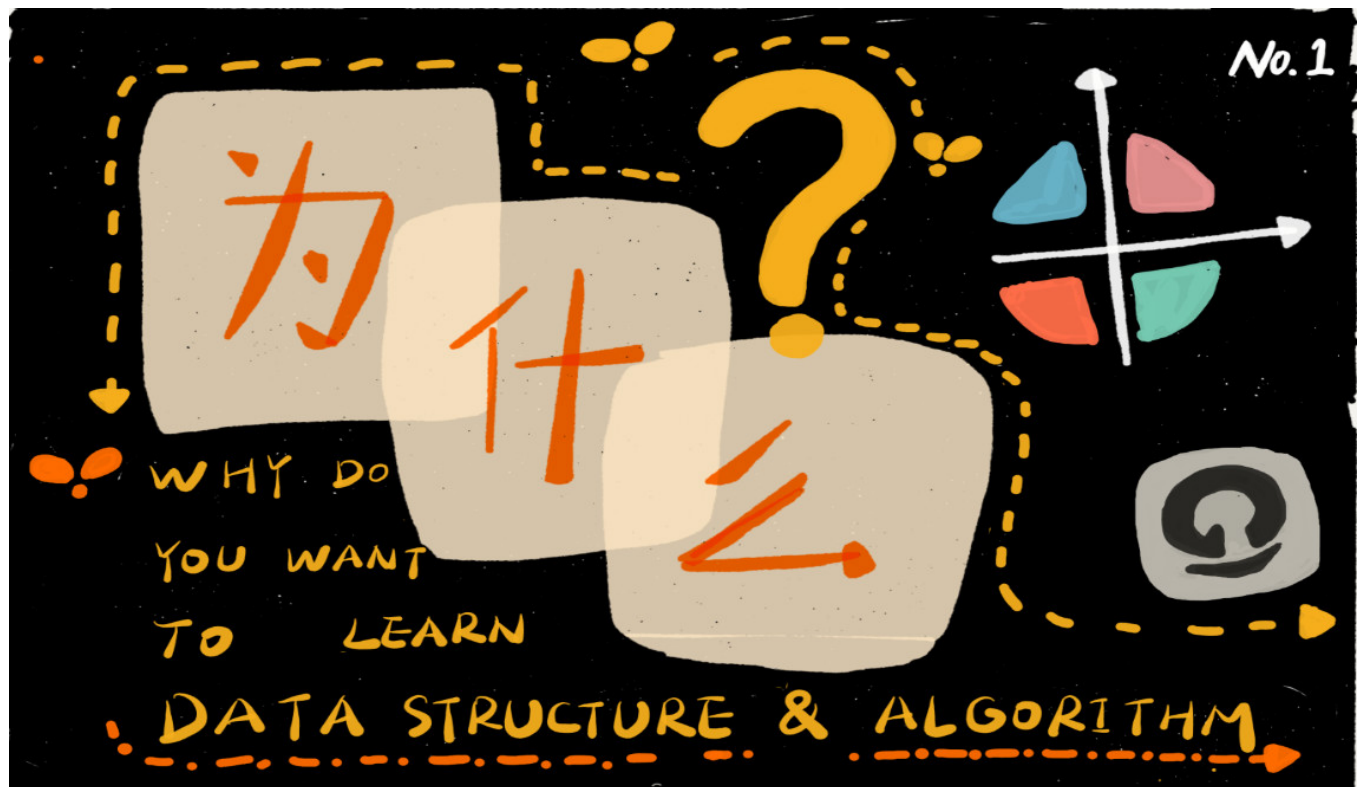


讲堂 > 数据结构与算法之美 > 文章详情

01 | 为什么要学习数据结构和算法？

2018-09-21 王争



01 | 为什么要学习数据结构和算法？

朗读人：修阳 08'59" | 4.13M

你是不是觉得数据结构和算法，跟操作系统、计算机网络一样，是脱离实际工作的知识？可能除了面试，这辈子也用不着？

尽管计算机相关专业的同学在大学都学过这门课程，甚至很多培训机构也会培训这方面的知识，但是据我了解，很多程序员对数据结构和算法依旧一窍不通。还有一些人也只听说过数组、链表、快排这些最最基本的数据结构和算法，稍微复杂一点的就完全没概念。

