的字符设置中不同的编码可以选择不同的子集。

拉丁文字的字体程序由 Adobe 公司使用的 Adobe 标准编码产生，通常被称为标准编码。

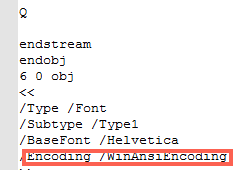
1. **简单字体与复杂字体**

**简单字体：**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，该符串由文本显示运算符给出，每个字形有一个单一的指标，包括水平位移或设置， 宽度。

**复杂字体**一种复合字体也称为类型 0 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。

**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。

**4Pdf文档中如何呈现Hello Word这个单词的（编码解释）**



从上面的helloword.pdf文件结构中可以看出采用的是winansiEncoding，从规范可以得出，如果编码项是 MacRomanEncoding 或 WinAnsiEncoding，该表是在附录 D中描述的初始化映射之一，查看附录D可以得出，hello word 这个单词的每个字符的16进制与表中的数字一一对应，所以通过解码能呈现出正确的字形。

**5复制pdf文档为什么能在其他编译器文档中显示**

复制pdf中的doc内容，在Txt格式中粘贴，默认是清除了字形的所有样式（在word是默认保存复制字体的样式）。首先对pdf内部的编码解析出字符对应相应的编码如helloword.Pdf对应的是WinAnsiEncoding的编码的值，然后在附录D标准查到相应的值转成Unicode对应的值，复制的内容就是转成Unicode的编码值，最后在呈现到其他编译器中。

**样例解析：HelloWord.pdf**

**TrueType简介：**TrueType字体格式是由苹果电脑公司开发的，并已作为微软Windows操作系统的标准字体格式。对TrueType字体文件格式规范苹果TrueType是可用的参考手册和微软的TrueType字体文件技术规范1。TrueType是作为一种对象流嵌套在pdf中的；

**TrueType字体 信息的描述可以和type1描述信息相同，不同的是：**

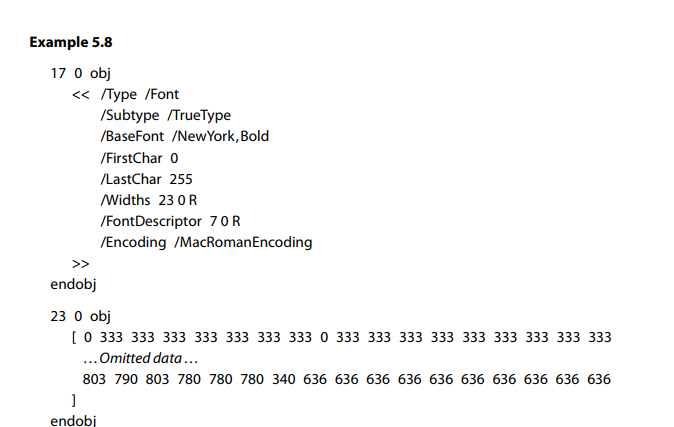
在subFont 的值为trueType 确定是trueType类型

在beseFont 赋值不同 有2中表示

1. 如果使用了trueType表格中的name就优先使用
2. 如果没有就用系统平台下的name来补充 如果系统中没有该类型的样式就会自动合成 如写成NewYork,Bold，找不到粗体类型，操作系统就会合成改样式的风格的字体。

在Ecoding受到约束条件的限制也就是说Differnces这个参数不推荐使用

**结合下面的实例描述绘制TureType的绘制流程：**



首先程序中Tj指令中获取到资源流的信息，得到当前图形状态下的类型是一个Font类型，

在通过SubType获取到字体的类型信息是TrueType类型；TrueType是简单的字体，每个字符对应一个字节；

从encoding编码中得到MacRomanEncoing；这个是预定义的编码模式之一，可以在附录D中查找到相应数值对应的name；

拿到name后，TrueType是字符直接映射到字形的描述，通过cmap表得到一个字符编码的映射，cmap表中有多个子表，优先从（3,1）表查找，如果查找到对应的name在从字形名称通过参考Adobe映射到Unicode值的字形列表找到相应的字形。

找到字形后在通过BaseFont中描述的NewYork，Bold取得字形的样式和类型进行绘制，假设没有找到NewYork对应的加粗字体，就好到当前的操作系统下找到类似的样式进行绘制。