involve
a journal of mathematics

Combinatorics of linked systems of quartet trees

Emili Moan and Joseph Rusinko

msp

2016 vol. 9, no. 1

INVOLVE 9:1 (2016)
dx.doi.org/10.2140/involve.2016.9.171

Combinatorics of linked systems of quartet trees

Emili Moan and Joseph Rusinko

(Communicated by Kenneth S. Berenhaut)

We apply classical quartet techniques to the problem of phylogenetic decisiveness and find a value k such that all collections of at least k quartets are decisive. Moreover, we prove that this bound is optimal and give a lower bound on the probability that a collection of quartets is decisive.

1. Overview

Evolutionary biologists represent relationships between groups of organisms with phylogenetic trees. Supertree methods were designed to handle the computationally difficult problem of reconstructing such trees for large data sets. Those methods generate a group of accurate, smaller input trees and combine them into a single supertree. Four-taxa trees, known as quartet trees, are commonly used as inputs in supertree methods.

Most quartet amalgamation algorithms use all quartet trees generated from sequencing data or only remove quartet trees that appear to be incorrect. As quartet trees may contain overlapping information, it is possible that a smaller number of trees may provide sufficient information for accurate reconstruction.

Böcker et al. [1999] developed a sufficient condition for a set of quartet trees to be definitive. For any tree on a taxon set $[n] = \{1, 2, \dots, n\}$, we develop a system of quartet trees that meets the criteria of Böcker et al., known as a linked system. Additionally, we develop collections of linked systems known as meshed systems.

Recently, Steel and Sanderson asked for which collections of sets of taxa do the corresponding induced subtrees determine a unique supertree. They called such collections decisive. The notion of decisiveness can be viewed as a generalization of definitiveness where no information is required about the particular subtrees that the subsets of taxa induce. This notion plays an important role in supertree reconstruction since it a priori addresses the question about which subsets of taxa must be analyzed to ensure that a unique supertree can be reconstructed.

MSC2010: 92B10.
Keywords: phylogenetics, quartets, decisiveness.

We use the term quartet to refer to any four-element taxon subset, and the term quartet tree when referencing a resolved four-taxa tree. Using meshed systems, we find a minimal number $k(n)$ such that every collection of at least k quartets is decisive. We use this number to find a lower bound on the probability that an arbitrary collection of quartets is decisive.

Finally, we find that meshed systems may be useful in amalgamation algorithms, such as maxcut [Snir and Rao 2008], that do not always find the correct supertree when given a definitive system of quartet trees.

2. Linked systems

We adopt the terminology in [Dress et al. 2012], except in noted instances when we follow [Semple and Steel 2003] or [Steel and Sanderson 2010]. Phylogenetic trees display relationships among a finite set of taxonomic units.

Definition 2.1. A *binary phylogenetic tree*, $T = (V, E, \varphi)$ on a finite set of taxa X, is a triple consisting of a finite set of vertices V, a set E of edges between vertices, and a labeling map $\varphi : X \to L$, where $L \subset V$ contains all vertices of degree one, or *leaves*, such that the graph (V, E) is an unrooted binary tree and the map φ induces a bijection between X and the set L of leaves of T.

An edge that contains a leaf is an *exterior edge*. The nonleaf vertex of an exterior edge is the *internal vertex of* e, denoted $v_{\text{int}}(e)$. Two exterior edges sharing an internal vertex form a *cherry*. Any edge that is not an exterior edge is an *interior edge*.

While edge length plays an important role in phylogenetics, we do not take it into account, and adopt instead a topological definition of tree isomorphism.

Definition 2.2. Phylogenetic trees, $T_1 = (V_1, E_1, \varphi_1)$ and $T_2 = (V_2, E_2, \varphi_2)$ on a taxon set X, are *isomorphic* if there exists a bijective map $f : V_1 \to V_2$, called an *isomorphism*, such that if $\{u, v\} \in E_1$ then $\{f(u), f(v)\} \in E_2$ and for every $x \in X$, we have $\varphi_2(x) = f(\varphi_1(x))$.

