一些搜索技巧

Claris

Hangzhou Dianzi University

2018 年 2 月 2 日

Overview

- 本节课主要介绍以下三种搜索方法：

Overview

- 本节课主要介绍以下三种搜索方法：
- 深度优先搜索 (DFS)。

Overview

- 本节课主要介绍以下三种搜索方法：
- 深度优先搜索 (DFS)。
- 广度优先搜索 (BFS)。

Overview

- 本节课主要介绍以下三种搜索方法：
- 深度优先搜索 (DFS)。
- 广度优先搜索 (BFS)。
- 双向搜索。

深度优先搜索

- 深度优先搜索是指递归地搜索。

深度优先搜索

- 深度优先搜索是指递归地搜索。
- 在递归的过程中，记录当前的状态，并加以适当的剪枝。

深度优先搜索

- 深度优先搜索是指递归地搜索。
- 在递归的过程中，记录当前的状态，并加以适当的剪枝。
- 实现起来较为方便，但是效率上一般不好。

深度优先搜索

- 深度优先搜索是指递归地搜索。
- 在递归的过程中，记录当前的状态，并加以适当的剪枝。
- 实现起来较为方便，但是效率上一般不好。
- 效率与遍历到的状态数个数有关，因此在用 DFS 之前需要分析状态数。

背包问题

给定 $n(n \leq 20)$ 个物品，每个物品有体积和价值，以及一个容量为 m 的背包。

背包问题

给定 $n(n \leq 20)$ 个物品，每个物品有体积和价值，以及一个容量为 m 的背包。

求最多能背走多少总价值的物品。

背包问题

给定 $n (n \leq 20)$ 个物品，每个物品有体积和价值，以及一个容量为 m 的背包。

求最多能背走多少总价值的物品。

- 记录当前考虑了前几个物品，选择物品的总体积和总价值。

背包问题

给定 $n(n \leq 20)$ 个物品，每个物品有体积和价值，以及一个容量为 m 的背包。

求最多能背走多少总价值的物品。

- 记录当前考虑了前几个物品，选择物品的总体积和总价值。
- 状态数：2^n。

全排列生成

生成 1 到 $n(n \leq 10)$ 的所有全排列。

全排列生成

生成 1 到 $n(n \le 10)$ 的所有全排列。

- 记录当前已经选了哪些数，然后枚举下一个要选择的数。

全排列生成

生成 1 到 $n(n \leq 10)$ 的所有全排列。

- 记录当前已经选了哪些数，然后枚举下一个要选择的数。
- 状态数：$n!$。

Consonant Fencity

本题中定义元音字母为:a,e,i,o,u,w,y，剩下的为辅音字母。

Consonant Fencity

本题中定义元音字母为:a,e,i,o,u,w,y，剩下的为辅音字母。

给定一个小写字符串 S，你需要给每种字母指定它是大写还是小写，使得 S 中相邻且大小写不同的辅音字母对数最多。

Consonant Fencity

本题中定义元音字母为:a,e,i,o,u,w,y，剩下的为辅音字母。

给定一个小写字符串 S，你需要给每种字母指定它是大写还是小写，使得 S 中相邻且大小写不同的辅音字母对数最多。

如：CoNsoNaNts 有 2 对：Ns、Nt；StRenGtH 有 5 对：St、tR、nG、Gt、tH。

Consonant Fencity

本题中定义元音字母为:a,e,i,o,u,w,y，剩下的为辅音字母。

给定一个小写字符串 S，你需要给每种字母指定它是大写还是小写，使得 S 中相邻且大小写不同的辅音字母对数最多。

如：CoNsoNaNts 有 2 对：Ns、Nt；StRenGtH 有 5 对：St、tR、nG、Gt、tH。

- $1 \leq |S| \leq 10^6$。

Consonant Fencity

本题中定义元音字母为:a,e,i,o,u,w,y，剩下的为辅音字母。

给定一个小写字符串 S，你需要给每种字母指定它是大写还是小写，使得 S 中相邻且大小写不同的辅音字母对数最多。

如：CoNsoNaNts 有 2 对：Ns、Nt；StRenGtH 有 5 对：St、tR、nG、Gt、tH。

- $1 \leq |S| \leq 10^6$。
- Source：NEERC 2017 Northern Subregional Contest

Solution

- 首先统计出 $f_{x,y}$ 表示 x 之后紧挨着 y 的对数，那么 S 就没有用了。

Solution

- 首先统计出 $f_{x,y}$ 表示 x 之后紧挨着 y 的对数，那么 S 就没有用了。
- 直接搜索所有字母的大小写情况？

Solution

- 首先统计出 $f_{x,y}$ 表示 x 之后紧挨着 y 的对数，那么 S 就没有用了。
- 直接搜索所有字母的大小写情况？
- 状态数 $= 2^{26}$，时间复杂度 $O(2^{26}26^2)$，不能承受。

Solution

- 首先统计出 $f_{x,y}$ 表示 x 之后紧挨着 y 的对数，那么 S 就没有用了。
- 直接搜索所有字母的大小写情况？
- 状态数 $= 2^{26}$，时间复杂度 $O(2^{26}26^2)$，不能承受。
- 注意到 7 个元音字母的大小写是没有用的，只搜索辅音字母即可。

Solution

- 首先统计出 $f_{x,y}$ 表示 x 之后紧挨着 y 的对数，那么 S 就没有用了。
- 直接搜索所有字母的大小写情况？
- 状态数 $= 2^{26}$，时间复杂度 $O(2^{26}26^2)$，不能承受。
- 注意到 7 个元音字母的大小写是没有用的，只搜索辅音字母即可。
- 状态数 $= 2^{19}$，时间复杂度 $O(2^{19}19^2)$。

点覆盖计数

给定一个 n 个点 m 条边的简单无向图，你需要选择其中的一些点，满足：

点覆盖计数

给定一个 n 个点 m 条边的简单无向图，你需要选择其中的一些点，满足：

对于任意一条边 (u, v)，u 和 v 中至少有一个点被选中。

点覆盖计数

给定一个 n 个点 m 条边的简单无向图，你需要选择其中的一些点，满足：

对于任意一条边 (u, v)，u 和 v 中至少有一个点被选中。

求可行的方案数。

点覆盖计数

给定一个 n 个点 m 条边的简单无向图，你需要选择其中的一些点，满足：

对于任意一条边 (u, v)，u 和 v 中至少有一个点被选中。

求可行的方案数。

- $1 \leq n \leq 40$。

点覆盖计数

给定一个 n 个点 m 条边的简单无向图，你需要选择其中的一些点，满足：

对于任意一条边 (u, v)，u 和 v 中至少有一个点被选中。

求可行的方案数。

- $1 \leq n \leq 40$。
- $0 \leq m \leq \frac{n(n-1)}{2}$。

Solution

- 从 1 到 n 依次考虑每个点。

Solution

- 从 1 到 n 依次考虑每个点。
- 对于当前考虑的点 i，如果 i 度数为 0，那么它选不选都不会对其它点造成影响，答案直接乘 2。

Solution

- 从 1 到 n 依次考虑每个点。
- 对于当前考虑的点 i，如果 i 度数为 0，那么它选不选都不会对其它点造成影响，答案直接乘 2。
- 否则 i 度数不为 0，说明至少与一个点相连，假设现在还有 n 个点。

Solution

- 从 1 到 n 依次考虑每个点。
- 对于当前考虑的点 i，如果 i 度数为 0，那么它选不选都不会对其它点造成影响，答案直接乘 2。
- 否则 i 度数不为 0，说明至少与一个点相连，假设现在还有 n 个点。
- 若 i 选，则还要考虑 $n-1$ 个点。

Solution

- 从 1 到 n 依次考虑每个点。
- 对于当前考虑的点 i，如果 i 度数为 0，那么它选不选都不会对其它点造成影响，答案直接乘 2。
- 否则 i 度数不为 0，说明至少与一个点相连，假设现在还有 n 个点。
- 若 i 选，则还要考虑 $n-1$ 个点。
- 若 i 不选，那么与 i 相连的点必选，最多还要考虑 $n-2$ 个点。

Solution

- 从 1 到 n 依次考虑每个点。
- 对于当前考虑的点 i，如果 i 度数为 0，那么它选不选都不会对其它点造成影响，答案直接乘 2。
- 否则 i 度数不为 0，说明至少与一个点相连，假设现在还有 n 个点。
- 若 i 选，则还要考虑 $n-1$ 个点。
- 若 i 不选，那么与 i 相连的点必选，最多还要考虑 $n-2$ 个点。
- 时间复杂度 $T(n) = T(n-1) + T(n-2) = Fib(n)$。

广度优先搜索

- 广度优先搜索利用队列来扩展状态。

广度优先搜索

- 广度优先搜索利用队列来扩展状态。
- 在搜索的过程中，加上判重来使得每个状态最多被扩展一次。

广度优先搜索

- 广度优先搜索利用队列来扩展状态。
- 在搜索的过程中，加上判重来使得每个状态最多被扩展一次。
- 同时因为队列的性质保证了到每个状态都是最小扩展次数，可以求最短路。

Distinctive Character

给定 n 个长度都为 k 的 01 串。

Distinctive Character

给定 n 个长度都为 k 的 01 串。

定义两个 01 串的相似度为它们对应位置字符相同的位置个数。

Distinctive Character

给定 n 个长度都为 k 的 01 串。

定义两个 01 串的相似度为它们对应位置字符相同的位置个数。

如：001110 和011011 的相似度为 3。请找到一个长度为 k 的 01 串，使得它到这 n 个串的相似度的最大值最小。

Distinctive Character

给定 n 个长度都为 k 的 01 串。

定义两个 01 串的相似度为它们对应位置字符相同的位置个数。

如：001110 和011011 的相似度为 3。请找到一个长度为 k 的 01 串，使得它到这 n 个串的相似度的最大值最小。

- $1 \le n \le 100000$。

Distinctive Character

给定 n 个长度都为 k 的 01 串。

定义两个 01 串的相似度为它们对应位置字符相同的位置个数。

如：001110 和011011 的相似度为 3。请找到一个长度为 k 的 01 串，使得它到这 n 个串的相似度的最大值最小。

- $1 \leq n \leq 100000$。
- $1 \leq k \leq 20$。

Distinctive Character

给定 n 个长度都为 k 的 01 串。

定义两个 01 串的相似度为它们对应位置字符相同的位置个数。

如：001110 和011011 的相似度为 3。请找到一个长度为 k 的 01 串，使得它到这 n 个串的相似度的最大值最小。

- $1 \leq n \leq 100000$。
- $1 \leq k \leq 20$。
- Source：NCPC 2017

Solution

- 相似度 $= k-$ 对应位置字符不同的位置个数。

Solution

- 相似度 $= k-$ 对应位置字符不同的位置个数。
- 定义对应位置字符不同的位置个数为两个串的“距离”。

Solution

- 相似度 $= k-$ 对应位置字符不同的位置个数。
- 定义对应位置字符不同的位置个数为两个串的“距离”。
- 相似度最大 $=$ 距离最小。

Solution

- 相似度 $= k-$ 对应位置字符不同的位置个数。
- 定义对应位置字符不同的位置个数为两个串的“距离”。
- 相似度最大 $=$ 距离最小。
- 那么求出离每个串距离最小的串就是一个多源最短路，BFS 即可。

Solution

- 相似度 $= k-$ 对应位置字符不同的位置个数。
- 定义对应位置字符不同的位置个数为两个串的“距离”。
- 相似度最大 $=$ 距离最小。
- 那么求出离每个串距离最小的串就是一个多源最短路，BFS 即可。
- 时间复杂度 $O(nk+2^k k)$。

双向搜索

- 双向搜索 (Meet in the Middle) 是指同时从起点终点开始搜索，在中途相遇。

双向搜索

- 双向搜索 (Meet in the Middle) 是指同时从起点终点开始搜索，在中途相遇。
- 如此一来复杂度可以开根号。

Ice Hockey World Championship

有 n 个物品，m 块钱，给定每个物品的价格，求买物品的方案数。

Ice Hockey World Championship

有 n 个物品，m 块钱，给定每个物品的价格，求买物品的方案数。

- $1 \leq n \leq 40$。

Ice Hockey World Championship

有 n 个物品，m 块钱，给定每个物品的价格，求买物品的方案数。

- $1 \le n \le 40$。
- $1 \le m \le 10^{16}$。

Ice Hockey World Championship

有 n 个物品，m 块钱，给定每个物品的价格，求买物品的方案数。

- $1 \le n \le 40$。
- $1 \le m \le 10^{16}$。
- Source：CEOI 2015

Solution

- 将 n 个物品平均分成前 $\frac{n}{2}$ 个和后 $\frac{n}{2}$ 个。

Solution

- 将 n 个物品平均分成前 $\frac{n}{2}$ 个和后 $\frac{n}{2}$ 个。
- 分别搜索出前后两部分所有可能的情况，每部分共 $2^{\frac{n}{2}}$ 种。

