

Pre-算法训练营-新百年

2018年6月 星期五 6:38

数据结构的新百年

****如约而至!**

第一批MOOC**四周年**。

——第一次线上直播。

不讲废话，直接开始~

&&对同学们的困难、解惑：

齐心协力。新百年的新教育。

- 年轻的力量，接下来如何发展？
 - 计算机科学之路漫长。无须考虑太多。知识层面分很多阶段。学完当前的知识，以后会自然而然有用。**踏实**下来。
 - 确认：① **兴趣**是否能保持、持久？② 比较系统的**数理基础**（离散数学、代数等）③ 计算机科学需要细致、**规范、系统**（面临大规模的问题，如何保持方法的持续有效）
- 关于课程：为什么强调C++？
 - 有点冤枉。语言上**没有很大的倾向**。计算机科学并不依赖于任何特定的语言[通用]，但在算法讲授的课程中，需要有一门具体的语言来实现。
 - 早年经历：**最开始讲课不用教材**——开放的学科，习惯于开书单、粗略的讲义。==> 后来做了调整，开始写教材。邓公的真正第一本数据结构的教材是用JAVA来写的【02年左右考察国外的学校，发现JAVA非常流行】。第一本(JAVA)其实不太好，形式上比较文绉绉，阅读难度较大。现在**为了国内、国情而转入C++**【c=>c++】。新书(C++)耗时整整七年，构思、动笔、调试、版面设计，耗费极大心血。邓公的特点：“**做事很精致**”。
 - 逐渐适应的过程。曾经收集语言上所不可或缺的内容，罗列书中那些语言知识是最必须的，只用到了很小很小的一部分，一篇很小的纸就能写完。虽然说是C++，但其实**真正用到的不多**。并且，讲课时会穿插其他语言(JAVA/Python/Perl)，作对比。
- 学习算法的基础
 - 至少学一门语言。基本的离散数学。
 - 不会有什么一定绕不过去的东西。中国的通识教育已经很发达了[比国外甚至还要强，没有什么硬伤]。泛泛地从数据结构的角度讲，完全没有问题：数据结构的学习有很多难度层次，“开车”，驾校[通用驾照]—>发烧友[DIY阶段]—>专业[设计/实现]。世界上没有万能的数据结构。算法已经发展到比较成熟的阶段了，现在的问题重心正在转移。
- 课本与讲义并非完全一致？
 - 教学需要分不同的层面。计算机系课堂/公选课堂/混合式课堂。统一教材、讲义，让不同的学生有不同的选择。注意**标记**。
- 依托MOOC资源，问题引导——如何设置和选择问题？
 - 问题来自于学生。每一届学生的关注点可能不一样。需要**建立反馈的渠道**。

- 与实际问题紧密相连。课堂上的讲课不如学生自己一点一点地实践。同时也需要注意变通、变化。
- 更加注重数学
 - 数据结构和数学的联系很高。凡是计算机科学做出大成就的人，在数学上都有很高的造诣。Knuth的书之所以称为Bible是因为书中的题目很经典，解决之后思路会方便很多。
 - 当初五年制，有2年的时间是跟数学系的同学一起学习。数学能告诉，除了什么能做之外，什么不能做[下界]。在任何一个特定的时期，总是会有一个极限；尽管整体上可能更乐观一些。“技术至上”？技术并不能解决一切问题。需要哲学、艺术……如果没有“下界”的意识，很可能会浪费时间——将宝贵的精力，放到有用的地方。
- 划重点？
 - 基本上都是重点。数据结构是一个整体。考试来说，可能会有知识点的选择。注重知识的继承、前驱。
- 算法导论、算法艺术等名书适合什么程度去看？
 - 到第二个级别时，“识优劣”，基本都需要看。比如：优化BiSearch到FibSearch。很多更细节、本质的知识，在这些书中会有很详细的介绍。
 - 注意到本门课的特点。名副其实的“图书”。理解不同算法的特点，优势和劣势===》需要反复阅读这些书。
- 大龄(男)青年无基础如何学？
 - 数据结构是一辈子的事情。基础上可以有所取舍。
- 以一位开课教师的角度，讲讲MOOC动机
 - 还有一门课。《计算几何》。邓公个人认为摄制的更好。
 - 国外的Coursera等，早年到中国游说，基本上反应平平。真正严肃的考虑，是03年，决定在MOOC上有所作为。明确表态想要做这件事，有出名的个人冲动，也有对个人责任、任务、使命的认知，在教育方面应该做一些事。
 - 事实与数字。事实：数据结构的教材大部分是过时的。数字：已经过了数十年。老师讲课的风格+**内容**为王。
 - 数据结构，实际上是潜能爆发的一个时期。保持完整性，比较实在。每学期1万注册人数的1%THU晋级，1% CST晋级。所有的教学资料基本都是共享的。
- 会开后续的课吗（更深入的课程）
 - 不太可能。基本的知识已经贯通了。
 - 怎么做都不可能做得完美。
 - 教材第4版，爽约了。身体原因。更考虑现代化的算法，包括竞赛、Online、系统缓存、大数据处理等。（研究气候，一个数据可以达到TB量级，数据的读取变得很重要）
- 如何检索课本中没有的数据结构？
 - 基本的骨架、子集要学会，方法和内涵要掌握。
 - 不要太在意什么东西没学，应该在意**学到了什么**，这时会有一种自信。
 - HandBook[推荐]或Wiki。
- GPU、量子计算机……
 - 量子计算机尚早。通用GPU，GPGPU……现在的数据结构与算法已经进入一个新的领域。GPU：单指令、多数据。软件与硬件的结合。流计算。不必

精确地获得所有的东西但本质上讲还未脱离图灵—>冯诺依曼体系[存储程序]。现在趋向于“**数据**”为重心，只是在基础上有所发展。“数字才是世界的本源”。

- 形式在变化，但万变不离其宗。
- 本科毕业，希望深造。想考THU
 - 看层面。**MOOC学好、学透**——是足够的。**【灵活运用，召之即来】**
 - 这几年有很多同学通过学习MOOC顺利进入清华计算机系研究生入学考试。专业课成绩也很好。完全可以这么走下去(MOOC)。
 - 四个方面：
 - ① 思维方式（严谨、套路==>计算思维!= 标新立异）
 - ② (实际)套路（知识点+解题方法，方法论[设计的套路，减治、分治、DP]）
 - ③ 综合运用能力（DIY）
 - ④ (没有现成轮子的基础上)造轮子
 - 对考研的帮助
 - 应该是有**帮助**的。只能说到这了。
- 一流的课程应该具有哪些特质
 - 获新百年成就奖时其实曾回答过。
 - 好的老师：风范。（学术基因）
 - 好的课程：**【严肃来讲】**有作业、有考试。应该是“**苦涩**”的。涩，就是不是很流畅，苦，有一些环节必须要过。享受这个过程。“茶”与白糖水。“回甘”。
 - “**精致的趣味**”。精致、趣[感觉]、味[体验]。
- 清华的情怀
 - **真诚**。不自欺，不相欺。
 - 人读书为了什么？

“士之读书治学，盖将以脱心志于俗谛之桎梏，真理因得以发扬。思想而不自由，毋宁死耳。”

 - ==》担当。

****什么是真实？你看到什么，听到什么，做什么和谁在一起，有一种从心灵深处漫溢出的不懊悔也不羞耻的平和与喜悦。 ——《无问西东》**

My Conclusion: **【Final】**

该算法课更面向于**能在短期内拿出足够时间**的同学。

历时5周，10节大课=20h，15节习题课=22.5h，41道习题≈80h。

内容充实。内容在第一周以后明显加强。但后期适应以后，会对整个算法设计实现能力的提升有极大的作用。

邓老师的讲课风格偏向于**联想、动态**。因此纯课件不适于复习强化，为了加强效率，最好能单独整理形成自己的笔记。并且，一定要复习，反复复习（实践也是复习的一种）。

算法训练营，“营”，一定要利用好学习的**环境、氛围**。充分利用合作学习的优势，积极分享好的想法，从同伴们的闪光点中得到激发。

入宝山不能空手而返！

人的潜能远远超乎想象。