CTT2024 Altar 2

Itst & asmend

2024 年 11 月 28 日

简要题意

交互题。给定一个 n 个点的竞赛图，使用尽可能少的边方向询问，找到一个点，使得其到达其他所有点的最短路最大值最小。

交互库不适应，300 组测试数据，$n \leq 300$。按照询问次数给分，平均不超过 700 次满分。

吐槽

TBA

子任务 1：边随机定向

返回任意一个点即可。

子任务 2：存在一个入度为 0 的点

找到这个点就行。每次询问两个目前入度为 0 的点之间的边的方向，可以删掉一个点。

~~这俩子任务好像都没啥用。~~

最短路最大值 $\min \leq 2$

我们先证明如下事实：存在一个点到其他点的距离都不超过 2。

更具体地，我们证明图上出度最大的点满足这个条件。

最短路最大值 min ≤ 2

我们先证明如下事实：存在一个点到其他点的距离都不超过 2。

更具体地，我们证明图上出度最大的点满足这个条件。

设 x 为出度最大的点，$N^+(x)$ 为 x 的出点。

若存在 x 无法两步到达的点 y，那么 $N^+(y)$ 既包含 x 又包含 $N^+(x)$，与 $|N^+(x)|$ 最大矛盾。

部分分做法 By asmend

度数最大的点一定是合法的答案，我们尝试找度数最大的点。

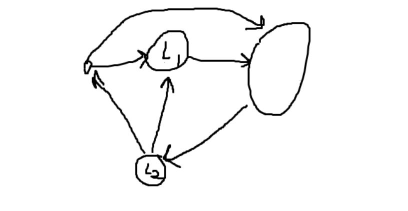
询问过程中，对每个点 x 维护 c_x 表示与 x 相邻的已知方向的边数（包括 x 自己，初始时 $c_x = 1$），s_x 表示这些边中起点是 x 的边数。直观上，$\frac{s_x}{c_x}$ 更大的点比较可能是度数最大的点。

每次考虑 $\frac{s_x}{c_x}$ 最大的点 x，询问 x 和 $\frac{s_x}{c_x}$ 次大的常数个点之间的边，如果这些边都已经被询问过了，就改为询问 x 和所有其它点之间的边。

设定一个阈值 C，在询问 C 次之后停止这个过程，输出所有 $c_x = n$ 的 x 中 s_x 最大的。

叉掉部分分做法 By asmend

想叉掉估计度数最大值的做法，只需要构造一种图，除了最大度点以外，其他大度数点都不满足条件就行。考虑以下构造。



其中三个圈表示三个边随机定向的团，L_1 和 L_2 的大小比较小。此时度数最大的是左边的单点，而 L_1 里的点尽管度数只比最大度点少大约 $\frac{L_1}{2}$，但不是满足条件的点。

实际生成的数据里 $|L_1|$ 和 $|L_2|$ 在满足条件的范围里均匀随机，上面做法的阈值 C 得大于 6000 才能通过该测试点。取 $C = 7000$ 结合前两个子任务约获得 52 分。

换种证法

我们换一种证法证明存在一个点到其他点的距离都不超过 2。

对 n 归纳。给定一个图，每次任选一个点 x。若 $N^{-}(x)$ 为空，那么 x 满足条件，否则由归纳假设，$N^{-}(x)$ 的导出子图里有一个点 y 到 $N^{-}(x)$ 的所有点最短路不超过 2。

而 y 到 x 的最短路为 1，y 到 $N^{+}(x)$ 的最短路也就不超过 2，因此 y 就是原图上的满足条件的点。

换种证法

我们换一种证法证明存在一个点到其他点的距离都不超过 2。

对 n 归纳。给定一个图，每次任选一个点 x。若 $N^{-}(x)$ 为空，那么 x 满足条件，否则由归纳假设，$N^{-}(x)$ 的导出子图里有一个点 y 到 $N^{-}(x)$ 的所有点最短路不超过 2。

而 y 到 x 的最短路为 1，y 到 $N^{+}(x)$ 的最短路也就不超过 2，因此 y 就是原图上的满足条件的点。

注意到以上归纳证明直接给了我们一个算法：任选一个点，把它的入点们递归下去。由于不适应，可以随机选一个点。

另外，若存在一个点入度为 0，这个点也一定是这个算法最终找到的点，所以其总是能找到满足题设条件的点，正确性可以保证。因此接下来我们分析以上算法的询问次数期望。

询问次数期望

直觉上来看，随机选一个点会让递归下去的图的大小减半，所以询问次数应该是 $2n$ 的。

询问次数期望

直觉上来看，随机选一个点会让递归下去的图的大小减半，所以询问次数应该是 $2n$ 的。

我们接下来对 $|G|$ 归纳，严谨证明任意一个 n 个点的图 G 的询问次数期望 $\mathbb{E}[Q(G)]$ 不超过 $2n$。

算法随机选图上的一个点，使用 $(n-1)$ 次询问划分为 n 个子问题 $G_1, G_2, \cdots, G_n$，它们的大小和为 $\binom{n}{2}$，且大小均小于 n。