当然，也有很多人说，自己实际工作中根本用不到数据结构和算法。所以，就算不懂这块知识，只要 Java API、开发框架用得熟练，照样可以把代码写得“飞”起来。事实真的是这样吗？

今天我们就来详细聊一聊，为什么要学习数据结构和算法。

想要通关大厂面试，千万别让数据结构和算法拖了后腿

很多大公司，比如 BAT、Google、Facebook，面试的时候都喜欢考算法、让人现场写代码。有些人虽然技术不错，但每次去面试都会“跪”在算法上，很是可惜。那你有没有想过，为什么

这些大公司都喜欢考算法呢？

校招的时候，参加面试的学生通常没有实际项目经验，公司只能考察他们的基础知识是否牢固。社招就更不用说了，越是厉害的公司，越是注重考察数据结构与算法这类基础知识。相比短期能力，他们更看中你的长期潜力。

你可能要说了，我不懂数据结构与算法，照样找到了好工作啊。那我是不是就不用学数据结构和算法呢？当然不是，你别忘了，**我们学任何知识都是为了“用”的，是为了解决实际工作问题的**，学习数据结构和算法自然也不例外。

业务开发工程师，你真的愿意做一辈子 CRUD boy 吗？

如果你是一名业务开发工程师，你可能要说，我整天就是做数据库 CRUD（增删改查），哪里用得到数据结构和算法啊？

是的，对于大部分业务开发来说，我们平时可能更多的是利用已经封装好的现成的接口、类库来堆砌、翻译业务逻辑，很少需要自己实现数据结构和算法。但是，**不需要自己实现，并不代表什么都不需要了解。**

如果不知道这些类库背后的原理，不懂得时间、空间复杂度分析，你如何能用好、用对它们？存储某个业务数据的时候，你如何知道应该用 ArrayList，还是 Linked List 呢？调用了某个函数之后，你又该如何评估代码的性能和资源的消耗呢？

作为业务开发，我们会用到各种框架、中间件和底层系统，比如 Spring、RPC 框架、消息中间件、Redis 等等。**在这些基础框架中，一般都揉和了很多基础数据结构和算法的设计思想。**

比如，我们常用的 Key-Value 数据库 Redis 中，里面的有序集合是用什么数据结构来实现的呢？为什么要用跳表来实现呢？为什么不用二叉树呢？

如果你能弄明白这些底层原理，你就能更好地使用它们。即便出现问题，也很容易就能定位。因此，**掌握数据结构和算法，不管对于阅读框架源码，还是理解其背后的设计思想，都是非常有用的。**

在平时的工作中，数据结构和算法的应用到处可见。我来举一个你非常熟悉的例子：如何实时地统计业务接口的 99% 响应时间？

你可能最先想到，每次查询时，从小到大排序所有的响应时间，如果总共有 1200 个数据，那第 1188 个数据就是 99% 的响应时间。很显然，每次用这个方法查询的话都要排序，效率是非常低的。但是，如果你知道“堆”这个数据结构，用两个堆可以非常高效地解决这个问题。

基础架构研发工程师，写出达到开源水平的框架才是你的目标！

现在互联网上的技术文章、架构分享、开源项目满天飞，照猫画虎做一套基础框架并不难。我就拿 RPC 框架举例。

不同的公司、不同的人做出的 RPC 框架，架构设计思路都差不多，最后实现的功能也都差不多。但是有的人做出来的框架，Bug 很多、性能一般、扩展性也不好，只能在自己公司仅有的几个项目里面用一下。而有的人做的框架可以开源到 GitHub 上给很多人用，甚至被 Apache 收录。为什么会有这么大的差距呢？

我觉得，高手之间的竞争其实就在细节。这些细节包括：你用的算法是不是够优化，数据存取的效率是不是够高，内存是不是够节省等等。这些累积起来，决定了一个框架是不是优秀。所以，如果你还不懂数据结构和算法，没听说过大 O 复杂度分析，不知道怎么分析代码的时间复杂度和空间复杂度，那肯定说不过去了，赶紧来补一补吧！

