受有根树子网络性质启发的一类无根系统发育网络

Leo van Iersel[1],Mark Jones[2],Simone Linz[3],和 Norbert Zeh[4]

[1]代尔夫特理工大学代尔夫特应用数学研究所，荷兰代尔夫特[2]英国伦敦密德萨斯大学计算机科学系[3]新西兰奥克兰大学计算机科学学院，奥克兰[4]加拿大哈利法克斯达尔豪斯大学计算机科学学院

arXiv:2603.07000v1 [math.CO] 7 Mar 2026

2026年3月10日

摘要

若每个非叶顶点都有一个不是网状节点的子节点，则称有向系统发育网络为 *tree-child* 树子网络。作为一类有向系统发育网络，树子网络从计算角度来看非常有用。例如，在树子网络这一限制下，系统发育学中的若干计算难题会变得容易求解。同时，这一类网络本身也足够丰富，能够包含相当复杂的网络。此外，判断一个有向网络是否为树子网络可以在多项式时间内完成。本文旨在寻找一类 *undirected* 系统发育网络，使其像树子网络 *directed*这一类网络一样，既丰富又具有计算上的实用价值。适合这一用途的一个自然候选类别是树子可定向网络，它包含所有边都可以定向为树子网络的无向系统发育网络。然而，我们证明，识别这类网络是NP困难的，即使对于二元网络也是如此，因此该类别不适合承担这一作用。为寻找一类能够发挥与有向树子网络类似作用的无向网络，我们提出了新的类别，称为 q-可割网络，其中 $q \geq 1$为任意整数。我们证明，对于所有 $q \geq 1$，这些类别都具有许多理想性质，与有根情形下的树子网络类似，包括可在多项式时间内识别。为展示这一类别在计算上的实用性，我们证明，NP困难问题 TreeContainment 在限制为包含 $q \geq 3$的 q-可割网络时，可以在多项式时间内求解。

1 引言

系统发育网络是表示一组分类单元（指被认为是进化过程产物的不同实体，例如生物物种）之间进化关系的图。这些分类单元通常由系统发育网络中度为1的顶点或叶节点 表示。人们已提出三种适用于不同情形的系统发育网络变体。有根系统发育网络是只有一个入度为0的顶点或根 的有向无环图，也是对进化情景最详尽的描述。由于进化方向并非总能还原，人们也在研究无根系统发育网络，即无向图。最后，半有向系统发育网络兼具无向边和有向边，是前两者的混合形式。本文探讨有根和无根系统发育网络。

研究最为深入的一类有根系统发育网络无疑是有根系统发育树。这类有根系统发育网络中，每个非根顶点的入度均为1。树适用于表示分化事件，但无法表示趋同事件，也就是谱系合并的事件，例如因杂交而发生的合并。有根系统发育网络可以通过入度至少为2的顶点来显示这类事件，这些顶点称为网状节点，这是对杂交、重组或基因转移等所有非树状进化过程的统称。

对于许多任务而言，所有有根系统发育网络构成的完整类过于庞大，会导致问题极难处理，并造成不切实际的情形。因此，人们正在研究一些子类，它们比树的类更大，但比完整类更小。特别是，有根系统发育网络中一个研究较为充分的类是树子网络类 [9]。这类网络中，每个非叶顶点都至少有一个不是网状节点的子节点。

有根树子系统发育网络因多种原因而颇具吸引力。在某些生物学情境中，树子情形符合现实 [35]。此外，关于这一类及其子类，已有一些积极的编码 [5, 6, 9, 10, 27,40], 可识别性 [1, 18], 计数 [8, 11, 21, 37, 43], 刻画 [15, 36, 41], 几何 [39] 和算法结果 [3, 4, 12, 13, 20, 28, 30, 31]。例如，问题在于

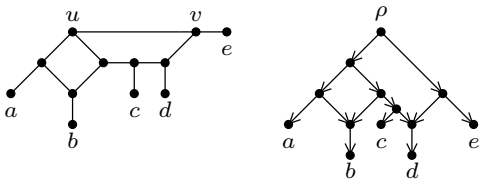


图1：一个具有五个叶节点的无根二叉系统发育网络 U，它是 2-可切但不是 3-可切（左），以及一个有根二叉系统发育网络 N，它是 U的树子定向（右）。注意，可以通过用一个新顶点 ρ 对边 $\{u, v\}$ 进行细分，并对所得图中的每条边进行定向，从而得到 N。

判断一个有根系统发育网络是否包含给定的有根系统发育树，这一问题称为 Rooted Tree Containment；一般情况下它是 NP完全的，而当网络为树子网络时则可在多项式时间内求解 [24, 28, 32, 34]。此外，有根树子网络是一个丰富的系统发育网络类，因为任何复杂程度任意高的系统发育网络都可以通过附加叶节点转化为树子网络。因此，有根树子网络的主要优势在于，它们能使许多问题在计算上变得可行，同时又不会对所研究的网络施加过多限制。

尽管进化是一个具有明确时间前进方向的过程，但人们仍在研究无根系统发育网络，因为进化方向并不总是能够从数据 [19,22, 23, 29, 38]中轻易恢复。将有根网络的类别扩展到无根情形的一种自然方法是考虑可定向性。具体而言，可以将无根系统发育网络定义为树子网络：如果能够为其边选择一种定向，使其成为有根树子网络（示例见图1）。对于给定的无根系统发育网络，判断是否可能实现这一点的计算问题称为 Tree-Child Orientation，其复杂性问题已经悬而未决一段时间 [14, 17, 25, 26, 44]。最近已有研究证明，对于无根非二叉系统发育网络，该问题是 NP困难的 [17]。在本文的前半部分，我们证明，即使对于无根二叉系统发育网络，该问题仍然是 NP困难的。因此，除非 P=NP，否则人们甚至无法在多项式时间内判断给定网络是否属于该类；由此可见，对于大多数用途而言，树子可定向网络并不是一个理想的无根系统发育网络类别。因此，有必要寻找新的无根系统发育网络类别，使其能够在多项式时间内被识别，并具备有根树子网络所具有的许多理想性质。

在论文的后半部分，我们针对任意整数 $q \geq 1$ 提出了 q-可切 网络这一类。若一个无根系统发育网络的每个环都包含一条路径，且该路径至少有 q 个顶点，并且路径上的每个顶点都与一条割边相接（移除该边会使网络断开），则称该网络为 q-可切。图1展示了一个 2-可切但非 3-可切的无根系统发育网络示例。 q-可切这一名称源自一个等价定义（见观察5）：该定义基本上指出，当且仅当切断每条至少包含 q 条割边的链后得到的是森林时，网络才是 q-可切的。我们证明，这些类别具有若干理想性质。首先，我们证明，对于任意整数 $q \geq 1$，可以在多项式时间内判定给定的无根系统发育网络是否为 q-可切。其次，我们证明 2-可切网络是树子可定向网络的子类（因而也是所谓“orchard”网络的子类 [14]）。第三，我们证明，该问题 Rooted Tree Containment 的无根对应版本虽然通常是 NP困难的 [29], ，但对于任意满足 $q \geq 3$ 的 q-可切网络，都可在多项式时间内求解。

2 预备知识

本节介绍本文余下部分将使用的符号和术语。全文中， X 表示一个非空有限集。下文中，我们考虑叶节点与 X中的元素一一对应标记的有向图和无向图。对于这些图，我们不区分叶节点与其标签。更确切地说，对于每个 $x \in X$，我们用 x 指代标记为 x 的叶节点，也指代该标签。设 G 和 G' 为有向图，且 $e = (u, w)$是 G的一条弧。将 e 替换为两条弧(u, v)和(v, w)的操作称为细分 e，其中 v 是一个新顶点。反过来，给定 G 的一个2度顶点 v，使得(u, v)和(v, w)是弧，则压缩 v 是指删除 v 并添加一条新弧(u, w)的操作。此外，如果 G可由 G' 压缩所有2度顶点得到，我们就称 G'是 G 的细分。需要指出的是，最后三个定义自然也适用于无向图。最后，对于无向图，$(v_1, v_2, \ldots, v_k)$表示路径 $P = \{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{k-1}, v_k\}$，其中 $k \geq 1$。如果 $k = 1$，那么 P只包含一个顶点，不含任何边。

系统发育网络。定义在 X 上的无根二叉系统发育网络 U 是一个简单连通无向图，其所有内部顶点的度均为3，且其度为1的顶点（称为叶节点）与 X中的元素一一对应标记。设 U的一条边为 $e = \{u, v\}$。如果删除 e会将 U 分成两个连通分量，则 e是 U 的一条割边。如果 e是割边，且 u 或 v 是 U的叶节点，则称 e 为平凡割边。此外， U 的块是 U 的一个极大子图，该子图没有割边且不只是单个顶点。最后，我们将 $r(U) = |E| - (|V| - 1)$称为 U的网状化数，其中 E和 V分别是 U的边集和顶点集。如果 $r(U) = 0$，则称 U为无根二叉系统发育 X-树。现在，设 U 是定义在 X上的无根二叉系统发育网络，且 T 是无根二叉系统发育 X-树。如果 U 的某个子图是 T的细分，我们就称 U 显示 T。最后，如果一棵子树可通过删除一条边从 T 中分离出来，则称其为 T 的悬挂子树。

转到有向图，有根二叉系统发育网络 N on X 是一个无平行弧的有向无环图，满足以下性质：

(i) 存在唯一的顶点 ρ，即根，其入度为 0，出度为 2；

(ii) 每个叶节点的入度为 1，出度为 0，且所有叶节点与 X 中的元素一一对应地标记；(iii) 所有其他顶点的入度和出度要么分别为 1 和 2，要么分别为 2 和 1。

设 N 是 X 上的有根二叉系统发育网络，(u, v) 是 N 的一条弧，且 w 是 N 的一个顶点。我们称 u 是 v 的父节点，称 v 是 u 的子节点。此外，若 w 的入度为 1 、出度为 2，则称 w 为树顶点；若 w 的入度为 2、出度为 1，则称 w 为网状节点。 N 的网状化数 $r(N)$ 等于 N 中网状节点的数量；如果 $r(N) = 0$，则称 N 为有根二叉系统发育 X-树。

由于本文中的所有系统发育网络都是二叉的，因此我们将无根二叉系统发育网络简称为无根系统发育网络，将有根二叉系统发育网络简称为有根系统发育网络。

摘樱序列。设 U 是一个定义在 X 上的无根系统发育网络，且 x 和 y 是 X 中两个不同的元素。如果 $|X| = 2$，或者 x 和 y 与同一个顶点相邻，则称 $\{x, y\}$ 为 U 的一对樱桃。此外，如果 U 中存在不同的内部顶点 u 和 v，使得 $\{x, u\}$、$\{u, v\}$ 和 $\{v, y\}$ 是 U 中的边，且 $\{u, v\}$ 是一个环的边，则 $\{x, y\}$ 是 U 中的网状樱桃；在这种情况下，我们称 $\{u, v\}$ 为 $\{x, y\}$ 的中心边。接下来，我们定义两个用于约简 U 中樱桃和网状樱桃的操作。设 (x, y) 是一对叶节点的有序对，且 $\{x, y\}$ 是 U 中的樱桃或网状樱桃。于是，约简 (x, y) 是按以下两种方式之一修改 U 的操作：

(i) 若 $\{x, y\}$ 是 U 中的樱桃，则删除 x；若 $|X| \geq 3$，还需抑制由此产生的 2 度顶点。

(ii) 若 $\{x, y\}$ 是 U 中的网状樱桃，则删除 $\{x, y\}$ 的中心边，并抑制由此产生的两个 2 度顶点。

我们将 (i) 称为樱桃约简，将 (ii) 称为网状樱桃约简。现在，设

$$\sigma = ((x_1, y_1), (x_2, y_2), \ldots, (x_s, y_s))$$

为 $X \times X$ 中的一个有序对序列。令 $U_0 = U$，并对所有 $i \in \{1, 2, \ldots, s\}$，令 U_i 为从 U_{i-1} 通过约简 (x_i, y_i) 得到的无根系统发育网络，其中若 $\{x_i, y_i\}$ 是 U_{i-1}中的樱桃或网状樱桃，则称 σ 为 U 的挑樱桃序列，当且仅当 U_s 由单个顶点组成。

系统发育网络的类别及其定向。设 N 为一个有根系统发育网络。若 N 的每个非叶顶点都有一个子节点是树顶点或叶，则称 N 为树子。此外，若两个网状节点之间有一条弧相连，则称 N 含有一个堆叠；若存在两个具有共同父节点的网状节点，则称 N 含有一对兄弟网状点。根据树子网络的定义，可得以下众所周知的等价关系。

引理1. 设 N 是一个有根系统发育网络。则 N 是树子，当且仅当它既没有堆叠，也没有一对兄弟网状点。

设 U 为一个无根系统发育网络。对 U 进行定向 的方法是：在 U 的一条边上细分出一个新顶点 ρ，从而由 U 得到一个图，然后为所得图的每条边指定方向。Janssen等人 [33] 指出，并非每个无根系统发育网络都存在一种定向，使其成为有根系统发育网络。好的一面是，可以在多项式时间内判定一个无根系统发育网络能否被定向为有根系统发育网络 [7, 33]。接下来讨论

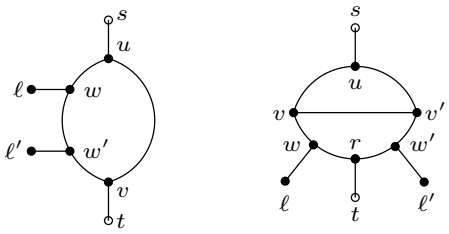


图2：连接构件（左）和网状构件（右）。

无根系统发育网络的类别。若 U 的每条割边都与一个叶相接，则称 U 为简单 的。此外，如果 U 具有一个挑樱桃序列，则称其为果园网络 。注意，无根系统发育网络是果园网络，当且仅当它存在一种定向，使其成为有根果园网络（有根果园网络的定义及证明见 [14, 定理2.1] ）。接下来，如果 U 的一种定向是以 ρ为根的树子网络，则称 U 具有树子定向。由此引出以下定义：若 U 具有树子定向，则称其为树子的。然而，正如我们将在第3节中看到的，判定 U 是否具有树子定向是一个NP完全问题。鉴于这一不太理想的结果，我们引入以下新的无根系统发育网络类别，并将在第4节和第5节中证明，这一类别具有若干有利性质，与有根情形下的树子网络类似。

定义 1。设 U 是一个无根系统发育网络，$q \geq 1$ 是一个整数。若 U 的每个环都包含一条至少含有 q 个顶点的路径，且该路径上的每个顶点都与一条割边相接，则称 U 为-可割 。

3 树子定向

本节将证明以下判定问题是 NP完全的。

Tree-Child Orientation**实例**。一个无根二叉系统发育网络 U。**问题**。U 是否存在树子定向?

我们首先介绍两个构件，分别称为连接构件 C和网状构件 R，如图2所示。我们将此类构件中的顶点 s （分别为 t）称为其 s-端点（分别为 t-端点）。稍后将看到，我们将多个 C 和 R 组合起来，构造一个无根系统发育网络 U，其子图同构于 C 或 R。图2–4中，C 和 R 中作为 U 内部顶点的叶节点以小空心圆表示。因此，C 和 R 中同时也是 U 叶节点的叶节点只有标记为 ℓ和 ℓ' 的顶点。其他顶点标签仅用于指代顶点，不应视为真正的标签。值得注意的是，连接构件以及下一条引理的证明最初见于预印本 [16]。由于发表版本 [17]，中未收录这些内容，为求完整，我们在此给出详细说明。

引理2. 设 U 是 X上的一个无根系统发育网络，并设 G是 U 的一个与连接构件同构的子图。假设 U 具有树子定向 N，使得N 的根不细分 G的一条边。则 N 满足以下两个性质之一：

（i）(s,u) 与 (v,t) 是 N中的弧，且 u 不是网状节点，或者

（ii）(u,s) 与 (t,v) 是 N中的弧，且 v 不是网状节点。

证明。首先，假设(u,s)与(v,t)是 N中的弧。也就是说，这两条弧都从该构件向外定向。由于 ℓ与 ℓ' 是叶节点，(w,ℓ)与(w',ℓ')是 N中的弧。由对称性，我们可以假设(u,v)是 N中的一条弧。进而，这意味着弧(w,u)存在，否则 u的入度为0。再次使用同一论证，(w',w)与(v,w')也是 N中的弧。于是(u,v)、(v,w')、(w',w)与(w,u)是 N中一个有向环的弧，从而与 N 无环相矛盾。

其次，假设（s，u）和（t，v）是 N中的弧。也就是说，这两条弧都指向该构件。根据对称性，我们仍可假设（u，v）是 N中的一条弧。于是，（v，w'）是 N中的一条弧，因为否则 v 的出度将为 0。由于 N 是树子，且 v 是网状节点，因此 w' 是树顶点，这意味着（w'，w）是 N中的一条弧。接着，（u，w）不可能是一条弧，否则 u的两个子节点都将是网状节点。因此，（w'，w）和（w，u）是 N中的弧。此时，（u，v）、（v，w'）、（w'，w）和（w，u）又构成了 N中的一个有向环，产生了另一个矛盾。结合这两种情况，我们可推知，在 N中，要么（s，u）和（v，t）是弧，要么（u，s）和（t，v）是弧。

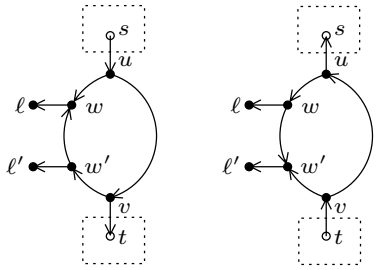


图3：一个有根系统发育网络 N，其以连接构件作为子图。如果 N 是树子的，则连接构件的边例如可以按所示两种方式之一定向。虚线矩形表示 N中被省略的部分。

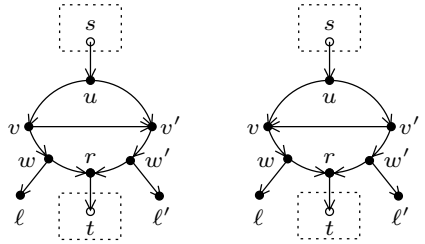


图4：一个有根系统发育网络 N，其以网状构件作为子图。如果 N 是树子的，则网状构件的边恰好按所示两种方式之一定向。虚线矩形表示 N中被省略的部分。

现在假设（s，u）和（v，t）是 N中的弧，且 u是网状节点。如果（u，v）是 N中的一条弧，那么（w，u）也必定是 N中的一条弧。此外，由于 N是树子，v是树顶点。利用每个非叶顶点都有一个非网状子顶点且没有顶点的出度为 3 这一事实，不难验证（u，v）、（v，w'）、（w'，w）和（w，u）构成了 N中的一个有向环，矛盾。另一方面，如果（v，u）是 N中的一条弧，那么（u，w）也必定是 N中的一条弧。由于 w是 N中的一个树顶点，因此可知（v，u）、（u，w）、（w，w'）和（w'，v）是有向环中的弧，产生了另一个矛盾。因此，如果（s，u）和（v，t）是 N中的弧，那么 u就不是网状节点。

我们通过指出以下事实来完成该证明：与上一段类似的论证可用于表明，如果$(u,\ s)$和$(t,\ v)$是 N中的弧，则 v不是网状节点。 □

承接上一引理，图3展示了当连接构件是树子网络的子图时，对其边进行定向的两种可能方式。

引理 3. 设 U 是 X上的无根系统发育网络，且设 G 是 U 的一个与网状构件同构的子图。假设 U 具有树子定向 N，使得 N的根不细分 G的一条边。则(w,r)、(w',r)、(r,t)和(s,u)是 N中的弧。

证明。根据对称性，在整个证明过程中，我们可以假设 (v,v') 是 N 中的一条弧。假设 v' 是 N 中的一个树顶点。因此，(v',u) 和 (v',w') 是 N 中的弧；又由于 N 无环，(v,u) 也是 N 中的一条弧。此外，由于 v 的入度至少为 1，弧 (w,v) 存在。对 w 应用相同的论证，弧 (r,w) 也存在。由此可知，(r,w') 是 N 中的一条弧；否则，N 会包含一个有向环。现在，v' 的两个子节点 u 和 w' 都是网状节点，这与 N 是树子网络相矛盾。因此，在证明余下部分中，我们可以假设 v' 是一个网状节点。接下来，假设 (w',v') 和 (v',u) 是 N 中的弧。由于 N 是树子网络，u 的入度为 1，这意味着 (u,v) 是一条弧。由此可得，(v,v')、(v',u) 和 (u,v) 构成 N 中的一个有向环，矛盾。因此，(u,v') 和 (v',w') 是 N 中的弧。由于 N 是树子网络，w' 是一个树顶点，这意味着弧 (w',r) 存在。如果 (w,v) 是一条弧，那么 (r,w) 也是一条弧，因为 w 的入度至少为 1。此时，弧 (v,v')、(v',w')、(w',r)、(r,w) 和 (w,v) 构成一个有向环，再次产生矛盾。因此，(v,w) 是 N 中的一条弧。由于 N 是树子网络，且 v' 是一个网状节点，所以 w 是一个树顶点，因此弧 (w,r) 存在于 N 中。容易验证，只要 N 是树子网络，图4就展示了对 G 的边进行定向的仅有两种可能。这就证明了该引理。

□

为了证明 Tree-Child Orientation 是NP完全的，我们从 Balanced 3-SAT 的一个变体进行归约；在该变体中，每个文字恰好出现两次，并且Berman等人 [2, 定理 1]已表明该变体仍然是NP完全的。在能够正式陈述这一变体之前，我们需要一些额外的定义。设 $V=\{x_1,x_2,\ldots,x_n\}$ 为一组变量。一个文字 是变量 x_i或其否定 $\bar{x}_i$，而一个子句 是 $\{x_i,\bar{x}_i$的某个子集的析取：$i\in\{1,2,\ldots,n\}\}$。此外，一个合取范式中的布尔公式 是若干子句的合取，即具有如下形式的表达式

$\Phi = \bigwedge_{j=1}^{m} c_j$，其中对所有 j， c_j 都是一个子句。 V的真值赋值 是一个映射 β： $V \to \{T, F\}$，其中 T 表示真值“真”， F 表示真值“假”。真值赋值 β满足 Φ，如果在 β下每个子句中至少有一个文字取值为 T。

2-Balanced 3-SAT
实例。 一组变量 $V = \{x_1, x_2, \ldots, x_n\}$和一组关于 V 的子句 $C = \{c_1, c_2, \ldots, c_m\}$，其中每个子句由三个文字组成，并且在所有子句中，每个变量恰好以否定形式出现两次，以非否定形式出现两次。**问题**。 是否存在一个满足 C的 V 真值赋值?

在本节余下部分中，我们用 ℓ 表示一个带标记的叶，用 l 表示一个文字。现在设 Φ 是一个 2-Balanced 3-SAT 实例，其变量为 $V = \{x_1, x_2, \ldots, x_n\}$，子句为 $C = \{c_1, c_2, \ldots, c_m\}$。此外，对于每个 $j \in \{1, 2, \ldots, m\}$，令 $c_j = (l_j^1 \vee l_j^2 \vee l_j^3)$。接下来，我们构造一个无根系统发育网络，其中包含 $n_r = 1 + 2n$ 个网状构件副本和 $n_c = 3m + 2n$ 个连接构件副本，并且文字 l_j^k（其中 $k \in \{1, 2, 3\}$）作为内部顶点。我们通过以下四个步骤来完成这一构造。

1. **根小工具**。令 $P = (p_1、\ p_2、\ \ldots\ldots、\ p_{n_r-2})$为一条路径。此外，令 R_r、R^1、R^2、.....、R^{n_r-1} 为网状构件的副本。现在，将 R_r 的 t 端点与 p_1 重合，并添加两条边，使其分别将 R_r 的 s 端点与新的叶相连，分别记为 ℓ_r 和 ℓ'_r。此外，对于每个 $k \in \{1, 2, \ldots, n_r - 2\}$，将 R^k 的 s 端点与 p_k 重合，并将 R^{n_r-1} 的 s 端点也与 p_{n_r-2} 重合。

2. **子句小工具**。对于每个子句 c_j，令 l_j^1、l_j^2、l_j^3、ℓ_j^1、ℓ_j^2、ℓ_j^3 和 z_j 为新的顶点，并令 C_j^1、C_j^2 和 C_j^3 为连接构件的副本。对于每个 $k \in \{1, 2, 3\}$，添加边 $\{l_j^k, t_j^k\}$、$\{t_j^k, \ell_j^k\}$，并将 s_j^k 与 z_j 重合，其中 s_j^k 和 t_j^k 分别是 C_j^k 的 s 端点和 t 端点。也就是说，子句小工具是通过将三个连接构件的 s 端点重合，并向三个 t 端点各添加两个相邻的新顶点而构造的。

3. **变量构件**。对于每个变量 x_i，令

$$X_i = \{l_j^k : j \in \{1, 2, \ldots, m\}, k \in \{1, 2, 3\}, \text{ and } l_j^k = x_i\}, \text{ and let}$$

$$\bar{X}_i = \{l_j^k : j \in \{1, 2, \ldots, m\}, k \in \{1, 2, 3\}, \text{ and } l_j^k = \bar{x}_i\}.$$

回顾 $|X_i| = |\bar{X}_i| = 2$。令 G_i^1 和 G_i^2 为连接构件的副本。添加一条边，将 G_i^1 的 s-端点与一个新顶点 r_i^2 连接起来；再添加一条边，将 r_i^2 与 G_i^2的 t-端点连接起来。此外，添加一条边，将 G_i^1 的 t-端点与一个新顶点 r_i^1 连接起来；再添加一条边，将 r_i^1 与 G_i^2的 s-端点连接起来。

4. **连接根、子句和变量构件**。对于每个变量 x_i，执行以下两个步骤。首先，为了将 x_i 的变量构件与根构件连接起来，将每个顶点 r_i^h与 $h \in \{1, 2\}$ 识别为某个网状构件的 t-端点，其中该网状构件不是 R_r。对于所有 n 变量构件，此识别过程应使每个网状构件的 t-端点（该网状构件不是 R_r）恰好被识别一次。其次，为了将 x_i 的变量构件与子句构件连接起来，将 G_i^1 和 G_i^2 的 s-端点与某个子句构件的顶点 l_j^k 识别，使得$l_j^k \in X_i$；并将 G_i^1 和 G_i^2 的 t-端点与某个子句构件的顶点 l_j^k 识别，使得 $l_j^k \in \bar{X}_i$。对于所有 n 变量构件，此识别过程应使每个顶点 l_j^k与 $j \in \{1, 2, \ldots, m\}$ 和 $k = \{1, 2, 3\}$ 恰好与某个连接构件的 s-端点或 t-端点识别一次。将所得的图记为 U_Φ。

