

编辑器的 数据结构设计

2006.11

qzy

1

编辑器的数据结构设计

编辑器要保存被编辑内容。虽说编辑内容就是一个字符串，但简单的字符串实现方法不适合编辑器的需要

实现一个编辑器，第一位的工作就是设计编辑区数据结构

设计数据结构，应考虑系统的实际需要，考虑所需要的操作，考虑数据结构对各种相关操作的支持

好的数据结构设计可能使程序结构清晰，操作实现简单，效率高，易于扩充其他操作，适应性强，等等

在本课程有关数据结构的方面，这些是最重要的东西

具体数据结构只是构造所需结构的构件，需要灵活使用

2

最明显的简单选择

最明显的简单选择常常不是好选择！

顺序表：



动态顺序表：



单链表：



注意：上述每种方案都有一些细节，可形成许多不同的设计

3

还有其他选择吗？

4

简单设计中缺少我们需要的重要概念吗？

这些结构里没有“行”的概念

因此，**行操作**的实现很麻烦

行操作在这里很重要，很多

最好能更直接地在数据结构中支持“行”的概念

5

连续存储

结构1：二维编辑区，用一个维表示行（固定最大行长）

变体：动态变化的两位编辑区（还可以有其他细节变化）

- 行数可以变化
- 行长可以变化

O(1)时间访问第*i*行第*j*个字符

行长统一，可能造成很大浪费

6

连续存储

下面回到一维连续存储，考虑对行的支持：

结构2：记录行的长度，加速行查找



每行增加一个字符表示行的长度，字符可以表示 0~255 的数

变体：去掉换行符；记录跳步值(行长+1)

7

连续存储

结构3：增加“行位置表”（外加速）



变体：去掉换行符；记录末端加1位置；等等

O(1)字符访问，跨行匹配（去换行符），有效支持任意行长

结构4：增加“行位置指针表”（外加速）



变体：去掉换行符；末端加1位置指针；等等

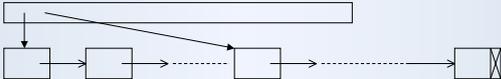
O(1)字符访问，跨行匹配（去换行符），有效支持任意行长

8

简单链接存储

变体：双链表（有利于找前一字符，前一行等）

结构5：链表+行表

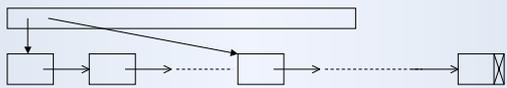


变体：行表可用顺序表、单链表、双链表等，存储区也可用双链表

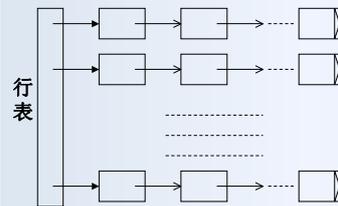
其他变体：给行表的结点增加指向前行最后字符的指针，这样，基本表和行表可以只用单链表实现，许多操作都更容易实现

9

简单链接存储



如果可以直接找到每个行，为什么还要把各行连在一起？



结构6：

行表可用顺序表，动态顺序表，链接表，双链表

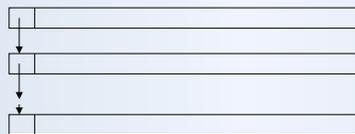
优点：一个行可以任意长，易于管理

缺点：存储效率低

10

分块存储

结构7：行的链表



变体：等长行，建立时确定长度（允许不等长，需要用特殊的程序技术），支持任意长的行（当行内操作空间不够时可换一个更长的行，用特殊程序技术）

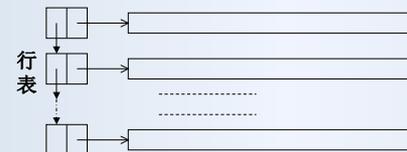
优点：一种结构，容易管理，行操作方便

缺点：链接结构和行内容放在一起，有不方便。查找行需线性时间（行数）

11

分块存储

结构8：行头链表



变体：等长行，分配时确定的不等长行，动态长度的行

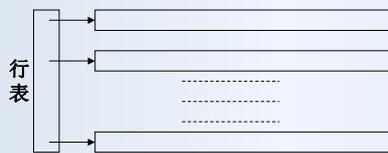
优点：行头和内容存储分开，比较规范，容易管理。根据需要调换行内容存储很方便

缺点：两种结构，查找行需线性时间（行数）

12

分块存储

结构9: 连续行表



变体: 等长行, 分配时确定的不等长行, 动态变动的不等长行, 固定行数和允许变动行数

优点: $O(1)$ 时间的任意字符访问, 常量时间确定行,

缺点: 插入删除行时的行移动

13

为什么要考虑现实的编辑器?

- 要为程序的修改扩充做准备
- 比较和参考有利于理解自己的工作成果和存在问题

14

常见编辑器与本实习的编辑器相比, 它们之间的最大概念差异在哪儿?

- 实习编辑器只有绝对操作
- 常见编辑器里的操作有许多相对操作, 有一个重要概念: 当前位置 (光标位置)。相关操作如:
 - 各种光标移动, 前后删除
 - 当前行操作
 - 从光标向下 (向上) 查找
 - 等等

15

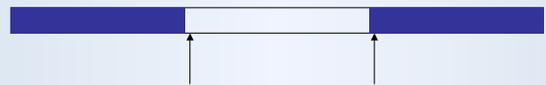
其他设计

结构10: 固定大小存储块的链表, 每个存储块存放 k 个字符 (例如32个), 基于这种存储块的各种组织方式

连续存储的变形 (为支持“当前位置”):

结构11: 用两个栈

结构12: 将编辑内容保存在两端, 中间空位是当前光标位置



前面各种连续存储结构都可以采用这种数据表示形式

16

其他问题

支持不同的编辑命令方式:

- 命令行式
- 所见即所得 (图形用户界面)

支持不同的编辑内容:

- 文字处理 (格式信息, 非文字内容的表示和处理等)
- 程序编辑器
- 表格编辑器, 电子报表和计算, 等等
- 其他情况

17