对编程还有追求？不想被行业淘汰？那就不要只会写凑合能用的代码！

何为编程能力强？是代码的可读性好、健壮？还是扩展性好？我觉得没法列，也列不全。但是，在我看来，**性能好坏起码是其中一个非常重要的评判标准**。但是，如果你连代码的时间复杂度、空间复杂度都不知道怎么分析，怎么写出高性能的代码呢？

你可能会说，我在小公司工作，用户量很少，需要处理的数据量也很少，开发中不需要考虑那么多性能的问题，完成功能就可以，用什么数据结构和算法，差别根本不大。但是你真的想“十年如一日”地做一样的工作吗？

经常有人说，程序员 35 岁之后很容易陷入瓶颈，被行业淘汰，我觉得原因其实就在此。有的人写代码的时候，从来都不考虑非功能性的需求，只是完成功能，凑合能用就好；做事情的时候，也从来没有长远规划，只把眼前事情做好就满足了。

我曾经面试过很多大龄候选人，简历能写十几页，经历的项目有几十个，但是细看下来，每个项目都是重复地堆砌业务逻辑而已，完全没有难度递进，看不出有能力提升。久而久之，十年的积累可能跟一年的积累没有任何区别。这样的人，怎么不会被行业淘汰呢？

如果你在一家成熟的公司，或者 BAT 这样的大公司，面对的是千万级甚至亿级的用户，开发的是 TB、PB 级别数据的处理系统。性能几乎是开发过程中时刻都要考虑的问题。一个简单的 ArrayList、Linked List 的选择问题，就可能会产生成千上万倍的性能差别。这个时候，数据结构和算法的意义就完全凸显出来了。

其实，我觉得，数据结构和算法这个东西，如果你不去学，可能真的这辈子都用不到，也感受不到它的好。但是一旦掌握，你就会常常被它的强大威力所折服。之前你可能需要费很大劲儿来优化的代码，需要花很多心思来设计的架构，用了数据结构和算法之后，很容易就可以解决了。

内容小结

我们学习数据结构和算法，并不是为了死记硬背几个知识点。我们的目的是建立时间复杂度、空间复杂度意识，写出高质量的代码，能够设计基础架构，提升编程技能，训练逻辑思维，积攒人生经验，以此获得工作回报，实现你的价值，完善你的人生。

所以，不管你是业务开发工程师，还是基础架构工程师；不管你是初入职场的初级工程师，还是工作多年的资深架构师，又或者是想转人工智能、区块链这些热门领域的程序员，数据结构与算法作为计算机的基础知识、核心知识，都是必须要掌握的。

掌握了数据结构与算法，你看待问题的深度，解决问题的角度就会完全不一样。因为这样的你，就像是站在巨人的肩膀上，拿着生存利器行走世界。数据结构与算法，会为你的编程之路，甚至人生之路打开一扇通往新世界的大门。

课后思考

你为什么要学习数据结构和算法呢？在过去的软件开发中，数据结构和算法在哪些地方帮到了你？

欢迎留言和我分享，我会第一时间给你反馈。

 极客时间

数据结构与算法之美

为工程师量身打造的数据结构与算法私教课

王争
前 Google 工程师



新版升级：点击「 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

©版权归极客邦科技所有，未经许可不得转载

上一篇 开篇词 | 从今天起，跨过“数据结构与算法”这道坎

下一篇 02 | 如何抓住重点，系统高效地学习数据结构与算法？

精选留言



姜威

👍 1526

为什么学习数据结构和算法？我认为有3点比较重要

- 1.直接好处是能够有写出性能更优的代码。
- 2.算法，是一种解决问题的思路和方法，有机会应用到生活和事业的其他方面。
- 3.长期来看，大脑思考能力是个人最重要的核心竞争力，而算法是为数不多的能够有效训练大脑思考能力的途径之一。

2018-09-21

| 作者回复

写的很好 同学们把这条回复顶上去

2018-09-21



HouShangLing

👍 680

一定要动手写

2018-09-21

| 作者回复

你掌握了学这门课的最有效的方法。看十遍也没自己实现一遍学的牢。同学们这条也帮忙顶上去

2018-09-21



wistbean

👍 148

其实问题的所在就是现在有很多现成的框架，器又大活又好，拿来就用，还不用太担心性能的问题。就好像那些建楼的工程师不需要懂砖瓦的构成，也能盖楼。司机不需要懂汽车引擎的原理，也能当顺风车司机载美女兜风。