由归纳假设，我们有

$$\begin{aligned}
E[Q(G)] &= (n-1) + \frac{1}{n}\sum \mathbb{E}(G_i) \\
&\le (n-1) + \frac{1}{n}\sum 2|G_i| \\
&= (n-1) + \frac{1}{n} \times 2 \times \binom{n}{2} \\
&= (n-1) + (n-1) = 2n - 2 \le 2n.
\end{aligned}$$

询问次数二阶矩*

选手只需要写出一个期望 $2n$ 的算法并通过本题就可以了，而出题人要想的就很多了。因为只能做有限次随机，为了搞清楚 $2n+\epsilon$ 的 ϵ 到底得多大，我们还需要分析询问次数的集中性。

我们使用切比雪夫不等式

$$\Pr[X - \mathbb{E}[X] > a\mathsf{Std}[X]] \leq \frac{1}{a^2}, \tag{1}$$

其中 $\mathsf{Std}[X]$ 是 X 的标准差。由 $\mathsf{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$，我们算出 $Q(G)$ 的二阶矩就可以了。

询问次数二阶矩*

我们再使用归纳证明 $\mathbb{E}[Q(G)^2] \le cn^2 + dn$，过程和上面对期望的证明基本一致。

$$
\begin{aligned}
\mathbb{E}[Q(G)^2] &= \mathbb{E}\left[\left((n-1) + \frac{1}{n}\sum_i Q(G_i)\right)^2\right] \\
&= (n-1)^2 + \frac{2(n-1)}{n}\mathbb{E}\left[\sum_i Q(G_i)\right] + \frac{1}{n^2}\mathbb{E}\left[\left(\sum_i Q(G_i)\right)^2\right] \\
&\le 3n^2 + \frac{1}{n^2}\left(\sum_i \mathbb{E}[Q(G_i)^2] + \sum_{i\neq j}\mathbb{E}[Q(G_i)]\mathbb{E}[Q(G_j)]\right) \\
&\le 3n^2 + \frac{1}{n^2}\left(\sum_i c|G_i|^2 + d|G_i| + 2|G_i|(n(n-1) - 2|G_i|)\right)
\end{aligned}
$$

询问次数二阶矩*

注意到 $\frac{1}{n^2}\sum_i (c-4)|G_i|^2 \le (c-4)n$，因为 $|G_i| \le n$ 且 $\sum |G_i| \le \binom{n}{2}$。因此

$$\begin{aligned}\mathbb{E}[Q(G)^2] &\le 3n^2 + \frac{1}{n^2}\left(\sum_i 2|G_i|n(n-1)\right) + (c-4)n + \frac{d}{2} \\ &\le 4n^2 + (c-3)n.\end{aligned}$$

取 $c=4$，$d=1$ 即可。

询问次数二阶矩*

注意到 $\frac{1}{n^2}\sum_i (c-4)|G_i|^2 \le (c-4)n$，因为 $|G_i| \le n$ 且 $\sum |G_i| \le \binom{n}{2}$。因此

$$\begin{aligned}\mathbb{E}[Q(G)^2] &\le 3n^2 + \frac{1}{n^2}\left(\sum_i 2|G_i|n(n-1)\right) + (c-4)n + \frac{d}{2} \\ &\le 4n^2 + (c-3)n.\end{aligned}$$

取 $c=4$，$d=1$ 即可。之前的期望归纳还可以证明 $\mathbb{E}[Q(G)] \ge 2n - \log_2 n$，故 $\mathrm{Var}[Q(G)] = 4n\log n + O(n)$，标准差这么算下来刚好是 100。

另外 $\mathbb{E}[G_i^2] \le 4|G_i|^2 + |G_i|$ 在和式中的贡献是 $O(n)$ 的，所以 G_i 怎么分，即图的结构，对 $E[Q(G)^2]$ 只有小量影响。

跑 300 次平均，方差变为 $\frac{1}{300}$，所以单组错误概率不超过 $\frac{1}{300}$，合并错误概率为 5%，实际跑概率还小一点，~~根本跑不出大于 600 的。~~

完结撒花