Solution

- 将 n 个物品平均分成前 $\frac{n}{2}$ 个和后 $\frac{n}{2}$ 个。
- 分别搜索出前后两部分所有可能的情况，每部分共 $2^{\frac{n}{2}}$ 种。
- 对于两部分，按总价格从小到大排序，双指针求出方案数。

Solution

- 将 n 个物品平均分成前 $\frac{n}{2}$ 个和后 $\frac{n}{2}$ 个。
- 分别搜索出前后两部分所有可能的情况，每部分共 $2^{\frac{n}{2}}$ 种。
- 对于两部分，按总价格从小到大排序，双指针求出方案数。
- 时间复杂度 $O(n2^{\frac{n}{2}})$。

Travelling to Random Cities

$n = 100000$ 个点，$m = 300000$ 条边的随机简单无向图，任意一条边长度都是 1。

Travelling to Random Cities

$n = 100000$ 个点，$m = 300000$ 条边的随机简单无向图，任意一条边长度都是 1。

k 次询问，每次随机指定起点和终点，求两点间的最短路。

Travelling to Random Cities

$n = 100000$ 个点，$m = 300000$ 条边的随机简单无向图，任意一条边长度都是 1。

k 次询问，每次随机指定起点和终点，求两点间的最短路。

- $1 \le k \le 10000$。

Travelling to Random Cities

$n = 100000$ 个点，$m = 300000$ 条边的随机简单无向图，任意一条边长度都是 1。

k 次询问，每次随机指定起点和终点，求两点间的最短路。

- $1 \leq k \leq 10000$。
- Source：XVII Open Cup named after E.V. Pankratiev. GP of Two Capitals

Solution

- 因为 $n = 100000, m = 300000$，且图随机，那么每个点平均连了 6 条边，且两点间最短路一般不会超过 8。

Solution

- 因为 $n = 100000, m = 300000$，且图随机，那么每个点平均连了 6 条边，且两点间最短路一般不会超过 8。
- 对于每个询问 S, T，若两个点不连通，那么显然无解，可以 $O(1)$ 判断。

Solution

- 因为 $n = 100000, m = 300000$，且图随机，那么每个点平均连了 6 条边，且两点间最短路一般不会超过 8。
- 对于每个询问 S, T，若两个点不连通，那么显然无解，可以 $O(1)$ 判断。
- 从 S 和 T 分别爆搜 4 步，那么除去最短路过来的那条边，平均还剩 5 条边。

Solution

- 因为 $n = 100000, m = 300000$，且图随机，那么每个点平均连了 6 条边，且两点间最短路一般不会超过 8。
- 对于每个询问 S, T，若两个点不连通，那么显然无解，可以 $O(1)$ 判断。
- 从 S 和 T 分别爆搜 4 步，那么除去最短路过来的那条边，平均还剩 5 条边。
- 如此可以得到所有不超过 8 的答案，时间复杂度 $O(6 \times 5^3)$。

Solution

- 因为 $n = 100000, m = 300000$，且图随机，那么每个点平均连了 6 条边，且两点间最短路一般不会超过 8。
- 对于每个询问 S, T，若两个点不连通，那么显然无解，可以 $O(1)$ 判断。
- 从 S 和 T 分别爆搜 4 步，那么除去最短路过来的那条边，平均还剩 5 条边。
- 如此可以得到所有不超过 8 的答案，时间复杂度 $O(6 \times 5^3)$。
- 若此时还得不到答案，那么答案超过 8，是小概率事件。

Solution

- 因为 $n = 100000, m = 300000$，且图随机，那么每个点平均连了 6 条边，且两点间最短路一般不会超过 8。
- 对于每个询问 S, T，若两个点不连通，那么显然无解，可以 $O(1)$ 判断。
- 从 S 和 T 分别爆搜 4 步，那么除去最短路过来的那条边，平均还剩 5 条边。
- 如此可以得到所有不超过 8 的答案，时间复杂度 $O(6 \times 5^3)$。
- 若此时还得不到答案，那么答案超过 8，是小概率事件。
- 直接 $O(n+m)$BFS 整张图即可。

题目提交

课上例题：

http://acm.hdu.edu.cn/diy/contest_show.php?cid=33128

课后习题：

http://acm.hdu.edu.cn/diy/contest_show.php?cid=33129

密码：

G*&GSF&*t387tr

Thank you!