java程序在编译成class文件之后还需要使用dx工具将所有的class文件整合到一个dex文件中，目的是使得其中的各个类都能够共享数据。

从宏观上看dex文件结构由多个不同的数据体拼接到一起，组成dex文件的结构体及其含义如下表所示

|  |  |
| --- | --- |
| 结构体 | 含义 |
| header | dex文件头部，记录整个dex文件的相关属性 |
| string\_ids | 字符串数据索引，记录了每个字符串在数据区的偏移量 |
| type\_ids | 类型数据索引，记录了每个类型的字符串索引 |
| proto\_ids | 原型数据索引，记录了方法声明的字符串，返回类型字符串，参数列表 |
| field\_ids | 字段数据索引，记录了所属类，类型以及方法名 |
| method\_ids | 方法id区，记录方法所属类名，方法声明以及方法名等信息 |
| class\_defs | 类定义区，记录指定类各类信息，包括接口，超类，类数据偏移量 |
| data | 数据区，保存了各个类的真是数据 |
| link\_data | 连接数据区 |

# header结构

header结构体中记录了dex文件的基本信息以及大致的数据分布，其长度固定为0x70，其中每个字段的具体意义如下表所示

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 偏移值 | 长度 | 说明 |
| magic | 0x0 | 8 | 魔数字段，值为"dex\n035\0" |
| checksum | 0x8 | 4 | 校验码 |
| signature | 0xc | 20 | sha-1签名 |
| file\_size | 0x20 | 4 | dex文件总长度 |
| header\_size | 0x24 | 4 | 文件头长度，009版本=0x5c,035版本=0x70 |
| endian\_tag | 0x28 | 4 | 标示字节顺序的常量 |
| link\_size | 0x2c | 4 | 链接段的大小，如果为0就是静态链接 |
| link\_off | 0x30 | 4 | 链接段的开始位置 |
| map\_off | 0x34 | 4 | map数据基址 |
| string\_ids\_size | 0x38 | 4 | 字符串列表中字符串个数 |
| string\_ids\_off | 0x3c | 4 | 字符串列表基址 |
| type\_ids\_size | 0x40 | 4 | 类列表里的类型个数 |
| type\_ids\_off | 0x44 | 4 | 类列表基址 |
| proto\_ids\_size | 0x48 | 4 | 原型列表里面的原型个数 |
| proto\_ids\_off | 0x4c | 4 | 原型列表基址 |
| field\_ids\_size | 0x50 | 4 | 字段个数 |
| field\_ids\_off | 0x54 | 4 | 字段列表基址 |
| method\_ids\_size | 0x58 | 4 | 方法个数 |
| method\_ids\_off | 0x5c | 4 | 方法列表基址 |
| class\_defs\_size | 0x60 | 4 | 类定义标中类的个数 |
| class\_defs\_off | 0x64 | 4 | 类定义列表基址 |
| data\_size | 0x68 | 4 | 数据段的大小，必须4k对齐 |
| data\_off | 0x6c | 4 | 数据段基址 |

# string\_ids结构

该区块是一个偏移量string\_data\_off列表，每个偏移量对应了一个真正的字符串资源，每个偏移量格式为uint占32位。可以通过偏移量找到对应的实际字符串数据。最终这个偏移的位置应该是落在数据区的

# type\_ids结构

区块是一个索引列表，每一项的格式为uint占32位，索引的值对应字符串id区域偏移量列表中的某一项如果我们要找到某个类型的值，需要先根据类型id列表中的索引值去字符串id列表中找到对应的项，这一项存储的偏移量对应的字符串资源就是这个类型的字符串描述。

# proto\_ids结构

方法原型id区是一个方法原型id的列表，列表中每一项的数据格式如下表所示

|  |  |  |
| --- | --- | --- |
| 数据结构 | 格式 | 意义 |
| shorty\_idx | uint | 一个字符串d区的索引，这个索引对应的字符串引列表项中的偏移量存储的字符串是这个方法原型的短格式描述符。 |
| return\_type\_idx | uint | 这个方法原型的返回值类型在类型d列表中索引 |
| parameters\_off | uint | 这个方法原型的参数值列表类型数据的偏移量。0代表没有参数。 |

# field\_ids结构

成员id区存储着原型id列表，其数据结构如下表所示

|  |  |  |
| --- | --- | --- |
| 数据结构 | 格式 | 意义 |
| class\_id | ushort | 这个成员的所在的类在类型id列表中的索引 |
| type\_id | ushort | 这个成员的类型在类型d列表中的索引 |
| name\_idx | uint | 这个成员的名字在字符串id列表的索引 |

# method\_ids结构

方法id区是一个保存方法id的列表，其列表结构如下所示

|  |  |  |
| --- | --- | --- |
| 数据结构 | 格式 | 意义 |
| class\_idx | ushort | 这个方法所在的类在类型d列表中的索引 |
| proto\_idx | ushort | 这个方法的方法原型在方法原型d列表中的索引 |
| name\_idx | uint | 这个方法的名称在字符串d列表中的英碧吴小污 |

# class\_defs结构

类定义区保存的是类定义的列表，其列表结构如下所示

|  |  |  |
| --- | --- | --- |
| 数据结构 | 格式 | 意义 |
| class\_idx | uint | 这个类的类型在类型id列表中的索引 |
| access\_flags | uint | 这个类的访问标记（如：public,final等） |
| superclass\_idx | uint | 父类对应的类型在类型id列表中的索引 |
| interfaces\_off | uint | 这个类使用的接口列表在文件中的偏移量。 |
| source\_file\_idx | uint | 这个类的源码文件的文件名称在字符串资源列表的索引位置 |
| annotations\_off | uint | 这个类的注解数数据在文件中的偏移量 |
| class\_data\_off | uint | 这个类的具体数据在文件中的偏移量 |
| static\_values\_off | uint | 静态成员的初始值列表在文件中的偏移量 |