遇到不会的上 Google，懒了就上 GitHub 找框架。所以写了这么多年代码，一直是个菜鸟。

或许是时候修炼自己的内功了。一直 CURD 有意思么？把设计模式，网络原理，数据结构和算法捡起来，或许就没那么容易菜鸟了。怼人也更加自信了。跳槽也不用畏畏缩缩的了。

为了不当菜鸟，我和我的小伙伴们就加入了哇！

2018-09-21

| 作者回复

写的太好了

2018-09-21



两颗小虎牙

👍 118

可不可以每次多发布几篇？

2018-09-21



tdytaylor

👍 91

老师，就像你说的，工作中其实好多都和业务挂钩，基本上都是针对业务做增删改查，很难把所学的算法应用起来，我平时也时不时学习算法，但总感觉学了就忘，忘了又学，如此反复，老师，这种到底是没了解算法的原理导致不会灵活应用，还是写的少导致的，感觉学习算法很少能应用起来

2018-09-21

作者回复

1. 客观的讲，有些项目确实涉及的数据结构和算法少一些，你可以再看下我文章里写的。
2. 你提到学了又忘，我觉得一方面你是没有掌握学习的方法，学习的重点，走马观花的看肯定比较容易忘；我们02节会具体讲；
3. 不会灵活应用？那估计还是没有好的教材教你如何应用，还有可能就是确实还没掌握太牢，只是懂点皮毛，很浅，灵活应用是一个比较的境界，需要一段时间的沉淀学习。
4. 学习算法并不是为了记住几个排序、二分查找、二叉树遍历，他还能锻炼你的逻辑思维、性能意识，而且，如果你写代码能力还有欠缺，你还可以通过把学到的数据结构和算法都实现一遍，这是一种很好很好的锻炼编程能力的方法。所以不要过度追求一定要在项目里手写快排、手写二叉树才能算是用上。

2018-09-21



裴海港

👍 81

年近40的大叔也有一颗积极向上的心

2018-09-21

作者回复

终身学习 多大年纪也不晚

2018-09-21



五岳寻仙

👍 74

老师好！看到专栏第一眼就果断订阅了。我是一个菜鸟程序员，半年的工作经验让我感受到数据结构和算法太重要了！讲一个自己亲身经历的例子。

入职不久，就遇到一个需求，需要建立一个3G(30亿)条键值对映射，已供后续检索。听上去很简单的问题，用python的字典就可以解决。但在实现的过程中，很快就遇到了问题：字典是基于hash的，对于每条键值对要多消耗50个字节的内存维持这种结构(即便使用redis也需要这个内存开销)，再加上键值存储消耗的内存，我大约需要 $3G \times 70 = 210G$ 内存，超过服务器内存了。

后来，想到了可以根据键排序后线性存储，使用二分查找，解决了这个问题，大约消耗的内存也就 $3G \times 10 = 30G$ 左右。

看似简单的问题，当规模大到一定程度，不借助算法和数据结构，就无法解决了。

2018-09-21

作者回复

哈哈 你要是看到我讲的散列那一篇你就知道了 像java里的hashmap是比较耗内存的 你用到解决方案是一种用时间复杂度换空间复杂度的思路 我们专栏也会讲的 不过你现在的解决办法还可以更高效 利用hash函数 我们专栏也会讲到 还有二分是 $\log n$ 的时间复杂度 是非常高效

的一种时间复杂度 2^{64} 次方个有序数据二分查找也顶多循环64次 有没有觉得logn这个复杂度很奇妙

2018-09-21



Haoz

👍 41

一、数据结构和算法是什么

- 1、数据结构是指一组数据的存储结构
- 2、算法就是操作数据的方法
- 3、数据结构和算法是相辅相成的，数据结构是为算法服务的，而算法要作用在特定的数据结构之上

二、学习的重点在什么地方

数据结构和算法解决的是如何更省、更快地存储和处理数据的问题，因此，我们就需要一个考量效率和资源消耗的方法，这就是复杂度分析方法。在学习数据结构和算法的过程中，要学习它的「来历」、「自身的特点」、「适合解决的问题」以及「实际的应用场景」。

