ON THE MAXIMUM AGREEMENT SUBTREE CONJECTURE FOR BALANCED TREES

MAGNUS BORDEWICH*, SIMONE LINZ†, MEGAN OWEN‡, KATHERINE ST. JOHN§, CHARLES SEMPLE¶, AND KRISTINA WICKE‖

Abstract. We give a counterexample to the conjecture of Martin and Thatte that two balanced rooted binary leaf-labelled trees on n leaves have a maximum agreement subtree (MAST) of size at least $n^{\frac{1}{2}}$. In particular, we show that for any $c > 0$, there exist two balanced rooted binary leaf-labelled trees on n leaves such that any MAST for these two trees has size less than $cn^{\frac{1}{2}}$. We also improve the lower bound of the size of such a MAST to $n^{\frac{1}{6}}$.

Keywords: balanced tree, caterpillar, maximum agreement subtree, phylogenetic tree

1. Introduction. Leaf-labelled trees, also called phylogenetic trees, are used to represent inferred evolutionary histories of sets of species or genes [18]. Due to different data sets and methods that are used to infer such histories, even phylogenetic trees that have been reconstructed for the same set of species or genes often differ. To quantify the dissimilarities between trees, several tree metrics are commonly used that compute distances between two phylogenetic trees (for a recent review, see [19]). Moreover, to summarise the information that two or more phylogenetic trees have in common, maximum agreement subtrees (MASTs) have become a popular tool. Historically, the concept of a MAST was introduced by Finden and Gordon [8] as a way of measuring similarities among an arbitrary number of phylogenetic trees. Since their introduction, MASTs as well as extensions of the concept of a MAST have also been used in related areas of phylogenetics such as the computation of consensus trees and supertrees [3, 4, 15]. Intuitively, an agreement subtree for a collection $\mathcal{P}$ of phylogenetic trees is a leaf-labelled tree $\mathcal{M}$ that can be embedded in each tree in $\mathcal{P}$. If $\mathcal{M}$ has the maximum number of leaves over all agreement subtrees for $\mathcal{P}$, then $\mathcal{M}$ is a MAST (formal definitions are given below).

Various aspects of MASTs have been well studied over the years. While the problem of computing a MAST for at least three phylogenetic trees is NP-hard in general [1], several polynomial-time algorithms have been developed to compute a MAST for two binary phylogenetic trees [6, 9, 21]. In terms of a lower bound on the size of a MAST, Martin and Thatte [13] have shown that two unrooted binary phylogenetic trees on n leaves have a MAST on $\Omega(\sqrt{\log n})$ leaves. Their result improves on a lower bound of $\Omega(\log \log n)$ leaves that was previously established by Steel and Székely [20]. Markin [12] has recently closed the gap between lower and upper bound asymptotics by showing that the minimum MAST for two unrooted binary phylogenetic trees has $\Theta(\log n)$ leaves.

The size of a MAST of two phylogenetic trees $\mathcal{S}$ and $\mathcal{T}$ is influenced by the shapes of $\mathcal{S}$ and $\mathcal{T}$, where the shape refers to the two trees obtained from $\mathcal{S}$ and $\mathcal{T}$ by ignoring their leaf labels. For instance, if $\mathcal{S}$ and $\mathcal{T}$ are rooted binary caterpillar trees on the same set of $n \geq 2$ leaves with their leaf labels reversed (we make this more precise in Section 2), they will only have a MAST of size two independent of n. We provide a formal justification for this statement in Lemma 3.1. Rooted binary caterpillar trees are usually considered to be the most 'imbalanced' trees among all rooted binary phylogenetic trees, and it is natural to ask how the size of a MAST changes with an increasing balance of the two rooted phylogenetic trees under consideration. The balance of a phylogenetic tree can, for example, be quantified by tree balance indices such as the Colless [7] or Sackin [16] index. In this regard, Martin and Thatte [13] have investigated the size of a MAST for two balanced rooted binary phylogenetic trees on $n = 2^m$ leaves, which—as the name suggests—are considered to be the most 'balanced' phylogenetic trees among all rooted binary phylogenetic trees. In particular, they have shown that a MAST for two such trees has at least $2^{\beta m}$ leaves, where $\beta \sim 0.149$ and, subsequently,

*Department of Computer Science, Durham University, United Kingdom.
†School of Computer Science, University of Auckland, Auckland, New Zealand.
‡Department of Mathematics, Lehman College, City University of New York, United States; Division of Invertebrate Zoology, American Museum of Natural History, New York, United States
§Department of Computer Science, Hunter College, City University of New York, United States; Division of Invertebrate Zoology, American Museum of Natural History, New York, United States
¶School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand.
‖Institute of Mathematics and Computer Science, University of Greifswald, Greifswald, Germany; Department of Mathematics, The Ohio State University, Ohio, United States

conjectured that two balanced rooted binary phylogenetic trees on n leaves have a MAST with at least $n^{\frac{1}{2}}$ leaves. The main result of this paper, disproves their conjecture for infinitely many values of n via a family of counterexamples.

For completeness, it is also worth mentioning that lower and upper bounds on the expected size of a MAST for two phylogenetic trees that are generated under the uniform or Yule-Harding distribution have been established in [2, 5, 14]. Interestingly, in the context of this paper, Misra and Sullivant [14] have shown that the expected size of a MAST for two balanced rooted binary phylogenetic trees on n leaves is $\Theta(n^{\frac{1}{2}})$ under the uniform distribution.

To formally state the main results, we require some terminology. A *rooted phylogenetic X-tree* $\mathcal{T}$ is a rooted tree with leaf set X and with no degree-two vertices, except for the root which has degree at least two. For technical reasons, if $|X| = 1$, we additionally allow $\mathcal{T}$ to consist of the single vertex in X, in which case, this vertex is the root as well as the leaf of $\mathcal{T}$. The *size* of $\mathcal{T}$ is $|X|$. If $|X| = 1$ or $\mathcal{T}$ has the property that the root has degree two and all other interior vertices have degree three, then $\mathcal{T}$ is *binary*. Furthermore, the *height* of $\mathcal{T}$ is the number of edges on the longest path from the root to a leaf. If $\mathcal{T}$ is binary, we say $\mathcal{T}$ is *balanced* if the size of $\mathcal{T}$ is 2^m for some non-negative integer m and the height of $\mathcal{T}$ is m.

Let $\mathcal{T}$ be a rooted binary phylogenetic X-tree, and let Y be a subset of X. Then, the *restriction of $\mathcal{T}$ to Y*, denoted by $\mathcal{T}|Y$, is the rooted phylogenetic Y-tree obtained from the minimal subtree of $\mathcal{T}$ that connects all leaves in Y by suppressing all non-root degree-two vertices. Now let $\mathcal{S}$ be a rooted binary phylogenetic X'-tree. If Y is a subset of $X \cap X'$ such that $\mathcal{S}|Y$ and $\mathcal{T}|Y$ are isomorphic, we call $\mathcal{S}|Y$ an *agreement subtree* of $\mathcal{S}$ and $\mathcal{T}$. If, amongst all agreement subtrees of $\mathcal{S}$ and $\mathcal{T}$, the restriction $\mathcal{S}|Y$ is of maximum size, then it is called a *maximum agreement subtree* (MAST) of $\mathcal{S}$ and $\mathcal{T}$, and $|Y|$ is denoted by $\mathrm{mast}(\mathcal{S}, \mathcal{T})$.

The first main result of the paper, Theorem 1.1, shows that two balanced rooted binary phylogenetic trees on n leaves do not necessarily have a MAST of size at least $n^{\frac{1}{2}}$.

THEOREM 1.1. *For any $c > 0$, there exist balanced rooted binary phylogenetic X-trees $\mathcal{S}$ and $\mathcal{T}$, where $n = |X|$, such that* $\mathrm{mast}(\mathcal{S}, \mathcal{T}) < cn^{\frac{1}{2}}$.

Theorem 1.1 disproves the aforementioned conjecture by Martin and Thatte [13, Conjecture 20] that we state next.

CONJECTURE 1.2 (Martin and Thatte [13]). *If $\mathcal{S}$ and $\mathcal{T}$ are two balanced rooted binary phylogenetic X-trees, where $n = |X|$, then* $\mathrm{mast}(\mathcal{S}, \mathcal{T}) \geq n^{\frac{1}{2}}$.

The second main result, Theorem 1.3, slightly improves the lower bound on the size of a MAST for a pair of balanced rooted binary phylogenetic trees given in [13] from n^{β}, where $\beta \sim 0.149$, to $n^{0.17}$.

THEOREM 1.3. *If $\mathcal{S}$ and $\mathcal{T}$ are two balanced rooted binary phylogenetic X-trees, where $n = |X|$, then* $\mathrm{mast}(\mathcal{S}, \mathcal{T}) \geq n^{0.17} > n^{\frac{1}{6}}$.

This paper is organised as follows. Section 2 details the notation and definitions used throughout, while Section 3 establishes Theorem 1.1 via a sequence of lemmas. Subsequently, Section 4 outlines the counterexample construction for when $n = 2048$. The last section, Section 5 gives the proof of the improved lower bound on the size of a MAST for a pair of balanced rooted binary phylogenetic trees (Theorem 1.3).

2. Notation and Preliminaries. In addition to the terminology given in the introduction, this section provides notation and terminology that is used in the remaining sections. Throughout the paper, X denotes a non-empty finite set and all logarithms are base 2.

Trees and subtrees. Since all phylogenetic trees in this paper are rooted and binary, we will refer to such a tree as simply a *phylogenetic tree*. Let $\mathcal{T}$ be a phylogenetic X-tree. We call X the *label set* of $\mathcal{T}$ and frequently denote it by $\mathcal{L}(\mathcal{T})$. A subtree of $\mathcal{T}$ is *pendant* if it can be detached from $\mathcal{T}$ by deleting a single edge. Observe that if $\mathcal{T}$ is balanced, then the size of $\mathcal{T}$ and each pendant subtree of $\mathcal{T}$ is a power of two. A balanced phylogenetic tree on n leaves has height $\log n$. As an example, Figure 1 shows such a tree on 8 leaves.

Of course, a phylogenetic tree is simply a certain type of rooted tree whose leaves are labelled. Thus, if we omit the adjective "phylogenetic" and say, for example, a balanced tree, we mean a balanced rooted binary tree whose leaf set is unlabelled. This occurs, in particular, in some of the proofs where we describe

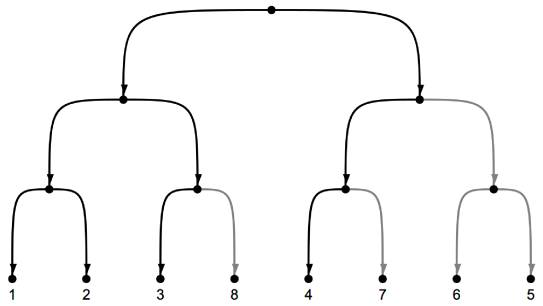

FIG. 1. *A balanced phylogenetic tree on 8 leaves. The black edges delineate the embedding of the restricted subtree on leaves* $\{1,2,3,4\}$ *which is the caterpillar* $(1,2,3,4)$*. The restricted subtree on the remaining four leaves also form a caterpillar, namely* $(5,6,7,8)$*.*

processes that start with an (unlabelled) tree and assign labels to its leaves so that the resulting tree is a phylogenetic tree.

Caterpillars. Let $\mathcal{T}$ be a phylogenetic X-tree with $n = |X|$. We call $\mathcal{T}$ an *n-caterpillar* or, simply, a *caterpillar* if $n = 1$, or $n \geq 2$ and we can order its leaf set X, say $l_1, l_2, \ldots, l_n$, so that l_1 and l_2 have the same parent and, for all $i \in \{2, 3, \ldots, n-1\}$, we have that (p_{i+1}, p_i) is an edge in $\mathcal{T}$, where p_{i+1} and p_i are the parents of l_{i+1} and l_i, respectively. We denote such a caterpillar $\mathcal{T}$ by $(l_1, l_2, \ldots, l_n)$ or, equivalently, $(l_2, l_1, l_3, \ldots, l_n)$, where l_1 and l_2 have been interchanged. For a caterpillar $\mathcal{C} = (l_1, l_2, \ldots, l_n)$ and a leaf l' with $l' \notin \{l_1, l_2, \ldots, l_n\}$, the caterpillar $(l_1, l_2, \ldots, l_n, l')$ is denoted by $\mathcal{C}||l'$. If $\mathcal{C} = (l_1, l_2, \ldots, l_n)$ and $\mathcal{C}' = (l'_1, l'_2, \ldots, l'_n)$ are two caterpillars on the same set of leaves such that $l_i = l'_{n-i+1}$ for all $i \in \{1, 2, \ldots, n\}$, we say that $\mathcal{C}$ and $\mathcal{C}'$ are a pair of *anti-caterpillars*. For example, if $\mathcal{C} = (l_1, l_2, \ldots, l_8)$ and $\mathcal{C}' = (l_8, l_7, \ldots, l_1)$, then $\mathcal{C}$ and $\mathcal{C}'$ are a pair of anti-caterpillars. Lastly, a caterpillar on 3 leaves, say (a, b, c), is called a *triple* and denoted by $ab|c$ or, equivalently, $ba|c$.

Embedding trees. Let $\mathcal{T}$ be a phylogenetic X-tree, and let $\mathcal{T}'$ be a phylogenetic Y-tree such that $Y \subseteq X$. We say that $\mathcal{T}$ *embeds* $\mathcal{T}'$ if $\mathcal{T}|\mathcal{L}(\mathcal{T}')$ is isomorphic to $\mathcal{T}'$, in which case we refer to the minimal rooted subtree of $\mathcal{T}$ that connects all leaves in Y as an *embedding* $\mathcal{E}$ of $\mathcal{T}'$ in $\mathcal{T}$. To illustrate, the subtree formed by the black edges in the balanced phylogenetic tree in Figure 1 is an embedding of the caterpillar $(1,2,3,4)$. Analogous to the label set of a phylogenetic tree, we use $\mathcal{L}(\mathcal{E})$ to refer to the label set of $\mathcal{E}$. More generally, let $\mathcal{P} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_k\}$ be a set of phylogenetic trees such that, for each $i \in \{1, 2, \ldots, k\}$, we have $\mathcal{L}(\mathcal{S}_i) \subseteq X$. Then the union $\mathcal{L}(S_1) \cup \mathcal{L}(\mathcal{S}_2) \cup \cdots \cup \mathcal{L}(S_k)$, denoted by $\mathcal{L}(\mathcal{P})$, is the *label* set of $\mathcal{P}$. If, for each tree $\mathcal{S}_i \in \mathcal{P}$ with $i \in \{1, 2, \ldots, k\}$, $\mathcal{T}$ embeds $\mathcal{S}_i$, we say that $\mathcal{T}$ *embeds* $\mathcal{P}$. In this case, an *embedding*, say $\mathcal{E}$, of $\mathcal{P}$ in $\mathcal{T}$ refers to the union of edges of an embedding of each tree of $\mathcal{P}$ in $\mathcal{T}$, and the *label* set of $\mathcal{E}$, denoted $\mathcal{L}(\mathcal{E})$, is $\mathcal{L}(\mathcal{P})$. Furthermore, if $\{\mathcal{L}(\mathcal{S}_1), \mathcal{L}(\mathcal{S}_2), \ldots, \mathcal{L}(\mathcal{S}_k)\}$ is a partition of X, then $\mathcal{T}$ *perfectly embeds* $\mathcal{P}$. Figure 1 shows a balanced phylogenetic tree on 8 leaves that perfectly embeds the two 4-caterpillars $(1,2,3,4)$ and $(5,6,7,8)$.

Maximum agreement subtrees. Let $\mathcal{S}$ and $\mathcal{T}$ be two phylogenetic trees, and let $\mathcal{S}_L, \mathcal{S}_R$ and $\mathcal{T}_L, \mathcal{T}_R$ denote the two maximal pendant subtrees of $\mathcal{S}$ and $\mathcal{T}$, respectively. Then, $\text{mast}(\mathcal{S}, \mathcal{T})$ can be computed recursively [21] as follows:

$$\begin{aligned} \text{mast}(\mathcal{S}, \mathcal{T}) = \max \big\{ &\text{mast}(\mathcal{S}_L, \mathcal{T}_L) + \text{mast}(\mathcal{S}_R, \mathcal{T}_R),\ \text{mast}(\mathcal{S}_L, \mathcal{T}_R) + \text{mast}(\mathcal{S}_R, \mathcal{T}_L), \\ &\text{mast}(\mathcal{S}, \mathcal{T}_L),\ \text{mast}(\mathcal{S}, \mathcal{T}_R),\ \text{mast}(\mathcal{S}_L, \mathcal{T}),\ \text{mast}(\mathcal{S}_R, \mathcal{T}) \big\}. \end{aligned} \tag{2.1}$$

Furthermore, a MAST for a pair of phylogenetic trees need not be unique. As an example, consider the two trees $\mathcal{S}$ and $\mathcal{T}$ on 16 leaves depicted in Figure 2. Any MAST of $\mathcal{S}$ and $\mathcal{T}$ has size 4. Two such MASTs for $\mathcal{S}$ and $\mathcal{T}$ are $\mathcal{M}_1 =$`((6,7),(10,11))` and $\mathcal{M}_2 =$`((5,6),(10,12))`.

3. Proof of Theorem 1.1. In this section, we prove Theorem 1.1 that there exist balanced phylogenetic trees on n leaves where the size of a MAST is $o(\sqrt{n})$. The key idea in establishing the result is to label the

leaves of two balanced trees $\mathcal{S}$ and $\mathcal{T}$ such that (1) each balanced pendant subtree of $\mathcal{S}$ of a certain size and each balanced pendant subtree of $\mathcal{T}$ of a certain size only have a fixed and small number of labels in common, and (2) each balanced pendant subtree of $\mathcal{S}$ (resp. $\mathcal{T}$) of a certain size perfectly embeds a set of label-disjoint caterpillars of a fixed size. Lemma 3.2 establishes Property (1), and Corollary 3.4 establishes Property (2). In fact, establishing both properties is also sufficient to show that the leaves common to a subtree of $\mathcal{S}$ and a subtree of $\mathcal{T}$ induce a pair of anti-caterpillars in $\mathcal{S}$ and $\mathcal{T}$, respectively. Once we have established that the leaves of $\mathcal{S}$ and $\mathcal{T}$ can be labeled in the desired way, we establish an upper bound on $\mathrm{mast}(\mathcal{S},\mathcal{T})$ in Lemma 3.5. Then the proof of Theorem 1.1 is an almost immediate consequence of Theorem 3.6 whose proof is an amalgamation of Lemmas 3.2 and 3.5 together with Corollary 3.4.

We start with a simple lemma that shows that the size of a MAST of any pair of anti-caterpillars is two.

LEMMA 3.1. *Let $\mathcal{C}$ and $\mathcal{C}'$ be a pair of anti-caterpillars on at least two leaves. Then* $\mathrm{mast}(\mathcal{C},\mathcal{C}') = 2$.

Proof. Evidently, $\mathrm{mast}(\mathcal{C},\mathcal{C}') \geq 2$. Let $ab|c$ be a triple that is embedded in $\mathcal{C}$. Since $\mathcal{C}$ and $\mathcal{C}'$ are anti-caterpillars, it follows that either $ac|b$ or $bc|a$ is a triple embedded in $\mathcal{C}'$, in particular, $ab|c$ is not embedded in $\mathcal{C}'$. Hence $\mathrm{mast}(\mathcal{C},\mathcal{C}') < 3$ and, therefore, $\mathrm{mast}(\mathcal{C},\mathcal{C}') = 2$. □

LEMMA 3.2. *Let $h_2 \geq h_1$ be non-negative integers. Let $\mathcal{S}$ (resp. $\mathcal{T}$) be a balanced tree on $2^{h_1+h_2}$ leaves consisting of 2^{h_1} balanced pendant subtrees $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{2^{h_1}}$ (resp. $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_{2^{h_1}}$) each of size 2^{h_2}. Then we can bijectively label the leaves of $\mathcal{S}$ and $\mathcal{T}$ with the elements in $\{1, 2, \ldots, 2^{h_1+h_2}\}$ so that, for all $i, j \in \{1, 2, \ldots, 2^{h_1}\}$, the pendant subtrees $\mathcal{S}_i$ and $\mathcal{T}_j$ have exactly $2^{h_2-h_1}$ common labels.*

Proof. For each $r \in \{1, 2, \ldots, 2^{2h_1}\}$, let

$$L_r = \{1 + (r-1)\cdot 2^{h_2-h_1},\ 2 + (r-1)\cdot 2^{h_2-h_1},\ \ldots,\ 2^{h_2-h_1} + (r-1)\cdot 2^{h_2-h_1}\}.$$

That is, $L_1 = \{1, 2, \ldots, 2^{h_2-h_1}\}$, $L_2 = \{1+2^{h_2-h_1}, 2+2^{h_2-h_1}, \ldots, 2^{h_2-h_1}+2^{h_2-h_1}\}$, and so forth. Note that $|L_r| = 2^{h_2-h_1}$ for all r. Moreover, $L_i \cap L_j = \emptyset$ for all $i \neq j$, and

$$\bigcup_{r=1}^{2^{2h_1}} L_r = \left\{1, 2, \ldots, 2^{h_1+h_2}\right\}.$$

Using the sets L_r, we now assign labels to the pendant subtrees $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{2^{h_1}}$ of $\mathcal{S}$ and the pendant subtrees $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_{2^{h_1}}$ of $\mathcal{T}$ as follows. For all i, the subtree $\mathcal{S}_i$ is assigned the union of label sets in row i and, for all j, the subtree $\mathcal{T}_j$ is assigned the union of label sets in column j:

	$\mathcal{T}_1$	$\mathcal{T}_2$	$\ldots$	$\mathcal{T}_j$	$\ldots$	$\mathcal{T}_{2^{h_1}}$
$\mathcal{S}_1$	L_1	L_2	$\ldots$	L_j	$\ldots$	$L_{2^{h_1}}$
$\mathcal{S}_2$	$L_{1+2^{h_1}}$	$L_{2+2^{h_1}}$	$\ldots$	$L_{j+2^{h_1}}$	$\ldots$	$L_{2^{h_1}+2^{h_1}}$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
$\mathcal{S}_i$	$L_{1+(i-1)\cdot 2^{h_1}}$	$L_{2+(i-1)\cdot 2^{h_1}}$	$\ldots$	$L_{j+(i-1)\cdot 2^{h_1}}$	$\ldots$	$L_{2^{h_1}+(i-1)\cdot 2^{h_1}}$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$		
$\mathcal{S}_{2^{h_1}}$	$L_{1+(2^{h_1}-1)\cdot 2^{h_1}}$	$L_{2+(2^{h_1}-1)\cdot 2^{h_1}}$	$\ldots$	$L_{j+(2^{h_1}-1)\cdot 2^{h_1}}$	$\ldots$	$L_{2^{h_1}+(2^{h_1}-1)\cdot 2^{h_1}} = L_{2^{2h_1}}$

Since $|L_r| = 2^{h_2-h_1}$, and $L_i \cap L_j = \emptyset$ for all $i \neq j$, it follows that each subtree $\mathcal{S}_i$ and $\mathcal{T}_j$ is assigned $2^{h_1} \cdot 2^{h_2-h_1} = 2^{h_2}$ distinct labels. Moreover, under this assignment of labels, $\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j) = L_r$ for some $r \in \{1, 2, \ldots, 2^{2h_1}\}$ (namely, the one that is placed in row i and column j). In particular, for all i and j, the pendant subtrees $\mathcal{S}_i$ and $\mathcal{T}_j$ have exactly $|L_r| = 2^{h_2-h_1}$ common labels. □

Consider the (infinite) sequence A054243 of integers from *The On-Line Encyclopedia of Integer Sequences* [11]. The first eighteen integers of this sequence are:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$a(n)$	1	1	1	2	2	4	8	16	16	32	64	128	256	512	1024	2048	2048	4096

and its closed form is $a(n) = 2^{\lfloor n - \log n - 1 \rfloor}$. Intuitively, the sequence is formed by successive powers of two, but it "stutters" if n is a power of two, that is, if n is a power of two, then $a(n) = a(n+1)$.