It is impossible to distinguish phylogenetic relationships from unrooted trees with fewer than four taxa; thus, supertree reconstruction algorithms frequently use four-taxa trees or *quartets trees* as inputs [Snir and Rao 2008; Snir et al. 2008; Strimmer and von Haeseler 1996]. *Quartet trees* are binary phylogenetic trees on four leaves. Such trees are in one-to-one correspondence with two-element subsets of X, such as $\{\{a, b\}, \{c, d\}\}$, according to the separation of the four leaves by the interior edge. The union of all four taxa is the *support of* q, denoted $\operatorname{supp}(q)$.

Quartet trees contain an interior edge which separates the taxa into two pairs. Similarly, removing an interior edge of a tree separates the graph into two connected components. An edge e *separates* taxa a and b from c and d if $\{a, b\}$ and $\{c, d\}$ are subsets of the vertex sets of different connected components of $T - \{e\}$. This separation points to a relationship between edges of a tree and quartet trees. A quartet tree $ab|cd$ is *displayed by a binary phylogenetic tree* T if there exists an edge $e \in E$ that separates a and b from c and d.

Denote the set of all quartet trees on a taxon set X by $Q(X)$. Any subset Q of $Q(X)$ is called a *system of quartet trees* on X with the support defined by $\operatorname{supp}(Q) = \bigcup_{q \in Q} \operatorname{supp}(q)$. Additionally, we denote the set of all quartet trees displayed by a tree T by Q_T. A system of quartet trees Q is *compatible* if there exists a tree T such that $Q \subseteq Q_T$.

Definition 2.3 [Semple and Steel 2003]. Let $T = (V, E, \varphi)$ be a binary phylogenetic tree and let $ab|cd \in Q_T$. An interior edge e of T is *distinguished* by $ab|cd$ if e is the only edge that separates a and b from c and d.

Quartet trees which distinguish edges are a powerful input to quartet amalgamation algorithms. These algorithms must handle noncompatible systems of quartet trees. However, even compatible systems may be difficult to resolve as multiple trees may display a particular collection of quartet trees.

Definition 2.4 [Steel and Sanderson 2010]. A system of quartet trees Q is *definitive* if up to isomorphism, there is a unique binary phylogenetic tree T for which $Q \subseteq Q_T$.

Böcker et al. [1999] described various criteria for a system of quartet trees of size $n-3$ to be definitive. We construct systems of quartet trees that meet this criteria and make note of some useful applications of these systems.

Proposition 2.5 [Böcker et al. 1999, Example 3.7]. *If T is a binary tree such that the interior edges of T are labeled $E = \{e_1, \dots, e_{n-3}\}$, and Q is a system of quartet trees such that each $q_i \in Q$ distinguishes a unique edge e_i in T with*

$$\left| \operatorname{supp}(q_i) \setminus \bigcup_{j<i} \operatorname{supp}(q_j) \right| = 1$$

for $i = 2, \dots, n-3$, then Q is definitive.

We create a system of quartet trees that satisfies the hypotheses of Proposition 2.5, known as a linked system, by imposing an ordering on the interior edges of a tree and the quartet trees which distinguish those edges. We define linked systems in terms of the associated graph.

Definition 2.6. For a compatible system of $n-3$ quartet trees Q on a taxon set X, define the associated graph $G_T(Q)$ with vertex set V and edge set E as follows:

- The vertex set V is the set of all quartet trees $q \in Q$ which distinguish a unique edge in T.
- Vertex pairs $\{q_i, q_j\}$ are connected by an edge $e \in E$ if the edge e_i that q_i distinguishes is adjacent to the edge e_j that q_j distinguishes and $|\operatorname{supp}(\{q_i, q_j\})| = 5$.

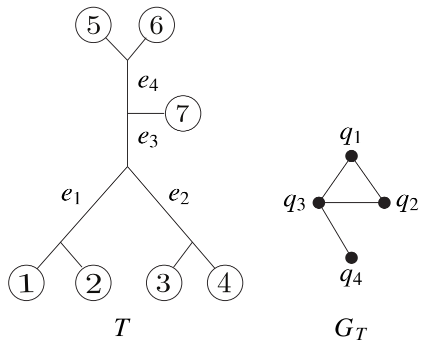

Figure 1. A binary phylogenetic tree T and the associated graph $G_T(Q)$ for the quartet trees $q_1 = 12|35$, $q_2 = 34|15$, $q_3 = 57|13$ and $q_4 = 56|71$.

Definition 2.7. Two quartet trees are *linked* if their vertices are connected in $G_T(Q)$. The system of quartet trees Q is *a linked system* if $G_T(Q)$ is connected. See Figure 1 for an example.

In Section 3 we prove that linked systems are definitive. Linked systems also help illuminate the broader concept of phylogenetic decisiveness, which we review here.

For a binary phylogenetic tree T and a subset Y of X, let $T|Y$ denote the induced binary phylogenetic tree on leaf set Y (the tree obtained from the minimal subtree connecting Y by suppressing any vertices of degree 2). Let $\mathcal{S}$ be the collection of subsets of a set X of size four; we refer to all such subsets as *quartets* [Steel and Sanderson 2010].

Definition 2.8 [Steel and Sanderson 2010]. We say that $\mathcal{S}$ is *phylogenetically decisive* if it satisfies the following property: if T and T' are binary phylogenetic trees, with $T|Y = T'|Y$ for all $Y \in \mathcal{S}$, then $T = T'$.

We will use collections of linked systems to find the minimal number $k(n)$ such that a collection of at least k quartets is phylogenetically decisive.

3. Applications of linked systems

We first show that linked systems meet the criteria of [Böcker et al. 1999] for defining a unique tree.

Theorem 3.1. *Every linked system of quartet trees is definitive.*

Proof. Let Q be a linked system of quartet trees on a tree T on a taxon set X. Linked systems are of size $n - 3$ and each quartet tree distinguishes a unique edge in T. Let T be a binary phylogenetic tree on a taxon set X and let e_1 be an interior edge adjacent to a cherry. The tree is connected, which implies we can label the remaining interior edges $\{e_2, \dots, e_{n-3}\}$ such that e_j is adjacent to some e_i with $i < j$. Moreover, because the support of each pair of quartet trees $\{q_i, q_j\}$ that distinguishes adjacent edges $\{e_i, e_j\}$ is five, each pair of quartet trees shares three taxa and each additional quartet tree in Q adds one new taxon to the support of Q. Thus, linked systems meet the criteria in Proposition 2.5 and are definitive. □

Though all linked systems are definitive, we find that not all definitive systems of size $n-3$ are linked.

Example. The system of quartet trees $Q = \{12|36, 23|45, 24|56\}$ meets the criteria established in Proposition 2.5, and is thus definitive. The graph $G_T(Q)$ contains the three vertices q_1, q_2, and q_3, where q_2 and q_3 are connected by an edge and q_1 is an isolated vertex. Thus, Q is not a linked system.

Since a system satisfying Proposition 2.5 and linked systems both contain $n-3$ quartet trees, one might surmise that all compatible systems of quartet trees of a modest size would be definitive. However, there are large systems of compatible quartet trees which are not definitive and large collections of quartets which are not decisive. A collection of quartets and the induced quartet trees on a caterpillar tree provides one such example.

Definition 3.2 [Semple and Steel 2003]. A *caterpillar* on n leaves is a binary phylogenetic tree for which there exists an induced subtree on a sequence of distinct interior vertices $v_1, v_2, \dots, v_k$ such that, for all $i \in \{1, 2, \dots, k-1\}$, v_i and v_{i+1} are adjacent.

The ordering of vertices in a caterpillar tree induces an ordering of interior edges e_i, where e_i connects v_i with v_{i+1}. We use this ordering to construct large families of quartet trees shared by several caterpillar trees.

Theorem 3.3. *The minimal number $k(n)$ such that every collection of quartets S with $|S| \geq k$ is decisive is greater than $\binom{n}{4} - (n-3)$.*

Proof. Let $\{a, b, c\}$ be a subset of X. We define T_1, T_2, and T_3 to be three distinct caterpillar trees of size $n \geq 4$ that differ only in the placement of three taxa a, b, and c such that for each tree, $v_{\text{int}}(a)$, $v_{\text{int}}(b)$ and $v_{\text{int}}(c)$ are incident to e_1. Denote by $\mathcal{S}$ the set of $n-3$ sets

$$\bigcup_{i=1}^{n-3} Y_i = \{a, b, c, y \mid y \in X - \{a, b, c\}\}$$

and let $\mathcal{S}'$ be the complement of $\mathcal{S}$. We observe that for all $Y \in \mathcal{S}'$, we have $T_1|Y = T_2|Y = T_3|Y$, but $T_1 \neq T_2 \neq T_3$. Therefore, S' is not decisive. Since $|\mathcal{S}'| = \binom{n}{4} - (n-3)$, the minimal number $k(n)$ such that every collection $\mathcal{S}$ of quartets with $|\mathcal{S}| \geq k$ is decisive is greater than $\binom{n}{4} - (n-3)$. □

To show that sets of quartets of size $\binom{n}{4} - (n-3)$ are decisive, we prove that Q_T contains at least $n-3$ disjoint linked systems, ensuring the removal of any $n-4$ quartet trees from a compatible system would leave at least one linked system. We introduce a process for building such systems by using a seed quartet tree which distinguishes an edge of a tree, and systematically constructing additional quartet trees which distinguish the same edge.