- 1、数据结构和算法学习的精髓-复杂度分析
- 2、最常用的、最基础的数据结构：数组、链表、栈、队列、散列表、二叉树、堆、跳表、图、Trie 树
- 3、最常用的算法：递归、排序、二分查找、搜索、哈希算法、贪心算法、分治算法、回溯算法、动态规划、字符串匹配算法

感谢老师的分析，一直以来数据结构和算法都是我的硬伤，学了很多次，也放弃了很多次，可能是方法不对，但更多的是因为自己之前的毅力不够。

在此立下 flag：从本篇文章开始，将老师的文章根据自己的理解进行输出相应的总结和思考，发表成留言，并将每周三篇文章中涉及的数据结构和算法用 Java 实现一次。

2018-09-25

作者回复

写的很好👍

2018-09-25



钳子钳子干

👍 24

大二上，刚刚学完第二章顺序表和链表（一切为了性能xd（半夜等更居然真的刷出来了（x

2018-09-21



小麦2018

👍 21

我还是觉得王争老师自己读的音频好，亲切，有听作者亲自讲解的感觉。

2018-09-21



邵峰

👍 19

记得以前看金庸小说，一个人练降龙十八掌，一个人练全真教内功，刚开始练内功的被吊打，十年后练内功的依然被吊打，但是20年后，降龙十八掌就怎么也打不过练内功的了，算法和数据结构就是内功，降龙十八掌我觉得就是各种框架，再怎么熟悉框架，也不如了解它内部的算法和数据结构

2018-09-21

作者回复

太形象了：)

2018-09-21



DDT

19

就是想升工资吧 我知道学技术必须要拿出十年磨一剑的精神 从头开始...

2018-09-21

作者回复

十年磨一剑说的太好了。我也是这么认为的。做技术就是不要浮躁。要耐得住寂寞。沉得下心。

2018-09-21



Liar

15

不做一个没有追求的码农，那些说过了35就不行的码农大部分都是不求上进的人。

2018-09-21

作者回复

说得好 不管哪个行业都会淘汰不求上进的人 it也不例外

2018-09-21



阿康

14

老师，我是一名iOS开发，怎么听完你这篇文章还是觉得算法对于我开发来说不是那么重要😄后面我有点怕自己听不懂，而且动手写，我该如何下笔，感觉自己很迷茫.....

2018-09-21

作者回复

1. 假设我们现在要做这样一个功能，我们希望在app上存储一个多级地址列表，用户可以一层一层的选择地址列表中的地址，来设置自己的所在的省份、市、区，如果这个地址列表不是经常变动的，我们希望保存在app端，这样就不需要每次操作地址列表都要跟后台交互，如果让你存储这个多级地址列表，你会怎么存储呢？

2. 如果地址列表也并不是一直都不变的，如果地址列表改变了，我们又不希望发新版，那如何更新app上的地址列表呢？如果地址列表比较大，我们不希望app重新全量的从服务器再拉一次，那又如何来做呢？

3. 我会用一种语言来实现 你可以翻译成object c

2018-09-21



小行家

11

为什么学习数据结构与算法？

1.作为一名程序员，我觉得这个基本功，就好比建房子一样，地基一定要牢固

2. 现在有如此多的开源项目，人们往往开发中只是照搬，调试，集成，然后很少会去注意性能和优化方面，慢慢的也就成了CRUD boy
3. 算法提供的是一种思想，这个思想真的可以让你超神，让你对待问题有着独特的见解，我觉得这个才是最重要的。不再是原来的，没有思考
4. 数据结构作为cs的基本专业，并且贯穿整个cs学习生涯，在cs方面的研究还是有很大的实用价值。

数据结构与算法的意义？

1. 在本科开发中，曾经就自己写了一个类似于目录搜索的算法，就是简单的for循环，性能极其低下，改为二叉树后，性能大大提升
2. 目前读研，接触机器学习理论和深度学习，原有的算法和数据结构的学习中，你所掌握的思想，确实会有很大的意义。

最后，我认为，学习最重要的是你能够自己成长，这个才是最棒的。

2018-09-21

作者回复

说的好

2018-09-21



lane

11

现实，都是面相领导编程.....一言难尽...

