arXiv:2305.05205v1 [cs.DM] 9 May 2023

R生成任务依赖图的随机过程 S

Jesse Geneson 和 Shen-Fu Tsai

摘要

我们研究了生成阶为 n 、具有 m条边以及指定数量初始顶点和终端顶点的任务依赖图的随机过程。为此，我们考虑了两种可结合起来完成该任务的任务依赖图生成随机过程。在（x， y）边删除过程中，我们从一个最大连通任务依赖图开始，只要删除边不会导致初始顶点数超过 x 或终端顶点数超过 y，就以均匀随机方式删除边。在（x， y）边添加过程中，我们从一个空任务依赖图开始，只要添加边不会导致初始顶点数少于x 或终端顶点数少于 y，就以均匀随机方式添加边。在（x， y）边添加过程中，如果恰好有 x个初始顶点和 y 个终端顶点，我们就停止。对于这两种过程，我们确定了使所得任务依赖图必然恰好具有 x 个初始顶点和 y 个终端顶点的 x 和 y 的取值；此外，我们还找到了所得任务依赖图中边数关于 x、 y和顶点数的极值。进一步地，我们渐近地界定了所得任务依赖图中边的期望数量。最后，我们定义了一个仅使用边添加和边删除的随机过程，并证明该随机过程以高概率生成一个阶为 n 、具有 m 条边的（x， y）任务依赖图。

1. 引言

任务依赖图是一种有向无环图，其中顶点表示任务，而边（a,b）表示任务 b 的完成取决于任务 a 的完成。任务依赖图是设计并行程序的有用工具。每个任务代表程序中由程序员指定的一部分，任务依赖图 G 则描述了程序各部分之间的相互依赖关系。理解 G 的结构有助于在多个处理器上调度相应程序的各个部分，并估算总运行时间的上限。针对使用多个处理器调度任务依赖图中的任务，人们已提出多种范式，包括工作共享 [5, 9, 19, 23], 、工作窃取 [6, 15, 3, 4, 11, 12, 14, 16, 18, 20, 21, 24], 以及并行深度优先调度 [1, 2, 7]。

本文聚焦于具有指定数量初始顶点和终端顶点的任务依赖图。(x， y) 任务依赖图是指具有 x 个初始顶点和 y 个终端顶点的任务依赖图，其中初始顶点和终端顶点分别是入度为零和出度为零的顶点。极小(x， y) 任务依赖图是指这样的 (x， y) 任务依赖图：删除其中任意一条边，所得图都不再是 (x， y) 任务依赖图。(1，1) 任务依赖图族曾被用于分析军方项目的项目管理方法 [22]。具体而言，Sentinel（原名陆基战略威慑）是美国空军的一项计划，涉及设计和生产 600 多枚导弹、翻新 450 个发射井、设计和生产至少 24 个发射中心，以及许多其他任务 [8]。一份近期报告 [22], 使用任务依赖图来表示 Sentinel 中相互依赖的任务。为分析 Sentinel 的项目管理方法，包括计划评审技术（PERT） [10] 和关键路径法（CPM） [17], ，该报告 [22] 用 (1，1) 任务依赖图对 Sentinel 进行建模，并将这些项目管理方法应用于这些任务依赖图。报告指出，对于任意任务依赖图 G，都可以将其转化为 (1，1) 任务依赖图。只需添加一个

新的初始顶点 u 和终端顶点 v，通过从 u 向 G 的所有初始顶点添加边，使 u 成为唯一的初始顶点；再通过从 G 的所有终端顶点向 v 添加边，使 v 成为唯一的终端顶点。

作为 [22], 随机 (1,1) 任务依赖图分析的一部分，使用以下过程生成了包含 n 个顶点和 m 条边的图。首先从一个最大连通任务依赖图开始，然后均匀随机地删除边，直到剩下 m 条边。每当删除一条边导致任务依赖图的初始顶点多于一个或终端顶点多于一个时，就取消这次边删除。请注意，如果该过程的任意一轮生成了一个极小 (1,1) 任务依赖图，那么就不能再删除更多边。显然，阶为 n 的 (1,1) 任务依赖图可能的最少边数是 $n-1$，但在 [22] 中指出，随机过程可以生成阶为 n 、边数多于 $n-1$ 的极小 (1,1) 任务依赖图。因此，如果选择一个接近 $n-1$ 的 m 值，那么 [22] 中的随机过程可能不会终止。基于这一观察，一个自然的问题是：阶为 n 的极小 (1,1) 任务依赖图可以有多少条边？更一般地，阶为 n 的极小 (x, y) 任务依赖图可以有多少条边？请注意，这些问题并不要求底层图是连通的，因此我们也研究了连通图中相应的问题。

为了回答这些问题，我们定义了另一种用于生成任务依赖图的随机过程，称为 (x, y) 边删除过程。它与上一段所述的随机过程几乎相同，唯一不同之处在于终止条件。首先取一个最大连通任务依赖图，然后随机均匀地删除边。每当删除一条边会导致任务依赖图中的初始顶点多于 x 个或终端顶点多于 y 个时，就取消此次边删除。无法再删除任何边时，过程终止。请注意，这与 [22] 中的过程不同：(x, y) 边删除过程在无法继续删除边时终止，而 [22] 中的过程则在边数达到指定值时终止。我们证明，对于 $n > x+y$，由 (x, y) 边删除过程生成的任意 (x, y) 任务依赖图，其可能的最少边数为 $n - \min(x, y)$。接着我们证明，对于 $n > x + y$，在 n 个顶点上运行 (x, y) 边删除过程所生成的任务依赖图，其可能的最多边数为 $2n - x - y - 2$；我们还证明，对于所有 $x, y \geq 1$，该数等于阶为 n 的极小 (x, y) 任务依赖图可能的最多边数。特别地，我们证明，阶为 n 的极小 (1, 1) 任务依赖图可能的最多边数为 $2n-4$。因此，当 m 小于 $2n-4$ 时，[22] 中的随机过程可能无法生成具有 n 个顶点和 m 条边的 (1, 1) 任务依赖图。另一方面，对于 $m \geq 2n-4$，可以保证 [22] 中的随机过程会生成具有 n 个顶点和 m 条边的 (1, 1) 任务依赖图。我们还刻画了达到最多边数的、阶为 n 的极小 (x, y) 任务依赖图，并特别证明，这些图中的任意有向路径长度至多为 2。

我们还考虑一种用于生成 (x, y) 任务依赖图的随机过程，该过程通过添加边而非删除边来生成图。假设我们从一个顶点集为 $\{1, 2, \ldots, n\}$ 的空有向图开始，然后以均匀随机的方式添加边 (a, b)，其中 $a < b$。如果添加某条边会导致初始顶点少于 x 个或终端顶点少于 y 个，就取消该边的添加。如果添加任意一条边后得到的结果是一个 (x, y) 任务依赖图，或者无法再添加边，我们就停止该过程。我们称之为 (x, y) 边添加过程。我们证明，对所有 $x, y \geq 1$ 和 $n > x + y$，在 n 个顶点上通过 (x, y) 边添加过程生成的任意 (x, y) 任务依赖图，其可能的最少边数为 $n - \min(x, y)$；而通过 (x, y) 边添加过程生成的任意任务依赖图，其可能的最多边数为 $\binom{n}{2} + 1 - \binom{\max(x,y)}{2} - \binom{\min(x,y)+1}{2}$。我们还证明，在 n 个顶点上进行 (x, y) 边添加过程时，边的期望数量为 $\Theta(n^2)$，其中下界中的常数取决于 (x, y)。对于 (x, y) 边添加过程和 (x, y) 边删除过程，我们还刻画了满足以下条件的参数对 (x, y)：当 $n > \max(x, y)$ 时，每个过程都能保证以一个 (x, y) 任务依赖图终止。对于每一对 (x, y)，我们证明，在 n 个顶点上进行 (x, y) 边添加过程，最终以一个 (x, y) 任务依赖图终止的概率会在 n 趋于无穷大时趋近于 1。

在第 2，我们确定，对于 x，$y \geq 1$，极小图可能的最多边数 1

阶为 n 的(x,y)任务依赖图中，若 $n=\max(x,y)$，则最多边数为0（此时 $x=y$）；若 $n=\max(x,y)+1$，则为 $2n-x-y-1$；若 $n>\max(x,y)+1$，则为 $2n-x-y-2$。此外，我们刻画了阶为 n 且边数达到可能最多值的极小(x,y)任务依赖图。作为推论，我们证明，任何阶为 n 且边数最多的极小(x,y)路径依赖图，都不存在长度超过2的有向路径。在第3节中，我们聚焦于 n 个顶点上的边添加过程。我们确定了使(x,y)边添加过程以(x,y)任务依赖图终止的(x,y)取值对；证明了由(x,y)边添加过程生成的任务依赖图中边数的极值界；证明孤立顶点的期望数量为 $o(1)$；并给出了边的期望数量的界。

在第 4 节中，我们介绍通过在 Python 中实现 (x,y) 边删除过程和 (x,y) 边添加过程所获得的实验结果。我们针对这两种过程，提出了关于边的期望数量、预期最大有向路径长度，以及生成 (x,y) 任务依赖图的概率的若干猜想。具体而言，我们的实验结果表明，随着顶点数增加，对于 (x,y) 边删除过程和 (x,y) 边添加过程，生成 (x,y) 任务依赖图的概率都趋近于 1；我们还证明，对于 (x,y) 边添加过程，这一结论成立。在第 5 节中，我们定义了一种仅使用边添加和边删除的随机过程，该过程以高概率生成一个阶为 n、含有 m 条边的 (x,y) 任务依赖图。此过程结合了 (x,y) 边添加过程与 (x,y) 边删除过程。请注意，此过程为随机生成阶为 n、含有 m 条边的 (1, 1) 任务依赖图提供了一种不同于 [22] 中方法的途径。在同一节中，我们确定了阶为 n 的 (1, 1) 任务依赖图中任务排序数的最大可能值。有趣的是，这一最大值由阶为 n 且边数达到可能的最多边数的极小 (1, 1) 任务依赖图取得。我们还确定了当 $x,y\geq 1$ 且 $\max(x,y)\leq n\leq \max(x,y)+2$ 时，阶为 n 的 (x,y) 任务依赖图中任务排序数的最大可能值。此外，我们还讨论了一些开放问题和未来研究方向。

2. 边删除过程

本节中，我们重点研究（x，y）边删除过程。我们证明，该过程并不总会生成（x，y）任务依赖图。我们还考察（x，y）边删除过程的极值行为。具体而言，对于所有 $n\geq \max(x,y)$，我们确定阶为 n 的极小（x，y）任务依赖图中可能的最多边数。这等价于确定对 n 个顶点执行（x，y）边删除过程所得到的（x，y）任务依赖图中可能的最多边数。我们还刻画了阶为 n 且边数达到可能最多的极小（x，y）任务依赖图。

为得到本节的结果，我们引入一些术语。给定一个任务依赖图 G，如果不存在满足（u，v）$\in E(G)$ 的 $u\in V(G)$，我们称顶点 $v\in V(G)$ 为初始顶点。如果不存在满足（v，u）$\in E(G)$ 的 $u\in V(G)$，我们称 $v\in V(G)$ 为终端顶点。初始顶点与根和源点相同，终端顶点与汇和汇点相同。注意，顶点 $v\in V(G)$ 可能同时是初始顶点和终端顶点，此时 v 是孤立的。若顶点 $v\in V(G)$ 是初始顶点或终端顶点，则称其为外部顶点；否则称其为内部的顶点。

显然，（1,1）边删除过程总是会得到一个（1,1）任务依赖图，因为极大连通有向无环图就是一个（1,1）任务依赖图。在下一个命题中，我们证明这一结论可推广至所有 $x\geq 1$ 的(x,x)边删除过程。然而，在再下一个命题中，我们将证明，当 $x\neq y$ 时，(x,y)边删除过程不一定会得到一个(x,y)任务依赖图。

命题 2.1。对于所有 $x\geq 1$ 和 $n\geq x$，$(,x)$ 边删除过程总是会得到一个(x,x) 任务依赖图。

证明。Supp 设我们完整执行（x、x）边删除过程，并得到任务依赖 cy

图 G，其中有 r 个初始顶点和 s 个终端顶点。为证明 $r = s = x$，反设 $r < x$ 或 $s < x$。不失一般性，设 $r < x$。注意，必有 $s = x$；否则，我们还可以删除一条边，而所得任务依赖图的初始顶点数仍至多为 x，终端顶点数仍至多为 x。我们首先论证，G 必须在同一个连通分量中至少有 2 个非孤立终端顶点。事实上，若 G 的每个连通分量至多有一个终端顶点，则必有 $r \geq s$，因为每个连通分量也至少有一个初始顶点，这与假设 $r < x$ 且 $s = x$ 矛盾。因此，G 的某个连通分量 C 有多个非孤立终端顶点。

由于 C 有多个非孤立终止顶点，因此必定存在某个顶点 u，它属于 C，且至少有两条出边。设 u 有边连接到顶点 v 和 w。如果从 G 中删除边（u，v），得到的任务依赖图至多有 $r+1$ 个初始顶点和 s 个终端顶点。这与我们的假设矛盾：假设（x，x）边删除过程在任务依赖图 G 上终止，因为此时仍可以删除一条边，同时保持初始顶点数至多为 x，终端顶点数至多为 x。因此，（x，x）边删除过程必定会得到一个（x，x）任务依赖图。

□

在下一个命题中，我们考察针对 $x \neq y$ 的（x，y）边删除过程，并证明该过程可能会在不具有 x 个初始顶点和 y 个终端顶点的任务依赖图上停止。

命题 2.2. 对于所有满足 $x \neq y$ 和 $n \geq \max(x, y)$ 的 $x, y \geq 1$，(x, y) 边删除过程不一定会生成(x, y) 任务依赖图。

证明。不失一般性，假设 $x > y$，那么 $x \geq 2$。令 $G_{y,\ n}$ 为定义在 $\{1, 2, \ldots, n\}$ 上的任务依赖图，其中 $\{1, \ldots, n-y+1\}$ 构成一条有向路径，并且有 $y-1$ 个孤立顶点 $n-y+2, \ldots, n$（若 $y = 1$，则没有孤立顶点）。那么 $G_{y,\ n}$ 有 y 个初始顶点和 y 个终端顶点。如果删除任意一条边，都会得到一个有 $y+1$ 个终端顶点的任务依赖图。因此，如果任意一轮产生 $G_{y,\ n}$，那么 $(x,\ y)$ 边删除过程不可能得到一个 $(x,\ y)$ 任务依赖图。 □

接下来，我们将重点求出在 n 个顶点上经过 (x, y) 边删除过程得到的 (x, y) 任务依赖图中边数的极值。我们先看最小值，显然，当 $x = y = 1$ 时，该值为 $n-1$。边数不可能少于 $n-1$，否则就会有多个连通分支，这意味着存在多个初始顶点和多个终端顶点。此外，阶为 n 的有向路径有 $n-1$ 条边。我们将这一结论推广到所有 $x, y \geq 1$，证明对于所有 $n > \max(x, y)$，通过 (x, y) 边删除过程生成的任意 (x, y) 任务依赖图的可能的最少边数为 $n - \min(x, y)$。我们先给出一个引理。

引理 2.3。对于所有 x、$y \geq 1$ 和 $n > \max(x$、$y)$，任意(x, y) 任务依赖图中可能的最少边数为 $n - \min(x$、$y)$。

证明。不失一般性，设 $x \geq y$。对于上界，令 G 为定义在 $1, \ldots, n$ 上的任务依赖图，其中孤立顶点为 $n-y+2, \ldots, n$（若 $y = 1$，则没有孤立顶点），包含一条经过 $x-y+2, \ldots, n-y+1$的有向路径，以及对每个满足 $1 \leq a \leq x-y+1$ 的 a，包含边 $(a, x-y+2)$。由于这$y-1$ 个孤立顶点既计作初始顶点，也计作终端顶点，因此 G 有 x 个初始顶点、y 个终端顶点和 $n-y$ 条边。因此，任何 (x, y) 任务依赖图的可能的最少边数至多为 $n-y$。

对于下界，假设 H是一个阶为 n且至多有 $n-y-1$条边的任务依赖图。那么 H 至少有 $y+1$ 个连通分支，因此 H 至少有 $y+1$ 个终端顶点。因此，H不可能是一个 (x, y) 任务依赖图，所以任何 (x, y) 任务依赖图的边数总是大于 $n-y-1$。

□

推论 2.4. 对于所有 x、$y \geq 1$ 和 $n > \max(x,\ y)$，由 $(x,\ y)$ 边删除过程生成的任意$(x,\ y)$ 任务依赖图中的可能的最少边数为 $n - \min(x,\ y)$。

证明。 下界由引理 2.3 得出。对于上界，注意我们可以使用引理 2.3 证明中的同一个任务依赖图 G 。 □

作为第二个推论，我们证明，对于通过（x， y）边删除过程生成的任意任务依赖图，同样的结果也成立。请注意，我们已去掉推论 2.4 中关于任务依赖图必须是（x， y）任务依赖图的要求。

推论 2.5. 对于所有 x、 $y \geq 1$ 和 $n > \max(x、\ y)$，通过 $(x、\ y)$ 边删除过程生成的任何任务依赖图中，可能的最少边数为 $n - \min(x、\ y)$。

证明。 假设 G 是通过在n 个顶点上执行（x, y）边删除过程生成的任务依赖图。由于边删除过程在任意轮次都不会减少初始顶点或终端顶点的数量， G 必定是某个 $r \leq x$ 和 $s \leq y$对应的（r, s）任务依赖图。如果$r = x$ 且 $s = y$，则根据推论 2.4， G 至少有 $n - \min(x, y)$ 条边。否则， $r \leq x - 1$ 或$s \leq y - 1$。在任一种情况下，根据引理 2.3， G 至少有 $n - \min(x, y)$ 条边。 □

在本节余下部分中，我们重点研究如何确定通过在 n 个顶点上执行 (x, y) 边删除过程所得到的 (x, y) 任务依赖图中可能的最多边数。这等价于确定阶为 n 的极小 (x, y) 任务依赖图中可能的最多边数。作为推论，我们获得了通过在 n 个顶点上执行 (x, y) 边删除过程所得到的随机任务依赖图中边的期望数量的渐近界。为了确定阶为 n的极小 (x, y) 任务依赖图中可能的最多边数，我们证明了若干引理。

引理 2.6。 图 G 是一个最小的(x, y) 任务依赖图，当且仅当对于每个$(u, v) \in E(G)$，出度$(u) = 1$ 或入度$(v) = 1$。

证明。假设 G 是一个极小的（x， y）任务依赖图，且对于某个（u， v） $\in E$（G），outdeg(u) > 1，indeg(v) >1。那么存在顶点 w、 $z \in V$（G），使得（u， w） $\in E$（G）且（z， v） $\in E$（G）。由于 G 是一个（x， y）任务依赖图，因此存在一条从 w 到 G 的至少一个终端顶点的路径，并且存在一条从 G 的至少一个初始顶点到 z 的路径。因此，如果我们从 E（G）中移除边（u， v），那么仍然存在一条从至少一个初始顶点到 v 且经过 z 的路径，并且仍然存在一条从 u 到至少一个终端顶点且经过 w 的路径。

如果对于每个 $(u, v) \in E(G)$，都有 outdeg(u) $= 1$ 或 indeg(v) $= 1$，那么移除 (u, v) 会产生 u 一个新的初始顶点或 v 一个新的终端顶点。因此， G 是一个极小的 (x, y) 任务依赖图。 □

我们称路径 $v_1, v_2, \ldots, v_k$ 满足 $k \geq 3$ 可删除 条件，如果对于 $1 < i < k$，indeg(v_i) =outdeg(v_i) $= 1$，且 outdeg(v_1) > 1,indeg(v_k) > 1。所谓删除 一条可删除路径，是指删除边 $(v_1, v_2) \ldots, (v_{k-1}, v_k)$ 和顶点 $v_2, \ldots, v_{k-1}$。

引理 2.7。 如果 G 是一个带有可删除路径 P 的极小（x, y）任务依赖图，那么从 G 中移除 P 后得到的图 G' 也是一个极小（x, y）任务依赖图。

证明。 显然，引理 2.6 中关于 G' 为极小（x， y）任务依赖图的充分条件仍然成立，因此 G' 是一个极小任务依赖图。 □

引理 2.8。 假设 G 是一个极小（x， y）任务依赖图，其中有一条路径（v_1，…， v_k），满足 $k \geq 3$，出度(v_1) > 1，且入度(v_k) > 1。那么 G 有一条可删除路径。

证明。 我们对 k使用归纳法证明该引理。如果 $k = 3$，则由引理 2.6 可立即得出，任何满足出度(v_1) > 1 且入度(v_3) $= 1$ 的路径（v_1， v_2， v_3）都必定是可删除路径。

现在，假设对于某个 $k \geq 3$，该引理对所有 $3 \leq j \leq k$ 都成立，并且 G 有一条路径（v_1，.，v_{k+1}），其中 outdeg(v_1) > 1 且 indeg(v_{k+1}) > 1。由引理 2.6，indeg(v_2) $= 1$。如果 outdeg(v_2) > 1，那么根据归纳假设，（v_2，.. .， v_{k+1}）包含一条可删除路径。

否则，如果 outdeg(v_2) $= 1$，那么令 $i > 2$ 为满足 indeg(v_i) > 1 或 outdeg(v_i) $= 1$的最小值。如果 indeg(v_i) $= 1$，那么（v_1，…，v_i）是一条可删除路径。如果 outdeg(v_i) > 1，那么根据引理 2.6，$i < k$。因此，根据归纳假设，（v_i，…，v_{k+1}）包含一条可删除路径。 □

推论 2.9. 如果 G 是一个没有可删除路径的极小（x，y）任务依赖图，那么 G 的底层无向图是树的不交并。

证明。假设存在矛盾，即 G 的底层图包含一个环 C。将该环 C 的边划分为极大有向路径，即长度尽可能长的有向路径。由于G 无环，该划分至少包含两条极大有向路径。如果任何一条极大有向路径仅包含一条边 $e =$（u，v），那么根据极大性，outdeg(u) > 1 且 indeg(v) > 1，这与引理 2.6 矛盾。因此，G存在一条路径（v_1，…，v_k），其中 $k \geq 3$，根据极大性，outdeg(v_1) > 1，且 indeg(v_k) > 1。根据引理 2.8，G 必须存在一条可删除路径，这与假设矛盾。 □

我们定义两族最小的（x, y）任务依赖图，供后续证明使用。对于每个 $x \geq 1$、$y \geq 1$ 和 $n \geq x+2$，定义 $S_{x,y,n}$ 为阶为 n 的最小（x, y）任务依赖图，其中有 $y-1$ 个孤立顶点、$x-y+1$ 个非孤立初始顶点 $u_1, \ldots, u_{x-y+1}$、1 个非孤立终端顶点 v、$n-x-1$ 个内部顶点 $i_1, \ldots, i_{n-x-1}$，以及边 (u_1, i_t)（$1 \leq t \leq n-x-1$）、(i_t, v)（$1 \leq t \leq n-x-1$）和 (u_t, v)（$2 \leq t \leq x-y+1$）。对于每个 $x \geq 1$、$y \geq 1$ 和 $n \geq x+y+1$，定义 $T_{x,y,n}$ 为阶为 n 的最小（x, y）任务依赖图，其中有 x 个初始顶点 $u_1, \ldots, u_x$、y 个终端顶点 $v_1, \ldots, v_y$、$n-x-y$ 个内部顶点$i_1, \ldots, i_{n-x-y}$，以及边 (u_1, i_t)（$1 \leq t \leq n-x-y$）、(i_t, v_1)（$1 \leq t \leq n-x-y$）、(u_1, v_t)（$2 \leq t \leq y$）和 (u_t, v_1)（$2 \leq t \leq x$）。