LEMMA 3.3. *Let n be a positive integer. Then a balanced phylogenetic tree on 2^{n-1} leaves embeds*

$$2^{\lfloor n-\log n-1\rfloor}$$

label-disjoint n-caterpillars.

Proof. The proof is by induction on n. Let $\mathcal{T}$ be a balanced phylogenetic tree on 2^{n-1} leaves. Evidently, if $n = 1$, then $\mathcal{T}$ embeds one 1-caterpillar, and so the lemma holds. Similarly, if $n = 2$ (resp. $n = 3$), then $\mathcal{T}$ embeds one 2-caterpillar (resp. one 3-caterpillar). Now suppose that $n \geq 4$, and that the lemma holds for all positive integers at most $n-1$. Let $\mathcal{T}_1$ and $\mathcal{T}_2$ denote the two maximal pendant subtrees of $\mathcal{T}$. For each $i \in \{1,2\}$, $\mathcal{T}_i$ is a balanced phylogenetic tree on 2^{n-2} leaves and so, by the induction assumption, $\mathcal{T}_i$ embeds

$$a(n-1) = 2^{\lfloor (n-1)-\log(n-1)-1\rfloor}$$

label-disjoint $(n-1)$-caterpillars. The remainder of the proof splits into two cases depending on whether or not $n-1$ is a power of two.

For the first case, assume that $n-1$ is a power of two. Then $a(n) = a(n-1)$, and so we need to show that $\mathcal{T}$ embeds $a(n-1)$ label-disjoint n-caterpillars. If $\mathcal{C}_{n-1}$ is an $(n-1)$-caterpillar embedded in $\mathcal{T}_i$, then $\mathcal{C}_{n-1}||l$ is an n-caterpillar embedded in $\mathcal{T}$, where $l \in \mathcal{T}_j$ and $\{i,j\} = \{1,2\}$. Thus, we can extend an embedding $\mathcal{E}_1$ of $\frac{1}{2}a(n-1)$ label-disjoint $(n-1)$-caterpillars in $\mathcal{T}_1$ and an embedding $\mathcal{E}_2$ of $\frac{1}{2}a(n-1)$ label-disjoint $(n-1)$-caterpillars in $\mathcal{T}_2$ to an embedding of $a(n-1)$ label-disjoint n-caterpillars in $\mathcal{T}$ provided, for each $i \in \{1,2\}$, we have

$$2^{n-2} - |\mathcal{L}(\mathcal{E}_i)| \geq \tfrac{1}{2}a(n-1).$$

That is,

$$\tfrac{1}{2} \cdot 2^{\lfloor (n-1)-\log(n-1)-1\rfloor} \cdot (n-1) + \tfrac{1}{2} \cdot 2^{\lfloor (n-1)-\log(n-1)-1\rfloor} \leq 2^{n-2}. \tag{3.1}$$

Since $n-1$ is a power of two and $n = 2^{\log n}$, the LHS of (3.1) is

$$\begin{aligned}\tfrac{1}{2} \cdot 2^{\lfloor (n-1)-\log(n-1)-1\rfloor} \cdot ((n-1)+1) &= 2^{-1} \cdot 2^{(n-1)-\log(n-1)-1} \cdot 2^{\log n} \\ &= 2^{n-3-\log(n-1)+\log n}.\end{aligned}$$

As $-\log(n-1) + \log n \leq 1$, we have

$$2^{n-3-\log(n-1)+\log n} \leq 2^{n-3+1} = 2^{n-2},$$

thereby establishing (3.1).

For the second case, assume that $n-1$ is not a power of two. Then $a(n) = 2a(n-1)$, and so we need to show that $\mathcal{T}$ embeds $2a(n-1)$ label-disjoint n-caterpillars. Analogous to the first case, we can extend an embedding $\mathcal{E}_1$ of $a(n-1)$ label-disjoint $(n-1)$-caterpillars in $\mathcal{T}_1$ and an embedding $\mathcal{E}_2$ of $a(n-1)$ label-disjoint $(n-1)$-caterpillars in $\mathcal{T}_2$ to an embedding of $2a(n-1)$ label-disjoint n-caterpillars in $\mathcal{T}$ provided, for each $i \in \{1,2\}$, we have

$$2^{n-2} - |\mathcal{L}(\mathcal{E}_i)| \geq a(n-1).$$

That is,

$$2^{\lfloor (n-1)-\log(n-1)-1\rfloor} \cdot (n-1) + 2^{\lfloor (n-1)-\log(n-1)-1\rfloor} \leq 2^{n-2}. \tag{3.2}$$

Since $n = 2^{\log n}$, the LHS of (3.2) is

$$\begin{aligned}2^{\lfloor (n-1)-\log(n-1)-1\rfloor} \cdot ((n-1)+1) &= 2^{\lfloor (n-1)-\log(n-1)-1\rfloor} \cdot 2^{\log n} \\ &= 2^{n-2+\lfloor -\log(n-1)\rfloor+\log n}.\end{aligned}$$

As $n-1$ is not a power of two, it follows that $\lfloor -\log(n-1)\rfloor + \log n \leq 0$, and so

$$2^{n-2+\lfloor -\log(n-1)\rfloor+\log n} \leq 2^{n-2}.$$

This establishes (3.2), and completes the proof of the lemma. ∎

COROLLARY 3.4. *Let k be a positive integer. Then a balanced phylogenetic tree $\mathcal{T}$ of height $2^k - 1$ perfectly embeds 2^{2^k-k-1} caterpillars each of size 2^k.*

Proof. Taking $n = 2^k$ in Lemma 3.3, it follows that $\mathcal{T}$ embeds

$$2^{\lfloor 2^k - \log(2^k) - 1 \rfloor} = 2^{\lfloor 2^k - k - 1 \rfloor} = 2^{2^k-k-1}$$

label-disjoint 2^k-caterpillars. Collectively, these caterpillars have

$$2^{2^k-k-1} \cdot 2^k = 2^{2^k-1}$$

labels and, since the caterpillars are label-disjoint and $\mathcal{T}$ has 2^{2^k-1} leaves, it follows that this embedding is perfect. □

To illustrate Corollary 3.4 for $k = 2$, a balanced phylogenetic tree of height 3 perfectly embeds two label-disjoint 4-caterpillars. For an example, see Figure 1, where the two perfectly embedded caterpillars are $(1, 2, 3, 4)$ and $(5, 6, 7, 8)$. Moreover, for $k = 3$, a balanced phylogenetic tree of height 7 perfectly embeds sixteen label-disjoint 8-caterpillars. Two such trees are shown in Figure 3, where the pair of anti-caterpillars $(1, 2, 3, 4, 5, 6, 7, 8)$ and $(8, 7, 6, 5, 4, 3, 2, 1)$ is highlighted in blue and green, respectively.

LEMMA 3.5. *Let r be a positive integer. Furthermore, let $\mathcal{S}$ and $\mathcal{T}$ be two balanced phylogenetic trees on $p \cdot 2^k$ and $q \cdot 2^k$ leaves, respectively, where k, p, and q are positive integers. Suppose that $\mathcal{S}$ consists of p pendant subtrees $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_p$ each of size 2^k and $\mathcal{T}$ consists of q pendant subtrees $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_q$ each of size 2^k. If*

(i) *for all i, j, we have $|\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j)| = r$ and*
(ii) *for all i, j, the restrictions $\mathcal{S}_i|(\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j))$ and $\mathcal{T}_j|(\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j))$ is a pair of anti-caterpillars,*

then, $\text{mast}(\mathcal{S}, \mathcal{T}) \leq 2 \cdot \max\{p, q\}$.

Proof. Let $f(p, q)$ denote the maximum size of a MAST for $\mathcal{S}$ and $\mathcal{T}$. We will show by induction on $p+q$ that $f(p, q) \leq 2 \cdot \max\{p, q\}$. Note that, since $\mathcal{S}$ and $\mathcal{T}$ are balanced, p and q are powers of two. Furthermore, by symmetry, $f(p, q) = f(q, p)$.

Without loss of generality, we may assume $p \leq q$. If $p = 1$, then, for all positive integers q, we have $f(1, q) = 2q$. To see this, for all j, the restricted subtrees $\mathcal{S}_1|(\mathcal{L}(\mathcal{S}_1) \cap \mathcal{L}(\mathcal{T}_j))$ and $\mathcal{T}_j|(\mathcal{L}(\mathcal{S}_1) \cap \mathcal{L}(\mathcal{T}_j))$ induce a pair of anti-caterpillars, and so, by Lemma 3.1, each pendant subtree $\mathcal{T}_j$ of $\mathcal{T}$ contributes exactly 2 leaves to a MAST between $\mathcal{S}$ and $\mathcal{T}$. It follows that the lemma holds for when $p = 1$ and q is a positive integer.

Now assume that $p \geq 2$ and the lemma holds for all smaller values of $p + q$. Let $\mathcal{S}_L, \mathcal{S}_R$ and $\mathcal{T}_L, \mathcal{T}_R$ be the maximal pendant subtrees of $\mathcal{S}$ and $\mathcal{T}$, respectively. Let $\mathcal{M}$ be a MAST of $\mathcal{S}$ and $\mathcal{T}$. If there exist $x, y \in \mathcal{L}(\mathcal{M})$ such that $x \in \mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_L)$ and $y \in \mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)$, then $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R)$ is empty. Otherwise, there is a triple $xy|z$ embedded in $\mathcal{M}$, where $z \in \mathcal{L}(\mathcal{S}_R)$, but $xy|z$ is not embedded in $\mathcal{T}$, a contradiction. Intuitively, if $\mathcal{M}$ connects one half of $\mathcal{S}$ to both halves of $\mathcal{T}$, then the other half of $\mathcal{S}$ is not used in $\mathcal{M}$. Using symmetric arguments, we obtain the following:

(A) if $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_L)$ and $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)$ are non-empty, then $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R) = \emptyset$;
(B) if $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_L)$ and $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_R)$ are non-empty, then $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_L) = \emptyset$;
(C) if $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_L)$ and $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_L)$ are non-empty, then $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{T}_R) = \emptyset$; and
(D) if $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)$ and $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_R)$ are non-empty, then $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{T}_L) = \emptyset$.

Thus,

(I) either $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_L)$ or $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R)$ is empty, or
(II) either $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{T}_L)$ or $\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{T}_R)$ is empty, or
(III) either

$$(\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)) \cup (\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_L))$$

or

$$(\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_L)) \cup (\mathcal{L}(\mathcal{M}) \cap \mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_R))$$

is empty.

Therefore, $f(p, q) \leq \max\{f(p/2, q),\ f(p, q/2),\ 2f(p/2, q/2)\}$.

If $p \leq q/2$, then, by induction, $f(p, q) \leq \max\{2q, q, 2q\} = 2q$ and so, as $p \leq q$, the lemma holds. If $p > q/2$, then, by induction, $f(p, q) \leq \max\{2q, 2p, 2q\} = 2q$ since $p \leq q$, and it again follows that the lemma holds. This completes the proof. □

Theorem 1.1 is an almost immediate consequence of the next theorem:

THEOREM 3.6. *Let k be a positive integer. Then, there exist balanced phylogenetic X-trees $\mathcal{S}$ and $\mathcal{T}$, where $|X| = 2^{2^{k+1}-k-2}$, such that* $\mathrm{mast}(\mathcal{S}, \mathcal{T}) = 2^{2^k-k}$.

