



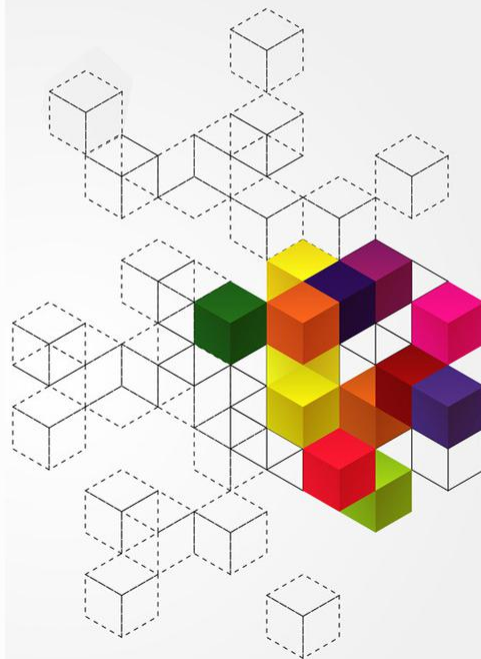
操作系统

Operating system

孔维强

大连理工大学

- 一、 什么是文件
- 二、 文件分类
- 三、 文件逻辑结构
- 四、 文件属性



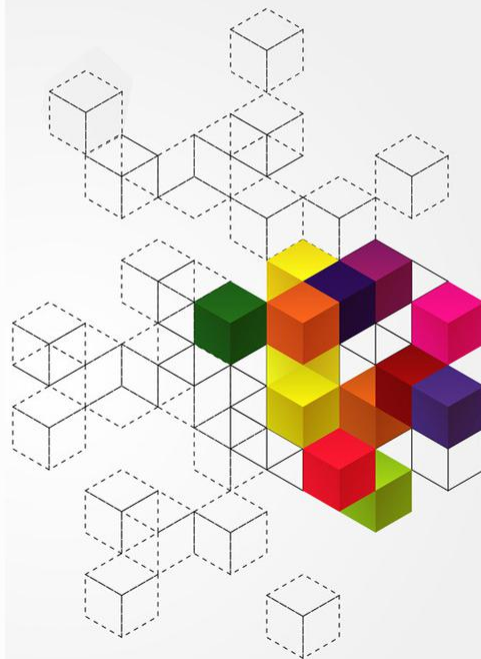
一、什么是文件



- 文件是什么？

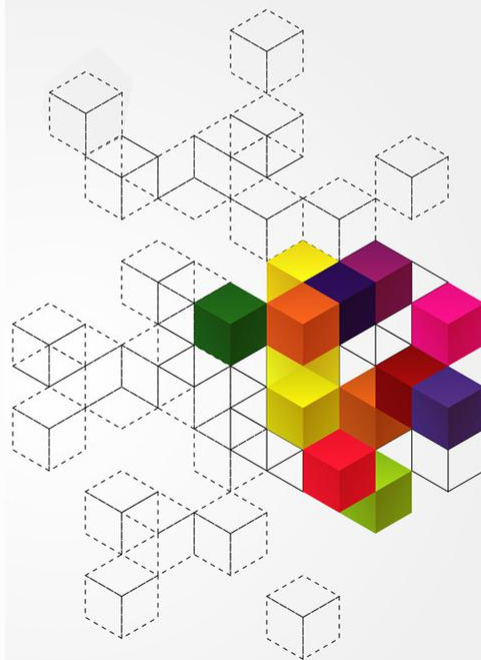
A uniform logical view of information storage