In a phylogenetic tree, each interior edge $e = (v_l, v_r)$ is adjacent to four edges e_i, e_j, e_h, and e_k, which divide the tree into four components and partition the set of taxa X into four distinct sets A_i, A_j, A_k, and A_h, with $x \in A_n$ if the unique path from x to v_l contains the edge e_n.

Definition 3.4. Let $q = ij\,|\,kh$ be a quartet tree that distinguishes an edge e and let $i \in A_i$, $j \in A_j$, $k \in A_k$, and $h \in A_h$, where A_i, A_j, A_k, and A_h are partitions of X induced by e. For $x \in X - \mathrm{supp}(q)$, define the *quartet-tree substitution* $q(x)$ to be the unique quartet tree in which the taxon $x \in A_n$ replaces the taxon in q that is in $\mathrm{supp}(q) \cap A_n$.

Notice $q(x)$ and q must distinguish the same edge of the tree.

Definition 3.5. Let the quartet tree q distinguish an edge e of a tree. Define the *vine of* q by $v(q) = \{q\} \cup \bigcup_{x \in X - \mathrm{supp}(q)} q(x)$. We refer to q as the *seed* of the vine.

The following shows that if two quartet trees are linked, then so are their vines.

Definition 3.6. Two vines $v(q_i)$ and $v(q_j)$ are *linked* if for each $q_i \in v(q_i)$ there exists a unique $q_j \in v(q_j)$ such that q_i and q_j are linked.

Theorem 3.7. *If q_i and q_j are linked quartet trees, then the associated vines $v(q_i)$ and $v(q_j)$ are linked.*

Proof. Assume that $\{q_i, q_j\}$ are the seeds of the adjacent edges e_i and e_j and are linked in T. Let $v(q_i)$ and $v(q_j)$ be the associated vines.

Since $\mathrm{supp}(q_i, q_j) = 5$, each quartet tree contains one taxon that the other does not. Let z be the taxon in $\mathrm{supp}(q_j) - \mathrm{supp}(q_i)$ and y be the taxon in $\mathrm{supp}(q_i) - \mathrm{supp}(q_j)$. Use quartet-tree substitution to construct the quartet trees $q_i(z)$ and $q_j(y)$. By construction, $\mathrm{supp}(q_i, q_j) = \mathrm{supp}(q_i(z), q_j(y))$ and $q_i(z)$ and $q_j(y)$ are linked.

Use quartet-tree substitution with each remaining taxon $x \in X - \mathrm{supp}(q_i, q_j)$ on q_i and q_j to construct the remaining quartet trees in $v(q_i)$ and $v(q_j)$. In the construction of each $\{q_i(x), q_j(x)\}$, one taxon (Case 1) or two taxa (Case 2) are removed and one taxon x is introduced. Thus, for $x \in X - (\mathrm{supp}(q_i, q_j))$, we have $4 \leq |\,\mathrm{supp}(q_i(x), q_j(x))| \leq 6$.

In Case 1, x replaces the same taxon in q_i and q_j and $\mathrm{supp}(q_i, q_j)$ remains the same.

In Case 2, x replaces one taxon in q_i and a different taxon in q_j.

Assume that x replaces two different taxa in $\mathrm{supp}(q_i, q_j) - \mathrm{supp}(q_i \cap q_j)$. Then, $|\,\mathrm{supp}(q_i(x), q_j(x))| = 4$ and $q_i(x) = q_j(x)$. This is not possible since

q_i and q_j distinguish different edges. Thus, x does not replace two different taxa in $\operatorname{supp}(q_i, q_j) - \operatorname{supp}(q_i \cap q_j)$ and $|\operatorname{supp}(q_i(x), q_j(x))| \neq 4$.

