前缀排序确定有限自动机：递归算法

Nicola Cotumaccio
意大利 GSSI、加拿大达尔豪斯
大学

摘要

过去三十年来，研究者提出了许多用于构建字符串后缀数组的算法。2021 年，后缀数组的概念从字符串扩展到了确定性有限自动机，并证明了由此得到的数据结构可以在 $O(m^2 + n^{5/2})$ 时间内构建，其中 n 是状态数，m 是边数 [SODA 2021]。最近，有研究提出了运行时间为 $O(mn)$ 和 $O(n^2 \log n)$ 的算法 [CPM 2023]。

本文受到 Farach 用于构建后缀树的算法 [FOCS 1997]启发，提出了一种 $O(n^2)$ 递归算法，从而改进了此前的界。为此，我们深入探讨图丰富的字典序结构和组合结构，为这段引人入胜的探索之旅贡献力量；这段旅程或许最终能帮助我们解决长期悬而未决的问题：构建图的后缀树。

2012 年 ACM 学科分类计算理论 → 图算法分析；计算理论 → 模式匹配

关键词和短语后缀数组、Burrows-Wheeler 变换、FM 索引、递归算法、图论、模式匹配。

资助本研究部分由 Dante Labs 资助。

致谢感谢 Nicola Prezza 指出了论文 [21]。

arXiv:2305.02526v2 [cs.DS] 6 Oct 2023

1 引言

字符串的后缀树 [30] 是一种由 Weiner 于 1973 年提出的多功能数据结构，可用于解决大量组合问题，例如确定某个模式是否出现在字符串中、计算匹配统计信息、搜索正则表达式、计算字符串的 Lempel-Ziv 分解以及查找回文。Gusfield 所著的书籍 [17] 用近 150 页介绍了后缀树的应用，并强调了这些应用在生物信息学中的重要性。然而，过去几十年中基因组数据的大幅增长要求使用能够高效支持模式匹配查询的节省空间数据结构，而后缀树的空间消耗过高。1990 年，Manber 和 Myers 发明了后缀数组 [24] ，将其作为后缀树的节省空间替代方案。虽然后缀数组不具备后缀树的全部功能，但仍然能够解决模式匹配查询。后缀数组开启了数据压缩的新篇章，并最终促成了与后缀数组密切相关的数据结构的发明，其中尤以 Burrows-Wheeler 变换 [4] 和 FM 索引 [13, 15], 为代表；这些结构对序列组装产生了深远影响 [29]。

后缀数组的影响促使人们投入大量精力，致力于设计高效的后缀数组构造算法；这里的“高效”涉及多种指标（最坏情况运行时间、平均运行时间、空间占用、在真实数据上的表现等）；关于这一主题的全面综述，请参见 [27] 。我们来关注最坏情况运行时间。Manber 和 Myers 通过一种前缀倍增算法 [24]，在 $O(n \log n)$ 时间内构造长度为 n 的字符串的后缀数组。1997 年，Farach 提出了一种递归算法，可在线性时间内为整数字母表上的字符串构造后缀树 [11]。在随后的几年里，Farach 算法的递归范式被用于开发大量线性时间算法，以构造后缀数组 [22, 19, 20, 25]。这些算法都充分利用了字符串后缀的字典序结构，

2 前缀排序确定有限自动机：一种递归算法

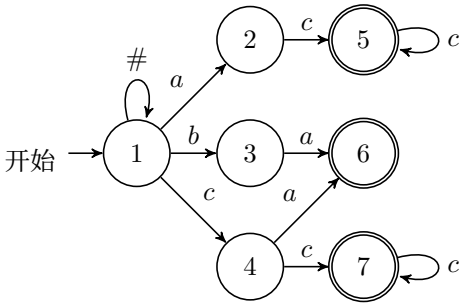
递归地将计算一个字符串的后缀数组的问题，化简为计算一个更短字符串的后缀数组的问题（诱导排序）。

不仅针对字符串，也针对标记图执行模式匹配查询，是一个活跃的研究课题。最近，Equi 等人证明，除非正交向量假设不成立，否则不存在能在 $O(m^{1-\epsilon}|P|)$ 时间或 $O(m|P|^{1-\epsilon})$ 内解决任意图上的 模式匹配查询的算法（其中 m 是边数， P 是模式， $\epsilon > 0$）。另一方面，多年来，将字符串的后缀按字典序排序这一思想已推广到图，从而产生了能够支持图上模式匹配查询的紧凑数据结构。后缀数组、Burrows-Wheeler 变换和 FM 索引背后的机制最初被推广到树[12, 14]；后来又推广到德布鲁因图 [3, 23]（可用于欧拉序列组装 [18]）。随后，该机制扩展到所谓的 Wheeler 图[16, 1]，，最终推广到任意图和自动机 [8, 6]。按字典序对到达自动机状态的字符串进行排序这一思想，在自动机理论中也有深远的理论意义：例如，它引出了幂集构造的一种参数化，由此可为 PSPACE 完全问题提供固定参数可处理算法，例如判定两个非确定性有限自动机（NFA）是否等价 [7]。

确定性有限自动机（DFA）的情况尤其值得关注，因为在这种情况下，DFA 的“后缀数组”概念可以通过字符串得到一种简单的解释。假设 DFA $\mathcal{A}$ 的字母表 Σ 上存在一个固定的全序 $\preceq$，并将 $\preceq$ 按字典序扩展到 Σ 上所有无限字符串的集合。假设每个状态至少有一条入边。对于一个状态 u，考虑从 u 开始并以反向方式沿着边读取的所有无限字符串构成的集合 I_u，并令 min_u 和 max_u 分别表示 I_u 中按字典序最小和最大的字符串（见图1）。考虑状态集合 Q 上的偏序 $\preceq_{\mathcal{A}}$，使得对于任意 u，$v \in Q$，当 $u \neq v$ 时，满足 $u \prec_{\mathcal{A}} v$ 当且仅当 $\mathsf{max}_u \preceq \mathsf{min}_v$。于是，偏序 $\preceq_{\mathcal{A}}$ 在所有状态的集合上诱导出一个（部分）排列，其作用与字符串 [8, 21], 的后缀数组所诱导的文本位置排列相同，因此， $\preceq_{\mathcal{A}}$ 本质上就是该 DFA 的“后缀数组”。如果我们能够计算偏序 $\preceq_{\mathcal{A}}$ （以及将 Q 划分为若干集合的一种最小规模划分，使得 $\preceq_{\mathcal{A}}$ 在每个集合上的限制都是一个全序），那么我们就可以通过将 Burrows-Wheeler 变换和 FM 索引从字符串扩展到图的技术，高效且紧凑地解决 DFA 上的模式匹配查询。因此，我们现在需要解决如何高效构建 DFA 的“后缀数组”这一问题，也就是计算 $\preceq_{\mathcal{A}}$ 的问题。

关于该主题的第一篇论文 [8] 提出了一种算法，可在 $O(m^2+n^{5/2})$ 时间内构建 $\preceq_{\mathcal{A}}$，其中 n 是状态数， m 是边数。在一篇近期论文 [21], 中，Kim 等人介绍了两种算法，其运行时间分别为 $O(mn)$ 和 $O(n^2 \log n)$。关键观察是，如果我们构建所有状态集合的 最小/最大划分，就可以通过归约到区间划分问题在线性时间内确定偏序 $\preceq_{\mathcal{A}}$。确定所有状态集合的最小/最大划分，意味着为每个状态 $u \in Q$选取字符串 min_u和 max_u，并对这些 $2|Q|$ 字符串排序（请注意，其中一些字符串可能相同，示例见图1）。因此，我们只剩下如何高效确定最小/最大划分这一问题。

$O(n^2 \log n)$ 算法通过将 Manber 和 Myers 针对字符串的 $O(n \log n)$ 算法推广到确定性有限自动机，构建最小/最大划分。然而，既然可以在 $O(n)$ 时间内构建字符串的后缀数组，自然会让人想知道是否能在 $O(n^2)$ 时间内确定最小/最大划分。



i	$\min_i$	$\max_i$
1	##### ...	##### ...
2	a####	a#### ...
3	b####	b#### .....
4	c####	c#### .....
5	ca### ...	$ccccc$ ...
6	ab### ...	ac### ...
7	cc### ...	$ccccc$ ...

图1 一个 DFA A，其中标出了到达每个状态的最小和最大字符串（我们假设 # $\prec a \prec b \prec c$）。最小/最大划分为 {(1, min), (1, max)} $<$ {(2, min), (2, max)} $<$ {(6,min)} $<$ {(6,max)} $<$ {(3, min), (3, max)} $<$ {(4, min), (4, max)} $<$ {(5, min)} $<$ {(7, min)} $<$ {(5,max), (7,max)}，这意味着 $\min_1 = \max_1 \prec \min_2 = \max_2 \prec \min_6 \prec \max_6 \prec \min_3 = \max_3 \prec \min_4 = \max_4 \prec \min_5 \prec \min_7 \prec \max_5 = \max_7$。

在这篇论文中，我们表明，确实可以通过采用一种递归方法，在 $O(n^2)$ 时间内构建最小/最大划分。这种方法受到我们先前提到的一种线性时间算法的启发，即 Ko 和 Aluru 的算法 [22]。因此，我们的算法在渐近意义上比所有先前的算法都更快。

一个长期悬而未决的问题是，是否可以定义图的后缀树。近期的一些研究 [5] 表明，通过仔细研究图的字典序结构，有可能定义能够模拟 后缀树行为的数据结构（ [5] 中的结果仅适用于 Wheeler 图，但我们相信它们可以扩展到任意图）。更具体地说，我们有理由相信，可以将字符串的压缩后缀树 概念 [28] 推广到图。压缩后缀树是后缀树的一种压缩表示，由若干组件构成，其中包括后缀数组。我们已经看到， $\preceq_{\mathcal{A}}$ 将后缀数组推广到了图结构，而 [5]表明其余组件也可能推广到图。字符串后缀树的补充结构大量利用了字符串的字典序结构和组合结构。由于本文提出的算法深度依赖字典序结构的丰富性（从字符串场景转向图场景后，这一问题更具挑战性，也更令人惊讶），我们相信，我们的结果也为将后缀树的功能扩展到图提供了坚实的概念性贡献。

值得指出的是，在我们将这篇论文提交至 arXiv 几天后，一篇新的 arXiv 预印本展示了如何在 $O(m \log n)$ 时间内确定最小/最大划分，其中 n 是状态数， m 是边数（这篇新的 arXiv 预印本也已获准发表 [2]）。如果 DFA 的底层图是稀疏图，那么 [2] 中的算法优于我们的 $O(n^2)$ 算法。由于 $O(m \log n)$ 算法采用了不同的技术（它是通过改造 Paige 和 Tarjan 的划分细化算法 [26]得到的），我们仍面临一个引人深思的开放问题：能否通过结合我们算法与 [2], 中算法背后的思想，在 $O(m)$ 时间内构造最小/最大划分?

由于篇幅限制，所有证明均可在附录中找到。

2 预备知识

2.1 与前人工作的关系

在此前关于这一主题的研究背景下 [1, 8, 21], , 我们要解决的问题定义如下。考虑一个确定性有限自动机（DFA），满足：（i）进入同一状态的所有边具有相同的标签；（ii）每个状态都可从初始状态到达；（iii）每个状态都可到达终态，也就是说，它本身是终态，或能够到达某个终态；（iv）初始状态没有入边。然后，确定状态集的最小/最大划分（最小/最大划分的形式化定义见第 2.2 节，示例见图1）。

- 假设 (ii) 和 (iii) 是自动机理论中的标准假设，因为可以移除所有不满足这些假设的状态，而不改变所接受的语言。

- 在这种设定下，所有非初始状态都有一条入边，但初始状态没有入边。这意味着，对于某个状态 u，可能有 $I_u = \emptyset$（请记住，I_u 是从 u 出发并以反向方式沿边行进时能够读出的所有无限字符串的集合，见引言），因此 Kim 等人 [21] 需要进行繁琐的分类讨论，并将有限字符串也考虑在内，才能定义最小/最大划分（尤其是，到达初始状态的最小字符串和最大字符串都等于空字符串）。不过，我们可以通过 [5] 中使用的同一技巧轻松避免这一复杂情况：给初始状态添加一条自环，其标签是一个特殊字符 #，它 小于 字母表中的任何字符。直观来说，# 的作用与字符串 Burrows-Wheeler 变换中的终止字符相同；而且由于 # 是最小的字符，添加这条自环不会影响最小/最大划分（详见 [21] ）。

- 请注意，初始状态和所有终态的集合在 min/max 划分的定义中不起任何作用；这说明更一般地说，考虑确定性图而不是确定性有限自动机会更方便（否则，我们在算法中递归构建图时，就需要人为添加一个初始状态和一组终态）。等价地，我们可以假设使用转移函数不一定是全函数的半自动机。这也为我们将在第 2.2 节中作出的假设提供了依据。

- 一些近期论文 [6, 7] 表明，假设 (i) 和 (iv) 可以去除。偏序 $\preceq_{\mathcal{A}}$ 以类似的方式定义，我们提到的所有用于构建 $\preceq_{\mathcal{A}}$ 的算法仍然适用。事实上，如果一个状态 u 通过带有不同标签的边到达，那么在计算 $\min_u$ 时，我们只需考虑所有带有最小标签的边；在计算 $\max_u$时，只需考虑所有带有最大标签的边。此外，我们再次假设初始状态有一条标记为 # 的自环。唯一的区别是，假设 (i) 意味着 $m \leq n^2$（n 为状态数，m 为边数），因为每个状态至多有 n 条入边，但去除假设 (i) 后，这一点不再成立。因此，我们算法的运行时间不再是$O(n^2)$，而是 $O(m + n^2)$（而 [21] 中的 $O(n^2 \log n)$算法的运行时间变为$O(m + n^2 \log n)$），因为我们仍然需要处理 DFA 中的所有边。

总而言之，所有用于计算 $\preceq_{\mathcal{A}}$ 的算法都适用于任意确定性有限自动机。

2.2 符号约定与首批定义

设 Σ 为有限字母表。我们考虑有限的边标记图 $G = (V, E)$，其中 V 是所有节点的集合，$E \subseteq V \times V \times \Sigma$ 是所有边的集合。通过取 Σ 的一个子集，我们假设所有 $c \in \Sigma$ 都标记图中的某条边。我们假设每个节点至少有一条入边，并且进入同一节点 u 的所有边都具有相同的标签 $\lambda(u)$（输入一致性）。这意味着边 $(u, v, a) \in E$ 可以简单地记为 (u, v)，因为它必定是 $a = \lambda(v)$。特别地，它必定是 $|E| \leq |V|^2$（因此，$O(|E|)$ 算法也是 $O(|V|^2)$ 算法）。如果我们不知道 $\lambda(u)$，可以通过扫描所有边轻松计算出来。此外，我们始终假设 G 是确定性的，也就是说，对于每个 $u \in V$ 和每个 $a \in \Sigma$，至多存在一个 $v \in V$，使得 $(u, v) \in E$ 且 $\lambda(v) = a$。

令 Σ^* 为 Σ 上所有有限字符串的集合，并令 Σ^ω 为 Σ 上所有（可数）右无限字符串的集合。若 $\alpha \in \Sigma^* \cup \Sigma^\omega$ 且 $i \geq 1$，则我们用 $\alpha[i] \in \Sigma$ 表示 α 的 i^{th} 个字符（即 $\alpha = \alpha[1]\alpha[2]\alpha[3]\ldots$）。若 $1 \leq i \leq j$，我们定义 $\alpha[i, j] = \alpha[i]\alpha[i+1]\ldots\alpha[j-1]\alpha[j]$；若 $j < i$，则 $\alpha[i, j]$ 是空字符串 ϵ。若 $\alpha \in \Sigma^*$，则 $|\alpha|$ 是 α 的长度；对于每个 $0 \leq j \leq |\alpha|$，字符串 $\alpha[1, j]$ 是 α 的一个前缀；若 $0 \leq j < |\alpha|$，它就是一个真前缀；类似地，也可以定义 α 的后缀和真后缀。从 $u \in V$ 开始、在 $u' \in V$ 结束的 $\alpha \in \Sigma^*$ 的一次出现，是 V 中节点 u_1、u_2、……、$u_{|\alpha|+1}$ 组成的序列，满足：(i) $u_1 = u$；(ii) $u_{|\alpha|+1} = u'$；(iii) 对每个 $1 \leq i \leq |\alpha|$，$(u_{i+1}, u_i) \in E$；(iv) 对每个 $1 \leq i \leq |\alpha|$，$\lambda(u_i) = \alpha[i]$。从 $u \in V$ 开始的 $\alpha \in \Sigma^\omega$ 的一次出现，是 V 中由节点 $(u_i)_{i \geq 1}$ 组成的序列，满足：(i) $u_1 = u$；(ii) 对每个 $i \geq 1$，$(u_{i+1}, u_i) \in E$；(iii) 对每个$i \geq 1$，$\lambda(u_i) = \alpha[i]$。直观而言，如果可以从 u 出发，在图上沿着边以反向方式读取 α，那么字符串 $\alpha \in \Sigma^* \cup \Sigma^\omega$ 就有一次从 $u \in V$ 开始的出现。

在论文中，Σ^ω 中字符串的出现将发挥关键作用，而 Σ^* 中字符串的出现将作为一种技术工具。对于每个 $u \in V$，我们用 I_u 表示 Σ^ω 中所有允许在 u 处开始出现的字符串的集合。由于每个节点至少有一条入边，因此 $I_u \neq \emptyset$。

如果一个集合 V 上的自反、反对称且传递的关系是 V，则称其为 **全序** $\leq$。如果$u, v \in V$，我们记作 $u < v$，如果 $u \leq v$ 且 $u \neq v$。

设 $\preceq$ 是 Σ 上的一个固定全序。我们将 $\preceq$ 按字典序扩展到 $\Sigma^* \cup \Sigma^\omega$ **按字典序**。容易证明，在每个 I_u 中都存在按字典序最小的字符串 $\min_u$ 和按字典序最大的字符串 $\max_u$（例如，这可由 [21, 观察 8] 推出）。

我们将经常使用以下直接观察结果。设 $u \in V$，并设 $(u_i)_{i \geq 1}$ 是 $\min_u$ 的一次出现。固定 $i \geq 1$。则，$(u_j)_{j \geq i}$ 是 $\min_{u_i}$ 的一次出现，并且 $\min_u = \min_u[1, i-1]\min_{u_i}$。

设 $V' \subseteq V$。令 A 为 V' 的唯一划分，并令 $\leq$ 为 A 上的唯一全序，使得对于任意 I、$J \in$ A 以及任意 $u \in I$ 和 $v \in J$，(i) 若 $I = J$，则 $\min_u = \min_v$；(ii) 若 $I < J$，则 $\min_u \prec \min_v$。于是，我们称 (A $\leq$)，或简写为 A，是 V' 的最小划分。V' 的最大划分类似定义。现在，考虑集合$V' \times \{\min, \max\}$，并对每个 $u \in V'$ 定义 $\rho((u, \min)) = \min_u$ 以及 $\rho((u, \max)) = \max_u$。令 B 为 $V' \times \{\min, \max\}$ 的唯一划分，并令 $\leq$ 为 B 上的唯一全序，使得对于任意 I、$J \in$ B 以及任意 $x \in I$ 和 $y \in J$，(i) 若 $I = J$，则 $\rho(x) = \rho(y)$；(ii) 若 $I < J$，则 $\rho(x) \prec \rho(y)$。于是，我们称 (B $\leq$)，或简写为 B，是 V' 的最小/最大划分。

本文的主要结果是证明，可以在 $O(|V|^2)$ 时间内确定 V 的最小/最大划分。

2.3 我们的方法

设 $G = (V, E)$ 是一个图。我们首先将展示如何在 $O(n^2)$ 时间内构造 V 的最小划分，其中 $n = |V|$（第 4 节）；随后，我们将展示如何调整该算法，使其能够在 $O(n^2)$ 时间内构造最小/最大划分（第5节）。

为了构建 V 的最小划分，我们将首先把所有极小值分为三种类别（第3节），以便将 V 拆分为三个两两不相交的集合 V_1, V_2, V_3。然后，我们将证明这一过程可以在 $O(n^2)$ 时间内完成：

- 我们可以计算 V_1、V_2、V_3（第4.1节）；
- 我们可以定义一个具有 $|V_3|$ 个节点的图 $G = (V$、$E)$（第 4.2 节）；
- 假设我们已经确定了 V 的最小划分，那么我们就可以确定 V 的最小划分（第4.3节）。

类似地，我们可以在 $O(n^2)$ 时间内，将确定V的最小划分问题归约为确定一个具有$|V_1|$（而非 $|V_3|$）个节点的图中所有节点集合的最小划分问题（第4.4节）。因此，由于 $\min\{|V_1|, |V_3|\} \leq |V|/2 = n/2$，我们得到一个递归算法，其运行时间由以下递推关系给出：

$$T(n) = T(n/2) + O(n^2)$$

由此我们得出，我们的算法的运行时间为 $O(n^2)$。

3 字符串分类

在 [22], Ko 和 Aluru 将一个字符串的后缀分为两组。这里我们采用一种完全基于字符串学的方法，而不是从一开始就固定一个字符串或一张图。我们将 Σ^ω 中的字符串分为三组，分别称为第1组、第2组和第3组（推论 3 为这一选择提供了直观解释）。

▶ **定义1**。令 $\alpha \in \Sigma^\omega$。令 $a \in \Sigma$ 和 $\alpha' \in \Sigma^\omega$，使得 $\alpha = a\alpha'$。然后，我们定义τ（α）如下：

1. τ（α）$= 1$ 若 $\alpha' \prec \alpha$。
2. τ（α）$= 2$ 若 $\alpha' = \alpha$。
3. τ（α）$= 3$ 若 $\alpha \prec \alpha'$。

我们将不断使用以下刻画。

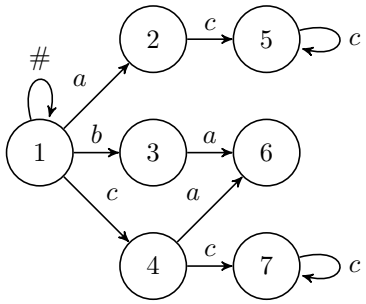
▶ **引理 2.** 令 $\alpha \in \Sigma^\omega$。令 $a \in \Sigma$ 和 $\alpha' \in \Sigma^\omega$ 满足 $\alpha = a\alpha'$。则：

1. $\tau(\alpha) = 2$ 当且仅当 $\alpha' = a^\omega$，当且仅当 $\alpha = a^\omega$。*2.* $\tau(\alpha) \neq 2$ 当且仅当 $\alpha' \neq a^\omega$，当且仅当 $\alpha \neq a^\omega$。假设 $\tau(\alpha) \neq 2$。那么，存在唯一的 $c \in \Sigma \setminus \{a\}$、$\alpha'' \in \Sigma^\omega$ 和 $i \geq 0$，使得$\alpha' = a^i c\alpha''$（因此 $\alpha = a^{i+1}c\alpha''$）。此外：*1.* $\tau(\alpha) = 1$ 当且仅当 $c \prec a$，当且仅当 $\alpha' \prec a^\omega$，当且仅当 $\alpha \prec a^\omega$。*2.* $\tau(\alpha) = 3$ 当且仅当 $a \prec c$，当且仅当 $a^\omega \prec \alpha'$，当且仅当 $a^\omega \prec \alpha$，

以下推论将成为我们递归方法中的关键组成部分。

▶ **推论 3**。设 $\alpha, \beta \in \Sigma^\omega$。设 $a, b \in \Sigma$ 和 $\alpha', \beta' \in \Sigma^\omega$，使得 $\alpha = a\alpha'$ 且 $\beta = b\beta'$。那么：

1. 如果 $a = b$ 和 $\tau(\alpha) = \tau(\beta) = 2$，那么 $\alpha = \beta$。*2.* 如果 $a = b$和 $\tau(\alpha) < \tau(\beta)$，那么 $\alpha \prec \beta$。等价地，如果 $a = b$和 $\alpha \preceq \beta$，那么 $\tau(\alpha) \leq \tau(\beta)$】 【。



i	$\min_i$	$\tau(i)$ $(=\tau(\min_i))$
1	##### ...	2
2	a#### ...	1
3	b#### ...	1
4	c#### ...	1
5	ca### ...	1
6	ab### ...	3
7	cc### ...	1

图 2 图1中的图，其中包含 $\min_i$'s 和 $\tau(i)$'s 的值。

4 计算最小划分

令 $G=(V,E)$ 为一个图。我们将证明，可以在 $O(|V|^2)$ 时间内计算出 V 的最小划分。在本节中，对于每个 $u\in V$，我们定义 $\tau(u)=\tau(\min_u)$（见图 2）。

令 $u\in V$，并设 $(u_i)_{i\geq 1}$ 为一个从 u 开始的 $\min_u$ 的出现。显然可以看出：(i) 如果 $\tau(u)=1$，那么 $\lambda(u_2)\preceq\lambda(u_1)$；(ii) 如果 $\tau(u)=2$，那么对于每个$k\geq 1$，都有 $\lambda(u_k)=\lambda(u_1)$；(iii) 如果 $\tau(u)=3$，那么 $\lambda(u_1)\preceq\lambda(u_2)$。

首先，我们来证明：不失一般性，我们可以从G 中删除一些边，而不影响最小/最大划分。这一预处理将在引理 23 中发挥作用。

▶ **定义 4.** 设 $G=(V,E)$ 为一个图。若 G 中不包含满足以下条件的边$(u,v)\in E$：$\tau(v)=1$ 且 $\lambda(v)\prec\lambda(u)$，则称 G 为修剪的 。

为使证明更易于阅读，我们不会直接从 $G=(V,E)$ 中删除一些边，而是先构造 G 的一个副本，其中每个节点 u 都映射到一个节点 u^*，然后再修剪该图。这样，当我们写 $\min_u$ 和 $\min_{u^*}$ 时，就能始终清楚我们指的是原图还是修剪图。在第 4.2 节定义用于递归步骤的图 $G=(V,E)$ 时，我们也会采用同样的约定。

▶ **引理 5.** 设 $G=(V,E)$ 是一个图。那么，我们可以在 $O(|E|)$ 时间内构造一个修剪图$G^*=(V^*,E^*)$，满足 $V^*=\{u^*\mid u\in V\}$，使得对每个 $u\in V$都有 最小值$_{u^*}$ = 最小值$_u$。特别地，对每个 $u\in V$，都有 $\tau(u^*)=\tau(u)$ 。

4.1 最小值分类

我们首先说明如何计算所有满足 $\tau(u)=1$ 的 $u\in V$。

▶ 引理 6. 设 $G=(V,E)$ 是一个图，且设 $u,v\in V$。1. 若 $(u,v)\in E$ 且 $\lambda(u)\prec\lambda(v)$，则 $\tau(v)=1$。2. 若 $(u,v)\in E$、$\lambda(u)=\lambda(v)$ 且 $\tau(u)=1$，则 $\tau(v)=1$。

▶ 推论 7。设 $G=(V,E)$ 是一个图，且设 $u\in V$。那么，$\tau(u)=1$ 当且仅当存在 $k\geq 2$ 和 $z_1,\dots,z_k\in V$，使得 (i) 对每个 $1\leq i\leq k-1$，$(z_i,z_{i+1})\in E$；(ii)$z_k=u$；(iii) $\lambda(z_1)\prec\lambda(z_2)$；以及 (iv) $\lambda(z_2)=\lambda(z_3)=\cdots=\lambda(z_k)$。

推论 7 给出了一个算法，用于判定 $u\in V$ 是否满足 $\tau(u)=1$。

▶ **推论 8.** 设 $G=(V,E)$ 为一个图。我们可以在 $O(|E|)$时间内确定所有满足 $\tau(u)=1$的 $u\in V$。

现在，我们来说明如何确定所有满足 $\tau(u) = 2$的 $u \in V$。我们可以假设，已经确定了所有满足 $\tau(u) = 1$的 $u \in V$。

▶ 引理 *9*。设 $G = (V, E)$ 是一个图，并设 $u \in V$使得 $\tau(u) \neq 1$。那么，我们有$\tau(u) = 2$ 当且仅当存在 $k \geq 2$和 $z_1, \ldots, z_k \in V$，使得 *(i)* $(z_{i+1}, z_i) \in E$对每个 $1 \leq i \leq k-1$，*(ii)* $z_1 = u$，*(iii)* $z_k = z_j$ 对某个 $1 \leq j \leq k-1$，以及 *(iv)*$\lambda(z_1) = \lambda(z_2) = \cdots = \lambda(z_k)$。

特别地，这样的 $z_1, \ldots, z_k \in V$必须满足 $\tau(z_i) = 2$ 对每个 $1 \leq i \leq k$。

▶ **推论 10**。设 $G = (V, E)$ 为一个图。我们可以在时间 $O(|E|)$内确定所有 $u \in V$，使得 $\tau(u) = 2$。

由推论 8 和推论 10，我们立即得到以下结果。

▶ **推论 11**。设 $G = (V, E)$ 为一张图。那么，我们可以在 $O(|E|)$ 时间内，为每个 $u \in V$ 计算出 $\tau(u)$。

4.2 递归步骤

我们概述一下递归步骤中构建较小图的一般思路。对于每个 $u \in V$，满足 $\tau(u) = 3$，我们以反向方式沿边前进，旨在确定 $\min_u$ 的前缀。因此，我们丢弃无法让 $\min_u$ 的任何出现通过的边；根据推论 3，我们可以将注意力限制在 v 这些节点上，使得 $\tau(v)$ 最小。我们按此方式继续，直到遇到满足 $\tau(v') = 3$ 的节点 v'。

让我们将直觉形式化。首先，我们将介绍字符串 $\min_u$ 的出现必须满足的一些性质。

▶ **引理 12.** 设 $G = (V, E)$ 为一个图。设 $u, v \in V$ 满足 $\min_u = \min_v$。设$(u_i)_{i \geq 1}$ 是 $\min_u$ 的一次出现，且$(v_i)_{i \geq 1}$ 是 $\min_v$ 的一次出现。那么：
1. $\lambda(u_i) = \lambda(v_i)$ 对于每个 $i \geq 1$。**2.** $\min_{u_i} = \min_{v_i}$ 对于每个 $i \geq 1$。**3.** $\tau(u_i) = \tau(v_i)$对于每个 $i \geq 1$。
特别地，如果 $u = v$ 和$(u_i)_{i \geq 1}$ 以及$(v_i)_{i \geq 1}$ 是 $\min_u$ 的两个不同出现，则前述结果成立。

▶ **引理13.** 设 $G = (V, E)$ 为一个图。设 $u \in V$，并设$(u_i)_{i \geq 1}$ 是从 u 开始的 $\min_u$ 的一次出现。设 $k \geq 1$ 满足 $\tau(u_1) = \tau(u_2) = \cdots = \tau(u_{k-1}) = \tau(u_k) \neq 2$。那么，$u_1$、…、$u_k$ 两两不同。特别地，$k \leq |V|$。

前述结果使我们能够给出以下定义。

▶ **定义14**。设 $G = (V, E)$ 为一个图。设 $u \in V$满足 $\tau(u) = 3$。令 ℓ_u 为满足 $\tau(u_k) \geq 2$ 的最小整数 $k \geq 2$，其中 $(u_i)_{i \geq 1}$ 是从 u 开始的 $\min_u$ 的一次出现。

注意，ℓ_u 是良定义的，因为 (i) 根据引理13，对于每个 $k \geq 2$，$\tau(u_k) = 1$ 都不成立（事实上，若 $\tau(u_2) = 1$，那么 $(u_i)_{i \geq 2}$ 就是一个从 u_2 开始的 $\min_{u_2}$ 的出现；根据引理 13，存在 $2 \leq k \leq |V| + 2$ 使得 $\tau(u_k) \neq 1$），且 (ii) 根据引理 12，ℓ_u 不依赖于 $(u_i)_{i \geq 1}$ 的选取。特别地，必有 $\ell_u \leq |V| + 1$，因为u_1，u_2，…，$u_{\ell_u - 1}$两两不同（u_1与 u_2，…，$u_{\ell_u - 1}$不同，因为 $\tau(u_1) = 3$且根据 ℓ_u 的极小性，$\tau(u_2) = \tau(u_3) = \ldots \tau(u_{\ell_u - 1}) = 1$）。

▶ **引理15**。设 $G = (V, E)$ 为一个图。设 $u \in V$满足 $\tau(u) = 3$。那么，对于每个 $2 \leq i \leq \ell_u - 1$，$\min_u[i+1] \preceq \min_u[i]$。特别地，如果 $2 \leq i \leq j \leq \ell_u$，那么 $\min_u[j] \preceq \min_u[i]$。

如果 $R \subseteq Q$ 是一个非空节点集，且对每个 $u, v \in R$，都有 $\lambda(u) = \lambda(v)$，则定义 $\lambda(R) = \lambda(u) = \lambda$。如果 $R \subseteq Q$ 是一个非空节点集，且对每个 $u, v \in R$，都有 $\tau(u) = \tau(v)$，则定义 $\tau(R) = \tau(u) = \tau(v)$。

设 $R \subseteq Q$ 是一个非空状态集合。令 $\mathtt{F}(R) = \arg\min_{u \in R'} \tau(u)$，其中 $R' = \arg\min_{v \in R} \lambda(v)$。注意，$\mathtt{F}(R)$ 非空，且 $\lambda(\mathtt{F}(R))$ 和 $\tau(\mathtt{F}(R))$ 都有定义。换言之，$\mathtt{F}(R)$ 的求法是：先考察所有满足 $\lambda(v)$ 尽可能小的节点组成的子集 $R' \subseteq \mathtt{F}(R)$，然后再考察 R' 中所有满足 $\tau(v)$ 尽可能小的节点组成的子集 v。这与我们寻找 $\min_u$ 的前缀时应采用的方法相符。

定义：

$$G_i(u) = \begin{cases} \{u\} & \text{if } i = 1; \\ \mathtt{F}(\{v' \in Q \mid (\exists v \in G_{i-1}(u))((v', v) \in E)\}) \setminus \bigcup_{j=2}^{i-1} G_j(u) & \text{if } 1 < i \leq \ell_u. \end{cases}$$

注意，我们还要求 $G_i(u)$ 中的节点此前未曾出现过。直观而言，这不会影响我们对 $\min_u$ 的前缀的搜索，因为如果同一个节点出现两次，那么就会形成一个环，其中所有边的标签都相同（因为根据引理15，标签只能递减）；而且，对于每个 $2 \leq i \leq \ell_u - 1$，都有 $\tau(G_i(u)) = 1$，因此，最小值的任何一次出现都不可能经过该环，因为从这次出现中删去该环后，根据引理 2，我们会得到一个更小的字符串。

以下技术引理对于证明我们的直觉正确至关重要。

▶ **引理16**。设 $G = (V, E)$ 为一个图。设 $u \in V$ 满足 $\tau(u) = 3$。*1.* 对于每个 $1 \leq i \leq \ell_u$，$G_i(u)$ 都是定义良好且非空的。*2.* 设 $(u_i)_{i \geq 1}$ 是从 u 开始的 $\min_u$ 的一次出现。那么，对于每个 $1 \leq i \leq \ell_u$，$u_i \in G_i(u)$。特别地，对于每个 $1 \leq i \leq \ell_u$，$\tau(u_i) = \tau(G_i(u))$，且 $\min_u[i] = \lambda(u_i) = \lambda(G_i(u))$。*3.* 对于每个 $1 \leq i \leq \ell_u$ 以及每个 $v \in G_i(u)$，都存在一次从 u 开始并在 v 结束的 $\min_u[1, i-1]$ 的出现。

令 $u \in V$ 满足 $\tau(u) = 3$。我们定义：

- $\gamma_u =$ 最小值$_u[1, \ell_u]$；$\mathtt{t}_u = \tau(G_{\ell_u}(u)) \in \{2、$
- $3\}$

现在，为了定义递归步骤所需的较小图，我们还需要一个新的字母表 $(\Sigma', \preceq')$，其定义必须与极小值之间的相互顺序保持一致。下一个引理给出了我们所需的全部信息。

▶ **引理 17**。设 $G = (V, E)$ 为一个图。设 $u, v \in V$ 满足 $\tau(u) = \tau(v) = 3$。假设以下陈述之一成立：

1. γ_u 不是 γ_v 的前缀，并且 $\gamma_u \prec \gamma_v$。
2. $\gamma_u = \gamma_v$，$\mathtt{t}_u = 2$ 并且 $\mathtt{t}_v = 3$。3. γ_v 是 γ_u的真前缀。

那么，$\min_u \prec \min_v$。

等价地，若 $\min_u \preceq \min_v$，则以下情形之一成立：*(i)* γ_u 不是 γ_v和 $\gamma_u \prec \gamma_v$的前缀；*(ii)* $\gamma_u = \gamma_v$ 和 $\mathtt{t}_u \leq \mathtt{t}_v$；*(iii)* γ_v 是 γ_u的真前缀。

现在，令 $\Sigma' = \{(\gamma_u, \mathtt{t}_u) \mid u \in V, \tau(u) = 3\}$，并令 $\preceq'$ 为 Σ' 上的全序，使得对于任意不同的 $(\alpha, x), (\beta, y) \in \Sigma'$，当且仅当以下条件之一成立时，有 $(\alpha, x) \prec' (\beta, y)$：

1. α 不是 β 的前缀，且 $\alpha \prec \beta$。
2. $\alpha = \beta$，$x = 2$ 且 $y = 3$。

3. β 是真前缀 α。

可以直接验证，$\preceq'$ 是一个全序：事实上，$\preceq'$ 是通过以下方式得到的：(i) 首先使用 Σ^* 上的（全）字典序的一种变体比较各个 γ_u，在这种变体中，一个字符串小于它的任何真前缀；(ii) 如果各个 γ_u相等，则通过比较 $\mathtt{t}_u$来确定顺序，而这些 $\mathtt{t}_u$是{2,3}中的元素。

从 $G=(V,E)$ 出发，我们按如下方式定义一个新的图 $G=(V,E)$：$V=\{\bar{u}\mid u\in V,\tau(u)=3\}$。新的全序字母表为 $(\Sigma',\preceq')$。对于每个 $\bar{u}\in V$，我们定义 $\lambda(\bar{u})=(\gamma_u,\mathtt{t}_u)$。$E=\{(\bar{v},\bar{v})\mid \mathtt{t}_v=2\}\cup\{(\bar{u},\bar{v})\mid \mathtt{t}_v=3,u\in G_{\ell_v}(v)\}$。

请注意，对于每个满足 $\mathtt{t}_v=3$ 的 $\bar{v}\in V$，以及对于每个 $u\in G_{\ell_v}(v)$，都有 $\tau(u)=\tau(G_{\ell_v}(v))=\mathtt{t}_v=3$，因此 $\bar{u}\in V$，且 $(\bar{u},\bar{v})\in E$。此外，$G=(V,E)$ 满足本文中使用的关于图的所有假设：(i) 进入同一节点的所有边具有相同的标签（根据定义）；(ii) 每个节点至少有一条入边（因为若 $\bar{v}\in V$，则根据引理16，$G_{\ell_v}(v)\neq\emptyset$）；(iii) G 是确定性的（因为若 $(\bar{u},\bar{v})$、$(\bar{u},\bar{v}')\in E$ 且 $\lambda(\bar{v})=\lambda(\bar{v}')$，则 $\gamma_v=\gamma_{v'}$ 且 $\mathtt{t}_v=\mathtt{t}_{v'}$；因此，根据 E 的定义，若 $\mathtt{t}_v=\mathtt{t}_{v'}=2$，我们立即得到 $\bar{v}=\bar{u}=\bar{v}'$；若 $\mathtt{t}_v=\mathtt{t}'_v=3$，则得到 $u\in G_{\ell_v}(v)\cap G_{\ell_{v'}}(v')$。由于根据引理16，存在两个从 v 和 v' 开始、都在 u结束的 $\min_v[1,\ell_v-1]=\gamma_v[1,\ell_v-1]=\gamma_{v'}[1,\ell_{v'}-1]=\min_{v'}[1,\ell_{v'}-1]$ 出现，G 的确定性意味着$v=v'$，因此 $\bar{v}=\bar{v}'$）。

注意，若 $\bar{v}\in V$ 满足 $\mathtt{t}_v=2$，则 $I_{\bar{v}}$ 恰好包含一个字符串，即 $\lambda(\bar{v})^\omega$；特别地，$\min_{\bar{v}}=\max_{\bar{v}}=\lambda(\bar{v})^\omega$。

实现 $G=(V,E)$ 和 $G=(V,E)$ 时，我们使用整数字母表 $\Sigma=\{0,1,\ldots,|\Sigma|-1\}$ 和 $\Sigma'=\{0,1,\ldots,|\Sigma'|-1\}$；特别地，我们不会用 $(\gamma_u,\mathtt{t}_u)$ 对来存储 Σ'，而会按照 $\preceq'$ 在 Σ' 上的全序，将 Σ' 重新映射到整数字母表中，使 $\min_{\bar{u}}$ 的相对顺序不受影响。

下面证明递归步骤中可以使用 $G=(V,E)$。我们先给出一些预备结果。

▶**引理18**。设 $G=(V,E)$是一个图。设 $u,v\in V$满足 $\tau(u)=\tau(v)=3$，$\gamma_u=\gamma_v$和 $\mathtt{t}_u=\mathtt{t}_v=2$。那么，极小值$_u$ = 极小值$_v$。

▶**引理19**。设 $G=(V,E)$ 是一个图。设 $u\in V$，并设 $(u_i)_{i\geq1}$ 是极小值$_u$ 的一个出现，起始于 u。那么，以下情况恰有一种成立：
1. 存在 $i_0\geq1$，使得对于每个 $1\leq i<i_0$，$\tau(u_i)\neq2$，并且对于每个$i\geq i_0$，$\tau(u_i)=2$。 *2.* 对于每个 $i\geq1$，$\tau(u_i)\neq2$，并且对于无穷多个i，$\tau(u_i)=1$和 $\tau(u_i)=3$都成立。

关键的是，下一个引理确立了 $G=(V,E)$中节点的极小值与 $G=(V,E)$中节点的极小值之间的对应关系。

▶引理*20*。设 $G=(V,E)$ 是一个图。设 $u\in V$，使得 $\tau(u)=3$。设 $(u_i)_{i\geq1}$是从 u 开始的 *min*$_u$ 的一次出现。设 $(u'_i)_{i\geq1}$ 是 V 中的节点无限序列，按如下方式构造。考虑 $L=\{k\geq1\mid\tau(u_k)=3\}$，并且对于每个 $i\geq1$，若存在，则令 $j_i\geq1$ 为 L 中最小的 i^{th}元素。对于每个使得 j_i 已定义的 $i\geq1$，令 $u'_i=u_{j_i}$；若$i\geq1$ 使得 j_i 未定义（因此 L 是有限集），则令 $u'_i=u'_{|L|}$。那么，$(\bar{u}'_i)_{i\geq1}$ 是 $G=(V,E)$中从 $\bar{u}$ 开始的 *min*$_{\bar{u}}$ 的一次出现。

以下定理表明，我们对 $G=(V,E)$ 的归约是正确的。

▶ 定理 21。设 $G=(V,E)$ 是一个图。设 $u,v\in V$ 满足 $\tau(u)=\tau(v)=3$。1. 如果 $min_u=min_v$，那么 $min_{\bar{u}}=min_{\bar{v}}$。2. 如果 $min_u \prec min_v$，那么 $min_{\bar{u}} \prec' min_{\bar{v}}$。

由于 $\preceq$ 是一个全序（因此 $\min_u \prec \min_v$、$\min_u=\min_v$ 和 $\min_v \prec \min_u$ 中恰有一个成立），由定理 21 可立即得到以下结果。

▶ 推论 22. 设 $G=(V,E)$ 是一个图。设 $u,v\in V$ 满足 $\tau(u)=\tau(v)=3$。1. $min_u=min_v$ 成立，当且仅当 $min_{\bar{u}}=min_{\bar{v}}$ 成立。2. $min_u \prec min_v$ 成立，当且仅当 $min_{\bar{u}} \prec' min_{\bar{v}}$ 成立。

特别地，如果我们已知 V 关于 G 的最小划分，那么也就已知 $\{u\in V \mid \tau(u)=3\}$ 关于 G 的最小划分。

最后，我们证明，将问题归约到 $\bar{G}=(\bar{V},\bar{E})$ 可以在 $O(n^2)$ 时间内完成。

▶ **引理 23.** 设 $G=(V,E)$ 是一个修剪图。那么，我们可以在 $O(|V|^2)$ 时间内构建 $\bar{G}=(\bar{V},\bar{E})$。

4.3 合并

我们希望确定 V 的最小划分 $\mathcal{A}$，并假设我们已经有了 $\{u\in V \mid \tau(u)=3\}$ 的最小划分 $\mathcal{B}$。

首先，注意到我们可以轻易地构建 $\{u\in V \mid \tau(u)=2\}$ 的最小划分 $\mathcal{B}'$。事实上，如果 $\tau(u)=2$，那么根据引理 2，$\min_u=\lambda(u)^\omega$。因此，如果 $\tau(u)=\tau(v)=2$，那么 (i) $\min_u=\min_v$ 当且仅当 $\lambda(u)=\lambda(v)$，且 (ii) $\min_u \prec \min_v$ 当且仅当 $\lambda(u)\prec\lambda(v)$，所以我们可以使用计数排序在 $O(|V|)$ 时间内构建 $\mathcal{B}'$。

对于每个 $c\in\Sigma$ 和 $t\in\{1,2,3\}$，令 $V_{c,t}=\{v\in V \mid \lambda(v)=c,\tau(v)=t\}$。考虑 $u,v\in V$：(i) 若 $\lambda(u)\prec\lambda(v)$，则 $\min_u \prec \min_v$；(ii) 若 $\lambda(u)=\lambda(v)$ 且 $\tau(u)<\tau(v)$，则根据推论 3，$\min_u \prec \min_v$。因此，为了构建 A，我们只需针对每个 $c\in\Sigma$ 和每个 $t\in\{1,2,3\}$，构建 $V_{c,t}$ 的最小划分 $A_{c,t}$。

实现每个 $\mathcal{A}_{c,t}$ 的一种可行方法是使用数组 $A_{c,t}$，存储 $V_{c,t}$ 的元素；此外，我们还使用一个特殊字符来标记 $A_{c,t}$ 中相邻元素之间的边界。

显然，对于每个 $c\in\Sigma$，可以从最小元素到最大元素逐步构建 $\mathcal{A}_{c,3}$。初始时，对于每个 $c\in\Sigma$，$A_{c,3}$ 均为空。然后，按从小到大的顺序扫描 $\mathcal{B}$ 中的元素 I，并将 I 加入 $\mathcal{A}_{c,3}$，其中对于任意 $u\in I$，$c=\lambda(u)$（c 的定义不依赖于 u 的选择）。我们只扫描一次 $\mathcal{B}$，因此这一步耗时 $O(|V|)$。类似地，对于每个 $c\in\Sigma$，我们可以利用 $\mathcal{B}'$ 构建 $\mathcal{A}_{c,2}$。

我们只需说明如何为每个 $c\in\Sigma$ 构造 $A_{c,1}$。初始时，每个 $A_{c,1}$ 都是空的，我们将从最小元素到最大元素逐步构造每个 $A_{c,1}$。在算法的这一步中，我们将逐渐标记满足 $\tau(u)=1$ 的节点 $u\in V$。在这一步开始时，没有这样的节点被标记；在这一步结束时，所有这些节点都会被标记。设 $\Sigma=\{c_1,c_2,\dots,c_\sigma\}$，其中 $c_1\prec c_2\prec\cdots\prec c_\sigma$。注意，必有 $V_{c_1,1}=\emptyset$，因为如果存在 $u\in V_{c_1,1}$，那么根据引理 2，它将是 $\min_u \prec c_1^\omega$，因此 c_1 就不会是 Σ 中最小的字符。现在考虑 $V_{c_1,2}$；我们已经完整计算出 $A_{c_1,2}$。从小到大处理 $A_{c_1,2}$ 中的每个 I，并对每个 $c_k\in\Sigma$，计算集合 J_k，其中包含所有未标记的节点 $v\in V$，使得 $\tau(v)=1$，$\lambda(v)=c_k$，且对于某个 $u\in I$，有 $(u,v)\in E$。然后，如果 $J_k\neq\emptyset$，就将 J_k 添加到 $A_{c_k,1}$，并标记 J_k 中的节点。处理完 $A_{c_1,2}$ 中的元素后，我们按此顺序依次处理 $A_{c_1,3}$、$A_{c_2,1}$、$A_{c_2,2}$、$A_{c_2,3}$、$A_{c_3,1}$，依此类推。每个 $A_{c_i,t}$ 都从其（当前）最小元素处理到其（当前）最大元素。我们从不删除或修改任何 $A_{c,t}$ 中的元素，而只会

向 $A_{c,1}$ 中添加元素。更准确地说，当我们处理 $A_{c,t}$ 中的 I 时，对于每个 $c_k \in \Sigma$，我们计算集合 J_k，其中包含所有未标记的节点 $v \in V$，使得 $\tau(v) = 1$, $\lambda(v) = c_k$，且对于某个 $u \in I$，有 $(u, v) \in E$；如果 $J_k \neq \emptyset$，那么我们就将 J_k 添加到 $A_{c_k,1}$，并标记 J_k 中的节点。

下面的引理表明，我们的方法是正确的。先来看看直观解释。最小划分 \(C\) 的一个前缀是 \(C\) 的子集 \(C'\)，并且若 \(\{v2\},\{v3\}C\)、\(\{v4\}\) 且 \(\{v5\}C\{v6\}\)，则 \(\{v7\}C\{v8\}\)。注意，\(A\) 的每个前缀都是按此顺序依次取 \(A\{v9\},2\)、\(A\{v10\},3\)、\(A\{v11\},1\)、\(A\{v12\},2\)、\(A\{v13\},3\)、\(A\{v14\},1\)、..... 的并集得到的，直到某个元素 \(A\{v15\},\{v16\}\) 为止；其中最后一个元素 \(A\{v17\},\{v18\}\) 可能只取其一个前缀。接下来，我们将证明，当我们在 \(\{v20\},\{v21\}\) 中处理 \(\{v19\}\) 时，已经构造出了最大元素为 \(\{v22\}\) 的 \(A\) 的前缀。这意味着，对于每个 \(\{v23\}\)，以及任何一个从 \(\{v27\}\) 开始的最小\(\{v26\}\)的出现 \((\{v24\})\{v25\}\)，都必须有 \(\{v28\}\) 属于 \(\{v29\}\)。

▶ **引理 24.** 设 $G = (V, E)$ 为一个图。如果我们已知 $\{u \in V \mid \tau(u) = 3\}$ 的最小划分，那么我们可以在 $O(|E|)$ 时间内构建 V 的最小划分。

4.4 互补情形

我们已经证明，可以在 $O(n^2)$ 时间内，将确定 V 的最小划分问题，归约为确定一个具有 $|\{u \in V \mid \tau(u) = 3\}|$ 个节点的图中所有节点的最小划分问题。现在，我们必须证明，同样地，可以在 $O(n^2)$ 时间内，将确定 V 的最小划分问题，归约为确定一个具有 $|\{u \in V \mid \tau(u) = 1\}|$ 个节点的图中所有节点的最小划分问题。这一次，构建较小的图 $G = (V, E)$ 将更加简单，因为在搜索前缀的极小值时，我们永远不会遇到已经遇到过的节点（因此，$G_i(u)$ 的定义更加简单，而且我们不需要对图进行裁剪来使其保持在 $O(n^2)$ 的界限之内）。另一方面，合并步骤将更加复杂，因为我们处理 $A_{c,t}$ 的顺序将从大到小排列（A_{c_σ},2、A_{c_σ},1、$A_{c_{\sigma-1}}$,3、$A_{c_{\sigma-1}}$,2、$A_{c_{\sigma-1}}$,1、$A_{c_{\sigma-2}}$,3，依此类推），因此，我们需要更新某些 $A_{c,t}$ 的部分元素，以纳入算法后续阶段可能推断出的有关极小值的信息。具体细节见附录。

5 计算最小/最大划分

设 $G = (V, E)$ 是一个图。只需考虑 $\preceq$ 的转置全序 $\preceq^*$（满足 $a \preceq^* b$ 当且仅当 $b \preceq a$ 的全序），并构建最小划分，即可构建 V 的最大划分。因此，构建最大划分的算法与构建最小划分的算法完全对称。

设 $G = (V, E)$ 是一张图。下面说明如何在 $O(|V|^2)$ 时间内构造 V 的最小/最大划分。假设我们有两张图 $G_1 = (V_1, E_1)$ 和 $G_2 = (V_2, E_2)$，它们使用相同的字母表 $(\Sigma, \preceq)$，且 $V_1 \cap V_2 = \emptyset$（允许 G_1 和 G_2 可能是空图，即没有顶点的图）。令 $V'_1 \subseteq V_1$, $V'_2 \subseteq V_2$, $W = V'_1 \cup V'_2$，并对每个 $u \in W$ 定义 $\rho(u) = \min_u$（若 $u \in V'_1$），以及 $\rho(u) = \max_u$（若 $u \in V'_2$）。令 A 为 W 的唯一划分，并令 $\leq$ 为 A 上的唯一全序，使得对任意 $I, J \in A$ 以及任意 $u \in I$ 和 $u \in J$，(i) 若 $I = J$，则 $\rho(u) = \rho(u)$；(ii) 若 $I < J$，则 $\rho(u) \prec \rho(u)$。于是，我们称 $(A \preceq)$，或更简单地称 A，是 (V'_1, V'_2) 的最小/最大划分。我们将证明，可以在 $O((|V_1| + |V_2|)^2)$ 时间内计算 (V_1, V_2) 的最小/最大划分。特别地，若 $G_1 = (V_1, E_1)$ 和 $G_2 = (V_1, E_2)$ 是同一张图 $G = (V, E)$ 的两个（不同的）副本，则可以在 $O(|V|^2)$ 时间内计算 V 的最小 / 最大划分。

我们对每个 $u \in V_1$ 计算 $\tau(\min_u)$，并对每个 $u \in V_2$ 计算 $\tau(\max_u)$。如果值为 3 的数量小于值为 1 的数量，那么（用时 $O(|V_1|^2 + |V_2|^2) = O((|V_1| + |V_2|)^2)$）我们按照前文定义构造图 $\bar{G}_1 = (\bar{V}_1, \bar{E}_1)$ 和 $\bar{G}_2 = (\bar{V}_2, \bar{E}_2)$，其中 $V_1 = \{\bar{u} \mid u \in V_1, \tau(\min_u) = 3\}$，$V_2 = \{\bar{u} \mid u \in V_2, \tau(\max_u) = 3\}$；否则我们考虑互补情形（其具有对称性）。构造 $G_1 = (V_1, E_1)$ 和 $G_2 = (V_2, E_2)$ 时，我们通过对 $(\gamma_{\min_u}, \mathtt{t}_{\min_u})$ 和 $(\gamma_{\max_u}, \mathtt{t}_{\max_u})$ 进行联合排序，定义唯一的字母表 $(\Sigma', \preceq')$。这是可行的，因为引理 17 对最大值也适用。注意，$|V_1| + |V_2| \leq (|V_1| + |V_2|)/2$。

假设我们已经递归地求得 $\bar{G}_1$ 和 $\bar{G}_2$ 的最小/最大划分。这将得到（$\{u \in \bar{V}_1 \mid \tau(\text{最小}_u) = 3\}$，$\{u \in V_2 \mid \tau(\text{最大}_u) = 3\}$）的最小/最大划分。接着，我们可以通过同时应用合并步骤，构造（V_1，V_2）的最小/最大划分。这是可行的，因为求极小值的合并步骤和求极大值的合并步骤都需要通过依次处理 $A_{c_1,2} A_{c_1,3}$、$A_{c_2,1}$、$A_{c_2,2}$、$A_{c_2,3}$、$A_{c_3,1}$ 等来构造 $A_{c,1}$。

由于我们得到与之前相同的递归关系，因此得出结论：我们可以在 $O((|V_1| + |V_2|)^2)$ 时间内计算 (V_1, V_2) 的最小/最大划分。

参考文献

1 Jarno Alanko、Giovanna D'Agostino、Alberto Policriti 和 Nicola Prezza。正则语言与前缀排序相遇。载于 Shuchi Chawla（编），第 31 届离散算法研讨会论文集（SODA'20），第 911–930 页。SIAM，2020。 doi:10.1137/1.9781611975994.55。2 Ruben Becker、Manuel Cáceres、Davide Cenzato、Sung-Hwan Kim、Bojana Kodric、Francisco Olivares 和 Nicola Prezza。通过划分细化对有限自动机排序。载于 Inge Li Gørtz、Martin Farach-Colton、Simon J. Puglisi 和 Grzegorz Herman（编），第 31 届欧洲算法年会（ESA 2023），Leibniz 信息学论文集（LIPIcs）第 274 卷，第 15:1–15:15 页，德国达格施图尔，2023。Schloss Dagstuhl – Leibniz 信息学中心。网址： https://drops.dagstuhl.de/opus/volltexte/2023/18668, doi:10.4230/LIPIcs.ESA.2023.15。3 Alexander Bowe、Taku Onodera、Kunihiko Sadakane 和 Tetsuo Shibuya。简洁 de Bruijn 图。载于 Ben Raphael 和 Jijun Tang（编），生物信息学算法，第 225–235 页，柏林、海德堡，2012。Springer Berlin Heidelberg。4 M. Burrows 和 D. J. Wheeler。一种块排序无损数据压缩算法。技术报告，1994。5 Alessio Conte、Nicola Cotumaccio、Travis Gagie、Giovanni Manzini、Nicola Prezza 和 Marinella Sciortino。在 Wheeler 确定性有限自动机上计算匹配统计量。载于 2023 年数据压缩会议（DCC），第 150–159 页，2023。 doi:10.1109/DCC55655.2023.00023。6 Nicola Cotumaccio。图可以在 $O(|E|^2+|V|^{5/2})$ 时间内以简洁方式建立索引以进行模式匹配。载于 2022 年数据压缩会议（DCC），第 272–281 页，2022。 doi:10.1109/DCC52660.2022.00035。7 Nicola Cotumaccio、Giovanna D'Agostino、Alberto Policriti 和 Nicola Prezza。按反字典序排列自动机和正则语言——第一部分。J. ACM，70(4)，2023 年 8 月。 doi:10.1145/3607471。8 Nicola Cotumaccio 和 Nicola Prezza。有限自动机的索引与压缩。载于 Dániel Marx（编），第 32 届离散算法研讨会论文集（SODA'21），第 2585–2599 页。SIAM，2021。 doi:10.1137/1.9781611976465.153。9 Massimo Equi、Veli Mäkinen 和 Alexandru I. Tomescu。除非 SETH 不成立，否则无法在多项式时间内为图建立索引，以实现亚二次时间的字符串匹配。载于 Tomáš Bureš、Riccardo Dondi、Johann Gamper、Giovanna Guerrini、Tomasz Jurdziński、Claus Pahl、Florian Sikora 和 Prudence W.H. Wong（编），SOFSEM 2021：计算机科学理论与实践，第 608–622 页，Cham，2021。Springer International Publishing。10 Massimo Equi、Veli Mäkinen、Alexandru I. Tomescu 和 Roberto Grossi。图上字符串匹配的复杂度。ACM Trans. Algorithms，19(3)，2023 年 4 月。 doi:10.1145/3588334。11 M. Farach。大字母表下的最优后缀树构造。载于第 38 届年度计算机科学基础研讨会论文集，第 137–143 页，1997。 doi:10.1109/SFCS.1997.646102。12 P. Ferragina、F. Luccio、G. Manzini 和 S. Muthukrishnan。构造带标签树结构以实现最优简洁性及更多应用。载于第 46 届 IEEE 计算机科学基础年会论文集（FOCS'05），第 184–193 页，2005。 doi:10.1109/SFCS.2005.69。13 P. Ferragina 和 G. Manzini。具有应用价值的机会主义数据结构。载于第 41 届计算机科学基础年会论文集（FOCS'00），第 390–398 页，2000。 doi:10.1109/SFCS.2000.892127。14 Paolo Ferragina、Fabrizio Luccio、Giovanni Manzini 和 S. Muthukrishnan。压缩和索引带标签的树及其应用。J. ACM，57(1)，2009 年 11 月。 doi:10.1145/1613676.1613680。15 Paolo Ferragina 和 Giovanni Manzini。压缩文本索引。J. ACM，52(4):552–581，2005 年 7 月。 doi:10.1145/1082036.1082039。16 Travis Gagie、Giovanni Manzini 和 Jouni Sirén。Wheeler 图：一种基于 BWT 的数据结构框架。《理论计算机科学》，698:67–78，2017。算法、字符串

以及大数据时代的理论方法（谨以此文献念拉斐尔·詹卡洛教授 60 岁寿辰）。 doi:10.1016/j.tcs.2017.06.016。17 Dan Gusfield。《字符串、树和序列算法：计算机科学与计算生物学》。剑桥大学出版社，1997 年。 doi:10.1017/CBO9780511574931。18 Ramana M. Idury 和 Michael S. Waterman。DNA 序列组装的新算法。《计算生物学杂志：计算分子细胞生物学期刊》，2 2:291–306，1995 年。19 Juha Kärkkäinen、Peter Sanders 和 Stefan Burkhardt。线性工作量的后缀数组构造。《ACM 杂志》，53(6):918–936，2006 年 11 月。 doi:10.1145/1217856.1217858。20 Dong Kyue Kim、Jeong Seop Sim、Heejin Park 和 Kunsoo Park。线性时间构造后缀数组。《离散算法杂志》，3(2):126–142，2005 年。组合模式匹配（CPM）特刊。网址： https://www.sciencedirect.com/science/article/pii/S1570866704000486, doi:https://doi.org/10.1016/j.jda.2004.08.019。21 Sung-Hwan Kim、Francisco Olivares 和 Nicola Prezza。确定性有限自动机的更快前缀排序算法。载于 Laurent Bulteau 和 Zsuzsanna Lipták 编，第三十四届年度组合模式匹配研讨会（CPM 2023），2023 年 6 月 26–28 日，法国马恩河谷，LIPIcs 第 259 卷，页码 16:1–16:16。Schloss Dagstuhl - Leibniz 信息学中心，2023 年。 doi:10.4230/LIPIcs.CPM.2023.16。22 Pang Ko 和 Srinivas Aluru。空间高效的线性时间后缀数组构造。《离散算法杂志》，3(2):143–156，2005 年。组合模式匹配（CPM）特刊。网址： https://www.sciencedirect.com/science/article/pii/S1570866704000498, doi:https://doi.org/10.1016/j.jda.2004.08.002。23 Veli Mäkinen、Niko Välimäki 和 Jouni Sirén。用于路径查询的图索引及其在基因组研究中的应用。《IEEE/ACM 计算生物学与生物信息学汇刊》，11:375–388，2014 年。24 U. Manber 和 G. Myers。后缀数组：一种在线字符串搜索的新方法。《SIAM 计算杂志》，22(5):935–948，1993 年。 doi:10.1137/0222058。25 Joong Chae Na。使用 o(n log n) 位工作空间为大字母表线性时间构造压缩后缀数组。载于 Alberto Apostolico、Maxime Crochemore 和 Kunsoo Park 编，《组合模式匹配》，页码 57–67，柏林、海德堡，2005 年。Springer Berlin Heidelberg。26 Robert Paige 和 Robert E. Tarjan。三种划分细化算法。《SIAM 计算杂志》，16(6):973–989，1987 年。 arXiv:https://doi.org/10.1137/0216062, doi:10.1137/0216062。27 Simon J. Puglisi、W. F. Smyth 和 Andrew H. Turpin。后缀数组构造算法分类。《ACM 计算调查》，39(2):4–es，2007 年 7 月。 doi:10.1145/1242471.1242472。28 Kunihiko Sadakane。功能完备的压缩后缀树。《理论计算机系统》，41(4):589–607，2007 年。 doi:10.1007/s00224-006-1198-x。29 Jared T. Simpson 和 Richard Durbin。使用 FM 索引高效构造组装字符串图。《生物信息学》，26(12):i367–i373，2010 年 6 月。 arXiv:https://academic.oup.com/bioinformatics/article-pdf/26/12/i367/48859513/bioinformatics_26_12_i367.pdf, doi:10.1093/bioinformatics/btq217。30 P. Weiner。线性模式匹配算法。载于第十四届 IEEE 年度开关与自动机理论研讨会论文集，页码 1–11，1973 年。 doi:10.1109/SWAT.1973.13。

第3节中的证明

引理 2 的陈述。设 $\alpha \in \Sigma^\omega$。设 $a \in \Sigma$ 和 $\alpha' \in \Sigma^\omega$ 使得 $\alpha = a\alpha'$。那么：
1. $\tau(\alpha) = 2$ 当且仅当 $\alpha' = a^\omega$，当且仅当 $\alpha = a^\omega$。
2. $\tau(\alpha) \neq 2$ 当且仅当 $\alpha' \neq a^\omega$，当且仅当 $\alpha \neq a^\omega$。

假设 $\tau(\alpha) \neq 2$。那么，存在唯一的 $c \in \Sigma \setminus \{a\}$、$\alpha'' \in \Sigma^\omega$ 和 $i \geq 0$，使得$\alpha' = a^i c\alpha''$（因此 $\alpha = a^{i+1}c\alpha''$）。此外：
1. $\tau(\alpha) = 1$ 当且仅当 $c \prec a$，当且仅当 $\alpha' \prec a^\omega$，当且仅当 $\alpha \prec a^\omega$。
2. $\tau(\alpha) = 3$ 当且仅当 $a \prec c$，当且仅当 $a^\omega \prec \alpha'$，当且仅当 $a^\omega \prec \alpha$，

证明。1。我们来证明 $\tau(\alpha) = 2$ 当且仅当 $\alpha' = a^\omega$。
($\Leftarrow$) 如果 $\alpha' = a^\omega$，我们有 $\alpha = a\alpha' = aa^\omega = a^\omega = \alpha'$，因此 $\tau(\alpha) = 2$。($\Rightarrow$) 假设 $\tau(\alpha) = 2$，于是 $\alpha = \alpha'$。为证明 $\alpha' = a^\omega$，只需证明对每个 $i \geq 1$，都有 $\alpha' = a^i\alpha'$。我们对 i 使用归纳法。注意到$\alpha' = \alpha = a\alpha'$，这就证明了我们关于 $i = 1$ 的断言。现在，假设 $i > 1$。根据归纳假设，我们可以假设 $\alpha' = a^{i-1}\alpha'$。因此，$\alpha' = \alpha = a\alpha' = a(a^{i-1}\alpha') = a^i\alpha'$。我们来证明 $\alpha' = a^\omega$ 当且仅当 $\alpha = a^\omega$。我们有 $\alpha' = a^\omega$ 当且仅当 $a\alpha' = aa^\omega$，当且仅当 $\alpha = a^\omega$。

2. 由否定第一点可得。
现在假设 $\tau(\alpha) \neq 2$。由于 $\alpha' \neq a^\omega$，因此存在唯一的 $c \in \Sigma \setminus \{a\}$、$\alpha'' \in \Sigma^\omega$和 $i \geq 0$，使得 $\alpha' = a^i c\alpha''$。此外：
<escape>1. 我们有 $\tau(\alpha) = 1$ 当且仅当 $\alpha' \prec \alpha$当且仅当 $\alpha' \prec a\alpha'$当且仅当$a^i c\alpha'' \prec a(a^i c\alpha'')$当且仅当 $a^i c\alpha'' \prec a^i(ac\alpha'')$当且仅当 $c\alpha'' \prec ac\alpha''$当且仅当 $c \prec a$当且仅当 $a^i c\alpha'' \prec a^\omega$当且仅当 $\alpha' \prec a^\omega$当且仅当 $a\alpha' \prec aa^\omega$当且仅当 $\alpha \prec a^\omega$。</escape>

2. 对前述各点取否定即可得到。

◀

推论 3 的陈述。设 α，$\beta \in \Sigma^\omega$。设 a，$b \in \Sigma$ 和 α'，$\beta' \in \Sigma^\omega$，使得 $\alpha = a\alpha'$ 且 $\beta = b\beta'$。则：
1. 如果 $a = b$ 且 $\tau(\alpha) = \tau(\beta) = 2$，那么 $\alpha = \beta$。2. 如果 $a = b$ 且 $\tau(\alpha) < \tau(\beta)$，那么 $\alpha \prec \beta$。等价地，如果 $a = b$ 且 $\alpha \preceq \beta$，那么 $\tau(\alpha) \leq \tau(\beta)$。

证明。1. 由引理 2，我们有 $\alpha = a^\omega = b^\omega = \beta$。
2. 下面证明，如果 $a = b$ 且 $\tau(\alpha) < \tau(\beta)$，那么 $\alpha \prec \beta$。我们分三种情况讨论。a. $\tau(\alpha) = 1$ 且 $\tau(\beta) = 2$。由引理 2，我们有 $\alpha \prec a^\omega = b^\omega = \beta$。b. $\tau(\alpha) = 2$ 且 $\tau(\beta) = 3$。由引理 2，我们有 $\alpha = a^\omega = b^\omega \prec \beta$。c. $\tau(\alpha) = 1$ 且 $\tau(\beta) = 3$。由引理 2，我们有 $\alpha \prec a^\omega = b^\omega \prec \beta$。最后，如果 $a = b$ 且 $\alpha \preceq \beta$，那么 $\tau(\alpha) \leq \tau(\beta)$，因为如果是 $\tau(\beta) \leq \tau(\alpha)$，则会推出 $\beta \prec \alpha$，这与矛盾相矛盾。

◀

B 第 4 节的证明

引理 5 的陈述。设 $G = (V, E)$ 是一个图。那么，我们可以在 $O(|E|)$ 时间内构造一个修剪图 $G^* = (V^*, E^*)$，其中 $V^* = \{u^* \mid u \in V\}$，使得对于每个 $u \in V$，都有 $\min_{u^*} = \min_u$。特别地，对于每个 $u \in V$，$\tau(u^*) = \tau(u)$。

证明。对于每个 $u^* \in V^*$，定义 $\lambda(u^*) = \lambda(u)$：$F = \{(u, v) \in E \mid \tau(v) = 1, \lambda(v) \prec \lambda(u)\}$，并定义 $E^* = \{(u^*, v^*) \mid (u, v) \in E \setminus F\}$。本质上，我们通过从 G 中移除所有违反修剪图定义的边来定义 G^*。显然，G^* 仍然满足我们在这篇论文中对图所作的假设。

1. 按照构造，V^* 中进入同一节点的所有边都具有相同的标签。2. G^*中的每个 $v^* \in G^*$ 都有一条入边。事实上，如果 $\tau(v) \neq 1$，那么若 $u \in V$ 是满足$(u, v) \in E$的任意节点，则$(u, v) \notin F$，因此$(u^*, v^*) \in E^*$。如果 $\tau(v) = 1$，那么存在 $u \in V$使得$(u, v) \in E$且 $\lambda(u) \preceq \lambda(v)$，因此$(u, v) \notin F$且$(u^*, v^*) \in E^*$。

3. G^* 是确定性的，因为它本质上是一个确定性图的子图。更准确地说，假设(u^*, v^*).$(u^*, z^*) \in E^*$ 满足 $\lambda(v^*) = \lambda(z^*)$。我们必须证明 $v^* = z^*$，或等价地，证明 $v = z$。由 (u^*, v^*)，$(u^*, z^*) \in E^*$，我们得到(u, v)，$(u, z) \in E$；因此，由 $\lambda(v) = \lambda(v^*) = \lambda(z^*) = \lambda(z)$，我们得到 $v = z$，因为 G 是确定性的。

现在，我们来证明，对于每个 $u \in V$，都有 $\min_{u^*} = \min_u$。为此，我们需要证明：(i) $\min_u \in I_{u^*}$；(ii) 若 $\alpha \in I_{u^*}$，则 $\min_u \preceq \alpha$。
1. 我们来证明 $\min_u \in I_{u^*}$。为此，只需证明：若 $(u_i)_{i \geq 1}$ 是从 u 开始的 $\min_u$ 的一次出现，那么 $(u_i^*)_{i \geq 1}$ 是从 u^* 开始的 $\min_u$ 的一次出现。对于每个 $i \geq 1$，都有 $\lambda(u_i^*) = \lambda(u_i) = \min_u[i]$。我们只需再证明，对于每个 $i \geq 1$，都有 $(u_{i+1}^*, u_i^*) \in E^*$。我们知道 $(u_{i+1}, u_i) \in E$，因此只需证明 $(u_{i+1}, u_i) \notin F$。假设 $\tau(u_i) = 1$。我们必须证明 $\lambda(u_{i+1}) \preceq \lambda(u_i)$。由于 $(u_j)_{j \geq i}$ 是从 u_i 开始的 $\min_{u_i}$ 的一次出现，这等价于证明 $\min_{u_i}[2] \preceq \min_{u_i}[1]$，而这确实可由 $\tau(u_i) = 1$ 得出。
2. 证明如果 $\alpha \in I_{u^*}$，那么 $\alpha \in I_u$（这意味着 $\min_u \preceq \alpha$）。令 $(v_i^*)_{i \geq 1}$ 为一个从 u^* 开始的 α 出现。只需证明 $(v_i)_{i \geq 1}$ 是一个从 u 开始的 α 出现。对于每个 $i \geq 1$，我们有 $(v_{i+1}, v_i) \in E$，因为 $(v_{i+1}^*, v_i^*) \in E$，并且 $\lambda(v_i) = \lambda(v_i) = \alpha[i]$，因此结论成立。

特别地，对于每个 $u \in V$，$\tau(u^*) = \tau(u)$，并且 $G^* = (V^*, E^*)$ 是一个修剪图。 ◀

B.1 第4.1节的证明

引理 6 的陈述。设 $G = (V, E)$ 是一个图，且 $u, v \in V$。
1. 若 $(u, v) \in E$and $\lambda(u) \prec \lambda(v)$，则 $\tau(v) = 1$. 2. 若 $(u, v) \in E$, $\lambda(u) = \lambda(v)$ 且 $\tau(u) = 1$，则 $\tau(v) = 1$.

证明。1. 由于 $\lambda(u) \prec \lambda(v)$，我们有 $\min_u \prec \lambda(v)^\omega$。此外，$\lambda(v) \min_u \in I_v$，因此 $\min_v \preceq \lambda(v) \min_u \prec \lambda(v)\lambda(v)^\omega = \lambda(v)^\omega$，故由引理 2 可得 $\tau(v) = 1$。2. 由于 $\tau(u) = 1$，因此由引理 2 可得 $\min_u \prec \lambda(u)^\omega$。由 $(u, v) \in E$ 可得 $\lambda(v) \min_u \in I_v$，因此 $\min_v \preceq \lambda(v) \min_u \prec \lambda(v)\lambda(u)^\omega = \lambda(v)\lambda(v)^\omega = \lambda(v)^\omega$，所以再次由引理 2 可得 $\tau(v) = 1$。

◀

推论 7 的陈述。设 $G = (V, E)$ 是一个图，且设 $u \in V$。那么，$\tau(u) = 1$当且仅当存在 $k \geq 2$ 和 $z_1, \ldots, z_k \in V$，使得 (i) 对每个$1 \leq i \leq k - 1$，$(z_i, z_{i+1}) \in E$；(ii) $z_k = u$；(iii) $\lambda(z_1) \prec \lambda(z_2)$；以及 (iv) $\lambda(z_2) = \lambda(z_3) = \cdots = \lambda(z_k)$。

证明。（⇐）设 $k \geq 2$和 $z_1, \ldots, z_k \in V$满足：(i) 对每个 $1 \leq i \leq k-1$，$(z_i, z_{i+1}) \in E$；(ii) $z_k = u$；(iii) $\lambda(z_1) \prec \lambda(z_2)$；(iv) $\lambda(z_2) = \lambda(z_3) = \cdots = \lambda(z_k)$。我们必须证明$\tau(u) = 1$。下面用归纳法证明，对每个 $2 \leq i \leq k$，都有 $\tau(z_i) = 1$（特别是 $\tau(u) = \tau(z_k) = 1$）。若 $i = 2$，则由 $(z_1, z_2) \in E$和 $\lambda(z_1) \prec \lambda(z_2)$，根据引理 6 可得$\tau(z_2) = 1$。现在，假设 $3 \leq i \leq k$。根据归纳假设，我们

有 $\tau(z_{i-1}) = 1$。由于 $(z_{i-1}, z_i) \in E$ 且 $\lambda(z_{i-1}) = \lambda(z_i)$，因此根据引理 6 可得$\tau(z_i) = 1$。

(⇒) 假设 $\tau(u) = 1$。那么，由引理 2 可知，存在 $c \in \Sigma \setminus \{\lambda(u)\}$、$\gamma' \in \Sigma^\omega$和 $j \geq 1$，使得 $\min_u = \lambda(u)^j c\gamma'$ 且 $c \prec \lambda(u)$。令 $(u_i)_{i\geq 1}$ 为从 u开始的 $\min_u$的一次出现。那么，(i) $u = u_1$，(ii) 对每个 $i \geq 1$，$(u_{i+1}, u_i) \in E$，(iii) 对$1 \leq i \leq j$，$\lambda(u_i) = \lambda(u)$，以及 (iv) $\lambda(u_{j+1}) = c$。因此，如果 $k = j + 1$ 且对每个 $1 \leq i \leq k$都有 $z_i = u_{k+1-i}$，那么 (i) 对每个 $1 \leq i \leq k-1$，$(z_i, z_{i+1}) = (u_{k-i+1}, u_{k-i}) \in E$，(ii)$z_k = u_1 = u$，(iii) $\lambda(z_1) = \lambda(u_k) = \lambda(u_{j+1}) = c \prec \lambda(u) = \lambda(u_j) = \lambda(u_{k-1}) = \lambda(z_2)$，以及 (iv)$\lambda(z_2) = \lambda(z_3) = \cdots = \lambda(z_k)$，因为 $\lambda(u_j) = \lambda(u_{j-1}) = \cdots = \lambda(u_1)$。◀

推论 8 的陈述。设 $G = (V, E)$ 为一个图。我们可以确定所有 $u \in V$ such that $\tau(u) = 1$ 在时间 $O(|E|)$。

证明。在我们的算法中，我们将通过图遍历恰好标记所有满足 $\tau(u) = 1$ 的节点 u。我们将使用一个（初始为空的）队列。初始时，没有节点被标记。首先，处理所有边 $(u, v) \in E$，如果 $\lambda(u) \prec \lambda(v)$ 且 v 之前尚未被标记，则标记v 并将 v 加入队列。此步骤耗时 $O(m)$。接下来，递归地从队列中取出一个元素 u，并考虑唯一满足 $(u, v) \in E$ 且 $\lambda(u) = \lambda(v)$ 的 $v \in Q$（如果存在）。如果 v 之前尚未被标记，则标记 v 并将 v 加入队列。由于每条边至多被考虑一次，此步骤也耗时 $O(|E|)$。算法结束时，根据推论 7，当且仅当 $\tau(u) = 1$ 时，节点 u 已被标记。◀

引理 9 的陈述。设 $G = (V, E)$ 是一个图，且 $u \in V$满足 $\tau(u) \neq 1$。那么，$\tau(u) = 2$ 当且仅当存在 $k \geq 2$ 和 $z_1, \ldots\ldots, z_k \in V$，使得 (i) 对每个 $1 \leq i \leq k-1$，$(z_{i+1}, z_i) \in E$；(ii) $z_1 = u$；(iii) 对某个 $1 \leq j \leq k-1$，$z_k = z_j$；以及 (iv) $\lambda(z_1) = \lambda(z_2) = \ldots\ldots = \lambda(z_k)$。

特别是，这样的 $z_1, \ldots\ldots, z_k \in V$必须对每个 $1 \leq i \leq k$都满足 $\tau(z_i) = 2$。

证明。（⇐）设 $k \geq 2$和 $z_1, \ldots, z_k \in V$满足 (i) 对每个 $1 \leq i \leq k-1$，$(z_{i+1}, z_i) \in E$；(ii) $z_1 = u$；(iii) 对某个 $1 \leq j \leq k-1$，$z_k = z_i$；以及 (iv) $\lambda(z_1) = \lambda(z_2) = \cdots = \lambda(z_k)$。我们必须证明 $\tau(u) = 2$。注意，$\min_u = \min_{z_1} \preceq \lambda(z_1)\lambda(z_2) \ldots \lambda(z_{j-1}) \min^{z_j} = \lambda(u)^{j-1} \min^{z_j}$。类似地，我们有 $\min_{z_j} \preceq \lambda(z_j)\lambda(z_{j+1}) \ldots \lambda(z_{k-1}) \min_{z_k} = \lambda(u)^{k-j} \min^{z_j}$。由于 $\min_{z_j} \preceq \lambda(u)^{k-i} \min^{z_j}$，根据归纳法可得，对每个$h \geq 1$，$\min_{z_i} \preceq (\lambda(u)^{k-j})^h \min^{z_i}$，因此 $\min_{z_j} \preceq \lambda(u)^\omega$。由此可知，$\min_u \preceq \lambda(u)^{j-1} \min_{z_j} \preceq \lambda(u)^{j-1}\lambda(u)^\omega = \lambda(u)^\omega$，因此根据引理 2，$\tau(u) \preceq 2$。由于 $\tau(u) \neq 1$，我们得出 $\tau(u) = 2$。

(⇒) 假设 $\tau(u) = 2$。特别地，$\lambda(u)^\omega \in I_u$，因此存在一个 $\lambda(u)^\omega$ 从 u 开始的出现 $(z_i)_{i\geq 1}$。这意味着 (i) (z_{i+1}, z_i) 对每个 $i \geq 1$ 都有 $\in E$，(ii) 对每个 $i \geq 1$，$z_1 = u$ 且 $\lambda(z_i) = \lambda(u)$。由于 V 是有限的，存在 $1 \leq j < k$ 使得 $z_j = z_k$。于是，$k \geq 2$ 和 $z_1, \ldots, z_k \in V$ 具有所需的性质。

现在，我们来证明，如果存在具有所需性质的 z_1、……、$z_k \in V$，那么对每个 $1 \leq i \leq k$ 都有 $\tau(z_i) = 2$。注意，对每个 $1 \leq i \leq k$ 都必须有 $\tau(z_i) \neq 1$，否则根据推论 7，我们就会得出 $\tau(u) = 1$。因此，根据我们刚刚证明的刻画，对每个 $1 \leq i \leq k$ 都必须有 $\tau(z_i) = 2$。◀

推论 10 的陈述。设 $G = (V, E)$ 是一个图。我们可以在 $O(|E|)$ 时间内确定所有满足 $\tau(u) = 2$ 的 $u \in V$。

证明。根据推论 8，我们可以假设已经计算出所有 $u \in V$使得$\tau(u) = 1$。为了计算所有 $u \in V$使得 $\tau(u) = 2$，我们使用广度优先搜索算法探索 G，并作如下修改：(i) 我们只从满足 $u \in V$使得 $\tau(u) \neq 1$的节点出发；(ii) 我们以反向方式而非正向方式沿边遍历；

(iii) 如果我们从 $u \in V$ 出发，就只沿标签为 $\lambda(u)$ 的边探索图（也就是说，我们首先只考虑所有满足 $(u', u) \in E$ 且 $\lambda(u') = \lambda(u)$ 的 $u' \in V$，然后以广度优先搜索的方式重复这一步）。注意，在搜索过程中，我们绝不会遇到满足 $\tau(v) = 1$ 的节点 v，否则根据推论 7，我们将得到 $\tau(u) = 1$。搜索过程中，对于遇到的每个节点 v，我们都会推断出值 $\tau(v)$。假设我们从节点 $u \in V$ 开始探索。如果在 u 处的探索过程中，我们遇到一个节点 v，而我们已经判定 $\tau(v) = 3$，那么就不考虑进入 v 的边，并进行回溯（因为引理 9 的刻画中的所有 z_i 都必须满足 $\tau(z_i) = 2$）。如果在 u 处的探索过程中，我们遇到一个节点 v，而我们已经判定 $\tau(v) = 2$，那么就直接回溯到 u；根据引理 9，我们还可以判定，回溯过程中遇到的所有 z（包括 u）都满足 $\tau(z) = 2$。如果在探索 u 的过程中，我们两次遇到同一个节点，那么就直接回溯到 u；根据引理 9，我们还可以判定，回溯过程中遇到的所有 z（包括 u）都满足 $\tau(z) = 2$。如果在探索 u 的过程中，我们从未遇到已判定 $\tau(v)$ 的节点 v，也从未两次遇到同一个节点（因此最终会陷入停滞），那么根据引理 9，我们可以判定，遇到的所有 v（包括 u）都满足 $\tau(v) = 3$。该算法的运行时间为 $O(|E|)$，因为我们不会重复遍历同一条边。◀

B.2 第 4.2 节的证明

引理 12 的陈述。设 $G = (V, E)$ 是一个图。设 $u, v \in V$ 满足 $\min_u = \min_v$。设 $(u_i)_{i \geq 1}$ 是 $\min_u$ 的一次出现，且 $(v_i)_{i \geq 1}$ 是 $\min_v$ 的一次出现。那么：

1. $\lambda(u_i) = \lambda(v_i)$ 对于每个 $i \geq 1$。
2. $\min_{u_i} = \min_{v_i}$ 对于每个 $i \geq 1$。
3. $\tau(u_i) = \tau(v_i)$ 对于每个 $i \geq 1$。

特别地，若 $u = v$ 和 $(u_i)_{i \geq 1}$ 以及 $(v_i)_{i \geq 1}$ 是 $\min_u$ 的两个不同出现，则前述结果成立。

证明。1. 对于每个 $i \geq 1$，我们有 $\lambda(u_i) = \min_u[i] = \min_v[i] = \lambda(v_i)$。
2. 固定 $i \geq 1$。我们有 $\min_u[1, i-1]\min_{u_i} = \min_u = \min_v = \min_v[1, i-1]\min_{v_i}$，因此 $\min_{u_i} = \min_{v_i}$3。3. 固定 $i \geq 1$。根据前面的各点，我们有 $\min_{u_i} = \min_{v_i}$，且 $\lambda(u_i) = \lambda(v_i)$，因此由推论 3 可得，必然有 $\tau(u_i) = \tau(v_i)$。

◀

引理13的陈述。设 $G = (V, E)$ 是一个图。设 $u \in V$，并设 $(u_i)_{i \geq 1}$ 是从 u 开始的 $\min_u$ 的一次出现。设 $k \geq 1$ 满足 $\tau(u_1) = \tau(u_2) = \cdots = \tau(u_{k-1}) = \tau(u_k) \neq 2$。那么，$u_1$、.....、$u_k$ 两两不同。特别地，$k \leq |V|$。

证明。我们假设 $\tau(u_1) = \tau(u_2) = \cdots = \tau(u_{k-1}) = \tau(u_k) = 1$，因为 $\tau(u_1) = \tau(u_2) = \cdots = \tau(u_{k-1}) = \tau(u_k) = 3$ 这一情形是对称的。注意，对于每个 $1 \leq l \leq k$，$(u_i)_{i \geq l}$ 都是从 u_l 开始的 $\min_{u_l}$ 的一次出现，且 $\tau(u_l) = 1$，因此 $\lambda(u_{l+1}) \preceq \lambda(u_l)$。特别地，对于每个 $1 \leq r \leq s \leq k+1$，都有 $\lambda(u_s) \preceq \lambda(u_r)$。

为导出矛盾，假设存在 $1 \leq i < j \leq k$ 使得 $u_i = u_j$。那么，$\min_{u_i} = \min_{u_i}[1, j-i]\min_{u_j} = \min_{u_i}[1, j-i]\min_{u_{i_0}}$。根据归纳法，我们得到 $\min_{u_i} = (\min_{u_i}[1, j-i])^t \min_{u_i}$，对每个 $t \geq 1$ 都成立，因此 $\min_{u_i} = (\min_{u_i}[1, j-i])^\omega$。对于每个 $1 \leq h \leq j$，我们有 $\lambda(u_j) \preceq \lambda(u_h) \preceq \lambda(u_i)$。由于 $u_i = u_j$，因此 $\lambda(u_i) = \lambda(u_j)$，所以 $\lambda(u_h) = \lambda(u_i)$，对

每个 $i \le h \le j$ 都成立，这意味着 $\min_{u_i}[1, j-i] = \lambda(u_i)\lambda(u_{i+1}) \dots \lambda(u_{j-1}) = \lambda(u_i)^{j-i}$，因此 $\min^{u_i} = (\min_{u_i}[1, j-i])^{\omega} = \lambda(u_i)^{\omega}$。根据引理 2，我们得出 $\tau(u_i) = 2$，这就产生了矛盾。◀

引理15的陈述。设 $G = (V, E)$ 为一个图。令 $u \in V$ 使得 $\tau(u) = 3$。那么，对于每个 $2 \le i \le \ell_u - 1$，$\min_u[i+1] \preceq \min_u[i]$。特别地，如果 $2 \le i \le j \le \ell_u$，则 $\min_u[j] \preceq \min_u[i]$。

证明。固定 $2 \le i \le \ell_u$。根据 ℓ_u) 的定义，我们有 $\tau(u_i) = 1$，因此 最小值$_u[i+1]$ = 最小值$_{u_i}[2] \preceq$最小值$_{u_i}[1]$ =最小值$_u[i]$。◀

引理16的陈述。令 $G = (V, E)$ 为一个图。令 $u \in V$ 使得 $\tau(u) = 3$。
1. 对每个 $1 \le i \le \ell_u$，$G_i(u)$ 都有定义且非空。 2. 设 $(u_i)_{i \ge 1}$ 是从 u 开始的 $\min_u$ 的一次出现。那么，对于每个 $1 \le i \le \ell_u$，$u_i \in G_i(u)$。特别地，对于每个 $1 \le i \le \ell_u$，$\tau(u_i) = \tau(G_i(u))$，且 $\min_u[i] = \lambda(u_i) = \lambda(G_i(u))$。 3. 对于每个 $1 \le i \le \ell_u$ 以及每个 $v \in G_i(u)$，存在一次从 u 开始并在 v 结束的 $\min_u[1, i-1]$ 的出现。

证明。我们对 $i \ge 1$ 进行归纳法。若 $i = 1$，则 $G_i(u) = \{u\}$ 定义良好且非空，并且 $u_1 = u \in G_i(u)$。此外，若 $v \in G_i(u)$，则 $v = u$，且存在 $\epsilon = \min[1, 0]$ 的一次出现，其起点和终点均为 u。现在，假设 $2 \le i \le \ell_u$。

首先，注意到 $\mathrm{F}(\{v' \in Q \mid (\exists v \in G_{i-1}(u))((v', v) \in E)\})$ 定义良好且非空。事实上，根据归纳假设，$G_{i-1}(u)$ 定义良好且非空，因此 $\{v' \in Q \mid (\exists v \in G_{i-1}(u))((v', v) \in E)\}$ 非空，因为每个节点都有一条入边。

我们来证明 $u_i \in \mathrm{F}(\{v' \in Q \mid (\exists v \in G_{i-1}(u))((v', v) \in E)\})$。令 $R = \{v' \in Q \mid (\exists v \in G_{i-1}(u))((v', v) \in E)\}$，并令 $R' = \arg\min_{v \in R} \lambda(v)$。我们必须证明 $u_i \in \arg\min_{u \in R'} \tau(u)$。1. 我们来证明 $u_i \in R$。根据归纳假设，我们知道 $u_{i-1} \in G_{i-1}(u)$。我们知道 $(u_i, u_{i-1}) \in E$，因此 $u_i \in R$。

2. 让我们证明 $u_i \in R'$。令 $v' \in R$。我们必须证明 $\lambda(u_i) \preceq \lambda(v')$。由于$v' \in R$，存在 $v \in G_{i-1}(u)$，使得 $(v', v) \in E$。由 $u_{i-1}, v \in G_{i-1}(u)$，我们得到 $\lambda(u_{i-1}) = \lambda(v)$。根据归纳假设，存在一个 $\min_u[1, i-2]$ 的出现，其起始于 u，结束于 v，因此存在一个 $\min_u[1, i-2]\lambda(v) = \min_u[1, i-2]\lambda(u_{i-1}) = \min_u[1, i-1]$ 的出现，其起始于 u，结束于 v'。由于 $\min_u = \min_u[1, i-1]\min^{u_i}$，必有 $\min_{u_i} \preceq \min_{v'}$，特别是 $\lambda(u_i) \preceq \lambda(v')$。

3. 让我们证明 $u_i \in \arg\min_{u \in R'} \tau(u)$。令 $v' \in R'$。我们必须证明 $\tau(u_i) \le \tau(v)$。特别地，$v' \in R$，因此如前所述，我们得到 $\min_{u_i} \preceq \min_{v'}$。此外，由 $u_i, v' \in R'$可得 $\lambda(u_i) = \lambda(v')$。由推论 3，我们得出 $\tau(u_i) \le \tau(v)$。

我们来证明 $u_i \notin \bigcup_{j=2}^{i-1} G_j(u)$。若 $i = \ell_u$，则有 $\tau(u_i) = \tau(u_{\ell_u}) \ge 2$，结论随之成立，因为对于每个 $2 \le j \le i-1$，$\tau(G_j(u)) = \tau(u_j) = 1$。因此，在下文中我们可以假设 $2 \le i \le \ell_u - 1$。特别地，$\tau(u_i) = 1$，所以由引理 2，存在 $k \ge 1, c \in \Sigma \setminus \{\lambda(u_i)\}$ 和 $\gamma' \in \Sigma^{\omega}$，使得 $\min_u = \lambda(u)^k c\gamma'$ 且 $c \prec \lambda(u_i)$。反设存在 $2 \le j \le i-1 \le \ell_u - 2$ 使得$u_i \in G_j(u)$。我们知道 $u_j \in G_j(u)$，特别是 $\lambda(u_i) = \lambda(G_j(u)) = \lambda(u_j)$。由引理15可得，对于每个 $j \le h \le i$，$\min_u[i] \preceq \min_u[h] \preceq \min_u[j]$，或者等价地，对于每个 $j \le h \le i$，$\lambda(u_i) \preceq \lambda(u_h) \preceq \lambda(u_j)$。由于 $\lambda(u_j) = \lambda(u_i)$，我们得出，对于每个 $j \le h \le i$，$\lambda(u_h) = \lambda(u_i)$。因此，$\min_u = \min_u[1, i-1]\min^{u_i} = \min_u[1, j-1]\min_u[j, i-1]\min^{u_i} = \min_u[1, j-1]\lambda(u_i)^{i-j}\lambda(u_i)^k c\gamma' = \min_u[1, j-1]\lambda(u_i)^{k+i-j}c\gamma'$。另一方面，由于 $u_i \in G_j(u)$ 且 $j < i$，根据归纳假设，存在一个从 u 开始、到 u_i 结束的 $\min_u[1, j-1]$ 的出现，因此 $\min_u[1, j-1]\min_{u_i} \in I_u$。因此，根据 $\min_u$ 的极小性，我们得到 $\min_u \preceq \min_u[1, j-1]\min^{u_i}$，或者在线观看中文字幕

等价地，$\min_u[1,j-1]\lambda(u_i)^{k+i-j}c\gamma' \preceq \min_u[1,j-1]\lambda(u_i)^k c\gamma'$。由于 $j<i$，我们得到$\lambda(u_i) \preceq c$，这与矛盾相矛盾。

我们来证明 $G_i(u)$ 定义良好且非空，并且 $u_i \in G_i(u)$。这由 $u_i \in \mathrm{F}(\{v' \in Q \mid (\exists v \in G_{i-1}(u))((v',v) \in E)\})$ 和 $u_i \notin \bigcup_{j=2}^{i-1} G_j(u)$ 得出。

最后，我们来证明，如果 $v' \in G_i(u)$，那么存在 $\min[1,i-1]$ 的一次出现，其起始位置为 u，结束位置为 v'。由于 $v' \in G_i(u)$，因此存在 $v \in G_{i-1}(u)$，使得 $(v',v) \in E$。特别地，$\lambda(v) = \lambda(G_{i-1}(u)) = \lambda(u_{i-1})$。根据归纳假设，存在 $\min_u[1,i-2]$ 的一次出现，其起始位置为 u，结束位置为 v，因此存在 $\min_u[1,i-2]\lambda(v) = \min_u[1,i-2]\lambda(u_{i-1}) = \min_u[1,i-1]$ 的一次出现，其起始位置为 u，结束位置为 v'。◀

引理 17 的陈述。设 $G=(V,E)$ 是一个图。设 $u,v \in V$，使得 $\tau(u) = \tau(v) = 3$。假设以下陈述之一成立：

1. γ_u 不是 γ_v 的前缀，并且 $\gamma_u \prec \gamma_v$。
2. $\gamma_u = \gamma_v$，$\mathtt{t}_u = 2$ 以及 $\mathtt{t}_v = 3$。
3. γ_v 是真前缀 γ_u。

那么，$\min_u \prec \min_v$。

等价地，如果 $\min_u \preceq \min_v$，那么以下情况之一成立：(i) γ_u 不是 γ_v和 $\gamma_u \prec \gamma_v$的前缀；(ii) $\gamma_u = \gamma_v$ 和 $\mathtt{t}_u \le \mathtt{t}_v$；(iii) γ_v 是 γ_u的真前缀。

证明。我们来证明，如果三个陈述 (1)–(3) 中有一个成立，那么 $\min_u \prec \min_v$。1. 如果 $\min_u[1,\ell_u] \prec \min_v[1,\ell_v]$，且 $\min_u[1,\ell_u]$ 不是 $\min_v[1,\ell_v]$的前缀，那么 $\min_u \prec \min_v$。
2. 假设 $\min_u[1,\ell_u] = \min_v[1,\ell_v]$（因此特别有 $\ell_u = \ell_v$），$\mathtt{t}_u = 2$ 且 $\mathtt{t}_v = 3$。令 $(u_i)_{i\ge1}$ 为 u 的一次出现，并令 $(v_i)_{i\ge1}$ 为 v 的一次出现。成立 $\min_u = \min_u[1,\ell_u-1]\min_{u_{\ell_u}}$，且 $\min_v = \min_v[1,\ell_v-1]\min_{v_{\ell_v}}$。由于 $\min_u[1,\ell_u-1] = \min_v[1,\ell_v-1]$，为了证明 $\min_u \prec \min_v$，我们只需证明 $\min_{u_{\ell_u}} \prec \min_{v_{\ell_v}}$。由引理16可得 $u_{\ell_u} \in G_{\ell_u}(u)$ 和 $v_{\ell_v} \in G_{\ell_v}(v)$，因此 $\lambda(u_{\ell_u}) = \min_u[\ell_u] = \min_v[\ell_v] = \lambda(v_{\ell_v})$。由于 $\tau(u_{\ell_u}) = \tau(G_{\ell_u}(u)) = \mathtt{t}_u = 2$，且 $\tau(v_{\ell_v}) = \tau(G_{\ell_v}(v)) = \mathtt{t}_v = 3$，由推论 3 得出 $\min_{u_{\ell_u}} \prec \min_{v_{\ell_v}}$。
3. 假设 γ_v 是 γ_u（因此也是 $\ell_v < \ell_u$）的真前缀。特别地，$\min_u[1,\ell_v] = \min_v[1,\ell_v]$，因此，与前一种情况类似，我们只需证明 $\min_{u_{\ell_v}} \prec \min_{v_{\ell_v}}$。同样，我们得到 $u_{\ell_v} \in G_{\ell_v}(u)$，$v_{\ell_v} \in G_{\ell_u}(v)$，$\lambda(u_{\ell_v}) = \lambda(v_{\ell_v})$ 以及 $\tau(v_{\ell_v}) = 3$。由于 $\ell_v < \ell_u$，ℓ_u 的极小性意味着 $\tau(u_{\ell_v}) = \tau(G_{\ell_v}(u)) = 1$。由推论 3 可得 $\min_{u_{\ell_v}} \prec \min_{v_{\ell_v}}$。

现在，我们来证明：如果 $\min_u \preceq \min_v$，那么以下情况至少有一种成立：(i) γ_u 不是 γ_v 和 $\gamma_u \prec \gamma_v$ 的前缀；(ii) $\gamma_u = \gamma_v$ 和 $\mathtt{t}_u \le \mathtt{t}_v$；(iii) γ_v 是 γ_u 的真前缀。如果这不成立，那么以下情况至少有一种成立：(1) γ_v 不是 γ_u 和 $\gamma_v \prec \gamma_u$ 的前缀；(2) $\gamma_v = \gamma_u$、$\mathtt{t}_v = 2$ 和 $\mathtt{t}_u = 3$；(3) γ_u 是 γ_v 的真前缀。于是我们会得出 $\min_v \prec \min_u$，这就产生了矛盾。◀

引理18 的陈述。设 $G=(V,E)$ 是一个图。设 $u,v \in V$ 满足 $\tau(u) = \tau(v) = 3$, $\gamma_u = \gamma_v$ 且 $\mathtt{t}_u = \mathtt{t}_v = 2$。那么，$\min_u = \min_v$。

证明。首先，注意到 $\gamma_u = \gamma_v$ 蕴含 $\ell_u = \ell_v$，且这等价于 $\min_u[1,\ell_u] = \min_v[1,\ell_v]$。现在，令 $(u_i)_{i\ge1}$ 为一个从 u 开始的 $\min_u$ 的出现，并令 $(v_i)_{i\ge1}$ 为一个从 v 开始的 $\min_v$ 的出现。我们有 $\min_u = \min_u[1,\ell_u-1]\min_{u_{\ell_u}}$，且 $\min_v = \min_v[1,\ell_v-1]\min_{v_{\ell_v}}$。由于 $\min_u[1,\ell_u-1] = \min_v[1,\ell_v-1]$，我们只需证明 $\min_{u_{\ell_u}} = \min_{v_{\ell_v}}$。注意到 $\lambda(u_{\ell_u}) = \min_u[\ell_u] = \min_v[\ell_v] = \lambda(v_{\ell_v})$，由引理16可得 $\tau(u_{\ell_u}) = \tau(G_{\ell_u}(u)) = \mathtt{t}_u = 2$，且 $\tau(v_{\ell_v}) = \tau(G_{\ell_v}(v)) = \mathtt{t}_v = 2$。由推论 3，我们得出 $\min_{u_{\ell_u}} = \min_{v_{\ell_v}}$。◀

22 前缀排序确定有限自动机：一种递归算法

引理19的陈述。设 $G=(V,\ E)$ 是一张图。设 $u\in V$，并令 $(u_i)_{i\geq 1}$ 为一个从 u 开始的 $\min_u$ 的出现。那么，以下情况恰有一种成立：
1. 存在 $i_0\geq 1$ 使得 $\tau(u_i)\neq 2$ 对每个 $1\leq i<i_0$ 且 $\tau(u_i)=2$ 对每个 $i\geq i_0$。 2. $\tau(u_i)\neq 2$ 对每个 $i\geq 1$，并且 $\tau(u_i)=1$ 和 $\tau(u_i)=3$ 对于无穷多个 i 都成立。

证明。如果对于每个 $i\geq 1$ 都有 $\tau(u_i)\neq 2$，那么 $\tau(u_i)=1$ 和 $\tau(u_i)=3$ 对无穷多个 i's 都成立，因为否则我们可以选择 $j\geq 1$，使得要么对于每个 $i\geq j$ 都有 $\tau(u_i)=1$，要么对于每个 $i\geq j$ 都有 $\tau(u_i)=3$；在这两种情况下，引理13都会导致矛盾，因为 $(u_j)_{j\geq i}$ 是从 u_j 开始的 $\min^{u_j}$ 的一次出现，而 V 是一个有限集。

现在，假设存在 $i_0\geq 1$，使得 $\tau(u_{i_0})=2$；不失一般性，我们可以假设 i_0 是具有此性质的最小整数。我们只需证明，如果 $i\geq i_0$，那么 $\tau(u_i)=2$。由于 $\tau(u_{i_0})=2$，根据引理 2，$\min^{u_{i_0}}=(\lambda(u_{i_0}))^\omega$。由于 $(u_j)_{j\geq i_0}$ 是从 i_0开始的 $\min^{u_{i_0}}$ 的一次出现，因此对于每个 $j\geq i_0$，都有 $\lambda(u_j)=\lambda(u_{i_0})$，所以对于每个 $i\geq i_0$，$\min^{u_i}=(\lambda(u_i))^\omega$，其中 $(u_j)_{j\geq i}$ 是从 u_i开始的 $\min^{u_i}$ 的一次出现。根据引理 2，我们得出，对于每个 $i\geq i_0$，$\tau(u_i)=2$。◀

引理20 的陈述。设 $G=(V,E)$ 是一个图。设 $u\in V$ 使得 $\tau(u)=3$。设 $(u_i)_{i\geq 1}$ 是 $\min_u$ 的一次出现，起始于 u。设 $(u'_i)_{i\geq 1}$ 是 V 中按如下方式得到的节点无限序列。考虑 $L=\{k\geq 1\mid \tau(u_k)=3\}$，并且对于每个 $i\geq 1$，若存在，则令 $j_i\geq 1$ 为 L 中最小的 i^{th} 个元素。对于每个使得 j_i 有定义的 $i\geq 1$，令 $u'_i=u_{j_i}$；并且若 $i\geq 1$ 使得 j_i 无定义（因此 L 是有限集），则令 $u'_i=u'_{|L|}$。那么，$\bar{}$ $\bar{}$ $\bar{}$$(\bar{u}'_i)_{i\geq 1}$ 是 $G=(V,E)$ 中 $\min_{\bar{u}}$ 的一次出现，起始于 $\bar{u}$。

$\bar{}$**证明**。首先，注意到 $\bar{u}'_i\in V$ 对于每个 $i\geq 1$ 因为 $\tau(u'_i)=3$，且 $u'_1=u_1=u$（因此 $\bar{u}'_1=\bar{u}_1=\bar{u}$）。现在，我们来证明 $(\bar{u}'_{i+1},\ \bar{u}'_i)\in E$ 对于每个 $i\geq 1$。固定 $i\geq 1$。我们分三种情况讨论。
1. j_i 和 j_{i+1} 都有定义。这意味着 $\tau(u'_i)=\tau(u'_{i+1})=3$，并且对于每个 $j_i<k<j_{i+1}$，$\tau(u_k)\neq 3$。根据引理 19，对于每个 $j_i<k<j_{i+1}$，必有 $\tau(u_k)=1$。由于 $(u_k)_{k\geq j_i}$ 是从 u'_i 开始的 $\min^{u'_i}$ 的一次出现，根据引理16，我们得到 $\mathtt{t}_{u'_i}=3$ 以及 $u'_{i+1}\in G_{\ell_{u'_i}}(u'_i)$，因此 $(\bar{u}'_{i+1},\bar{u}'_i)\in E$。2. j_i 有定义，而 j_{i+1} 没有定义。这意味着 $i=|L|$，$\tau(u_{j_{|L|}})=3$，并且根据引理 19，存在 $h>j_{|L|}$ 使得 $\tau(u_h)=2$，且对于每个 $j_{|L|}<k<h$，$\tau(u_k)=1$。由于 $(u_k)_{k\geq j_{|L|}}$ 是从 $u'_{|L|}$开始的 $\min^{u'_{|L|}}$ 的一次出现，根据引理16，我们得到 $\mathtt{t}_{u'_{|L|}}=2$，$\bar{}$，因此 $(\bar{u}'_{i+1},\bar{u}'_i)=(\bar{u}'_{|L|},\bar{u}'_{|L|})\in E$。3. j_i 和 j_{i+1} 都没有定义。于是，根据前一种情况，我们可以得出 $(\bar{u}'_{i+1},\bar{u}'_i)=\bar{}(\bar{u}'_{|L|},\bar{u}'_{|L|})\in E$。

接下来要证明，对于每个 $i\geq 1$，都有 $\min_{\bar{u}}[i]=\lambda(\bar{u}'_i)$。设 $\alpha\in(\Sigma')^\omega$ 满足：对于每个 $i\geq 1$，都有 $\bar{}\alpha[i]=\lambda(\bar{u}'_i)$。我们必须证明 $\alpha=\min_{\bar{u}}$。由于对于每个 $i\geq 1$，都有 $(\bar{u}'_{i+1},\bar{u}'_i)\in E$，因此我们有 $\alpha\in I_{\bar{u}}$，并且 $(\bar{u}'_i)_{i\geq 1}$ 是一个从 $\bar{u}$ 开始的 α 的出现。只要证明对于每个 $\beta\in I_{\bar{u}}$，都有 $\alpha\preceq'\beta$，即可得出结论。固定 $\beta\in I_{\bar{u}}$；只需证明，对于每个 $k\geq 1$，若 $\alpha[1,k-1]=\beta[1,k-1]$，则 $\alpha[k]\preceq'\beta[k]$。固定 $k\geq 1$，并设 $(\bar{v}_i)_{i\geq 1}$ 是一个从 $\bar{u}$ 开始的 β 的出现。注意，对于每个 $1\leq h\leq k-1$，都有 $\alpha[h]=\beta[h]$，等价地，$\lambda(\bar{u}'_h)=\lambda(\bar{v}_h)$，又等价地，$\gamma_{u'_h}=\gamma_{v_h}$（因此 $\ell_{u'_h}=\ell_{v_h}$）且 $\mathtt{t}_{u'_h}=\mathtt{t}_{v_h}$。注意，$(\bar{v}_{i+1},\bar{v}_i)\in E$ 对于每个 $i\geq 1$ 都成立。我们分两种情况讨论：
1. 存在 $1\leq h\leq k-1$ 使得 $\mathtt{t}_{u'_h}=\mathtt{t}_{v_h}=2$。在此情况下，$u'_h$意味着 $u'_h=u'_{h+1}=u'_{h+2}=\ldots$，特别是 $u'_h=u'_k$。此外，由于 $\mathtt{t}_{v_h}=2$

和 $(\bar{v}_{i+1}, \bar{v}_i) \in E$，$G = (V, E)$ 的定义意味着 $v_h = v_{h+1} = v_{h+2} = \dots$，尤其是 $v_h = v_k$。由此得出 $\alpha[k] = \lambda(u'_k) = \lambda(u'_h) = \lambda(v_h) = \lambda(v_k) = \beta[k]$。

2. 对于每个 $1 \le h \le k-1$，都有 $\mathtt{t}_{u'_h} = \mathtt{t}_{v_h} = 3$。在这种情况下，对于每个 $1 \le h \le k-1$，我们有 $\mathtt{t}_{v_h} = 3$，且 $(\bar{v}_{h+1}, \bar{v}_h) \in E$，因此 $v_{h+1} \in G_{\ell_{v_h}}(v_h)$；根据引理16，存在一个 $\min_{v_h}[1, \ell_{v_h} - 1] = \gamma_{v_h}[1, \ell_{v_h} - 1] = \gamma_{u'_h}[1, \ell_{u'_h} - 1] = \min_{u'_h}[1, \ell_{u'_h} - 1]$ 的出现，它从 v_h 开始，到 v_{h+1} 结束。因此，存在一个 $\min_{u'_1}[1, \ell_{u'_1} - 1] \min_{u'_2}[1, \ell_{u'_2} - 1] \dots \min_{u'_{k-1}}[1, \ell_{u'_{k-1}} - 1] = \min_u[1, j_k - 1]$ 的出现，它从u开始，到 v_k结束；于是，若 $\beta' = \min_u[1, j_k - 1] \min_{v_k}$，则 $\beta' \in I_u$。这意味着 $\min_u \preceq \beta'$，因此由 $\min_u = \min_u[1, j_k - 1] \min_{u'_k}$ 可得 $\min_{u'_k} \preceq \min_{v_k}$。根据引理 17 以及 $\min_{u'_k} \preceq \min_{v_k}$，以下陈述必有一个成立：

a. $\gamma_{u'_k}$ 不是 γ_{v_k}的前缀，γ_{v_k} 不是$\gamma_{u'_k}$ 的前缀，并且$\gamma_{u'_k} \prec \gamma_{v_k}$。
b. $\gamma_{u'_k} = \gamma_{v_k}$ 和 $\mathtt{t}_{u'_k} \le \mathtt{t}_{v_k}$。 **c.** γ_{v_k} 是$\gamma_{u'_k}$的真前缀。在这三种情况下，我们得出$(\gamma_{u'_k}, \mathtt{t}_{u'_k}) \preceq' (\gamma_{v_k}, \mathtt{t}_{v_k})$，等价地，$\lambda(u'_k) \preceq' \lambda(v_k)$，等价地，$\alpha[k] \preceq' \beta[k]$。

◀

定理 21 的陈述。设 $G = (V, E)$ 是一个图。设 $u, v \in V$满足条件$\tau(u) = \tau(v) = 3$。

1. 若$\min_u = \min_v$，则$\min_{\bar{u}} = \min_{\bar{v}}$。
2. 若$\min_u \prec \min_v$，则$\min_{\bar{u}} \prec' \min_{\bar{v}}$。

证明。设 $(u_i)_{i \ge 1}$ 是从 u 开始的 $\min_u$ 的一个出现，并设 $(v_i)_{i \ge 1}$ 是从 v 开始的 $\min_v$ 的一个出现。设 $(\bar{u}'_i)_{i \ge 1}$ 是根据引理20中的 $(u_i)_{i \ge 1}$ 定义的 $\min_{\bar{u}}$ 的出现，并设 $(\bar{v}'_i)_{i \ge 1}$ 是根据引理20中的 $(v_i)_{i \ge 1}$ 定义的 $\min_{\bar{v}}$ 的出现。令 $L = \{i \ge 1 \mid \tau(u_i) = 3\}$，若 L 存在，则令 $j_i \ge 1$ 为 L 的最小元素。此外，令 $M = \{i \ge 1 \mid \tau(v_i) = 3\}$，若 K 存在，则令 $k_i \ge 1$ 为 K 的最小元素。注意到 $j_1 = k_1 = 1$。

在证明的其余部分，如果整数 $h \ge 1$ 满足以下性质，我们就称其为良好的：

- $j_1, \dots, j_h$和 $k_1, \dots, k_h$均已定义，且 $j_i = k_i$对每个 $1 \le i \le h$ 都成立。
- $\gamma_{u'_i} = \gamma_{v'_i}$对每个 $1 \le i \le h-1$均成立。

请注意以下性质：

- 1 总是良好的，因为 $j_1 = k_1 = 1$。如果 h 是良好的，那么每个 $1 \le h' \le h$ 都是良好的。
- 这意味着，要么每个 $h \ge 1$ 都是良好的，要么存在唯一的 $h^* \ge 1$，使得每个 $h \le h^*$ 都是良好的，而每个 $h > h^*$ 都不是良好的。如果 h 是良好的，且 $\gamma_{u'_h} = \gamma_{v'_h}$ 和 $\mathtt{t}_{u'_h} = \mathtt{t}_{v'_h} = 3$，那么 $h+1$ 是良好的。事实上，$\gamma_{u'_h} = \gamma_{v'_h}$意味着 $\ell_{u'_h} = \ell_{v'_h}$。此外，由于 $(u_i)_{i \ge j_h}$ 是从 u'_h开始的 $\min_{u'_h}$ 的一次出现，因此根据 $\ell_{u'_h}$ 的极小性和引理16，对于每个 $1 \le i \le \ell_{u'_h} - 2$，都有 $\tau(u_{j_h+i}) = \tau(G_{i+1}(u'_h)) = 1$，且 $\tau(u_{j_h+\ell_{u'_h}-1}) = \tau(G_{\ell_{u'_h}}(u'_h)) = \mathtt{t}_{u'_h} = 3$，这意味着$j_{h+1}$ 有定义。类似地，可以得到，对于每个 $1 \le i \le \ell_{v'_h} - 2$，都有 $\tau(u_{k_h+i}) = 1$，且$\tau(v_{k_h+\ell_{v'_h}-1}) = 3$，因此 k_{h+1} 有定义，且 $j_{h+1} = k_{h+1}$，这意味着 $h+1$ 是良好的。如果 h 是良好的，那么对于每个 $1 \le i \le h-1$，都有 $\mathtt{t}_{u'_i} = \mathtt{t}_{v'_i} = 3$。事实上，反设对于某个 $1 \le l \le h-1$，$\mathtt{t}_{u'_l} = 2$ 成立（$\mathtt{t}_{v'_l} = 2$ 的情形类似）。由于 $(u_i)_{i \ge j_l}$ 是从 u'_l开始的 $\min_{u'_l}$ 的一次出现，因此根据引理16，对于每个 $1 \le i \le \ell_{u'_l} - 2$，都有 $\tau(u_{j_l+i}) = \tau(G_{i+1}(u'_l)) = 1$，且 $\tau(u_{j_l+\ell_{u'_l}-1}) = \tau(G_{\ell_{u'_l}}(u'_l)) = \mathtt{t}_{u'_l} = 2$；由引理19可知，对于每个 $i \ge j_l + \ell_{u'_l} - 1$，都有 $\tau(u_i) = 2$。这意味着 j_{l+1}无定义，这与 $1 \le l \le h-1$矛盾。

24 前缀排序确定有限自动机：一种递归算法

下面证明，若 $h \geq 1$ 是良好的，则：

- $\min_u[1, j_h - 1] = \min_v[1, k_h - 1]$.
- $\min_{\bar{u}}[1, h-1] = \min_{\bar{v}}[1, h-1]$.

我们将分别证明这两个性质。

- 让我们证明 $\min_u[1, j_h - 1] = \min_v[1, j_h - 1]$。由于 h 是良好的，因此对于每个 $1 \leq i \leq h$，j_i 和 k_i 都有定义，且 $j_i = k_i$。因此，只需证明对于每个 $1 \leq i \leq h-1$，$\min_u[j_i, j_{i+1} - 1] = \min_v[k_i, k_{i+1} - 1]$。这等价于证明对于每个 $1 \leq i \leq h-1$，$\min_{u'_i}[1, \ell_{u'_i} - 1] = \min_{v'_i}[1, \ell_{v'_i} - 1]$。这可由对每个 $1 \leq i \leq h-1$成立的 $\gamma_{u'_i} = \gamma_{v'_i}$ 得出。

- 让我们证明 $\min_{\bar{u}}[1, h-1] = \min_{\bar{v}}[1, h-1]$。固定 $1 \leq i \leq h-1$。我们必须证明 $\min_{\bar{u}}[i] = \min_{\bar{v}}[i]$，或者等价地，$\lambda(\bar{u}'_i) = \lambda(\bar{v}'_i)$。这由 $\gamma(u'_i) = \gamma(v'_i)$ 和 $\mathtt{t}_{u'_i} = \mathtt{t}_{v'_i}$ 得出。

现在我们可以证明该定理的两个结论。我们将使用以下观察结果：如果每个 $h \geq 1$ 都是良好的，那么 $\min_u = \min_v$ 且 $\min_{\bar{u}} = \min_{\bar{v}}$，因为对于每个 $h \geq 1$，都有 $\min_u[1, j_h - 1] = \min_v[1, k_h - 1]$ 和 $\min_{\bar{u}}[1, h-1] = \min_{\bar{v}}[1, h-1]$。

1. 假设 $\min_u = \min_v$；我们必须证明 $\min_{\bar{u}} = \min_{\bar{v}}$。如果每个 $h \geq 1$ 都是良好的，我们就完成了，因此可以假设 $h^* \geq 1$ 是最大的良好整数。根据引理 12，对于每个 $i \geq 1$，都有 $\tau(u_i) = \tau(v_i)$，因此 $L = M$。特别地，对于每个 $i \geq 1$，j_i 有定义，当且仅当 k_i 有定义；若有定义，则 $j_i = k_i$。由于 h^* 是良好的，因此 j_{h^*} 和 k_{h^*} 有定义，$j_{h^*} = k_{h^*}$，且 $\min_{\bar{u}}[1, h^* - 1] = \min_{\bar{v}}[1, h^* - 1]$。我们知道 $\min_{\bar{u}} = \min_{\bar{u}}[1, h^* - 1] \min^{\bar{u}'_{h^*}}$ 且 $\min_{\bar{v}} = \min_{\bar{v}}[1, h^* - 1] \min^{\bar{v}'_{h^*}}$，所以 $\min_{\bar{u}}[1, h^* - 1] = \min_{\bar{v}}[1, h^* - 1]$ 意味着，为了证明 $\min_{\bar{u}} = \min_{\bar{v}}$，我们只需证明 $\min^{\bar{u}'_{h^*}} = \min^{\bar{v}'_{h^*}}$。

根据引理 12，我们有 $\min^{u'_{h^*}} = \min^{v'_{h^*}}$。由于对于每个 $i \geq 1$，$\tau(u_i) = \tau(v_i)$，$(u_i)^{i \geq j_{h^*}}$ 是从 u'_{h^*} 开始的 $\min^{u'_{h^*}}$ 的一次出现，而 $(v_i)^{i \geq j_{h^*}}$ 是从 v'_{h^*} 开始的 $\min^{v'_{h^*}}$ 的一次出现，因此我们得到 $\ell_{u'_{h^*}} = \ell_{v'_{h^*}}$。因此，$\gamma_{u'_{h^*}} = \min_{u'_{h^*}}[1, \ell_{u'_{h^*}}] = \min_{v'_{h^*}}[1, \ell_{v'_{h^*}}] = \gamma_{v'_{h^*}}$。特别地，它必须是 $\mathtt{t}_{u'_{h^*}} = \mathtt{t}_{v'_{h^*}}$，所以由 $\gamma_{u'_{h^*}} = \gamma(v'_{h^*})$ 我们得出 $\lambda(\bar{u}'_{h^*}) = \lambda(\bar{v}'_{h^*})$。此外，它必须是 $\mathtt{t}_{u'_{h^*}} = \mathtt{t}_{v'_{h^*}} = 2$，否则我们将得出 $h^* + 1$ 是良好的，这与 h^* 的极大性相矛盾。因此，根据 $G = (V, E)$ 的定义，我们得出 $\min^{\bar{u}'_{h^*}} = (\lambda(\bar{u}'_{h^*}))^\omega = \lambda(\bar{v}'_{h^*}))^\omega = \min^{\bar{v}'_{h^*}}$。

2. 假设 $\min_u \prec \min_v$；我们必须证明 $\min_{\bar{u}} \prec' \min_{\bar{v}}$。注意，不可能所有 $h \geq 1$ 都是良好的，否则我们将得到 $\min_u = \min_v$，这与矛盾相矛盾。令 $h^* \geq 1$ 为最大的良好整数。特别地，j_{h^*} 和 k_{h^*} 是有定义的，$j_{h^*} = k_{h^*}$，$\min_u[1, j_{h^*} - 1] = \min_v[1, k_{h^*} - 1]$ 且 $\min_{\bar{u}}[1, h^* - 1] = \min_{\bar{v}}[1, h^* - 1]$。我们知道 $\min_u = \min_u[1, j_{h^*} - 1] \min^{u'_{h^*}}$ 且 $\min_v = \min_v[1, k_{h^*} - 1] \min^{v'_{h^*}}$，因此由 $\min_u \prec \min_v$ 可得 $\min^{u'_{h^*}} \prec \min^{v'_{h^*}}$。由于 $\min_{\bar{u}} = \min_{\bar{u}}[1, h^* - 1] \min^{\bar{u}'_{h^*}}$，$\min_{\bar{v}} = \min_{\bar{v}}[1, h^* - 1] \min^{\bar{v}'_{h^*}}$ 且 $\min_{\bar{u}}[1, h^* - 1] = \min_{\bar{v}}[1, h^* - 1]$，为了证明 $\min_{\bar{u}} \prec' \min_{\bar{v}}$，我们只需证明 $\min_{\bar{u}'_{h^*}} \prec' \min^{\bar{v}'_{h^*}}$。

注意，它不可能是 $\gamma_{u'_{h^*}} = \gamma_{v'_{h^*}}$ and $\tau(u'_{h^*}) = \tau(v'_{h^*})$，因为：

a. 如果 $\tau(u'_{h^*}) = \tau(v'_{h^*}) = 2$，那么根据引理18，我们可以得出 $\min^{u'_{h^*}} = \min^{v'_{h^*}}$，这就产生了矛盾。

b. 如果 $\tau(u'_{h^*}) = \tau(v'_{h^*}) = 3$，那么 $h^* + 1$ 将是一个较好的整数，这与 h^* 的极大性矛盾。

由此观察、引理 17 以及 $\min^{u'_{h^*}} \prec \min^{v'_{h^*}}$，我们可以得出以下结论之一必然成立：

a. $\gamma_{u'_{h^*}}$ 不是 $\gamma_{v'_{h^*}}$ 的前缀，$\gamma_{v'_{h^*}}$ 不是 $\gamma_{u'_{h^*}}$ 和 $\gamma_{u'_{h^*}} \prec \gamma_{v'_{h^*}}$ 的前缀。b. $\gamma_{u'_{h^*}} = \gamma_{v'_{h^*}}$、$\mathtt{t}_{u'_{h^*}} = 2$ 和 $\mathtt{t}_{v'_{h^*}} = 3$。c. $\gamma_{u'_{h^*}}$ 是 $\gamma_{v'_{h^*}}$ 的真前缀。

在这三种情况下，我们得出 $(\gamma_{u'_{h^*}}, \mathtt{t}_{u'_{h^*}}) \prec' (\gamma_{v'_{h^*}}, \mathtt{t}_{v'_{h^*}})$，或者等价地，$\lambda(\bar{u}'_{h^*}) \prec' \lambda(\bar{v}'_{h^*})$，这意味着 $\min_{\bar{u}'_{h^*}} \prec' \min_{\bar{v}'_{h^*}}$。

◀

引理23的陈述。设 $G = (V, E)$ 是一个修剪图。那么，我们可以在 $O(|V|^2)$ 时间内构造 $\bar{G} = (\bar{V}, \bar{E})$。

证明。的定义 G 表明，为了构造 G，只需计算 γ_u、$\mathtt{t}_u$以及 $G_{\ell_u}(u)$ 对于每个 $u \in V$ 满足 $\tau(u) = 3$；此外，我们还需要计算全序 $\preceq'$，以便根据全序将 Σ' 一致地重新映射到整数字母表$\preceq'$ 上的 Σ'。

固定 $u \in V$ 使得 $\tau(u) = 3$。对于每个 $1 \le i \le \ell_u - 1$，令 $E_i(u)$ 为所有进入 $G_i(u)$ 中某个节点的边的集合，并令 $m_i(u) = |E_i(i)|$。注意，由于 $G_i(u)$'s 两两不同，因此 $E_u(i)$'s 两两不相交。下面证明 $\sum_{i=1}^{\ell_u - 1} m_i(u) \in O(n)$。为此，只需证明，对于每个 $v \in V$，$\bigcup_{i=1}^{\ell_u - 1} E_i(u)$ 中至多存在两条边，其起始节点等于 v。固定 $v \in V$，令 $1 \le i \le \ell_u - 1$ 为满足以下条件的最小整数：v 是某条边 $(v, v') \in E_i(u)$ 的第一个状态，其中 $v' \in G_i(u)$；如果这样的 v' 存在。注意，v' 是唯一确定的，因为值 $\lambda(z)$ 不依赖于 $z \in G_i(u)$ 的选择，并且 G 是确定性的。这意味着 $(v, v') \in E_i(u)$ 是唯一确定的。为了证明我们的断言，只需说明：如果对于某个 $i + 1 \le j \le \ell_u - 1$ 以及 $v'' \in G_j(u)$，有 $(v, v'') \in E_j(u)$，那么 $\lambda(v'') = \lambda(v)$；因为这样我们就能得出 j 和 v'' 是唯一确定的，这是由于 G 是确定性的。由于 $v' \in G_i(u)$ 且 $(v, v') \in E_i(u)$，根据 $G_{i+1}(u)$ 的定义，我们有 $\lambda(G_{i+1}(u)) \preceq \lambda(v)$。由于 $2 \le i + 1 \le j \le \ell_u - 1$，由引理15可得 $\min_u[j] \preceq \min_u[i + 1]$，根据引理16，这等价于 $\lambda(G_j(u)) \preceq \lambda(G_{i+1}(u))$，因此我们得到 $\lambda(G_j(u)) \preceq \lambda(v)$，或者等价地，$\lambda(v'') \preceq \lambda(v)$。由于 $v'' \in G_j(u)$ 且 $2 \le j \le \ell_u - 1$，因此 $\tau(v'') = \tau(G_j(u)) = 1$，所以由 $(v, v'') \in E$ 我们得出必有 $\lambda(v) \preceq \lambda(v'')$，因为 G 是修剪的，因此 $\lambda(v'') = \lambda(v)$。

下面证明，对于每个满足 $\tau(u) = 3$ 的 $u \in V$，我们可以在 $O(|V|^2)$ 中计算 $G_{\ell_u}(u)$、γ_u 和 $\mathtt{t}_u$。只需证明，对于一个固定的、满足 $\tau(u) = 3$ 的 $u \in V$，我们可以在 $O(|V|)$ 时间内计算 $G_{\ell_u}(u)$、γ_u 和 $\mathtt{t}_u$。

下面说明如何递归计算每个集合 $G_i(u)$，其中 $1 \le i \le \ell_u$。在算法执行过程中，我们会标记一些节点。初始时，没有节点被标记；执行步骤 $i \ge 1$ 后，V 中的一个节点被标记，当且仅当它属于 $\bigcup_{j=2}^{i-1} G_j(u)$。如果 $i = 1$，我们只需令 $G_1(u) = \{u\}$，并定义 $c_1 = \lambda(u)$。现在，假设 $i \ge 2$，且我们已经计算出 $G_{i-1}(u)$。首先扫描 $E_{i-1}(u)$ 中的所有边，并收集集合 R_i，即这些边的所有起始节点组成的集合。然后，通过扫描 R_i，我们先确定 $c_i = \min\{\lambda(v) \mid v \in R_i\}$，再确定 $R'_i = \{v \in R_i \mid \lambda(v) = c_i\}$。接下来，确定 $t_i = \min\{\tau(v) \mid v \in R'_i\}$ 和 $R''_i = \{v \in R'_i \mid \lambda(v) = t_i\}$。通过检查 R''_i 中哪些节点已被标记，我们确定 $G_i(u)$。最后，标记 $G_i(u)$ 中的所有节点。所有这些步骤都可以在 $O(m_{i-1}(u))$ 时间内完成，其中渐近表示法中的隐藏常数不依赖于 i。现在，我们检查 $t_i \ge 2$ 是否成立。如果 $t_i \ge 2$，那么 $i = \ell_u$，至此完成。否则，我们有 $i < \ell_u$，并继续执行步骤 $i + 1$。

我们得出，可以在时间 $\sum_{i=1}^{\ell_u - 1} O(m_i(u)) = O(\sum_{i=1}^{\ell_u - 1} m_i(u)) = O(n)$内确定 $G_{\ell_u}(u)$；这里我们利用了 $O(m_i(u))$中的隐藏常数与 i无关这一点。此外，我们有 $\gamma_u = c_1 c_2 \dots c_{\ell_u}$ 和 $\mathtt{t}_u = t_{\ell_u}$。

现在只需说明如何计算 $\preceq'$。本质上，我们只需对字符串 $\gamma(u)$'s 进行基数排序，同时注意到 $\preceq'$ 是通过考虑字典序的一种细微变体定义的。更确切地说，我们按如下方式进行。我们知道，对于满足 $\tau(u) = 3$ 的每个 $u \in V$，都有 $|\gamma(u)| \le |V| + 1$，因此首先在每个 $\gamma(u)$ 的末尾填充一种特殊的

字符，该字符大于字母表中的所有字符，直至每个字符串的长度恰好为$|V|+1$。接下来，我们考虑两个额外的特殊字符 d_2 和 d_3，满足 $d_2 \prec' d_3$，并将这两个字符中的恰好一个追加到每个 $\gamma(u)$：如果 $\mathtt{t}_u = 2$，则追加 d_2；如果 $\mathtt{t}_u = 3$，则追加d_3。现在，我们在 $O(|V|^2)$ 时间内对（修改后的） $\gamma(u)$'s 进行基数排序，从而得到 $\preceq'$。◀

B.3 第4.3节的证明

引理 24 的陈述。设 $G = (V, E)$ 是一张图。如果我们知道$\{u \in V \mid \tau(u) = 3\}$的最小划分，那么就能在 $O(|E|)$ 时间内构造 V 的最小划分。

证明。我们已经看到，对于每个 $c \in \Sigma$，算法都能正确构造$\mathcal{A}_{c,3}$和$\mathcal{A}_{c,2}$。注意，当算法构造$\mathcal{A}_{c,1}$时，它从不修改$A_{c,2}$和$A_{c,3}$（因为我们只向$A_{c,1}$中添加元素）。

我们来证明，当考虑 I 在 $A_{c_i,t}$ 中时（且在计算 J_k 之前），以下结论成立：**1.** I 是 $\mathcal{A}$ 的一个元素，并且我们已经构造了 $\mathcal{A}$ 中最大元素为 I 的前缀。**2.** 每个 $A_{c,1}$ 都包含 $\mathcal{A}_{c,1}$ 的一个前缀。

在算法开始时，我们考虑 I 属于 $A_{c_1,t}$，其中 $t \in \{2, 3\}$，因此我们的断言成立，因为我们已经构造了 $\mathcal{A}_{c_1,2}$ 和 $\mathcal{A}_{c_1,3}$，而所有 $A_{c,1}$ 都为空。

现在，假设考虑 $I \in \mathcal{A}_{c_i,t}$ 时，我们的断言成立。我们要证明，处理下一个元素（如果存在）时，断言仍然成立。考虑 I 时，我们计算非空的 J_k，并将每个这样的 J_k 添加到 $A_{c_k,1}$ 中。现在我们要证明，J_k 被正确识别为 $A_{c_k,1}$ 中的下一个元素。

- 首先，我们证明，如果 $v_1, v_2 \in J_k$，那么 $\min_{v_1} = \min_{v_2}$。由于 $v_1, v_2 \in J_k$，存在 $u_1, u_2 \in I$，使得 $(u_1, v_1), (u_2, v_2) \in E$。根据归纳假设，$I$ 是 $\mathcal{A}$ 的一个元素，而且我们已经构造了 $\mathcal{A}$ 中最大元素为 I 的前缀；又因为 v_1 和 v_2 之前没有被标记，所以必有 $\min_{v_1} = c_k\min_{u_1}$、$\min_{v_2} = c_k\min_{u_2}$ 且 $\min_{u_1} = \min_{u_2}$，因此得出 $\min_{v_1} = \min_{v_2}$。
- 我们只需证明：如果在将 J_k 添加到 $A_{c_k,1}$ 后，$v_1 \in J_k$ and $v_2 \in V_{c_k,1}$ 不属于 $A_{c_k,1}$ 的某个元素，则 $\min_{v_1} \prec \min_{v_2}$。令 $u_1 \in I$ 使得 $(u_1, u_2) \in E$；我们已经证明 $\min_{v_1} = c_k \min_{u_1}$。根据 v_2 的定义，对于某个我们尚未处理的 u_2，有 $\min_{v_2} = c_k \min_{u_2}$。由于根据归纳假设，$I$ 是 $\mathcal{A}$ 的一个元素，并且我们已经构建了 $\mathcal{A}$ 的前缀，其最大元素为 I，因此必有 $\min_{u_1} \prec \min_{u_2}$，从而得到 $\min_{v_1} \prec \min_{v_2}$。

现在，让我们证明：如果在考察了 I in $A_{c_i,t}$ 后，下一个要考察的元素不在 $A_{c_i,t}$ 中，那么 $\mathrm{A}_{c_i,t}$ 的构造就完成了。如果 $t \in \{2, 3\}$，那么结论显然成立，因为 $\mathrm{A}_{c_i,2}$ 和 $\mathrm{A}_{c_i,3}$ 都已经完全构造出来了。现在假设 $t = 1$。反设存在尚未被添加到 $A_{c_i,1}$ 中某个元素的 $v \in V_{c_i,1}$。不失一般性地，我们选择 v，使得 $\min_v$ 尽可能小。由于 $\tau(v) = 1$，因此存在 $u \in V$，使得 $(u, v) \in E$ 且 $\min_u \prec \min_v$（从而 $\lambda(u) \preceq \lambda(v) = c_i$）。如果 $\lambda(u) \prec c_i$，那么我们立即得到 u 已经被处理过（因为我们已经构造出了 A 的前缀，其最大元素为 I），因此在处理 u 时，v 应该已经被处理过了，这就导致了矛盾。如果 $\lambda(u) = c_i$，那么根据推论 3，我们有 $\tau(u) \leq \tau(v)$，所以 $\tau(u) = 1$ 和 $u \in V_{c_i,1}$；v 的极小性意味着 u 之前已经被添加到了 $A_{c_i,1}$ 的某个元素中，因此 v 应该已经被处理过了，再次导致矛盾。

因此，当我们处理 $A_{c_i,t}$ 中 I 之后的下一个元素时，我们的断言将成立（无论下一个元素是否在 $A_{c_i,t}$ 中）。特别地，如果没有下一个元素，我们就可以得出结论：所有 $\mathcal{A}_{c,1}$ 都已构建完成，

