

不定期福利第四期 | 刘超：我是怎么学习《数据结构与算法之美》的？

time.geekbang.org/column/article/75197



你好，我是刘超，是隔壁《趣谈网络协议》专栏的作者。今天来“串个门儿”，讲讲我学习《数据结构与算法之美》这个专栏的一些体会和感受。

《数据结构与算法之美》是目前“极客时间”订阅量最多的专栏，我也是其中最早购买的一员。我之所以一看就心动了，源于王争老师在开篇词里面说的那段话：

基础知识就像是一座大楼的地基，它决定了我们的技术高度。那技术人员究竟都需要修炼哪些“内功”呢？我觉得，无外乎就是大学里的那些基础课程，操作系统、计算机网络、编译原理等等，当然还有数据结构和算法。

这个也是我写《趣谈网络协议》的时候，在开篇词里反复强调的观点。我为什么这么说呢？因为，我们作为面试官，在招人的时候，往往发现，使用框架速成的人很多，基础知识扎实的人少见，而基础不扎实会影响你以后学习新技术的速度和职业发展的广度。

和“极客时间”编辑聊的时候，我也多次表达，希望我们讲的东西和一般的培训机构有所区别，希望“极客时间”能做真正对程序员的技能提升和职业发展有价值的内容，希望“极客时间”能够成为真正帮助程序员成长的助手。

所以，当“极客时间”相继推出《Java 核心技术 36 讲》《零基础学 Python》《从 0 开始学架构》《MySQL 实战 45 讲》这些课程的时候，我非常开心。我希望将来能够继续覆盖到编译原理、操作系统、计算机组成原理等等。在这些课程里，算法是基础的基础，也是我本人很想精进的部分。

当然，除了长远的职业发展需要，搞定算法还有一个看得见、摸得着的好处，面试。

我经常讲，越是薪资低的企业，面试的时候，它们往往越注重你会不会做网站，甚至会要求你现场做出个东西来。你注意了，这其实是在找代码熟练工。相反，越是薪资高的企业，越是重视考察基础知识。基础好，说明可塑性强，培养起来也比较快。而最牛的公司，考的往往是算法和思路。

相信很多购买《数据结构与算法之美》专栏的同学，下单的时候，已经想象自己面试的时候，在白板上挥洒代码，面试官频频点头的场景，想着自己马上就能“进驻牛公司，迎娶白富美”了。

然而，事实却是，**武功套路容易学，扎马步基本功难练**，编程也是一样。框架容易学，基本功难。你没办法讨巧，你要像郭靖学习降龙十八掌那样，一掌一掌劈下去才行。

于是，咱们这个专栏就开始了，你见到的仍然是困难的复杂度计算，指针指来指去，烧脑的逻辑，小心翼翼的边界条件判断。你发现，数据结构和算法好像并不是你上下班时间顺便听一听就能攻克的问题。你需要静下心来仔细想，拿个笔画一画，甚至要写一写代码，Debug 一下，才能够理解。是的，的确不轻松，那你坚持下来了吗？

我在这里分享一下我的学习思路，我将这个看起来困难的过程分成了几部分来完成。

第一部分，数据结构和算法的基础知识部分。如果在大学学过这门课，在专栏里，你会看到很多熟悉的描述。有些基础比较好的同学会质疑写这些知识的必要性。这大可不必，因为每个人的基础不一样，为了专栏内容的系统性和完整性，老师肯定要把这些基础知识重新讲述一遍的。对于这一部分内容，如果你的基础比较好，可以像学其他课程一样，在上下班或者午休的时候进行学习，主要是起到温习的作用。

第二部分，需要代码练习的部分。由于王争老师面试过很多人，所以在专栏里，他会列举一些他在面试中常常会问的题目。很多情况下，这些题目需要当场就能在白板上写出来。这些问题对于想要提升自己面试能力的同学来说，应该是很帮助的。

我这里列举几个，你可以看看，是不是都能回答出来呢？

- 在链表这一节：单链表反转，链表中环的检测，两个有序的链表合并，删除链表倒数第 n 个结点，求链表的中间结点等。
- 在栈这一节，在函数调用中的应用，在表达式求值中的应用，在括号匹配中的应用。
- 在排序这一节，如何在 $O(n)$ 的时间复杂度内查找一个无序数组中的第 K 大元素？

- 在二分查找这一节，二分查找的四个变体。

这些问题你都应该上手写写代码，或者在面试之前拿来练练手，而且，不仅仅只是实现主要功能。大公司的面试很多情况下都会考虑边界条件。只要被面试官抓住漏洞，就会被扣分，所以你最好事先写写。

第三部分，对于海量数据的处理思路问题。现在排名靠前的大公司，大都存在海量数据的处理问题。对于这一类问题，在面试的时候，也是经常会问到的。由于这类问题复杂度比较高，很少让当场就写代码，但是基本上会让你说一个思路，或者写写伪代码。想要解决海量数据的问题，你会的就不能只是基础的数据结构和算法了，你需要综合应用。如果平时没有想过这部分问题，临时被问，肯定会懵。

在专栏里，王争老师列举了大量这类问题，你要重点思考这类问题背后的思路，然后平时自己处理问题的时候，也多想想，如果这个问题数据量大的话，应该怎么办。这样多思考，面试的时候，思路很容易就来了。

比如，我这里随便列了几个，都是很经典的问题。你要是想不起来，就赶紧去复习吧！

- 比如说，我们有 10GB 的订单数据，我们希望按订单金额（假设金额都是正整数）进行排序，但是我们的内存有限，只有几百 MB，没办法一次性把 10GB 的数据都加载到内存中。这个时候该怎么办呢？
- 如果你所在的省有 50 万考生，如何通过成绩快速排序得出名次呢？
- 假设我们有 10 万个手机号码，希望将这 10 万个手机号码从小到大排序，你有什么比较快速的排序方法呢？
- 假设我们有 1000 万个整型数据，每个数据占 8 个字节，如何设计数据结构和算法，快速判断某个整数是否出现在这 1000 万数据中？我们希望这个功能不要占用太多的内存空间，最多不要超过 100MB，你会怎么做呢？

第四部分，工业实践部分。在每种数据结构的讲解中，老师会重点分析一些这些数据结构在工业上的实践，封装在库里面的，一般人不注意的。

我看王争老师也是个代码分析控。一般同学可能遇到问题，查一查有没有开源软件或者现成的库，可以用就完了。而王争老师会研究底层代码的实现，解析为什么这些在工业中大量使用的库，应该这样实现。这部分不但对于面试有帮助，对于实际开发也有很大的帮助。普通程序员和高手的差距，就是一个用完了就完了，一个用完了要看看为啥这样用。

例如，老师解析了 Glibc 中的 `qsort()` 函数，Java 中的 `HashMap` 如何实现工业级的散列表，Redis 中的有序集合（Sorted Set）的实现，工程上使用的红黑树等等。

尤其是对于哈希算法，老师解析了安全加密、数据校验、唯一标识、散列函数，负载均衡、数据分片、分布式存储等应用。如果你同时订阅了架构、微服务的课程，你会发现这些算法在目前最火的架构设计中，都有使用。

师傅领进门，修行在个人。尽管老师只是解析了其中一部分，但是咱们在平时使用开源软件和库的时候，也要多问个为什么。写完了程序，看看官方文档，看看原理解析的书，看看源代码，然后映射到算法与数据结构中，你会发现，这些知识和思路到处都在使用。

最后，我还想说一句，坚持，别放弃，啃下来。基础越扎实，路走得越远，走得越宽。加油！

 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。