对于 2-Balanced 3-SAT 实例，构造如图5所示

$$\Phi = (x \vee \bar{y} \vee z) \wedge (\bar{x} \vee y \vee \bar{z}) \wedge (x \vee y \vee \bar{z}) \wedge (\bar{x} \vee \bar{y} \vee z).$$

不失一般性，我们可以假设 U_Φ 的所有叶节点都具有不同的标签。随后，直接检查可知， U_Φ 是一个具有 $|X| = 2 + 2(n_r + n_c) + 3m$ 的 X 无根系统发育网络。

定理 4. Tree-Child Orientation是NP完全的。

证明。 设 U 是一个无根系统发育网络， N 是 U的一种定向。由于可以在多项式时间内检查 N 是否为树子，因此 Tree-Child Orientation属于 NP。

为了证明 NP难性，设 Φ 是 2-Balanced 3-SAT的一个实例，并设 U_Φ 是根据紧接在 2-Balanced 3-SAT的正式陈述之后的构造，从 Φ 重建得到的无根系统发育网络。此外，设 $V = \{x_1, x_2, \ldots, x_n\}$是 Φ的变量集合， $C = \{c_1, c_2, \ldots, c_m\}$是 Φ的子句集合。容易验证， U_Φ的规模关于 m 和 n是多项式的。我们通过证明 Φ 是一个是实例，当且仅当 U_Φ 具有树子定向，来完成证明。

假设 Φ 是一个是实例。令 β 为满足 Φ的真值赋值。接下来我们构造一个定向 N_Φ 的 U_Φ。

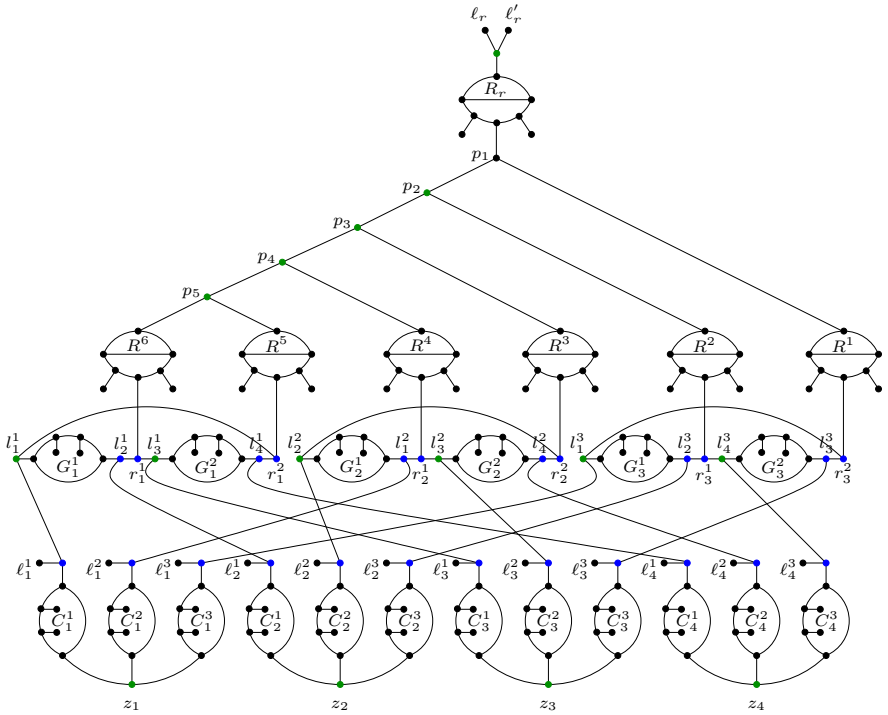


图5：无根系统发育网络 U_Φ 对应于 $\Phi = (l_1^1 \vee l_1^2 \vee l_1^3) \wedge (l_2^1 \vee l_2^2 \vee l_2^3) \wedge (l_3^1 \vee l_3^2 \vee l_3^3) \wedge (l_4^1 \vee l_4^2 \vee l_4^3) = (x \vee \bar{y} \vee z) \wedge (\bar{x} \vee y \vee \bar{z}) \wedge (x \vee y \vee \bar{z}) \wedge (\bar{x} \vee \bar{y} \vee z)$。为简化展示，省略了部分叶标签。每个连接构件和网状构件的 s端（相应地， t端）以绿色（相应地，蓝色）标示。注意， p_1 是 R_r 的 t端，也是 R^1的 s端。

1. 在与 ℓ_r 相连的边上插入一个新顶点 ρ，并将所有与 ρ 相连的边从 ρ向外定向。

2. 对于每个叶 ℓ，将与 ℓ 关联的边定向为指向 ℓ。

3. 按图4左侧所示，将每个网状构件 $R_r, R^1, R^2, \ldots, R^{n_r-1}$ 的边进行定向。

4. 对每个 $k \in \{1, 2, \ldots, n_r - 3\}$，将 P 的边 $\{p_k, p_{k+1}\}$ 定向为 (p_k, p_{k+1})。

5. 将与网状构件 $R^1, R^2, \ldots, R^{n_r-1}$ 的 t-端以及未包含在任何网状构件中的一个顶点相连的每条边，都从 t-端向外定向。

6. 对每个满足 $\beta(x_i) = T$ 的 $i \in \{1, 2, \ldots, n\}$，按图3左侧所示，将连接构件 G_i^1 和 G_i^2 的边定向。同样，对每个满足 $\beta(x_i) = F$ 的 $i \in \{1, 2, \ldots, n\}$，按图3右侧所示，将连接构件 G_i^1 和 G_i^2 的边定向。

7. 对每个 $j \in \{1, 2, \ldots, m\}$ 和 $k \in \{1, 2, 3\}$，将边 $\{l_j^k, t_j^k\}$ 定向为 (l_j^k, t_j^k)，并应用以下两种情况之一：

   (i) 如果 c_j 至少包含一个在 β 下取值为 F 的文字，那么，如果 l_j^k 在 β 下取值为 F，则按图3右侧所示为 C_j^k 的边定向；如果 l_j^k 在 β 下取值为 T，则按图3左侧所示定向。

   (ii) 如果 c_j 包含三个在 β 下分别取值为 T 的文字，那么按图3左侧所示为 C_j^1 和 C_j^2 的边定向，并按图3右侧所示为 C_j^3 的边定向。

例如， U_Φ 的定向 N_Φ 对于

$$\Phi = (x \vee \bar{y} \vee z) \wedge (\bar{x} \vee y \vee \bar{z}) \wedge (x \vee y \vee \bar{z}) \wedge (\bar{x} \vee \bar{y} \vee z)$$

如图6所示。

根据 N_Φ 的构造可立即得出， N_Φ 的每个内部顶点要么入度为1且出度为2，要么入度为2且出度为1。接下来，我们对 N_Φ中的网状化事件作四点观察。首先，每个连接构件恰好包含一个网状节点，而每个网状构件恰好包含两个网状节点（见图3和图4）。其次，根据步骤6，对于每个满足 $\beta(x_i) = T$的 $i \in \{1, 2, \ldots, n\}$，连接构件 G_i^1和 G_i^2的 t端点是一个网状节点，

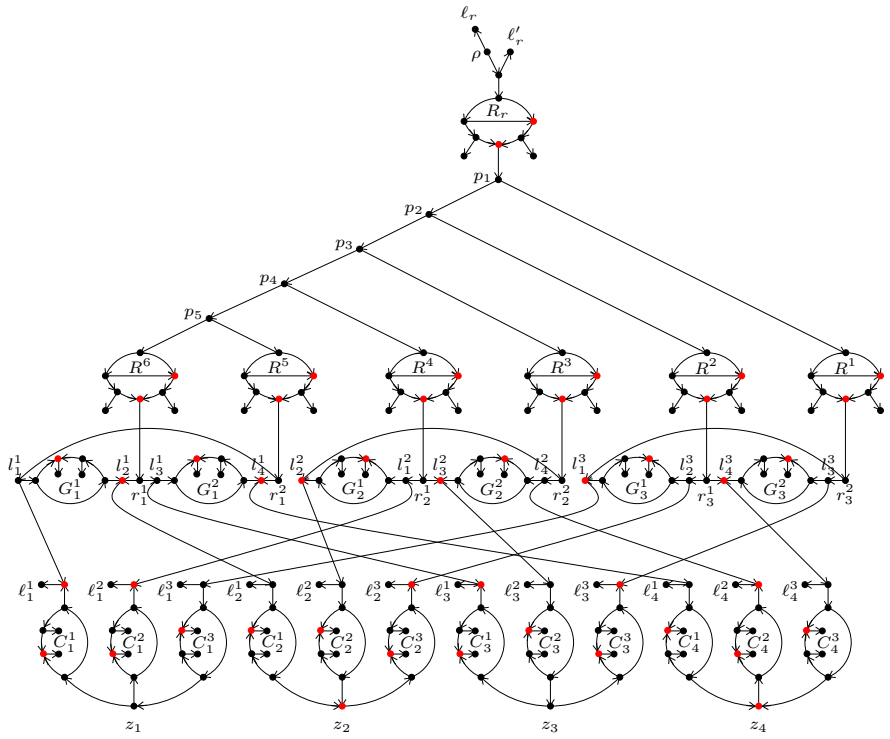


图6：具有真值赋值 $\beta(x) = T, \beta(y) = F$, and $\beta(z) = F$且满足 Φ 的有根系统发育网络 N_Φ for $\Phi = (l_1^1 \vee l_1^2 \vee l_1^3) \wedge (l_2^1 \vee l_2^2 \vee l_2^3) \wedge (l_3^1 \vee l_3^2 \vee l_3^3) \wedge (l_4^1 \vee l_4^2 \vee l_4^3) = (x \vee \bar{y} \vee z) \wedge (\bar{x} \vee y \vee \bar{z}) \wedge (x \vee y \vee \bar{z}) \wedge (\bar{x} \vee \bar{y} \vee z)$。请注意，$N_\Phi$ 是图5所示无根系统发育网络 U_Φ 的一个树子定向。为简化说明，省略了部分叶标签。网状化事件以红色标出。

而对于每个满足 $\beta(x_i) = F$的 $i \in \{1, 2, \ldots, n\}$，连接构件 G_i^1和 G_i^2的 s端点是一个网状节点。第三，根据步骤7，若 l_j^k是一个网状节点，且 $j \in \{1, 2, \ldots, m\}$和 $k \in \{1, 2, 3\}$满足相应条件，则 t_j^k是一个树顶点。第四，当且仅当 c_j 包含两个在 β下均取值为 F的文字时，z_j 才是一个网状节点。综合以上观察，经过简单验证可知，N_Φ 中的每个非叶顶点都有一个子节点是叶或网状节点。最后，我们证明 N_Φ 无环。若 N_Φ 包含一个有向环，则该环包含某个 $j \in \{1, 2, \ldots, m\}$对应的 z_j，以及某个连接构件 G_i^1或 G_i^2中的一个顶点，且满足$i \in \{1, 2, \ldots, n\}$。观察到 U_Φ 中的每条边 $\{l_j^k, t_j^k\}$ 在 N_Φ中都按(l_j^k，t_j^k)的方向定向，可知这样的环不可能存在。总之，N_Φ 是一个以 ρ为根的树子网络。

反过来，假设 U_Φ 具有树子定向 N_Φ。令 ρ 为 N_Φ 的根。由于存在一条从 ρ 到 N_Φ 每个叶的有向路径，根据 ℓ_r 和 ℓ_r' 的存在以及引理 3，可知 ρ 将 R_r 的一条边细分，或将一条与 ℓ_r 或 ℓ_r' 相接的边细分。

现在令 i 为 $\{1, 2, \ldots, n\}$ 中的一个元素。对每个 $i \in \{1, 2, \ldots, n\}$，考虑顶点 r_i^1 和 r_i^2。根据构造，每个这样的顶点都与一个网状构件的 t 端相识别。由于 N_Φ 是树子网络，根据引理 3，r_i^1 和 r_i^2 均为某个网状节点的子节点，因此是树顶点。令 u_i^2 和 v_i^1（分别为 u_i^1 和 v_i^2）表示 r_i^1（分别为 r_i^2）的两个邻点，且它们不是网状构件的顶点。不失一般性，我们可以假设，对每个 $h \in \{1, 2\}$，u_i^h 是 G_i^h 的 s 端，而 v_i^h 是 G_i^h 在 U_Φ 中的 t 端。根据引理2，G_i^h 的边在 N_Φ 中的定向满足以下二者之一：存在一条从 u_i^h 到 v_i^h 的有向路径，或存在一条从 v_i^h 到 u_i^h 的有向路径，且这些路径仅包含 G_i^h 的边。回顾 r_i^1 和 r_i^2 均为树顶点，可知 u_i^h 或 v_i^h 是网状节点。特别地，如果存在一条从 u_i^1 到 v_i^1 的有向路径，那么 v_i^1 是网状节点；否则，u_i^1 是 N_Φ 中的网状节点。假设 v_i^1 是网状节点。由于 N_Φ 是树子网络，且 v_i^1 和 u_i^2 以 r_i^2 为共同父节点，因此 u_i^2 是树顶点。于是，在 N_Φ 中存在一条从 u_i^2 到 v_i^2 的有向路径，特别地，v_i^2 是网状节点。现在假设 u_i^1 是网状节点。由于 N_Φ 是树子网络，v_i^2 是树顶点。因此，在 N_Φ 中存在一条从 v_i^2 到 u_i^2 的有向路径，特别地，u_i^2 是网状节点。因此，u_i^1 和 u_i^2，或 v_i^1 和 v_i^2，必为网状节点。

