数据结构与算法综述与复杂度分析基础

陈宇琪

2020 年 3 月 10 日

摘要

主要内容包括数据结构综述和算法复杂度分析。

预备知识：复杂度分析

提交方式：在超星上提交第二和第三部分作业。

参考链接：https://www.jianshu.com/p/eb8ed6032f98

ddl：2020-03-14

目录

1 数据结构综述

1.1 数据结构到底是什么？

“数据结构是数据对象，以及存在于该对象的实例和组成实例的数据元素之间的各种联系。这些联系可以通过定义相关的函数来给出。” — Sartaj Sahni，《数据结构、算法与应用》

“数据结构是 ADT（抽象数据类型 Abstract Data Type）的物理实现。” — Clifford A.Shaffer，《数据结构与算法分析》

“数据结构（data structure）是计算机中存储、组织数据的方式。通常情况下，精心选择的数据结构可以带来最优效率的算法。” — 中文维基百科

1.2 到底什么是数据结构？

- 数据对象在计算机中的组织方式

  逻辑结构

  物理存储结构：逻辑结构在机器的内存里面怎么放

- 数据对象必定与一系列加在其上的操作相关联
- 完成这些操作所用的方法就是算法

所有的数据都是由数据元素构成，数据元素是数据的基本单位。数据元素也称元素、结点、顶点、记录。一个数据元素可以由若干个数据项（也可称为字段、域、属性）组成，数据项是数据的不可分割的最小单位。

1.3 存在的目的

让我们操作里面的数据，快速。

- 方便我们增加
- 方便我们快速查找到 (查是删改的前提)
- 节省空间 (存储)

当然，操作的时候，不光与数据结构有关，与算法的精妙程度也有关系。

1.4 什么是逻辑结构？

简单说，逻辑结构就是数据之间的关系。

而按数据之间的关系来说，逻辑结构大概可以分为两种:

1.4.1 线性结构

- 线性表：顺序表、链表（线性链表、循环链表、双向链表）
- 栈（顺序栈、链栈）
- 队列（顺序队列、链队列 (循环链等)）
- ...

1.4.2 非线性结构

- 集合
- 树 (二叉树)
- 图 (网)
- ...

有且只有一个开始结点和一个终端结点，并且所有结点都最多只有一个直接前驱和一个直接后继。它们共同的特点就是数据之间的线性关系，除了头结点和尾结点之外，每个结点都有唯一的前驱和唯一的后继，也就是所谓的一对一的关系。

队列的先进先出、栈的后进先出

1.5 存储结构

存储结构：逻辑结构用计算机语言的实现。

1.5.1 常见的存储结构

- 顺序存储 – (应用：数组)

  优点：节省空间，可以实现随机存取；

  缺点：插入、删除时需要移动元素，效率低。

- 链式存储 – (应用：指针)

  优点：插入、删除灵活 (不必移动节点，只要改变节点中的指针)

  缺点：不能随机存取，查找速度慢。

- 索引存储：除建立存储结点信息外，还建立附加的索引表来标识结点的地址。索引表由若干索引项组成。

  特点：索引存储结构是用结点的索引号来确定结点存储地址，其优点是检索速度快，缺点是增加了附加的索引表, 会占用较多的存储空间。

- 散列存储：hash 存储，是一种力图将数据元素的存储位置与关键码之间建立确定对应关系的查找技术。

  特点：散列是数组存储方式的一种发展，相比数组，散列的数据访问速度要高于数组，因为可以依据存储数据的部分内容找到数据在数组中的存储位置，进而能够快速实现数据的访问。

  理想的散列访问速度是非常迅速的，而不像在数组中的遍历过程，采用存储数据组中内容的部分元素作为映射函数的输入，映射函数的输出就是存储数据的位置，这样的访问速度就省去了遍历数组的实现，**因此时间复杂度可以认为为 O(1)，而数组遍历的时间复杂度为 O(n)**。

2 选择题

1、以下与数据的存储结构无关的术语是（）。

A．顺序队列 B. 链表 C. 有序表 D. 链栈

2、与数据元素本身的形式、内容、相对位置、个数无关的是数据的（）。

A．存储结构 B．存储实现 C．逻辑结构 D．运算实现

3、以下数据结构中，（）是非线性数据结构

A．树 B．字符串 C．队 D．栈

3 简答题

1、给出在复杂度分析中 $O, o, \Omega, \omega, \Theta$ 的定义。

2、Relative asymptotic growths

Indicate, for each pair of expressions (A, B) in the table below, whether A is O, o, Ω, ω, or Θ of B. Assume that $k \geq 1$, $\epsilon > 0$, and $c > 1$ are constants. Your answer should be in the form of the table with "yes" or "no" written in each box.