定理 2.10。对于 x、$y \geq 1$，阶为 n 的最小 (x, y) 任务依赖图的可能的最多边数为：当 $n = \max(x, y)$ 时为 0（此时 $x = y$）；当 $n = \max(x, y) + 1$ 时为 $2n - x - y - 1$；当 $n > \max(x, y) + 1$ 时为 $2n - x - y - 2$。

证明。不失一般性，假设 $x \geq y$。若 $n = x$，则任意阶为 n 的 (x, y) 任务依赖图中的所有顶点既是初始顶点，也是终端顶点，因此 $x = y$，且没有边。若 $y < x$，则任意阶为 n 的 (x, y) 任务依赖图 G 至多有 $y-1$ 个孤立顶点，每个孤立顶点既计作初始顶点，也计作终端顶点，即 G 至少有 $x+1$ 个外部顶点。因此，若 G 的阶为 $x+1$，则 G 有 $y-1$ 个孤立顶点、一个非孤立终端顶点、$x-y+1$ 个非孤立初始顶点、没有内部顶点，并且有 $x-y+1 = 2n-x-y-1$ 条边；每个非孤立初始顶点都通过一条边连接到唯一的非孤立终端顶点。

对于 $n > x+1$，我们对 n 使用归纳法证明该命题。若 G 是一个阶为 $n = x+2$ 的 (x, y) 任务依赖图，那么 G 要么 (a) 有 $y-1$ 个孤立顶点、$x-y+1$ 个非孤立初始顶点、一个非孤立终端顶点和一个内部顶点；要么 (b) 有 $y-2$ 个孤立顶点、$x-y+2$ 个非孤立初始顶点、两个非孤立终端顶点且没有内部顶点。注意，只有当 $y > 1$ 时，情况 (b) 才会出现。两种情况下，G 都有 $2n-x-y-2$ 条边。

现在假设 $n > x+2$，并令 G 为一个阶为 n 的极小 (x, y) 任务依赖图。若 G 有一条可删除路径 $P = (v_1, \ldots, v_k)$，则根据引理 2.7，将其删除，得到一个新的 (x, y) 任务依赖图 G'。G' 不可能恰有 $x+1$ 个顶点，否则 G' 中每个非孤立顶点都是外部顶点，且 $(v_1, v_k) \in E(G)$，也就是说，删除 (v_1, v_k) 后，G 仍是一个 (x, y) 任务依赖图，这与其极小性矛盾。因此，G' 至少有 $x+2$ 个顶点，故根据归纳假设，至多有 $2(n-k+2)$ 条边。由此，图 G 至多有 $2(n-k+2)$ 条边，因为 $k \geq 3$。

如果 G 没有可删除路径，那么根据推论 2.9，G 的底层无向图必定是树的不交并。如果 G 有 z 个孤立顶点，那么由于 $n \geq x+3$，它必定至少有 $z+1$个连通分量。

假设 G 有 $z+1$ 个连通分量，并令 C 为 G 中不是孤立顶点的那个连通分量。那么，C 的底层无向图必定是一棵树。因此，由于 $n \geq x+3$ 和 $x \geq y$，这个连通分量必定至少有一个内部顶点。此外，G 有 $n-$（$z+1$）条边。显然，在这种情况下 $x+y-z \leq n-1$，所以 G 至多有 $2n-x-y-2$ 条边。

否则，G 至少有 $z+2$ 个连通分支。在这种情况下，G 至多有 $n-(z+2)$ 条边。由于 $x+y-z \leq n$，我们得出结论：G 至多有 $2n-x-y-2$ 条边。

注意， $2n-x-y-2$ 边的上界可由图族 $S_{x、y、n}$ 对所有 $n \geq x+2$ 达到。

□

利用上一个定理，我们还可以确定对含有 n 个顶点的任何任务依赖图执行 (x, y) 边删除过程后，可能产生的任务依赖图的可能的最多边数。请注意，在下一个推论中，我们不要求生成的任务依赖图具有 x 个初始顶点和 y 个终端顶点。

推论 2.11. 对于 $x、y \geq 1$ 和 $n > x+y$，通过在 n 个顶点上进行(x, y) 边删除过程生成的任务依赖图中，可能的最多边数为 $2n-x-y-2$。

证明。 假设 G 是通过在n 个顶点上执行 (x, y) 边删除过程得到的任务依赖图，因此 G 必定是某个 $r \leq x$ 和 $s \leq y$ 对应的 (r, s) 任务依赖图。若 $r = x$ 和 $s = y$，则 G 至多有 $2n-x-y-2$ 条边。否则，有 (a) $r \leq x-1$ 和 $s = y$，或 (b) $r = x$ 和 $s \leq y-1$。在情况 (a) 中，不可能存在顶点 u, v, w，使得边 (u, v) 和 (u, w) 存在；否则，我们可以删除边 (u, v)，使初始顶点的数量最多增加一个，而不增加终端顶点的数量。于是 G 的底层无向图是树的不交并，因此 G 至多有 $n-1 \leq 2n-x-y-2$ 条边，因为 $n > x+y$。情况 (b) 的证明类似。

□

请注意，通过 (x, y) 边删除过程得到的任何任务依赖图，都必定是某些 $r \leq x$ and $s \leq y$ 对应的最小 (r, s) 任务依赖图。如果它不是最小的，那么我们就可以再删除一条边，同时仍然得到一个初始顶点不超过 x 个、终端顶点不超过 y 个的任务依赖图。因此，根据定理 2.10，在 n 个顶点上进行 (x, y) 边删除过程，所得任务依赖图的边数至多为 $2n$，且根据推论 2.5，至少为 $n-\min(x, y)$。这就得到了关于边的期望数量的以下推论。

推论 2.12. 对于所有 $x、y \geq 1$，在 n 个顶点上进行 $(x、y)$ 边删除过程后得到的随机任务依赖图中的边的期望数量为 $\Theta(n)$。

在下一个定理中，我们刻画了对于所有 $n \geq \max(x, y)$，边数尽可能多的极小 (x, y) 任务依赖图。

定理 2.13. 一个阶为 n 的图 G，当且仅当它至多有两个不是孤立顶点的连通分量时，才在所有阶为 n 的极小(x, y) 任务依赖图中具有可能的最多边数，并且

1. 如果 G 由所有孤立顶点组成，那么 $n = x = y$，
2. *if* G 具有两个不是孤立顶点的连通分支，则 G 没有内部顶点，并且
3. 如果 G 恰好有一个非孤立顶点的连通分量 C，那么要么

   (a) C 没有内部顶点，*(b)* C中存在一个初始顶点 u、一个终端顶点 v以及多个内部顶点，使得 C 的所有内部顶点均与 u 和 v相邻， u 可能与更多终端顶点相邻，而 v 可能与更多初始顶点相邻；或者 *(c)* C 恰好有一个内部顶点，该顶点与 $p > 0$ 个初始顶点和 $q > 0$ 个终端顶点相邻。如果 $p = 1$，则该初始顶点可能与更多终端顶点相邻。如果 $q = 1$，则该终端顶点可能与更多初始顶点相邻。如果$p > 1$，则这些初始顶点仅与该内部顶点相邻。如果 $q > 1$，则这些终端顶点仅与该内部顶点相邻。

证明。 我们通过归纳法证明正向，沿用定理 2.10 证明中的归纳过程。假设 G 是一个阶为 n 的图，在所有阶为 n 的极小（x， y）任务依赖图中具有可能的最多边数。不失一般性，我们假设$x \geq y$。如果 $n = x$，那么 $x = y$，且该图由所有孤立顶点组成，即结构（1）。如果

$n = x+1$，那么 G 恰好有一个不是孤立顶点的连通分量 C，并且 C 没有内部顶点，即结构（3）（a）。如果 $n = x+2$，那么 G 要么恰好有两个非孤立终端顶点且没有内部顶点，即结构（2）；要么恰好有一个非孤立终端顶点和恰好一个内部顶点。这两个顶点必须属于同一个连通分量，因此 G 具有结构（3）（c），其中 $q = 1$。

现在假设 $n > x + 2$。如果 G 没有可删除路径，那么根据定理 2.10 G 证明的最后两段， G 要么没有内部顶点，且恰有两个不是孤立顶点的连通分量，即结构（2）；要么恰有一个连通分量 C 不是孤立顶点，并且恰有一个内部顶点。如果 C 有多个与其内部顶点相邻的初始顶点，那么所有这些初始顶点都只与该内部顶点相邻；否则 G 就不会是一个极小的（x， y）任务依赖图。同样地，如果 C 有多个与其内部顶点相邻的终端顶点，那么所有这些终端顶点都只与该内部顶点相邻；否则 G 就不会是一个极小的（x， y）任务依赖图。因此， G 具有结构（3）（c）。

如果 G 有一条可删除路径 $P =$（v_1，.....， v_k），删除该路径后得到 G'，那么根据定理 2.10 的证明，我们知道 $k = 3$。此外， v_1 和 v_k 不相邻，因为 G 是一个极小的（x， y）任务依赖图，因此在 G' 中，它们要么属于不同的连通分量，要么在 G 的底层无向图中属于一个环。如果它们在 G的底层无向图中属于一个环，那么我们将该环的边划分为极大有向路径。我们断言，这个划分恰好由两条含三个顶点的有向路径组成，因此根据归纳假设， G'具有结构（3）（b）或（3）（c）。首先注意到，该划分至少由两条极大有向路径组成，因为 G 是无环的。如果这些有向路径中有任何一条仅包含两个顶点，那么连接这两个顶点的边就违反了引理 2.6，因此每条有向路径至少包含三个顶点，从而至少有一个内部顶点。由于一个环中不可能恰好有三条极大有向路径，假设该划分由四条或更多有向路径组成，以导出矛盾。于是， G' 在一个（3）（b）和（3）（c）中均不存在的结构中具有三个内部顶点，这与归纳假设矛盾。因此，该划分由两条极大有向路径组成。两条有向路径中没有任何一条包含超过三个顶点，否则 G' 将包含一条超过三个顶点的有向路径，这与归纳假设矛盾。因此，两条极大有向路径都恰好包含三个顶点。由此， G' 具有结构（3）（b）或（3）（c）。

如果 G' 具有 (3)(b) 中的结构，那么与 G' 中内部顶点相邻的初始顶点必须是 v_1 ，与 G' 中内部顶点相邻的终端顶点必须是 v_k，因为 G' 中距离为 2 的顶点只有 v_1 和 v_k。由于 G 是在 G' 的基础上添加一个仅与 v_1 和 v_k 相邻的顶点 v_2 得到的，因此我们得出 G 具有 (3)(b) 中的结构。现在假设 G' 具有 (3)(c) 中的结构。如果多个初始顶点都与 G' 中的内部顶点相邻，那么 G 就不是一个极小的 (x， y) 任务依赖图，因为我们可以移除从 v_1 到 G' 中内部顶点的边，且所得图仍然是一个 (x， y) 任务依赖图。因此，只有一个初始顶点与 G' 中的内部顶点相邻。类似地，只有一个终端顶点与 G' 中的内部顶点相邻。因此， G 具有 (3)(b) 中的结构。

如果在 G' 中，顶点 v_1 和 v_k 属于不同的连通分支，那么这两个连通分支都不是孤立顶点，因为在 G中，outdeg(v_1) > 1 且 indeg(v_k) > 1 。因此，根据归纳假设， G'没有内部顶点， v_1 是初始顶点，而 v_k 是终端顶点。由此可知， G 恰好有一个非孤立顶点的连通分支，且该连通分支具有 (3)(c)中描述的结构，其中 $p = q = 1$。

对于反向方向，首先假设 G 由所有孤立顶点组成，并且 $n = x = y$。那么显然， G是一个阶为 n 的图，在所有阶为 n的极小（x， y）任务依赖图中，它具有可能的最多边数。接下来，假设 G 有两个非孤立连通分量，并且 G 没有内部顶点。如果 G 中非孤立连通分量的大小分别为 j和 k，那么 G有 $n-j-k$ 个孤立顶点和 $j+k-2$ 条边，并且 $x+y = 2n-j-k$。因此， $n \geq x+2$ 且 G 有 $2n-x-y-2$ 条边，所以 G是一个阶为 n的图，在所有阶为 n的极小（x， y）任务依赖图中，它具有可能的最多边数。接下来，假设G 恰好有一个非孤立连通分量 C ，并且 C 没有内部顶点，即 G具有（3）（a）中的结构。如果 C 的大小为 j，那么 G 有 $n-j$ 个孤立顶点和 $j-1$ 条边，并且

$x+y=2n-j$。因此， $n=x+1$ 和 G 有 $2n-x-y-1$ 条边，所以 G 是一个阶为 n 的图，在所有阶为 n的最小（x， y）任务依赖图中具有可能的最多边数。

接下来，假设 G 具有 (3)(b) 中的结构，即它有 $i \geq 0$ 个孤立顶点，并且恰好有一个连通分量 C 不是孤立顶点； C 有 $j>1$ 个内部顶点，每个都与一个初始顶点 u 和终端顶点 v相邻， u 与 $k \geq 0$ 个其他终端顶点相邻，而 v 与 $l \geq 0$ 个初始顶点相邻。于是 $x=i+1+l$、 $y=i+1+k$ 和 $n=i+2+k+l+j$。显然$n \geq \max(x,y)+3$，并且 G有 $2j+k+l=2n-x-y-2$条边。因此，根据定理 2.10， G是一个阶为 n 的图，在所有阶为 n的最小（x,y）任务依赖图中具有可能的最多边数。

最后，假设 G 具有 (3)(c) 中的结构，因此连通分量 C 恰好有一个内部顶点，该顶点与 $p>0$ 个初始顶点和 $q>0$ 个终端顶点相邻。我们考虑三种情况。首先，假设 $p>1$、 $q>1$ 以及 C 中的初始顶点和终端顶点都只与该内部顶点相邻。那么 G 有 $n-p-q-1$ 个孤立顶点， $x=n-q-1$、 $y=n-p-1$，并且 G 有 $p+q$ 条边。因此 G 有 $2n-x-y-2$ 条边，所以 G 是一个阶为 n 的图，在所有阶为 n 的极小（x， y）任务依赖图中具有可能的最多边数。

其次，假设 $p=1$、 $q=1$中与 C 内部顶点相邻的初始顶点与 $r \geq 0$ 个终端顶点相邻，并且 C中与内部顶点相邻的终端顶点与 $s \geq 0$ 个初始顶点相邻。那么 G 有 $n-r-s-3$ 个孤立顶点， $x=n-r-2$、$y=n-s-2$，且 G 有 $r+s+2$ 条边。那么 G 有 $2n-x-y-2$ 条边，因此，在所有阶为 n 的极小（x， y）任务依赖图中， G 是一个阶为n 且具有可能的最多边数的图。

第三，假设 $p=1$、 $q>1$，且 C 中的初始顶点与 $r \geq 0$ 个终端顶点相邻。那么 G 有 $n-q-r-2$ 个孤立顶点， $x=n-q-r-1$、 $y=n-2$，且 G 有 $q+r+1$ 条边。于是 G 有 $2n-x-y-2$ 条边，因此，在所有阶为 n的极小（x， y）任务依赖图中， G 是具有可能的最多边数的一个 n 阶图。当 $q=1$ 且 $p>1$时，证明过程类似。 □

在定理 2.13 中，请注意，情形 (1) 中具有该结构的有向图的最大有向路径长度为 0，而情形 (2) 和 (3)(a) 中具有相应结构的有向图的最大有向路径长度为 1。情形 (3)(b) 和 (3)(c) 中具有相应结构的有向图的最大有向路径长度为 2。因此，我们得到以下推论。

推论 2.14。 每个边数达到可能最多值的极小(x,y) 路径依赖图，其阶为 n，都不含长度大于2的有向路径。

对于仅有一个初始任务的极小任务依赖图，我们根据定理 2.10 的推论，得到了边数可能最多值关于图的阶和终端任务数的函数表达式。我们还得到了仅有一个终端任务的任务依赖图的相应结果。

推论 2.15。 阶为n 的极小$(x,1)$ 任务依赖图的最大边数为：若 $n=x$，则为0 （此时 $x=1$）；若 $n=x+1$，则为 x；若 $n>x+1$，则为 $2n-x-3$。阶为 n 的极小$(1,y)$ 任务依赖图的最大边数为：若 $n=y$，则为0 （此时$y=1$）；若 $n=y+1$，则为 y；若 $n>y+1$，则为 $2n-y-3$。

下一个推论涵盖了在 [22] 中研究的任务依赖图。

推论 2.16. 阶数为n 的极小$(1,1)$ 任务依赖图的最大边数为0 （当 $n=1$时），1 （当 $n=2$时），以及 $2n-4$ （当 $n>2$时）。

将连通极小(x,y) 任务依赖图定义为底层无向图连通的极小 (x,y) 任务依赖图。

定理 2.17. 对于 x、 $y \geq 1$，阶为 n 的连通极小 (x,y) 任务依赖图的边数最大值为 $\max(x,y)$，若 $n=x+y$ （此时 $\min(x,y)=1$）；若 $n>x+y$，则为 $2n-x-y-2$。

证明。不失一般性，假设 $x \geq y$。首先注意，若 $n < x + y$，则不存在阶为 n 的连通极小 (x, y) 任务依赖图。这是因为，若 $n < x + y$，则某个初始顶点也必定是终端顶点，因此必定存在一个孤立顶点，所得有向图不是连通图。

接下来，假设 $n = x + y$。如果 $y > 1$，那么根据定理 2.10，阶为 n 的（不一定连通的）极小（x，y）任务依赖图可能的最多边数为 $2n - x - y - 2 = x + y - 2$。这个有向图不可能连通，因为它的边数不足，因此当 $n = x + y$ 且 $y > 1$ 时，不存在阶为 n 的连通极小（x，y）任务依赖图。如果 $y = 1$，那么根据定理 2.10，阶为 n 的（不一定连通的）极小（x，y）任务依赖图可能的最多边数为 $2n - x - y - 1 = x + y - 1$。请注意，这个上界可以由阶为 n、有 x 个初始顶点和一个终端顶点，并且每个初始顶点都连接到该终端顶点的连通极小（x，y）任务依赖图达到。

最后，假设 $n > x + y$。根据定理 2.10，阶为 n 的（不一定连通的）极小 (x, y) 任务依赖图中可能的最多边数为 $2n - x - y - 2$。注意，有向图 $T_{x,y,n}$ 在 $n \geq x + y + 1$ 的情况下达到了这个 $2n - x - y - 2$ 条边的上界。

□

3. 边添加过程

边添加过程显然会在 $x = y = 1$ 时生成一个 (x, y) 任务依赖图，因为每次添加边只会使初始顶点和终端顶点的数量减少或保持不变，直到两者都为 1。然而，当 $y \neq x$ 时，该过程不一定会生成一个 (x, y) 任务依赖图。

命题 3.1。对于 $x \neq y$ 和 $n > \max(x, y)$，(x, y) 边添加过程不一定会生成(x, y) 任务依赖图。

证明。不失一般性，假设 $n > x > y$。令 $G_{x,y,n}$ 为 $\{1, 2, \ldots, n\}$ 上的任务依赖图，其中顶点 $n - y + 1, \ldots, n$ 是孤立的，$\{x - y + 1, \ldots, n - y\}$ 是最大连通的，并且对于所有 $a \in \{1, \ldots, x - y\}$ 和 $b \in \{x - y + 1, \ldots, n - y\}$，都存在边 (a, b)。那么$G_{x,y,n}$有 x 个初始顶点和 $y + 1$ 个终端顶点，因为这 y 个顶点 $n - y + 1, \ldots, n$ 既是初始顶点，也是终端顶点。如果添加任何一条边，结果都会是一个仅有$x - 1$ 个初始顶点的任务依赖图。因此，如果任何一轮产生 $G_{x,y,n}$，(x, y) 边添加过程就不可能产生一个 (x, y) 任务依赖图。 □

根据上一个结果，很自然会问：对于 $x > 1$，(x, x) 边添加过程是否不一定会得到 (x, x) 任务依赖图。我们将在下一个命题中证明，与 $x = y = 1$ 时的情况一样，对于所有 $x > 1$，(x, x) 边添加过程总会得到 (x, x) 任务依赖图。请注意，这与 (x, x) 边删除过程的情况相同：对于所有 $x \geq 1$，该过程总会得到 (x, x) 任务依赖图。

命题 3.2。对于所有 $x \geq 1$，(x, x) 边添加过程总会得到(x, x) 任务依赖图。

证明。假设我们完整地执行了 (x, x) 边添加过程，并得到任务依赖图 G，其中有 r 个初始顶点和 s 个终端顶点。为了证明 $r = s = x$，反设 $r > x$ 或 $s > x$。不失一般性，设 $r > x$。注意，我们必须有 $s = x$，否则我们就可以再添加一条边，而所得任务依赖图仍会至少有 x 个初始顶点和至少 x 个终端顶点。与命题 2.1 类似，我们首先论证，G 必须在同一个连通分量中至少有 2 个非孤立初始顶点。事实上，如果 G 的每个连通分量至多有一个初始顶点，那么我们就有 $r \leq s$，因为每个连通分量也必须至少有一个终端顶点，这与假设 $r > x$ 和 $s = x$ 矛盾。因此，G 的某个连通分量有多个非孤立初始顶点。

将这些非孤立初始顶点 u, v 标记为 $u < v$。如果我们向 G添加边 (u, v)，那么我们就得到一个有 $r - 1$ 个初始顶点和 s 个终端顶点的任务依赖图 G'。由于$r > x$ 和 $s = x$， G' 至少有 x 个初始顶点和 x 个终端顶点。这与我们的假设矛盾：(x, x) 边添加过程在任务依赖图 G上终止了，因为仍然可以添加一条边，同时保持初始顶点数量至少为 x，且终端顶点数量至少为 x。因此，(x, x) 边添加过程必然会得到一个 (x, x) 任务依赖图。

□

当 $x = y = 1$ 时，在 n 个顶点上进行 (x, y) 边添加过程，直至其终止所需的可能最少边数为 $n - 1$，因为要包含所有顶点，至少需要 $n - 1$ 条边，而阶为 n 的有向路径有 $n - 1$ 条边。注意，这与在 n 个顶点上进行 (x, y) 边删除过程直至终止所需的可能最少边数相同，理由也相同。采用与命题 2.4 相同的证明，只需将 边添加 替换为 边删除，即可得到以下推广。

命题 3.3. 对于所有 $n > \max(x, y)$，由在 n 个顶点上进行的 (x, y) 任务依赖图生成过程 (x, y) 边添加过程所生成的任意任务依赖图，其可能的最少边数为 $n - \min(x, y)$。

不难利用上一个命题证明，由 (x, y) 边添加过程生成的任意任务依赖图，其可能的最少边数为 $n - \min(x, y)$，其中 $n > \min(x, y) + 1$。注意，在下一个结果中，我们不再要求任务依赖图必须是 (x, y) 任务依赖图。

命题 3.4。 对于所有 $n > \max(x, y) + 1$，在 n 个顶点上通过 (x, y) 边添加过程生成的任何任务依赖图，其可能的最少边数为 $n - \min(x, y)$。

证明。 设 G 是通过 (x, y) 边添加过程生成的任务依赖图。那么 G 是某些 $r \geq x$ 和 $s \geq y$对应的 (r, s) 任务依赖图。若 $r = x$ 和 $s = y$，则根据命题 3.3， G 至少有$n - y$条边。否则，假设 $r \geq x + 1$或 $s \geq y + 1$。若 G有任何内部顶点 u，那么我们可以在两个与 u 相邻的顶点之间添加一条边，同时保持初始顶点和终端顶点的数量不变，因此 G 就不是通过 (x, y) 边添加过程生成的任务依赖图。因此， G 没有内部顶点，所以 G 中的所有顶点都是孤立顶点、非孤立初始顶点或非孤立终止顶点。

首先考虑 $r \geq x + 1$ 的情况。在这种情况下，我们必有 $s = y$。 G 不能有多个非孤立初始顶点，否则我们可以在两个非孤立初始顶点之间添加一条边，使初始顶点的数量减少 1，而终端顶点的数量保持不变。因此， G 中的顶点总数至多为 $y + 1$，这与我们的假设 $n > \max(x, y) + 1$ 矛盾。现在考虑 $s \geq y + 1$ 的情况。在这种情况下，我们必有 $r = x$。与前一种情况类似， G 不能有多个非孤立终端顶点，否则我们可以在两个非孤立终端顶点之间添加一条边，使终端顶点的数量减少 1，而初始顶点的数量保持不变。因此，G 中的顶点总数至多为 $x + 1$，这与我们的假设 $n > \max(x, y) + 1$ 矛盾。

□

在 n 个顶点上运行的 (1, 1) 边添加过程终止前，可能的最多边数为 $\binom{n}{2}$，这远大于在 n 个顶点上运行的 (1, 1) 边删除过程终止前可能的最多边数。显然，上界成立，因为在顶点集为 $\{1, 2, \ldots, n\}$ 的有向图中，形如 (a, b) 且 $a < b$ 的边只有 $\binom{n}{2}$ 条。为证明 $\binom{n}{2}$ 是下界，考虑从顶点为 1, 2, . . ., n 的完全任务依赖图中删除边 $(n - 1, n)$ 后得到的任务依赖图 G。 G 有两个终端顶点 $n - 1$ 和 n，因此，任何产生图 G 的随机过程实例都会在具有 $\binom{n}{2}$ 条边时终止。在接下来的几个证明中，我们将其推广到所有 $x, y \geq 1$，证明由 (x, y) 边添加过程产生的任意 (x, y) 任务依赖图中，可能的最多边数对于所有 $n > x + y$ 均为 $\binom{n}{2} + 1 - \binom{\max(x,y)}{2} - \binom{\min(x,y)+1}{2}$。我们先给出一个定义和一个引理。定义 $Q_{x,y,n}$ 为如下任务依赖图：其顶点为 1, 2, . . ., n，包含 $k =$ 个孤立顶点 1, . . ., k（其中数量为 $\max(0, x + y - n)$），以及所有形如 (a, b) 且 $k + 1 \leq a < b \leq n$ 的可能边，但 $k + 1 \leq a < b \leq x$ 或 $n - y + k + 1 \leq a < b \leq n$ 的情况除外。

引理 3.5。对于所有 $x, y \geq 1$ 以及 $n \geq \max(x, y)$，阶为 n 的 (x, y) 任务依赖图中的可能的最多边数为 $\binom{n-\max(0,x+y-n)}{2} - \binom{x-\max(0,x+y-n)}{2} - \binom{y-\max(0,x+y-n)}{2}$。

证明。下界可通过使用 $Q_{x,y,n}$ 得出。对于 $k = \max(0, x+y-n)$，$Q_{x,y,n}$ 中的非孤立初始顶点为 $k+1, \ldots, x$，而非孤立终止顶点为 $n-y+k+1, \ldots, n$。$Q_{x,y,n}$ 有 $\binom{n-k}{2} - \binom{x-k}{2} - \binom{y-k}{2}$ 条边。对于上界，注意在任何任务依赖图中，初始顶点之间都不能有边，终端顶点之间也不能有边。假设一个阶为 n 的 (x, y) 任务依赖图 G 有 z 个孤立顶点，那么它有 $x-z$ 个非孤立初始顶点和 $y-z$ 个非孤立终止顶点。注意，$z \geq x+y-n$，因为 G 中内部顶点的数量为 $n-(x+y-z)$。G 中可能的最多边数为 $f(n, x, y, z) = \binom{n-z}{2} - \binom{x-z}{2} - \binom{y-z}{2}$。由于

$$f(n, x, y, z) - f(n, x, y, z+1) = n - z - 1 - (x - z - 1) - (y - z - 1) = z - (x + y - n) + 1 > 0,$$

当 z 最小时，边数达到最大。 □

命题 3.6。对于所有 $x \geq y \geq 1$ 和 $n > x + y$，由 (x, y) 边添加过程生成的任意阶为 n 的任务依赖图中，可能的最多边数为 $\binom{n}{2} + 1 - \binom{x}{2} - \binom{y+1}{2}$。

证明。对于下界，考虑 $Q_{x,y+1,n}$。$Q_{x,y+1,n}$ 中的 x 个初始顶点为 $1, \ldots, x$，而 $y+1$ 个终端顶点为 $n-y, \ldots, n$。$Q_{x,y+1,n}$ 有 $\binom{n}{2} - \binom{x}{2} - \binom{y+1}{2}$ 条边。若随机向 G 添加一条新边 (a, b)，则要么 $1 \leq a < b \leq x$，要么 $n-y \leq a < b \leq n$。在前一种情况下，由于这会导致任务依赖图的初始顶点少于 x 个，因此取消添加这条边。在后一种情况下，生成的任务依赖图有 x 个初始顶点和 y 个终端顶点，这将导致 (x, y) 边添加过程停止。因此，(x, y) 边添加过程在含有 n 个顶点的图上终止之前，可能的最多边数至少为 $\binom{n}{2} + 1 - \binom{y+1}{2} - \binom{x}{2}$。

对于上界，假设我们正在执行 (x, y) 边添加过程，并且该过程刚刚以任务依赖图 H 终止。设 H' 为该过程终止前一轮中的任务依赖图。那么 H' 或者具有 (a) x 个初始顶点和 $y+1$ 个终端顶点，(b) $x+1$ 个初始顶点和 y 个终端顶点，或者 (c) $x+1$ 个初始顶点和 $y+1$ 个终端顶点。如果 I 是 H' 中初始顶点的集合，而 T 是其中终端顶点的集合，那么根据引理 3.5，H' 中的边数至多为 $\binom{n}{2} - \binom{|I|}{2} - \binom{|J|}{2}$。在情况 (a)、(b) 和 (c) 中，当 $|I| = x$ 且 $|J| = y+1$ 时，该数量在情况 (a) 下达到最大值，这一结论可由以下事实推出：$\binom{x+1}{2} - \binom{x}{2} \geq \binom{y+1}{2} - \binom{y}{2}$。因此，$H$ 至多有 $\binom{n}{2} + 1 - \binom{y+1}{2} - \binom{x}{2}$ 条边。 □

根据上一个结果，即使不要求任务依赖图必须是 (x, y) 任务依赖图，由 (x, y) 边添加过程生成的任意任务依赖图，其可能的最多边数也可以得到相同的值。

定理 3.7. 对于所有 $x \geq y \geq 1$ 和 $n > x + y$，由 (x, y) 边添加过程生成的任意任务依赖图，其可能的最多边数为 $\binom{n}{2} + 1 - \binom{x}{2} - \binom{y+1}{2}$。

证明。假设 G 是对 n 个顶点执行 (x, y) 边添加过程所生成的任务依赖图。由于边添加过程在任何一轮中都不会增加初始顶点数或终端顶点数，因此，对于某些 $r \geq x$ 和 $s \geq y$，G 必定是一个 (r, s) 任务依赖图。若 $r = x$ 和 $s = y$，则根据命题 3.6，G 至多有 $\binom{n}{2} + 1 - \binom{x}{2} - \binom{y+1}{2}$ 条边。否则，$r \geq x+1$ 或 $s \geq y+1$。那么根据引理 3.5，G 至多有 $\binom{n}{2} - \binom{x}{2} - \binom{y+1}{2}$ 条边或至多有 $\binom{n}{2} - \binom{x+1}{2} - \binom{y}{2}$ 条边。注意

$$\binom{n}{2} - \binom{x}{2} - \binom{y+1}{2} \geq \binom{n}{2} - \binom{x+1}{2} - \binom{y}{2}$$

因此，自 $\binom{x+1}{2} - \binom{x}{2} \geq \binom{y+1}{2} - \binom{y}{2}$ 以来。由此可知，如果 G 不是一个 (x, y) 任务依赖图，那么 G 中的边数严格小于 $\binom{n}{2} + 1 - \binom{x}{2} - \binom{y+1}{2}$。 □

对于在 n 个顶点上进行的 $(x,\ y)$ 边添加过程，不难看出，所得任务依赖图中的边的期望数量为 $\Theta(n^2)$，我们将在下面的命题中证明这一点。请注意，这远大于通过 $(x,\ y)$ 边删除过程得到的任务依赖图中边的期望数量 $\Theta(n)$。

命题 3.8。对于固定的正整数 x 和 y，在 n 个顶点上的 (x,y)边添加过程中，边的期望数量为 $\Theta(n^2)$，其中下界中的常数取决于 (x,y)。

证明。 $O(n^2)$的上界是显而易见的，因为在一个阶为 n的有向有序图中，只有 $\binom{n}{2}$ 条形如(a,b)且 $a<b$ 的边。对于$\Omega(n^2)$的下界，令 S 表示顶点集为 $\{1,2,\ldots,n\}$的有向图中所有形如(a,b)且 $a<b$ 的边所构成的集合，令 P_x 表示所有形如(r,s)且 $1\le r<s\le x+1$的边所构成的集合，令 P_y 表示所有满足$n-y\le p<q\le n$且形如(p,q)的边所构成的集合，并令 $P=P_x\cup P_y$。首先，我们注意到，在 P中包含某条边之前，该过程不可能终止，因为否则， $[1,x+1]$ 中的所有顶点都将是初始顶点，而 $[n-y,n]$ 中的所有顶点都将是终端顶点。如果情况确实如此，那么我们将有 $x+1$ 个初始顶点和 $y+1$ 个终端顶点，因此，(x,y)边添加过程尚不会终止，因为每次添加边最多只能使初始顶点和终端顶点的数量各减少一个。为了得到该过程终止时边的期望数量的下界，我们可以计算向图中添加 P 中的某条边之前所经历的轮次数的期望值。

注意，在我们的随机边添加过程中，直到将 P 中的一个元素添加到任务依赖图的第一轮为止，S 中各元素的任意排列出现的概率都相同。为了界定将 P 中的第一个元素添加到任务依赖图之前的预期轮次数，我们考虑一种修改后的过程：以相同的概率生成S 中各元素的任意排列，然后界定排列中 P 中第一个元素的预期位置。注意，在修改后的过程中生成的随机排列中， P 中第一个元素的预期位置，等于在(x,y)边添加过程中将 P中的第一个元素添加到任务依赖图的预期轮次，因为直到将 P 中的一个元素添加到任务依赖图的第一轮为止， S 中各元素的任意排列出现的概率都相同。

给定 S 的任意一个排列， P 中的边将 $S-P$ 的元素划分为 $|P|+1$ 个序列，每个序列由 $(a_0,$ $a_1,\ldots,{}^{a_{|P|}})$ 条连续的边组成，其中

$$\sum_{i=0}^{|P|} a_i = |S|-|P|.$$

我们的目标是确定 a_0 的期望。由对称性可知，对于 $|P|+1$ 个元素的任意排列 π， P 中的边将 $S-P$ 中的元素划分为由 $(a_0,\ a_1,\ \ldots\ldots,\ {}^{a_{|P|}})$ 条连续边组成的 $|P|+1$ 个序列的概率，等于 P 中的边将 $S-P$ 中的元素划分为由 $\pi(a_0,\ a_1,\ \ldots\ldots,\ {}^{a_{|P|}})$ 条连续边组成的 $|P|+1$ 个序列的概率。因此， a_0 的期望为

$$\frac{|S|-|P|}{|P|+1} = \Omega(n^2).$$

□

在上一个证明中，对于所有 $x,\ y\ge 1$，上界的首项系数为 $\frac{1}{2}$。另一方面，下界的首项系数为

$$\frac{1}{(x+1)x+(y+1)y+2}.$$

例如，当 $x=y=1$ 时，上界中 $\frac{1}{2}$ 的首项系数与下界中 $\frac{1}{6}$ 的首项系数之间存在 3 倍的乘法间隙。

根据有向图的握手引理，我们得到以下推论，讨论在 n 个顶点上进行边添加过程时，预期最大入度和预期最大出度。

推论 3.9. 对于固定的正整数 x 和 y，在 n 个顶点上通过 (x,y) 边添加过程生成的任务依赖图，其预期最大入度为 $\Theta(n)$。此外，在 n 个顶点上通过 (x,y) 边添加过程生成的任务依赖图，其预期最大出度为 $\Theta(n)$。

证明。由于任务依赖图的阶为 n，最大入度和最大出度的上界 $O(n)$ 均显然成立。对于下界，注意到平均入度和平均出度的期望均为 $\frac{m}{n}$，其中 m 是边的期望数量。根据命题 3.8，平均入度和平均出度的期望均为 $\Theta(n)$，因此最大入度和最大出度的期望均为 $\Theta(n)$。

□

接下来，我们证明，在包含 n 个顶点的图上，通过 (x,y) 边添加过程生成的任务依赖图中，孤立顶点的期望数量为 $o(1)$。换言之，随着 n 趋于无穷大，孤立顶点的期望数量趋于 0。

定理 3.10。 对于固定的正整数 x 和 y，在包含 n 个顶点的图上，通过 (x,y) 边添加过程生成的任务依赖图中，孤立顶点的期望数量为 $o(1)$。

证明。假设 n 足够大，使得 $\binom{n}{2} - (n-1) - \lceil n\sqrt{n} \rceil > \binom{x+1}{2} + \binom{y+1}{2}$ 且 $n > \binom{x+1}{2} + \binom{y+1}{2} + 1$。与命题 3.8 的证明一样，令 P_x 为形如 (r,s) 的边的集合，其中 $1 \le r < s \le x+1$，令 P_y 为形如 (p,q) 的边的集合，其中 $n-y \le p < q \le n$，并令 $P = P_x \cup P_y$。注意，在具有 n 个顶点的 (x,y) 边添加过程中，在首次添加 P 中某个元素之前的每一轮中，每条剩余边被添加的概率都相同。我们首先证明，在前 $\lceil n\sqrt{n} \rceil$ 轮中添加一条属于 P 的边的概率为 $o(1)$。然后利用这一结果证明，孤立顶点的期望数量为 $o(1)$。

首先，注意到 $|P| = \binom{x+1}{2} + \binom{y+1}{2} = O(1)$。因此，在前 $\lceil n\sqrt{n} \rceil$ 轮中没有向 P 添加边的概率为

$$\prod_{i=1}^{\lceil n\sqrt{n} \rceil} \frac{\binom{n}{2} - |P| + 1 - i}{\binom{n}{2} + 1 - i} \ge \left(1 - \frac{|P|}{\binom{n}{2} + 1 - \lceil n\sqrt{n} \rceil}\right)^{\lceil n\sqrt{n} \rceil} = 1 - o(1).$$

因此，在前 $\lceil n\sqrt{n} \rceil$ 轮中向 P 添加边的概率为 $o(1)$。

孤立顶点的期望数量为 $p_1 E_1 + p_2 E_2$，其中 p_1 是在前 $\lceil n\sqrt{n} \rceil$ 轮中向 P 添加边的概率，E_1 是在前 $\lceil n\sqrt{n} \rceil$ 轮中向 P 添加边这一条件下孤立顶点的期望数量，p_2 是在前 $\lceil n\sqrt{n} \rceil$ 轮中未向 P 添加边的概率（因此 $p_2 = 1 - p_1$），而 E_2 是在前 $\lceil n\sqrt{n} \rceil$ 轮中未向 P 添加边这一条件下孤立顶点的期望数量。我们证明了 $p_1 = o(1)$ 和 $p_2 = 1 - o(1)$，并且 $E_1 \le \max(x, y) = O(1)$，因此 $p_1 E_1 = o(1)$。

假设在前 $\lceil n\sqrt{n} \rceil$ 轮中没有添加 P 中的任何边，因此，除 P 中的边以外，其余所有边在前 $\lceil n\sqrt{n} \rceil$ 轮中的每一轮被添加的可能性都相同。下面我们用前 $\lceil n\sqrt{n} \rceil$ 轮后孤立顶点的期望数量给 E_2 一个上界。对于任务依赖图中的任意顶点 v，将 P 中与 v 相连的边集记为 Q_v。那么，在前 $\lceil n\sqrt{n} \rceil$ 轮中没有添加任何与 v 相连的边的概率至多为

$$\prod_{i=1}^{\lceil n\sqrt{n} \rceil} \frac{\binom{n}{2} - (n-1) - |P| + |Q_v| + 1 - i}{\binom{n}{2} - |P| + 1 - i} \le \left(1 - \frac{n - 1 - |Q_v|}{\binom{n}{2} - |P|}\right)^{\lceil n\sqrt{n} \rceil} \le e^{-\frac{\lceil n\sqrt{n} \rceil (n-1-|P|)}{\binom{n}{2} - |P|}}.$$

因此，根据期望的线性性，

$$E_2 \le n e^{-\frac{\lceil n\sqrt{n} \rceil (n-1-|P|)}{\binom{n}{2} - |P|}} = o(1).$$

因此，$p_1 E_1 + p_2 E_2 = o(1)$。

□

(x,y)	5	6	7	8	9	10	11	12	13	14
(1, 2)	0.947	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
(1, 3)	0.507	0.748	0.908	0.965	0.992	0.998	1.000	1.000	1.000	1.000
(1, 4)	0.051	0.229	0.484	0.733	0.883	0.971	0.985	0.998	0.998	1.000
(2, 3)	0.687	0.870	0.968	0.994	0.998	1.000	1.000	1.000	1.000	1.000
(2, 4)	0.086	0.332	0.572	0.806	0.926	0.978	0.990	0.996	1.000	1.000
(3, 4)	0.258	0.590	0.796	0.908	0.981	0.992	0.999	0.999	1.000	1.000

表1：此表展示了在 n 个顶点上运行 1000 次 (x,y) 边删除过程的结果，并计算该过程生成 (x,y) 任务依赖图的试验次数占总试验次数的比例。顶点数 n 从 5 到 14 不等。

(x,y)	5	6	7	8	9	10	11	12	13	14
(1, 2)	0.923	0.962	0.964	0.98	0.993	0.989	0.998	0.995	1.000	1.000
(1, 3)	0.715	0.828	0.903	0.914	0.954	0.968	0.978	0.988	0.982	0.992
(1, 4)	0.382	0.616	0.727	0.825	0.864	0.916	0.931	0.937	0.958	0.963
(2, 3)	0.958	0.986	0.988	0.998	0.998	0.999	1.000	1.000	1.000	1.000
(2, 4)	0.706	0.890	0.954	0.981	0.985	0.988	0.994	0.994	0.999	0.999
(3, 4)	0.907	0.982	0.994	0.999	1.000	0.999	1.000	1.000	1.000	1.000

表2：此表显示了在 n 个顶点上运行 1000 次（x,y）边添加过程的结果，针对（x,y）$=$ (1,2),(1,3),(1,4),(2,3),(2,4),(3,4)，并计算该过程生成（x,y）任务依赖图的试验次数占总试验次数的比例。顶点数 n 的范围为 5 到 14。

4. 实验结果

我们使用 Python [13] 实现了 (x,y) 边删除过程和 (x,y) 边添加过程。针对每种过程进行了数千次试验后，我们对这两种过程中边的期望数量、预期最大有向路径长度，以及生成 (x,y) 任务依赖图的概率提出了若干猜想。表1展示了对 (x,y) 边删除过程进行 1000 次试验的结果。具体而言，对于 (x,y) $=$(1, 2)、(1, 3)、(1, 4)、(2, 3)、(2, 4) 和 (3, 4)，我们计算了该过程生成 (x,y) 任务依赖图的试验次数占总试验次数的比例。我们将顶点数量设定在 5 至 14 之间。回顾一下，对于每个 $x \geq 1$，(x,x) 边删除过程都会生成一个 (x,x) 任务依赖图；但是，当 $x \neq y$ 时，(x,y) 边删除过程可能生成初始顶点数量少于 x 或终端顶点数量少于 y 的任务依赖图。

根据表1中的结果，我们猜想，对于每个固定的数对 $x,y \geq 1$，当顶点数量趋于无穷时，在 (x,y) 边删除过程结束时获得 (x,y) 任务依赖图的概率趋于 1。

猜想4.1。对于固定的 $x,y \geq 1$，若 $r_{x,y,n}$ 表示在 n 个顶点上进行的(x,y) 边删除过程生成(x,y) 任务依赖图的概率，那么

$$\lim_{n\to\infty} r_{x,y,n} = 1.$$

与前一个实验类似，表2展示了对（x，y）边添加过程进行1000次试验的结果；该过程针对（x，y）$=$(1，2)，(1，3)，(1，4)，(2，3)，(2，4)，(3，4)运行，并计算该过程生成（x，y）任务依赖图的试验次数占总试验次数的比例。与上一个表格一样，我们让顶点数从5变化到14。回顾一下，对于每个 $x \geq 1$，（x，x）边添加过程都会生成一个（x，x）任务依赖图；但当 $x \neq y$时，（x，y）边添加过程可能生成初始顶点多于 x 或终端顶点多于 y 的任务依赖图。

与先前关于（x， y）边删除过程的实验类似，我们猜想，对于每个固定的数对 x， $y \geq 1$，随着顶点数趋于无穷大，在（x， y）边添加过程结束时得到（x， y）任务依赖图的概率趋近于1。

猜想 4.2。 对于固定的 x、 $y \geq 1$，如果 $a_{x,\ y,\ n}$ 表示在 n 个顶点上进行 (x, y) 边添加过程生成一个(x, y) 任务依赖图的概率，那么

$$\lim_{n\to\infty} a_{x,y,n} = 1.$$

我们能够通过一种类似于命题 2.1 和 3.2 的证明方法，证明猜想 4.2。

定理 4.3. 猜想 *4.2* 是正确的。

证明。 假设 G 是通过在 n 个顶点上执行 (x, y) 边添加过程生成的一个 (r, s) 任务依赖图，且满足 $(r, s) \neq (x, y)$。那么我们必有 $r > x$ 或 $s > y$，因此不失一般性，设 $r > x$。注意，我们必有 $s = y$，否则我们还可以添加一条边，而所得任务依赖图仍然至少具有 x 个初始顶点和 y 个终端顶点。我们首先断言， G 中任意一对不同的非终端顶点 $u < v$ 之间都必须存在一条边。如果存在这样一对顶点，那么我们就可以添加边 (u, v)，而不减少终端顶点的数量，因此 (x, y) 边添加过程就不会在 G 上终止。

接下来，我们断言 G 必须有一个孤立顶点。为说明这一点，假设 G 没有孤立顶点。那么 G 只有一个初始顶点，因为每一对非终端顶点之间都必须有一条边。因此 $r = 1$，这与 $r > x$矛盾，所以 G 必须有一个孤立顶点。

根据马尔可夫不等式， G含有孤立顶点的概率至多为孤立顶点的期望数量，根据定理 3.10，该数量为 $o(1)$。因此， G 不是一个（x， y）任务依赖图的概率为 $o(1)$，所以 $a_{x,\ y,\ n} = 1 - o(1)$。□

我们还对 (1,1) 边删除过程和 $(1, 1)$ 边添加过程进行了 1000 次试验，并计算了平均边数和平均最大有向路径长度。我们针对顶点数从 3 到 40 的情况运行了这些过程。图 4 展示了平均值的曲线图及其拟合曲线。

根据第2节的结果，我们知道，在 n 个顶点上通过 (1,1) 边删除过程生成的任务依赖图中，边的期望数量为 $\Theta(n)$，其上界为$2n - 4$，下界为 $n - 1$。根据实验结果，我们猜想该期望更接近下界。

