

搜索顺序的选择

福州三中 王知昆

例：【间隔排列】问题

题意简述：有 $2N$ 个木块，每个木块标上 1 到 N 的自然数中的一个，每个数字会出现在两个木块上。把这些木块排成一排，要求是：标号为 i 的两个木块之间要间隔 i 个其它木块。比如说 $N=3$ 的情况，下面就是一个可行的排列：

3, 1, 2, 1, 3, 2。

编程实现，对给定的 $N(n \leq 40)$ ，求出一个可行排列。

运行效果比较

N	从大到小搜索	从小到大搜索
11	0.00 sec	0.00 sec
15	0.00 sec	0.11 sec
20	0.00 sec	36.32 sec
32	0.00 sec	超时
40	0.00 sec	超时

选择搜索顺序的基本原则

- 1、取值范围小的搜索元素先搜索。
- 2、一个搜索元素确定以后对其它搜索元素取值范围的影响称为约束力。约束力较大的搜索元素先搜索。
- 3、先搜索对解影响较大的元素可以使剪枝时的估价函数更准确，使剪枝更加有效。

例：【算符破译】（NOI2000）

题意简述：古梅算符由小写字母 a 到 m 组成，分别对应于现代算符中的 0,1,2,3,4,5,6,7,8,9,+,* 中的一个。给出一组古梅算符表示的等式，若存在满足等式的对应关系，则输出所有能够确定的古梅算符和现代算符的对应关系；否则输出 'noway'。

三个最特殊的元素

本题中有三个算符最特殊：'='、'*'、'+'，它们要满足以下条件：

- 1、这三个算符不能出现在等式的最左端和最右端。
- 2、这三个算符两两不能相邻。
- 3、'='，这是最特殊的算符，它在任何一个等式中必须出现且仅出现一次。

确定搜索顺序

从取值范围方面考虑， $'='$ ， $'+'$ ， $'*'$ 的取值范围在所有算符中是最小的；从约束力方面考虑， $'='$ 和 $'+'$ ， $'*'$ 的约束力无疑都强于 $'0'$ 到 $'9'$ 这十个数字；从对剪枝有利的角度考虑，这三个算符对解的影响最大，因此 $'='$ ， $'+'$ ， $'*'$ 这三个算符应当放在搜索序列的前面。对于这三个算符，由于 $'='$ 受到的限制更多，取值范围更小，所以应当优先搜索。

由此得出的最优搜索顺序：先搜索 $'='$ ，其次是 $'+'$ ， $'*'$ ，最后是 10 个数字。

减小搜索树规模的具体实现方法

1、静态优化搜索顺序

例【**质数方阵**】(IOI94),

【**修建栅栏**】(USACO TRAINING)

2、动态调整搜索顺序

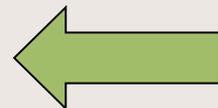
例【**棋盘遍历**】，

【**篮球锦标赛**】(BOI98)



静态优化搜索顺序

在一些问题中，搜索元素的约束力和取值范围在搜索过程中变化不大，或变化对搜索效率影响不大。如果要动态判断元素的取值范围和约束力需要花费较大的代价，而且优化效果不好。在这种情况下只需在搜索开始前确定搜索顺序，而不必在搜索过程中再改变搜索顺序。



动态调整搜索顺序

有时在搜索过程中元素的取值范围和约束力会有较大的变化，而且这些变化直接影响到搜索树的规模，因此需要动态的调整搜索顺序，也就是启发式搜索。启发式搜索继承了回溯法占用空间少，编程简单的优点，而启发式搜索的最大优点是可以在较短的时间内找到一组可行解，这最适合解决一类只需要求出一组可行解的搜索问题。



例：【质数方阵】（ IOI94 ）

题意简述：在一个 5×5 的方阵中填入数字，使得沿行，沿列及两个对角线的 5 个数字都能构成一个 5 位质数（5 位质数的首位不为 0）。所有质数的各位数字之和必须等于一个常数。这个常数和方阵左上角中的数字预先给出。若存在多个解，必须全部得出。

搜索元素的性质

- 1、最后一行和最后一列：可以填的数字只有 $\{1, 3, 7, 9\}$ 。
- 2、两条对角线：与方阵中的所有五位素数有关。
- 3、其他行列：特殊性取决于行列中已经确定的格子个数。

根据元素取值范围和制约力确定搜索顺序

- 1、最后一行和最后一列是取值范围最小的搜索元素，而且它们对其他所有的元素都有一定的制约力，因此要放在搜索序列的最前面。
- 2、两条对角线同样影响到其他所有的搜索元素，制约力在剩下的格子中是最大的，因此也应当优先搜索。
- 3、剩下的行列依据它们取值范围的大小确定搜索顺序。



例：【篮球锦标赛】（BOI98）

题意简述：有 5 支球队参加一次篮球锦标赛。比赛采用单循环，每两支球队比赛一次。已知每个队最终的获胜场次数，失败场次数，总得分以及总失分，并给出所有场次得分的列表，要求出可能的比赛的情况。（每一队每场比赛的得分组成的战绩表）。

动态调整搜索顺序的依据

搜索元素（即每场的比分）之间的制约力大小很难确定。

本题中，元素的取值范围定义为某一场中一个队的比分可以有几种可能的取值。由于比分的取值范围在搜索中会有较大的变化，而且这种变化直接影响到搜索树的规模，因此改变搜索顺序的主要依据就是每一个比分取值范围的大小。

动态判断元素取值范围的方法

首先，我们要进行一次预处理，求出所有 n 次得分的总和等于 m 的情况（可以用 HASH 表存储以提高检索效率）。在搜索的每一步中，根据已经填过的比分，我们可以判断出剩余的比分中哪些是不能填在某个位置的。

此外，在已知 i 和 j 比赛的输赢情况和 I 在这场比赛中得到的分数，则这场比赛中 j 对 i 赢的比分也要受到限制。

由此，可以确定出每一个元素的取值范围。

运行情况比较

数据	普通搜索算法 + 剪枝 (求一组解)	动态调整搜索顺序 (求一组解)	普通搜索算法 + 剪枝 (求所有解)	动态调整搜索顺序 (求所有解)
1	2.09 sec	0.00 sec	12.97 sec	0.11 sec
2	2.69 sec	0.00 sec	15.22 sec	0.00 sec
3	超时	0.16 sec	超时	0.11 sec
4	超时	0.05 sec	超时	0.27 sec
5	超时	0.27 sec	超时	1.10 sec



搜索顺序与剪枝优化

剪枝优化始终是搜索优化中最重要的一步，但剪枝优化的效果依赖于搜索顺序，只有正确选择对剪枝效果有益的搜索顺序，才能使剪枝充分发挥作用。

例、【生日蛋糕】（NOI99）

题意简述：问题要求设计一个 m 层的蛋糕，蛋糕的体积为 $N\pi$ ，蛋糕中每一层都是一个圆柱体，而且每一层的高度 H 和半径 R 均要比下一层的小。对于给定的 n 和 m ，求蛋糕最少可能的表面积（不包括最下一层的下底面）大小。

搜索顺序对剪枝的影响

由于 R 和 H 是从下到上递减，所以体积也是从下到上递减，因此选择从下到上的搜索顺序有利于判断当前情况是否可行。

此外，注意到表面积包括每层蛋糕的侧面积和裸露的顶面积。确定了最底层的蛋糕就确定了每层蛋糕裸露的顶面积之和，需要考虑的只剩下每层蛋糕的侧面积。因此这样的搜索顺序有利于判断是否可能出现比已知最优解更优的解。

小结

在实际解题和应用中，搜索问题总是离不开优化，而优化的第一步就是选择正确的搜索顺序。从上面的讨论中可以看出选择搜索顺序对搜索算法的效率起着重要作用。

但搜索顺序的选择仅依靠一些基本的方法和思想是不够的，它需要对问题进行全面、透彻的分析，深入挖掘问题的本质，找出解决问题的突破口。

谢谢！