令 β为针对 V的真值赋值：对于每个 $i \in \{1, 2, \ldots, n\}$，若 v_i^1和 v_i^2在 N_Φ中均为网状节点，则赋值为 $x_i = T$；若 u_i^1 和 u_i^2 在 N_Φ中均为网状节点，则赋值为 $x_i = F$。注意，根据从 Φ构造 U_Φ的步骤4，$\{v_i^1, v_i^2, u_i^1, u_i^2\}$ 中的每个元素也是一个顶点 l_j^k，它对应于 Φ中的文字 l_j^k。特别地，当且仅当 l_j^k是 N_Φ中的网状节点时，l_j^k 才由 β 赋值为 F。

我们通过证明 β 满足 Φ 来完成证明。反设 β 不满足 Φ。那么，对于某个 $j \in \{1, 2, \ldots, m\}$，存在一个子句 $c_j = (l_j^1 \vee l_j^2 \vee l_j^3)$，使得在 β 下 $l_j^1 = l_j^2 = l_j^3 = F$。此外，根据上述论证可知，$l_j^1$、$l_j^2$ 和 l_j^3 在 N_Φ 中都是网状节点。

反过来，由于 N_Φ 是树子，每个满足 $k \in \{1,2,3\}$ 的 t_j^k 都是树顶点。因此，根据引理2，三个连接构件 C_j^k 的方向均如图 3 右侧所示。重要的是， z_j 在 N_Φ 中的入度为 3，这就导出了最后的矛盾。至此，定理得证。 □

4 q-可割网络的性质

在本节中，我们建立 q-可割网络的若干性质。设 U 为一个无根系统发育网络。

q-可割这一名称源于如下刻画。一条链 是一条路径，其顶点均位于同一块中，且均与一条割边关联。链的长度 是该路径中的顶点数。链的边是该路径的边，由定义可知它们不是割边。一条长度为q 的链是极大 的，如果它不包含在一条长度为q' 且满足 $q' > q$的链中。

观察5. 若 U 是一个无根系统发育网络，且 q 是满足 $q \geq 1$的整数，则下列陈述等价。

(i) U 是 q-可割的，即每个环都包含一条 q-链。

(ii) 删除位于长度为q 的链中的每个顶点后，得到一个森林。

(iii) 从每条长度至少为 q 的极大链中任意删除一条边，所得结果是一个森林。

以下定理指出，可以在多项式时间内判定一个无根系统发育网络是否为 q-可割，这可直接由观察5得出。

定理 6. 设 U 是一个无根系统发育网络，且 q 是满足 $q \geq 1$的正整数。可以在多项式时间内验证 U 是否为 q-可割。

为证明下面的命题，设 U 是一个无根系统发育网络。 U 的部分定向 U' 是一个由（无向）边和（有向）弧构成的图，可通过将 U中的部分边定向而从 U得到。如果 e在 U中是割边，且在 U'中是弧，则称 e为 U'中的一条割弧 。

命题7. 设 U 是一个以 X为对象的无根系统发育网络。如果 U 是2-可割的，那么它具有一个树子定向。

证明。 假设 U 是2-可切的。设 S 是 U 的边子集，它恰好包含每一条边$e = \{u, v\}$，使得 e 不是割边，并且在 U中存在割边 $\{u, u'\}$ 和 $\{v, v'\}$，即 S由所有位于长度至少为2的极大链中的边组成。由于 U 是2-可切的，根据观察5，存在 S 的一个子集 S'，使得从 U 中删除 S' 中的每一条边后，得到 U 的一棵生成树 T，其叶集为 X。见图7中的例子。现在用根 ρ 细分 U 的任意一条割边，并通过将与 ρ 关联的两条边以及 T的所有边定向为远离 ρ （相对于 T），得到所得网络的一个部分定向 U' 。

尚需证明， U'的其余边，即 S'中的边，可以按满足树子条件和无环条件的方式定向。考虑 U 的一个块 B 以及 U'的对应块 B' 。设 r 是 B' 中具有一条入割弧的唯一顶点。我们首先给出两个观察，它们都由 S' 的定义以及以下事实得出： U 是二叉的，并且从B 中删除 S' 中的边后所得的图是连通的。

(OB1) 在 B 中不存在满足 $\{u, v\}, \{v, w\} \in S'$ 的路径 (u, v, w)。

(OB2) 在 B 中不存在满足 $\{t, u\}, \{v, w\} \in S'$ 且满足 $\{u, v\} \notin S'$ 的路径 (t, u, v, w)。

现在构造一个辅助多重图 G_B ，其顶点集由 S' 中属于 B 的边组成；对于 B 中每条满足 $e = \{s, t\}$、$f = \{v, w\}$ 和 $e \neq f$ 的路径 (s, t, u, v, w)，图中都有一条边 $\{e, f\}$。注意，若 (s, t, u, v, w) 是 B 中的一条路径，则反向路径 (w, v, u, t, s) 也是 B 中的一条路径。这在 G_B 中仅对应一条边。此外， G_B 是一个多重图，因为可能存在另一条路径 (t, s, z, w, v)，从而在 G_B 中引入第二条边 $\{e, f\}$。

假设 G_B 有一个度至少为 3 的顶点。那么，存在一个顶点 u 属于 B ，使得 (u, v_1, w_1)、(u, v_2, w_2) 和 (u, v_3, w_3) 是 U中的三条边不相交路径，并且每条满足$i \in \{1, 2, 3\}$ 的边 $\{v_i, w_i\}$ 都是 S' 中的一条边。由此可知，从 B中删除 S'中的所有边后得到的图是不连通的；这与假设矛盾。因此， G_B 中的每个顶点的度至多为 2，这意味着 G_B由两两顶点不相交的路径和环构成。

现在按如下方式构造 B' 的一个定向 N_B 。对于 G_B 的每个连通分量 C ，设其顶点集为 V_C，我们在 B' 中定义一条路径 P_C ，该路径包含边和弧，并且可以沿任一方向遍历弧。具体而言， B' 中的路径 P_C 会遍历 S' 中与 V_C 中的顶点对应的所有边，而不遍历 S'中的其他边。此外，如果 P_C 包含 r （即 B' 中唯一一个具有一条

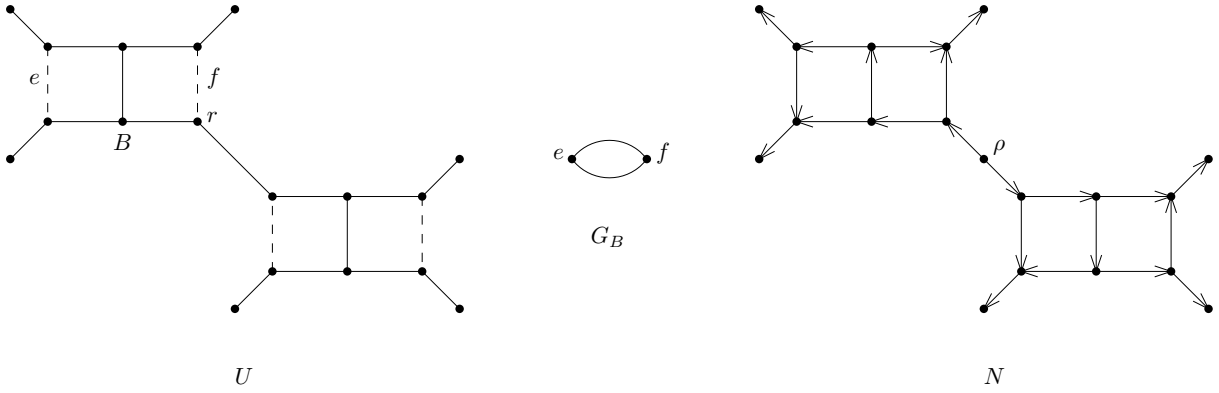


图7：命题7证明示例。左：虚线边标示集合 S' 的一个 2-可切网络。中：块 B 的辅助多重图 G_B 。右：U 的一个树子定向 N。为简化表述，省略了叶标签。

入向割弧的顶点），并且遍历一条边 $\{r, v\} \in S'$，那么 P_C 会先于 v 遍历 r （否则，我们可以反向遍历该路径）。根据构造， P_C 总是存在。接着，对于每条满足 P_C中的 $\{g, h\}$ 是一条边的边 $\{g, h\} \in S'$，如果 P_C 先于遍历 h 而遍历 g，则将边 $\{g, h\}$ 替换为弧（g， h）；否则，替换为弧（h， g）。接下来反设 N_B中存在一个有向环 C 。那么 C 包含一个网状节点；根据构造，该节点是 S'中某条弧的头。事实上，如果 C中的 r 是一个顶点，那么 C 中指向 r 的弧属于 S'，但这是不可能的，因为遍历这条边的路径 P_C 会先于 r遍历 v 。因此， r 不是 C中的顶点。所以， C中存在一个顶点 v，它有一条入向弧（u， v），该弧不是 T中从 r到 v的路径上的 C中的弧。由于v 也有一条入向弧（u'， v），且该弧属于 C，因此 v 是一个网状节点。此外，我们有$\{u', v\} \in S'$。因此，（u， v）是一条割弧。然而， B中只有顶点 r具有入向割弧，而我们刚刚发现 $v \neq r$。这就是所需的矛盾，说明 N_B 中不含有向环。

现在证明 N_B 没有堆叠，也没有一对兄弟网状点。由 (OB1) 和 (OB2) 可直接推出 N_B 没有堆叠。接下来，假设 N_B 有一对兄弟网状点 v_1 和 v_2。在 N_B 中，令 p 为 v_1 和 v_2 的共同父节点，并令 u_1 （分别为 u_2）为 v_1 （分别为 v_2）除 p 之外的另一个父节点。于是 (u_1, v_1, p, v_2, u_2) 是 B 中的一条路径。由于 N_B 的每个网状节点都与一条来自 S' 的定向边相接，路径 (u_1, v_1, p, v_2, u_2) 包含来自 S' 的两条边。由 (OB1) 和 (OB2) 可知 $u_1 \neq u_2$，且 $e = \{u_1, v_1\} \in S'$ 和 $f = \{u_2, v_2\} \in S'$。然而，这与 N_B 的构造矛盾，因为在这种情况下， $\{e, f\}$ 是 G_B 的一条边，而边 e、f 本应定向为 (u_1, v_1)、(v_2, u_2)，或定向为 (v_1, u_1)、(u_2, v_2)。（即使 e 和 f 是 G_B 中某个环的一部分，且包含它们的路径从 e 出发并以 f 结束，或反之，结论仍然成立。）因此，我们得出 N_B 满足树子条件。

对每个块重复上述过程。由此得到一个定向 N of U。我们已证明，每个块内部均满足树子条件和无环条件。此外，由于每个块的根不是网状节点，根据引理1可知， N is U的树子定向。

□

下一个推论可由命题7直接推出。

推论8. 设 U 为无根系统发育网络。若 U 是2-可割的，则它是果园网络。

推论 9. 设 U 为2-可割网络。则 U 中的网状化事件数量至多为 $|X| - 1$。

证明。 由命题7可知， U 是树子的。因此，根据定义， U 具有一个树子定向 N。由于 $|X| - 1 \geq r(N) = r(U)$，其中不等式已在 $[9, \text{命题 } 1]$ 中得到证明，等式则由从 U构造 N 得出，因此推论成立。

□

设 U 是叶集为 X的无根系统发育网络。我们称 U 是层-k 网络，如果从每个块中至多删除 k 条边，即可由 U 得到一棵树。此外，若 U 是层-k但不是层-$(k-1)$，则 U 是严格层-k 。下述观察容易看出，因为任意无根系统发育网络都可转化为一个 q-可割网络，其叶集是 X 的超集，做法是在每个环中插入一条由 q 个顶点构成的路径，且这些顶点都与不在 X 中的叶节点相邻。

观察 10。 对于所有 $k \geq 0$ 和 $q \geq 1$，都存在一个严格的 *level-k* q-可切网络。

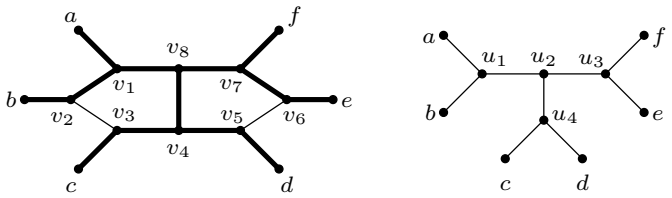


图 8: 无根系统发育网络 U（左）显示无根系统发育树 T（右）。嵌入 ϕ of T in U 中的边比不属于 ϕ的边更粗，其中 $\phi(x) = x$ 对于每个 $x \in \{a, b, c, d, e, f\}$，$\phi(u_1) = v_1$，$\phi(u_2) = v_8$，$\phi(u_3) = v_7$，$\phi(u_4) = v_4$，$\phi(u_1b) = (v_1, v_2, b)$，$\phi(u_3e) = (v_7, v_6, e)$，$\phi(u_4c) = (v_4, v_3, c)$，$\phi(u_4d) = (v_4, v_5, d)$，以及 $\phi(uu') = (\phi(u)\phi(u'))$；对于 T 中所有其他边 uu'亦然。

5 3-可切网络的无根树包含问题 s

在本节中，我们证明 Unrooted Tree Containment 对于 3-可切网络可在多项式时间内求解。

Unrooted Tree Containment (T, U)**实例**。一个定义在 X 上的无根二叉系统发育网络 U，以及一个无根二叉系统发育X-树。**问题**。T 是否由 U 显示?