2018-09-21

作者回复

没事的 如果工作不满意 不顺心 更要卧薪尝胆 提高能力 沉淀自己。总有一天 机会会到来 你做好充分的抓住它的准备就好。人在职场中 只要抓住一两次大的机会 就能做到很高的职位 就怕的是机会来了我们也抓不住

2018-09-21



grandcool

10

没理解统计业务接口99%响应时间啥意思

2018-09-21

作者回复

举一个例子 你写了一个接口 每天有成千上万的访问 你如何知道这个接口够不够快？响应时间是1s还是5s？如何统计度量？用平均值？显然不是太适合？那用什么值来统计度量呢？你可以自己搜索研究下

2018-09-21



SelfImpr李

8

三年小开发来说一波：上周生产系统上遇到了Java的内存溢出情况（可以买彩票了，OOM事件一直在我们的书上），经过分析dump文件，发现Java的类和Java的动态语言生成的对象没有及时释放，我尽全力就修改了代码上，该释放释放；再过了一个星期分析dump文件，发现少了很多，但是还有一些遗留着，我想是时候需要《数据结构和算法》来帮我了！PS：之前也看过红壳得《算法导论》，看了一周，然后又用来垫显示器了，来来回回几次，最终还是垫显示器了。

2018-09-21

| 作者回复

oom这种问题看似很遥远 实际上我还真遇到过很多次 多数是从数据库捞数据时 一不小心捞了太多的数据到内存 就oom了 而且这种情况用mat等工具也无法分析 因为这种oom是一过性的 heap dump的时候可能数据已经不可达 而mat这些工具只能分析可达对象

2018-09-21



Pardo

7

大学的时候没意识，工作之后才发现，不熟悉数据结构和算法，写出的代码虽然能用，但完全经不起推敲。随着项目复杂度的提高，没有算法护体，代码基本上会是一塌糊涂，迷失在细节调整和障碍修复中不能自拔。

2018-09-21



Mortimey

7

作为一枚iOS开发者，感觉里面好多的专业术语是我没听过的，请问老师，后期的课程也是这样的吗？

2018-09-21

| 作者回复

嗯嗯 后面会好点

2018-09-21



辞酒轻尘

7

开始是因为要转行数据挖掘和人工智能这一块才想起来系统的学习下数据结构和算法，结果看了之后就着了迷。。。觉得算法真的非常有意思，现在完全是因为兴趣在学习

2018-09-21



_Breeze.

7

看了下目录，没有AVL、红黑树的内容吗？

2018-09-21

| 作者回复

二叉树那节会讲

2018-09-21



松鼠君

7

有一个机电学长和我说，学习数据结构和算法可以改变你的思维方式，我信了，然后学了。

2018-09-21



NoN

6

我个人觉得学习数据结构不说其他，真的非常锻炼一个人的思维能力，大学时候参加过学校的ACM集训队，学了一年多算法(上一届获过奖大牛学长带，基本属于纯自学)，我们班加上我有五个人，给我的感觉就是，凡是参加过集训的，学习能力都要比班里边其他人强，也是那时被算法吸引，由于是自学，有时候一个问题想不明白就得困扰很长时间，清晰记得当初

被动态规划折磨的日子，这次看到这个课程，就选择了，希望能跟着老师系统学习一下这门课！

2018-09-21



demo

👍 6

双非，不是CS专业的本科生，学校里边学的太少了。自己虽然有坚持刷leetcode，但是好多都是先看人家的思路才可以写出来的。自己一个自学，压力大，效率低吧。希望通过这个课程开阔自己的视野。

之后不管考研还是找工作，数据结构都是重中之重。所以 我一定要好好学。

2018-09-21



夏洛克的救赎

👍 6

精度问题举例：实现超出JS整数存储范围的两个大整数(含负数)相加。

补充说明：js计算超过2的53次方的数字计算精度都会有问题

2018-09-21

作者回复

这块我有经验 因为我是做金融开发的 精度尤为重要 我之前还写了一篇文章 专门说金融系统中精度问题如何解决 我想办法贴出来 你稍等几天

2018-09-21



建国

👍 5

老师您好，问您个问题，我们公司现在有一个问题已经持续几个月了，仍然没有找到解决办法：我们做考试系统的，问答题需要根据关键词判分，关键词是客户根据每道问答题录入的；目前的判分规则是先用分词工具根据关键词对答案进行分词生成一个集合，然后与关键词对比，由于分词是需要读写文件，但是每次都关闭了文件流，结果是：服务器刚重启完分词判分还管用，过一个礼拜不行了，线上日志显示linux服务器打开文件过多，至今不知道怎么优化，我想问的是分词有没有好的办法自己写一套而不用工具呢！