猜想 4.4. 如果 e_n 表示在 n 个顶点上通过$(1,1)$ 边删除过程生成的任务依赖图中的边的期望数量，那么

$$\lim_{n\to\infty} \frac{e_n}{n} \approx 1.35.$$

引理 4.5. 对于任意 $1 \leq r < s \leq n$，边 (r, s) 保留在通过$(1,1)$ 边删除过程生成的$(1,1)$ 任务依赖图中的概率至多为

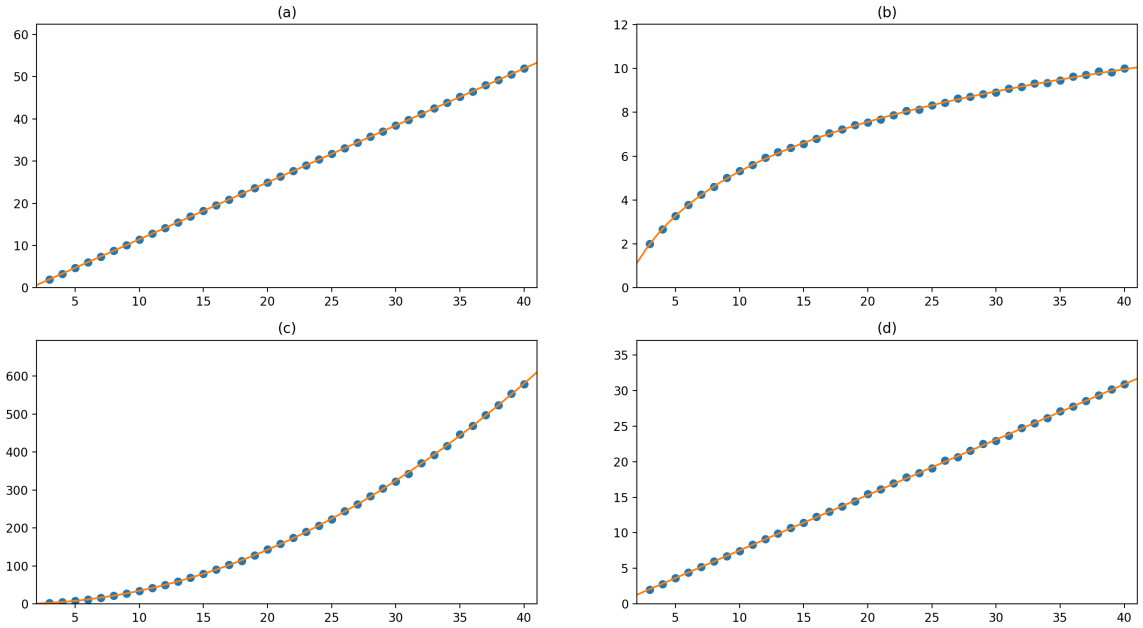
$$\frac{1}{s-1} + \frac{1}{n-r} - \frac{1}{n-2+s-r}.$$

证明。 回想一下，从一个阶为 n 的最大连通 (1，1) 任务依赖图 G 开始，边删除过程会随机排列全部 $\binom{n}{2}$ 条边，然后逐条删除，但如果删除某条边会使任何内部顶点变为外部顶点，则不予删除。考虑边集 $A_s = \{(p,\ s)\colon\ p < s\}$ 以及$B_r = \{(r,\ t)\colon\ r < t\}$。边 $(r,\ s)$ 要保留在最终任务依赖图中的必要条件是：在边的排列中，它必须在 B_r 或 A_s 中最后出现。否则，当考虑删除边 $(r,\ s)$ 时，顶点 r 的出度将大于 1，顶点 s 的入度将大于 1，而 $(r,\ s)$ 将已被删除。根据容斥原理，这一必要条件发生的概率为

$$\frac{1}{|A_s|} + \frac{1}{|B_r|} - \frac{1}{|A_s \cup B_r|} = \frac{1}{s-1} + \frac{1}{n-r} - \frac{1}{n-2+s-r}.$$

图1：图(a)显示了当 n 从 3 变化到 40 时， n 个顶点上的 (1，1) 边删除过程相对于 n 的平均边数。最佳拟合直线为 $y = 1.349x - 2.029$。
图 (b) 显示了在 n 个顶点上执行 (1,1) 边删除过程时， n 从 3 变化到 40 的情况下，关于 n 的平均最大有向路径长度。最佳拟合对数曲线为 y =3.679 ln($x + 1.799$) $- 3.776$。

图（c）显示了当 n 从 3 变化到 40 时，在 n 个顶点上进行（1,1）边添加过程时，相对于 n 的平均边数。最佳拟合二次函数为 $y = 0.369x^2 - 0.258x + 0.007$。
图 (d) 显示了在 n 个顶点上进行 (1,1) 边添加过程时，平均最大有向路径长度相对于 n 的变化，其中 n 从 3 变化到 40。最佳拟合直线为 $y = 0.779x - 0.273$。所有系数均已四舍五入到小数点后三位。



□

在上一个证明中，请注意，边 (r, s) 最后出现在 B_r 或A_s 的边中，只是 (r, s) 保留在最终任务依赖图中的必要条件。具体来说，这一条件并不足以保证 (r, s) 保留在最终任务依赖图中。例如，假设存在边 $(r, t) \in B_r$和 $(p, s) \in A_s$，它们分别不是B_r 和 A_s 的边中最后出现的边。如果 (r, t) 最后出现在 A_t的边中，且 (p, s) 最后出现在 B_p的边中，那么这两条边都有可能保留在最终任务依赖图中。在这种情况下，即使边 (r, s) 最后出现在 B_r 或 A_s的边中，它也会从任务依赖图中删除，因为删除它不会使任何顶点变成外部顶点。

命题 4.6. 若 e_n 表示在 n 个顶点上通过 $(1,1)$ 边删除过程生成的任务依赖图中的边的期望数量，则

$$\lim_{n\to\infty} \frac{e_n}{n} \leq 3 - 2\ln 2 \approx 1.614.$$

证明。 我们通过对引理 4.5 中的概率求和来证明：

$$\frac{e_n}{n} \leq \frac{1}{n}\left(\sum_{r=1}^{n-1}\sum_{s=r+1}^{n} \frac{1}{s-1} + \frac{1}{n-r} - \frac{1}{n-2+s-r}\right) = \frac{1}{n}\left(n-1+n-1-\sum_{d=1}^{n-1}\frac{n-d}{n-2+d}\right).$$

通过设置 $z = \frac{n-2+d}{n}$，最后一项趋近于 $-\int_1^2 \frac{2-z}{z} dz = 1 - 2\ln 2$。因此

$$\lim_{n\to\infty} \frac{e_n}{n} \leq 3 - 2\ln 2.$$

□

根据第 3 节的结果，我们知道，在 n 个顶点上通过 (1，1) 边添加过程生成的任务依赖图，其边的期望数量为 $\Theta(n^2)$，上界约为 $\frac{1}{2}n^2$，下界约为 $\frac{1}{4}n^2$。根据实验结果，我们推测该期望略微接近下界。

猜想 4.7。 若 f_n 表示在 n 个顶点上通过 $(1，1)$ 边添加过程生成的任务依赖图的边的期望数量，则

$$\lim_{n\to\infty} \frac{f_n}{n^2} \approx 0.37.$$

关于预期最大有向路径长度的实验结果最为有趣。对于 (1，1) 边添加过程，随着 n 变化，最大有向路径长度平均值的最佳拟合曲线显然是线性的。然而，对于 (1，1) 边删除过程，预期最大有向路径长度似乎是 n 的对数函数。

猜想 4.8。 对于在 n 个(1,1) 顶点上的边添加过程，所得任务依赖图的预期最大有向路径长度为 $\Theta(n)$。

猜想 4.9。 对于具有 n 个(1,1) 边删除过程，所得任务依赖图的预期最大有向路径长度为 $\Theta(\log n)$。

我们没有证明猜想 4.9，但针对另一种生成任务依赖图的随机过程，找到了一个相关结果。令顶点数为 n 的随机有向树过程 为一种随机过程，它从 n 个孤立顶点 1，2，.....，n开始，并在第 $s > 0$轮从所有满足条件的边中均匀随机地添加一条边（r，$s+1$），其中 $r \leq s$。注意，对于随机有向树过程在 n 个顶点上生成的任意任务依赖图，其底层无向图都是一棵树。

引理 4.10。 设 G 是由随机有向树过程在 n 个顶点上生成的阶为 n 的任务依赖图。那么对于 $k > 1$，从顶点 1 到顶点k的唯一路径的期望长度为 $1 + \frac{1}{2} + \ldots + \frac{1}{k-1} = \Theta(\text{对数 } k)$。

证明。 用 $f(k)$ 表示顶点 1 到 k 之间唯一一条路径的期望长度。顶点 $r < s$ 相邻的概率为 $\frac{1}{s-1}$，因此我们有 $f(1) = 0$，对于 $k > 1$

$$f(k) = 1 + \frac{1}{k-1}\left(f(1) + \ldots + f(k-1)\right).$$

化简后得

$$f(k+1) - f(k) = \frac{1}{k}.$$

□

5. 未来方向与未解决问题

我们的实验结果表明，随着顶点数量增加，(x, y) 边添加过程和 (x, y) 边删除过程都以趋近于 1 的概率生成 (x, y) 任务依赖图。我们证明了这一结论对 (x, y) 边添加过程成立，但对于 (x, y) 边删除过程，仍有待证明。

我们确定了对于所有 x，$y \geq 1$ 以及 $n \geq \max(x, y)$，阶为 n 的最小（x，y）任务依赖图中可能的最多边数，并证明该数等于最大

通过在 n 个顶点上执行（x, y）边删除过程所生成的任意任务依赖图中的可能的最多边数。我们还刻画了极值任务依赖图。

还有许多相关问题尚待解答。首先，给定一个任务依赖图，任务有多少种排序方式？具体而言，阶为 n 的 (x, y) 任务依赖图中，任务排序方式的最大可能数量是多少？

我们在以下定理中对这个问题取得了一些进展。具体来说，我们解决了所有 $n \geq 1$ when $x = y = 1$ 的情形。我们还解决了所有 $x, y \geq 1$ when $\max(x, y) \leq n \leq \max(x, y) + 2$ 的情形。

命题 5.1。对于 $x \geq y \geq 1$，阶为 n 的（x, y）任务依赖图中，任务排序方式的最大可能数量为：当 $n = x$ 时为 $n!$（此时 $x = y$）；当 $n = x + 1$ 时为 $(x+1)! \, 2^{x-y+2}$；当 $n = x + 2$ 且 $y > 1$ 时为 $(x+2)! \, 2^{(x-y+2)}$；当 $n = x + 2$ 且 $y = 1$ 时为 $(x+2)! \, 2^{(x-y+3)}$。阶为 n 的（1，1）任务依赖图中，若 $n \geq 2$，任务排序方式的最大可能数量为（$n - 2$）！。

证明。对于 $n = x$，上界显然，因为一个阶为 n的有向图的顶点共有 $n!$ 种排列方式。阶为 n 且所有顶点均为孤立的有向图达到了这个上界。

对于 $n = x + 1$，我们在定理 2.10 的证明中指出，任何 (x, y) 任务依赖图都有$y - 1$ 个孤立顶点、$x - y + 1$ 个非孤立初始顶点、一个非孤立终端顶点，以及从每个非孤立初始顶点指向该非孤立终端顶点的边。因此，任何使非孤立终端顶点排在所有非孤立初始顶点之后的顶点排序，都与该任务依赖图一致。随机排列顶点时，非孤立终端顶点排在所有非孤立初始顶点之后的概率为 $\frac{1}{x-y+2}$，而总共有 $n! = (x + 1)!$ 种排序，因此与该任务依赖图一致的排序数量为

$$\frac{n!}{x - y + 2} = \frac{(x + 1)!}{x - y + 2}.$$

对于 $n = x + 2$，我们在定理 2.10 的证明中指出，阶为 n 的 (x, y) 任务依赖图 G 有两种可能：G 要么有 (a) $y - 1$ 个孤立顶点、$x - y + 1$ 个非孤立初始顶点、一个非孤立终端顶点和一个内部顶点；要么有 (b) $y-2$ 个孤立顶点、$x - y + 2$ 个非孤立初始顶点、两个非孤立终止顶点，且没有内部顶点。两种情况下，G 都有 $2n - x - y - 2$ 条边。注意，只有当 $y > 1$ 时，情况 (b) 才可能发生。

对于情况 (a)，假设 G 有 t 个非孤立初始顶点，它们与内部顶点相连，还有 $x - y + 1 - t$ 个非孤立初始顶点，它们与非孤立终端顶点相连，其中$1 \leq t \leq x - y + 1$。孤立顶点有 $(y - 1)!$ 种排序方式，孤立顶点的位置有 $\binom{x+2}{y-1}$ 种选择方式，非孤立终端顶点的位置有 1 种选择方式，内部顶点及与其相连的 t 个非孤立初始顶点的位置有 $\binom{x-y+2}{t+1}$ 种选择方式，与内部顶点相连的 t 个非孤立初始顶点有 $t!$ 种排序方式，与非孤立终端顶点相连的非孤立初始顶点有 $(x - y + 1 - t)!$ 种排序方式。总计为

$$(y - 1)!\binom{x + 2}{y - 1}\binom{x - y + 2}{t + 1}t!(x - y + 1 - t)!$$

顶点的排列，这可简化为

$$\frac{(x + 2)!}{(x - y + 3)(t + 1)}.$$

当 $t = 1$ 时，该值达到最大，从而得到

$$\frac{(x + 2)!}{2(x - y + 3)}.$$

因此，对于 $n = x + 2$ 和 $y = 1$，该量是排列的最大数量。

对于情况 (b)，假设 t 非孤立初始顶点与第一个非孤立终端顶点之间存在边，且 $1 \leq t \leq x - y + 1$。这形成了一个大小为 $t + 1$ 的第一个连通分量。此外，

假设 $x-y+2-t$ 非孤立初始顶点与第二个非孤立终端顶点之间有边。这构成了一个大小为 $x-y+3-t$ 的第二个连通分量。孤立顶点有 $(y-2)!$ 种排列方式，$\binom{x+2}{y-2}$ 种选择孤立顶点位置的方式，$\binom{x-y+4}{t+1}$ 种选择第一个连通分量中顶点位置的方式，第一个连通分量中的顶点有 $t!$ 种排列方式，第二个连通分量中的顶点有 $(x-y+2-t)!$ 种排列方式。因此总共有

$$(y-2)!\binom{x+2}{y-2}\binom{x-y+4}{t+1}t!(x-y+2-t)!$$

顶点的阶数排列，简化为

$$\frac{(x+2)!}{(t+1)(x-y+3-t)}.$$

当 $t=1$ 时，该值达到最大值，因此得到

$$\frac{(x+2)!}{2(x-y+2)}.$$

因此，对于 $n=x+2$ 和 $y>1$，该数量是排列的最大数量。

如果 $x=y=1$ 且 $n\geq 2$，那么有 $n-2$ 个内部顶点、一个初始顶点和一个终端顶点。初始任务必须先于所有其他任务，最后一个任务必须晚于所有其他任务，而内部任务的阶数最多有 $(n-2)!$ 种，因为如果边不对内部任务之间的先后顺序造成限制，那么内部任务就有 $(n-2)!$ 种排序方式。这个 $(n-2)!$ 的上界由 $T_{1,\ 1,\ n}$ 达到。 □

对于上一条定理所涵盖的每种情况，值得注意的是，在一个阶为 n 的 (x,y) 任务依赖图中，任务的可能排列方式的最大数量，是由一个阶为 n 且具有可能的最多边数的最小 (x,y) 任务依赖图达到的。确定这一模式是否对所有 $x,y\geq 1$ 和 $n\geq \max(x,y)$ 都继续成立，将会很有趣。

还有一些与我们的结果相关的有趣概率问题。我们证明了，对于所有$x,y\geq 1$，在具有 n 个顶点的图上执行 (x,y) 边删除过程所得到的随机任务依赖图中的边的期望数量为 $\Theta(n)$。我们的上界的首项系数为 2，而下界的首项系数为 1，因此，缩小这两个界之间的差距将是一个值得研究的问题。我们还证明了，对于所有 $x,y\geq 1$，在具有 n 个顶点的图上执行 (x,y) 边添加过程所得到的随机任务依赖图中的边的期望数量为 $\Theta(n^2)$。上界的首项系数为$_{1\ 2}$，而下界的首项系数为$^{1}{}_{(x+1)x+(y+1)y+2}$。

基于我们的结果，对于任意 $x,y\geq 1$，我们描述了一种随机过程，仅通过添加边和删除边来生成阶为 n、具有 m 条边的 (x,y) 任务依赖图。当 $x=y=1$ 时，这一随机过程为 [22] 中用于随机生成阶为 n、具有 m 条边的 (1,1) 任务依赖图的方法提供了一种替代方案。

命题 5.2. 对于任意的 x、$y\geq 1$、$n>\max(x$、$y)+1$，以及 $2n-x-y-2\leq m\leq \binom{n}{2}-(x+y)n$，存在一个仅涉及边添加和边删除的随机过程，它以高概率生成一个阶为 n、具有 m 条边的$(x$、$y)$ 任务依赖图。

证明。考虑以下随机过程。首先，在 n 个顶点上运行 (x,y) 边添加过程，直至过程终止。以高概率，该过程会生成一个 (x,y) 任务依赖图G。令 r 为 G 中的边数。如果 $r=m$，那么 G 就是最终任务依赖图。如果$r<m$，那么 $r<\binom{n}{2}-(x+y)n$。由于至多有 $(x+y)n$ 条缺失边，其添加会减少初始顶点或终端顶点的数量，因此必定存在一条缺失边，添加后不会减少初始顶点或终端顶点的数量。因此在这种情况下，我们可以从G 中均匀随机地添加边，同时不减少初始顶点和终端顶点的数量，直至所得任务依赖图具有 m 条边。

如果 $r>m$，则从 G 开始执行 $(x,\ y)$ 边删除过程，直到得到的任务依赖图具有 m 条边。请注意，给定假设，针对 n 个顶点完整执行 $(x,\ y)$ 边删除过程，始终会得到至多具有 $2n-x-y-2$ 条边的任务依赖图

即 $n > \max(x, y) + 1$。此外，在从 G 开始执行 (x, y) 边删除过程的每一轮中，只要 G 是一个 (x, y) 任务依赖图，就能保证得到的任务依赖图是一个 (x, y) 任务依赖图。因此，只要 $2n - x - y - 2 \leq m \leq \binom{n}{2} - (x + y)n$，这个随机过程就会以很高的概率生成一个阶为 n 、具有 m 条边的 (x, y) 任务依赖图。 □

关于预期值，还有其他一些自然的问题，同时涉及 (x, y) 边添加过程和 (x, y) 边删除过程。

1. 生成的任务依赖图的预期最大有向路径长度是关于 x、 y 和 n 的什么函数？2. 生成的任务依赖图中任务的排列数的预期值是关于 x、 y 和 n 的什么函数？3. 生成的任务依赖图中给定有向无环图 H 的副本数的预期值是关于 x、 y 和 n 的什么函数?

除了预期值之外，还有其他一些自然的概率问题。例如，求阶为 n 的生成任务依赖图恰有 m 条边的概率、其最大有向路径长度为 m 的概率、排列数为 m 的概率，或给定有向无环图 H 的副本数为 m 的概率。

致谢

Shen-Fu Tsai 获得台湾科技部项目 MOST 111-2115-M-008-010-MY2 的资助。我们感谢 Lance Menthe 对本文提出的宝贵意见。

参考文献

[1] G.E. Blelloch 和 P.B. Gibbons, 《在线程间有效共享缓存》。载于《第十六届 ACM 并行算法与体系结构年度研讨会论文集》（2004），235-244。[2] G.E. Blelloch、P.B. Gibbons 和 Y. Matias, 《为具有细粒度并行性的语言提供可证明高效的调度》。J. ACM 46（1999），281-321。[3] R.D. Blumofe 和 P.A. Lisiecki, 《在工作站网络上进行自适应且可靠的并行计算》。载于《USENIX 1997 UNIX 与高级计算系统年度技术会议论文集》（1997），133-147。[4] R.D. Blumofe 和 C.E. Leiserson, 《通过工作窃取调度多线程计算》。J. ACM 46（1999），720-748。[5] R.D. Blumofe 和 C.E. Leiserson, 《多线程计算的空间高效调度》。SIAM J. Comput. 27（1998），202-229。[6] F. W. Burton 和 M.R. Sleep, 《在处理器虚拟树上执行函数式程序》。载于《1981 年函数式编程语言与计算机体系结构会议论文集》（1981），187-194。[7] S. Chen、P.B. Gibbons、M. Kozuch、V. Liaskovitis、A. Ailamaki、G.E. Blelloch、B. Falsafi、L. Fix、N. Hardavellas、T.C. Mowry 和 C. Wilkerson, 《为 CMP 上的建设性缓存共享调度线程》。载于《第十九届 ACM 并行算法与体系结构年度研讨会论文集》（2007），105-115。

[8] 美国国会研究服务处，《国防入门：LGM-35A“哨兵”洲际弹道导弹》。2023年1月10日。[9] D.L. Eager、E.D. Lazowska 和 J. Zahorjan，《接收方发起与发送方发起的自适应负载共享比较》。Perform. Eval. 6（1986），53-68。[10] M. Engwall，《PERT：北极星与项目执行的现实》。《国际项目管理期刊》5（2012），595-616。[11] R. Feldmann、P. Mysliwietz 和 B. Monien，《大规模并行系统上的博弈树搜索》。Adv. Comput. Chess 7（1993），203-219。[12] R. Finkel 和 U. Manber，《DIB——回溯的分布式实现》。ACM Trans. Program. Lang. Syst. 9（1987），235-256。[13] J. Geneson 和 S. Tsai，《边添加与边删除过程的代码》。https://github. com/jgeneson/tdgraphs。[14] M. Halbherr、Y. Zhou 和 C.F. Joerg，《使用延续传递线程进行 MIMD 式并行编程》。载于《第二届大规模并行国际研讨会：硬件、软件与应用论文集》（1994），1-10。[15] R.H. Halstead Jr.，《Multilisp 的实现：多处理器上的 Lisp》。载于《1984 年 ACM LISP 与函数式编程研讨会会议录》（1984），9-17。[16] R.M. Karp 和 Y. Zhang，《用于回溯搜索和分支定界计算的随机化并行算法》。J. ACM 40（1993），765-789。[17] J.E. Kelley, Jr. 和 M.R. Walker，《关键路径法的起源：个人史》。PM Network 3（1989），7-22。[18] B.C. Kuszmaul，《同步 MIMD 计算》。博士学位论文，电气工程与计算机科学系，麻省理工学院，剑桥，马萨诸塞州（1994），1-162。[19] R. Mirchandaney、D. Towsley 和 J.A. Stankovic，《异构分布式系统中的自适应负载共享》。J. Parallel Distrib. Comput. 9（1990），331-346。[20] E. Mohr、D.A. Kranz 和 R.H. Halstead Jr.，《惰性任务创建：提高并行程序粒度的一种技术》。IEEE Trans. Parall. Dist. Syst. 2（1991），264-280。[21] L. Rudolph、M. Slivkin-Allalouf 和 E. Upfal，《一种用于在并行机器上分配任务的简单负载均衡方案》。载于《第三届 ACM 并行算法与体系结构年度研讨会论文集》（1991），237-245。[22] D. Snyder、C. Johnson、P. Rockstroh、L. Menthe 和 B.E. Bennett，《陆基战略威慑计划的图论算法：优先级排序与调度》。加利福尼亚州圣莫尼卡：兰德公司（2021），1-68。[23]B. Van Houdt，《非指数型作业大小的大规模系统中，随机工作窃取与共享的比较》。《IEEE/ACM 网络汇刊》27（2019），2137-2149。[24] M.T. Vandevoorde 和 E.S. Roberts，《WorkCrews：一种控制并行性的抽象》。《国际并行编程期刊》17（1988），347-366。