Contiguous logical address space



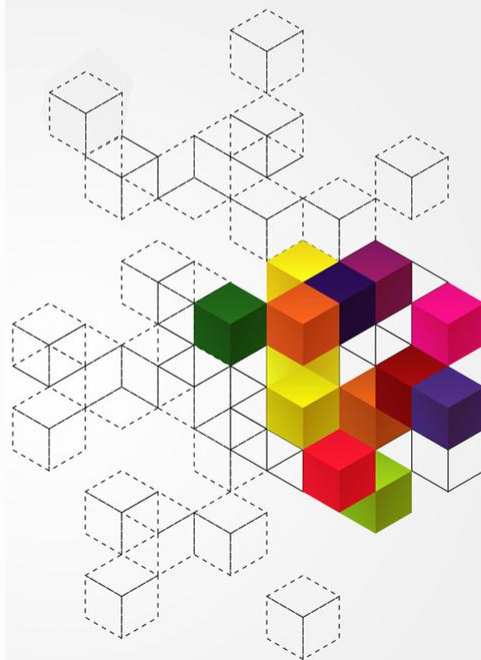
一、什么是文件

- 文件是操作系统在外部持久性存储设备上存储信息的基本单位
- 每个文件代表一段连续的逻辑数据



二、文件分类

file type	usual extension	function
executable	exe, com, bin or none	ready-to-run machine-language program
object	obj, o	compiled, machine language, not linked
source code	c, cc, java, pas, asm, a	source code in various languages
batch	bat, sh	commands to the command interpreter
text	txt, doc	textual data, documents
word processor	wp, tex, rtf, doc	various word-processor formats
library	lib, a, so, dll	libraries of routines for programmers
print or view	ps, pdf, jpg	ASCII or binary file in a format for printing or viewing
archive	arc, zip, tar	related files grouped into one file, sometimes compressed, for archiving or storage
multimedia	mpeg, mov, rm, mp3, avi	binary file containing audio or A/V information



二、文件分类

按内容性质划分：

- 数据文件

.txt

.pdf

.doc

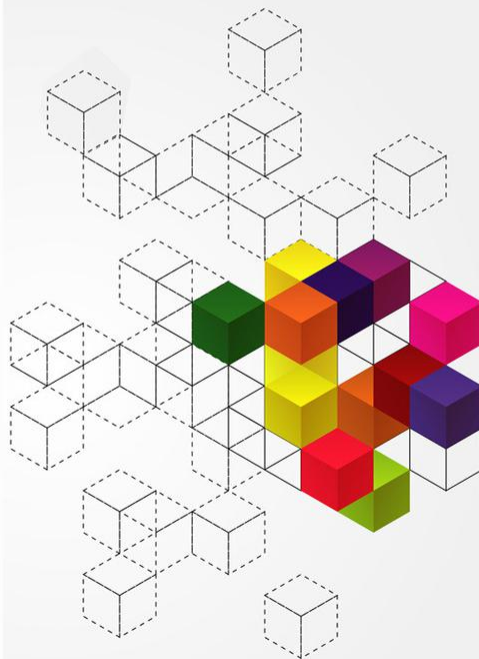
...

- 代码文件

.c, .cpp, .java, .py, ...

.obj, .o

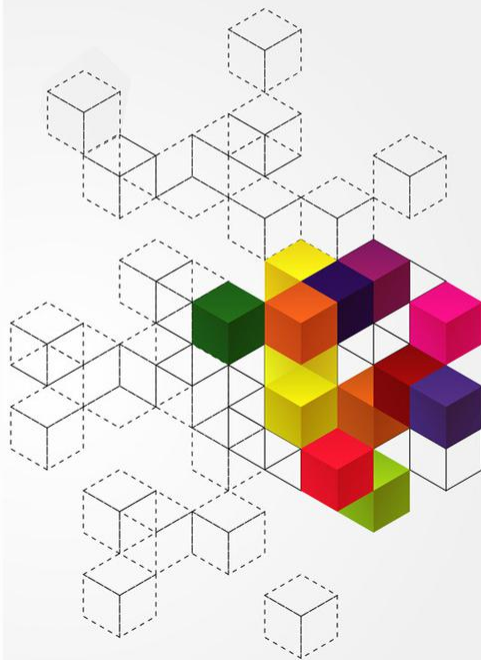
.lib, .dll, .a, .so



二、文件分类

UNIX文件类型

文件类型	说明
1.普通文件	Ordinary or regular files
2.目录文件	a binary file used to track and locate other files and directories.
3.特殊文件（设备文件）	used for device I/O on UNIX and Linux systems
4.链接	a tool used for having multiple filenames that reference a single file on a physical disk
5.管道文件	tools that allow two or more system processes to communicate with each other
6.套接字	tools used for interprocess communication via network.