设 T 是一个无根系统发育 X-树，其顶点集为 $V(T)$，边集为 $E(T)$。此外，设U 是一个定义在 X 上的无根系统发育网络，其顶点集为 $V(U)$，边集为 $E(U)$，并用 $P(U)$ 表示 U 中所有路径构成的集合。我们注意到，U 显示 T 当且仅当存在一个函数

$$\phi : V(T) \cup E(T) \to V(U) \cup P(U)$$

使得以下所有性质均成立:

(i) 对所有 $v \in V(T)$，$\phi(v) \in V(U)$；(ii) 对所有 $x \in X$，$\phi(x) = x$；(iii) 对任意两个不同的顶点 u，$v \in V(T)$，$\phi(u) \neq \phi(v)$；(iv) 对所有边 $\{u, v\} \in E(T)$，$\phi(\{u, v\})$ 是一条连接 $\phi(u)$ 与 $\phi(v)$ 的路径；(v) 对任意两条不同的边 e，$e' \in E(T)$，$\phi(e)$ 与 $\phi(e')$ 边不相交。

如果存在 ϕ，我们称其为 T 在 U 中的嵌入（或者在上下文已明确 T 和 U 时，简称为嵌入）。因此，判定 Unrooted Tree Containment (T, U) 等价于判定是否存在 T 在 U 中的嵌入。作为说明，图 8 展示了一棵由无根系统发育网络 U 显示的无根系统发育树 T，以及 T 在 U 中的嵌入。

本节的主要结果如下定理所示。

定理 11. 设 T 是一棵无根系统发育 X-树，且 U 是定义在 X 上、具有 $q \geq 3$ 的 q-可切网络。那么，Unrooted Tree Containment (T, U) 可在多项式时间内求解。

本节余下部分安排如下。第 5.1 节介绍一种操作：将其反复应用于一个 3-可切网络和一棵定义在相同叶集上的无根系统发育树，可得到一组简单的 3-可切网络和一组规模更小的无根系统发育树。随后，第 5.2 节将通过考察一棵由 U 显示的树的嵌入，确定 3-可切网络 U 中若干特定路径的性质。接着，第 5.3 节详细介绍四条归约规则，并在第 5.4 节反复使用这些规则，建立一个多项式时间算法，以求解 Unrooted Tree Containment 的实例 (T, U)，其中 U 是一个 3-可切网络；由此也证明了对于任意 $q \geq 3$，定理 11 均成立。

5.1 将一个 3-可切网络约简为一个简单的 3-可切网络 k

我们先给出一些额外的定义，并称 $X_1|X_2$ 为 X-划分，如果（X_1，X_2）是 X 的一个划分（即，$X_1 \cup X_2 = X$、$X_1 \cap X_2 = \emptyset$和 $|X_i| \geq 1$ 满足 $i \in \{1, 2\}$）。此外，如果 $|X_1| = 1$ 或 $|X_2| = 1$成立，则称 X-划分 $X_1|X_2$ 为平凡的。最后，如果四个交集 $X_1 \cap Y_1$、$X_1 \cap Y_2$、$X_2 \cap Y_1$和 $X_2 \cap Y_2$ 中至少有一个为空，则称两个 X-划分 $X_1|X_2$ 和 $Y_1|Y_2$ 是相容的。如果这四个交集都非空，则称 $X_1|X_2$ 和 $Y_1|Y_2$ 不相容

现在设 U 是 X上的一个无根系统发育网络，并设 e 是 U的一条割边。我们称 e诱导划分 $X_1|X_2$，如果 $X_1|X_2$ 是一个 X-划分，且 U 中介于 X_1 中的一个叶与 X_2中的一个叶之间的所有路径

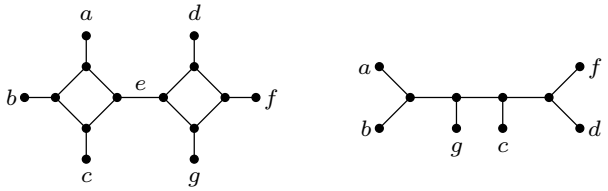


图 9：无根系统发育网络 U（左）具有由 e诱导的划分 $\{a,b,c\}|\{d,f,g\}$，而无根系统发育树 T（右）具有划分 $\{a,b,g\}|\{c,d,f\}$。因此，$(T,\ U)$具有冲突划分，而 U 不显示 T。

都经过 e。类似地，我们称 $X_1|X_2$ 是一个划分于 U，如果 U 中存在一条诱导$X_1|X_2$的割边。注意，无根系统发育树的每条边都诱导一个划分。

对于定义在 X 上的无根系统发育网络 U 和无根系统发育 X 树 T，如果存在 U中的划分 $X_1|X_2$和 T中的划分 $Y_1|Y_2$，使得 $X_1|X_2$和$Y_1|Y_2$ 不相容，则称 $(T,\ U)$具有冲突划分 。图 9 展示了冲突划分的一个例子。

下一个观察结果源于这样一个事实：由带有割边 e 的无根系统发育网络显示的每棵树，都具有一个与 e诱导的划分相等的划分。

观察 12. 设 T 是一棵无根系统发育 X-树，U 是定义在 X 上的无根系统发育网络。如果 (T,U) 存在冲突划分，则 U 不显示 T。此外，可以在多项式时间内判定 (T,U) 是否存在冲突划分。

因此，在第5节余下部分中，我们始终作如下假设。

(A1) Unrooted Tree Containment 的一个实例（T,U）不存在冲突划分。特别地，对于 U 中的每条割边 e，都存在 T 中的一条边 e'，使得 e 和 e' 诱导相同的划分。

我们还有以下结果。

引理 13. 设 U 是一个2-可切网络，e 是 U 的一条割边。那么 e 诱导一个划分。

证明. 假设 U 有一条不诱导划分的割边。那么存在 U 中的某个块 B，使得只有一条割边 uv 满足：u是 B 的一个顶点，而 v 不是 B的顶点，并且 B 中没有顶点与叶相邻。由于 B 至少包含一个环，根据 B 的选取可知，该环至多关联一条割边；这与 U 为2-可切网络矛盾。 □

引理 14. 设 U 是一个2-可切网络，且 e 和 e' 是 U 的割边。那么，e 和 e' 所诱导的划分不同。

证明. 我们首先证明，U 的每个块 B 至少有三条关联割边。如果 B 由单个环构成，那么它至少有三条关联割边，否则就会出现平行边。如果 B 不由单个环构成，那么由于 U 是 2-可切的，B 至少有两条互不相交的长度为 2 的链，因此至少有 4 条关联割边。考虑将 U 的每个块收缩为单个顶点后得到的叶标记树 T。众所周知，T 的任意两条不同割边都会诱导不同的划分 [42]。因此，U 的任意两条不同割边都会诱导不同的划分。 □

分支操作。设 T 和 U 分别是定义在 X 上的无根系统发育 X-树和 3-可切网络，且 U 不是简单的。设 $e=\{u,v\}$ 是 U 的一条非平凡割边。由于 U 是 3-可切的，根据引理 13，e 会诱导一个划分 $X_1|X_2$。对 e 进行分支操作后，得到如下两个新的实例 (T_1,U_1) 和 (T_2,U_2)，它们均属于 Unrooted Tree Containment：首先，从 U 中删除 e，并添加新边 $\{u,x_1\}$ 和 $\{v,x_2\}$，使得 $x_1,x_2\notin X$，从而得到 U_1 和 U_2。其次，设 T 中诱导划分 $X_1|X_2$ 的边为 $e'=\{u',v'\}$。然后从 T 中删除 e'，并添加新边 $\{u',x_1\}$ 和 $\{v',x_2\}$，从而得到 T_1 和 T_2。不失一般性，我们可以假设 $X_1\cup\{x_1\}$（分别为 $X_2\cup\{x_2\}$）是 T_1 和 U_1（分别为 T_2 和 U_2）的叶集。图 10 展示了分支操作的一个示例。下一个引理的证明很直接，故省略；该引理表明，判断 U 是否显示 T，等价于判断 U_1 是否显示 T_1，以及 U_2 是否显示 T_2。

引理 15。 设 T 和 U 分别是 X上的一棵无根系统发育 X-树和一个3-可割网络，且 U 不是简单的。此外，设 T_1 和 T_2是对 T 应用分支操作后得到的两棵无根系统发育树，U_1 和 U_2 是对 U应用分支操作后得到的两个无根系统发育网络；该操作均作用于 U的一条非平凡割边。那么，U 显示 T，当且仅当 U_1显示 T_1 且 U_2 显示 T_2。

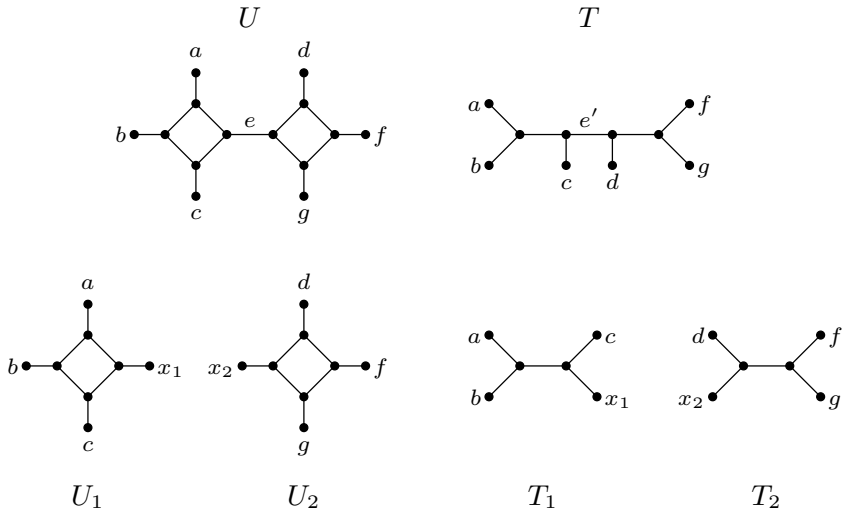


图 10：上：一个带有割边 e 的 3-可切网络 U （左）以及一棵无根系统发育树 T（右），使得 (T, U) 没有冲突划分。下：对 e 应用分支操作后得到的 3-可切网络 U_1 和 U_2（左）以及无根系统发育树 T_1 和 T_2 （右）。

除（A1）外，鉴于引理 15，在第5节余下部分中我们还可以作出下一个假设。

(A2) 设 (T, U) 是 Unrooted Tree Containment 的一个实例，且 U 是 3-可切的。那么 U 是简单的。特别地， U 中的每条割边都是平凡的，因此会导出一个平凡分裂。

5.2 纠缠路径

接下来，我们研究无根系统发育树在无根系统发育网络 U 中的嵌入所形成路径的性质。设 $P = (v_1, v_2, \ldots\ldots, v_k)$ 是 U 中的一条路径。顶点 $v_2, v_3, \ldots\ldots, v_{k-1}$ 称为 P 的内部顶点 。此外，如果 P 的所有内部顶点都不与 P 以外的割边关联，我们就称 P 为纠缠的 。

引理 16. 设 U 是 X 上一个3-可割网络，且 T 是由 U 显示的无根系统发育 X-树。此外，设 ϕ 是 T 在 U 中的一个嵌入。那么，对于 T 中的每条边 $\{u, w\}$， U 中的路径 $\phi(\{u, w\})$ 都是纠缠的。

证明。设 $e = \{u, w\}$。反设 $\phi(e)$ 并非纠缠。则 $\phi(e)$ 有一个内部顶点 v ，它与 U 的一条割边 $\{v, v'\}$ 相接，且 $\{v, v'\}$ 不是 $\phi(e)$ 的边。由于 U 显示 T ，并且根据引理 13， U 中的每条割边都会诱导一个划分，因此边 $\{v, v'\}$ 是某条边 e' （ T 中）的 $\phi(e')$ 的像。由于 ϕ 满足嵌入定义中的性质（v），且 U 是二叉的， v 是 $\phi(e')$ 的第一个或最后一个顶点。令 t 为 T 中满足 $\phi(t) = v$ 的顶点。显然， t 与 e' 和 $t \notin X$ 相接。因此， T 中存在一条边 e'' ，它也与 t 和$e'' \notin \{e, e'\}$相接。由于 U 是二叉的，这意味着 $\phi(e'')$ 和 $\phi(e)$ 有一条公共边，从而与嵌入定义中的性质（v）矛盾。 □

下一个引理表明，在简单 3-可切网络中，纠缠路径是唯一的。

引理17。 设 U 是一个简单的 3-可切网络， u 和 v 是 U 的两个顶点。那么， u 和 v 之间在 U 中至多存在一条纠缠路径；若该路径存在，则可在多项式时间内找到。

证明。 反设在 U 中， u 和 v 之间存在两条不同的纠缠路径 P_1 和 P_2 ，也就是说， P_1 中至少有一个顶点不是 P_2 的顶点。由于 P_1 和 P_2 是 u 和 v 之间两条不同的路径，恰好包含 P_1 和 P_2 的边的 U 子图中有一个环 C。如果 u 和 v 都是叶，且 C 包含 u 的邻点和 v 的邻点，那么 C 中恰有两个顶点与割边关联。其他情况下， C 至多与一条割边关联。综上，这与 U 是 3-可切网络矛盾。