最后，我们可以在 $O(|E|)$ 时间内构建每个 $\mathcal{A}_{c,\,1}$，因为我们只需要扫描每个节点和每条边一次。

◀

C 补情形的详细说明

下面说明如何将计算$\{u \in V \mid \tau(u) = 1\}$的最小划分问题，化简为确定 V 的最小划分问题，其中 $G = (V, E)$是一个图，满足$|V| = \{u \in V \mid \tau(u) = 1\}$。我们采用与前几节相同的记号（但含义互补），以便于理解我们的论证。

C.1 递归步骤

对于每个满足 $\tau(u) = 1$的 $u \in V$，令 ℓ_u为满足 $\tau(u_k) \leq 2$的最小整数 $k \geq 2$，其中$(u_i)_{i \geq 1}$ 是一个从 u开始的 $\min_u$ 出现。对于每个 $1 \leq i \leq \ell_u$，我们可以像之前那样定义每个 $G_i(u)$；现在，对于每个 $2 \leq i \leq \ell_u - 1$，有 $\tau(G_i(u)) = 3$。事实上，当我们以反向方式从 u 开始探索图，并（隐式地）构建 $\min_u$的一个前缀时，在每一步中，我们仍然必须考虑值 $\lambda(v)$最小的节点，并在这些节点中考虑值 $\tau(v)$最小的节点。这一次，我们不可能将先前已纳入的节点再次纳入 $G_i(u)$：在每一步中，$\lambda(G_i(u))$只会增大（因为 $\tau(G_i(u)) = 3$），因此，如果存在一个环，那么环上的所有边都将具有相同的标签，由引理 9 可知，对某个$2 \leq i \leq \ell_u - 1$，$\tau(G_i(u)) \leq 2$，这就导致了矛盾。特别是，这意味着现在我们甚至不需要设法修剪图来确保算法在 $O(|V|^2)$时间内运行，因为我们不可能两次访问同一个节点。γ_u和 $\mathtt{t}_u$的定义与之前相同，而引理17仍然意味着 Σ' 和 $\preceq'$ 必须按之前的方式定义。当我们定义 $G = (V, E)$时，我们定义 $\bar{V} = \{\bar{u} \mid \tau(u) = 1\}$，并按之前的方式定义 $\bar{E}$。不难验证，我们可以像之前那样在 $O(|V|^2)$时间内构造 $\bar{G}$；如果我们有 $\bar{V}$ 的最小划分（相对于 $\bar{G}$），那么也就有$\{u \in V \mid \tau(u) = 1\}$的最小划分（相对于 G）。

C.2 合并

设 $u \in V$ 满足 $\tau(u) = 3$。由引理 2，存在 $k \geq 1$、$c \in \Sigma$ 和 $\gamma' \in \Sigma^\omega$，使得 $\min_u = \lambda(u)^k c\gamma'$ 且 $\lambda(u) \prec c$。然后，我们定义 $\psi(u) = k$。

下文中，我们将经常使用以下观察。设 $v \in V$ 满足 $\tau(v) = 3$。设 $u \in V$ 满足 $(u, v) \in E$ 且 $\min_v = \lambda(v) \min_u$。那么，(i) $\lambda(v) \preceq \lambda(u)$；且 (ii) 若$\lambda(v) = \lambda(u)$，则 $\tau(u) = 3$ 且 $\psi(v) = \psi(u) + 1$。

注意，若 $u, v \in V$ 满足 $\tau(u) = \tau(v) = 3$、$\lambda(u) = \lambda(v)$ 且 $\psi(u) \neq \psi(v)$，则 $\min_u \prec \min_v$ 当且仅当 $\psi(v) < \psi(u)$。

在 $O(|E|)$ 时间内计算所有 $\psi(u)$'s 是很容易的。从每个满足 $\tau(u) = 3$ 的 $u \in V$ 开始，以反向方式遍历图，只沿着标记为 $\lambda(u)$ 的边前进，直到遇到一个其 $\psi(u)$ 已经确定的节点，或者无法再继续遍历图为止。对于遇到的所有节点 u'，必须有 $\tau(u') = 3$（否则它将是 $\tau(u) \neq 3$），而如果从遇到的某个 u' 出发无法再继续遍历，则必须有 $\psi(u') = 1$。我们不可能两次遇到同一个节点，因为 G 是确定性的；也不可能存在环，否则根据引理 9，它将是 $\tau(u) \neq 3$。因此，我们构建了一棵树（其根节点 u 没有出边），而计算 $\psi(u)$'s 等价于计算每个节点的高度。

28 前缀排序确定性有限自动机：递归算法

让我们说明如何利用 $\psi(u)$ 来确定 V 的最小划分 $\mathcal{A}$，假设我们已经有 $\{u \in V \mid \tau(u) = 1\}$ 的最小划分 $\mathcal{B}$。和之前一样，我们有 $\{u \in$ 顶点集 $\mid \tau(u) = 2\}$ 的最小划分 $\mathcal{B}'$。对于每个 $c \in \Sigma$ 以及每个 $t \in \{1, 2, 3\}$，我们像之前一样定义 $V_{c,t}$、$\mathcal{A}_{c,t}$ 和 $A_{c,t}$，于是问题就归结为计算每个 $\mathcal{A}_{c,t}$。

通过使用 $\mathcal{B}$ 和 $\mathcal{B}'$，我们可以像之前一样计算 $\mathcal{A}_{c,\,1}$ 和 $\mathcal{A}_{c,\,2}$，因此具有挑战性的部分是计算 $\mathcal{A}_{c,\,3}$。令 $\Sigma = \{c_1,\ c_2,\ \ldots,\ c_\sigma\}$，其中有 $c_1 \prec c_2 \prec \cdots \prec c_\sigma$。在算法过程中，我们会为每个元素 I 在某个时刻属于某个 $A_{c,\,3}$ 的元素分配一个数值 $\psi(I) \geq 1$。

注意，$A_{c_\sigma,3}$ 必须为空，否则我们将得出结论：c_σ 不是最大字符。这表明，这一次我们处理 $A_{c,t}$ 的顺序必须是 $A_{c_\sigma,2}$、$A_{c_\sigma,1}$、$A_{c_{\sigma-1},3}$、$A_{c_{\sigma-1},2}$、$A_{c_{\sigma-1},1}$、$A_{c_{\sigma-2},3}$，依此类推。此外，我们将按照从最大元素到最小元素的顺序，逐步构建每个 $A_{c,t}$（因此我们将考虑最小划分的后缀，而不是前缀）。这一次我们不会标记节点，而是标记 $A_{c,t}$ 的条目，以表示某个节点已从 $A_{c,t}$ 中的一个元素移除。直观地说，这一次我们需要移除节点，因为当我们处理 I 在 $A_{c_i,t}$中时，我们已经构建了 A 的后缀（而不是前缀），其最大元素为 I；但我们现在正按照从最大元素到最小元素的顺序构建 A，因此，沿着离开 I 中某个节点 u 的边到达的节点，也可能沿着离开尚未处理的某个节点 u' 的边到达，并且该节点满足 $\min_{u'} \prec \min_u$。

假设我们在 $A_{c_i,t}$ 中处理 I。注意，如果 $v \in V$ 满足 $\tau(v) = 3$、$\lambda(v) = c_k$，且对某个 $u \in I$ 有 $(u, v) \in E$，那么它必定是 $k \leq i$，否则 $\tau(v) = 1$；如果 I 是 $A_{c,1}$ 中的元素，或是 $A_{c,2}$ 中的元素，那么它必定始终是 $k < i$，因为如果它是 $k = i$，我们就会得出 $\tau(v) \neq 3$。对每个 $k \leq i$，定义 $\psi_{I,k} = 1$，若 $k < i$；若 $k = i$，则定义 $\psi_{I,k} = \psi(I) + 1$。我们计算集合 J_k，其中包含所有满足以下条件的节点 $v \in V$：$\tau(v) = 3$、$\lambda(v) = c_k$、$\psi(v) = \psi_{I,\,k}$，并且对某个 $u \in I$ 有 $(u, v) \in E$；此外，若 $J_k \neq \emptyset$，则 (i) 标记 $A_{c_k,3}$ 中包含 J_k 中某个元素的条目；(ii) 将 J_k 添加到 $A_{c_k,3}$ 中，并令 $\psi(J_k) = \psi_{I,k}$。我们假设维护一个额外的数组，对于 $\{u \in V \mid \tau(u) = 3\}$ 中已经出现在某个 $A_{c,3}$ 中的每个节点，该数组都存储其当前（唯一）位置。这样，操作 (i) 就能在常数时间内完成，而不会影响算法的运行时间（仍为 $O(|E|)$）。

请注意，在任何时刻，对于每个 $A_{c,\,3}$，以下各点都成立：(a) 如果 K 属于 $A_{c,\,3}$，则 $\psi(K) \geq 1$；(b) 如果 $A_{c,\,3}$ 非空，则我们添加的第一个 K 满足 $\psi(K) = 1$；(c) 如果在 K 之后立即将 K' 添加到 $A_{c,\,3}$ 中，则 $\psi(K) \leq \psi(K') \leq \psi(K) + 1$；(d) 如果对于 $A_{c,\,3}$ 中的某个 K 有 $v \in K$，则 $\psi(v) = \psi(K)$。

让我们证明，当我们考虑 I 在 $A_{c_i,\,t}$（并且在计算 J_k 之前）时，I 是 $\mathcal{A}$ 的一个元素，并且我们已经构造出了 $\mathcal{A}$ 的后缀，其最大元素为 I。

在算法开始时，我们考虑 I 在 $A_{c_\sigma,\,t}$，其中 $t \in \{1,\ 2\}$，因此我们的断言成立，因为我们已经构造出了 $\mathcal{A}_{c_\sigma,\,2}$ 和 $\mathcal{A}_{c_\sigma,\,1}$。

现在，假设当我们考虑 I in $A_{c_i,t}$ 时，命题成立。下面证明，如果存在下一个元素 I' in $A_{c'_i,t'}$，那么我们可以考虑它；此时 I' 是 $\mathcal{A}$ 中的一个元素，并且我们已经构造出了 $\mathcal{A}$ 的一个后缀，其最大元素为 I'。注意，$i' = i$ 或 $i' = i - 1$ 必有其一。

1. 假设 $i' = i$ 且 $t' = t$。如果 $t' = t \neq 3$，那么我们就完成了，因为我们已经构造出了 $\mathcal{A}_{c,2}$'s 和 $\mathcal{A}_{c,1}$'s。现在假设 $t' = t = 3$。这意味着 $\psi(I) + 1 \leq \psi(I') \leq \psi(I)$。

a. 假设 $\psi(I') = \psi(I)$。首先，我们来证明，如果 $v_1, v_2 \in I'$，那么 $\min_{v_1} = \min_{v_2}$。令 u_1, $u_2 \in V$ 满足 (u_1, v_1)，$(u_2, v_2) \in E$，$\min_{v_1} = c_i \min_{u_1}$ 且 $\min_{v_2} = c_i \min_{u_2}$。由于 $\tau(v) = 3$，我们得到以下两种情况之一：(i) $c_i \prec \lambda(u_1)$，或 (ii) $\lambda(u_1) = c_i, \tau(u_1) = 3$ 且 $\psi(u_1) = \psi(I') - 1 = \psi(I) - 1$。在这两种情况下，我们都可以得出 u_1 是

后缀中 A 的一个元素 K，其最大元素为 I。根据归纳假设，该元素已被正确构建，因此 v_1 被添加到 A_{c_i} 的一个元素中,t；否则，v_1 永远不会从该元素中被移除，因为在 I_{v_1} 中，否则就会存在一个小于 $\min_{v_1}$ 的字符串。因此，它也必定是 $u_2 \in K$。根据归纳假设，$\min_{u_1} = \min_{u_2}$，由此我们得出 $\min_{v_1} = \min_{v_2}$。

我们只需证明：如果 v_1 既不在 I' 中，也不在 A 的最大元素为 I 的后缀中，并且 $v_2 \in I'$，那么 $\min_{v_1} \prec \min_{v_2}$。如果 $\tau(v_i) \neq 3$，结论立即成立，因此我们可以假设 $\tau(v_i) = 3$。不可能有 $c_i = \lambda(v_2) \prec \lambda(v_1)$，否则 v_1 将位于 A 的最大元素为 I 的后缀中。由于 $\tau(v_i) = 3$，不可能有 $\lambda(v_1) = c_i$ 且 $\psi(v_1) < \psi(I)$，否则 v_1 同样会位于 A 的最大元素为 I 的后缀中。因此，必有 $\lambda(v_1) \preceq c_i$，并且如果 $\lambda(v_1) = c_i$，那么 $\psi(I') = \psi(I) \leq \psi(v_1)$。如果 $\lambda(v_1) \prec c_i$，或者 $\lambda(v_1) = c_i$ 且 $\psi(I') < \psi(v_1)$，我们立即得到 $\min_{v_1} \prec \min_{v_2}$。因此，我们可以假设 $\lambda(v_1) = c_i$、$\tau(v_1) = 3$ 且 $\psi(v_1) = \psi(I')$。设 u_1、$u_2 \in V$ 满足 (u_1, v_1)、$(u_2, v_2) \in E$，$\min_{v_1} = c_i \min_{u_1}$ 且 $\min_{v_2} = c_i \min_{u_2}$。如前所述，我们必然已经考察过 u_1 和 u_2，但由于 v_1 不在 I' 中，按照构造，它必有 $\min_{u_1} \prec \min_{u_2}$，因此我们得到 $\min_{v_1} \prec \min_{v_2}$。

b. 假设 $\psi(I') = \psi(I) + 1$。按照上文的论证，我们推断所有满足 $\psi(v) = \psi(I)$ 的 $v \in V_{c_i,3}$ 都已位于 $\mathcal{A}$ 后缀中的某个元素内，该后缀的最大元素为 I。利用这一信息，前一种情形的同一证明表明 I' 是正确的。

2. 假设 $i' = i$ 和 $t' \leq 2$。由于我们已经构造了 $\mathcal{A}_{c,2}$ 和 $\mathcal{A}_{c,1}$，我们只需证明：若 $t = 3$，则所有 $v \in V_{c_i,3}$ 都已在 $\mathcal{A}$ 的某个后缀元素中，该元素的最大元素为 I。反设对某个 $v \in V_{c_i,3}$，此结论不成立，并选取 v，使得 $\psi(v)$ 尽可能小。像前面一样论证，我们可得（由 $\psi(v)$ 的极小性）v 必须在某个 $A_{c_i,3}$ 中；但由于 $t' \leq 2$，我们可得 v 必须在 $\mathcal{A}$ 的某个后缀元素中，该元素的最大元素为 I，矛盾。

3. 假设 $i' = i - 1$ 和 $t' = 3$。特别地，$\psi(I') = 1$。与前一种情况一样，我们得到：若 $t = 3$，则所有 $v \in V_{c_i,3}$ 都已经位于 A 的某个后缀元素中，该元素的最大元素是 I。此外，$A_{c_i,2}$ 和 $A_{c_i,1}$ 必然为空，因为我们已经构造了它们，而 I' 不包含在其中任何一个中。首先，我们证明，若 $v_1, v_2 \in I'$，则 $\min_{v_1} = \min_{v_2}$。设 $u_1, u_2 \in V$ 满足 (u_1, v_1)、$(u_2, v_2) \in E$、$\min_{v_1} = c_{i-1} \min_{u_1}$ 且 $\min_{v_2} = c_{i-1} \min_{u_2}$。由于 $\tau(v) = 3$，我们得到以下两种情况之一：(i)$c_{i-1} \prec \lambda(u_1)$；或 (ii) $\lambda(u_1) = c_{i-1}$, $\tau(u_1) = 3$ 且 $\psi(u_1) = \psi(I') - 1$。然而，情况 (ii) 不可能成立，因为 $\psi(I') = 1$，因此必有 $c_{i-1} \prec \lambda(u_1)$，由此可得 u_1 是 A 的后缀的一个元素 K，其最大元素是 I。与前面一样，我们可得 u_2 也在 K 中，$\min_{u_1} = \min_{u_2}$ 且 $\min_{v_1} = \min_{v_2}$。

现在只需证明：若 v_1 既不在 I'中，也不在 A 的后缀中（该后缀的最大元素为 I），且 $v_2 \in I'$，则 $\min_{v_1} \prec \min_{v_2}$。若 $\tau(v_i) \neq 3$，结论显然成立，因此可以假设 $\tau(v_i) = 3$。不可能有 $c_{i-1} = \lambda(v_2) \prec \lambda(v_1)$，否则 v_1 就会在 A 的后缀中，而该后缀的最大元素为 I。也不可能有 $\lambda(v_1) = c_{i-1}$ 且 $\psi(v_1) < \psi(I')$，因为 $\psi(I') = 1$。由此可知，必有 $\lambda(v_1) \preceq c_{i-1}$，并且若 $\lambda(v_1) = c_{i-1}$，则 $1 = \psi(I') \leq \psi(v_1)$。若 $\lambda(v_1) \prec c_{i-1}$，或 $\lambda(v_1) = c_{i-1}$ 且 $1 = \psi(I') < \psi(v_1)$，则可立即得出 $\min_{v_1} \prec \min_{v_2}$。因此，可以假设 $\lambda(v_1) = c_{i-1}$、$\tau(v_1) = 3$ 且$\psi(v_1) = \psi(I') = 1$。取 u_1、$u_2 \in V$使得 (u_1, v_1)、$(u_2, v_2) \in E$，$\min_{v_1} = c_{i-1} \min_{u_1}$且 $\min_{v_2} = c_{i-1} \min_{u_2}$。与之前一样，我们必定已经考虑过 u_1 和 u_2；但由于 v_1 不在 I'中，根据构造，必有 $\min_{u_1} \prec \min_{u_2}$，因此得出

$\min_{v_1} \prec \min_{v_2}$。

4. 假设 $i' = i - 1$ 和 $t' \leq 2$。与之前一样，我们可以得出，若 $t = 3$，则所有 $v \in V_{c_i}$,3 都已位于 A 的后缀中某个最大元素为 I 的元素内。此外，A_{c_i},2 和 A_{c_i},1 必为空，因为我们已经构造了它们，而 I' 不包含在其中任何一个集合里。既然我们已经构造了 $A_{c_{i-1}}$,2 和 $A_{c_{i-1}}$,1，我们只需证明 $A_{c_{i-1}}$,3 为空。若它非空，则特别地，存在 $v \in V_{c_{i-1}}$,3，使得 $\psi(v) = 1$。若 $u \in V$ 满足 $(u, v) \in E$ 且 $\min_v = c_{i-1} \min_u$，那么我们可以得出，它必定满足 $c_{i-1} = \lambda(v) \prec \lambda(u)$。于是，$u$ 必须位于 A 的后缀中某个最大元素为 I 的元素内，因此 v 会位于 $A_{c_{i-1}}$,3 的某个元素内；但这与 $t' \leq 2$ 相矛盾。

最后，如果考察 I in A_{c_i},t 后，没有元素 I' 可供考察，我们可以验证，每个 $v \in V$ 都属于 A 的后缀中某个最大元素为 I 的元素（因此 A 已构建完成）。事实上，反设这一结论不成立。考虑由所有不满足这些性质的节点 v 组成的集合 S（它必定是 $\tau(v) = 3$）；令 $S' \subseteq S$ 为 S 中所有使 $\lambda(v)$ 取最大值的节点 v 的集合，并令 $S'' \subseteq S'$ 为 S' 中所有使 $\psi(v)$ 取最小值的节点 v 的集合。任取一个 $v \in S''$，并令 $u \in V$ 满足 $(u, v) \in E$ 且 $\min_v = \lambda(v) \min_u$。与之前一样，我们可以得出，$u$ 必属于 A 的后缀中某个最大元素为 I 的元素，因此 v 应属于某个 A_c,3，并且会有一个元素 I' 可供考察，这就产生了矛盾。