三、文件逻辑结构

无结构

顺序文件

记录按时间或关键字
字典序顺序排列

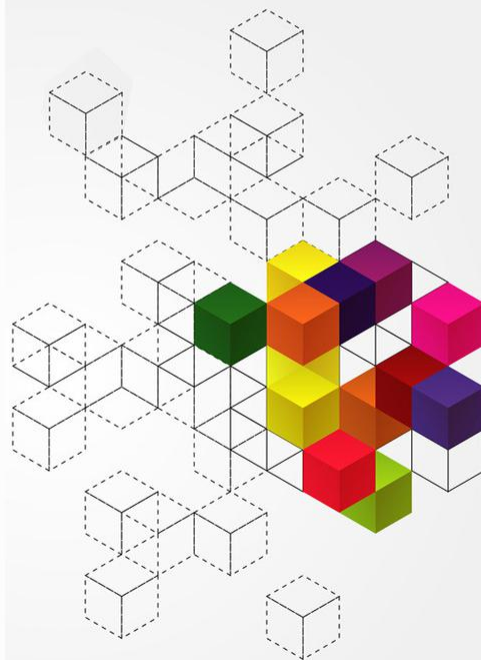
有结构

索引文件

为记录建立索引表

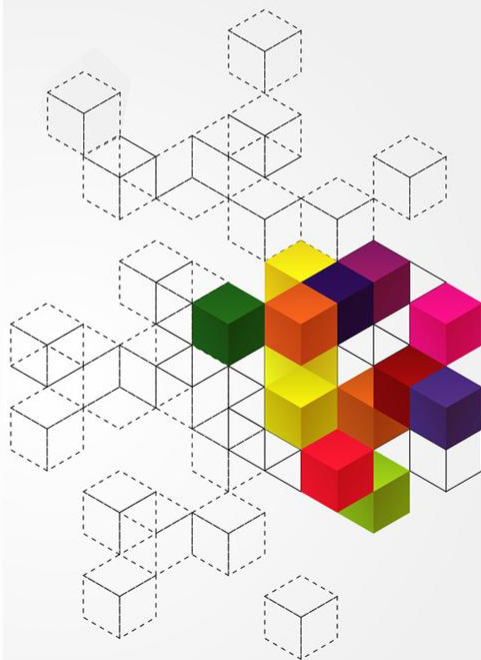
索引顺序文件

记录成组组织，为每
个记录组建立索引



三、文件逻辑结构

- The design of logical structure of the file, is mainly about the organization of file content according to user requirements
- 文件逻辑结构的分类
 - 无结构文件（或称流式文件）
 - 有结构文件（或称记录式文件）



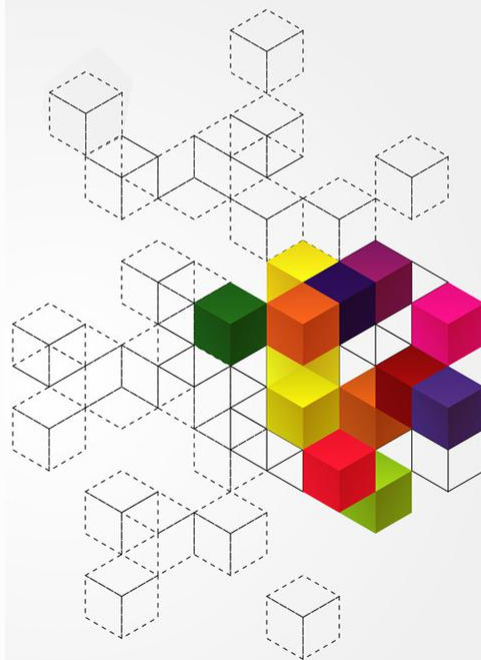
三、文件逻辑结构

- 无结构文件（字符流）

- 无结构文件可以看成是一个字节流，其文件元素为一个“字符”或“字节”，有时称为字节流文件，或流文件。
- 适用于源程序、目标代码等；

- 有结构文件（记录式）

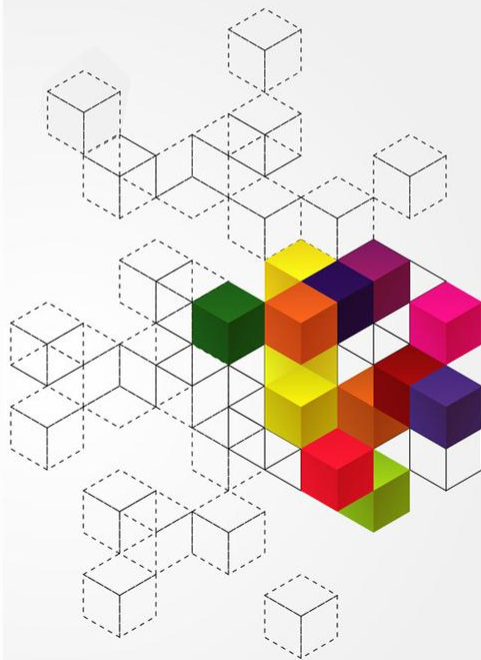
- 有结构文件的文件元素是一条记录，文件是由若干条相关记录组成。
 - 记录式文件主要用于信息管理，例如数据库系统等。
-
- 字符流式的无结构文件实质上是记录长度为一个字符的连续结构文件。