接下来需要证明：如果存在 u和 v 之间的纠缠路径，那么可以在多项式时间内于 U 中找到该路径。事实上，令$E_{u,\,v}$ 为 U 中所有不与 u 或 v关联的割边组成的集合，并令$U_{u,\,v}$由 U 删除与$E_{u,\,v}$中的边关联的所有顶点后得到。那么， U 中存在 u 和 v 之间的纠缠路径，当且仅当$U_{u,\,v}$中存在 u 和 v 之间的路径。由于$U_{u,\,v}$中任意一条连接 u和 v的路径都是 U 中的纠缠路径，且可以在多项式时间内于$U_{u,\,v}$ 中找到这样的路径，因此引理得证。 □



图11：一个无根系统发育网络 U，其中 $e=\{u,v\}$ 是非割边（左图），而图 G是通过消去 e从 U 得到的。

设 U 是一个简单 3-可切网络， T 是一个无根系统发育树，且 T 由 U显示。设 ϕ是 T在 U中的一个嵌入。由引理 16 和 17，乍一看似乎 ϕ 是 T 在 U 中的唯一嵌入，因为对于 T 中的任意边 $\{u,v\}$ ， ϕ（$\{u,v\}$）都是连接 ϕ（u）和ϕ（v）的纠缠路径，并且U中任意一对顶点之间至多存在一条纠缠路径。然而，事实未必如此，因为对于 T中的一个非叶顶点 w ， U 中的 T 的两个不同嵌入 ϕ 和 ϕ' 可能满足 ϕ（w）$\neq\phi'$（w）。尽管如此，我们仍能以一种有用的方式限制 U 中 T 的可能嵌入的行为，下一条引理将对此进行说明。

引理18。 设 U 是 X上的一个简单3-可割网络，且 T 是一个无根系统发育 X-树，使得 T 由 U显示。设 T 中的 p 是一个内部顶点，其邻点为 q、 u和 v。此外，设 ϕ是 T在 U中的一个嵌入。若 q不是叶，则 $\phi(\{u,p\})$与 $\phi(\{p,v\})$的拼接是 $\phi(u)$ 与 $\phi(v)$之间的一条纠缠路径。

证明。 设 $(u_1,u_2,\ldots,u_s)$ 是路径 $\phi(\{p,u\})$，并设 $(v_1,v_2,\ldots,v_t)$ 是路径 $\phi(\{p,v\})$，其中 $u_1=v_1=\phi(p)$、$u_s=\phi(u)$ 和 $v_t=\phi(v)$。因此，这两条路径拼接而成的路径为

$$P=(u_s,u_{s-1},\ldots,u_1=v_1,v_2,\ldots,v_t).$$

令 C 为 U 中不与 $\phi(u)$ 或 $\phi(v)$ 关联的割边集合。由于 U 是简单的，回想一下，若 $\phi(u)$（分别地，$\phi(v)$）与一条割边关联，则 $\phi(u)$（分别地， $\phi(v)$）是叶。由引理 16， $\phi(\{p,u\})$ 和 $\phi(\{p,v\})$ 是 U 中的纠缠路径。此外，由于 U 是简单的，根据纠缠路径的定义可知， $\{u_2,u_3,\ldots,u_{s-1},v_2,v_3,\ldots,v_{t-1}\}$ 中没有顶点与 C 中的割边关联。因此，只需证明 U 中与 $\phi(p)$ 关联且不属于 P 的边 e 不是割边。由以下事实即可得出这一点： U 是简单的， p 不是叶，且 e 是路径 $\phi(\{p,q\})$ 的一条边。 □

5.3 归约规则

在本节中，我们提出四条归约规则。给定一个 Unrooted Tree Contain-ment 的实例 $(T,\ U)$，每条归约规则都能正确判定 $(T,\ U)$ 是是实例还是否实例，或者通过删除 U 中的一条边，从 $(T,\ U)$ 得到一个更小的实例 $(T,\ U')$，并且 $(T,\ U')$ 是是实例当且仅当 $(T,\ U)$ 也是是实例。对于每条归约规则，我们先陈述规则，然后证明其正确性。

归约规则1。 设 $(T,\ U)$ 是 Unrooted Tree Containment 的一个实例。如果 U至多有三个叶节点，则返回 Yes。

引理 19。 设 (T,U) 是 Unrooted Tree Containment 的一个实例。如果 U至多有三个叶节点，则 (T,U) 是是实例。

证明。 设 X 是 T 和 U（其根为 $|X|\leq 3$）的叶集。无论 X 的大小如何，都只有一棵无根系统发育 X-树。显然， T由 U 显示，且 $(T,\ U)$ 是是实例。 □

接下来的三条归约规则涉及删除非割边，下面我们将对其进行形式化定义。设 U是一个无根系统发育网络， $e=\{u,v\}$ 是 U的一条非割边。设 G是从 U 中删除 e并压缩由此产生的两个度为2的顶点后得到的图。我们称 G 是通过消除 e 从 U 得到的（图11展示了这一操作）。

引理20. 设 T 是一个无根系统发育 X-树， U 是 X上的一个3-可割网络，且 e是 U的一条非割边。设 U' 是从 U中消除 e后得到的图。则以下各项均成立：

(i) U' 是 X上的一个3-可割网络；

(ii) 若 U'显示 T，则 U显示 T；并且

(iii) 若 U显示 T，且存在 ϕ的一个嵌入，使得 T嵌入于 U，并且对于 T中的每条边 f， $\phi(f)$均不包含 e，则 U'显示 T。

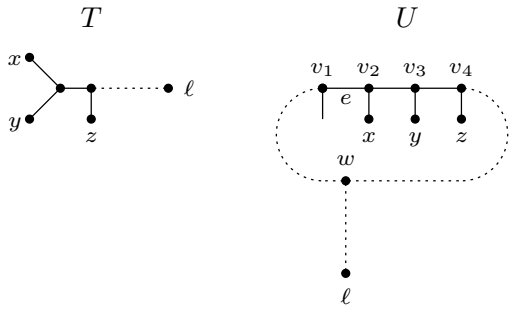


图 12：如引理 21的证明中所述的设置，用以确立规则2的正确性。实线（相应地，虚线）表示边（相应地，路径）。为清晰起见，省略了 T 和 U 的部分。

证明。设 $e=\{u,v\}$，令 u_1 和 u_2 表示 U 中与 u 相邻且不等于 v的两个邻点；类似地，令v_1 和 v_2 表示 U 中与 v 相邻且不等于 u的两个邻点。注意，$\{u_1,u_2\}$ 和 $\{v_1,v_2\}$ 是 U'中仅有的两条不属于 U的边。

首先，我们证明 U'是3-可切的。设 C'是 U'中的一个环。设 C是 U中的环，它由 C' 通过将边 $\{u_1,u_2\}$ 的任何出现替换为两条边 $\{u_1,u\}$ 和$\{u,u_2\}$，并将边 $\{v_1,v_2\}$的任何出现替换为两条边 $\{v_1,v\}$和 $\{v,v_2\}$而得到。由构造可知，C 不包含 e。此外，若 C 包含 u （相应地，v），则 u （相应地，v）在 U中不与割边关联。由于 U是3-可切的，由此可知 C包含一条路径（w_1,w_2,w_3），使得 $\{u,v\}\cap\{w_1,w_2,w_3\}=\emptyset$以及 $\{w_1,w_2,w_3\}$中的每个元素都与一条割边关联。由从 U构造U'可知，$\{w_1,w_2,w_3\}$中的顶点及其关联的割边也存在于 U'中。由此可知U' 是一个3-可切网络。这就确立了(i)。

接下来，假设 U'显示 T。那么，U'包含一个子图 G'，它同构于 T的一个细分。通过将 G'中边 $\{u_1,u_2\}$的任意出现替换为边 $\{u_1,u\}$和 $\{u,u_2\}$，并类似地将边 $\{v_1,v_2\}$的任意出现替换为边 $\{v_1,v\}$和$\{v,v_2\}$，得到 G。显然，G也是 T的一个细分，且是 U的子图，由此证明 (ii)。

最后，假设 U 显示 T，并且存在 T 在 U 中的一个嵌入 ϕ，使得对于 T中的每条边 f，$\phi(f)$都不包含 e。那么，对于 T中的每个顶点 u'，都有 $\phi(u')\neq u$，因为u' 要么是叶，要么度为 3。因此，要么存在 T 中的一条边 $\{u',v'\}$，使得 u是 $\phi(\{u,'v'\})$的内部顶点；要么对于 T中的任意一条边 $\{u',v'\}$，u都不是 $\phi(\{u',v'\})$的顶点。现在，如果u是 T中的某条边 $\{u',v'\}$对应的 $\phi(\{u',v'\})$的内部顶点，那么 $\phi(\{u',v'\})$也包含 u_1和u_2。此外，将边 $\{u_1,u\}$和 $\{u,u_2\}$替换为 $\{u_1,u_2\}$后，可得到 U'中连接$\phi(u')$与 $\phi(v')$的一条路径。再看 v，我们可以用类似的论证推得：如果 T中的某条边$\{p',q'\}$使得 v是 $\phi(\{p',q'\})$的内部顶点，那么 $\phi(\{p',q'\})$也包含 v_1和 v_2；将边 $\{v_1,v\}$和 $\{v,v_2\}$替换为 $\{v_1,v_2\}$后，可得到 U'中连接 $\phi(p')$与 $\phi(q')$的一条路径。由此可知，我们可以将 ϕ 转化为 T在 U'中的一个嵌入。这证明了 (iii)，从而证明了该引理。

□

约简规则 2。设 $(T,\ U)$ 是 Unrooted Tree Containment 的一个实例。设 X 是 T 和 U 的叶集，且满足 $|X|>3$。假设 U 有三个叶节点 x、y 和 z，以及一条路径 $P=(v_1,\ v_2,\ v_3,\ v_4)$，使得 $\{v_2,x\}$、$\{v_3,y\}$ 和 $\{v_4,z\}$ 是 U 中的割边，并且连接 P 中两个顶点的边都不是割边。此外，假设 T 有一个叶集为 $\{x,y,z\}$ 的悬挂子树，且 $\{x,y\}$是一个樱桃。通过在 U 中消去 $e=\{v_1,v_2\}$，从 $(T,\ U)$ 得到 Unrooted Tree Containment 的一个实例 $(T,\ U')$。

引理 21。设 $(T,\ U)$ 和 $(T,\ U')$ 是 Unrooted Tree Containment 的两个实例，满足 U 是一个3-可割网络，并且 $(T,\ U')$ 是通过对 $(T,\ U)$ 单次应用规则2得到的。那么$(T,\ U)$ 是是实例，当且仅当 $(T,\ U')$ 是是实例。

证明。证明过程如图 12 所示，并且我们使用与规则2中相同的术语。根据引理 20(i)，回顾可知，U'是一个 3-可切网络。假设 U显示 T。令 U_T为 U的一个子图，该子图是 T的细分。如果 U_T不包含 e，则根据引理 20(iii)，U'显示 T。因此，我们可以假设 U_T 包含 e。由于 T 具有樱桃 $\{x,y\}$，U_T 也包含边 $\{v_2,v_3\}$，但不包含边 $\{v_3,v_4\}$。令 $\ell\in X\setminus\{x,y,z\}$。此外，令 w为 U_T中的唯一顶点，使得在 U_T中存在三条两两边不相交的路径，分别从 w 连接到 ℓ、x和 z。$w=v_1$可能存在，也可能不存在。令 P'为 U_T中从 w到 v_2的路径。令 U_T^*由 U_T删除 P'的边和内部顶点，并添加边 v_3v_4得到。于是，U_T^*也是 T的一个细分。重要的是，U_T^*不使用 e。因此，我们处于前一种情况，从而根据引理 20(iii)，U'显示 T，如所需。反之，根据引理 20(ii)，如果 U' 显示 T，则 U 也显示 T。引理得证。

□

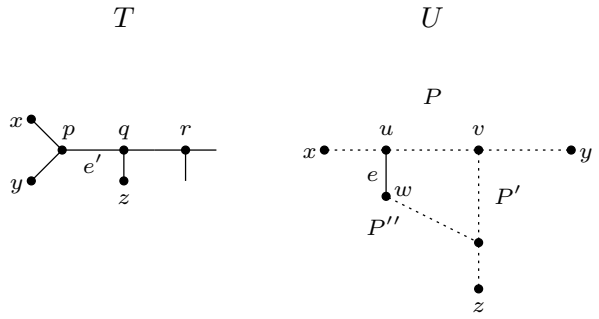


图 13：如引理22的证明中所述的设置，用以确立规则3的正确性。实线（相应地，虚线）表示边（相应地，路径）。为清晰起见，省略了 T 和 U 的部分。

继规则2的陈述之后，并且在 U 是简单且 3-可切的条件下，注意到如果 $|X| = 4$，则该规则总是适用。事实上，由于 U 只有平凡分裂，因此可知， U 是一个无根系统发育网络，其中恰好包含四个构成一个环的内部顶点，并且每个这样的顶点都与一个叶相邻。此外，在任意无根系统发育 X-树中，四个叶中的任意三个都构成一个悬挂子树。因此，规则2适用。在第5节的其余部分，我们始终可以依据规则1和规则2作出如下假设。

(A3) 设（T， U）为 Unrooted Tree Containment的一个实例。 T和 U各自至少有五个叶节点。

归约规则3。设（T， U）为 Unrooted Tree Containment 的一个实例，使得 U是简单的。设 X 为 T和 U 的叶集，且带有 $|X| > 4$。假设 T 有一个叶集为 $\{x, y, z\}$ 的悬挂子树，使得 $\{x, y\}$ 是一个樱桃。则应用以下三者之一：

（I）若 U中在 x 与 y 之间不存在纠缠路径，则返回 No。

（II）若 U中在 x 与 y 之间存在一条纠缠路径 P，但 U中在 z 与 P 的某个内部顶点之间不存在纠缠路径，则返回 No。

(III) 如果 (i) 和 (ii) 均不适用，令 P 为 U 中连接 x 与 y 的唯一纠缠路径，并令 P' 为一条连接 z 与 P 上内部顶点 v 的纠缠路径，且 P 与 P' 边不相交。然后，通过从 (T, U) 中删除 U 中的一条边 $e = \{u, w\}$，得到 Unrooted Tree Containment 的一个实例 (T, U')；其中 u 是 P 的一个顶点，满足 $u \neq v$，且 w 不是 P 的顶点。

假设 (I) 和 (II) 均不适用。为证明 u 存在，假设 v 是 P 唯一的内部顶点。那么 $\{x, y\}$ 是 U 中的一个樱桃。特别地， U 有一条不与叶相接的割边，这与 U 是简单的这一假设矛盾。此外， e 不是割边，因为 P 是纠缠的。因此，如果 (I) 和 (II) 均不适用，则 (III) 定义良好且始终适用。

引理22. 设(T, U) 是 Unrooted Tree Containment 的一个实例，且规则*2*不适用， U 是 X上的简单3-可割网络，T 有一个包含三个叶节点 x、 y和 z的悬挂子树，且 $\{x, y\}$ 是 T中的一个樱桃。则以下两个陈述成立：

(i) 如果规则 *3(I)* 或 *(II)* 适用于(T, U)，那么(T, U)是一个否实例。

(ii) 如果规则 *3(III)* 适用于(T, U)，那么(T, U)是一个是实例，当且仅当(T, U')是一个是实例 e.