Proof. Let $h_1 = 2^k - k - 1$ and $h_2 = 2^k - 1$. Let $\mathcal{S}$ and $\mathcal{T}$ be two (unlabelled) balanced trees of height $h = h_1 + h_2$. Thus, $\mathcal{S}$ and $\mathcal{T}$ each has size $2^h = 2^{2^{k+1}-k-2}$ and, each furthermore, has 2^{h_1} balanced pendant subtrees of size 2^{h_2}. Let $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{2^{h_1}}$ denote the pendant subtrees of $\mathcal{S}$ of size 2^{h_2} and let $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_{2^{h_1}}$ denote the pendant subtrees of $\mathcal{T}$ of size 2^{h_2}.

By Lemma 3.2, we can bijectively label the leaves of $\mathcal{S}$ and $\mathcal{T}$ with the elements in $\{1, 2, \ldots, 2^{h_1+h_2}\}$ so that, for all $i, j \in \{1, 2, \ldots, 2^{h_1}\}$, the pendant subtrees $\mathcal{S}_i$ and $\mathcal{T}_j$ have exactly $2^{h_2-h_1} = 2^k$ common labels. Moreover, it follows by Corollary 3.4 that, under such a bijection, we can label the leaves of $\mathcal{S}$ and $\mathcal{T}$ so that, for all i, j, the restrictions $\mathcal{S}_i|(\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j))$ and $\mathcal{T}_j|(\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j))$ is in fact a pair of anti-caterpillars. With this labelling of $\mathcal{S}$ and $\mathcal{T}$, Lemma 3.5 says that:

$$\mathrm{mast}(\mathcal{S}, \mathcal{T}) \leq 2 \cdot \max\left\{2^{h_1}, 2^{h_1}\right\} = 2 \cdot 2^{2^k-k-1} = 2^{2^k-k}.$$

Since $\mathcal{S}$ and $\mathcal{T}$ have the same label sets, it is easily checked that this inequality is an equality, that is,

$$\mathrm{mast}(\mathcal{S}, \mathcal{T}) = 2^{2^k-k},$$

completing the proof of the theorem. □

We now prove Theorem 1.1.

Proof of Theorem 1.1. Pick an integer k such that $k > 2\log(1/c) + 2$. By Theorem 3.6, there exist balanced phylogenetic trees $\mathcal{S}$ and $\mathcal{T}$ on the same label set of size $n = 2^{2^{k+1}-k-2}$ such that

$$\mathrm{mast}(\mathcal{S}, \mathcal{T}) = 2^{2^k-k}.$$

Since

$$\frac{2^{2^k-k}}{2^{2^k-\frac{k}{2}-1}} = 2^{-\frac{k}{2}+1} < c,$$

where $2^{2^k-\frac{k}{2}-1} = n^{\frac{1}{2}}$, we have

$$2^{2^k-k} = \frac{2^{2^k-k}}{2^{2^k-\frac{k}{2}-1}} \cdot 2^{2^k-\frac{k}{2}-1} < cn^{\frac{1}{2}},$$

and the theorem follows. □

In particular, if we take $c = 1$, then following the proof of Theorem 1.1 and choosing $k = 3$ implies that there exist two balanced phylogenetic trees on $2^{11} = 2048$ leaves such that their MAST has size $2^5 = 32$, whereas $\sqrt{2048} \approx 45.25$. Using the results presented in this section, the next section gives an explicit construction of two balanced phylogenetic trees on 2048 leaves whose MAST has size 32.

4. The Counterexample Construction. In this section, we give an explicit construction of a counterexample to Conjecture 1.2. We note that the results in Section 3 are more general than what we need for the counterexample that is presented in this section. In particular, Theorem 3.6 shows that, for any positive integer k, there exists a pair of balanced phylogenetic trees on $2^{2^{k+1}-k-2}$ leaves such that the size of a MAST for these two trees is 2^{2^k-k}. Thus, for $k = 3$, this theorem states that there exist two balanced trees $\mathcal{S}$ and $\mathcal{T}$ on $2^{2^{3+1}-3-2} = 2^{11} = 2048$ leaves that can be bijectively labelled with the elements in $\{1, 2, \ldots, 2048\}$ such that the size of a MAST for the resulting two balanced phylogenetic trees has size $2^{2^3-3} = 32 < 2048^{\frac{1}{2}}$.

Let $\mathcal{S}$ and $\mathcal{T}$ be two balanced trees on $2048 = 2^{4+7}$ leaves. In what follows, we label the leaves of $\mathcal{S}$ and $\mathcal{T}$ such that $\mathrm{mast}(\mathcal{S}, \mathcal{T}) = 32$. Let $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{16}$ be the sixteen leaf-disjoint balanced pendant subtrees of $\mathcal{S}$ that each have size 2^7, and let $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_{16}$ be the sixteen leaf-disjoint balanced pendant subtrees of $\mathcal{T}$ that each have size 2^7. By Lemma 3.2 and Corollary 3.4, we can bijectively label the leaves of $\mathcal{S}$ and $\mathcal{T}$ with the elements in $\{1, 2, \ldots, 2048\}$ such that the following three properties are satisfied:

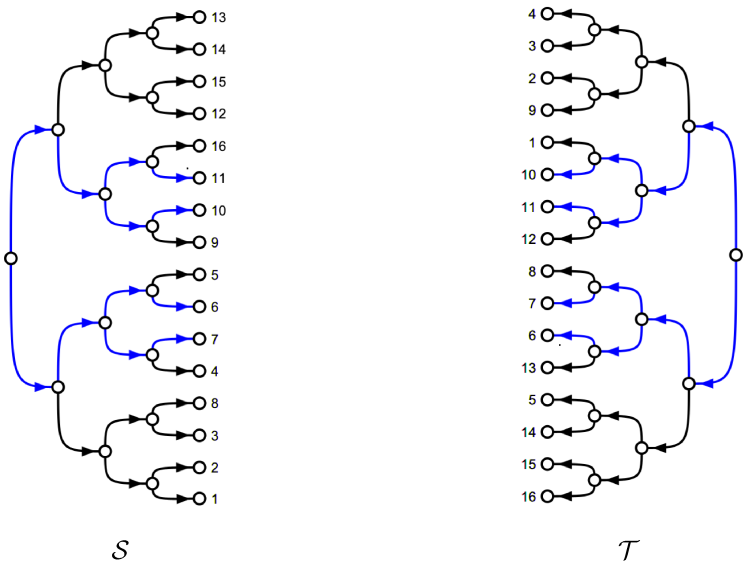

FIG. 2. *Two phylogenetic trees $\mathcal{S}$ and $\mathcal{T}$ on 16 leaves, drawn with PhyloSketch [10]. An embedding of the MAST $\mathcal{M}_1 =$ ((6,7),(10,11)) for $\mathcal{S}$ and $\mathcal{T}$ is shown in blue.*

(i) for all $i, j \in \{1, 2, \ldots, 16\}$, the pendant subtrees $\mathcal{S}_i$ and $\mathcal{T}_j$ have exactly $|\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j)| = 2^{7-4} = 8$ labels in common,
(ii) each balanced pendant subtree in $\{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{16}\}$ and $\{\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_{16}\}$ perfectly embeds sixteen 8-caterpillars, and
(iii) the restrictions $\mathcal{S}_i|(\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j))$ and $\mathcal{T}_j|(\mathcal{L}(\mathcal{S}_i) \cap \mathcal{L}(\mathcal{T}_j))$ are a pair of anti-caterpillars.

Now, if $\mathcal{S}$ and $\mathcal{T}$ have their leaves labelled such that (i), (ii), and (iii) hold, then applying Lemma 3.5 for when $r = 8$, $k = 7$, and $p = q = 16$ implies that

$$\text{mast}(\mathcal{S}, \mathcal{T}) = 2 \cdot \max\{p, q\} = 2 \cdot 16 = 32 < 2048^{\frac{1}{2}},$$

which disproves Conjecture 1.2. To illustrate the labelling of $\mathcal{S}$ and $\mathcal{T}$, the two subtrees $\mathcal{S}_1$ and $\mathcal{T}_1$ with their leaves labelled are shown in Figure 3. Note that $\mathcal{S}_1$ and $\mathcal{T}_1$ embed the pair $(1, 2, 3, 4, 5, 6, 7, 8)$ and $(8, 7, 6, 5, 4, 3, 2, 1)$ of anti-caterpillars as indicated by the edges highlighted in blue and green, respectively. The full trees in Newick format are given in Appendix A and their MAST may be verified to have size 32 by standard computations [17].

5. Lower Bounds. The main result of this paper has been to disprove Conjecture 1.2 of [13]. What then can be said about the size of a MAST of two balanced phylogenetic trees on the same label set? Corollary 16 of the same paper states that two balanced phylogenetic trees on the same leaf set of cardinality n have a MAST on at least n^β leaves, where β is not explicitly given but is defined by some formulae. Specifically, for any $\delta \in \left(0, \frac{1}{3} - \frac{1}{3\sqrt{2}}\right)$, we can take

$$\beta = \left(\frac{1 + 2\log(1 - 3\delta)}{\log(1 - 3\delta) - \log \delta}\right).$$

Numerically, the largest value of β that may be inferred from their approach is approximately 0.149. We have observed that this lower bound can be slightly improved by analyzing the lower bounds of the possible cases in the computation of the MAST and careful choice of constants to yield $n^{0.17}$. We can get a slightly better lower bound of $n^{0.1721}$ by using 0.22166 and 0.02478 in Equation 5.1, and similar increases to the lower bound are in the thousands place. Significant improvement in the lower bound will require a new approach.

THEOREM 5.1. *If $\mathcal{S}$ and $\mathcal{T}$ are two balanced phylogenetic X-trees, where $n = |X|$, then*

$$\text{mast}(\mathcal{S}, \mathcal{T}) \geq n^{0.17} > n^{\frac{1}{6}}.$$

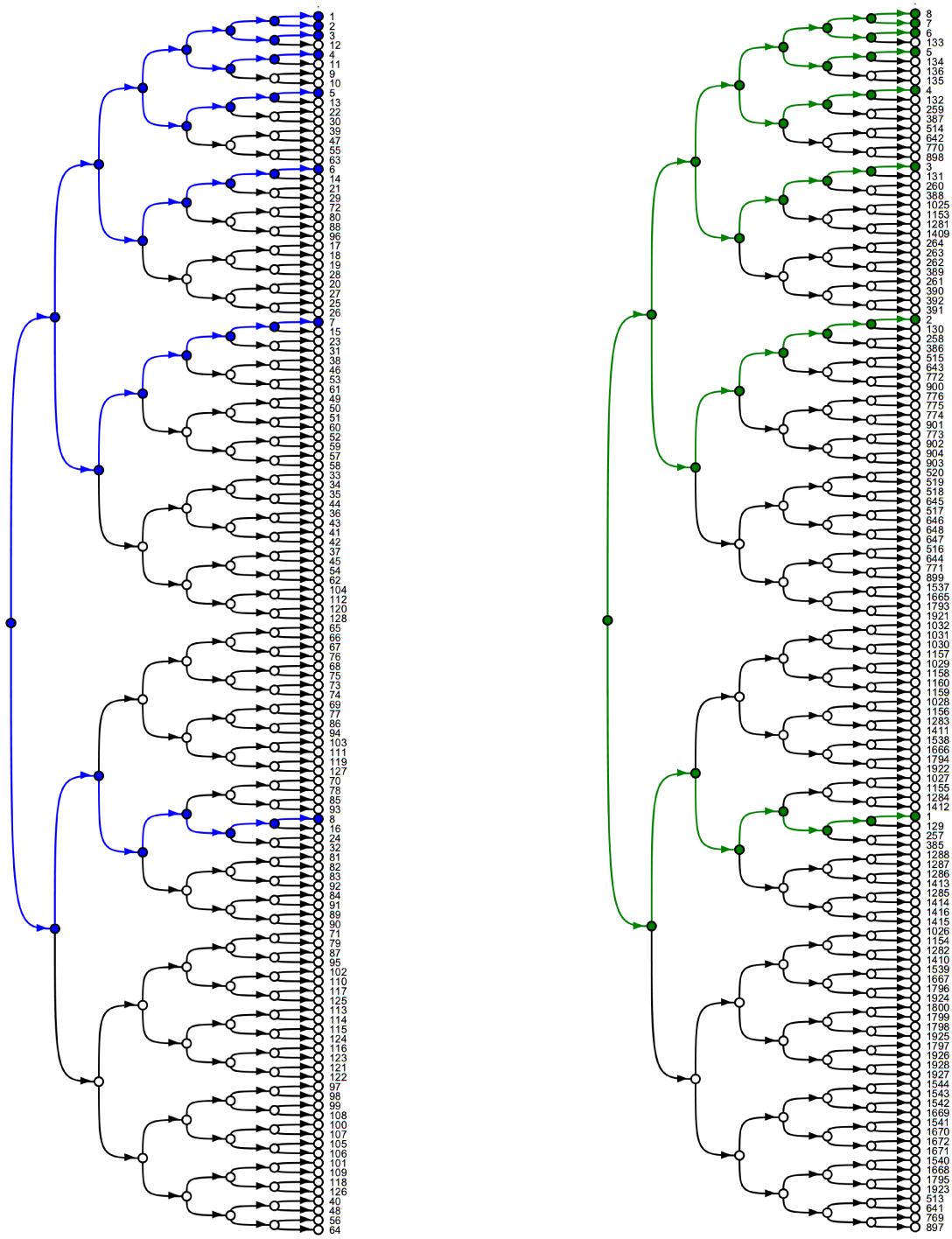

FIG. 3. *Subtrees $\mathcal{S}_1$ (left) and $\mathcal{T}_1$ (right) that are used to construct a counterexample for Conjecture 1.2. A pair of anti-caterpillars on leaves $\{1, 2, \ldots, 8\}$ is highlighted in blue and green in $\mathcal{S}_1$ and $\mathcal{T}_1$, respectively. Each of $\mathcal{S}_1$ and $\mathcal{T}_1$ perfectly embeds sixteen 8-caterpillars.*

Proof. The proof closely follows the approach of Lemma 15 in [13] for which Corollary 16 in [13] is an almost immediate consequence. Let $g(h_1, h_2, t)$ be the minimum size of a MAST over all pairs of balanced phylogenetic trees $\mathcal{S}$ and $\mathcal{T}$ such that $\mathcal{S}$ has 2^{h_1} leaves, $\mathcal{T}$ has 2^{h_2} leaves, and the leaf sets overlap in exactly

$t > 0$ leaves. We first show by induction on $h_1 + h_2$ that

$$g(h_1, h_2, t) \geq 2^{0.22 \log t - 0.025(h_1+h_2)}. \tag{5.1}$$

For the base case, if $h_1 + h_2 \in \{0, 1\}$, then at least one of the trees consists of a single leaf, and so $g(h_1, h_2, 1) = 1 \geq 2^0$ and Equation (5.1) holds. Now, assume that Equation (5.1) holds for all pairs of balanced phylogenetic trees whose heights sum to at most $h_1 + h_2 - 1$. Let $\mathcal{S}_L$ and $\mathcal{S}_R$ be the two maximal pendant subtrees of $\mathcal{S}$, and let $\mathcal{T}_L$ and $\mathcal{T}_R$ be the two maximal pendant subtrees of $\mathcal{T}$. By Equation (2.1),

$$\begin{aligned}\operatorname{mast}(\mathcal{S}, \mathcal{T}) \geq \max \big\{ & g(h_1 - 1, h_2 - 1, |\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_L)|) + g(h_1 - 1, h_2 - 1, |\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_R)|), \\ & g(h_1 - 1, h_2 - 1, |\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)|) + g(h_1 - 1, h_2 - 1, |\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_L)|), \\ & g(h_1, h_2 - 1, |\mathcal{L}(\mathcal{S}) \cap \mathcal{L}(\mathcal{T}_L)|),\ g(h_1, h_2 - 1, |\mathcal{L}(\mathcal{S}) \cap \mathcal{L}(\mathcal{T}_R)|), \\ & g(h_1 - 1, h_2, |\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T})|),\ g(h_1 - 1, h_2, |\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T})|) \big\}.\end{aligned}$$

We freely use this inequality in the remainder of the proof.

Without loss of generality, we may assume that the largest overlap between the leaf sets $\mathcal{S}_L, \mathcal{S}_R$ and $\mathcal{T}_L, \mathcal{T}_R$ is between $\mathcal{S}_L$ and $\mathcal{T}_L$. By the pigeonhole principle, $|\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_L)| \geq t/4$, and so one of the following cases must occur by exhaustion:

(i) $|\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_R)| \geq 0.037t$;
(ii) $|\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)|, |\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_L)| \geq 0.037t$;
(iii) $|\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_R)|, |\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)|, |\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_L)| < 0.037t$;
(iv) $|\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)| \geq 0.037t$ and $|\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_R)|, |\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_L)| < 0.037t$;
(v) $|\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_R)|, |\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_R)| < 0.037t$ and $|\mathcal{L}(\mathcal{S}_R) \cap \mathcal{L}(\mathcal{T}_L)| \geq 0.037t$.

In Case (i), it follows by the induction assumption that

$$\begin{aligned}\operatorname{mast}(\mathcal{S}, \mathcal{T}) &\geq g(h_1 - 1, h_2 - 1, 0.25t) + g(h_1 - 1, h_2 - 1, 0.037t) \\ &> 2 \cdot 2^{0.22 \log 0.037t - 0.025(h_1 - 1 + h_2 - 1)} \\ &= 2^{1 + 0.22 \log 0.037 + 0.22 \log t + 0.05 - 0.025(h_1+h_2)} \\ &> 2^{0.22 \log t - 0.025(h_1+h_2)}.\end{aligned}$$

In Case (ii), by a similar calculation to that in Case (i), we obtain

$$\operatorname{mast}(\mathcal{S}, \mathcal{T}) \geq 2g(h_1 - 1, h_2 - 1, 0.037t) > 2^{0.22 \log t - 0.025(h_1+h_2)}.$$

For Case (iii), since the total overlap in leaf sets is t, we must have $|\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T}_L)| \geq 0.889t$. So,

$$\begin{aligned}\operatorname{mast}(\mathcal{S}, \mathcal{T}) &\geq 2^{0.22 \log 0.889t - 0.025(h_1 - 1 + h_2 - 1)} \\ &= 2^{0.22 \log 0.889 + 0.22 \log t + 0.05 - 0.025(h_1+h_2)} \\ &> 2^{0.22 \log t - 0.025(h_1+h_2)}.\end{aligned}$$

Finally, in Case (iv), we have $|\mathcal{L}(\mathcal{S}_L) \cap \mathcal{L}(\mathcal{T})| \geq 0.926t$ and, in Case (v), we have $|\mathcal{L}(\mathcal{S}) \cap \mathcal{L}(\mathcal{T}_L)| \geq 0.926t$. In both cases, we have

$$\begin{aligned}\operatorname{mast}(\mathcal{S}, \mathcal{T}) &\geq 2^{0.22 \log 0.926t - 0.025(h_1 + h_2 - 1)} \\ &= 2^{0.22 \log 0.926 + 0.22 \log t + 0.025 - 0.025(h_1+h_2)} \\ &> 2^{0.22 \log t - 0.025(h_1+h_2)}.\end{aligned}$$

Thus, Equation (5.1) holds by induction. Substituting $h_1 = h_2 = \log n$ and $t = n$ into Equation (5.1), we obtain that any two balanced phylogenetic trees on the same leaf set of size n have a MAST on at least

$$g(\log n, \log n, n) \geq 2^{0.22 \log n - 0.05 \log n} = n^{0.17} > n^{\frac{1}{6}}$$

leaves. This completes the proof of the theorem. ∎

6. Acknowledgements. All authors thank Schloss Dagstuhl—Leibniz Centre for Informatics—for hosting the Seminar 19443 *Algorithms and Complexity in Phylogenetics* in October 2019, where this work was initiated and Prof. Mike Steel for hosting a workshop in Sumner, New Zealand in March 2020. The first, second and fifth authors thank the New Zealand Marsden Fund for financial support. The first and fourth authors thank the Erskine Visiting Fellowship Programme for supporting their extended visits to the University of Canterbury, New Zealand. The third author was partially supported by the US National Science Foundation (DMS 1847271) and by a grant from the Simons Foundation (355824). The fourth author thanks the Simons Foundation for their grant to support collaboration and travel. The sixth author thanks the German Academic Scholarship Foundation for a doctoral scholarship, under which parts of this work were conducted. All authors thank Sean Cleary for helpful discussions.

REFERENCES

[1] A. Amir and D. Keselman. Maximum agreement subtree in a set of evolutionary trees: metrics and efficient algorithms. *SIAM Journal on Computing*, 26(6):1656–1669, 1997.
[2] D. I. Bernstein, L. S. Tung Ho, C. Long, M. Steel, K. St. John, and S. Sullivant. Bounds on the expected size of the maximum agreement subtree. *SIAM Journal on Discrete Mathematics*, 29(4):2065–2074, 2015.
[3] V. Berry and F. Nicolas. Maximum agreement and compatible supertrees. *Journal of Discrete Algorithms*, 5(3):564–591, 2007.
[4] D. Bryant. A classification of consensus methods for phylogenetics. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 61:163–184, 2003.
[5] D. Bryant, A. McKenzie, and M. Steel. The size of a maximum agreement subtree for random binary trees. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 61:55–66, 2003.
[6] R. Cole, M. Farach-Colton, R. Hariharan, T. Przytycka, and M. Thorup. An $O(n \log n)$ algorithm for the maximum agreement subtree problem for binary trees. *SIAM Journal on Computing*, 30(5):1385–1404, 2000.
[7] D. Colless. Review of "Phylogenetics: the theory and practice of phylogenetic systematics". *Systematic Zoology*, 31(1):100–104, 1982.
[8] C. Finden and A. Gordon. Obtaining common pruned trees. *Journal of Classification*, 2(1):255–276, 1985.
[9] W. Goddard, E. Kubicka, G. Kubicki, and F. R. McMorris. The agreement metric for labeled binary trees. *Mathematical Biosciences*, 123(2):215–226, 1994.
[10] D. H. Huson and M. A. Steel. PhyloSketch: interactive program for sketching rooted phylogenetic trees and networks. https://software-ab.informatik.uni-tuebingen.de/download/phylosketch/welcome.html, 2020.
[11] M. LeBrun. The On-Line Encyclopedia of Integer Sequences, Sequence A054243. https://oeis.org/A054243, 2000.
[12] A. Markin. On the extremal maximum agreement subtree problem. *Discrete Applied Mathematics*, 285:612–620, 2020.
[13] D. M. Martin and B. D. Thatte. The maximum agreement subtree problem. *Discrete Applied Mathematics*, 161(13-14):1805–1817, 2013.
[14] P. Misra and S. Sullivant. Bounds on the expected size of the maximum agreement subtree for a given tree shape. *SIAM Journal on Discrete Mathematics*, 33(4):2316–2325, 2019.
[15] A. Ramu, T. Kahveci, and J. G. Burleigh. A scalable method for identifying frequent subtrees in sets of large phylogenetic trees. *BMC bioinformatics*, 13(1):1–15, 2012.
[16] M. J. Sackin. "Good" and "Bad" phenograms. *Systematic Biology*, 21(2):225–226, 1972.
[17] K. P. Schliep. Phangorn: phylogenetic analysis in R. *Bioinformatics*, 27(4):592–593, 2011.
[18] C. Semple and M. Steel. *Phylogenetics*, volume 24 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, 2003.
[19] K. St. John. The shape of phylogenetic treespace. *Systematic Biology*, 66(1):e83–e94, 2017.
[20] M. Steel and L. A. Székely. An improved bound on the maximum agreement subtree problem. *Applied Mathematics Letters*, 22(11):1778–1780, 2009.
[21] M. Steel and T. Warnow. Kaikoura tree theorems: computing the maximum agreement subtree. *Information Processing Letters*, 48(2):77–82, 1993.

Appendix A. Example on 2048 leaves. The following two balanced phylogenetic trees on label set $\{1, 2, \ldots, 2048\}$ may be shown to have a MAST of size 32 by standard computational means [17]. They are presented in standard Newick format below.

$\mathcal{S}$ =(((((((((((1,2),(3,12)),((4,11),(9,10))),(((5,13),(22,30)),((39,47),(55,63)))),((((6,14),(21,29)),((72
,80),(88,96))),(((17,18),(19,28)),((20,27),(25,26)))),(((((7,15),(23,31)),((38,46),(53,61))),(((49,50),(51,60
)),((52,59),(57,58)))),((((33,34),(35,44)),((36,43),(41,42))),(((37,45),(54,62)),((104,112),(120,128))))),(((
(((65,66),(67,76)),((68,75),(73,74))),(((69,77),(86,94)),((103,111),(119,127)))),((((70,78),(85,93)),((8,16),(
24,32))),(((81,82),(83,92)),((84,91),(89,90))))),(((((71,79),(87,95)),((102,110),(117,125))),(((113,114),(115,
124)),((116,123),(121,122)))),((((97,98),(99,108)),((100,107),(105,106))),(((101,109),(118,126)),((40,48),(56,
64)))))),(((((((129,130),(131,140)),((132,139),(137,138))),(((133,141),(150,158)),((167,175),(183,191)))),(((
(134,142),(149,157)),((200,208),(216,224))),(((145,146),(147,156)),((148,155),(153,154))))),(((((135,143),(151
,159)),((166,174),(181,189))),(((177,178),(179,188)),((180,187),(185,186)))),((((161,162),(163,172)),((164,171
),(169,170))),(((165,173),(182,190)),((232,240),(248,256))))),((((((193,194),(195,204)),((196,203),(201,202))
),(((197,205),(214,222)),((231,239),(247,255)))),((((198,206),(213,221)),((136,144),(152,160))),(((209,210),(2
11,220)),((212,219),(217,218))))),(((((199,207),(215,223)),((230,238),(245,253))),(((241,242),(243,252)),((244
,251),(249,250)))),((((225,226),(227,236)),((228,235),(233,234))),(((229,237),(246,254)),((168,176),(184,192))
))))),(((((((257,258),(259,268)),((260,267),(265,266))),(((261,269),(278,286)),((295,303),(311,319)))),((((2
62,270),(277,285)),((328,336),(344,352))),(((273,274),(275,284)),((276,283),(281,282))))),(((((263,271),(279,2
87)),((294,302),(309,317))),(((305,306),(307,316)),((308,315),(313,314)))),((((289,290),(291,300)),((292,299),
(297,298))),(((293,301),(310,318)),((360,368),(376,384))))),((((((321,322),(323,332)),((324,331),(329,330))),
(((325,333),(342,350)),((359,367),(375,383)))),((((326,334),(341,349)),((264,272),(280,288))),(((337,338),(339
,348)),((340,347),(345,346))))),(((((327,335),(343,351)),((358,366),(373,381))),(((369,370),(371,380)),((372,3
79),(377,378)))),((((353,354),(355,364)),((356,363),(361,362))),(((357,365),(374,382)),((296,304),(312,320))))
))),(((((((385,386),(387,396)),((388,395),(393,394))),(((389,397),(406,414)),((423,431),(439,447)))),((((390,3
98),(405,413)),((456,464),(472,480))),(((401,402),(403,412)),((404,411),(409,410))))),(((((391,399),(407,415))
,((422,430),(437,445))),(((433,434),(435,444)),((436,443),(441,442)))),((((417,418),(419,428)),((420,427),(425
,426))),(((421,429),(438,446)),((488,496),(504,512))))),((((((449,450),(451,460)),((452,459),(457,458))),(((4
53,461),(470,478)),((487,495),(503,511)))),((((454,462),(469,477)),((392,400),(408,416))),(((465,466),(467,476
)),((468,475),(473,474))))),(((((455,463),(471,479)),((486,494),(501,509))),(((497,498),(499,508)),((500,507),
(505,506)))),((((481,482),(483,492)),((484,491),(489,490))),(((485,493),(502,510)),((424,432),(440,448)))))))
),(((((((((513,514),(515,524)),((516,523),(521,522))),(((517,525),(534,542)),((551,559),(567,575)))),((((518,5
26),(533,541)),((584,592),(600,608))),(((529,530),(531,540)),((532,539),(537,538))))),(((((519,527),(535,543))
,((550,558),(565,573))),(((561,562),(563,572)),((564,571),(569,570)))),((((545,546),(547,556)),((548,555),(553
,554))),(((549,557),(566,574)),((616,624),(632,640))))),((((((577,578),(579,588)),((580,587),(585,586))),(((5
81,589),(598,606)),((615,623),(631,639)))),((((582,590),(597,605)),((520,528),(536,544))),(((593,594),(595,604
)),((596,603),(601,602))))),(((((583,591),(599,607)),((614,622),(629,637))),(((625,626),(627,636)),((628,635),
(633,634)))),((((609,610),(611,620)),((612,619),(617,618))),(((613,621),(630,638)),((552,560),(568,576)))))))
,(((((((641,642),(643,652)),((644,651),(649,650))),(((645,653),(662,670)),((679,687),(695,703)))),((((646,654),
(661,669)),((712,720),(728,736))),(((657,658),(659,668)),((660,667),(665,666))))),(((((647,655),(663,671)),((6
78,686),(693,701))),(((689,690),(691,700)),((692,699),(697,698)))),((((673,674),(675,684)),((676,683),(681,682
))),(((677,685),(694,702)),((744,752),(760,768))))),((((((705,706),(707,716)),((708,715),(713,714))),(((709,7
17),(726,734)),((743,751),(759,767)))),((((710,718),(725,733)),((648,656),(664,672))),(((721,722),(723,732)),(
(724,731),(729,730))))),(((((711,719),(727,735)),((742,750),(757,765))),(((753,754),(755,764)),((756,763),(761
,762)))),((((737,738),(739,748)),((740,747),(745,746))),(((741,749),(758,766)),((680,688),(696,704)))))))),(((
(((((769,770),(771,780)),((772,779),(777,778))),(((773,781),(790,798)),((807,815),(823,831)))),((((774,782),(7
89,797)),((840,848),(856,864))),(((785,786),(787,796)),((788,795),(793,794))))),(((((775,783),(791,799)),((806
,814),(821,829))),(((817,818),(819,828)),((820,827),(825,826)))),((((801,802),(803,812)),((804,811),(809,810))
),(((805,813),(822,830)),((872,880),(888,896))))),((((((833,834),(835,844)),((836,843),(841,842))),(((837,845
),(854,862)),((871,879),(887,895)))),((((838,846),(853,861)),((776,784),(792,800))),(((849,850),(851,860)),((8
52,859),(857,858))))),(((((839,847),(855,863)),((870,878),(885,893))),(((881,882),(883,892)),((884,891),(889,8
90)))),((((865,866),(867,876)),((868,875),(873,874))),(((869,877),(886,894)),((808,816),(824,832)))))))),((((((
(897,898),(899,908)),((900,907),(905,906))),(((901,909),(918,926)),((935,943),(951,959)))),((((902,910),(917,9
25)),((968,976),(984,992))),(((913,914),(915,924)),((916,923),(921,922))))),(((((903,911),(919,927)),((934,942

),(949,957))),(((945,946),(947,956)),((948,955),(953,954))),((((929,930),(931,940)),((932,939),(937,938))),((
(933,941),(950,958)),((1000,1008),(1016,1024))))),((((((961,962),(963,972)),((964,971),(969,970))),(((965,973
),(982,990)),((999,1007),(1015,1023)))),((((966,974),(981,989)),((904,912),(920,928))),(((977,978),(979,988)),
((980,987),(985,986))))),(((((967,975),(983,991)),((998,1006),(1013,1021))),(((1009,1010),(1011,1020)),((1012,
1019),(1017,1018)))),((((993,994),(995,1004)),((996,1003),(1001,1002))),(((997,1005),(1014,1022)),((936,944),(
952,960)))))))))),((((((((((1025,1026),(1027,1036)),((1028,1035),(1033,1034))),(((1029,1037),(1046,1054)),((10
63,1071),(1079,1087)))),((((1030,1038),(1045,1053)),((1096,1104),(1112,1120))),(((1041,1042),(1043,1052)),((10
44,1051),(1049,1050))))),(((((1031,1039),(1047,1055)),((1062,1070),(1077,1085))),(((1073,1074),(1075,1084)),((
1076,1083),(1081,1082)))),((((1057,1058),(1059,1068)),((1060,1067),(1065,1066))),(((1061,1069),(1078,1086)),((
1128,1136),(1144,1152))))),((((((1089,1090),(1091,1100)),((1092,1099),(1097,1098))),(((1093,1101),(1110,1118)
),((1127,1135),(1143,1151)))),((((1094,1102),(1109,1117)),((1032,1040),(1048,1056))),(((1105,1106),(1107,1116)
),((1108,1115),(1113,1114))))),(((((1095,1103),(1111,1119)),((1126,1134),(1141,1149))),(((1137,1138),(1139,114
8)),((1140,1147),(1145,1146)))),((((1121,1122),(1123,1132)),((1124,1131),(1129,1130))),(((1125,1133),(1142,115
0)),((1064,1072),(1080,1088)))))),(((((((1153,1154),(1155,1164)),((1156,1163),(1161,1162))),(((1157,1165),(11
74,1182)),((1191,1199),(1207,1215)))),((((1158,1166),(1173,1181)),((1224,1232),(1240,1248))),(((1169,1170),(11
71,1180)),((1172,1179),(1177,1178))))),(((((1159,1167),(1175,1183)),((1190,1198),(1205,1213))),(((1201,1202),(
1203,1212)),((1204,1211),(1209,1210)))),((((1185,1186),(1187,1196)),((1188,1195),(1193,1194))),(((1189,1197),(
1206,1214)),((1256,1264),(1272,1280))))),((((((1217,1218),(1219,1228)),((1220,1227),(1225,1226))),(((1221,122
9),(1238,1246)),((1255,1263),(1271,1279)))),((((1222,1230),(1237,1245)),((1160,1168),(1176,1184))),(((1233,123
4),(1235,1244)),((1236,1243),(1241,1242))))),(((((1223,1231),(1239,1247)),((1254,1262),(1269,1277))),(((1265,1
266),(1267,1276)),((1268,1275),(1273,1274)))),((((1249,1250),(1251,1260)),((1252,1259),(1257,1258))),(((1253,1
261),(1270,1278)),((1192,1200),(1208,1216)))))))),(((((((1281,1282),(1283,1292)),((1284,1291),(1289,1290))),(
((1285,1293),(1302,1310)),((1319,1327),(1335,1343)))),((((1286,1294),(1301,1309)),((1352,1360),(1368,1376))),(
((1297,1298),(1299,1308)),((1300,1307),(1305,1306))))),(((((1287,1295),(1303,1311)),((1318,1326),(1333,1341)))
,(((1329,1330),(1331,1340)),((1332,1339),(1337,1338)))),((((1313,1314),(1315,1324)),((1316,1323),(1321,1322)))
,(((1317,1325),(1334,1342)),((1384,1392),(1400,1408))))),((((((1345,1346),(1347,1356)),((1348,1355),(1353,135
4))),(((1349,1357),(1366,1374)),((1383,1391),(1399,1407)))),((((1350,1358),(1365,1373)),((1288,1296),(1304,131
2))),(((1361,1362),(1363,1372)),((1364,1371),(1369,1370))))),(((((1351,1359),(1367,1375)),((1382,1390),(1397,1
405))),(((1393,1394),(1395,1404)),((1396,1403),(1401,1402)))),((((1377,1378),(1379,1388)),((1380,1387),(1385,1
386))),(((1381,1389),(1398,1406)),((1320,1328),(1336,1344)))))),(((((((1409,1410),(1411,1420)),((1412,1419),(
1417,1418))),(((1413,1421),(1430,1438)),((1447,1455),(1463,1471)))),((((1414,1422),(1429,1437)),((1480,1488),(
1496,1504))),(((1425,1426),(1427,1436)),((1428,1435),(1433,1434))))),(((((1415,1423),(1431,1439)),((1446,1454)
,(1461,1469))),(((1457,1458),(1459,1468)),((1460,1467),(1465,1466)))),((((1441,1442),(1443,1452)),((1444,1451)
,(1449,1450))),(((1445,1453),(1462,1470)),((1512,1520),(1528,1536))))),((((((1473,1474),(1475,1484)),((1476,1
483),(1481,1482))),(((1477,1485),(1494,1502)),((1511,1519),(1527,1535)))),((((1478,1486),(1493,1501)),((1416,1
424),(1432,1440))),(((1489,1490),(1491,1500)),((1492,1499),(1497,1498))))),(((((1479,1487),(1495,1503)),((1510
,1518),(1525,1533))),(((1521,1522),(1523,1532)),((1524,1531),(1529,1530)))),((((1505,1506),(1507,1516)),((1508
,1515),(1513,1514))),(((1509,1517),(1526,1534)),((1448,1456),(1464,1472)))))))),(((((((((1537,1538),(1539,154
8)),((1540,1547),(1545,1546))),(((1541,1549),(1558,1566)),((1575,1583),(1591,1599)))),((((1542,1550),(1557,156
5)),((1608,1616),(1624,1632))),(((1553,1554),(1555,1564)),((1556,1563),(1561,1562))))),(((((1543,1551),(1559,1
567)),((1574,1582),(1589,1597))),(((1585,1586),(1587,1596)),((1588,1595),(1593,1594)))),((((1569,1570),(1571,1
580)),((1572,1579),(1577,1578))),(((1573,1581),(1590,1598)),((1640,1648),(1656,1664))))),((((((1601,1602),(16
03,1612)),((1604,1611),(1609,1610))),(((1605,1613),(1622,1630)),((1639,1647),(1655,1663)))),((((1606,1614),(16
21,1629)),((1544,1552),(1560,1568))),(((1617,1618),(1619,1628)),((1620,1627),(1625,1626))))),(((((1607,1615),(
1623,1631)),((1638,1646),(1653,1661))),(((1649,1650),(1651,1660)),((1652,1659),(1657,1658)))),((((1633,1634),(
1635,1644)),((1636,1643),(1641,1642))),(((1637,1645),(1654,1662)),((1576,1584),(1592,1600)))))),(((((((1665,1
666),(1667,1676)),((1668,1675),(1673,1674))),(((1669,1677),(1686,1694)),((1703,1711),(1719,1727)))),((((1670,1
678),(1685,1693)),((1736,1744),(1752,1760))),(((1681,1682),(1683,1692)),((1684,1691),(1689,1690))))),(((((1671
,1679),(1687,1695)),((1702,1710),(1717,1725))),(((1713,1714),(1715,1724)),((1716,1723),(1721,1722)))),((((1697
,1698),(1699,1708)),((1700,1707),(1705,1706))),(((1701,1709),(1718,1726)),((1768,1776),(1784,1792))))))),((((((
1729,1730),(1731,1740)),((1732,1739),(1737,1738))),(((1733,1741),(1750,1758)),((1767,1775),(1783,1791)))),((((
1734,1742),(1749,1757)),((1672,1680),(1688,1696))),(((1745,1746),(1747,1756)),((1748,1755),(1753,1754))))),(((
((1735,1743),(1751,1759)),((1766,1774),(1781,1789))),(((1777,1778),(1779,1788)),((1780,1787),(1785,1786)))),((

((1761,1762),(1763,1772)),((1764,1771),(1769,1770))),(((1765,1773),(1782,1790)),((1704,1712),(1720,1728)))))))
),(((((((1793,1794),(1795,1804)),((1796,1803),(1801,1802))),(((1797,1805),(1814,1822)),((1831,1839),(1847,185
5)))),((((1798,1806),(1813,1821)),((1864,1872),(1880,1888))),(((1809,1810),(1811,1820)),((1812,1819),(1817,181
8)))),(((((1799,1807),(1815,1823)),((1830,1838),(1845,1853))),(((1841,1842),(1843,1852)),((1844,1851),(1849,1
850)))),((((1825,1826),(1827,1836)),((1828,1835),(1833,1834))),(((1829,1837),(1846,1854)),((1896,1904),(1912,1
920)))))),((((((1857,1858),(1859,1868)),((1860,1867),(1865,1866))),(((1861,1869),(1878,1886)),((1895,1903),(19
11,1919)))),((((1862,1870),(1877,1885)),((1800,1808),(1816,1824))),(((1873,1874),(1875,1884)),((1876,1883),(18
81,1882)))),(((((1863,1871),(1879,1887)),((1894,1902),(1909,1917))),(((1905,1906),(1907,1916)),((1908,1915),(
1913,1914)))),((((1889,1890),(1891,1900)),((1892,1899),(1897,1898))),(((1893,1901),(1910,1918)),((1832,1840),(
1848,1856)))))),(((((((1921,1922),(1923,1932)),((1924,1931),(1929,1930))),(((1925,1933),(1942,1950)),((1959,1
967),(1975,1983)))),((((1926,1934),(1941,1949)),((1992,2000),(2008,2016))),(((1937,1938),(1939,1948)),((1940,1
947),(1945,1946)))),(((((1927,1935),(1943,1951)),((1958,1966),(1973,1981))),(((1969,1970),(1971,1980)),((1972
,1979),(1977,1978)))),((((1953,1954),(1955,1964)),((1956,1963),(1961,1962))),(((1957,1965),(1974,1982)),((2024
,2032),(2040,2048)))))),((((((1985,1986),(1987,1996)),((1988,1995),(1993,1994))),(((1989,1997),(2006,2014)),((
2023,2031),(2039,2047)))),((((1990,1998),(2005,2013)),((1928,1936),(1944,1952))),(((2001,2002),(2003,2012)),((
2004,2011),(2009,2010)))),(((((1991,1999),(2007,2015)),((2022,2030),(2037,2045))),(((2033,2034),(2035,2044)),
((2036,2043),(2041,2042)))),((((2017,2018),(2019,2028)),((2020,2027),(2025,2026))),(((2021,2029),(2038,2046)),
((1960,1968),(1976,1984)))))))))));

$\mathcal{T} =$ ((((((((((8,7),(6,133)),((5,134),(136,135))),(((4,132),(259,387)),((514,642),(770,898)))),((((3,131)
,(260,388)),((1025,1153),(1281,1409))),(((264,263),(262,389)),((261,390),(392,391))))),(((((2,130),(258,386)),
((515,643),(772,900))),(((776,775),(774,901)),((773,902),(904,903)))),((((520,519),(518,645)),((517,646),(648,
647))),(((516,644),(771,899)),((1537,1665),(1793,1921)))))),((((((1032,1031),(1030,1157)),((1029,1158),(1160,1
159))),(((1028,1156),(1283,1411)),((1538,1666),(1794,1922)))),((((1027,1155),(1284,1412)),((1,129),(257,385)))
,(((1288,1287),(1286,1413)),((1285,1414),(1416,1415))))),(((((1026,1154),(1282,1410)),((1539,1667),(1796,1924)
)),(((1800,1799),(1798,1925)),((1797,1926),(1928,1927)))),((((1544,1543),(1542,1669)),((1541,1670),(1672,1671)
)),(((1540,1668),(1795,1923)),((513,641),(769,897)))))),(((((((16,15),(14,141)),((13,142),(144,143))),(((12,1
40),(267,395)),((522,650),(778,906)))),((((11,139),(268,396)),((1033,1161),(1289,1417))),(((272,271),(270,397)
),((269,398),(400,399))))),(((((10,138),(266,394)),((523,651),(780,908))),(((784,783),(782,909)),((781,910),(9
12,911)))),((((528,527),(526,653)),((525,654),(656,655))),(((524,652),(779,907)),((1545,1673),(1801,1929))))))
,((((((1040,1039),(1038,1165)),((1037,1166),(1168,1167))),(((1036,1164),(1291,1419)),((1546,1674),(1802,1930))
)),((((1035,1163),(1292,1420)),((9,137),(265,393))),(((1296,1295),(1294,1421)),((1293,1422),(1424,1423))))),((
(((1034,1162),(1290,1418)),((1547,1675),(1804,1932))),(((1808,1807),(1806,1933)),((1805,1934),(1936,1935)))),(
(((1552,1551),(1550,1677)),((1549,1678),(1680,1679))),(((1548,1676),(1803,1931)),((521,649),(777,905)))))))),(
(((((((24,23),(22,149)),((21,150),(152,151))),(((20,148),(275,403)),((530,658),(786,914)))),((((19,147),(276,4
04)),((1041,1169),(1297,1425))),(((280,279),(278,405)),((277,406),(408,407))))),(((((18,146),(274,402)),((531,
659),(788,916))),(((792,791),(790,917)),((789,918),(920,919)))),((((536,535),(534,661)),((533,662),(664,663)))
,(((532,660),(787,915)),((1553,1681),(1809,1937)))))),((((((1048,1047),(1046,1173)),((1045,1174),(1176,1175)))
,(((1044,1172),(1299,1427)),((1554,1682),(1810,1938)))),((((1043,1171),(1300,1428)),((17,145),(273,401))),(((1
304,1303),(1302,1429)),((1301,1430),(1432,1431))))),(((((1042,1170),(1298,1426)),((1555,1683),(1812,1940))),((
(1816,1815),(1814,1941)),((1813,1942),(1944,1943)))),((((1560,1559),(1558,1685)),((1557,1686),(1688,1687))),((
(1556,1684),(1811,1939)),((529,657),(785,913)))))),(((((((32,31),(30,157)),((29,158),(160,159))),(((28,156),(
283,411)),((538,666),(794,922)))),((((27,155),(284,412)),((1049,1177),(1305,1433))),(((288,287),(286,413)),((2
85,414),(416,415))))),(((((26,154),(282,410)),((539,667),(796,924))),(((800,799),(798,925)),((797,926),(928,92
7)))),((((544,543),(542,669)),((541,670),(672,671))),(((540,668),(795,923)),((1561,1689),(1817,1945))))),((((
((1056,1055),(1054,1181)),((1053,1182),(1184,1183))),(((1052,1180),(1307,1435)),((1562,1690),(1818,1946)))),((
((1051,1179),(1308,1436)),((25,153),(281,409))),(((1312,1311),(1310,1437)),((1309,1438),(1440,1439))))),(((((1
050,1178),(1306,1434)),((1563,1691),(1820,1948))),(((1824,1823),(1822,1949)),((1821,1950),(1952,1951)))),((((1
568,1567),(1566,1693)),((1565,1694),(1696,1695))),(((1564,1692),(1819,1947)),((537,665),(793,921)))))))),((((
(((((40,39),(38,165)),((37,166),(168,167))),(((36,164),(291,419)),((546,674),(802,930)))),((((35,163),(292,420
)),((1057,1185),(1313,1441))),(((296,295),(294,421)),((293,422),(424,423))))),(((((34,162),(290,418)),((547,67
5),(804,932))),(((808,807),(806,933)),((805,934),(936,935)))),((((552,551),(550,677)),((549,678),(680,679))),(
((548,676),(803,931)),((1569,1697),(1825,1953)))))),((((((1064,1063),(1062,1189)),((1061,1190),(1192,1191))),(
((1060,1188),(1315,1443)),((1570,1698),(1826,1954)))),((((1059,1187),(1316,1444)),((33,161),(289,417))),(((132
0,1319),(1318,1445)),((1317,1446),(1448,1447))))),(((((1058,1186),(1314,1442)),((1571,1699),(1828,1956))),(((1
832,1831),(1830,1957)),((1829,1958),(1960,1959)))),((((1576,1575),(1574,1701)),((1573,1702),(1704,1703))),(((1
572,1700),(1827,1955)),((545,673),(801,929)))))),(((((((48,47),(46,173)),((45,174),(176,175))),(((44,172),(29
9,427)),((554,682),(810,938)))),((((43,171),(300,428)),((1065,1193),(1321,1449))),(((304,303),(302,429)),((301
,430),(432,431))))),(((((42,170),(298,426)),((555,683),(812,940))),(((816,815),(814,941)),((813,942),(944,943)
))),((((560,559),(558,685)),((557,686),(688,687))),(((556,684),(811,939)),((1577,1705),(1833,1961)))))),((((((
1072,1071),(1070,1197)),((1069,1198),(1200,1199))),(((1068,1196),(1323,1451)),((1578,1706),(1834,1962)))),((((
1067,1195),(1324,1452)),((41,169),(297,425))),(((1328,1327),(1326,1453)),((1325,1454),(1456,1455))))),(((((106
6,1194),(1322,1450)),((1579,1707),(1836,1964))),(((1840,1839),(1838,1965)),((1837,1966),(1968,1967)))),((((158

4,1583),(1582,1709)),((1581,1710),(1712,1711))),(((1580,1708),(1835,1963)),((553,681),(809,937)))))))),(((((((
(56,55),(54,181)),((53,182),(184,183))),(((52,180),(307,435)),((562,690),(818,946)))),((((51,179),(308,436)),(
(1073,1201),(1329,1457))),(((312,311),(310,437)),((309,438),(440,439)))))),(((((50,178),(306,434)),((563,691),(
820,948))),(((824,823),(822,949)),((821,950),(952,951)))),((((568,567),(566,693)),((565,694),(696,695))),(((56
4,692),(819,947)),((1585,1713),(1841,1969))))))),((((((1080,1079),(1078,1205)),((1077,1206),(1208,1207))),(((10
76,1204),(1331,1459)),((1586,1714),(1842,1970)))),((((1075,1203),(1332,1460)),((49,177),(305,433))),(((1336,13
35),(1334,1461)),((1333,1462),(1464,1463))))),(((((1074,1202),(1330,1458)),((1587,1715),(1844,1972))),(((1848,
1847),(1846,1973)),((1845,1974),(1976,1975)))),((((1592,1591),(1590,1717)),((1589,1718),(1720,1719))),(((1588,
1716),(1843,1971)),((561,689),(817,945)))))))),(((((((64,63),(62,189)),((61,190),(192,191))),(((60,188),(315,44
3)),((570,698),(826,954)))),((((59,187),(316,444)),((1081,1209),(1337,1465))),(((320,319),(318,445)),((317,446
),(448,447))))),(((((58,186),(314,442)),((571,699),(828,956))),(((832,831),(830,957)),((829,958),(960,959)))),
((((576,575),(574,701)),((573,702),(704,703))),(((572,700),(827,955)),((1593,1721),(1849,1977))))))),((((((1088
,1087),(1086,1213)),((1085,1214),(1216,1215))),(((1084,1212),(1339,1467)),((1594,1722),(1850,1978)))),((((1083
,1211),(1340,1468)),((57,185),(313,441))),(((1344,1343),(1342,1469)),((1341,1470),(1472,1471))))),(((((1082,12
10),(1338,1466)),((1595,1723),(1852,1980))),(((1856,1855),(1854,1981)),((1853,1982),(1984,1983)))),((((1600,15
99),(1598,1725)),((1597,1726),(1728,1727))),(((1596,1724),(1851,1979)),((569,697),(825,953)))))))))),(((((((((
(72,71),(70,197)),((69,198),(200,199))),(((68,196),(323,451)),((578,706),(834,962)))),((((67,195),(324,452)),(
(1089,1217),(1345,1473))),(((328,327),(326,453)),((325,454),(456,455))))),(((((66,194),(322,450)),((579,707),(
836,964))),(((840,839),(838,965)),((837,966),(968,967)))),((((584,583),(582,709)),((581,710),(712,711))),(((58
0,708),(835,963)),((1601,1729),(1857,1985))))))),((((((1096,1095),(1094,1221)),((1093,1222),(1224,1223))),(((10
92,1220),(1347,1475)),((1602,1730),(1858,1986)))),((((1091,1219),(1348,1476)),((65,193),(321,449))),(((1352,13
51),(1350,1477)),((1349,1478),(1480,1479))))),(((((1090,1218),(1346,1474)),((1603,1731),(1860,1988))),(((1864,
1863),(1862,1989)),((1861,1990),(1992,1991)))),((((1608,1607),(1606,1733)),((1605,1734),(1736,1735))),(((1604,
1732),(1859,1987)),((577,705),(833,961)))))))),(((((((80,79),(78,205)),((77,206),(208,207))),(((76,204),(331,45
9)),((586,714),(842,970)))),((((75,203),(332,460)),((1097,1225),(1353,1481))),(((336,335),(334,461)),((333,462
),(464,463))))),(((((74,202),(330,458)),((587,715),(844,972))),(((848,847),(846,973)),((845,974),(976,975)))),
((((592,591),(590,717)),((589,718),(720,719))),(((588,716),(843,971)),((1609,1737),(1865,1993))))))),((((((1104
,1103),(1102,1229)),((1101,1230),(1232,1231))),(((1100,1228),(1355,1483)),((1610,1738),(1866,1994)))),((((1099
,1227),(1356,1484)),((73,201),(329,457))),(((1360,1359),(1358,1485)),((1357,1486),(1488,1487))))),(((((1098,12
26),(1354,1482)),((1611,1739),(1868,1996))),(((1872,1871),(1870,1997)),((1869,1998),(2000,1999)))),((((1616,16
15),(1614,1741)),((1613,1742),(1744,1743))),(((1612,1740),(1867,1995)),((585,713),(841,969)))))))),(((((((88,
87),(86,213)),((85,214),(216,215))),(((84,212),(339,467)),((594,722),(850,978)))),((((83,211),(340,468)),((110
5,1233),(1361,1489))),(((344,343),(342,469)),((341,470),(472,471))))),(((((82,210),(338,466)),((595,723),(852,
980))),(((856,855),(854,981)),((853,982),(984,983)))),((((600,599),(598,725)),((597,726),(728,727))),(((596,72
4),(851,979)),((1617,1745),(1873,2001))))))),((((((1112,1111),(1110,1237)),((1109,1238),(1240,1239))),(((1108,1
236),(1363,1491)),((1618,1746),(1874,2002)))),((((1107,1235),(1364,1492)),((81,209),(337,465))),(((1368,1367),
(1366,1493)),((1365,1494),(1496,1495))))),(((((1106,1234),(1362,1490)),((1619,1747),(1876,2004))),(((1880,1879
),(1878,2005)),((1877,2006),(2008,2007)))),((((1624,1623),(1622,1749)),((1621,1750),(1752,1751))),(((1620,1748
),(1875,2003)),((593,721),(849,977)))))))),(((((((96,95),(94,221)),((93,222),(224,223))),(((92,220),(347,475)),
((602,730),(858,986)))),((((91,219),(348,476)),((1113,1241),(1369,1497))),(((352,351),(350,477)),((349,478),(4
80,479))))),(((((90,218),(346,474)),((603,731),(860,988))),(((864,863),(862,989)),((861,990),(992,991)))),((((
608,607),(606,733)),((605,734),(736,735))),(((604,732),(859,987)),((1625,1753),(1881,2009))))))),((((((1120,111
9),(1118,1245)),((1117,1246),(1248,1247))),(((1116,1244),(1371,1499)),((1626,1754),(1882,2010)))),((((1115,124
3),(1372,1500)),((89,217),(345,473))),(((1376,1375),(1374,1501)),((1373,1502),(1504,1503))))),(((((1114,1242),
(1370,1498)),((1627,1755),(1884,2012))),(((1888,1887),(1886,2013)),((1885,2014),(2016,2015)))),((((1632,1631),
(1630,1757)),((1629,1758),(1760,1759))),(((1628,1756),(1883,2011)),((601,729),(857,985)))))))),(((((((((104,1
03),(102,229)),((101,230),(232,231))),(((100,228),(355,483)),((610,738),(866,994)))),((((99,227),(356,484)),((
1121,1249),(1377,1505))),(((360,359),(358,485)),((357,486),(488,487))))),(((((98,226),(354,482)),((611,739),(8
68,996))),(((872,871),(870,997)),((869,998),(1000,999)))),((((616,615),(614,741)),((613,742),(744,743))),(((61
2,740),(867,995)),((1633,1761),(1889,2017))))))),((((((1128,1127),(1126,1253)),((1125,1254),(1256,1255))),(((11
24,1252),(1379,1507)),((1634,1762),(1890,2018)))),((((1123,1251),(1380,1508)),((97,225),(353,481))),(((1384,13
83),(1382,1509)),((1381,1510),(1512,1511))))),(((((1122,1250),(1378,1506)),((1635,1763),(1892,2020))),(((1896,
1895),(1894,2021)),((1893,2022),(2024,2023)))),((((1640,1639),(1638,1765)),((1637,1766),(1768,1767))),(((1636,
1764),(1891,2019)),((609,737),(865,993)))))))),(((((((112,111),(110,237)),((109,238),(240,239))),(((108,236),(3
63,491)),((618,746),(874,1002)))),((((107,235),(364,492)),((1129,1257),(1385,1513))),(((368,367),(366,493)),((
365,494),(496,495))))),(((((106,234),(362,490)),((619,747),(876,1004))),(((880,879),(878,1005)),((877,1006),(1
008,1007)))),((((624,623),(622,749)),((621,750),(752,751))),(((620,748),(875,1003)),((1641,1769),(1897,2025)))
))),((((((1136,1135),(1134,1261)),((1133,1262),(1264,1263))),(((1132,1260),(1387,1515)),((1642,1770),(1898,202
6)))),((((1131,1259),(1388,1516)),((105,233),(361,489))),(((1392,1391),(1390,1517)),((1389,1518),(1520,1519)))
)),(((((1130,1258),(1386,1514)),((1643,1771),(1900,2028))),(((1904,1903),(1902,2029)),((1901,2030),(2032,2031)
))),((((1648,1647),(1646,1773)),((1645,1774),(1776,1775))),(((1644,1772),(1899,2027)),((617,745),(873,1001))))
)))),(((((((((120,119),(118,245)),((117,246),(248,247))),(((116,244),(371,499)),((626,754),(882,1010)))),((((11
5,243),(372,500)),((1137,1265),(1393,1521))),(((376,375),(374,501)),((373,502),(504,503))))),(((((114,242),(37
0,498)),((627,755),(884,1012))),(((888,887),(886,1013)),((885,1014),(1016,1015)))),((((632,631),(630,757)),((6
29,758),(760,759))),(((628,756),(883,1011)),((1649,1777),(1905,2033))))))),((((((1144,1143),(1142,1269)),((1141
,1270),(1272,1271))),(((1140,1268),(1395,1523)),((1650,1778),(1906,2034)))),((((1139,1267),(1396,1524)),((113,

```
241),(369,497))),(((1400,1399),(1398,1525)),((1397,1526),(1528,1527)))))),(((((1138,1266),(1394,1522)),((1651,1
779),(1908,2036))),(((1912,1911),(1910,2037)),((1909,2038),(2040,2039)))),((((1656,1655),(1654,1781)),((1653,1
782),(1784,1783))),(((1652,1780),(1907,2035)),((625,753),(881,1009)))))),(((((((128,127),(126,253)),((125,254
),(256,255))),(((124,252),(379,507)),((634,762),(890,1018)))),((((123,251),(380,508)),((1145,1273),(1401,1529)
)),(((384,383),(382,509)),((381,510),(512,511))))),(((((122,250),(378,506)),((635,763),(892,1020))),(((896,895
),(894,1021)),((893,1022),(1024,1023)))),((((640,639),(638,765)),((637,766),(768,767))),(((636,764),(891,1019)
),((1657,1785),(1913,2041))))))),((((((1152,1151),(1150,1277)),((1149,1278),(1280,1279))),(((1148,1276),(1403,1
531)),((1658,1786),(1914,2042)))),((((1147,1275),(1404,1532)),((121,249),(377,505))),(((1408,1407),(1406,1533)
),((1405,1534),(1536,1535))))),(((((1146,1274),(1402,1530)),((1659,1787),(1916,2044))),(((1920,1919),(1918,204
5)),((1917,2046),(2048,2047)))),((((1664,1663),(1662,1789)),((1661,1790),(1792,1791))),(((1660,1788),(1915,204
3)),((633,761),(889,1017)))))))))));
```