2018-09-23

作者回复

你的分词工具每次分词都要打开文件吗？我觉得应该是一次性加载资源文件，多次使用，这样的呀。是你们使用的方法不对吗？或者分词工具有bug吗？

2018-09-24



Chevins

👍 5

今天去面试。数据结构与算法。6道题。一道都不会。我说我3年开发经验，呵呵，自己都觉得可笑。不过无所谓，我会追上来的，加油，菜鸟。

2018-09-22



ZX

👍 5

老师，想问一下计算机基础除了数据结构和算法，还有哪些是非常重要的需要掌握的

2018-09-22

作者回复

操作系统 计算机网络 编译原理 这些是比较基础的计算机课程

2018-09-22



五岳寻仙

5

还想请教老师一个问题，对于前面提到的30亿键值对的问题，之所以倾向于线性存储，是因为这样的话可以直接把数据加载到共享内存中，多个程序都可以通过一个指针使用，省去了每次读数据到内存中的时间。

老师提到的散列，是不是需要基于链表，这样数据之间靠指针连接。这种结构可以持续放在共享内存中吗？

2018-09-21

作者回复

你用什么结构存储的这30亿条数据呢？二分查找需要顺序表结构存储的 30亿条数据占内存30g 需要连续的30g的内存空间 对内存要求是很高的。如果我现在数据量增大了 变成60亿条 120亿条 显然你的这个办法就不work了 你再想想该怎么办

2018-09-21



嘻哈

5

场景：给文章加锚链，最多匹配5个词语，优先匹配最长的词，文章根据后台配置的词和链接，将匹配到的词动态用a标签包含，并跳转到匹配词的链接上。

分析：如果循环关键词匹配，考虑如果词无限多的话，那么一次匹配太浪费资源，不利于匹配词的扩展。因此考虑下文的结构，并启动加载到内存中，同时将匹配词的结果放到redis缓存（1个小时），功能已上线，关键字2w,系统平稳，CPU表示毫无压力

将每个词拆成单个字，如"你好吗"数据结构是

```
{
  "depth": 0,
  "matcherMap": {
    "你": {
      "depth": 0,
      "matcherMap": {
        "你好": {
          "depth": 1,
          "matcherMap": {
            "你好吗": {
              "result": "xxx",
              "depth": 2,
              "matcherMap": {},
              "keyword": "你好吗"
            }
          }
        }
      }
    }
  }
}
```

```
}  
}
```

每一层对象都放到HashMap中，然后拿整篇文章逐个字从hashMap中匹配对象，此时每次截取的步长为1，如果单个字获能取到对象，将步长加1，继续匹配，通过查result判断是否查找到结果存到缓存中，下次遇到自己拿出匹配的结果，直接替换即可。

问题：

- 1.我的这种方式是否可行，有什么推荐的算法，有那些需要优化的地方
- 2.如果关键词是海量的，每次启动加载肯定特别慢，遇到这种场景我是否考虑将这些数据外置，如放到redis的Map中，还是可以有其他方式？

2018-09-21

作者回复

1. 你需要不要考虑分词呢？从你的描述感觉不需要。
2. 从实战的角度来说，2w的关键词，是可以直接放到内存中的，并不多，即便每个关键词用的你结构存储1KB的话，那也用不了20MB
3. 如果有太多关键词，内存都放不下，这种情况不大可能吧，你想想一个按照我们上面的假设推理，8GB内存可以放下800万的关键词，会有那么多吗？
4. 如果真的要放不下，我建议通过hash分片存储，多机器并行匹配说的很粗略，不知道你能听懂吗？
5. 还有，对于你的算法，从实战的角度，没有问题，但是还有一些解决思路：trie树、多模式串匹配等，这个要具体看你的关键词的特点。
6. 你要自己做些benchmark，然后再benchmark的基础上做优化，否则，再换个解决思路，你也不知道是不是更好了

2018-09-21



太空土豆

👍 5

我做了三年工程，教了10年数据结构，每次上课前都会有如坐针毡的感觉。怎样讲解基本结构与算法基本能挥洒自如，怎样与工程结合应用更好结合还在探索之中。看了你的课程列表，我想应该正是我需要的。谢谢你分享你的积累。

2018-09-21

作者回复

嗯嗯 感谢你的认可 我跟大学老师讲数据结构和算法的不同点就是我有大量的软件开发的实战经验 而大学老师往往连编程都不怎么会 当然这里没有鄙视的意思 术业有专攻而已

2018-09-21



Yayu

👍 4

为什么学已经不重要了，开始干货吧！

2018-09-22



依为

👍 4

我是大三的一名学生，现在在自学大数据和JAVA。其实老师说的很好，我更希望老师在讲解这些知识的时候，可以教我们一些学习方法，有时候觉得自己的学习方法不行，效率低下，有时看书一句话都读不懂，很无奈，希望老师授人以鱼不如授人以渔！

2018-09-21



熏墨染

👍 4

我竟然把所有留言看完了(˘ ˘˘)

王老师这篇开篇文很好，吊起了我想学好数据结构和算法的欲望。C++开发快3年了，但一直是个CRUD girl，最近决定考个第二专业（交大计算机），通过考试的学习顺便把这些基础知识夯实，但我知道算法和数据结构这一块不容易掌握，毕竟我自己之前也来来回回折腾过几次，所以这个课程来得很及时。

我也立个flag：每次更新后都要留言！

2018-09-21

| 作者回复

赞flag

2018-09-21



qpm

👍 4

大学里守恩师指点，沉下心写了半年的数据结构学习。现在毕业两年了，这部分的经历让我在工作中受益匪浅。数据结构是生活的抽象，算法是工作的流程，两者的原理就是数学，学习数据结构和算法，重要的一点是懂得抽象需求以契合到这些算法工具上，高效地解决问题。

数据结构，计算机网络，操作系统三者的融通非常重要，可以为我们技术人员观察计算机科学工程提供不可缺少的视觉。

在工作中引用过不少算法工具去高效地解决问题，不过还需要继续补充这方面的知识，立个flag，在这个专栏里留言覆盖四分之三的文章。

2018-09-21

| 作者回复

你这个flag很好 我02也讲到立flag的学习法 到时候再聊

2018-09-21



韦爷

👍 3

作为一个pm，除了基本的产品能力，多了解一下程序实现的算法更有助于自己规避一些触及算法极限的坑，也能避免被程序员忽悠这个或那个实现不了.....

2018-09-25



oday you

👍 3

工作多年却时常有这种感觉:代码实现很啰嗦 层层遍历，意识上觉得可以优化，却又看不出需要优化的地方；有时候就以为代码越少就越好性能；有时候所谓的优化代码就是从同样功能的api中选择一个好看、易用的，殊不知可能把代码性能变得更糟糕；有时候监控性能只是简单的在一段代码的首尾输出时间内存使用..... 这些问题发生在工作中往往影响别人对自己的正向评价，会少掉很多被委以重任的机会；而且如果基础不牢底层不懂的话那种感觉和兜里没钱挺不起腰板是一样样的，所以数据结构、算法认知的短板，概念缺失，这一块恶补一定是破局的关键。

2018-09-24





有钱才可以为所欲为

👍 3

关于业务接口99%响应时间那个。文章里说，如果有1200个数据，那么第1188个就是99%响应时间。这样说的话，那么这里的99%，就是一个类似下标的概念，特指某一个数据。而你在留言回复里面，又说这个99%响应时间，是为了反映接口的响应速度。我不是特别理解，百度谷歌也没找到有效解释，能具体讲一下么？谢谢了。

2018-09-23

作者回复

我后面讲到堆那节会讲的 如果你现在感兴趣 你可以谷歌下 percentile response time 或者 percentile vs average 或者百度下90%响应时间。注意 前面的90这两个数字可以换成任意数字。比如99 95 50

2018-09-23



xuan zhu

👍 3

年近40的大妈也有一颗积极向上的心

2018-09-22