证明。在 T中，设 p 是 x 和 y的公共邻居， q 是 p 和 z的公共邻居，并设r 是 q 的邻点，且不等于 p 或 z。由于 T 至少有五个叶节点， r 不是叶节点。此结构如图 13 左侧所示。

首先，我们证明 (i) 成立。假设 (T, U) 是一个是实例。那么，存在 T 在 U 中的一个嵌入 ϕ。根据引理18， $\phi(\{x, p\})$ 与 $\phi(\{p, y\})$ 的拼接是 U 中一条从 x 到 y 的纠缠路径 P^*。因此，规则 3(I) 不适用。根据引理17， P^* 是唯一的。此外，同样根据引理18， $\phi(\{p, q\})$ 与 $\phi(\{q, z\})$ 的拼接是 U 中一条连接 $\phi(p)$ 和 z 的纠缠路径，而 $\phi(p)$ 是 P^* 的一个内部顶点。因此，规则 3(II) 也不适用。由此证明了 (i)。

接下来，我们证明 (ii) 成立。证明的这一部分如图 13 所示，并且我们使用与规则 3(III) 的陈述中相同的术语。由于由引理 20(ii) 可知，若 U' 显示 T，则 U 显示 T，因此尚需证明：若 U 显示 T，则 U' 显示 T。设 U 显示 T。为了得出矛盾，假设 U' 不显示 T。令 ϕ 为 T 在 U 中的一个嵌入。由引理 20(iii)，我们可以假设 T 中存在一条边 e'，使得 $\phi(e')$ 包含 e。此外，由引理17和引理18可知， $\phi(\{x, p\})$ 与 $\phi(\{p, y\})$ 的拼接等于 P。因此，在 U 中与 u 关联的全部三条边都位于 ϕ 的像中，特别地， $\phi(p) = u$ 和 $e' = \{p, q\}$。令

P''为 $\phi(\{p,q\})$ 与 $\phi(\{q,z\})$ 的拼接。注意， P'' 至少有两条边。于是，再次由引理18，并注意到 r 是一个内部顶点，可知 P'' 是 U 中介于 u 与 z 之间的一条纠缠路径。此外，由 ϕ 的定义可知， P 与 P'' 边不相交。现在考虑由三条纠缠路径 P、 P' 和 P'' 的边所诱导的 U 的子图G。由于 z 是 P' 和 P'' 的终端顶点，这两条路径相交。由此可知， G 是连通的，包含一个环 C，并且在 U 中不存在割边 st，使得 s 是 G 的一个顶点而 t 不是 G 的一个顶点。回想 U 是简单且 3-可切的。由于 C 也是 U 的一个环，在 U 中存在一条由三个顶点 s_1、 s_2、 s_3 组成的路径，使得每个满足 $i \in \{1,2,3\}$ 的 s_i 都与一条割边 $\{s_i\ell_i\}$ 和 $\{\ell_1,\ell_2,\ell_3\} = \{x,y,z\}$ 相关联。由于P和 P'' 的边都在 ϕ 和 $|X| > 4$ 的像中，在 P'' 上存在一个内部顶点，其三条关联边全都在 ϕ 的像中，并且这三条边都不与 x 或 y 相关联。因此 $\ell_2 \neq z$。在互换 x 与 y 的角色的意义下，现在可知规则2适用于 (T,U)；矛盾。这就证明了(ii)，从而证明了该引理。 □

承接规则3的陈述，可以观察到，若 $|X|$=5，则 T 总是诱导出一个划分$X_1|X_2$，其中 $|X_1| = 3$ 且 $|X_2| = 2$。由此可知， T 具有一个含三个叶节点的悬挂子树，且规则3适用。因此，综合规则1、2和3，我们可以在第5节的其余部分始终作如下假设。

(A4) 设（T， U）为 Unrooted Tree Containment的一个实例。 T和 U各自至少有六个叶节点。

归约规则4。设（T， U）为 Unrooted Tree Containment 的一个实例，使得 U是简单的。设 X 为 T的叶集，且 U 满足 $|X| > 5$。假设 T具有一个叶集为 $\{w,x,y,z\}$ 的悬挂子树，使得 $\{x,y\}$ 和 $\{w,z\}$ 是两个樱桃。则应用以下四者之一：

(I) 如果在 U 中， x 与 y 之间不存在纠缠路径，或 w 与 z 之间不存在纠缠路径，则返回 No。(II) 如果在 U 中，存在一条连接 x 与 y 的纠缠路径 P_1，以及一条连接 w 与 z 的纠缠路径 P_2，但 P_1 与 P_2 不边不相交，则返回 No。(III) 如果在U 中，存在一条连接 x 与 y 的纠缠路径 P_1，以及一条连接 w 与 z 的纠缠路径 P_2，且 P_1 与 P_2 边不相交，但不存在一条从 P_1 的某个内部顶点到 P_2 的某个内部顶点的纠缠路径，该路径至少包含两条边，且与P_1 和 P_2分别边不相交，则返回 No。(IV) 如果 (I)–(III) 均不适用，令 P_1 为 U 中连接 x 与 y 的一条纠缠路径，令 P_2 为 U 中连接 w 与 z 的一条纠缠路径，并令 P_3 为一条纠缠路径，连接 P_1的某个内部顶点 v_1 与 P_2 的某个内部顶点 v_2，且 P_3 至少包含两条边，并分别与 P_1 和 P_2 边不相交。然后，通过消去 U 中的一条边 $e = \{u,w\}$，从 (T,U) 得到 Unrooted Tree Containment 的一个实例 (T,U')，其中 u 是 P_1 的一个顶点，且 $u \neq v_1$，而 w 不是 P_1的顶点。

假设 (I)–(III) 均不适用。与规则3陈述之后的段落类似，可知 e 存在，且 (IV) 适用。

引理23。设$(T$、$U)$是 Unrooted Tree Containment 的一个实例，使得规则*2*不适用， U 是 X上的一个简单 3-可割网络，并且 T 有一个包含四个叶节点 w、 x、 y 和z 的悬挂子树，其中 $\{x,y\}$ 和 $\{w,z\}$ 是 T中的两个樱桃。则以下两个陈述成立：

(i) 如果规则*4(I)*、*(II)*或*(III)*适用于$(T$、$U)$，则$(T$、$U)$是否实例。

(ii) 如果规则 *4(IV)*适用于$(T$、$U)$，则$(T$、$U)$是是实例，当且仅当$(T$、$U')$是是实例 e.

证明。证明采用与引理22类似的策略。在 T中，设 p_1 为 x 和 y的公共邻居，设 p_2 为 w 和 z的公共邻点，设 q 为 p_1 和 p_2的公共邻居，并设 r 为 q 的、不是 p_1 或 p_2的邻点。由于 T 至少有六个叶节点，因此 r 不是叶节点。该设置如图14左侧所示。

首先，我们证明 (i) 成立。假设 (T,U) 是一个是实例。那么，存在 U 中 T 的一个嵌入ϕ。由引理18可知， $\phi(\{x,p_1\})$ 与 $\phi(\{p_1,y\})$ 的拼接是 U 中一条从x 到 y的纠缠路径 P；同样由该引理可知， $\phi(\{w,p_2\})$ 与 $\phi(\{p_2,z\})$ 的拼接是 U 中一条从 w到 z的纠缠路径 P'。因此，规则4(I) 不适用。此外，由于根据引理17， P和 P'是唯一的，根据 ϕ 的定义可知， P和 P' 也边不相交。因此，规则4(II) 也不适用。最后，再次应用引理18可知， $\phi(\{p_1,q\})$ 与 $\phi(\{q,p_2\})$ 的拼接是 U 中一条从 P上的内部顶点 $\phi(p_1)$ 到 P'上的内部顶点 to$\phi(p_2)$ 的纠缠路径 P''。此外，由于 P''是两条边不相交路径的拼接， P''至少有两条边。这意味着规则4(III) 也不适用。综上，(i) 得证。

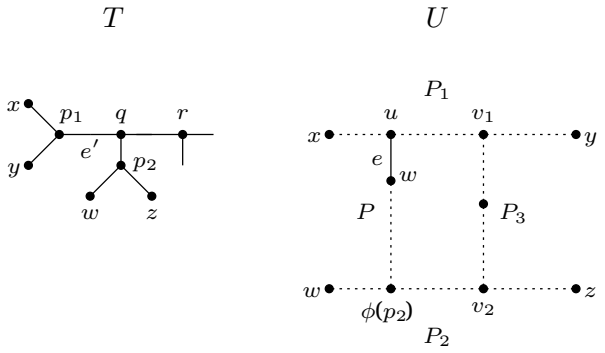


图14：如引理23的证明中所述，为证明规则4的正确性所作的设置。注意，$u = \phi(p_1)$。实线（分别为点线）表示边（分别为路径）。为清晰起见，省略了 T 和 U 的部分内容。

接下来，我们证明 (ii) 成立。证明过程见图14，且我们使用规则4(IV) 的陈述中的相同术语。由于由引理 20(ii) 可知，若 U' 显示 T，则 U 也显示 T，因此只需证明：若 U 显示 T，则 U' 也显示 T。假设 U 显示 T。为导出矛盾，假设 U'不显示 T。令 ϕ为 U 中 T 的一个嵌入。由引理 20(iii)，不妨设 T 中存在一条边 e'，使得 $\phi(e')$ 包含 e。此外，由引理17和引理18可知，$\phi(\{x, p_1\})$ 与 $\phi(\{p_1, y\})$ 的拼接等于 P_1，且 $\phi(\{w, p_2\})$ 与 $\phi(\{p_2, z\})$ 的拼接等于 P_2。因此，在 U 中与u 相接的三条边都属于 ϕ 的像，特别是 $\phi(p_1) = u$ 和 $e' = \{p_1, q\}$。令 P为 $\phi(\{p_1, q\})$ 与 $\phi(\{q, p_2\})$ 的拼接。显然，P 至少有两条边。此外，由引理18可知，P 是 U 中一条从 u 到 $\phi(p_2)$ 的纠缠路径，而 $\phi(p_2)$ 是 P_2的一个内部顶点。现在考虑 U 中由 P、P_1、P_2和 P_3的边诱导的子图 G。由于 $\phi(p_2)$ 位于 P_2 上，且 $\{P, P_1, P_2, P_3\}$ 中的每个元素都是 U中的一条纠缠路径，因此 G是连通的，包含一个环 C；并且不存在 U 中的割边 $\{s, t\}$，使得 s 是 G 的一个顶点，而 t 不是 G的顶点。注意，U 是简单的。因此，U 中与割边相接的 C 的每个顶点，都是$\{w, x, y, z\}$中某个元素的邻点。由于 P和 P_3各自至少有两条边，且 P（分别为 P_3）上存在一条不属于 P_3（分别为 P）的边，因此很容易验证，U 中不存在由三个顶点s_1、s_2、s_3 构成的路径，使得每个满足 $i \in \{1, 2, 3\}$ 的 s_i 都与一条割边相接，从而与U 是3-可切的这一点矛盾。由此 (ii) 得证，进而引理得证。

□

本节最后给出一个引理，其证明显而易见，故省略。该引理表明，具有至少四个叶节点的无根系统发育树 T 总是含有某种悬挂子树。因此，由该引理可知，规则 2–4 中总有一条适用于 T。

引理 24. 设 T 是具有 $|X| \geq 4$的无根系统发育 X树。那么，T 或者有一个包含三个叶节点 x、y和 z 的悬挂子树，使得 $\{x, y\}$ 是一个樱桃；或者有一个包含四个叶节点 w、x、y和z 的悬挂子树，使得 $\{x, y\}$ 和 $\{w, z\}$ 是樱桃。