Now assume that x replaces two different taxa in $\operatorname{supp}(q_i \cap q_j)$. Then, we have that $|\operatorname{supp}(q_i(x), q_j(x))| = 6$. Recall that q_i and q_j distinguish the edges e_i and e_j. Thus, in order for x to replace two different taxa in $\operatorname{supp}(q_i \cap q_j)$, x would have to be in two different sets of the partition that e_i induces on X. This is not possible because x cannot be in two different sets of a partition. Thus, x does not replace two different taxa in $\operatorname{supp}(q_i \cap q_j)$ and $|\operatorname{supp}(q_i(x), q_j(x))| \neq 6$.

Thus, in this case, x replaces one taxon in $\operatorname{supp}(q_i \cap q_j)$ and one taxon in $\operatorname{supp}(q_i, q_j) - \operatorname{supp}(q_i \cap q_j)$ and $|\operatorname{supp}(\{q_i(x), q_j(x)\})| = 5$. Therefore the vines $v(q_i)$ and $v(q_j)$ are linked. □

A linking between vines allows us to construct multiple disjoint linked systems of quartet trees. We refer to these systems as *meshed systems* and use them to show that any set of quartets of sufficient size is decisive.

Definition 3.8. A *meshed system* on a tree T with taxon set X is an $(n-3)$ by $(n-3)$ array of quartet trees, where each row is a linked system and each column is a vine.

Note that the existence of a meshed system ensures that the removal of up to $n-4$ quartet trees from Q_T must leave at least one definitive set.

Theorem 3.9. *For any binary phylogenetic tree T on a taxon set X, the system Q_T of all quartet trees displayed by T contains a meshed system.*

Proof. Let T be a binary phylogenetic tree on a taxon set X and let e_1 be an interior edge adjacent to a cherry. The tree is connected, which implies we can label the remaining interior edges $\{e_2, \dots, e_{n-3}\}$ such that e_j is adjacent to some e_i with $i < j$.

Let e_j be adjacent to e_i with $i < j$. We know that e_i separates T into two connected components T_i^a and T_i^b. Moreover, e_j separates T into two connected components T_j^a and T_j^b. Since e_i and e_j are adjacent, $\operatorname{supp}(T_i^b) \cap \operatorname{supp}(T_j^a) \neq \varnothing$. Let $q_i = ab|cd$ and $q_j = ac|de$ such that $a \in \operatorname{supp}(T_i^a)$, $c \in \operatorname{supp}(T_j^b)$, and $d \in \operatorname{supp}(T_i^b) \cap \operatorname{supp}(T_j^a)$. Thus, q_i and q_j are linked for all e_i and e_j in T, and we have a linked system that makes up the first row of our matrix.

Using quartet-tree substitution, construct vines $v(q_i)$ and $v(q_j)$. By Theorem 3.7 the vines $v(q_j)$ and $v(q_i)$ are linked. Thus, we have $n-3$ disjoint columns of quartet trees in our matrix. Additionally, for each pair of linked quartet trees $\{q_i, q_j\}$ in row one, there exists a pair $\{q_i(x), q_j(x)\}$ in the remaining rows of the matrix that are linked. Thus, we have $n-3$ rows of linked systems.

Therefore, Q_T contains a meshed system. □

The existence of a meshed system allows us to find the minimal number, $k(n)$, such that every collection $\mathcal{S}$ of quartets with $|\mathcal{S}| \geq k$ is decisive.

Theorem 3.10. *The number $k(n) = \binom{n}{4} - (n-4)$ is the smallest number such that every collection of quartets $\mathcal{S}$ on a taxon set $X = [n]$ with $\mathcal{S} \geq k$ is decisive.*

Proof. Let $\mathcal{S}$ be a collection of quartets on a taxon set $X = [n]$ with $|\mathcal{S}| \geq \binom{n}{4} - (n-4)$. Let T and T' be two phylogenetic trees such that $T|Y = T'|Y$ for all $Y \in \mathcal{S}$. We define $Q \subset Q_T$ to be the collection of $T|Y$ for all $Y \in \mathcal{S}$. By Theorem 3.9, Q_T contains a meshed system M. By the pigeon hole principle, if $|Q| \geq \binom{n}{4} - (n-4)$ then Q must contain one of the linked systems in M, and by Theorem 3.1, Q is definitive. Thus, T is the unique tree which displays Q. However, since Q is also $T'|Y$ for all $Y \in \mathcal{S}$, we must have $T = T'$. Therefore, $\mathcal{S}$ is decisive. Moreover, Theorem 3.3 shows that $k \geq \binom{n}{4} - (n-4)$. Therefore $k(n) = \binom{n}{4} - (n-4)$ is the minimal number such that every collection of quartets with $|\mathcal{S}| \geq k$ is decisive. □

