

线性表

张海鹏 2017-03-26

程序?



程序?

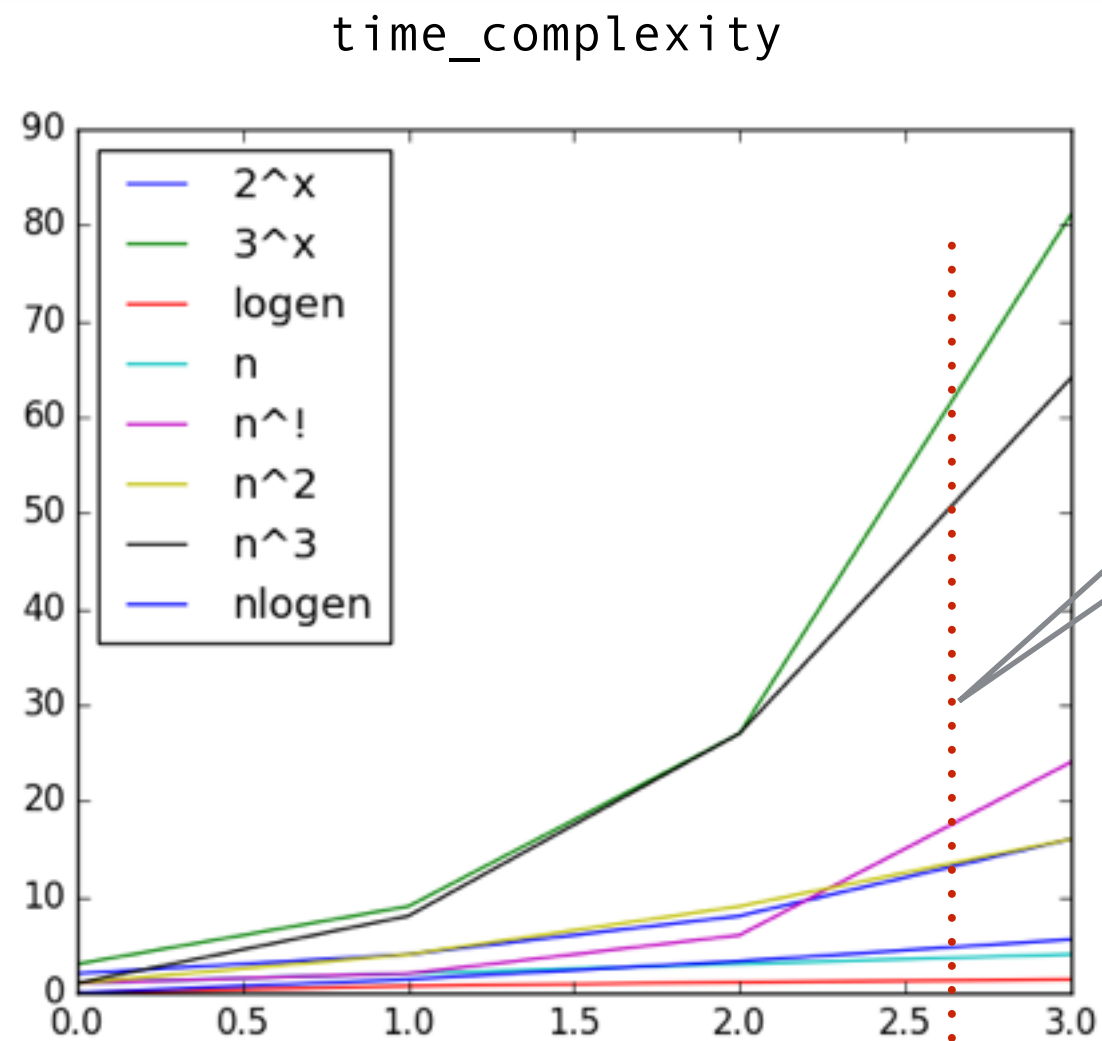
数据结构	算法
数组	直接访问
单链表	遍历元素

程序=数据结构+算法

数据结构: 逻辑和物理

算法(复杂度): 时间和空间

时间复杂度?



渐近

$O(1) < O(\log_2 N) < O(N) < O(N \log_2 N) < O(N^2) < O(N^3) < O(2^N) < O(3^N) < O(N!)$

渐近界?

$O(g(n)) = \{ f(n) \mid \text{存在正常数 } c \text{ 和 } n_0 \text{ 使得对所有 } n \geq n_0 \text{ 有: } 0 \leq f(n) \leq cg(n) \}$

$\Omega(g(n)) = \{ f(n) \mid \text{存在正常数 } c \text{ 和 } n_0 \text{ 使得对所有 } n \geq n_0 \text{ 有: } 0 \leq cg(n) \leq f(n) \}$

$\Theta(g(n)) = \{ f(n) \mid \text{存在正常数 } c_1, c_2 \text{ 和 } n_0 \text{ 使得对所有 } n \geq n_0 \text{ 有: } c_1g(n) \leq f(n) \leq c_2g(n) \}$

$$T(n) = aT(n/b) + f(n)$$

Theorem 4.1 (Master theorem)

Let $a \geq 1$ and $b > 1$ be constants, let $f(n)$ be a function, and let $T(n)$ be defined on the nonnegative integers by the recurrence

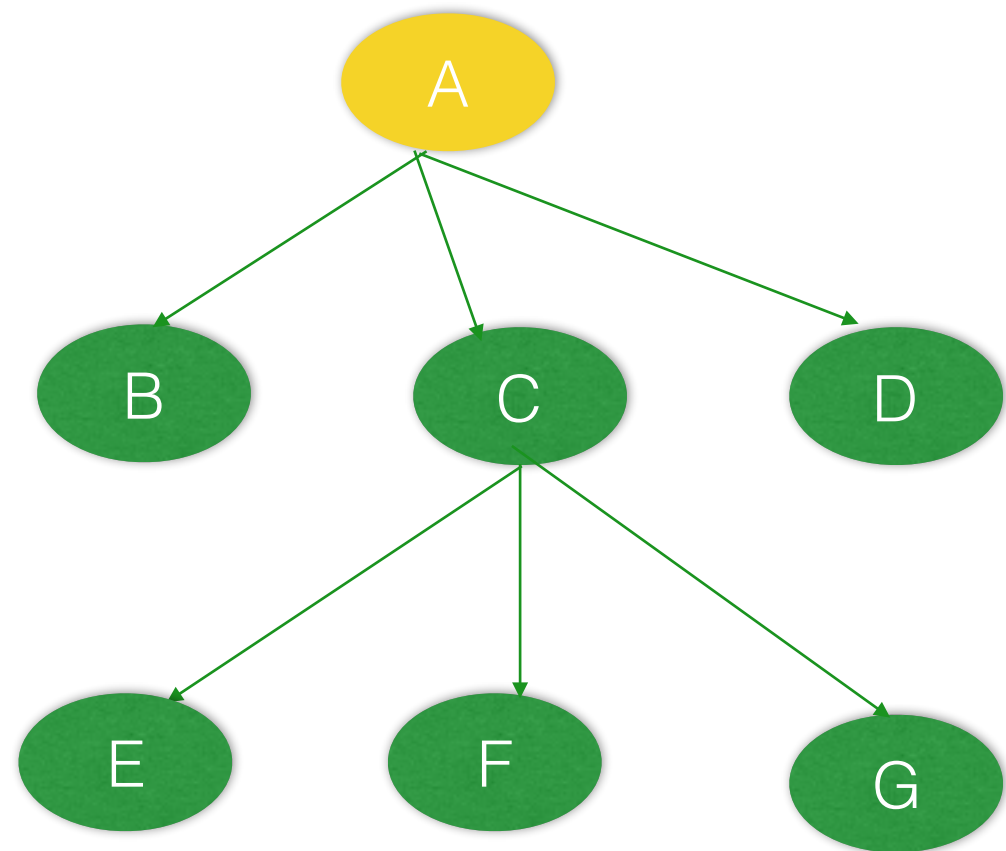
$$T(n) = aT(n/b) + f(n),$$

where we interpret n/b to mean either $\lfloor n/b \rfloor$ or $\lceil n/b \rceil$. Then $T(n)$ has the following asymptotic bounds:

1. If $f(n) = O(n^{\log_b a - \epsilon})$ for some constant $\epsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$.
 2. If $f(n) = \Theta(n^{\log_b a})$, then $T(n) = \Theta(n^{\log_b a} \lg n)$.
 3. If $f(n) = \Omega(n^{\log_b a + \epsilon})$ for some constant $\epsilon > 0$, and if $af(n/b) \leq cf(n)$ for some constant $c < 1$ and all sufficiently large n , then $T(n) = \Theta(f(n))$. ■
-

$$T(n) = aT(n/b) + f(n)$$

- Substitution method
- Recursion-tree method
- Master method



RoadMap

- 线性表

顺序表，链表(单链表，双链表，循环链表，十字链表)等

- 栈

- 队列

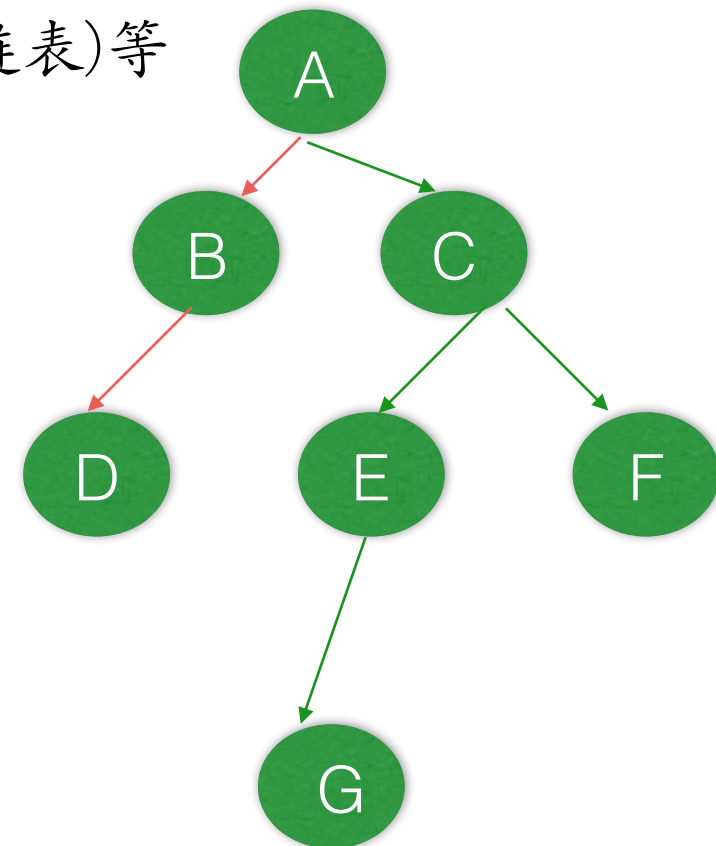
双端队列，优先级队列，循环队列等

- 树

二叉树，森林，平衡树，B树，B+树，红黑树等

- 图

有向图，无向图等



Application

单链表：内存分配

栈/堆：OS等

队列：OS进程调度等

B/+树：MySQL等

图：导航地图等



Java/C++/Python

```
import java.util.ArrayList;

public class ArrayListTest{

    public static void main(String[] args){

        ArrayList<String> list = new ArrayList<String>();

        list.add("a");
        list.add("b");
        list.set(1,"c");

        String val = list.get(1);

    }
```



LinkedList
Map
Set

Java/C++/Python

```
import pandas as pd
```

```
import numpy as np
```

```
a = []
```

```
a.append('a')
```

```
a.append('b')
```

```
a.remove('a')
```

```
del a[0]
```

```
# del the last element of a
```

```
a.pop()
```



Series
DataFrame
Dict and so on

interview

- 中序遍历二叉树，利用 $O(1)$ 空间统计遍历每个节点的层次？
（今日头条，2016）
 - python中list删除操作的代价？（某安全公司，2016）
 - 用最小时间复杂度求数组中第K大的数？（今日头条，2016）
 - 给一个query，如何快速从10亿个query中找出和它最相似的？（网易杭州研究院，2016）
 - 堆排序为什么是 $O(n\log n)$ 的效率？（京东，2016）
-

知识点

顺序表：地址连续的存储空间

实现：数组

缺点：更新操作费时

单链表：存储空间地址不连续

实现：指针

优点：内存节约

缺点：查找操作费时

循环链表：尾结点指针指向头结点

优点：头结点和尾结点使用频繁的场景

静态链表

SL=0

AV=6

index	data	next
0		4
1	a4	5
2	a2	3
3	a3	1
4	a1	2
5	a5	-1
6		7
7		7

双链表：每个节点两个指针

优点：两个方向移动

总结

- A. 什么是程序? 程序=算法+数据结构
 - B. 算法复杂度? 时间复杂度+空间复杂度
 - C. 数据结构的基本知识体系? 线性表, 栈, 队列, 树, 图等
 - D. 数据结构的具体应用场景? 具体技术+面试问题
 - E. 数据结构的封装实现? 实际应用
 - F. 线性表的基本知识点? 顺序表+链表(单, 循环, 双, 静态)
-