Advances in Applied Mathematics 78 (2016) 114–141

Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama

Reticulation-visible networks ☆

Magnus Bordewich [a,*], Charles Semple [b]

[a] School of Engineering and Computing Sciences, Durham University, Durham DH1 3LE, United Kingdom

[b] Biomathematics Research Centre, School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand

ARTICLE INFO

Article history:
Received 12 August 2015
Received in revised form 13 April 2016
Accepted 13 April 2016
Available online 27 April 2016

MSC:
05C85
68R10

Keywords:
Phylogenetic network
Reticulation-visible network
Tree Containment problem

ABSTRACT

Let X be a finite set, $\mathcal{N}$ be a reticulation-visible network on X, and $\mathcal{T}$ be a rooted binary phylogenetic tree. We show that there is a polynomial-time algorithm for deciding whether or not $\mathcal{N}$ displays $\mathcal{T}$. Furthermore, for all $|X| \geq 1$, we show that $\mathcal{N}$ has at most $8|X| - 7$ vertices in total and at most $3|X| - 3$ reticulation vertices, and that these upper bounds are sharp.

1. Introduction

Phylogenetic networks have become increasingly more prominent in the literature as they correctly allow the evolution of certain collections of present-day species to be

☆ Part of this work was conducted while the first author was an Erskine Visiting Fellow at the University of Canterbury, New Zealand. The second author was supported by the Allan Wilson Centre, and the New Zealand Marsden Fund.

* Corresponding author.

E-mail addresses: m.j.r.bordewich@durham.ac.uk (M. Bordewich), charles.semple@canterbury.ac.nz (C. Semple).

http://dx.doi.org/10.1016/j.aam.2016.04.004
0196-8858/

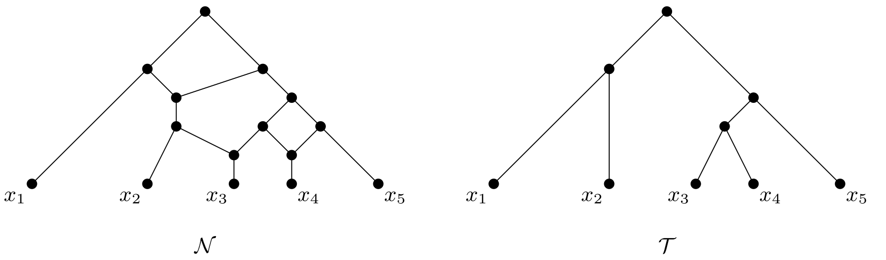

Fig. 1. The phylogenetic network $\mathcal{N}$ displays the rooted binary phylogenetic tree $\mathcal{T}$.

described with reticulation (non-tree-like) events. However, the evolution of a particular gene can generally be described without reticulation events. As a result, analysing the tree-like information in a phylogenetic network has become a common task. Central to this task is that of deciding if a given phylogenetic network $\mathcal{N}$ infers a given rooted binary phylogenetic tree on the same collection of taxa. In this paper, we show that if $\mathcal{N}$ is a so-called reticulation-visible network, then there is a polynomial-time algorithm for making this decision. This resolves a problem left open in [2] and [4]. In the rest of the introduction, we formally state this result as well as the other main result which concerns the number of vertices in a reticulation-visible network.

Throughout the paper, X denotes a non-empty finite set. A *phylogenetic network on* X is a rooted acyclic digraph with no parallel arcs and the following properties:

(i) the root has out-degree two;
(ii) vertices of out-degree zero have in-degree one, and the set of vertices with out-degree zero is X; and
(iii) all other vertices either have in-degree one and out-degree two, or in-degree two and out-degree one.

For technical reasons, if $|X| = 1$, then we additionally allow $\mathcal{N}$ to consist of the single vertex in X. The vertices in $\mathcal{N}$ of out-degree zero are called *leaves*. Furthermore, the vertices of $\mathcal{N}$ with in-degree two and out-degree one are called *reticulations*, while the vertices of in-degree one and out-degree two are called *tree vertices*. The arcs directed into a reticulation are *reticulation arcs*; all other arcs are called *tree arcs*. Note that, what we have called a phylogenetic network is sometimes referred to as a *binary* phylogenetic network. A *rooted binary phylogenetic X-tree* is a phylogenetic network on X with no reticulations.

Let $\mathcal{N}$ be a phylogenetic network on X and let $\mathcal{T}$ be a rooted binary phylogenetic X-tree. We say that $\mathcal{N}$ *displays* $\mathcal{T}$ if $\mathcal{T}$ can be obtained from $\mathcal{N}$ by deleting arcs and vertices, and contracting degree-two vertices. To illustrate, in Fig. 1, the phylogenetic network $\mathcal{N}$ on $X = \{x_1, x_2, x_3, x_4, x_5\}$ displays the rooted binary phylogenetic X-tree $\mathcal{T}$. The particular problem of interest is the following: