

历年机试题大数据分析

历年题目及题型

年份	题目	类型
2008	素数	数学问题
	旋转矩阵	矩阵问题
	字符匹配	字符串问题
2009	立方根逼近	数学问题，公式计算
	数组排序	基本算法
	字符串的查找删除	字符串问题
2010	泰勒公式	数学问题，公式计算
	归并字符串	基本算法，归并去重
	数组是否相等	基本算法，多重集相等问题
2011	求孪生数	数学问题
	矩阵替换	矩阵问题
	字符串扩展	字符串问题
2012	某数分解成若干连续整数的和	数学问题
	小岛面积	矩阵问题
	统计关键字出现位置	字符串问题
2013	真分数约分	数学问题

	简单八皇后问题	矩阵问题
	科学计数法	字符串问题
2014	阶乘数	数学问题
	五子棋	矩阵问题
	排版	字符串问题
2015	相亲数	数学问题
	窗口模拟点击	简单模拟
	文章识别	基本算法、排序去重
2016	逆序数	数学问题
	enum 定义语句字符串解析	字符串问题
2017	词法分析，查找未定义变量	字符串问题
	找家谱成员	树的问题
2018	最长折线段	图的问题
	三叉树分支最多的节点高度	树的问题
2019	素数等差数列	数学问题
	三叉树最短路径	树的问题
2020	无	无

命题规律：题目数量方面，从 2008 至 2015 年，每年三道题，从 2016 至 2019 年，每年两道题。可见题量是 2-3 道，若是 3 道题，可能有送分题，若是 2 道题，则难度都比较大。题型方面，2008 至 2016 年，题型主要是一道数学问题，一道字符串问题，再加一道矩阵问题或者

基本算法问题，没有涉及图或者树问题。从 2017 年开始，每年都有一道树的问题，也出现了图的问题。题型归类：数学问题，字符串问题，矩阵问题和简单模拟问题，这几类可以归为经典题型，因为这些都是过去很长一段时间常考的，树和图问题是近来的趋势，因为最近三年开始出现并且每年必然出现，说明加大对树和图的考察是一个趋势。

题型频率统计

题型	出现频数	年份
数学问题	10	2008 素数、2009 立方根逼近、2010 泰勒公式、2011 求孪生数、2012 某数分解成若干连续整数的和、2012 真分数约分、2014 阶乘数、2016 逆序数、2019 素数等差数列
字符串问题	8	2008 字符匹配、2009 字符串的查找删除、2011 字符串扩展、2012 统计关键字出现的位置、2013 科学计数法、2014 排版、2016 enum 定义语句字符串解析、2017 词法分析查找未定义变量
矩阵问题	5	2008 旋转矩阵、2011 矩阵替换、2012 小岛面积、2013 简单八皇后问题、2014 五子棋
基本算法	4	2009 数组排序、2010 归并字符串、2010 数组是否相等、2015 文章识别
树问题	3	2017 找家谱成员、2018 三叉树分支最多的

		节点高度、2019 三叉树最短路径
图问题	1	2018 最长折线段
简单模拟	1	2015 窗口模拟点击

命题规律：数学问题占的比重最大，其次是字符串问题，这两类问题不涉及很深的算法，在效率要求不高的前提下，以朴素算法就可以解决，因此难点不在算法，而在于程序满足题意。矩阵问题和基本算法问题的频率次之，这些问题需要一定的算法，但通常是具体问题具体分析，没有特定规律；树问题近三年都考察了，可见是近来的大热，今年很可能再考。和树问题一样，图问题有许多经典算法，非常容易命题。简单模拟出现频率不高，因为这类题在理解了题意后比较简单，难点在于理解题意，用适当的数据结构进行模拟。

重点复习题型：

1. 树问题。不出意外今年也会出一题，涉及的算法比较困难，问题本身倒是不难理解。
2. 图问题。与树问题有亲缘关系，2018 年出现了树和图一起考的情况，今后不排除再次出现。而且图的算法比较经典，容易命题，而不懂算法又难以解题，所以复习的性价比较高。
3. 字符串问题。这类问题在历史上频率很高，而且本身具有很大难度，考察编码的细心，由于本身不考察特定的算法，因此题目可以很灵活，比如可以考察解析某种语法的字符串，JSON，HTML，URL 等。特别注意字符串的解析和在此基础上的问题，都是比较棘手的。
4. 数学问题。考虑到 2019 年才考过，不会这么快再考，但是此类问

题再历史上频率极高，通常作为有些难度的送分题。要熟悉各种常见数学公式、算法，如素数判定、GCD、同余定理等。有可能在常见算法的基础上加深难度，例如 2019 的素数等差数列。也有可能不涉及常见算法，如求孪生数，阶乘数，这类问题只有按题目来就比较好办了，近 3 年已经没有这么简单的题了。