

GRAPHCNT 解题报告

谷晟

1 题目大意

给定一个 N 个点（从 1 到 N 标号） M 条边的有向图。请你统计无序对 (X, Y) 的个数，其中 (X, Y) 满足存在一条从点 1 到点 X 的路径，和一条从点 1 到点 Y 的路径，且两条路径除了点 1 以外没有公共点。

- $1 \leq n \leq 100000$
- $0 \leq m \leq 500000$
- 数据集 1（30 分）：图是一个有向无环图（DAG）
- 数据集 2（20 分）： $N \times M \leq 5 \times 10^7$
- 数据集 3（50 分）：无其他限制

2 算法分析

2.1 基本概念

要解决这个问题，先要了解几个概念。

对于一个有向图，如果从给定的起点到某个节点 X 的所有路径都经过节点 Y ，我们称 Y 是 X 的必经点（dominator）。这些必经点之间存在层次关系，我们把 X 的必经点中离起点最远（离 X 最近）的必经点称为 X 的最近必经点（immediate dominator），记为 $idom[X] = Y$ 。最近必经点是唯一的。将除起点外所有节点与最近必经点连接构成一棵树，称为支配树（Dominator Tree）。

存在从点 1 到点 X 和 Y 的路径，且两条路径没有公共点，说明 X 和 Y 的公共必经点只有 1。如果已经构建出了支配树，只要将根节点的所有子树的大小两两相乘并累加即是满足条件的无序对个数。因此本题的重点在于如何快速求出给定有向图的支配树。

2.2 DAG 上的算法

DAG 上的节点有严格的拓扑序，即只能从一个拓扑序靠前的点到达一个拓扑序靠后的点。在这种情况下，某个节点的最近必经点一定是拓扑序更小的点，我们可以按照拓扑序增量构造支配树。

假设已经构造出了拓扑序为 $1 \sim X-1$ 的节点组成的支配树，现在要加入节点 X 。考虑所有能直接到达节点 X 的节点，这些节点的最近公共祖先（LCA）就是节点 X 的最近必经点，据此将节点 X 加入支配树中。

LCA 使用倍增法求，时间复杂度 $O((N+M) \log N)$ ，能得到 30 分。

2.3 NM 较小时的算法

上述算法不能在有环图上工作。考虑一种通用算法。每次删除一个节点，从起点开始 DFS，如果能到达某个节点，则被删除的点不是这个节点的必经点。重复对每个节点进行删除尝试，可以得到图中所有点之间的必经关系。

忽略所有不可达点，时间复杂度 $O(NM)$ ，能得到 20 分。结合 DAG 上的算法可以得到 50 分。

2.4 快速算法

对于 NM 比较大的图，可以使用 Lengauer-Tarjan 算法 [1] 计算支配树。Lengauer-Tarjan 算法分为以下三个步骤 [2]：

1. 对图进行 DFS，计算出搜索树及其 DFS 序。
2. 根据搜索树计算每个节点的半必经点，作为计算最近必经点的依据。
3. 对半必经点进行修正，计算出每个节点的最近必经点。

2.4.1 半必经点定理

对于一个节点 Y ，存在祖先 X 能通过一系列节点 p_i （可以为 0 个）到达节点 Y ，且 $\forall i, dfn[p_i] > dfn[Y]$ ，我们称节点 X 是节点 Y 的半必经点 (semidominator)，记为 $semi[Y] = X$ 。

半必经点可以通过半必经点定理计算。半必经点定理描述如下：

- 对于一个节点 Y ，考虑所有能直接到达它的节点，设其中一个为 X ；
- 若 $dfn[X] < dfn[Y]$ ， X 为可能的半必经点；
- 若 $dfn[X] > dfn[Y]$ ，此时 X 在搜索树中的祖先 Z ，满足 $dfn[Z] > dfn[Y]$ 时， $semi[Z]$ 为可能的半必经点；
- 以上所有可能的节点中，DFS 时间戳 (dfn) 最小的节点为 Y 的半必经点。

2.4.2 必经点定理

半必经点不一定是必经点，需要通过必经点定理修正。必经点定理描述如下：

- 考虑节点 X 与其半必经点 $semi[X]$ 的路径上的其它节点，记其中半必经点时间戳最小的节点为 Z ；
- 如果 $semi[Z] = semi[X]$ ，则 $idom[X] = semi[X]$ ；
- 如果 $semi[Z] \neq semi[X]$ ，则 $idom[X] = idom[Z]$ 。

2.4.3 实现

我们注意到，在计算半必经点的过程中如果某个节点的前驱节点的时间戳比它大，需要考察这个前驱节点的所有时间戳比当前考虑节点大的祖先。计算最近必经点的过程中也要考虑它和半必经点之间的路径上的所有节点。我们可以按照时间戳逆序计算每个节点的半必经点，此时所有时间戳更小的节点仍未计算，在考察时间戳大的祖先时只要考察它在已计算节点森林中的所有祖先。维护已计算节点森林可以用并查集¹来实现。

¹不能按秩合并，只路径压缩

时间复杂度 $O((N + M) \log N)$ ，空间复杂度 $O(N + M)$ ，能得到本题所有分数。

参考资料

- [1] Thomas Lengauer. A fast algorithm for finding dominators in a flowgraph.
- [2] 李煜东. 图连通性若干拓展问题探讨.