5.4 算法

在本节中，我们给出算法 3-CuttableTC(T, U) 的伪代码，该算法求解 Unrooted Tree Containment 的一个实例 (T, U)，其中 U 是一个 3-可切网络，并证明该算法是正确的且在多项式时间内运行。应用第 5.1–5.3 节中建立的结果和归约规则，算法首先检查 (T, U) 是否具有冲突划分。如果存在这样的划分，则根据观察 12，U 不显示 T。如果不存在这样的划分，算法接着检查U 是否为简单。如果不是，则它对 U 的一条非平凡割边应用分支操作，并对由此得到的两个较小实例递归调用自身。如果不存在这样的割边，则继续应用规则 1–4 之一。如果规则 1 适用，则算法停止并返回 Yes。否则，算法应用规则 2–4 之一，这总是得到一个边数少于 U 的 3-可切网络，并对 T 和所得网络递归调用自身。3-CuttableTC(T, U) 的详细伪代码见算法 1。

引理 25。 设(T, U) 是 Unrooted Tree Containment 的一个实例，其中 U 是 X 上的一个3-可割网络，而 T 是一棵无根系统发育 X- 树。算法 3-CuttableTC 会根据 T 是否由 U 显示，正确地返回 Yes 或 No，并且运行时间为多项式时间。

证明。由观察 12 以及引理 15、19 和 21–23 可知，3-CuttableTC 当且仅当 U 显示 T 时返回 Yes。

我们完 成证明，只需说明该算法的运行时间为多项式时间。单次调用 3-CuttableTC 返回 Yes 或 No，在U 的非平凡割边上执行分支操作

算法 1： 3-CuttableTC（T，U）

输入： 一个无根系统发育 X-树 T 和一个定义在 X 上的 3-可切网络 U。
结果： Yes 若 U 显示 T；否则为 No。
若 (T, U) 具有冲突划分 **则**
  返回 No; //观察 12
若 U 具有一条非平凡割边 e**则**
  (T_1, U_1) 和 $(T_2, U_2) \leftarrow$ 对 e 应用分支操作;
  若 3-CuttableTC(T_1, U_1) 且 3-CuttableTC(T_2, U_2) 返回 Yes **则**
    返回 Yes ; //引理 15
  else
    返回 No ; //引理 15
如果 $|X| \leq 3$ **那么**
  返回 Yes; //引理 19
如果 规则*2*适用于(T, U) **那么**
  $(T, U') \leftarrow$将规则2应用于 (T, U);
否则，如果规则*3*适用于(T, U) **则**
  $(T, U') \leftarrow$对(T, U)应用规则3;
否则，如果规则*4*适用于(T, U) **则**
  $(T, U') \leftarrow$对(T, U)应用规则4;
返回 3-CuttableTC(T, U')

并对由此产生的两个较小实例分别递归调用自身，或者应用规则 2–4 中的一条并递归调用自身一次。事实上，如果 U 有一条非平凡割边 e，那么如果（T，U）存在冲突划分，算法就返回No；如果（T，U）不存在冲突划分，算法则在 e 上执行分支操作。现在，假设 U 中的每条割边都是平凡的。首先，如果 $|X| \leq 3$，算法就返回 Yes。其次，如果 $|X| = 4$，则应用规则2（见引理 21 证明之后的段落）。第三，如果 $|X| \geq 5$，则根据引理 24 应用规则 2–4 中的一条。假设 U 只有平凡割边，由此可知算法在多项式时间内运行，因为每次应用规则 2–4 都会减少 U 中的边数，而且由于 U 中任意两个顶点之间的纠缠路径唯一，并且根据引理17可在多项式时间内找到，每次应用规则 2–4 所需的时间也是多项式时间。

因此，我们可以假设 U 有一条非平凡割边。令 s 为 U中的非平凡割边数量。由于 T有 $|X| - 3$ 条边不与叶相邻，根据假设（A1）以及引理13和14，可得 $s \leq |X| - 3$。令 e 为 U中的一条非平凡割边。在 e 上应用分支操作会产生两个实例(T_1, U_1)和(T_2, U_2)。此外，U_1 和 U_2 中的非平凡割边总数为 $s-1$。因此，与递归调用相关联的递归树（这些递归调用由反复应用分支操作产生，直至 U 被化简为一系列简单 3-可切网络）恰好有 s 个非叶顶点。由此可知，该搜索树的规模关于 $|X|$是多项式级的。我们得出结论，3-CuttableTC(T, U)的总运行时间为多项式级。

□

由于任何满足 $q \geq 4$ 的 q-可切网络尤其也是 3-可切的，因此定理 11 可直接由引理 25 得出。

6 讨论

我们证明了，判定给定的无根二叉系统发育网络是否能定向为有根树子网络是 NP困难的。因此，可定向为树子网络的网络这一类可能并不是无根网络的一个理想子类，因为除非 P = NP，否则无法在多项式时间内判定给定网络是否属于该类。

因此，我们引入了 q-可割网络这一类，并证明它具有若干理想性质，如下表所示。

特征	***q*-可切**
多项式时间可识别	定理 6
Unrooted Tree Containment 可在多项式时间内求解	定理 11，对于 $q \geq 3$
由Blob决定	定义 1
大小与叶节点数量呈线性关系	推论 9，对于 $q \geq 2$
包含任意层级的网络	观察 10

因此，我们认为，q-可割网络这一类值得进一步研究，并且可能在无根情形中发挥类似于树子网络在有根情形中持续发挥的作用。未来仍有一些值得探讨的问题。首先，除了 Unrooted Tree Containment之外，是否还有其他 NP困难问题，在限制为某个（较小）q值的 q-可割网络时，可以在多项式时间内求解？其次，对于某个 q，q-可割网络是否可由某些子结构编码，例如分类单元间距离、四元组（嵌入的四叶树）或四元网络（嵌入的四叶网络）？

致谢。 本文基于作者在罗得岛州普罗维登斯的数学计算与实验研究所驻留期间，参加定量系统发育基因组学的理论、方法与应用 学期项目时获得美国国家科学基金会资助（项目编号：DMS-1929284）所开展的工作。前两位作者获得荷兰研究委员会（NWO）资助（项目编号：OCENW.M.21.306）。第三位作者获得新西兰马斯登基金资助；该基金来自政府拨款，由新西兰皇家学会 Te Ap¯arangi 管理。最后一位作者获得加拿大自然科学与工程研究委员会（NSERC）探索基金资助，项目编号为 RGPIN-2025-06235。

参考文献

[1] E. S. Allman、C. An´e、H. Banos与J. A. Rhodes。超越一级：一类仙人掌树子网络的可识别性。数学生物学通报，87:166，2025。[2] P. Berman、M. Karpinski与A. D. Scott。MAX-3-SAT短对称实例的近似难度。计算复杂性电子讨论会，报告第49号，2003。[3] F. Bienvenu、A. Lambert与M. Steel。排序树子网络的组合与随机性质。随机结构与算法，60:653–689，2022。[4] M. Bordewich与N. Tokac。由分类单元间距离重构超度量树子网络的算法。离散应用数学，213:47–59，2016。[5] M. Bordewich、K. T. Huber、V. Moulton与C. Semple。由最短分类单元间距离信息恢复正规网络。数学生物学杂志，77:571–594，2018。[6] M. Bordewich、C. Semple与N. Tokac。由距离矩阵构造树子网络。Algorithmica，80:2240–2259，2018。[7] L. Bulteau、M. Weller与L. Zhang。论将图转化为系统发育网络。HAL开放科学，hal-04085424，2023。[8] G. Cardona与L. Zhang。树子网络及其子类的计数与枚举。计算机与系统科学杂志，114:84–104，2020。[9]G. Cardona、F. Rossell´o与G. Valiente。树子系统发育网络的比较。IEEE/ACM计算生物学与生物信息学汇刊，6:552–569，2008。[10] G. Cardona、M. Llabr´es、F. Rossell´o与G. Valiente。论约简系统发育网络的Nakhleh度量。IEEE/ACM计算生物学与生物信息学汇刊，6:629–638，2009。[11] G. Cardona、J. C. Pons与C. Scornavacca。二叉树子系统发育网络的生成。PLoS计算生物学，15:e1007347，2019。[12] G. Cardona、G. Ribas与J. C. Pons。果园网络与树子网络的生成。数学生物学通报，86:10，2024。[13] Y.-B. Chan与C. Robin。基因网络与物种树的调和。理论生物学杂志，472:54–66，2019。[14] J. Dempsey、L. v. Iersel、M. Jones、Y. Murakami与N. Zeh。果园中的野羊追逐。arXiv预印本 arXiv:2408.10769，2024。[15] J. D¨ocker与S. Linz。刻画有根与无根树子网络。arXiv预印本 arXiv:2408.17105，2024。[16] J. D¨ocker与S. Linz。论若干类有根系统发育网络的漏斗定向的存在性。arXiv预印本 arXiv:2401.05611，2024。[17] J. D¨ocker与S. Linz。论若干类有根系统发育网络的漏斗定向的存在性。理论计算机科学，1023:114908，2025。

[18] A. Francis与V. Moulton。概率重组-突变进化模型下树子系统发育网络的可识别性。理论生物学杂志，446:160–167，2018。[19] A. Francis、K. T. Huber与V. Moulton。基于树的无根系统发育网络。数学生物学通报，80:404–416，2018。[20] M. Frohn与S. Kelk。二叉树子系统发育网络上软接线简约问题的2-近似算法。运筹学年鉴，345:125–145，2025。[21] M. Fuchs、G.-R. Yu与L. Zhang。论树子网络数目的渐近增长。欧洲组合学杂志，93:103278，2021。[22] P. Gambette、V. Berry与C. Paul。四元组与无根系统发育网络。生物信息学与计算生物学杂志，10:1250004，2012。[23] P. Gambette、K. T. Huber与G. E. Scholz。去根系统发育网络。数学生物学通报，79:2022–2048，2017。[24] P. Gambette、A. D. Gunawan、A. Labarre、S. Vialette与L. Zhang。以线性时间求解近稳定系统发育网络的树包含问题。离散应用数学，246: 62–79，2018。[25] J. Garvardt、M. Renken、J. Schestag与M. Weller。寻找度约束无环定向。载于第18届参数化与精确计算国际研讨会（IPEC 2023）。达格斯图尔–Leibniz-Zentrum f¨ur Informatik，2023。[26] K. T. Huber、L. v. Iersel、R. Janssen、M. Jones、V. Moulton、Y. Murakami与C. Semple。为无向系统发育网络定向。计算机与系统科学杂志，140:103480，2024。[27] L. v. Iersel与V. Moulton。三元网络编码树子网络与二级系统发育网络。数学生物学杂志，68:1707–1729，2014。[28] L. v. Iersel、C. Semple与M. Steel。在系统发育网络中定位一棵树。信息处理快报，110:1037–1043，2010。[29] L. v. Iersel、S. Kelk、G. Stamoulis、L. Stougie与O. Boes。论若干著名系统发育网络问题的无根与根不确定变体。Algorithmica，80:2993–3022，2018。[30] L. v. Iersel、R. Janssen、M. Jones、Y. Murakami与N. Zeh。由多棵二叉树构造树子网络的实用固定参数算法。Algorithmica，84:917–960，2022。[31] R. Janssen与Y. Murakami。树子网络包含的线性时间算法。载于计算生物学算法国际会议，第93–107页。施普林格，2020。[32] R. Janssen与Y. Murakami。论樱桃挑选与网络包含。理论计算机科学，856:121–150，2021。[33] R. Janssen、M. Jones、P. L. Erd˝os、L. v. Iersel与C. Scornavacca。利用尾移动探索有根系统发育网络空间的层级。数学生物学通报，80:2177–2208，2018。[34] I. A. Kanj、L. Nakhleh、C. Than与G. Xia。在网络中辨识树及其分支是困难的。理论计算机科学，401:153–164，2008。[35] S. Kong、J. C. Pons、L. Kubatko与K. Wicke。显式系统发育网络的类别及其生物学与数学意义。数学生物学杂志，84:47，2022。[36] S. Linz与C. Semple。通过添加叶节点与挑选樱桃刻画一组系统发育的杂交数。应用数学进展，105:102–129，2019。[37] C. McDiarmid、C. Semple与D. Welsh。系统发育网络的计数。组合学年鉴，19:205–224，2015。[38] D. A. Morrison。系统发育分析中的网络：种群生物学的新工具。国际寄生虫学杂志，35:567–582，2005。[39] V. Moulton与A. Spillner。排序树子网络的空间。数学生物学杂志，91:1–26，2025。[40] Y. Murakami、L. v. Iersel、R. Janssen、M. Jones与V. Moulton。由删除网状边的子网络重构树子网络。数学生物学通报，81:3823–3863，2019。

[41] C. Semple. 每棵嵌入系统发育树都是基树的系统发育网络。数学生物学通报，78:132–137，2016。[42] C. Semple 和 M. Steel。系统发育学。牛津大学出版社，2003。[43] C. Semple 和 K. Wicke。树子网络所显示的系统发育树数量的紧下界。*arXiv* 预印本 *arXiv:2508.13414*，2025。[44] T. Urata、M. Yokoyama 和 M. Hayamizu。将无向系统发育网络定向为目标类别的可定向性：实用算法及其在树子定向中的应用。载于第*24*届国际生物信息学算法研讨会（*WABI 2024*）。达格斯图尔–莱布尼茨信息学中心，2024。