In addition to establishing requirements for collections of subsets of $[n]$ to be decisive, [Steel and Sanderson 2010] provides a formula for the probability that a particular collections of subsets of $[n]$ will be decisive for an arbitrarily sampled phylogenetic tree. In this section, we prove a similar result by finding a lower bound for the probability that a collection of subsets of $[n]$ of a particular size will be phylogenetically decisive. This bound is independent of the underlying tree topology.

Theorem 3.11. *The probability $p(X, k)$ that an arbitrary collection of k quartets is decisive has the property*

$$p(X,k) \geq \frac{\sum_{i=1}^{n-3} (-1)^{i+1} \binom{n-3}{i} \binom{|Q_T| - i(n-3)}{|Q_T| - k}}{\binom{|Q_T|}{k}}.$$

Proof. Let $\mathcal{S}$ be a collection of k quartets. Let T and T' be two phylogenetic trees such that $T|Y = T'|Y$ for all $Y \in \mathcal{S}$. We define $Q \subset Q_T$ to be the collection of $T|Y$ for all $Y \in \mathcal{S}$. Following Theorem 3.10, if Q contains a definitive set of quartet trees, then S is decisive. By Theorem 3.1, if a collection of compatible quartet trees contains a linked system of quartets, then it is definitive. Thus, the probability that a collection $\mathcal{S}$ is decisive is at least the probability that Q contains one of the $n-3$ disjoint linked systems constructed in Theorem 3.9. The formula follows from applying the inclusion-exclusion principle to count the number of subsets of size k which contain one of the disjoint systems of linked quartets. □

To illustrate the utility of the formula, we express the lower bound probability versus the number of quartets selected in Figure 2. In Figure 3, we plot the number of quartets required to ensure a fixed accuracy as a power of n. Notice the number of quartets needed to ensure the sample is decisive with accuracy of 25% is on the order of n^c with $c \sim 3.3$ and is almost indistinguishable from the number required to ensure 99% accuracy.

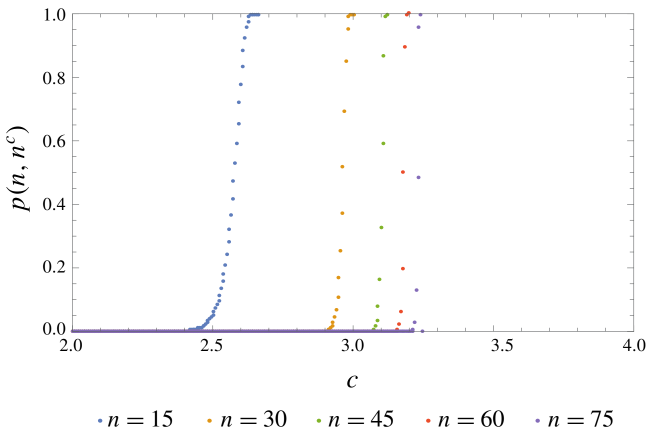

Figure 2. A lower bound on the probability that a set of n^c quartets is decisive.

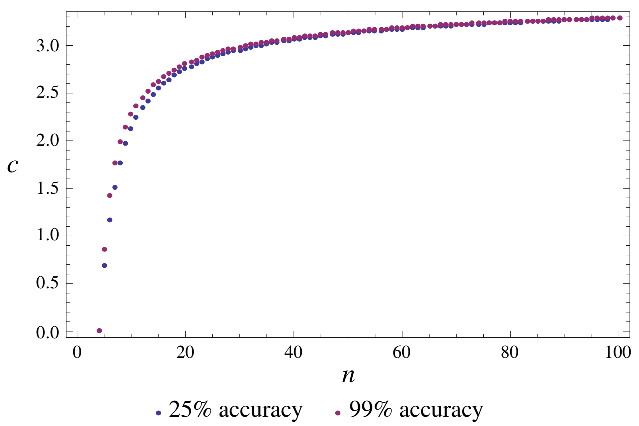

Figure 3. For $|\mathcal{S}| = n^c$, this shows the size of compatible quartets as a power of n required to ensure a decisive subset with fixed probability.

4. Conclusion

Using the criteria established in [Böcker et al. 1999], we have developed a new type of definitive system of quartet trees, linked systems. We have also developed groups of linked systems, known as meshed systems. We have used meshed systems to show that the number of quartets required to ensure decisiveness is on the order of $O(n^4)$. Moreover, we have used meshed systems to show the probability that an arbitrary collection of quartets contains a decisive system. These results lend credence to sampling quartets on the order of $n^{3.3}$.

It has been suggested that smaller sets of representative quartet trees will play a crucial role in developing efficient scalable supertree methods, as the use of all quartet tree samples may be computationally inefficient [Swenson et al. 2011]. Thus, linked systems may be useful inputs in such algorithms. However, some supertree methods, such as quartets maxcut, do not always return a fully resolved tree even when the input sets contain small definitive systems. For example, maxcut does not return a fully resolved tree for the linked system $Q_1 = \{12|35, 13|45, 14|56\}$, but returns the correct tree for the meshed system $M = \{Q_1, Q_2, Q_3\}$, where $Q_2 = \{12|34, 23|45, 24|56\}$ and $Q_3 = \{12|36, 23|46, 34|56\}$. Therefore, we anticipate that both linked and meshed systems will serve as efficient inputs for future supertree algorithms, as these algorithms could be reformulated to emphasize small definitive units.

Acknowledgements

Both authors were supported by grants from the National Center for Research Resources (5 P20 RR016461) and the National Institute of General Medical Sciences (8 P20 GM103499) from the National Institutes of Health. We would also like to thank Dr. Mike Steel for introducing us to the concept of decisiveness and giving input throughout the writing process.

References

[Böcker et al. 1999] S. Böcker, A. W. M. Dress, and M. A. Steel, "Patching up X-trees", *Ann. Comb.* **3**:1 (1999), 1–12. MR 2001d:05038 Zbl 0933.05039

[Dress et al. 2012] A. Dress, K. T. Huber, J. Koolen, V. Moulton, and A. Spillner, *Basic phylogenetic combinatorics*, Cambridge University Press, 2012. MR 2893879 Zbl 1298.92008

[Semple and Steel 2003] C. Semple and M. Steel, *Phylogenetics*, Oxford Lecture Series in Mathematics and its Applications **24**, Oxford University Press, 2003. MR 2005g:92024 Zbl 1043.92026

[Snir and Rao 2008] S. Snir and S. Rao, "Quartets maxcut: a divide and conquer quartets algorithm", *IEEE/ACM Trans. Comput. Biol. Bioinform.* **7**:4 (2008), 701–718.

[Snir et al. 2008] S. Snir, T. Warnow, and S. Rao, "Short quartet puzzling: a new quartet-based phylogeny reconstruction algorithm", *J. Comput. Biol.* **15**:1 (2008), 91–103. MR 2009b:92035

[Steel and Sanderson 2010] M. Steel and M. J. Sanderson, "Characterizing phylogenetically decisive taxon coverage", *Appl. Math. Lett.* **23**:1 (2010), 82–86. MR 2011c:05351 Zbl 1181.92068

[Strimmer and von Haeseler 1996] K. Strimmer and A. von Haeseler, "Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies", *Mol. Biol. Evol.* **13**:7 (1996), 964–969.

[Swenson et al. 2011] M. S. Swenson, R. Suri, C. R. Linder, and T. Warnow, "An experimental study of quartets maxcut and other supertree methods", *Algorithms for Molecular Biology* **6**:7 (2011).

Received: 2014-11-17 Revised: 2015-01-14 Accepted: 2015-02-02

pricee4@winthrop.edu *Department of Mathematics, Winthrop University, Rock Hill, SC 29733, United States*

rusinkoj@winthrop.edu *Department of Mathematics, Winthrop University, Rock Hill, SC 29733, United States*

involve

msp.org/involve

MANAGING EDITOR

Kenneth S. Berenhaut, Wake Forest University, USA, berenhks@wfu.edu

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2016 is US $160/year for the electronic version, and $215/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve

2016 vol. 9 no. 1

1944-4176(2016)9:1;1-4