	A	B	O	o	Ω	ω	Θ
a.	$\lg^k n$	n^ϵ					
b.	n^k	c^n					
c.	$\sqrt{n}$	$n^{\sin n}$					
d.	2^n	$2^{n/2}$					
e.	$n^{\lg c}$	$c^{\lg n}$					
f.	$\lg(n!)$	$\lg(n^n)$					

图 1: Relative asymptotic growths

4 LeetCode 准备

1、给定一个非空整数数组，除了某个元素只出现一次以外，其余每个元素均出现两次。找出那个只出现了一次的元素。

说明：复杂度要求是 $O(n)$，且不能使用额外的存储空间。

提交链接：https://leetcode-cn.com/problems/single-number/。

2、求 $n!$ 末位 0 的个数。

说明：复杂度要求是 $O(log(n))$。

提交链接：https://leetcode-cn.com/problems/factorial-trailing-zeroes/

3、给定一个数组，它的第 i 个元素是一支给定股票第 i 天的价格。设计一个算法来计算你所能获取的最大利润。你可以尽可能地完成更多的交易（多次买卖一支股票）。

说明：这个问题求一些进阶的版本供大家挑战！

提交链接：https://leetcode-cn.com/problems/best-time-to-buy-and-sell-stock-ii/

5 参考答案

5.1 选择题

CCA

5.2 简答题

$f(x) \in O(g(x)) : \exists c > 0, \exists x_0 >> 2, \forall x \geq x_0, 0 \leq f(x) \leq c * g(x)$

$f(x) \in o(g(x)) : \forall c > 0, \exists x_0 >> 2, \forall x \geq x_0, 0 \leq f(x) < c * g(x)$

$f(x) \in \Omega(g(x)) : \exists c > 0, \exists x_0 >> 2, \forall x \geq x_0, 0 \leq c * g(x) \leq f(x)$

$f(x) \in \omega(g(x)) : \forall c > 0, \exists x_0 >> 2, \forall x \geq x_0, 0 \leq c * g(x) < f(x)$

$f(x) \in \Theta(g(x)) : \exists c_1, c_2 > 0, \exists x_0 >> 2, \forall x \geq x_0, 0 \leq c_1 * g(x) \leq f(x) \leq c_2 * g(x)$

	A	B	O	o	Ω	ω	θ
a)	$\lg^k n$	n^ϵ	YES	YES	NO	NO	NO
b)	n^k	c^n	YES	YES	NO	NO	NO
c)	$\sqrt{n}$	$n^{\sin n}$	NO	NO	NO	NO	NO
d)	2^n	$2^{n/2}$	NO	NO	YES	YES	NO
e)	$n^{\lg c}$	$c^{\lg n}$	YES	NO	YES	NO	YES
f)	$\lg(n!)$	$\lg(n^n)$	YES	NO	YES	NO	YES

图 2: Answer

a. Applying L'Hopital's Rule k times, we get

$$\begin{aligned}&\lim_{n\to\infty}\frac{(\lg n)^k}{n^\varepsilon}\\ &=\frac{k!}{\epsilon^k \ln^k 2}\lim_{n\to\infty}\frac{1}{n^\epsilon}\\ &=0\end{aligned}$$

So, $\lg^k n = O(n^\varepsilon)$

图 3: Answer(a)

b. Again, by $\lceil k \rceil$ applications of L'Hopital's rule,

$$\lim_{n\to\infty}\frac{n^k}{c^n}=0,$$

Hence, $n^k = O(c^n)$.

图 4: Answer(b)

c. $\sin n$ is a periodic function and takes the values in the range [-1, 1].
When $\sin n$=1 we have,

$$\lim_{n\to\infty}\frac{n}{n^{1/2}}=\infty$$

When $\sin n$ = -1 we have,

$$\lim_{n\to\infty}\frac{n^{-1}}{n^{1/2}}=0$$

图 5: Answer(c)

d. $\lim_{n\to\infty}\frac{2^n}{2^{n/2}}=\lim_{n\to\infty}2^{n/2}$
$=\infty$
Therefore, $2^n=\omega(2^{n/2})$.

图 6: Answer(d)

e. $n^{\lg c}=c^{\lg n}$
Take log on both sides:

$$\log\left(n^{\lg c}\right)=\log\left(c^{\lg n}\right)$$
$$(\log c).(\log n)=(\log n).(\log c)$$

图 7: Answer(e)

f. $\lg(n^n) = n\lg(n)$,

$\log! = \sum_{i=1}^{n} \lg i \le \sum_{i=1}^{n} \lg n = n\lg n$ and thus $\lg! = O(n\lg n)$

$$\begin{aligned}\log! &= \sum_{i=1}^{n} \lg i \ge \sum_{i=n/2+1}^{n} \lg n/2 \\ &= n/2 \lg n/2 \\ &= (n/2-1)\lg n/2\end{aligned}$$

And thus $\lg n! = \Omega(n\lg n)$

Finally $\boxed{\lg n! = o(n\lg n) \quad \text{and} \quad \lg n! = \omega(n\lg n)}$

The limit $\boxed{\lim_{n\to\infty} \frac{\lg n!}{\lg n^n}}$ lies in the interval [1/2, 1].

图 8: Answer(f)