三、文件逻辑结构

字符流文件

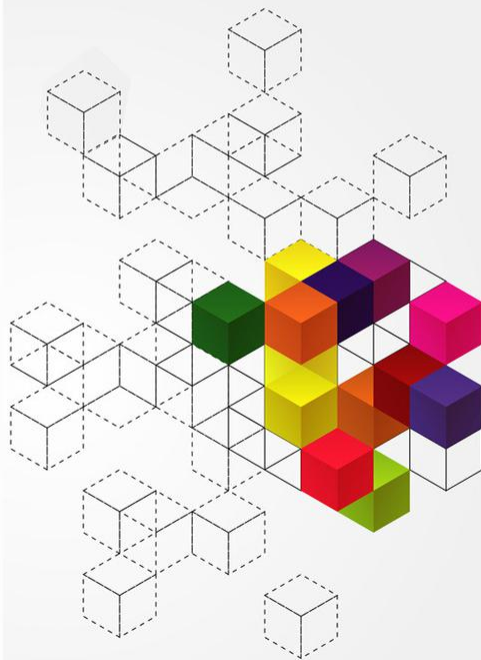
- **优点:** 管理简单, 用户可以方便地对其进行操作。
- **缺点:** 查找文件中的基本信息单位, 例如某个单词, 比较困难的。
- **适用范围:** 适用于对基本信息单位操作不多的文件, 例如, 源程序文件、目标代码文件等。
- Windows和Unix采用无结构的文件形式



三、文件逻辑结构

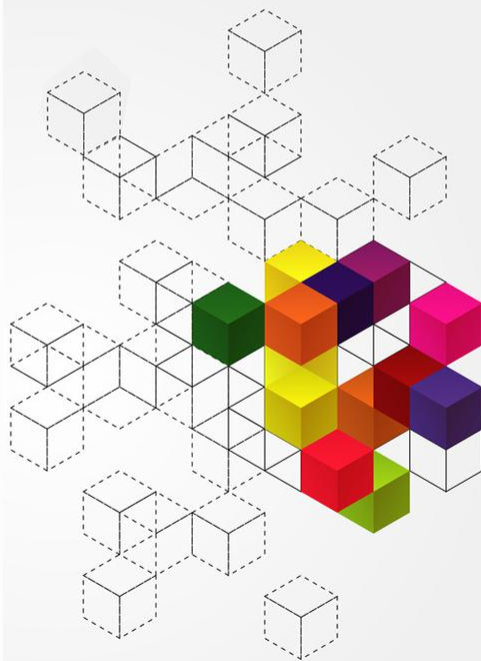
记录式文件

- 把文件中的记录按各种不同的方式排列，构成不同的逻辑结构，以便用户对文件中的记录进行修改、追加、查找和管理等操作。
- **优点:** 便于用户对文件中的记录进行修改、追加、查找和管理。
- **缺点:** 文件管理相对于字符流文件复杂。
- **适用范围:** 适用于需要经常对基本信息单位进行修改、追加、查找的文件。例如: 用于信息系统管理的数据文件。



三、文件逻辑结构

- 大多数的文件系统是无结构文件系统
- 实际应用需要处理各种数据结构，这些数据结构由相应的应用程序提供。
- 例如，Unix的文件系统，本身不含任何数据结构。但是，Unix却能支持使用有结构文件的应用程序，如各种基于Unix操作系统的数据库管理系统，电子邮件系统等。

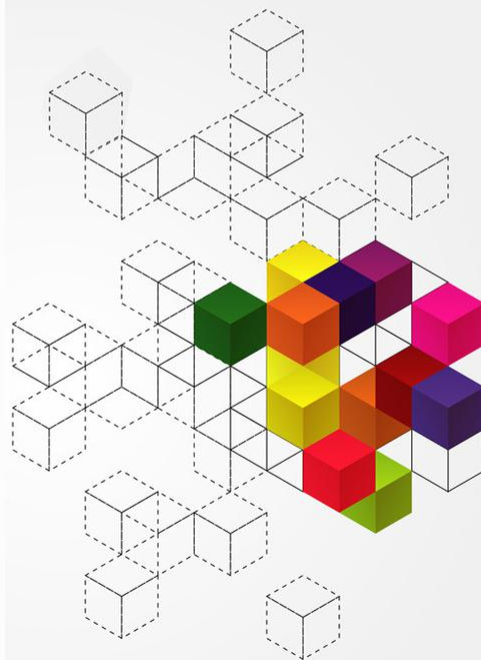


四、文件属性

文件基本属性表 (FCB = directory entry)

属性名称	说明
文件名	File Name
文件ID	Unique File Identifier
文件类型	File Type
文件大小	File Size
文件保护信息	Protection Information
时间信息	Date of Creation, Last Modified, etc.
文件所有者	File Owner

Information about files are kept in the directory structure, which is also maintained on the disk



本讲小结

- 什么是文件
- 文件分类
- 文件逻辑结构
- 文件属性

